

## 7.01 Overview and Introduction

Tadhg P. Begley, Texas A&M University, College Station, TX, USA

© 2010 Elsevier Ltd. All rights reserved.

The use of cofactors by proteins greatly augments the limited catalytic potential of the amino acid side chains and is a general motif in biocatalysis. Many of the complex reactions found in living systems are catalyzed by cofactor-dependent enzymes. Natural products chemists have been intensely interested in cofactor chemistry and the isolation, structural characterization, and synthesis of the cofactors constituted an important chapter in the history of organic chemistry. Cofactor chemistry has also served as a bridge between organic chemistry and biochemistry because cofactors were tractable experimental systems for chemical studies long before organic chemists were comfortable dealing with the complexity of enzymes. Model studies, using simple organic mimics of the enzyme–substrate complex, have played a major role in elucidating the biochemical functions of the cofactors. Largely based on these model studies, our current understanding of the mechanistic role of the cofactors is now advanced and the identification of the cofactor content of a new enzyme is often sufficient to propose a function for the enzyme and to generate a reasonable mechanistic hypothesis.

The mechanistic enzymology of cofactor biosynthesis has lagged far behind the biosynthesis of other primary metabolites; there are two major reasons for this delay. First, since cofactors are required in catalytic amounts, the biosynthetic enzymes are present at levels that are too low for direct mechanistic and structural characterization using conventional methods of protein purification. It was therefore not possible, in most cases, to proceed beyond the identification of the cofactor precursors, using labeling studies, until the methodologies for gene cloning and overexpression were readily available in chemistry laboratories. Second, since cofactors are present at low concentrations in the cell, complex, slow reactions on the biosynthetic pathways have survived evolutionary pressure toward simplification and the chemistry of cofactor biosynthesis is considerably more complex than the chemistry generally found in primary metabolism. Therefore, the successful reconstitution of many cofactor biosynthetic enzymes required sophisticated techniques and methodologies that have only become available recently. As a consequence, the mechanistic enzymology of cofactor biosynthesis is a relatively young field that has blossomed over the past 15 years. During this time, most of the genes involved in cofactor biosynthesis in bacteria have been identified and most of the enzymes involved have been mechanistically and structurally characterized.

Volume 7 of *Comprehensive Natural Products Chemistry* provides a comprehensive review of the biosynthesis and catabolism of the major cofactors (biotin, coenzyme A, folate, heme, lipoic acid, menaquinone, methanogenic cofactors, NAD, pterins, pyridoxal phosphate (PLP), riboflavin, thiamin, ubiquinone, and vitamin B<sub>12</sub>), as well as an overview of the catalytic motifs found in most cofactor-dependent enzymes. In contrast to the complexity of the cofactor biosynthetic pathways, an entirely new set of cofactors, formed by autocatalytic posttranslational modification reactions, has recently emerged. This topic is covered in Chapter 7.19. The use of comparative genomics to survey cofactor biosynthesis across all sequenced genomes is a powerful strategy for identifying new cofactor biosynthetic pathways and variations in existing pathways. This topic is reviewed in Chapters 7.5 and 7.8. Finally, the spectacularly successful process for the production of riboflavin by fermentation is described in Chapter 7.4. The regulation of cofactor biosynthetic operons by cofactor-binding riboswitches is covered in Volume 6, Chapter 6.21.

While recent progress has been impressive, there are still many unsolved problems and challenges in cofactor metabolism; some examples are as follows: The mechanistic enzymology of cofactor biosynthesis in the methanogens is still in its infancy, cofactor catabolism is poorly understood, and cofactors have not

yet been developed as probes in activity-based proteomics. Our understanding of cofactor biosynthesis also has practical applications. Genetic engineering of bacteria for the commercial production of vitamins by fermentation is an active area of current research and cofactor biosynthetic enzymes have potential as targets for new antibiotics. We hope that this volume stimulates investigators to contribute to these important and interesting problems.

### Biographical Sketch



Tadhg P. Begley obtained his B.Sc. degree from The National University of Ireland in Cork in 1977 and his Ph.D. degree from the California Institute of Technology under P. Dervan in 1982. He carried out postdoctoral studies at the University of Geneva with W. Oppolzer and at MIT with C. Walsh. After 23 years in the Cornell Chemistry Department, he recently moved to Texas A&M University where he is the Derek H. R. Barton Professor of Chemistry. Begley's research is focused on the mechanistic enzymology of complex organic transformations, particularly those found on the cofactor biosynthetic pathways.

Dr. Begley has a strong interest in the development of chemical biology. He has helped organize several major conferences and is the co-organizer (with John Schwab, NIH) of the NIH Young Faculty Mentoring Workshop. He is a member of the editorial boards for *Molecular Biosystems*, *Vitamins and Hormones*, *Bioorganic Chemistry*, *Chemical Biology and Drug Design*, the *Wiley Encyclopedia of Chemical Biology*, and has recently coauthored, with John McMurry, *The Organic Chemistry of Biological Pathways*.



## 7.02 Riboflavin Biosynthesis

**Markus Fischer**, University of Hamburg, Hamburg, Germany

**Adelbert Bacher**, Technical University of Munich, Garching, Germany

© 2010 Elsevier Ltd. All rights reserved.

---

7.02.1	Introduction	3
7.02.2	GTP Cyclohydrolase II	4
7.02.3	Deaminase/Reductase	8
7.02.4	3,4-Dihydroxy-2-Butanone 4-Phosphate Synthase	11
7.02.5	Lumazine Synthase	12
7.02.6	Riboflavin Synthase	15
7.02.7	Pentameric Riboflavin Synthases of Archaea	22
7.02.8	Lumazine Protein	23
7.02.9	Regulation of Riboflavin Biosynthesis	23
7.02.10	Biotechnology	28
7.02.11	Riboflavin Biosynthesis Genes as Potential Drug Targets	28
7.02.12	Riboflavin Kinase and FAD Synthetase	29
7.02.13	Biosynthesis of Deazaflavin Cofactors	29
7.02.14	Riboflavin as Substrate for Other Biosynthetic Pathways	29
	References	32

---

### 7.02.1 Introduction

The discovery of vitamin B<sub>2</sub> and the functions of the flavocoenzymes derived from it are rife with famous names. Paul Karrer and Richard Kuhn were awarded the Nobel Prize in 1937 and 1938, respectively, for their independent (and highly competitive) work on the isolation and structure elucidation of vitamins A and B<sub>2</sub>.

Even before the structure elucidation of the vitamin, Otto Warburg, who had just been awarded the Nobel Prize in 1931 for his work on a respiratory enzyme (i.e., cytochrome oxidase), had isolated a yellow-colored enzyme that was subsequently shown to use the 5'-phosphate of riboflavin as catalytic cofactor.<sup>1</sup>

When Hugo Theorell worked in Warburg's laboratory between 1933 and 1935, he became familiar with that 'yellow enzyme'; his further work on that protein was part of the achievements for which he was awarded the Nobel Prize in 1951. According to a reference in the *Encyclopedia Britannica*, Warburg had been shortlisted in 1944 for a second Nobel Prize based on his work on flavoproteins, but he could not receive it due to the political situation in Nazi Germany; however, the authenticity of that report is doubtful.<sup>2</sup>

The Nobel laudation for Karrer specifically mentions that his work served as the basis for the technical production of vitamin B<sub>2</sub>. In fact, a modification of that work that was introduced by Tishler was actually used for decades for the bulk production of the vitamin. However, the discovery of the vitamin and its biological functions was rapidly followed by the first attempts directed at its production by biotechnological approaches that were favored by the occurrence of naturally flavinogenic species of bacteria as well as fungi. That work prompted the first studies on the biosynthesis of the vitamin (a review on this has been done by Demain<sup>3</sup>).

More specifically, in the 1950s, McLaren reported the pioneering discovery that the fermentation yield of vitamin B<sub>2</sub> in cultures of *Eremothecium asbyii* (then in the range of milligram per liter) could be increased by the addition of naturally occurring purines such as adenine to the culture medium.<sup>4</sup> A flurry of papers subsequently demonstrated that the atoms of the purine ring system, with the exception of C-8, could actually become part of the vitamin by biochemical transformation.

Early work on flavinogenic ascomycetes by Masuda afforded a green fluorescent substance (initially designated G-compound), which was identified as 6,7-dimethyl-8-ribityllumazine that is structurally similar to riboflavin, although the benzenoid ring is missing.<sup>5-7</sup> Subsequent work showed that the compound could be

transformed into riboflavin by treatment with cell extracts. More surprisingly, that transformation could also be achieved without enzyme catalysis by boiling the compound in aqueous solution.<sup>8–11</sup> Riboflavin synthase catalyzing the transformation of G-compound into riboflavin was the first enzyme, and for a considerable period of time the only enzyme of the riboflavin biosynthesis pathway to be studied in detail.<sup>12–14</sup>

It has been known for a long time that numerous microorganisms can produce riboflavin in amounts far exceeding their metabolic requirements. Flavinogenic organisms are found among ascomycetes (e.g., *Asbyya gosypii*, *E. asbyyii*), yeasts (e.g., *Candida guilliermondii*, *Candida flaveri*), and bacteria. Early work on riboflavin oversynthesis has been reviewed by Demain.<sup>3</sup>

More recently, the elucidation of the biosynthetic pathway has resulted in highly efficient biotechnical methods that have replaced the chemical synthesis of vitamin B<sub>2</sub>.<sup>15–18</sup>

The present information on the riboflavin biosynthetic pathway is summarized in **Figure 1**. Briefly, the pathway starts from GTP (**1**), which is converted into the first committed intermediate **2** by the hydrolytic release of pyrophosphate and of C-8 of the imidazole ring that are both catalyzed by a single enzyme, GTP cyclohydrolase II (reaction I). In Archaea and in fungi, that compound is transformed into 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione phosphate (**5**) by a reduction (reaction IV) that transforms the ribosyl side chain into the ribityl side chain (**4**) and by subsequent deamination (reaction V) of the pyrimidine ring yielding compound **5**. In plants and in eubacteria (reactions II and III), these reaction steps occur in inverse order via the ribosylaminopyrimidine derivative **3**.

It is still unknown how the pyrimidine intermediate **5** is dephosphorylated (reaction VI). However, it is well established that the dephosphorylation product **6** is condensed with 3,4-dihydroxy-2-butanone 4-phosphate (**8**) by the catalytic action of lumazine synthase (reaction VIII). The carbohydrate substrate **8** is in turn obtained from ribulose phosphate (**7**) by a complex reaction sequence that is catalyzed by a single enzyme, 3,4-dihydroxy-2-butanone 4-phosphate synthase (reaction VII). As mentioned above, the lumazine **9** is converted to riboflavin (**10**) by the catalytic action of riboflavin synthase (reaction IX). Ultimately, riboflavin is converted to the coenzymes, riboflavin 5'-phosphate (flavin mononucleotide (FMN), **11**) and flavin adenine dinucleotide (FAD, **12**) by the catalytic action of riboflavin kinase (reaction X) and FAD synthase (reaction XI). These reaction steps are required in all organisms, irrespective of their acquisition of riboflavin from nutritional sources or by endogenous biosynthesis.

Earlier work on riboflavin biosynthesis has been reviewed repeatedly.<sup>19–27</sup> The present article is focused on recent work on the structure and mechanism of riboflavin biosynthetic enzymes. For historic aspects of this research area, the reader is directed to review the articles cited.

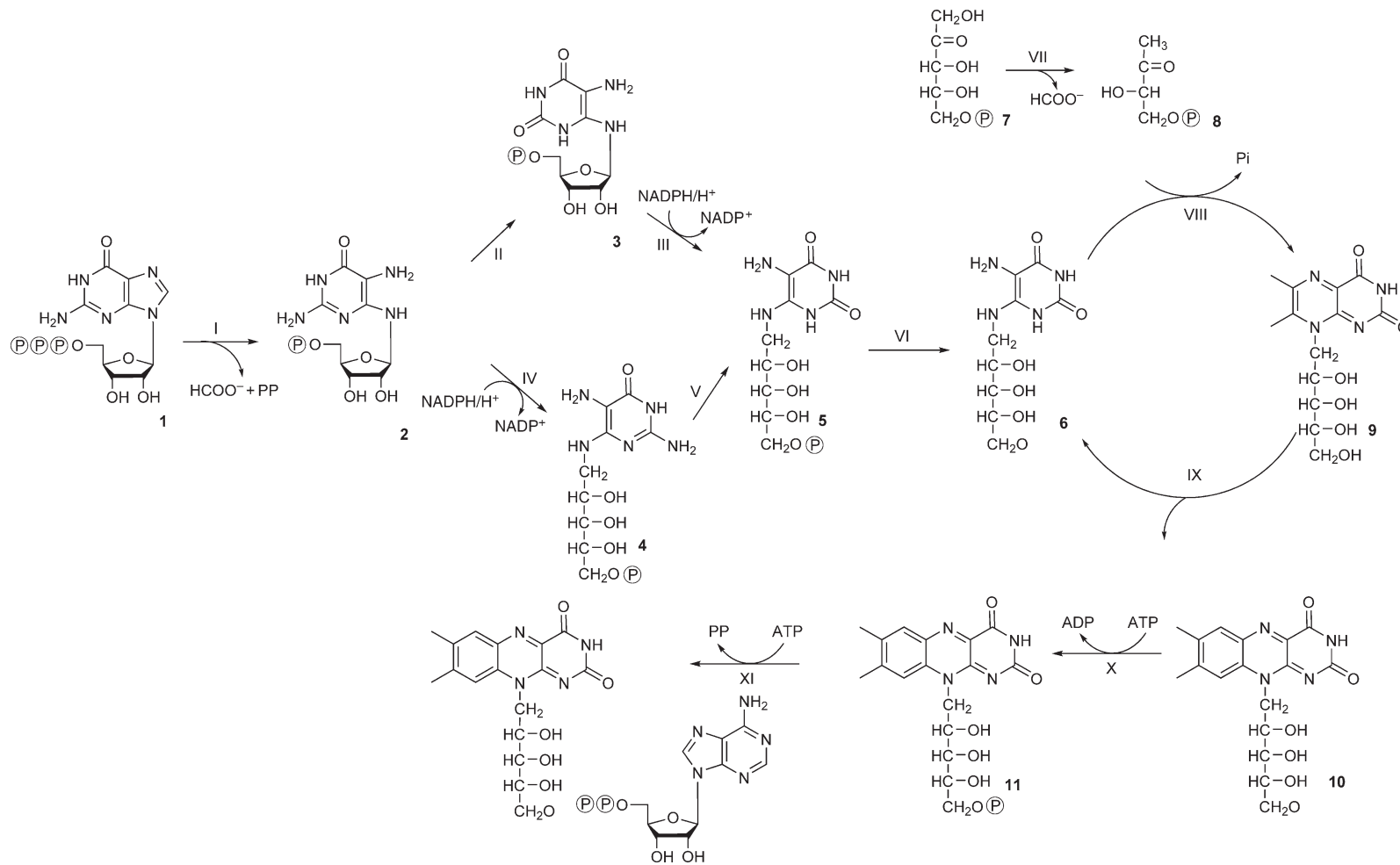
## 7.02.2 GTP Cyclohydrolase II

A biosynthetic relationship between purines and flavin had been initially conjectured on the basis of fermentation experiments where the addition of certain purine derivatives increased the yield of riboflavin produced by flavinogenic microorganisms. Ample confirmation for that hypothesis was obtained by isotope incorporation studies performed in parallel by various research groups using different microorganisms. Specifically, it was shown that the pyrimidine ring of purines can be contributed to biosynthetic riboflavin in its entirety, whereas C-8 of the imidazole ring of purine precursors is lost. The early work has been reviewed repeatedly and will not be discussed in detail.<sup>19,28–31</sup>

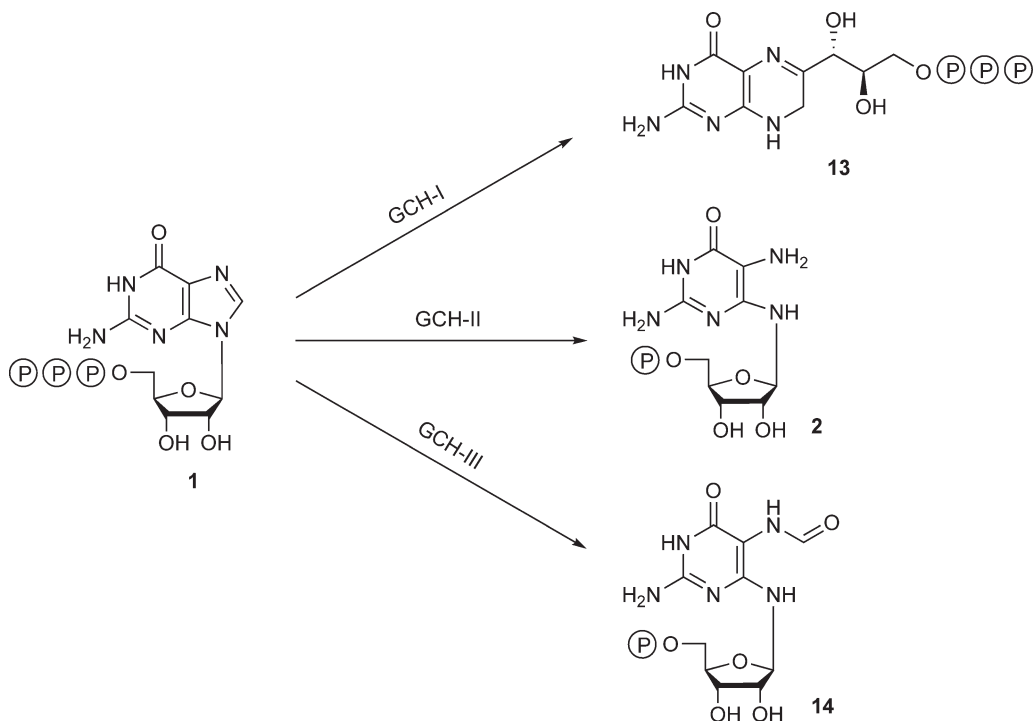
Studies with purine-deficient mutants narrowed the purine precursor down to a compound at the biosynthetic level of guanine.<sup>32–34</sup> This implicated that the amino group of the cognate guanine precursor had to be replaced by oxygen in the course of biosynthesis.

Work with a multiple purine mutant could also show that the pyrimidine ring and the ribose side chain of a guanosine nucleoside or nucleotide were incorporated *en bloc* into riboflavin; thus, it was obviously necessary to convert the ribosyl moiety of the guanosine-type precursor into the ribityl side chain of the vitamin by a reduction step.<sup>34</sup>

GTP cyclohydrolase II (reaction I) catalyzing the first committed step in the biosynthesis of riboflavin was discovered in the wake of work on GTP cyclohydrolase I (**Figure 2**), the enzyme catalyzing the first reaction step in the biosynthesis of tetrahydrofolate.<sup>35–39</sup>



**Figure 1** Biosynthesis of riboflavin and flavocoenzymes.



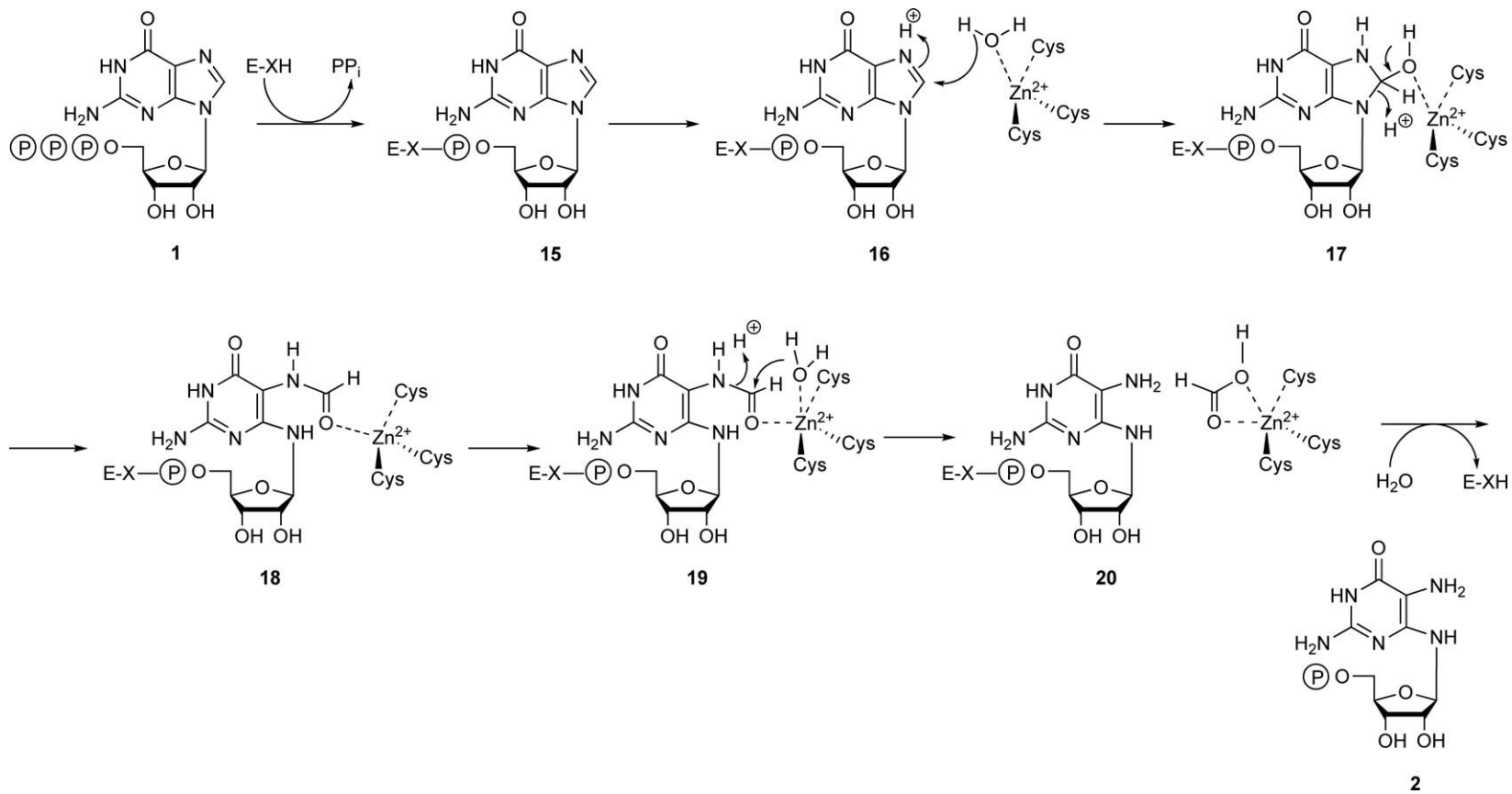
**Figure 2** GTP cyclohydrolases.

The latter enzyme enables a complex series of reaction steps involving the hydrolytic opening of the imidazole ring of GTP between C-8 and N-9, hydrolytic deformylation of the resulting formamide-type intermediate, Amadori rearrangement of the ribosyl side chain, and finally ring closure under formation of a dihydropterine derivative (**13**).<sup>40–44</sup> Since C-8 of the GTP precursor is liberated as formate, the reaction could be monitored by the detection of radiolabeled formate from [8-<sup>14</sup>C]GTP.<sup>35</sup> While GTP cyclohydrolase I activity is not affected by the addition of EDTA, cell extracts of *Escherichia coli* were found to contain a second enzyme activity that was inhibited by EDTA, but the activity could be restored by the addition of Mg<sup>2+</sup> ions. The Mg<sup>2+</sup>-dependent enzyme was then shown to catalyze not only the release of formate but also of pyrophosphate from GTP, and the product was assigned as 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (**2**). More recently, that structure assignment could be confirmed directly by nuclear magnetic resonance (NMR) spectroscopy using [U-<sup>13</sup>C<sub>10</sub>]GTP as substrate for recombinant GTP cyclohydrolase II in order to increase detection sensitivity.<sup>45</sup> The direct NMR observation showed that the  $\beta$ -anomer, which is the direct product of the enzyme reaction can undergo spontaneous isomerization affording the  $\alpha$ -anomer in neutral aqueous solution at room temperature.<sup>45</sup>

Recombinant GTP cyclohydrolase II of *E. coli* is a homodimer of 22 kDa subunits.<sup>46</sup> The protein contains one tightly bound Zn<sup>2+</sup> ion per monomer.<sup>47</sup> The enzyme converts GTP into a mixture of about 90% of the specific product **2** and about 10% of GMP.<sup>45</sup> It was also shown that the 5'-triphosphates of 8-oxo-7,8-dihydro-2'-deoxyguanosine and 8-oxo-7,8-dihydroguanosine can be converted into the respective monophosphates, although the enzyme is unable to open the imidazole ring of the structurally modified guanine residues of these nucleotides.<sup>48</sup> Surprisingly, the formation of the covalent phosphoguanosyl derivative of the enzyme is the rate-determining step of the overall reaction sequence.<sup>45</sup> The rate of formation of the specific product **2** is 182 nmol mg<sup>-1</sup> min<sup>-1</sup> at 37 °C.<sup>47</sup>

Mutation analysis identified three absolutely conserved cysteine residues (Cys54, Cys65, and Cys67); the involvement of these residues in the complexation of the essential zinc ion was later confirmed by X-ray structure analysis.<sup>46,47</sup>

**Figure 3** shows the hypothetical reaction mechanism of GTP cyclohydrolase II. The replacement of any of the cysteine residues that complex the catalytic Zn<sup>2+</sup> ion afforded mutant proteins that were unable to open the



**Figure 3** Hypothetical mechanism for release of formate by GTP cyclohydrolase II.

imidazole ring of GTP or to hydrolyze the formamide motif of the reaction intermediate **18** but were capable of generating inorganic pyrophosphate from GTP under production of GMP. Moreover, the mutants could convert 2-amino-5-formylamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-triphosphate (**18**) into the corresponding monophosphate.<sup>47</sup>

In H<sub>2</sub><sup>18</sup>O as solvent, <sup>18</sup>O is specifically incorporated into the phosphate residue of the 5-aminopyrimidine derivative **2** and of GMP but not into inorganic pyrophosphate.<sup>45</sup> This is well in line with the hypothesis that a guanylate residue is covalently linked to the enzyme in an early reaction step.<sup>45</sup> Hydrolytic cleavage of the covalent bond between the GMP moiety and the enzyme, either before or after hydrolytic removal of C-8 from the imidazole moiety of GMP, should result in the observed incorporation of <sup>18</sup>O from the bulk solvent into the products **2** and GMP.<sup>45</sup>

Kinetic analysis showed that the rate-determining step is located in the early part of the reaction trajectory; under single-turnover (quenched flow) conditions, the rates for the consumption of the substrate and the formation of the product **2** were identical within the limits of experimental accuracy.<sup>49</sup>

2-Amino-5-formylamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-triphosphate can serve as a substrate but does not fulfill the criteria for a kinetically competent intermediate. The kinetic data suggest that the initial formation of the covalent guanyl adduct (**15**) is the rate-determining step, although the hydrolytic removal of C-8 from the imidazole ring might intuitively appear to have a higher free energy barrier than the formation of the covalent guanyl adduct.<sup>49</sup> The zinc ion is apparently required for both hydrolytic reaction steps involved in the formation of formate; mutants that are unable to bind the essential Zn<sup>2+</sup> ion are unable to form **2** from either GTP (**1**) or the nonnatural substrate, 2-amino-5-formylamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-triphosphate.<sup>47</sup> Notably, the initial formation of a covalent enzyme adduct is absolutely required for product formation, since the 5'-monophosphate of 2-amino-5-formylamino-6-ribosylamino-4(3*H*)-pyrimidinone cannot be converted to product.

In summary, GTP cyclohydrolase II appears to catalyze an ordered reaction that starts with the formation of a covalent guanyl adduct (**15**). This is followed by the hydrolytic opening of the imidazole ring (**16**, **17**) and the hydrolysis of the resulting formamide-type intermediate (**18**, **19**); the latter two reactions depend on the Zn<sup>2+</sup> ion acting as a Lewis acid, which sequentially activates two water molecules (**17**, **19**) that attack first the imidazole carbon atom 8 of the covalent guanyl adduct **17** and then the formamide motif of the covalent intermediate **19**.

X-ray structure analysis of GTP cyclohydrolase from *E. coli* revealed a c<sub>2</sub>-symmetric dimer whose folding pattern and substrate-binding topology has no similarity with GTP cyclohydrolase I.<sup>37,46,50</sup> Based on topological arguments, arginine 128 is likely to serve as the binding partner for covalent phosphoguanilate attachment. Tyrosine 105 is believed to cooperate with the Zn<sup>2+</sup> ion in the activation of water.<sup>46</sup>

Recent studies have identified three different cyclohydrolase II genes/proteins in the *Streptomyces coelicolor* genome.<sup>51</sup> Two orthologues catalyze the formation of 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate (**2**). In the third protein, a tyrosine residue that is homologous to tyrosine 105 of the *E. coli* protein is replaced by methionine.<sup>52</sup> This isoenzyme stops short of the production of **2** and generates the formamide-type compound **14** (2-amino-5-formylamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate) instead. The biological role of this enzyme, if any, is currently unknown. The protein has been designated cyclohydrolase III.

More importantly, Archaea have no apparent homologues of the eubacterial GTP cyclohydrolase II. However, a cyclohydrolase III-type protein with an unrelated sequence type has been reported to produce the formamide-type product **14**.<sup>53</sup> It is unknown how that compound is deformylated in Archaea.

Plants and certain eubacteria specify bifunctional proteins where a cyclohydrolase II domain is fused with a domain catalyzing the formation of 3,4-dihydroxy-2-butanone 4-phosphate.<sup>54,55</sup> Jointly, these two domains produce both committed starting materials of the convergent riboflavin pathway (see below).

### 7.02.3 Deaminase/Reductase

Two somewhat different pathways exist for the conversion of 2,5-diamino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate (**2**) into the substrate of lumazine synthase, 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-dihydropyrimidine-dione (**6**). In fungi and Archaea (**Figure 1**, reaction IV), the ribosyl side chain of **2** is reduced under formation of

2,5-diamino-6-ribitylamino-4(3*H*)-pyrimidinone 5'-phosphate (**4**).<sup>56,57</sup> Subsequent deamination of the pyrimidine ring affords 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione 5'-phosphate (**5**). In eubacteria and plants (**Figure 1**, reaction III), the product of cyclohydrolase II (**2**) first undergoes deamination affording 5-amino-6-ribosylamino-4(3*H*)-pyrimidinone 5'-phosphate (**3**), which is then converted into 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione 5'-phosphate (**5**) by side-chain reduction.<sup>58–60</sup>

By comparison with other aspects of riboflavin biosynthesis, the reduction and deamination steps obtained less extensive coverage in the literature. Recently, however, the topic has been addressed more actively by structural biologists.

Historically, the first evidence for the fungal pathway was the isolation of reduced and deaminated products of **4** and **5** from the culture fluid of mutants of *Saccharomyces cerevisiae*.<sup>61–63</sup> Similar results have been obtained with *C. guilliermondii*<sup>64–67</sup> and *Aspergillus nidulans*.<sup>68</sup> On the other hand, the bacterial pathway was discovered by direct enzyme studies in *E. coli*.<sup>58</sup> Although the initial studies suggested separate proteins with deaminase and reductase activity, the *ribG* gene of *Bacillus subtilis* is now known to specify a bifunctional protein with an N-terminal deaminase domain and a C-terminal reductase domain; recombinant expression of the separate domains afforded only highly unstable proteins.<sup>60</sup> Bifunctional deaminase/reductase domains appear to be the rule in eubacteria. However, one known exception is *Buchnera aphidicola*, which carries two adjacent open reading frames for deaminase and reductase.

The deaminase from the plant, *Arabidopsis thaliana*, comprises an N-terminal plastid-targeting sequence, a catalytically competent deaminase domain, and a putative domain of unknown function. The deaminase domain shows remarkably close similarity to those of eubacteria.<sup>59</sup>

A second family of plant genes with similarity to bacterial *ribG* genes predicts a domain with an N-terminal domain that has significant homology to deaminases, but without the canonical zinc-binding motif, and with a C-terminal domain that may well represent a catalytically active reductase domain, but studies at the protein level are yet to be performed.<sup>69</sup>

A monofunctional reductase has been expressed from *Methanocaldococcus jannaschii* and has been shown to selectively use **2** as substrate.<sup>57</sup> The deaminase of Archaea is yet to be discovered.

Yeasts, fungi, and Archaea use monofunctional reductases. Although the deaminase of Archaea is as yet unknown, the presence of a monofunctional reductase<sup>57</sup> suggests that the deaminases could also be monofunctional. The deaminase specified by the RIB2 gene of *S. cerevisiae* has been partially purified, but no detailed studies have been performed.<sup>70,71</sup> Recently, the RIB2 gene of *S. cerevisiae* has been expressed in *E. coli*.<sup>72</sup> The resulting protein possesses two distinct sequence domains, namely a C-terminal deaminase domain and an N-terminal part related to RNA:pseudouridine (psi)-synthases.<sup>73</sup> The deaminase domain is implicated in the riboflavin biosynthesis, while the exact function of the RNA:Psi-synthase domain remains obscure.

X-ray structures have been reported for the bifunctional deaminase/reductase proteins from *E. coli* and *B. subtilis* and for the monofunctional reductase of *M. jannaschii*.<sup>56,69,74</sup> The *E. coli* protein is a homodimer and the *B. subtilis* protein is a heterotetramer.<sup>69,74</sup>

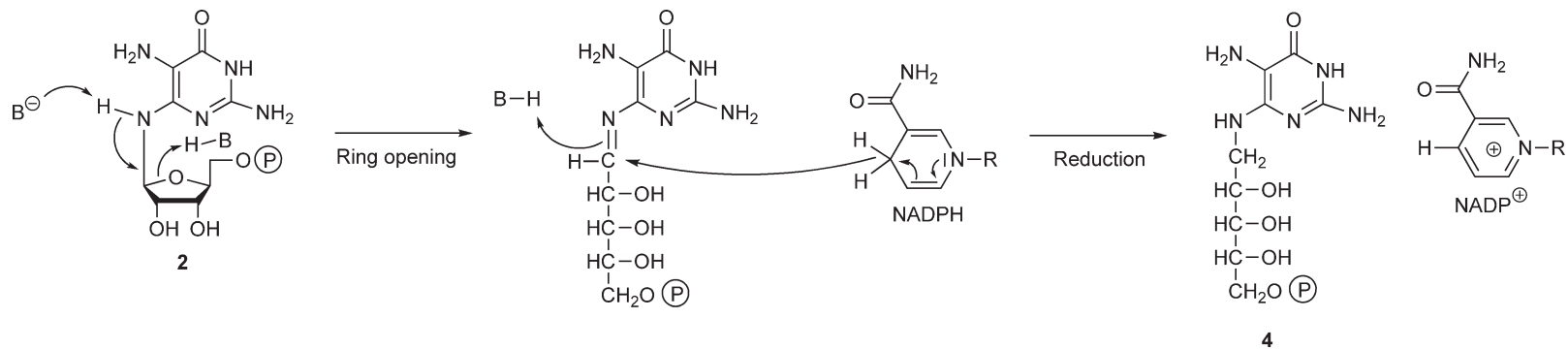
The deaminase domains are members of the cytosine deaminase superfamily.<sup>59,60</sup> They bind catalytically essential Zn<sup>2+</sup> ions through complexation by one histidine and two cysteine residues. The reductase domains are members of the dihydrofolate reductase superfamily.<sup>56</sup> In fact, the reactions catalyzed by pyrimidine reductase and dihydrofolate reductase are closely similar. Early *in vivo* work had shown that the hydride ion is incorporated into the 1' position.<sup>75</sup>

Thus, it follows that the substrate for reduction is a Schiff base as opposed to an Amadori compound. Thus, the substrates of dihydrofolate reductase and riboflavin-type pyrimidine reductase differ mainly by the inclusion of the Schiff base into a ring system in case of dihydrofolate.

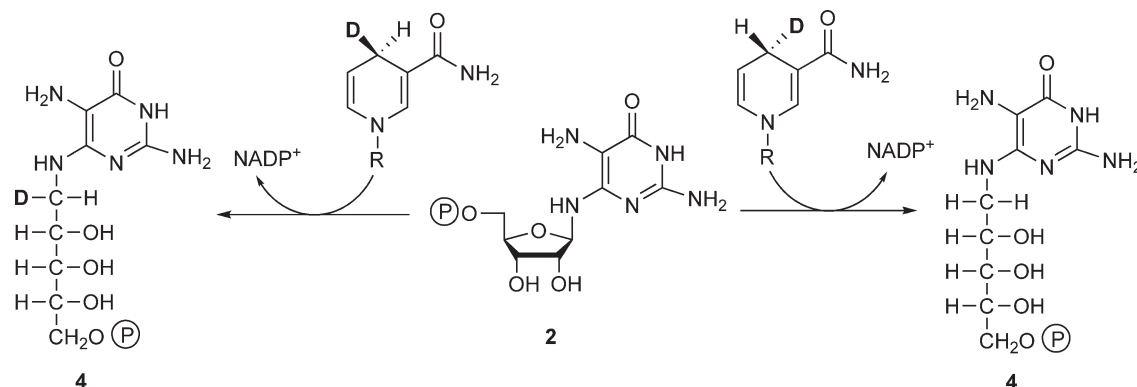
The X-ray structure of the enzyme from *M. jannaschii* in complex with NADPH suggests the transfer of the proR hydrogen of C-4 of NADPH to C-1' of the substrate.<sup>56</sup> *In vivo* studies indicated that the hydride ion is implemented in the proS position of C-1' of the substrate (**Figures 4 and 5**).<sup>76–79</sup>

The available X-ray structures provide no additional information on that aspect of the reaction since enzyme/substrate complex structures have not yet been reported. The crystallization of a reductase/deaminase from *Shigella* has been reported, but the structure of the protein remains to be determined.<sup>80</sup>





**Figure 4** Proposed reaction mechanism for reduction of the ribosylamino to the ribitylamino moiety in riboflavin biosynthesis.

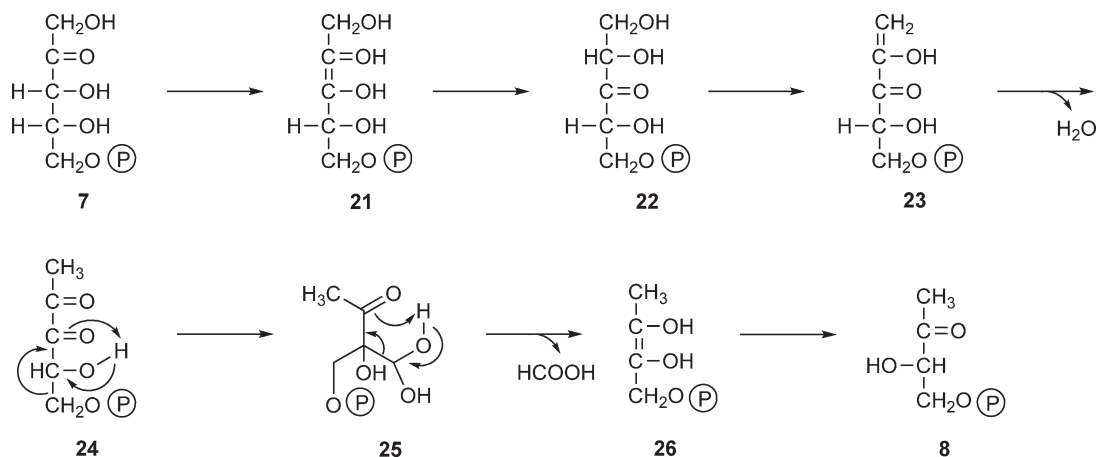


**Figure 5** Stereochemistry of hydride transfer from NADPH to the product of the *Methanocaldococcus jannaschii* reductase. Only deuterium from the 4(*R*)-position of NADPH is transferred into the 1'-proS position of the product **5** (left side).<sup>79</sup>

### 7.02.4 3,4-Dihydroxy-2-Butanone 4-Phosphate Synthase

Since the discovery of riboflavin synthase more than four decades ago, it was understood that all eight carbon atoms of the benzenoid ring of the vitamin are derived from carbon atoms 6 $\alpha$ , 6, 7, and 7 $\alpha$  of 6,7-dimethyl-8-ribityllumazine via the strange dismutation catalyzed by riboflavin synthase.<sup>12–14,81–83</sup> However, the biosynthetic origin of the respective carbon atoms of the lumazine derivative remained controversial for a long time. The early studies on this problem have been reviewed repeatedly.<sup>20,24,84</sup> *In vivo* studies with various <sup>13</sup>C-labeled precursors then revealed the formation of a 4-carbon precursor by an intramolecular rearrangement of a pentose or pentulose derivative that involves the loss of carbon 4.<sup>85–88</sup> Later studies established ribulose 5-phosphate as the direct precursor of the 4-carbon unit.<sup>89,90</sup> To make a long story short<sup>90,91</sup> it was then shown that a single enzyme converts ribulose 5-phosphate into 3,4-dihydroxy-2-butanone 4-phosphate (**Figure 6**).

The reaction involves the elimination of C-4 of the pentulose substrate and proceeds by a strictly intramolecular mechanism.<sup>91</sup> Deuterium from solvent water is incorporated into the position 1 methyl group and at C-3. A hypothetical reaction mechanism that is consistent with all experimental data is summarized in **Figure 6**.<sup>91</sup> The reaction is believed to begin with the formation of an endiol (**21**) from the substrate, ribulose 5-phosphate (**7**). Protonation at C-1 of the intermediate **22** and the subsequent elimination of water affords an enol-type intermediate that can tautomerize under formation of the diketone **24**. A sigmatropic rearrangement is then believed to generate the branched aldose intermediate **25** that can fragment under release of formate.



**Figure 6** Hypothetical reaction mechanism of 3,4-dihydroxy-2-butanone 4-phosphate synthase.<sup>89–91</sup>

**Table 1** Specific activities of enzymes involved in the biosynthesis of riboflavin of *Escherichia coli*

Enzyme	Gene	Activity (nmol mg <sup>-1</sup> min <sup>-1</sup> ); 37 °C	Reference
GTP cyclohydrolase II	<i>ribA</i>	182	47
Pyrimidine deaminases	<i>ribD</i>	360	60
Pyrimidine reductases	<i>ribD</i>	110 (NADPH) 125 (NADH)	60
3,4-Dihydroxy-2-butanone 4-phosphate synthase	<i>ribB</i>	283	84
Lumazine synthase	<i>ribE</i>	197	93
Riboflavin synthase	<i>ribC</i>	430	94

The resulting enol intermediate **26** is reprotonated in a stereospecific mode affording the L(*S*) enantiomer **8**,<sup>90,91</sup> the proton used for reprotonation at C-3 is introduced from bulk water. The stereochemistry of the rearrangement reaction has been studied in some detail.<sup>92</sup>

The rate constants of 3,4-dihydroxy-2-butanone 4-phosphate synthases (reaction VII in **Figure 1**) from various microbial organisms are generally low (**Table 1**).<sup>84,95–97</sup>

In *E. coli*, the unusual reaction is catalyzed by a homodimer of 47 kDa.<sup>98,99</sup> The reaction requires Mg<sup>2+</sup> but no other cofactors. The structure of the *E. coli* enzyme has been elucidated by X-ray crystallography and by NMR spectroscopy.<sup>98,99</sup> The homodimer is characterized by c<sub>2</sub> symmetry. Each of the two topologically active sites is located at a subunit interface. The reaction cavity can be occluded by a flexible loop that contains several acidic amino acids. The replacement of any of the conserved acidic amino acid residues in that loop inactivates the enzyme. Several other charged amino acid residues at the active site are likewise essential for enzymatic activity.<sup>96</sup>

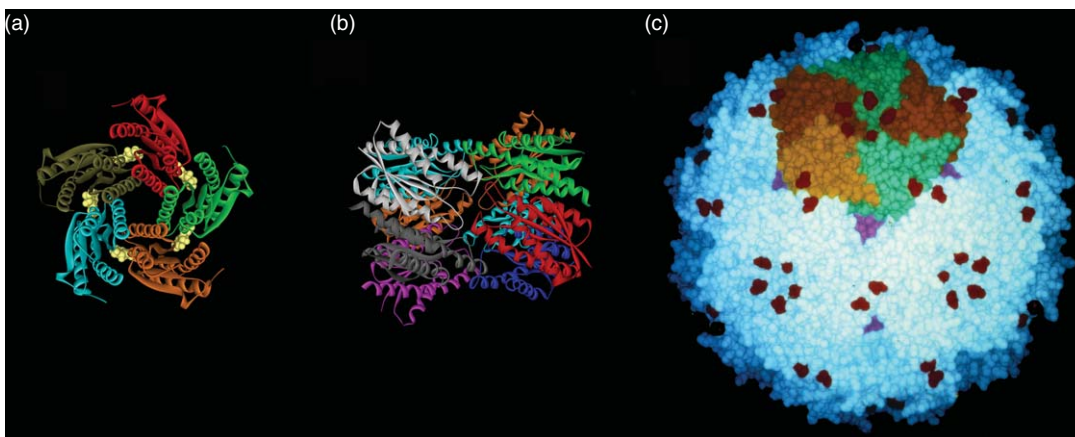
X-ray structures were also reported for the enzymes from *Magnaporthe grisea*, *M. jannaschii*, and *Candida albicans*.<sup>95,100–102</sup>

In plants and in many microorganisms (but not in *E. coli*), 3,4-dihydroxy-2-butanone-4-phosphate synthase and GTP cyclohydrolase II are expressed as a fusion protein with GTP cyclohydrolase II as the C-terminal domain.<sup>54</sup> The ratio of product formation of the two initial reactions of the convergent biosynthetic pathway is thus rigidly coupled.

## 7.02.5 Lumazine Synthase

Studies in the 1970s revealed the presence of two different proteins with riboflavin synthase activity in *B. subtilis*. One was a homotrimer of 25 kDa subunits, the other was a complex protein where a homotrimer of the 25 kDa subunits was associated with sixty 16 kDa subunits.<sup>103–105</sup> More precisely, the trimer of 25 kDa subunits was enclosed in a capsid of sixty 15 kDa subunits with icosahedral 532 symmetry.<sup>103</sup> The complex protein had a mass of about 1 MDa and was designated heavy riboflavin synthase, whereas the homotrimer with a mass of about 75 kDa was designated light riboflavin synthase.<sup>103</sup> The 25 kDa subunits (which were subsequently designated  $\alpha$  subunits) were apparently the carriers of the riboflavin synthase activity in both proteins, and the function of 16 kDa subunits that were designated  $\beta$  subunits remained unknown for two decades.

Subsequent to the identification of 3,4-dihydroxy-2-butanone 4-phosphate (**8**) as the 4-carbon precursor of 6,7-dimethyl-8-ribityllumazine in the late 1980s,<sup>89</sup> it was not difficult to show that the  $\beta$  subunits catalyzed the condensation of **6** with **8** under formation of the lumazine derivative **9**. The ‘heavy riboflavin synthase’ can now be correctly addressed as a riboflavin synthase/lumazine synthase complex. However, the designation ‘ $\beta$  subunit of riboflavin synthase’ clung tenaciously to the 16 kDa peptide that had ultimately been assigned a function as lumazine synthase (reaction VIII in **Figure 1**); notably, this somewhat unfortunate nomenclature survives in databases, despite the fact that in most organisms, lumazine synthase and riboflavin synthase (reaction IX in **Figure 1**) are separate proteins.



**Figure 7** Crystal structures of lumazine synthases of eubacteria and fungi. (a) Pentameric lumazine synthase of *Schizosaccharomyces pombe*. (b) Decameric lumazine synthase from *Brucella abortus*. (c) Icosahedral lumazine synthase of *Bacillus subtilis*.

As a further semantic complication, recent work has identified a riboflavin synthase type in Archaea that is paralogous to lumazine synthase.<sup>106–108</sup> This unexpected finding will be described in more detail in the chapter on riboflavin synthase.

The sequence evolution of lumazine synthase has been relatively conservative. In a wide variety of prokaryotes and eukaryotes that engage in the biosynthesis of riboflavin, the lumazine synthases (including the mature lumazine synthases in plant chloroplasts that have lost their N-terminal targeting peptides by processing<sup>109</sup>) are small peptides of about 150 amino acids, which obey the same folding pattern with remarkable fidelity.<sup>110,111</sup> Despite that similarity, lumazine synthases from different organisms show widely different molecular weights due to their different quaternary structures involving pentamers, decamers, and 60-mers (**Figure 7**). The three-dimensional structures of lumazine synthase have been studied in considerable detail and will be described below.

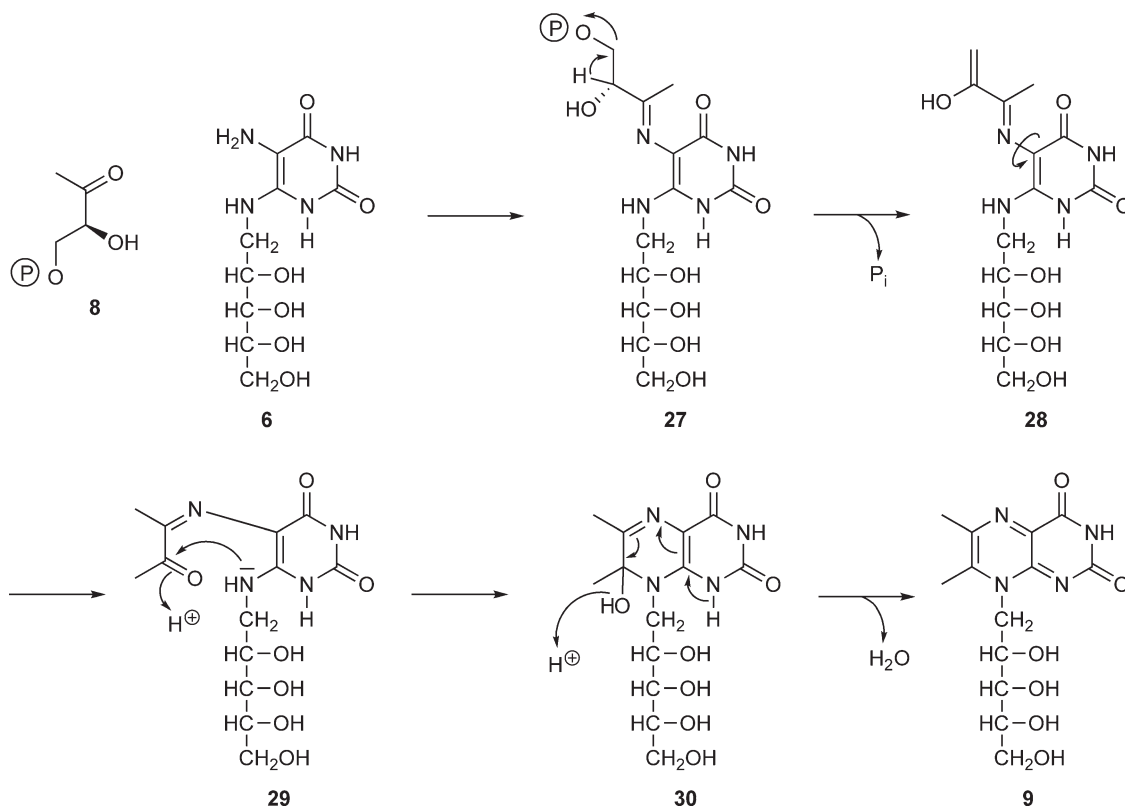
Lumazine synthase catalyzes the condensation of 5-amino-6-ribitylamino-2,4-(1*H*,3*H*)-pyrimidinedione (**6**) with 3,4-dihydroxy-2-butanone 4-phosphate (**8**).<sup>85,89,112</sup> Lumazine synthase is unable to use the phosphoric acid ester **5** as substrate. It follows that the phosphate residue of **5** must be removed, but it is still unknown how this is achieved.

The condensation of the pyrimidine derivative **6** with the 3,4-dihydroxy-2-butanone 4-phosphate (**8**) is regiospecific as shown by experiments with isotope-labeled substrate.<sup>89</sup> The regiospecificity is best explained by the hypothesis that the carbonyl group of **8** is initially attacked by the position 5 amino group of **6** (**Figure 8**).<sup>85,112</sup>

The resulting Schiff base (**27**) is believed to undergo phosphate elimination affording the enolate **28**. Direct evidence for the Schiff base intermediate was obtained by presteady-state kinetic analysis.<sup>113,114</sup> Based on X-ray structure data, the Schiff base and probably the enolate **28** are formed in an extended conformation that must undergo rotation prior to ring closure. This rotation could occur before or after the tautomerization converting the enolate **28** into the ketone **29**. The reaction is terminated by an intramolecular nucleophilic attack of the newly formed carbonyl group in **29** by the ribitylamino group in the 6 position of the pyrimidine intermediate **29**.<sup>112</sup>

The reaction mechanism shown in **Figure 6** suggests the involvement of several proton transfer reactions and one would expect the participation of amino acid residue as acid/base catalysts. In a systematic mutagenesis study, the replacement of amino acid residues lining the active site cavity had only minor impact on catalytic rates.<sup>115</sup> This suggests that the contribution of enzyme catalysis is mainly entropic, via the generation of a favorable reaction topology.

This hypothesis is well in line with the finding that the formation of **9** by condensation of **6** and **8** can proceed in the absence of a catalyst under mild conditions.<sup>116</sup> Specifically, the reaction occurs with appreciable velocity at room temperature in aqueous solution of neutral pH with reactant concentrations in the low  $\text{mmol}^{-1}$  range.<sup>116</sup> The regiospecificity of the uncatalyzed reaction is significantly lower as compared to the strictly regiospecific enzyme-catalyzed reaction. Most likely, the spontaneous reaction is subject to partitioning



**Figure 8** Hypothetical mechanism of lumazine synthase.

with one reaction pathway that corresponds to the mechanism of the enzyme-catalyzed reaction and a second pathway that involves the formation of diacetyl by elimination of phosphate from **8**; the inherently symmetric diacetyl could then react in a nonregiospecific mode with the pyrimidine intermediate **6**.<sup>116</sup>

The riboflavin precursor **9** can also be obtained by heating a solution of **6** and ribulose 1,5-bisphosphate at 120 °C and pH 7.3 with a relatively high yield.<sup>117,118</sup>

The Schiff base (**27**) formed in the initial phase of the reaction catalyzed by lumazine synthase can be observed in single-turnover experiments as an optical transient with an absorption maximum at 330 nm.<sup>113</sup> A later optical transient with an absorption maximum at 445 nm has been assigned to the product resulting from phosphate elimination. Surprisingly, the ring closure reaction at the end of the reaction sequence appears as the rate-determining step.<sup>113</sup>

Crystal structures have been reported for eubacteria, Archaea, fungi, and plants.<sup>110,111,119–130</sup> The structure of the protein from *Aquifex aeolicus* has been determined to a resolution of 1.6 Å.<sup>126</sup> The protomer folding pattern is characterized by a four-stranded  $\beta$  sheet that is flanked on both sides by two helices. The protomers of yeasts, fungi, and certain bacteria associate under formation of  $c_5$ -symmetric homopentamers.<sup>110,111,119–122</sup> The topologically equivalent active sites are all located at interfaces of adjacent subunits in the pentameric assembly.

The 60-mers with icosahedral symmetry that are observed in many bacteria and in plants are best described as dodecamers of pentamers, although attempts to directly observe pentamer precursors of the icosahedral 60-mers have met with little success. The molecules have the shape of hollow capsids with a diameter of about 160 Å and a wall thickness of about 50 Å.<sup>123</sup> Relatively narrow channels through the capsid wall are observed along the five-fold symmetry axes. A considerable number of X-ray structures with bound substrate and intermediate analogues define the large active site cavity. The capsid wall is rather tightly knit and it is not entirely clear how substrates and products can traverse between solvent and active sites.

A decameric lumazine synthase that can be described as a  $d_5$ -symmetric aggregate of two pentamers has been found in the pathogenic *Brucella abortus*.<sup>128–130</sup> Besides the decameric species, the bacterium specifies a

second orthologue that forms a homopentamer.<sup>131</sup> The functional implications of two different lumazine synthase orthologues in *Brucella* and related bacteria are currently unknown. Notably, pentameric *Brucella* lumazine synthase is an immunogenic protein during infection and has been proposed as a useful antigen for serological diagnosis and a potential candidate for the design of acellular vaccines.<sup>132</sup>

In Bacillaceae, the icosahedral lumazine synthase capsid can trap a homotrimeric riboflavin synthase molecule in the central core.<sup>103,105,124,133–136</sup> That unusual complex was originally designated ‘heavy riboflavin synthase’ in the absence of knowledge of the catalytic function of the 15 kDa subunits that form the icosahedral capsid. The structure of the complex is not known in atomic detail. As described below, the homotrimeric riboflavin synthase is, surprisingly, devoid of trigonal symmetry;<sup>137,138</sup> moreover, although the molecule has two different types of two-fold pseudosymmetry, no strict two-fold symmetry is present. In the absence of symmetry elements that could be used in a symmetry-related interaction between the inherently asymmetric, homotrimeric kernel and the 532 symmetric capsid, computer modeling could show that the trimer can be fit into the capsid without steric clashes, but the topological relationship remains unknown.

Steady-state kinetic analysis afforded evidence for intermediate channeling in the lumazine synthase/riboflavin synthase complex.<sup>139</sup> Briefly, the conversion of 6,7-dimethyl-8-ribityllumazine molecules that have been newly formed by the lumazine synthase module of the protein complex are more rapidly converted to riboflavin than molecules from the bulk solvent. The topological constraint by the capsid is believed to cause this phenomenon. It has been proposed that the channels along the five-fold axes could serve as port of entry and exit for substrates and products.<sup>124</sup>

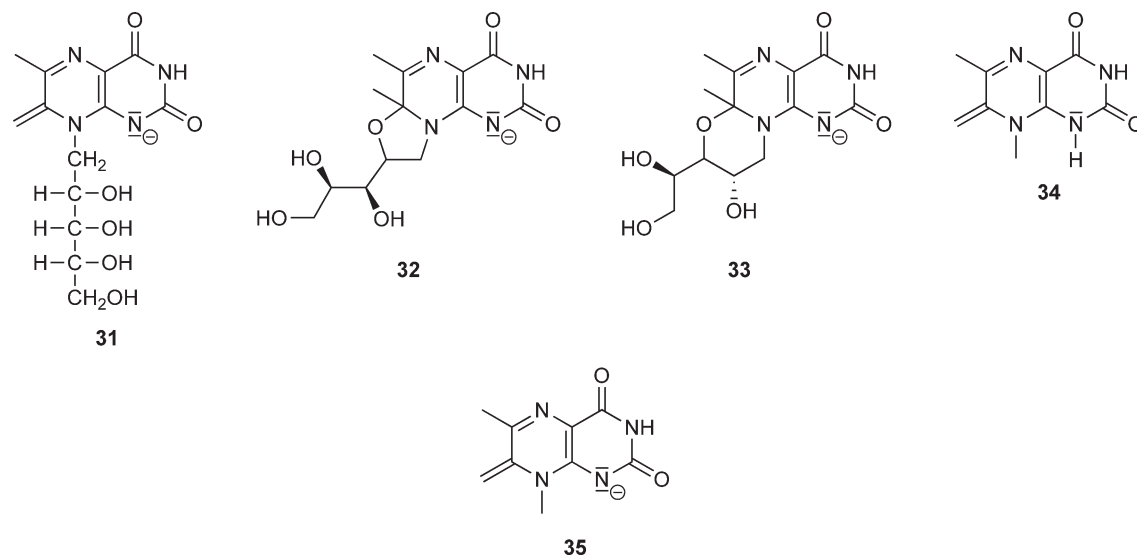
After dissociation of the native enzyme complex at elevated pH values, the  $\beta$  subunits form high molecular weight aggregates. These aggregates form a complex mixture of different molecular species, which sediment at velocities of about 48 and 70 S. Electron micrographs show hollow, spherical vesicles with diameters of about 29 nm.<sup>140</sup>

These irregular capsids can be converted back into the strictly icosahedral native structure with a diameter of 150 Å by ligand-driven renaturation.<sup>140</sup> The stability of the large capsids can be increased by certain amino acid replacements; the structure of the mutant capsids has been investigated by small-angle X-ray scattering and electron microscopy.<sup>141</sup>

## 7.02.6 Riboflavin Synthase

The enzyme was first purified to near homogeneity (about 4000-fold) from bakers yeast.<sup>13,14</sup> The substrate of riboflavin synthase (reaction IX in **Figure 1**), 6,7-dimethyl-8-ribityllumazine (**9**), was discovered in the late 1950s as a green fluorescent spot on chromatograms of culture fluid of the flavinogenic ascomycete *A. gossypii*.<sup>5</sup> Its structure was more than superficially similar to that of riboflavin, and its occurrence in a flavinogenic organism immediately prompted the hypothesis of its implication in the biosynthesis of the vitamin. That hypothesis was rapidly confirmed by studies with cell extracts that converted the green fluorescent compound into riboflavin. Surprisingly, however, that transformation that increases the mass of the substrate by four carbon atoms and two hydrogen atoms requires neither a second substrate nor a cofactor (with the exception of the pentameric riboflavin synthases from Archaea that require  $Mg^{2+}$ ).<sup>106</sup> This apparent paradox was resolved by the discovery that the substrate 6,7-dimethyl-8-ribityllumazine serves both as donor and as acceptor of a 4-carbon unit in a highly unusual dismutation reaction whose mechanism is still incompletely understood.<sup>9,142–144</sup> The second product of that reaction, 5-amino-2,4(1*H*,3*H*)-pyrimidinedione (**6**), can be reutilized as a substrate of lumazine synthase. Most surprisingly, the complex reaction can proceed without catalysis under mild conditions; boiling of 6,7-dimethyl-8-ribityllumazine in aqueous neutral or acidic solution under an inert atmosphere affords riboflavin.<sup>8,10,11</sup>

It is generally assumed that the unusual CH acidity of 6,7-dimethyl-8-ribityllumazine is important for the understanding of the enzyme-catalyzed as well as the uncatalyzed reaction. Briefly, the position 7 methyl groups of 8-substituted 7-methylumazine have p*K* values in the range of 8–9.<sup>145</sup> The anions are characterized by their 7-exomethylene structure. In cases where the position 8 substituent carries hydroxyl groups in the 2' or 3' position, a nucleophilic attack of C-7 of the exomethylene anion by the hydroxyl group is conducive to the formation of tricyclic systems (**Figure 9**).



**Figure 9** Anionic species of 6,7-dimethyl-8-ribityllumazine (**31–33**), 6,7,8-trimethylalumazine (**34**), and anionic form of **34** (**35**).<sup>146–148</sup>

Specifically, an alkaline aqueous solution of 6,7-dimethyl-8-ribityllumazine contains a mixture of at least five anionic species, namely the exomethylene-type anion with an open chain structure of the position 8 substituent and two diastereomers each of the 5-ring and 6-ring forms arising by intramolecular reaction of the 2' and 3' hydroxyl groups. The hypothesis of an involvement of one of the tricyclic species as intermediates in the biosynthesis of riboflavin<sup>149</sup> is not supported by more recent data. The exomethylene form appears likely to be a late intermediate in the mechanism of lumazine synthase as documented by stopped flow analysis.<sup>113,114</sup> Moreover, the exomethylene form (**31**) has been proposed as an early intermediate of the riboflavin synthase reaction. This is supported by the early observation that riboflavin synthase accelerates proton exchange at the position 7 methyl group.

An isotope effect of 5.0 has been found for [ $6\alpha\text{-}^2\text{H}_3$ ]6,7-dimethyl-8-ribityllumazine. Hence, the release of a proton(s) from the position 6 methyl group could involve a relatively high energy barrier.<sup>14</sup>

Early studies also showed that riboflavin synthase catalyzes a strictly regiospecific transfer of a 4-carbon unit between the two substrate molecules.<sup>10,11,142</sup> Initial experiments with  $^2\text{H}$ -substituted **9** showed that C-6 and C-7 of the donor molecule are connected to C-7 $\alpha$  and C-6 $\alpha$  of the acceptor molecule, respectively. These findings were later confirmed with  $^{13}\text{C}$ -labeled substrate samples.<sup>144</sup> Notably, it was also shown that the uncatalyzed reaction follows the same regiospecificity.<sup>8,10,11</sup> This suggested that the two substrate molecules should have an antiparallel arrangement at the active site. Since this could implicate a  $c_2$ -symmetric or pseudo- $c_2$ -symmetric arrangement of the two substrate molecules at the active site, it was speculated that the enzyme, *per se*, could have  $c_2$  or pseudo- $c_2$  symmetry, thus imposing the reaction geometry on the substrate.<sup>104</sup> This hypothesis was confirmed for the riboflavin synthases from plants, eubacteria, and fungi,<sup>137,138</sup> whereas the riboflavin synthases of Archaea achieve a pseudo- $c_2$ -symmetric arrangement of the substrates in a protein environment that is devoid of  $c_2$  symmetry properties (see below).<sup>108</sup>

While the stereochemically favored addition of a 2' or 3' hydroxyl group to C-8 of 6,7-dimethyl-8-ribityllumazine has been documented in considerable detail, the formation of a position 7 hydrate has only been observed after the introduction of fluorine to increase the electronegativity.<sup>150,151</sup> Indeed, the bistrifluoromethyl-8-ribityllumazine is so stable that the two diastereomeric forms A and B show no evidence of racemization whatsoever. Notably, a hypothetical covalent hydrate of 6,7-dimethyl-8-ribityllumazine has been proposed as an intermediate at the donor site of riboflavin synthase.<sup>149</sup>

Single-turnover kinetic experiments provided evidence for a transient species with an absorption maximum at 312 nm.<sup>152</sup> The intensity of that transient was increased by a S41A mutation. Rapid quenching of reaction



mixtures containing the mutant protein 6,7-dimethyl-8-ribityllumazine afforded a compound with absorption maxima at 256, 306, and 412 nm (pH 3.4) that was designated Compound Q.<sup>153,154</sup>

<sup>13</sup>C NMR spectroscopy of multiply <sup>13</sup>C-labeled Compound Q samples (obtained by rapid quench experiments with various <sup>13</sup>C-labeled samples of 6,7-dimethyl-8-ribityllumazine) revealed the pentacyclic structure **36** that is clearly an adduct of two 6,7-dimethyl-8-ribityllumazine (**9**) molecules (**Figures 10 and 11**).<sup>156</sup> The optical absorption of the compound is well in line with the presence of a lumazine motif and a pyrimidine motif.

Compound Q can be converted by riboflavin synthase into a mixture of 6,7-dimethyl-8-ribityllumazine, riboflavin, and 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione.<sup>153,154</sup> This is best explained by the hypothesis that Compound Q is an intermediate of the enzyme-catalyzed reaction than can undergo either a forward reaction affording riboflavin (**10**) and **6** or a reverse reaction affording 6,7-dimethyl-8-ribityllumazine (**9**) with similar reaction velocity. Moreover, it was shown by presteady-state kinetic analysis that Compound Q fulfills the criteria for a kinetically competent reaction intermediate.<sup>154</sup> As described in more detail below, it can be deduced from X-ray structure data that the two lumazine components in the pentacyclic adduct **36** must be fused in *cis*.<sup>137</sup>

The experimentally observed intermediate **36** can be incorporated into earlier mechanistic proposals<sup>8,10,11</sup> without difficulty.<sup>153,154</sup> The reaction sequence shown in **Figure 10** suggests that the reaction is initiated by a nucleophilic attack of an exomethylene anion (**31**) on a covalent 4a adduct between a lumazine molecule and an amino acid side chain or a covalent 4a hydrate.

Following a tautomerization step that affords a position 6 exomethylene structure, the central ring of the pentacyclic adduct could be closed.

The proposed mechanism for the formation of Compound Q is not without difficulties. The tautomer **31** with the position 6 exomethylene structure (an exocyclic enamine) would be expected to have a higher free energy than its hypothetical precursor, which has a more extended conjugated  $\pi$  electron system. Moreover, if we assume that the spontaneous reaction proceeds via the same mechanism as the enzyme-catalyzed reaction, the fact that the spontaneous reaction can proceed under acidic conditions is hardly in favor of the proposed anion intermediate **31**. However, a mechanism has been proposed by Beach and Plaut<sup>8</sup> that starts with a protonated substrate molecule.

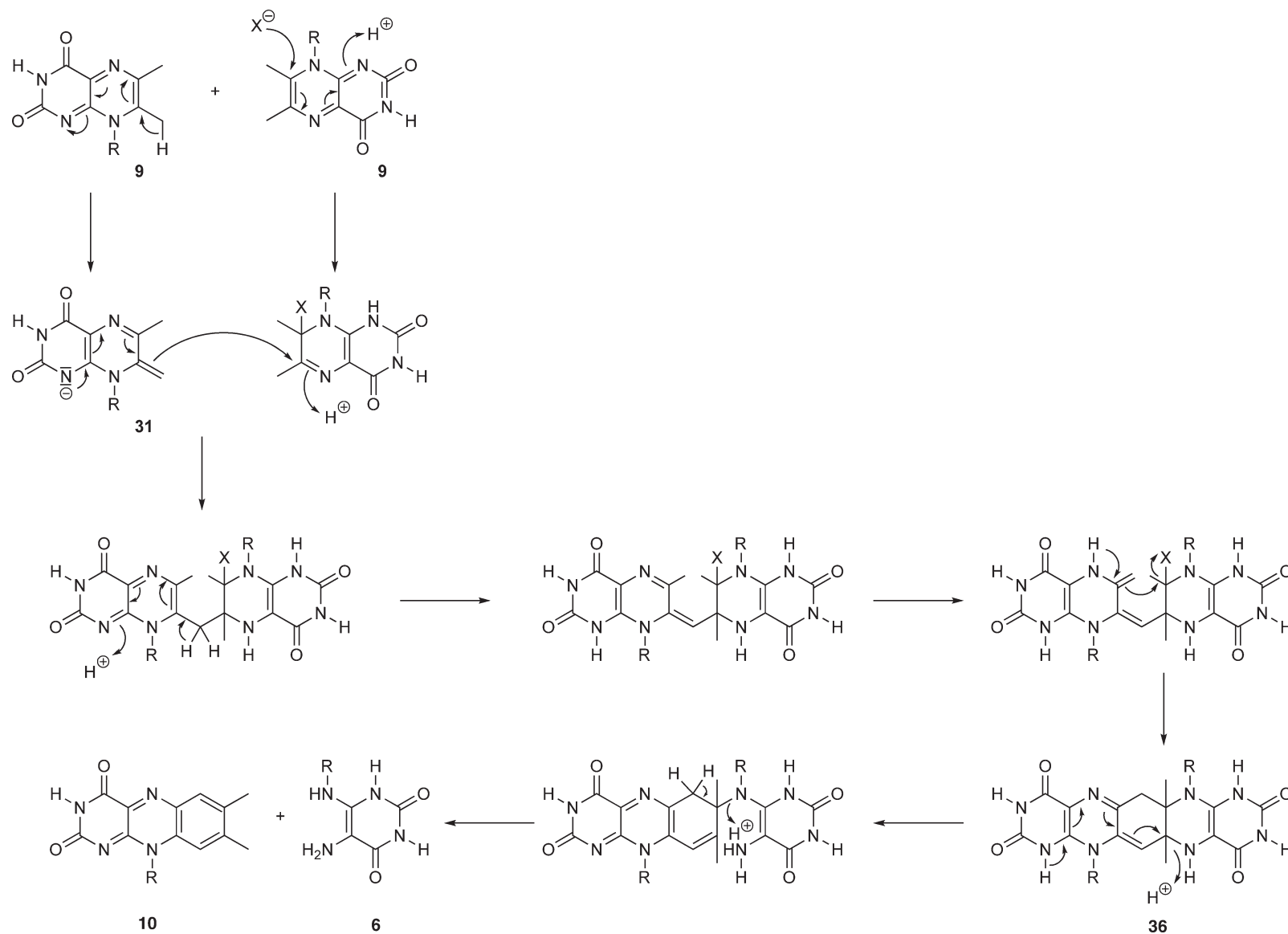
In contrast to the mechanistic difficulties with the formation of Compound Q, its cleavage under formation of the products riboflavin (**10**) and **6** by a sequence of two elimination steps appears straightforward.

The sequence of riboflavin synthase of *B. subtilis*, the first that became available, was determined at the protein level by Edman sequencing and, independently, by DNA sequencing.<sup>17,104,157</sup> In line with the hypothesis that the enzyme could provide a pseudo-*c*<sub>2</sub>-symmetric environment, it was gratifying to observe that the 215 amino acid peptide shows strong internal sequence homology; specifically, 26 amino acids in the internal sequence repeat of the *B. subtilis* enzyme are identical and 23 are similar (**Figure 12**).<sup>104</sup>

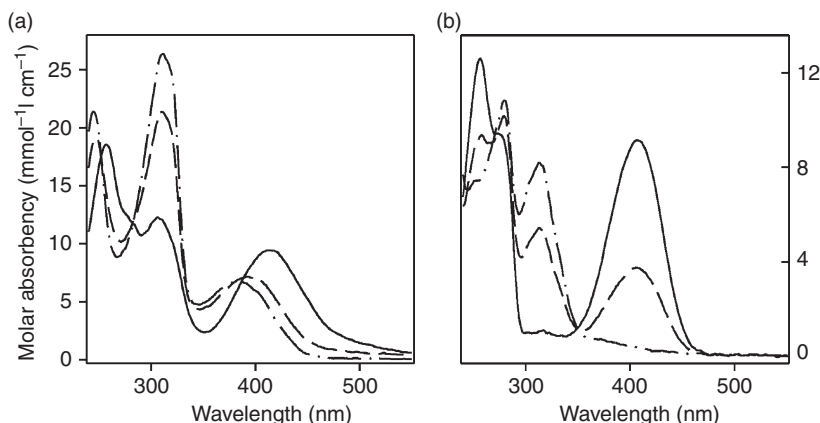
Notably, however, the sequence similarity does not extend to a segment of about 22 amino acids at the C-terminus. Meanwhile, that internal homology has been found in more than hundred riboflavin synthase sequences of bacterial, fungal, or plant origin. It should be noted, however, that the riboflavin synthases of Archaea have no intramolecular sequence similarity (see below).<sup>107,108</sup>

The N-terminus of riboflavin synthases from eubacteria, fungi, and plants is marked by an extraordinary degree of sequence conservation. The canonical amino acid sequence is MFTG, and no N-terminal extensions beyond that motif are known. The replacement of the phenylalanine residue by any amino acid except tyrosine nullifies the catalytic activity of the protein.<sup>158</sup>

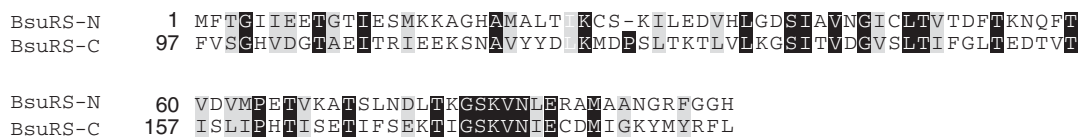
The N-terminal domain of riboflavin synthase can be expressed as a stable protein in recombinant *E. coli* strains.<sup>159</sup> Surprisingly, the recombinant protein forms a dimer with strict *c*<sub>2</sub> symmetry (as opposed to the pseudo-*c*<sub>2</sub> symmetry relations that were described for the full-length protein). The recombinant protein binds 6,7-dimethyl-8-ribityllumazine as well as several analogues and riboflavin with relatively high affinity. The three-dimensional structure of the artificial dimer has been determined by X-ray crystallography and by NMR structure analysis.<sup>160,161</sup> The intersubunit interface differs from the interface areas that are involved in pseudo-*c*<sub>2</sub> symmetry involved in the two types (intrasubunit and intersubunit) of complexation between the N-terminal and the C-terminal domain in the full-length protein. Attempts to express the C-terminal domain afforded only rather unstable protein that could not be studied in closer detail (**Figure 12**).<sup>159</sup>



**Figure 10** Hypothetical reaction mechanism of riboflavin synthase: **9** (D), donor lumazine molecule; **9** (A), acceptor lumazine molecule; X, proposed nucleophile, which neutralizes the carbonium center at C-7 of **9** (A) and enables carbanion attack at C-6 of **36** by the 7-exomethylene carbon of **9** (D); R, ribityl chain.<sup>155</sup>



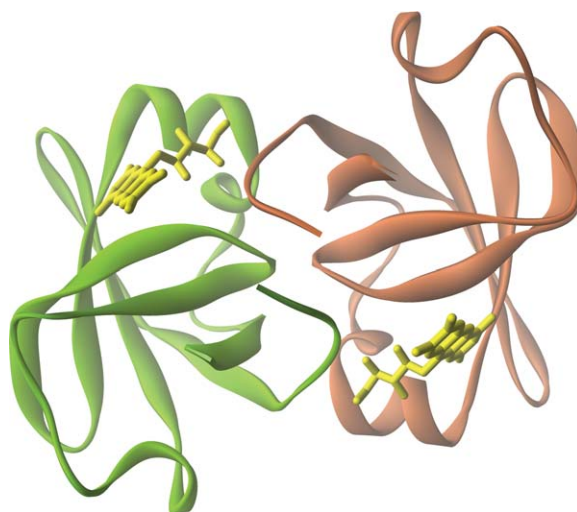
**Figure 11** Optical spectra of Compound Q (a) and 6,7-dimethyl-8-ribityllumazine (b) at pH 3.4 (solid lines), pH 8.8 (dashed lines), and pH 11.1 (dotted lines).<sup>153</sup>



**Figure 12** Internal sequence homology of riboflavin synthase from *Bacillus subtilis*.<sup>104</sup>

When X-ray structures of riboflavin synthases from *E. coli* and from *Schizosaccharomyces pombe* became available,<sup>137,138</sup> they confirmed certain expectations that had been formulated earlier on the basis of mechanistic and sequence arguments but also provided unexpected surprises.

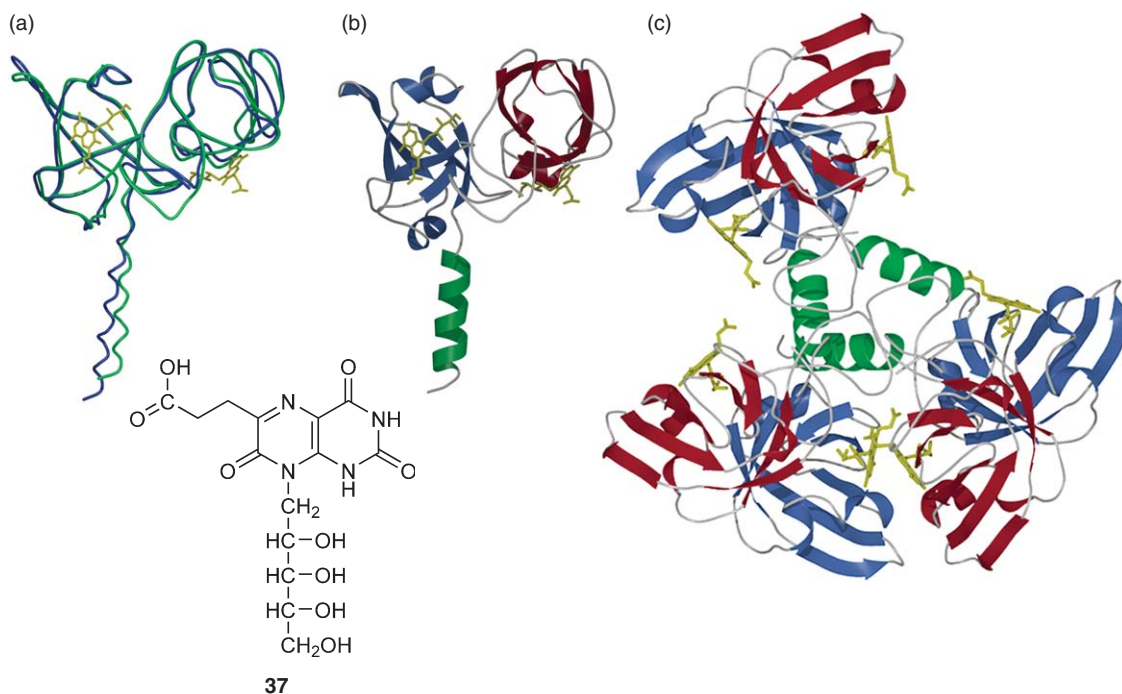
As predicted, each subunit of the homotrimer folds into two closely similar domains that are connected by a short linker sequence. Also in line with the earlier hypotheses, the two domains of each subunit are related by pseudo- $c_2$  symmetry. However, contrary to expectation, the *E. coli* protein is devoid of trigonal symmetry. In the *E. coli* protein, the N-terminal domain of one subunit and the C-terminal domain of a second subunit form a pseudo-symmetric pair, and the single (!) active site is located at the interface of that particular domain pair.<sup>137,138</sup> The evidence for that conjecture is to a significant part based on a comparison between the X-ray structures of the enzymes from *E. coli* and *S. pombe* (Figure 13).<sup>137,138</sup> Notably, the *E. coli* protein was crystallized



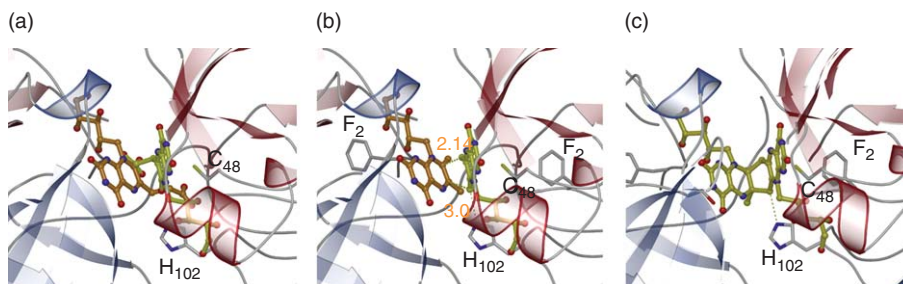
**Figure 13** Structure of the N-terminal domain dimer of riboflavin synthase from *Escherichia coli* (PDB entry #1PKV) with bound riboflavin viewed along the twofold noncrystallographic symmetry axis.

without low molecular weight ligands. The *S. pombe* protein was crystallized with the substrate analogue 6-carboxyethyl-7-oxo-8-ribityllumazine (**37**), but the crystal structure shows only the structure of the monomer because the trimeric quaternary structure (that has been shown to exist in solution by hydrodynamic experiments)<sup>162</sup> was obviously broken under the crystallization conditions (with 2-methyl-2,4-pentanediol (MPD) as the precipitating agent). However, the substrate 6,7-dimethyl-8-ribityllumazine (**9**) can be modeled into the structure of the *E. coli* enzyme by emulation of the intermediate analogue in the crystal structure of the *S. pombe* enzyme (Figures 14(b) and 15).

The model shows unequivocally that the N-terminal domain hosts the acceptor substrate-binding site, and that the linkage between the two substrate moieties should result in syn linkage of the ring systems of donor and the acceptor molecule.<sup>137</sup> Moreover, the absolute stereochemistry can be deduced from the active site topology (cf. Figure 14(c)); the arguments in favor of a trans linkage<sup>100</sup> have been addressed elsewhere and will not be discussed here in detail.



**Figure 14** X-ray structures of riboflavin synthases from *Escherichia coli* (PDB entry #118D) and *Schizosaccharomyces pombe* (PDB entry #1KZL). (a) Diagram of the superposition of one subunit of riboflavin synthase from *E. coli* (blue) and *S. pombe* (green). (b) *S. pombe* riboflavin synthase monomer with bound 6-carboxyethyl-7-oxo-8-ribityllumazine (**37**, yellow). (c) Trimeric model of *S. pombe* riboflavin synthase with bound 6-carboxyethyl-7-oxo-8-ribityllumazine (yellow).<sup>137,138</sup>



**Figure 15** Active site of riboflavin synthase of *Schizosaccharomyces pombe* formed by two adjacent monomers with bound 6-carboxyethyl-7-oxo-8-ribityllumazine (**37**). (a) The ligand bound to the N barrel (red) is drawn in yellow, whereas the 6-carboxyethyl-7-oxo-8-ribityllumazine in the adjacent C barrel (blue) is shown in dark yellow. (b) Proposed binding of 6,7-dimethyl-8-ribityllumazine at the active site. (c) Model for the pentacyclic reaction intermediate.<sup>137</sup>

The N-terminus of the protein is located close to the heterocyclic ring system of the acceptor substrate; the earlier finding that the deletion or replacement of phenylalanine 2 (except the replacement by tyrosine) results in complete loss of catalytic activity.<sup>158</sup> Tentatively, it is assumed that the phenylalanine residue is crucial for the fine-tuning of the substrate topology. Similarly, the decrease of catalytic activity by the replacement of serine 41 and histidine 102 may be caused by indirect conformational influences on substrate topology rather than direct participation of the amino acid residues in catalysis. No evidence has been obtained for covalent catalysis (formation of a covalent adduct of an amino acid side chain at C-4a of the acceptor substrate as proposed in early mechanistic publications). Moreover, no specific amino acid side chain could be shown to participate directly in noncovalent catalysis, other than by fine-tuning of the topology. This may be surprising in light of the extraordinary complexity of the reaction trajectory. On the other hand, it should be noted that the formation of riboflavin can proceed in aqueous solution without any catalysts. In fact, the role of the enzyme may be limited to controlling and fine-tuning the reaction topology; this would implicate that, as in the case of lumazine synthase, the enzyme operates predominantly through modulation of the activation entropy. It should be noted that the rate of the catalyzed reaction is modest (cf. **Table 1**) and the catalytic acceleration, by comparison with the uncatalyzed reaction, is modest. It is obvious that this reaction sequence is still riddled with puzzles, even after five decades of investigation.

As if the mechanisms of lumazine synthase and riboflavin synthase were not complex enough, *per se*, an additional layer of complexity is introduced by the existence of ‘heavy riboflavin synthase’, a complex that comprises a homotrimeric riboflavin synthase module that is enclosed in the central cavity of an icosahedral lumazine synthase capsid.<sup>97,133,140</sup>

Up to now, this 1 MDa protein has been observed only in Bacillaceae. Computer modeling has confirmed that the central cavity of the icosahedral capsid is large enough to accommodate the riboflavin synthase trimer (**Figure 16**). However, there are no corresponding symmetry properties between the two modules that could suggest a potential orientation. The molecular structure is not known in closer detail. The import and export of substrates and products is also an open problem. It is hard to see how the channels in the capsid wall could allow the passage of 6,7-dimethyl-8-ribityllumazine, and, even less, of riboflavin. Dynamic fluctuations of the capsid structure have been proposed but remain speculative.



**Figure 16** Computer-generated model of a heavy riboflavin synthase complex. The capsid was generated using the coordinates from the lumazine synthase 60-mer of *Bacillus subtilis* (PDB entry #1RVV) and from the riboflavin synthase trimer of *Escherichia coli* (PDB entry #118D).



Intermediate channeling resp. substrate channeling has been reported for heavy riboflavin synthase.<sup>139</sup> At low substrate concentrations, 6,7-dimethyl-8-ribityllumazine formed *in situ* by the lumazine synthase capsid is more rapidly converted by the resident riboflavin synthase as compared to 6,7-dimethyl-8-ribityllumazine in the bulk solvent. Retarded evasion of the newly formed lumazine synthase product from inside the capsid has been proposed as a tentative explanation.

### 7.02.7 Pentameric Riboflavin Synthases of Archaea

*In vivo* studies with <sup>13</sup>C-labeled precursors established that Archaea generate riboflavin via 6,7-dimethyl-8-ribityllumazine. A protein with riboflavin synthase activity was purified from *Methanothermobacter thermoautotrophicus*, and the cognate gene was cloned by a marker rescue strategy.<sup>106</sup> The peptide of 154 amino acid residues showed no intramolecular sequence similarity and no sequence relationships with the riboflavin synthases of eubacteria and eukarya.

While the riboflavin synthases of Archaea are devoid of similarity with those of eubacteria and eukaryotes, they have significant sequence similarity with 6,7-dimethyl-8-ribityllumazine synthases (Figure 17).<sup>107</sup>

The riboflavin synthase of *M. jannaschii* shows 20 and 17% sequence identity with the lumazine synthases of *M. jannaschii* and *E. coli*, respectively.<sup>107</sup>

Single-turnover kinetic analysis revealed a transient with similar absorption as that of Compound Q, the pentacyclic intermediate of eubacterial riboflavin synthase (Figure 19).<sup>156</sup>

The compound could be isolated after rapid quenching of reaction mixtures.<sup>156</sup> The <sup>13</sup>C NMR spectrum was similar, but not identical with that of Compound Q. Moreover, the compound (designate Q') produced by the *M. jannaschii* enzyme could serve as substrate for the archaeal enzyme by which it had been produced, but not for eubacterial riboflavin synthase. Vice versa, Compound Q produced by eubacterial riboflavin synthase could serve as a kinetically competent substrate for the eubacterial enzyme (see above) but could not be used by the enzyme from *M. jannaschii*.

The hypothesis, based on the kinetic findings, that Compounds Q and Q' are diastereomers was confirmed by circular dichroism (CD) spectroscopy (Figure 18) and X-ray structure analysis of pentameric riboflavin synthase from *M. jannaschii* in complex with the substrate analogue, 8-ribityl-6,7-dioxo-8-ribityllumazine (38).

Although the protein environment has no local symmetry properties, the heterocyclic moieties of two substrate analogue molecules are arranged with pseudo-*c*<sub>2</sub> symmetry in the relatively large active site cavity. The topology predicts a pentacyclic adduct whose heterocyclic part is the mirror image of the heterocyclic moiety of Compound Q (Figure 19).

The substrate molecule that serves as 4-carbon acceptor in the pentameric riboflavin synthase is almost in the same position as riboflavin bound to lumazine synthase (Figure 20).

6,7-Dimethyl-8-ribityllumazine synthase and the pentameric, archaeal riboflavin synthase appear to have diverged from a common ancestor at an early time point in the evolution of Archaea. Genes that may specify riboflavin synthases of the eubacterial type in a few Archaea may have been acquired relatively recently by horizontal gene transfer.

```

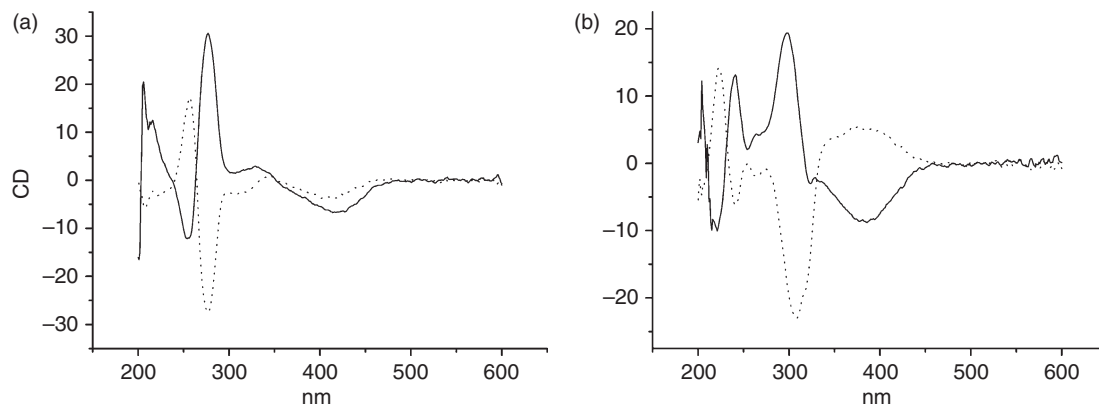
MthRS   ---MKVGCIDTTFAR--YDMGGAIDELKKHATGKIKTIRRTVPGIKDLPVACKKLIIEEE
MjaRS   --MTKKVGIIVDTTFAR--VDMASIAIKKLELSEPNIKIIRKTVPGIKDLPVACKKLEEE
MjaLS   MVLMLVNLGFVITAEFNRDITYMMEKVAEELHAEFLGATVVKYKIVVPGVFDMLPLAVKKLLEKD

MthRS   GCEVMVALG--MPGPPEKDKVCAHEASTGLIQALMTNTHILEVFVHDEDEDDPEDLKVL
MjaRS   GCDIVMALG--MPGKA EKDKVCAHEASTGLMLAOLMTNKHILEVFVHDEDAKDDKELDVL
MjaLS   DVDAVVTIGCVIEGETEHEDEIVVHNAARKIADLALQYDKPVTLGISGPGMTRLQAQERVD

MthRS   ADNRRAREHAQNLIMMLFRPERLTRDAGMGMRECKPDVGPGL--
MjaRS   AKRRAREHAENVYLLKPEVLTMRACKGLRQCFEDAGPARE
MjaLS   YGKRAVEAAVKMVKRLKALE-----

```

**Figure 17** Sequence comparison of lumazine synthase and archaeal-type riboflavin synthase. MthRS, riboflavin synthase of *Methanothermobacter thermoautotrophicus* delta H (accession no. Q59587); MjaRS, riboflavin synthase of *Methanocaldococcus jannaschii* (accession no. Q58584); MjaLS, lumazine synthase of *M. jannaschii* (accession no. Q57751). Identical residues are shaded in black.<sup>107</sup>



**Figure 18** Circular dichroism spectra of Compound Q ( $147 \mu\text{mol l}^{-1}$ , dotted line) and Compound Q' ( $181 \mu\text{mol l}^{-1}$ , solid line): (a) at pH 1.0 and (b) at pH 11.0.<sup>156</sup>

X-ray analysis identified a  $c_5$ -symmetric homopentamer whose folding pattern and quaternary structure closely resembles that of pentameric lumazine synthase (Figure 21).<sup>108</sup>

### 7.02.8 Lumazine Protein

Paralogues of eubacterial riboflavin synthase (designated as lumazine protein, yellow fluorescent protein, and blue fluorescent protein) without enzymatic activity have been isolated from several luminescent bacteria.<sup>163–166</sup> These proteins are highly fluorescent and are believed to modulate to serve as optical transponders that modulate the spectral characteristics of bacterial bioluminescence.<sup>167</sup> Energy is believed to be transferred from activated bacterial luciferase to the fluorescent proteins by radiation-less transfer.

The fluorescent proteins share considerable sequence similarity with riboflavin synthases over most of their length, that is, with the exception of the C-terminus of riboflavin synthase that is missing in the fluorescent proteins (Figure 22).

The C-terminal segment of riboflavin synthase is believed to provide the driving force for trimer association via the formation of a triple helix. The absence of the segment in the fluorescent proteins is probably the reason for their monomeric structure.

The fluorescent proteins share with riboflavin synthase the intramolecular sequence similarity<sup>168</sup> and, like riboflavin synthase, fold into two closely similar domains (Figure 23).

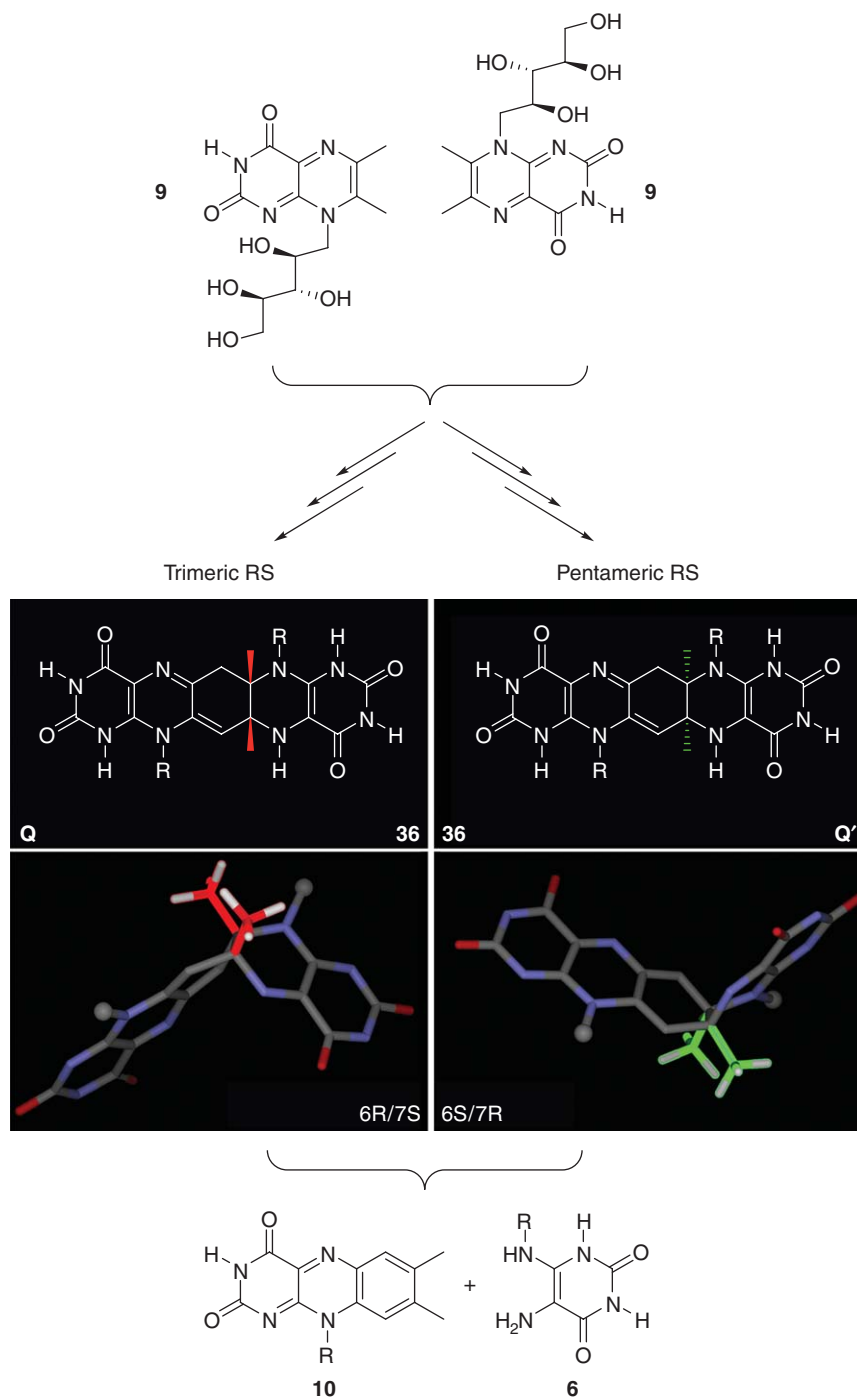
The presence of two closely similar folding domains and the overall similarity to the riboflavin synthase fold has been confirmed recently by X-ray structure analysis.<sup>170</sup> Despite that two-domain architecture, the fluorescent proteins bind only one ligand molecule per protein molecule, whereas riboflavin synthases have two ligand-binding sites per protomer.<sup>171,172</sup> A recent mutagenesis study has located the single binding site to the N-terminal domain.<sup>173</sup>

### 7.02.9 Regulation of Riboflavin Biosynthesis

In *B. subtilis*, the expression of riboflavin biosynthesis enzymes can be regulated over a wide dynamic range (least 30-fold).<sup>32,133</sup> Riboflavin oversynthesis can be caused by mutations in a region designated *ribO* that is located at the 5' end of the *rib* operon and in a gene designated *ribC* that is not linked to the *rib* operon. The *ribC* gene was later shown to specify a bifunctional riboflavin kinase/FAD synthetase.<sup>174</sup> Overexpression caused by *ribC* mutations in *B. subtilis* can be suppressed by mutations in the *ribR* gene, which specifies a dihydroriboflavin kinase.<sup>175–178</sup> These findings indicated that the active regulatory species is not riboflavin *per se* but one of its coenzyme forms.

FMN is now believed to regulate flavin biosynthesis enzyme expression in *B. subtilis* and other bacteria by binding to an RFN motif located upstream of the *ribG* gene (designated *ribO*) that acts as a riboswitch

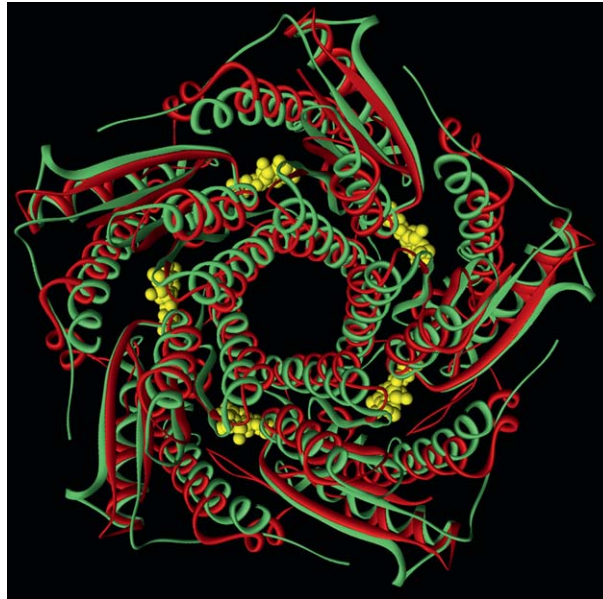




**Figure 19** Stereochemistry of 6,7-dimethyl-8-ribityllumazine conversion into riboflavin catalyzed by eubacterial (trimeric) and archaeal (pentameric) riboflavin synthase (R, ribityl).

enabling transcription termination/antitermination (Figure 24).<sup>179,180</sup> Mutations within the RFN element inactivated FMN-based regulation of the *rib* operon.<sup>181,182</sup> It was reported that the RFN domain serves as the receptor for a metabolite-dependent riboswitch that directly binds FMN in the absence of proteins, controlling gene expression by causing premature transcription termination of the *rib* operon. Besides this, it has also been shown that artificial RNA sequences (aptamers) can bind FMN with very high affinity.<sup>183</sup>





**Figure 21** Structural comparison of pentameric riboflavin synthase from *Methanocaldococcus jannaschii* (red, PDB entry #2B99) and lumazine synthase from *Schizosaccharomyces pombe* with bound riboflavin (yellow, green, PDB entry #1KZL).<sup>108,137</sup>

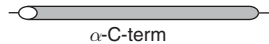
(a)

```
LUXP-PHOPO 1 MFKGIVQGVCI IKKISKND D TORHGIT FPKDILDSV E KDTV MLVNGCS V TVVRITGDVVY
LUXY-VIBFI 1 MFKGIVEGIGI I EKIDITYD LDKYAIR F PENMNLNGI K KESSI MFNGC E LTVTVSVNSNIVM
LUXP-PHOLE 1 MFKGIVQGRGVIRS I SKSEDSQRHGIA F PEGM FQLVDV D TVMLVNGCS L TVVRILGDMVY
RISA-SCHPO 1 MFTG L VEA I G V V K D V Q G T I D N G F A M K I E A P O I L D D C H T G D S I A V N G T C L T V T D F D R Y H F T
```

```
LUXP-PHOPO 61 FDI--DQALNTTFRKLEVCNKVNLEVRP C FGSLLCKGAL TGNIKGVA TVDNI T E E D D L L
LUXY-VIBFI 61 FDI FEKEARKLDTFREYKVGDRVNLGTFPKFGAASGGHILSARHSCVASIIEITENEDVYQ
LUXP-PHOLE 61 FDI--DQALGTTTFDGLK E G D Q V N L E I H P K F G E V V C R G C L T G N I K G T A L V A A I E N D A G F
RISA-SCHPO 61 VGTA-PE S L R L T N L G Q C K A G D P V N L E R A V L S S T R M G G H F V Q G H V D T V A E L V E K K Q D G E A I
```

```
LUXP-PHOPO 119 KVV I K I P K D L I - E N I S S E D H I G I N G V S N S I E E V S N D I I C I N Y P K N L S I T T N L G L E T C S E
LUXY-VIBFI 121 Q M W I Q I P E N F T - E F L I D K D Y I A V D G I S L T I D T L K N N Q P F I S L P L K T A Q N T N M K W R K K G D K
LUXP-PHOLE 119 S V L I D I P K G L A - E N L T V K D D I G I D G I S L P I T D M S D S I I T L N Y S R D L L A S T N I A S L A K D V K
RISA-SCHPO 120 D F T F R P R D P F V L K Y I V Y K S V I A L D G T S L I T H V D D S T F S T M M I S Y T Q S K V I M A K K N V G D L
```

```
LUXP-PHOPO 178 V N V E L L N V S N E W -----
LUXY-VIBFI 180 V N V E L S N K I N A N Q C W -----
LUXP-PHOLE 178 V N V E L L N E W -----
RISA-SCHPO 180 V N V E V D Q I G K Y T E K L V E A H I A D W I K K T Q A
```



(b)

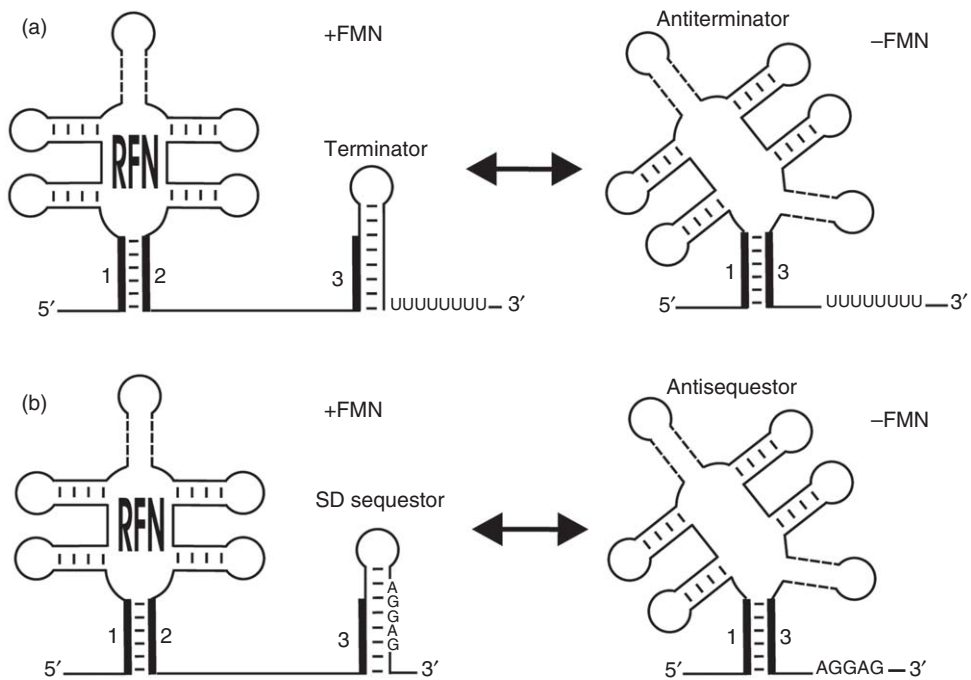
```
LUXP-PHOLE-N 1 MFRGIVQGRGVIRS I SKSEDSQRHGIA F PEGM FQLVDV D TVMLVNGCS L TVVRILGDMVY
LUXP-PHOLE-C 97 GLTGNIKGTALVAA I EENDAGFSVL I D I P K G L A E N L T V K D D I G I D C I S L P I T D M S D S I I T
```

```
LUXP-PHOLE-N 61 FDID-QALGTTTFDGLKEGDQVNL E I H P K F G E V V G R G
LUXP-PHOLE-C 157 LNYSRDLLAS T N I A S L A K D V K V N V E I L N E W
```

**Figure 22** Amino acid sequence similarity between fluorescent proteins of photobacteria and riboflavin synthase from *Schizosaccharomyces pombe* (a) and N- and C-terminal parts of LUXP-PHOLE (b). LUXP-PHOLE, lumazine protein of *P. leiognathi* (accession number Q06877); LUXP-PHOPO, lumazine protein of *Photobacterium phosphoreum* (accession number P25082); LUXY-VIBFI, yellow fluorescent protein of *Vibrio fischeri* (accession number P21578); RISA-SCHPO, riboflavin synthase of *S. pombe* (accession number Q9Y7P0). Identical amino acid residues are in shadowed typeface. Red boxes indicate amino acid residues in the *S. pombe* riboflavin synthase structure that contact substrate in the N- or C-terminal domain. The cylinder represents the C-terminal  $\alpha$ -helix responsible for trimerization of riboflavin synthases.<sup>173</sup>



**Figure 23** Structural superposition of lumazine protein (green, PDB entry code: 3DDY<sup>169</sup>) and riboflavin synthase from *Schizosaccharomyces pombe* (yellow, PDB entry code: 1KZL<sup>137</sup>). The C-terminal segment of riboflavin synthase is marked by red color.



**Figure 24** The predicted mechanism of the RFN-mediated regulation of riboflavin genes: (a) transcription attenuation and (b) translation attenuation.<sup>180</sup>

## 7.02.10 Biotechnology

Riboflavin is a bulk commodity that is used predominantly for animal husbandry. Pigs and poultry require supplements of riboflavin for efficient food utilization, whereas ruminants can obtain the vitamin from their intestinal microbial symbionts. Smaller amounts of the vitamin are used for the production of drugs and as food colorant. Up to the 1990s, the vitamin was synthesized chemically from ribose that was predominantly produced by fermentation. The potential to manufacture the compound by fermentation was in principle appreciated since the early discovery of naturally occurring flavinogenic microorganisms (certain ascomycetes, yeasts, and eubacteria). However, during several decades, fermentation contributed only minor amounts of industrial riboflavin. However, several technologies based on fermentation with production strains of *B. subtilis*, *A. gossypii*, and *Candida famata* have by now replaced the chemical process.<sup>15,16</sup>

Riboflavin production with *B. subtilis* is based on a recombinant strain that has been engineered for the overexpression of the *rib* operon. Four genes of that operon specify the two bifunctional and two monofunctional proteins (*ribA*, GTP cyclohydrolase/3,4-dihydroxy-2-butanone 4-phosphate synthase; *ribB*, riboflavin synthase; *ribG*, deaminase/reductase; *ribH*, lumazine synthase). The operon also contains an unannotated hypothetical gene *ribT* whose function is unknown.<sup>17,185</sup> As mentioned above, it is still unknown how **5** is dephosphorylated (Figure 1); remarkably, however, dephosphorylation appears not to be a limiting step for the technical process.

Briefly, the production of riboflavin with *B. subtilis* could be increased by the deregulation of purine synthesis and a mutation in the gene coding for the riboflavin kinase. The *B. subtilis* process is further based on the recombinant, homologous overexpression of multiple chromosomal copies of the riboflavin operon under the control of a strong synthetic promoter. The introduction of an additional copy of the *ribA* gene has also been shown to increase the productivity of the strain.<sup>18,174,176,186</sup>

Industrially utilizable strains have been obtained by multiple rounds of mutagenetic treatment and selection of improved colonies. Inhibitor molecules like itaconate, which inhibits the isocitrate lyase in *A. gossypii*<sup>187</sup> and tubercidin (7-deaza-adenosine),<sup>188</sup> which inhibits purine biosynthesis in *C. famata* have been applied successfully for mutant selections. In *B. subtilis*, three purine analogues (8-azaguanine, decoyinine, and methionine sulfoxide plus roseoflavin have been used to obtain resistant mutant strains.<sup>185</sup> Roseoflavin-resistant *B. subtilis* mutants were shown to contain mutations in regulatory genes and/or in the 5' region of the *rib* operon.

## 7.02.11 Riboflavin Biosynthesis Genes as Potential Drug Targets

Flavocoenzymes are indispensable in all cells. While vertebrates depend on nutritional sources (including the microbial gut flora of ruminants), many pathogenic bacteria are equipped for endogenous synthesis. Moreover, numerous pathogens cannot absorb riboflavin (or its cofactor forms) easily from the environment; thus, riboflavin-deficient mutants of *Escherichia*, *Salmonella*, and *Candida* strains require very high concentrations for growth. Hence, enzymes of the riboflavin biosynthetic pathway are believed to be potential targets for anti-infective chemotherapy.<sup>189,190</sup>

Recently, it has also been shown by *in vivo* studies that riboflavin biosynthesis genes are virulence factors for several human pathogens including *Escherichia*, *Salmonella*, and *Haemophilus*. Specifically, the *ribB* and *ribF* genes specifying 3,4-dihydroxy-2-butanone 4-phosphate synthase and a bifunctional riboflavin kinase/FAD synthetase have been reported to be essential in *Haemophilus influenzae*.<sup>191</sup> Interestingly, *ribB* is also necessary for colonization of the medicinal leech, *Hirudo verbana* by *Aeromonas veronii*.<sup>192</sup> Thus, riboflavin genes may be essential for the establishment of parasitic as well as symbiotic relationships. For unknown reasons, lumazine synthase of *B. abortus* is the dominant antigen for the formation of *B. abortus* antibodies and may have potential for vaccine development.

While the genome of *Mycobacterium leprae* has undergone extensive gene fragmentation in relatively recent times with the result that about half of the typical gene arsenal of mycobacteria were reduced to the status of pseudogenes in the transition to a strictly intracellular lifestyle, the genes specifying the enzymes of riboflavin are all intact. At least tentatively, this suggests that riboflavin biosynthesis genes are virulence factors for pathogenic mycobacteria.

On this basis, the proteins of riboflavin biosynthesis may be relevant targets for the development of novel antibacterial agents, notably with selectivity for mycobacteria, Enterobacteriaceae, or pathogenic yeasts. Since these proteins are absent in human and animal hosts, the risks of off-target toxicity should be comparatively low.

The synthesis and characterization of inhibitors for lumazine synthase and riboflavin synthase have been pursued by several research groups over extended periods, and compounds with high *in vitro* activity have been reported.<sup>193–204</sup> Unfortunately, none of the known compounds had significant *in vivo* activity. Most probably, the synthetic compounds, being structural analogues of substrates, products, or intermediates, had insufficient drug-like properties. They may have typically failed to penetrate into the bacterial cells.

Methods for high-throughput screening of riboflavin biosynthesis have been reported recently and should enable the screening of large libraries of compounds with drug-like properties.<sup>205,206</sup>

### 7.02.12 Riboflavin Kinase and FAD Synthetase

Although the early steps of the riboflavin biosynthetic pathway proceed via phosphorylated intermediates, the product is unphosphorylated riboflavin. The coenzyme forms, riboflavin 5'-phosphate (FMN, **11**) and FAD (**12**) must be obtained by enzymatic conversion of riboflavin in all organisms, no matter whether riboflavin (**10**) is obtained by synthesis or from extraneous sources.

Research on riboflavin kinase and FAD synthetase has a long history that has been addressed in an earlier review. Notably, the substrate specificities of enzymes from various sources have been studied in some detail using collections of riboflavin analogues. This chapter focuses on recent developments. In light of the general importance of these enzymes in all organisms, the number of recent studies appears low.

In eubacteria, the bifunctional riboflavin kinase/FAD synthetase type appears to be widely distributed.<sup>207,208</sup> A monofunctional kinase that is specific for dihydroriboflavin has been found in *B. subtilis*, which also has the bifunctional enzyme type that operates with oxidized flavin substrates.

In yeasts and animals, riboflavin kinase and FAD synthetase are separate enzymes. Structures have been reported for human riboflavin kinase, for riboflavin synthase of the yeast, *S. pombe*, and for FAD synthetase of *Thermotoga maritima*.<sup>127,209–211</sup>

Little is known about the conversion of riboflavin into FMN and FAD by plants. Early studies reported on the isolation of riboflavin kinases from various plant species. Recently, a bifunctional protein with a C-terminal riboflavin kinase domain and an N-terminal FMN-hydrolase domain has been cloned from *A. thaliana*.<sup>212</sup>

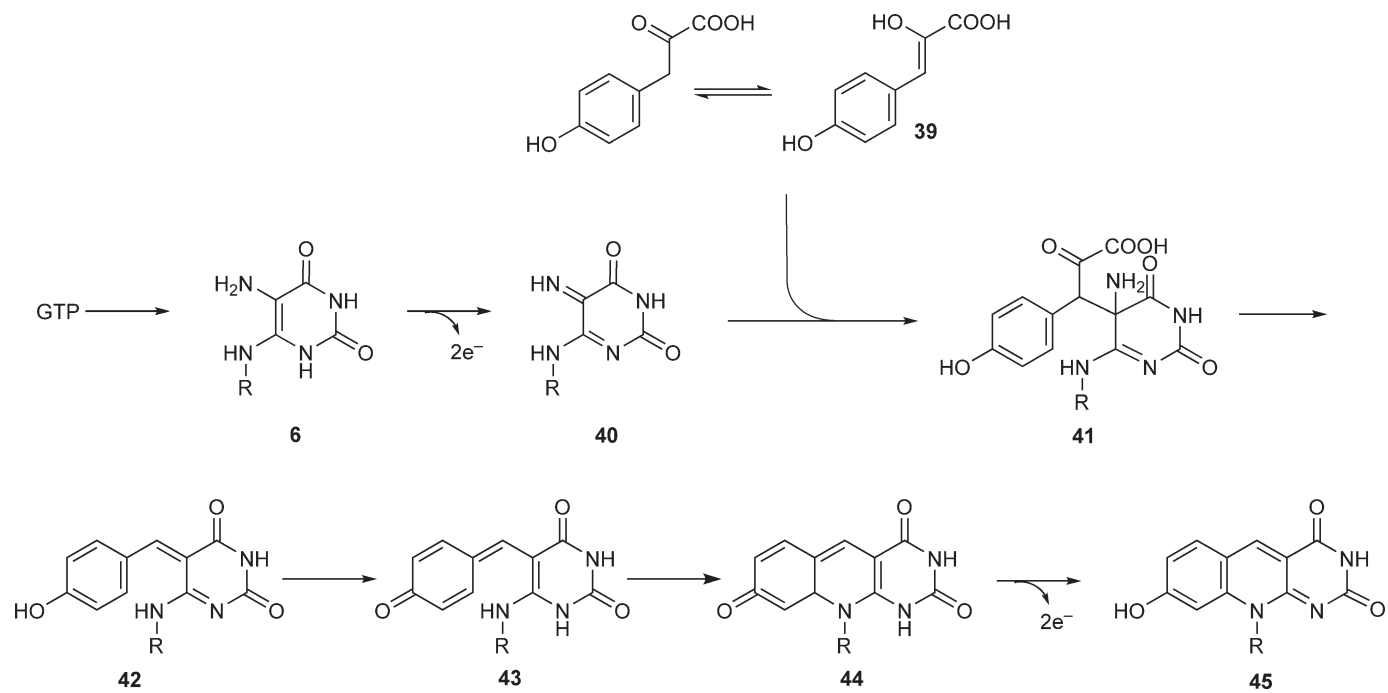
### 7.02.13 Biosynthesis of Deazaflavin Cofactors

The study of methanogenic bacteria from the archaeobacterial kingdom revealed an impressive number of hitherto unknown coenzymes, including the deazariboflavin derivative, coenzyme F<sub>420</sub>. In methanogenic bacteria, they serve an important role in the conversion of CO<sub>2</sub> and of low molecular weight organic acids into methane. More recently, deazaflavin coenzymes were also found in Streptomycetes and in mycobacteria. In Streptomycetes, they function as antenna chromophore of DNA photolyase.

The biosynthesis of the deazaflavin chromophore, 5-deaza-7,8-desmethyl-8-hydroxyriboflavin (**45**) branches off the riboflavin pathway at the level of 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione (**6**) as initially suggested on the basis of <sup>13</sup>C incorporation studies that also suggested that the benzenoid ring is a shikimate derivative. This was later confirmed by *in vivo* studies with isotope-labeled **6** and 4-hydroxyphenylpyruvate (**39**). These studies afforded the hypothetical mechanism shown in **Figure 25**. Dehydrogenation is believed to transform the pyrimidine precursor **6** into the quinoid species **40** that is then nucleophilically attacked by the enolate of 4-hydroxyphenylpyruvate.

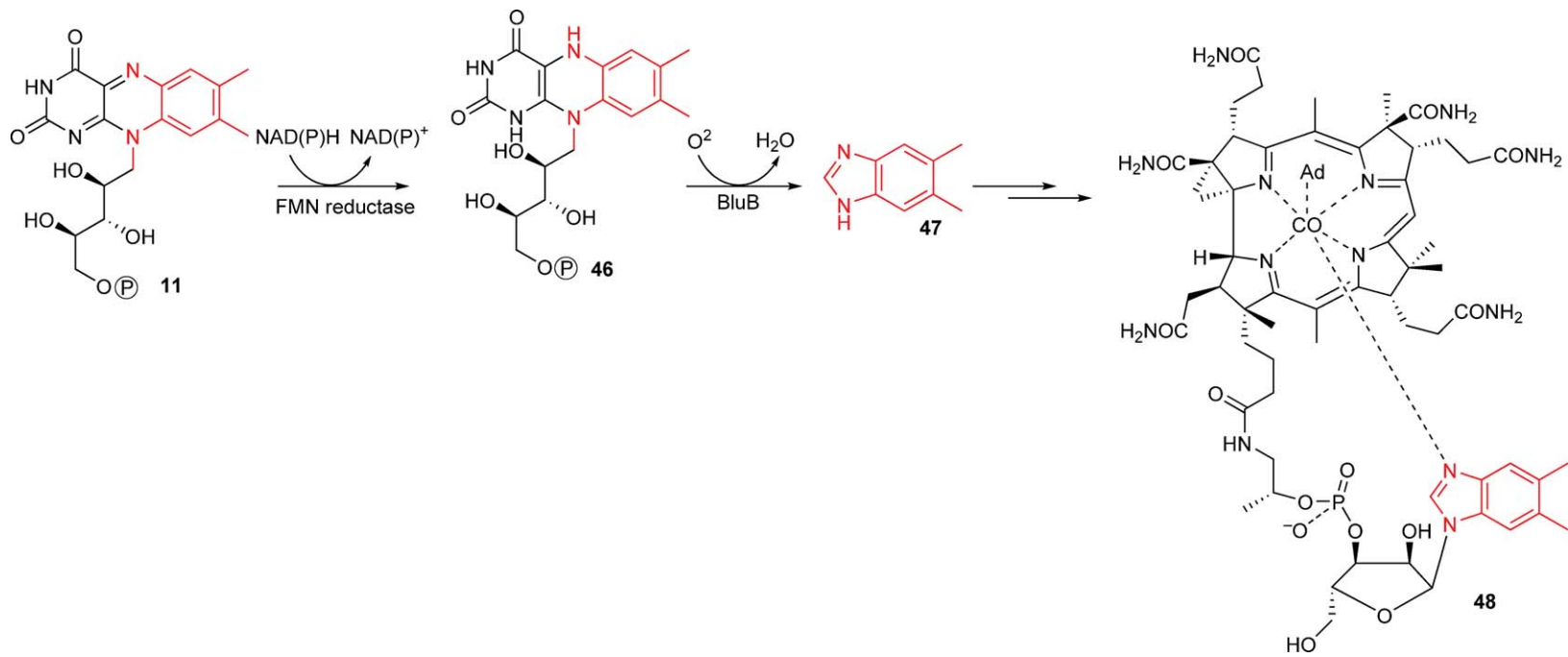
### 7.02.14 Riboflavin as Substrate for Other Biosynthetic Pathways

Isotope incorporation studies showed that the dimethylbenzimidazole moiety of vitamin B<sub>12</sub> can be biosynthesized under aerobic conditions from riboflavin (**10**). Specifically, all carbon atoms of dimethylbenzimidazole are derived from the xylene ring and the 1'-carbon of riboflavin. Recently, this transformation was shown to be catalyzed by the BluB protein. Surprisingly, the reaction requires only FMNH (**46**) and molecular oxygen. The structure of BluB protein of *Sinorhizobium meliloti* has been reported (**Figure 26**).<sup>213</sup>



**Figure 25** Biosynthesis of 5-deaza-7,8-desmethyl-8-hydroxyriboflavin.





**Figure 26** BluB catalyzes DMB production.<sup>213</sup>

The antibiotic roseoflavin (8-dimethylamino-8-demethylriboflavin) produced by *Streptomyces davavensis* is active against Gram-positive bacteria and has been used in the selection of riboflavin-producing strains.<sup>16,214,215</sup> The compound is biosynthesized from riboflavin, but the mechanistic details are as yet unknown.

## References

1. O. Warburg; W. Christian, *Naturwissenschaften* **1938**, *26*, 235.
2. O. Warburg, *Encyclopaedia Britannica [Online]* **2008**. <http://www.britannica.com>.
3. A. L. Demain, *Annu. Rev. Microbiol.* **1972**, *26*, 369–388.
4. J. A. MacLaren, *J. Bacteriol.* **1952**, *63*, 233–241.
5. T. Masuda, *Pharm. Bull.* **1956**, *4*, 71–72.
6. T. Masuda, *Pharm. Bull.* **1957**, *5*, 136–141.
7. T. Masuda, *Pharm. Bull.* **1957**, *5*, 28–30.
8. R. L. Beach; G. W. E. Plaut, *Tetrahedron Lett.* **1969**, *40*, 3489–3492.
9. T. Paterson; H. C. S. Wood, *J. Chem. Soc. Commun.* **1969**, 290–291.
10. T. Rowan; H. C. Wood, *J. Chem. Soc. Perkin Trans. I* **1968**, *4*, 452–458.
11. T. Rowan; H. C. S. Wood, *Proc. Chem. Soc.* **1963**, 21–22.
12. H. Wacker; R. A. Harvey; C. H. Winestock; G. W. Plaut, *J. Biol. Chem.* **1964**, *239*, 3493–3497.
13. R. A. Harvey; G. W. Plaut, *J. Biol. Chem.* **1966**, *241*, 2120–2136.
14. G. W. Plaut; R. L. Beach; T. Aogaichi, *Biochemistry* **1970**, *9*, 771–785.
15. K. P. Stahmann; H. N. Arst, Jr.; H. Althofer; J. L. Revuelta; N. Monschau; C. Schlupen; C. Gatgens; A. Wiesenburg; T. Schlosser, *Environ. Microbiol.* **2001**, *3*, 545–550.
16. K. P. Stahmann; J. L. Revuelta; H. Seulberger, *Appl. Microbiol. Biotechnol.* **2000**, *53*, 509–516.
17. J. B. Perkins; J. G. Pero; A. Sloma, Riboflavin Overproducing Bacteria Expressing the Rib Operon of *Bacillus*. Eur. Pat. Appl. EP 405370 A1 910102, **1991**.
18. M. Humbelin; V. Griesser; T. Keller; W. Schurter; M. Haiker; H. P. Hohmann; H. Ritz; G. Richter; A. Bacher; A. P. G. M. Van Loon, *J. Ind. Microbiol. Biotechnol.* **1999**, *22*, 1–7.
19. G. W. Plaut; C. M. Smith; W. L. Alworth, *Annu. Rev. Biochem.* **1974**, *43*, 899–922.
20. A. Bacher, Biosynthesis of Flavins. In *Chemistry and Biochemistry of Flavoenzymes*; F. Müller, Ed.; CRC Press: Boca Raton, FL, 1991; Vol. 1, pp 215–259.
21. A. Bacher; S. Eberhardt; W. Eisenreich; M. Fischer; S. Herz; B. Illarionov; K. Kis; G. Richter, *Vitam Horm* **2001**, *61*, 1–49.
22. A. Bacher; S. Eberhardt; M. Fischer; K. Kis; G. Richter, *Annu Rev. Nutr.* **2000**, *20*, 153–167.
23. A. Bacher; S. Eberhardt; G. Richter, Biosynthesis of Riboflavin. In *Escherichia coli and Salmonella*, 2nd ed.; F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, H. E. Umbarger, Eds.; ASM Press: Washington: DC, 1996; 2 vols., Vol. 1, pp 657–664.
24. A. Bacher; W. Eisenreich; K. Kis; R. Ladenstein; G. Richter; J. Scheuring; S. Weinkauff, Biosynthesis of Flavins. In *Bioorganic Chemistry Frontiers*; H. Dugas, F. P. Schmidtchen, Eds.; Springer: Berlin, 1993; Vol. 3, pp 147–192.
25. M. Fischer; A. Bacher, *Nat. Prod. Rep.* **2005**, *22*, 324–350.
26. M. Fischer; A. Bacher, *Physiologia Plantarum* **2006**, *126*, 304–318.
27. J. B. Perkins; J. Pero, Biosynthesis of Riboflavin, Biotin, Folic Acid, and Cobalamin. In *Bacillus Subtilis and Other Gram-Positive Bacteria. Biochemistry, Physiology, and Molecular Genetics*; A. L. Sonensheim, Ed.; American Society for Microbiology: Washington, DC, 1993, pp 319–334.
28. A. Bacher, *Chemistry and Biochemistry of Flavoproteins*; F. Müller, Ed.; CRC Press: Boca Raton, FL, 1990; Vol. 1, p 215.
29. D. W. Young, *Nat. Prod. Rep.* **1986**, *3*, 395–419.
30. G. M. Brown; J. J. Reynolds, *Annu. Rev. Biochem.* **1969**, *32*, 419–462.
31. G. W. E. Plaut, *Annu. Rev. Biochem.* **1961**, *30*, 409–446.
32. A. Bacher; B. Mailaender, *J. Biol. Chem.* **1973**, *248*, 6227–6231.
33. C. M. Baugh; C. L. Krumdieck, *J. Bacteriol.* **1969**, *98*, 1114–1119.
34. B. Mailaender; A. Bacher, *J. Biol. Chem.* **1976**, *251*, 3623–3628.
35. A. W. Burg; G. M. Brown, *J. Biol. Chem.* **1968**, *243*, 2349–2358.
36. W. A. Wolf; G. M. Brown, *Biochim. Biophys. Acta* **1969**, *192*, 468–478.
37. G. Auerbach; A. Herrmann; A. Bracher; G. Bader; M. Gutlich; M. Fischer; M. Neukamm; M. Garrido-Franco; J. Richardson; H. Nar; R. Huber; A. Bacher, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13567–13572.
38. J. Rebelo; G. Auerbach; G. Bader; A. Bracher; H. Nar; C. Hösl; N. Schramek; J. Kaiser; A. Bacher; R. Huber; M. Fischer, *J. Mol. Biol.* **2003**, *326*, 503–516.
39. J. J. Yim; G. M. Brown, *J. Biol. Chem.* **1976**, *251*, 5087–5094.
40. A. Bracher; M. Fischer; W. Eisenreich; H. Ritz; N. Schramek; P. Boyle; P. Gentili; R. Huber; H. Nar; G. Auerbach; A. Bacher, *J. Biol. Chem.* **1999**, *274*, 16727–16735.
41. N. Schramek; A. Bracher; M. Fischer; G. Auerbach; H. Nar; R. Huber; A. Bacher, *J. Mol. Biol.* **2002**, *316*, 829–837.
42. A. Bracher; W. Eisenreich; N. Schramek; H. Ritz; E. Gotze; A. Herrmann; M. Gutlich; A. Bacher, *J. Biol. Chem.* **1998**, *273*, 28132–28141.
43. A. Bracher; N. Schramek; A. Bacher, *Biochemistry* **2001**, *40*, 7896–7902.
44. N. Schramek; A. Bracher; A. Bacher, *J. Biol. Chem.* **2001**, *276*, 2622–2626.
45. H. Ritz; N. Schramek; A. Bracher; S. Herz; W. Eisenreich; G. Richter; A. Bacher, *J. Biol. Chem.* **2001**, *276*, 22273–22277.
46. J. Ren; M. Kotaka; M. Lockyer; H. K. Lamb; A. R. Hawkins; D. K. Stammers, *J. Biol. Chem.* **2005**, *280*, 36912–36919.
47. J. Kaiser; N. Schramek; S. Eberhardt; S. Puttmer; M. Schuster; A. Bacher, *Eur. J. Biochem.* **2002**, *269*, 5264–5270.

48. M. Kobayashi; Y. Ohara-Nemoto; M. Kaneko; H. Hayakawa; M. Sekiguchi; K. Yamamoto, *J. Biol. Chem.* **1998**, *273*, 26394–26399.
49. N. Schramek; A. Bracher; A. Bacher, *J. Biol. Chem.* **2001**, *276*, 44157–44162.
50. H. Nar; R. Huber; W. Meining; C. Schmid; S. Weinkauff; A. Bacher, *Structure* **1995**, *3*, 459–466.
51. J. E. Spoonamore; A. L. Dahlgren; N. E. Jacobsen; V. Bandarian, *Biochemistry* **2006**, *45*, 12144–12155.
52. J. E. Spoonamore; V. Bandarian, *Biochemistry* **2008**, *47*, 2592–2600.
53. D. E. Graham; H. Xu; R. H. White, *Biochemistry* **2002**, *41*, 15074–15084.
54. S. Herz; S. Eberhardt; A. Bacher, *Phytochemistry* **2000**, *53*, 723–731.
55. F. Fassbinder; M. Kist; S. Bereswill, *FEMS Microbiol. Lett.* **2000**, *191*, 191–197.
56. L. Chatwell; T. Krojer; A. Fidler; W. Romisch; W. Eisenreich; A. Bacher; R. Huber; M. Fischer, *J. Mol. Biol.* **2006**, *359*, 1334–1351.
57. M. Graupner; H. Xu; R. H. White, *J. Bacteriol.* **2002**, *184*, 1952–1957.
58. R. B. Burrows; G. M. Brown, *J. Bacteriol.* **1978**, *136*, 657–667.
59. M. Fischer; W. Römisch; S. Saller; B. Illarionov; G. Richter; F. Rohdich; W. Eisenreich; A. Bacher, *J. Biol. Chem.* **2004**, *279*, 36299–36308.
60. G. Richter; M. Fischer; C. Krieger; S. Eberhardt; H. Lüttgen; I. Gerstenschläger; A. Bacher, *J. Bacteriol.* **1997**, *179*, 2022–2028.
61. A. Bacher; F. Lingens, *J. Biol. Chem.* **1970**, *245*, 4647–4652.
62. A. Bacher; F. Lingens, *J. Biol. Chem.* **1971**, *246*, 7018–7022.
63. O. Oltmanns; F. Lingens, *Z. Naturforsch. B* **1967**, *22*, 751–754.
64. E. M. Logvinenko; G. M. Shavlovskii; L. V. Koltun; G. P. Ksheminskaia, *Mikrobiologiya* **1975**, *44*, 48–54.
65. E. M. Logvinenko; G. M. Shavlovskii; A. E. Zakal'skii, *Mikrobiologiya* **1979**, *48*, 756–768.
66. E. M. Logvinenko; G. M. Shavlovskii; A. E. Zakal'skii; E. Z. Seniuta, *Biokhimiya* **1980**, *45*, 1284–1292.
67. G. M. Shavlovskii; A. A. Sibirnyi; B. V. Kshanovskaia; L. V. Koltun; E. M. Logvinenko, *Genetika* **1979**, *15*, 1561–1568.
68. J. Sadique; R. Shanmugasundaram; E. R. Shanmugasundaram, *Biochem. J.* **1966**, *101*, 2C–3C.
69. P. Stenmark; M. Moche; D. Gurmu; P. Nordlund, *J. Mol. Biol.* **2007**, *373*, 48–64.
70. G. Klein; A. Bacher, *Z. Naturforsch.* **1980**, *35b*, 482–484.
71. P. Nielsen; A. Bacher, *Biochim. Biophys. Acta* **1981**, *662*, 312–317.
72. A. Urban; I. Ansmant; Y. Motorin, *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2003**, *786*, 187–195.
73. E. V. Koonin, *Nucleic Acids Res.* **1996**, *24*, 2411–2415.
74. S. C. Chen; Y. C. Chang; C. H. Lin; C. H. Lin; S. H. Liaw, *J. Biol. Chem.* **2006**, *281*, 7605–7613.
75. P. J. Keller; Q. Le Van; S. U. Kim; D. H. Bown; H. C. Chen; A. Köhne; A. Bacher; H. G. Floss, *Biochemistry* **1988**, *27*, 1117–1120.
76. M. Fischer; L. Chatwell; W. Römisch; W. Eisenreich; R. Huber; T. Krojer; A. Bacher, Biosynthesis of Riboflavin. Stereochemistry and X-ray Structure of 2,5-Diamino-6-Ribitylamino-4(3H)-Pyrimidinone 5'-Phosphate Synthase of *Methanococcus jannaschii*. In *Flavins Flavoproteins*, Proceedings of the 15th International Symposium. ARchiTect Inc., 2005; pp 725–730.
77. W. Eisenreich; M. Joshi; B. Illarionov; G. Richter; W. Romisch-Margl; F. Muller; A. Bacher; M. Fischer, *FEBS J.* **2007**, *274*, 5876–5890.
78. M. L. B. Magalhaes; A. Argyrou; S. M. Cahill; J. S. Blanchard, *Biochemistry* **2008**, in press.
79. W. Roemisch-Margl; W. Eisenreich; I. Haase; A. Bacher; M. Fischer, *FEBS J.* **2008**, in press.
80. D. Yuan; Q. Wang; W. Gao; F. Sheng; Z. Zhang; Q. Lu; H. Cang; R. Bi, *Protein Pept. Lett.* **2007**, *14*, 925–927.
81. G. W. Plaut, *J. Biol. Chem.* **1960**, *235*, PC41–PC42.
82. G. W. E. Plaut, *J. Biol. Chem.* **1963**, *238*, 2225–2243.
83. G. W. E. Plaut; R. A. Harvey, *Meth. Enzymol.* **1971**, *18*, 515–538.
84. G. Richter; C. Krieger; R. Volk; K. Kis; H. Ritz; E. Gotze; A. Bacher, *Meth. Enzymol.* **1997**, *280*, 374–382.
85. P. Nielsen; G. Neuberger; I. Fujii; D. H. Bown; P. J. Keller; H. G. Floss; A. Bacher, *J. Biol. Chem.* **1986**, *261*, 3661–3669.
86. A. Bacher; Q. Le Van; P. J. Keller; H. G. Floss, *J. Biol. Chem.* **1983**, *258*, 13431–13437.
87. A. Bacher; Q. L. Van; P. J. Keller; H. G. Floss, *J. Am. Chem. Soc.* **1985**, *107*, 6380–6385.
88. Q. Le Van; P. J. Keller; D. H. Bown; H. G. Floss; A. Bacher, *J. Bacteriol.* **1985**, *162*, 1280–1284.
89. R. Volk; A. Bacher, *J. Am. Chem. Soc.* **1988**, *110*, 3651–3653.
90. R. Volk; A. Bacher, *J. Biol. Chem.* **1990**, *265*, 19479–19485.
91. R. Volk; A. Bacher, *J. Biol. Chem.* **1991**, *266*, 20610–20618.
92. E. Goetze; K. Kis; W. Eisenreich; N. Yamauchi; K. Kakinuma; A. Bacher, *J. Org. Chem.* **1998**, *63*, 6456–6457.
93. S. Mörtl; M. Fischer; G. Richter; J. Tack; S. Weinkauff; A. Bacher, *J. Biol. Chem.* **1996**, *271*, 33201–33207.
94. S. Eberhardt; G. Richter; P. W. Gimbel; T. Werner; A. Bacher, *Eur. J. Biochem.* **1996**, *242*, 712–719.
95. S. Echt; S. Bauer; S. Steinbacher; R. Huber; A. Bacher; M. Fischer, *J. Mol. Biol.* **2004**, *341*, 1085–1096.
96. M. Fischer; W. Römisch; S. Schiffmann; M. Kelly; H. Oschkinat; S. Steinbacher; R. Huber; W. Eisenreich; G. Richter; A. Bacher, *J. Biol. Chem.* **2002**, *277*, 41410–41416.
97. D. I. Liao; P. V. Viitanen; D. B. Jordan, *Acta Crystallogr. D Biol. Crystallogr.* **2000**, *56*, 1495–1497.
98. D. I. Liao; J. C. Calabrese; Z. Wawrzak; P. V. Viitanen; D. B. Jordan, *Structure (Camb)* **2001**, *9*, 11–18.
99. M. J. Kelly; L. J. Ball; C. Krieger; Y. Yu; M. Fischer; S. Schiffmann; P. Schmieder; R. Kuhne; W. Bermel; A. Bacher; G. Richter; H. Oschkinat, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13025–13030.
100. D. I. Liao; Y. J. Zheng; P. V. Viitanen; D. B. Jordan, *Biochemistry* **2002**, *41*, 1795–1806.
101. S. Steinbacher; S. Schiffmann; A. Bacher; M. Fischer, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1338–1340.
102. S. Steinbacher; S. Schiffmann; G. Richter; R. Huber; A. Bacher; M. Fischer, *J. Biol. Chem.* **2003**, *278*, 42256–42265.
103. A. Bacher; H. Schneppe; B. Mailaender; M. K. Otto; Y. Ben-Shaul, Structure and Function of the Riboflavin Synthase Complex of *Bacillus subtilis*. In *Flavins Flavoproteins*, Proceedings of the 6th International Symposium, 1980, pp 579–586.
104. K. Schott; J. Kellermann; F. Lottspeich; A. Bacher, *J. Biol. Chem.* **1990**, *265*, 4204–4209.
105. A. Bacher; S. Eberhardt; M. Fischer; S. Mörtl; K. Kis; K. Kugelbrey; J. Scheuring; K. Schott, *Meth. Enzymol.* **1997**, *280*, 389–399.
106. S. Eberhardt; S. Korn; F. Lottspeich; A. Bacher, *J. Bacteriol.* **1997**, *179*, 2938–2943.
107. M. Fischer; A. K. Schott; W. Römisch; A. Ramsperger; M. Augustin; A. Fidler; A. Bacher; G. Richter; R. Huber; W. Eisenreich, *J. Mol. Biol.* **2004**, *343*, 267–278.

108. A. Ramsperger; M. Augustin; A. K. Schott; S. Gerhardt; T. Krojer; W. Eisenreich; B. Illarionov; M. Cushman; A. Bacher; R. Huber; M. Fischer, *J. Biol. Chem.* **2006**, *281*, 1224–1232.
109. D. B. Jordan; K. O. Bacot; T. J. Carlson; M. Kessel; P. V. Viitanen, *J. Biol. Chem.* **1999**, *274*, 22114–22121.
110. S. Gerhardt; I. Haase; S. Steinbacher; J. T. Kaiser; M. Cushman; A. Bacher; R. Huber; M. Fischer, *J. Mol. Biol.* **2002**, *318*, 1317–1329.
111. W. Meining; S. Mortl; M. Fischer; M. Cushman; A. Bacher; R. Ladenstein, *J. Mol. Biol.* **2000**, *299*, 181–197.
112. K. Kis; R. Volk; A. Bacher, *Biochemistry* **1995**, *34*, 2883–2892.
113. N. Schramek; I. Haase; M. Fischer; A. Bacher, *J. Am. Chem. Soc.* **2003**, *125*, 4460–4466.
114. I. Haase; M. Fischer; A. Bacher; N. Schramek, *J. Biol. Chem.* **2003**, *278*, 37909–37915.
115. M. Fischer; I. Haase; K. Kis; W. Meining; R. Ladenstein; M. Cushman; N. Schramek; R. Huber; A. Bacher, *J. Mol. Biol.* **2003**, *326*, 783–793.
116. K. Kis; K. Kugelbrey; A. Bacher, *J. Org. Chem.* **2001**, *66*, 2555–2559.
117. A. Eschenmoser; E. Loewenthal, *Chem. Soc. Rev.* **1992**, *21*, 1–16.
118. C. J. Strupp, ETH Zuerich **1992**.
119. M. Koch; C. Breithaupt; S. Gerhardt; I. Haase; S. Weber; M. Cushman; R. Huber; A. Bacher; M. Fischer, *Eur. J. Biochem.* **2004**, *271*, 3208–3214.
120. E. Morgunova; W. Meining; B. Illarionov; I. Haase; G. Jin; A. Bacher; M. Cushman; M. Fischer; R. Ladenstein, *Biochemistry* **2005**, *44*, 2746–2758.
121. E. Morgunova; S. Saller; I. Haase; M. Cushman; A. Bacher; M. Fischer; R. Ladenstein, *J. Biol. Chem.* **2007**, *282*, 17231–17241.
122. K. Persson; G. Schneider; B. J. Douglas; P. V. Viitanen; T. Sandalova, *Plant Prot. Sci.* **1999**, *8*, 2355–2365.
123. R. Ladenstein; K. Ritsert; R. Huber; G. Richter; A. Bacher, *Eur. J. Biochem.* **1994**, *223*, 1007–1017.
124. R. Ladenstein; M. Schneider; R. Huber; H. D. Bartunik; K. Wilson; K. Schott; A. Bacher, *J. Mol. Biol.* **1988**, *203*, 1045–1070.
125. K. Ritsert; R. Huber; D. Turk; R. Ladenstein; K. Schmidt-Base; A. Bacher, *J. Mol. Biol.* **1995**, *253*, 151–167.
126. X. Zhang; W. Meining; M. Fischer; A. Bacher; R. Ladenstein, *J. Mol. Biol.* **2001**, *306*, 1099–1114.
127. S. Karthikeyan; Q. Zhou; F. Mseeh; N. V. Grishin; A. L. Osterman; H. Zhang, *Structure (Camb)* **2003**, *11*, 265–273.
128. S. Klinke; V. Zylberman; H. R. Bonomi; I. Haase; B. G. Guimaraes; B. C. Braden; A. Bacher; M. Fischer; F. A. Goldbaum, *J. Mol. Biol.* **2007**, *373*, 664–680.
129. S. Klinke; V. Zylberman; D. R. Vega; B. G. Guimaraes; B. C. Braden; F. A. Goldbaum, *J. Mol. Biol.* **2005**, *353*, 124–137.
130. V. Zylberman; P. O. Craig; S. Klinke; B. C. Braden; A. Cauert; F. A. Goldbaum, *J. Biol. Chem.* **2004**, *279*, 8093–8101.
131. V. Zylberman; S. Klinke; I. Haase; A. Bacher; M. Fischer; F. A. Goldbaum, *J. Bacteriol.* **2006**, *188*, 6135–6142.
132. F. A. Goldbaum; C. A. Velikovskiy; P. C. Baldi; S. Mortl; A. Bacher; C. A. Fossati, *J. Med. Microbiol.* **1999**, *48*, 833–839.
133. A. Bacher; B. Mailänder, *J. Bacteriol.* **1978**, *134*, 476–482.
134. R. Ladenstein; H. C. Ludwig; A. Bacher, *J. Biol. Chem.* **1983**, *258*, 11981–11983.
135. A. Bacher; H. C. Ludwig; H. Schnepfle; Y. Ben-Shaul, *J. Mol. Biol.* **1986**, *187*, 75–86.
136. R. Ladenstein; B. Meyer; R. Huber; H. Labischinski; K. Bartels; H. D. Bartunik; L. Bachmann; H. C. Ludwig; A. Bacher, *J. Mol. Biol.* **1986**, *187*, 87–100.
137. S. Gerhardt; A. K. Schott; N. Kairies; M. Cushman; B. Illarionov; W. Eisenreich; A. Bacher; R. Huber; S. Steinbacher; M. Fischer, *Structure (Camb)* **2002**, *10*, 1371–1381.
138. D. I. Liao; Z. Wawrzak; J. C. Calabrese; P. V. Viitanen; D. B. Jordan, *Structure (Camb)* **2001**, *9*, 399–408.
139. K. Kis; A. Bacher, *J. Biol. Chem.* **1995**, *270*, 16788–16795.
140. A. Bacher; H. C. Ludwig; H. Schnepfle; Y. Ben-Shaul, *J. Mol. Biol.* **1986**, *187*, 75–86.
141. X. Zhang; P. V. Konarev; M. V. Petoukhov; D. I. Svergun; L. Xing; R. H. Cheng; I. Haase; M. Fischer; A. Bacher; R. Ladenstein; W. Meining, *J. Mol. Biol.* **2006**, *362*, 753–770.
142. R. L. Beach; G. W. E. Plaut, *J. Am. Chem. Soc.* **1970**, *92*, 2913–2916.
143. T. Paterson; H. C. Wood, *J. Chem. Soc. Perkin Trans. I* **1972**, *8*, 1051–1056.
144. H. Sedlmaier; F. Muller; P. J. Keller; A. Bacher, *Z. Naturforsch [C]* **1987**, *42*, 425–429.
145. W. Pfeleiderer; W. Hutzenlaub, *Chem. Ber.* **1973**, *106*, 3149–3174.
146. R. L. Beach; G. W. E. Plaut, *J. Org. Chem.* **1971**, *36*, 3937–3943.
147. D. H. Bown; P. J. Keller; H. G. Floss; H. Sedlmaier; A. Bacher, *J. Org. Chem.* **1986**, *51*, 2461–2467.
148. W. Pfeleiderer; J. W. Bunting; D. D. Perrin; G. Nuebel, *Chem. Ber.* **1966**, *99*, 3503–3523.
149. G. W. Plaut; R. L. Beach, Interaction of Riboflavine Synthetase with Analogs of 6,7-Dimethyl-8-Ribityllumazine. In *Chemistry and Biology of Pteridines*, Proceedings of the 5th International Symposium; 1975; W. Pfeleiderer, Ed.; de Gruyter: Berlin; pp 101–124.
150. M. Cushman; W. C. Wong; A. Bacher, *J. Chem. Soc. Perkin Trans. I* **1986**, *6*, 1051–1053.
151. M. Cushman; D. A. Patrick; A. Bacher; J. Scheuring, *J. Org. Chem.* **1991**, *56*, 4603–4608.
152. B. Illarionov; I. Haase; A. Bacher; M. Fischer; N. Schramek, *J. Biol. Chem.* **2003**, *278*, 47700–47706.
153. B. Illarionov; W. Eisenreich; A. Bacher, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7224–7229.
154. B. Illarionov; I. Haase; M. Fischer; A. Bacher; N. Schramek, *Biol. Chem.* **2005**, *386*, 127–136.
155. B. Illarionov; W. Eisenreich; A. Bacher, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7224–7229.
156. B. Illarionov; W. Eisenreich; N. Schramek; A. Bacher; M. Fischer, *J. Biol. Chem.* **2005**, *280*, 28541–28546.
157. V. N. Mironov; A. S. Kraev; B. K. Chernov; A. V. Ul'ianov; B. Golova lu, *Dokl. Akad. Nauk. SSSR* **1989**, *305*, 482–487.
158. B. Illarionov; K. Kemter; S. Eberhardt; G. Richter; M. Cushman; A. Bacher, *J. Biol. Chem.* **2001**, *276*, 11524–11530.
159. S. Eberhardt; N. Zingler; K. Kemter; G. Richter; M. Cushman; A. Bacher, *Eur. J. Biochem.* **2001**, *268*, 4315–4323.
160. W. Meining; S. Eberhardt; A. Bacher; R. Ladenstein, *J. Mol. Biol.* **2003**, *331*, 1053–1063.
161. V. Truffault; M. Coles; T. Diercks; K. Abelmann; S. Eberhardt; H. Luttgen; A. Bacher; H. Kessler, *J. Mol. Biol.* **2001**, *309*, 949–960.
162. M. Fischer; A. K. Schott; K. Kemter; R. Feicht; G. Richter; B. Illarionov; W. Eisenreich; S. Gerhardt; M. Cushman; S. Steinbacher; R. Huber; A. Bacher, *BMC Biochem.* **2003**, *4*, 18.
163. R. Gast; J. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 833–837.
164. E. D. Small; P. Koka; J. Lee, *J. Biol. Chem.* **1980**, *255*, 8804–8810.

165. P. Macheroux; K. U. Schmidt; P. Steinerstauch; S. Ghisla; P. Colepiccolo; R. Buntic; J. W. Hastings, *Biochem. Biophys. Res. Commun.* **1987**, *146*, 101–106.
166. V. N. Petushkov; J. Lee, *Eur. J. Biochem.* **1997**, *245*, 790–796.
167. J. Lee, *Biophys. Chem.* **1993**, *48*, 149–158.
168. D. J. O’Kane; B. Woodward; J. Lee; D. C. Prasher, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1100–1104.
169. L. Chatwell; V. Illarionova; B. Illarionov; W. Eisenreich; R. Huber; A. Skerra; A. Bacher; M. Fischer, *J. Mol. Biol.* **2007**, submitted.
170. L. Chatwell; V. Illarionova; B. Illarionov; W. Eisenreich; R. Huber; A. Skerra; A. Bacher; M. Fischer, *J. Mol. Biol.* **2008**, in press.
171. J. Lee; I. B. C. Matheson; F. Mueller; D. J. O’Kane; J. Vervoort; A. J. W. G. Visser, The Mechanism of Bacterial Bioluminescence. In *Chemistry and Biochemistry of Flavoenzymes*; F. Müller, Ed.; CRC Press: Boca Raton, FL, 1991; Vol. 2.
172. J. Scheuring; J. Lee; M. Cushman; H. Patel; D. A. Patrick; A. Bacher, *Biochemistry* **1994**, *33*, 7634–7640.
173. B. Illarionov; W. Eisenreich; M. Wirth; C. L. Lee; Y. E. Woo; A. Bacher; M. Fischer, *Biol. Chem.* **2007**, *388*, 1313–1323.
174. M. Mack; A. P. van Loon; H. P. Hohmann, *J. Bacteriol.* **1998**, *180*, 950–955.
175. I. M. Solovieva; A. Iomantas Iu; R. A. Kreneva; I. Kozlov Iu; D. A. Perumov, *Genetika* **1997**, *33*, 739–743.
176. I. M. Solovieva; R. A. Kreneva; D. J. Leak; D. A. Perumov, *Microbiology* **1999**, *145* (Pt. 1), 67–73.
177. I. M. Solovieva; R. A. Kreneva; B. M. Polanuer; I. Kozlov Iu; D. A. Perumov, *Genetika* **1998**, *34*, 839–842.
178. Y. Higashitsuji; A. Angerer; S. Berghaus; B. Hobl; M. Mack, *FEMS Microbiol. Lett.* **2007**, *274*, 48–54.
179. M. S. Gelfand; A. A. Mironov; J. Jomantas; Y. I. Kozlov; D. A. Perumov, *Trends Genet.* **1999**, *15*, 439–442.
180. A. G. Vitreschak; D. A. Rodionov; A. A. Mironov; M. S. Gelfand, *Nucleic Acids Res.* **2002**, *30*, 3141–3151.
181. I. Gusarov; R. A. Kreneva; D. A. Podcharniaev; V. Iomantas Iu; E. G. Abalagina; N. V. Stoinova; D. A. Perumov; I. Kozlov Iu, *Mol. Biol. (Mosk.)* **1997**, *31*, 446–453.
182. Y. V. Kil; V. N. Mironov; I. Gorishin; R. A. Kreneva; D. A. Perumov, *Mol. Gen. Genet.* **1992**, *233*, 483–486.
183. W. C. Winkler; S. Cohen-Chalamish; R. R. Breaker, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15908–15913.
184. A. G. Vitreschak; D. A. Rodionov; A. A. Mironov; M. S. Gelfand, *Trends Genet.* **2004**, *20*, 44–50.
185. J. B. Perkins; A. Sloma; T. Hermann; K. Theriault; E. Zachgo; T. Erdenberger; N. Hannett; N. P. Chatterjee; V. Williams, II; G. A. Rufo; R. Hatch; J. Pero, *J. Ind. Microbiol. Biotechnol.* **1999**, *22*, 8–18.
186. D. Coquard; M. Huecas; M. Ott; J. M. van Dijk; A. P. van Loon; H. P. Hohmann, *Mol. Gen. Genet.* **1997**, *254*, 81–84.
187. G. Schmidt; K.-P. Stahmann; H. Sahn, *Microbiology (Reading, UK)* **1996**, *142*, 411–417.
188. D. L. Heefner; C. A. Weaver; M. J. Yarus; L. A. Burdzinski; D. C. Gyure; E. W. Foster, Riboflavin-Producing Strains of Microorganisms, Method for Selecting, and Method for Fermentation. PCT Int. Appl., 1988, p 54.
189. D. Becker; M. Selbach; C. Rollenhagen; M. Ballmaier; T. F. Meyer; M. Mann; D. Bumann, *Nature* **2006**, *440*, 303–307.
190. C. Rollenhagen; D. Bumann, *Infect. Immun.* **2006**, *74*, 1649–1660.
191. J. E. Mott; C. M. Trepod; S. Arvidson. Critical Genes and Polypeptides of *Haemophilus influenzae*. Division of U.S. Ser. No. 274,586, USA, abandoned, 2005, p 158.
192. A. C. Silver; N. M. Rabinowitz; S. Küffer; J. Graf, *J. Bacteriol.* **2007**, *189*, 6763–6772.
193. J. Chen; T. Sambaiiah; B. Illarionov; M. Fischer; A. Bacher; M. Cushman, *J. Org. Chem.* **2004**, *69*, 6996–7003.
194. M. Cushman; G. Jin; T. Sambaiiah; B. Illarionov; M. Fischer; R. Ladenstein; A. Bacher, *J. Org. Chem.* **2005**, *70*, 8162–8170.
195. M. Cushman; F. Mavandadi; K. Kugelbrey; A. Bacher, *Bioorg. Med. Chem.* **1998**, *6*, 409–415.
196. M. Cushman; F. Mavandadi; D. Yang; K. Kugelbrey; K. Kis; A. Bacher, *J. Org. Chem.* **1999**, *64*, 4635–4642.
197. M. Cushman; J. T. Mihalic; K. Kis; A. Bacher, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 39–42.
198. M. Cushman; T. Sambaiiah; G. Jin; B. Illarionov; M. Fischer; A. Bacher, *J. Org. Chem.* **2004**, *69*, 601–612.
199. M. Cushman; D. Yang; S. Gerhardt; R. Huber; M. Fischer; K. Kis; A. Bacher, *J. Org. Chem.* **2002**, *67*, 5807–5816.
200. M. Cushman; D. Yang; K. Kis; A. Bacher, *J. Org. Chem.* **2001**, *66*, 8320–8327.
201. M. Cushman; D. Yang; J. T. Mihalic; J. Chen; S. Gerhardt; R. Huber; M. Fischer; K. Kis; A. Bacher, *J. Org. Chem.* **2002**, *67*, 6871–6877.
202. Y. Zhang; B. Illarionov; A. Bacher; M. Fischer; G. I. Georg; Q. Z. Ye; D. Vander Velde; P. E. Fanwick; Y. Song; M. Cushman, *J. Org. Chem.* **2007**, *72*, 2769–2776.
203. Y. Zhang; B. Illarionov; E. Morgunova; G. Jin; A. Bacher; M. Fischer; R. Ladenstein; M. Cushman, *J. Org. Chem.* **2008**, *73*, 2715–2724.
204. Y. Zhang; G. Jin; B. Illarionov; A. Bacher; M. Fischer; M. Cushman, *J. Org. Chem.* **2007**, *72*, 7176–7184.
205. J. Chen; B. Illarionov; A. Bacher; M. Fischer; I. Haase; G. Georg; Q. Z. Ye; Z. Ma; M. Cushman, *Anal. Biochem.* **2005**, *338*, 124–130.
206. J. Kaiser; B. Illarionov; F. Rohdich; W. Eisenreich; S. Saller; J. V. den Brulle; M. Cushman; A. Bacher; M. Fischer, *Anal. Biochem.* **2007**, *365*, 52–61.
207. S. Nakagawa; A. Igarashi; T. Ohta; T. Hagihara; T. Fujio; K. Aisaka, *Biosci. Biotechnol. Biochem.* **1995**, *59*, 694–702.
208. D. J. Manstein; E. F. Pai, *J. Biol. Chem.* **1986**, *261*, 16169–16173.
209. S. Bauer; K. Kemter; A. Bacher; R. Huber; M. Fischer; S. Steinbacher, *J. Mol. Biol.* **2003**, *326*, 1463–1473.
210. W. Wang; R. Kim; J. Jancarik; H. Yokota; S. H. Kim, *Proteins* **2003**, *52*, 633–635.
211. W. Wang; R. Kim; H. Yokota; S. H. Kim, *Proteins* **2004**.
212. F. J. Sandoval; S. Roje, *J. Biol. Chem.* **2005**, *280*, 38337–38345.
213. M. E. Taga; N. A. Larsen; A. R. Howard-Jones; C. T. Walsh; G. C. Walker, *Nature* **2007**, *446*, 449–453.
214. W. Sybesma; C. Burgess; M. Starrenburg; D. van Sinderen; J. Hugenholtz, *Metab. Eng.* **2004**, *6*, 109–115.
215. C. Burgess; M. O’Connell-Motherway; W. Sybesma; J. Hugenholtz; D. van Sinderen, *Appl. Environ. Microbiol.* **2004**, *70*, 5769–5777.



**Biographical Sketches**

Markus Fischer received his Ph.D. degree from the Technical University of Munich. In 2003, he completed his 'habilitation' for food chemistry and biochemistry. In 2004, he received the Kurt-Taeufel Award from the Gesellschaft Deutscher Chemiker. He served as a guest lecturer at the German Institute of Science and Technology, Singapore from 2002 to 2006. Since 2006, he is full professor and director of the Institute of Food Chemistry at the University of Hamburg. His research interest lies in the biosynthesis of riboflavin and folic acid. His research deals with the characterization of biochemical reactions at the level of genes, proteins, metabolites, and mechanisms. The development of drugs and evaluation of drug targets are major objectives of his research group.



Adelbert Bacher received his Ph.D. and M.D. degrees from the University of Tübingen. He served as an associate professor of microbiology at the University of Frankfurt and is presently professor of organic chemistry and biochemistry at the Technical University of Munich.



## 7.03 Flavin-Dependent Enzymes

Rebecca L. Fagan and Bruce A. Palfey, University of Michigan Medical School, Ann Arbor, MI, USA

© 2010 Elsevier Ltd. All rights reserved.

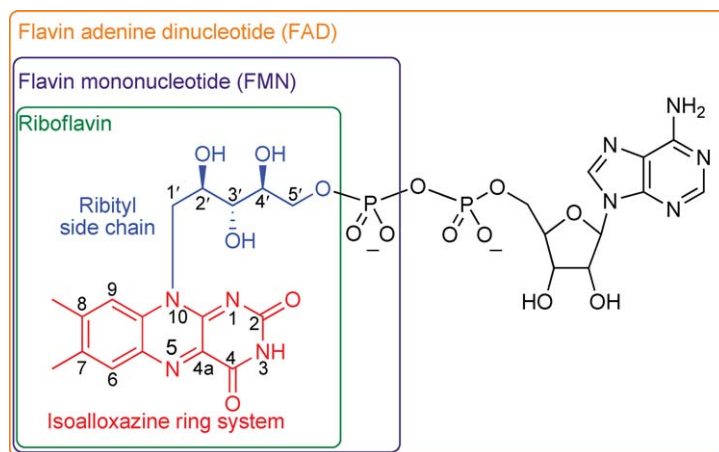
<b>7.03.1</b>	<b>Introduction</b>	38
<b>7.03.2</b>	<b>Oxidation of Carbon–Heteroatom Bonds</b>	42
7.03.2.1	Carbon–Nitrogen Oxidation	43
7.03.2.1.1	D-Amino acid oxidase	43
7.03.2.1.2	Monoamine oxidase	45
7.03.2.1.3	Lysine-specific demethylase 1	48
7.03.2.1.4	Monomeric sarcosine oxidase	48
7.03.2.1.5	Proline dehydrogenases	49
7.03.2.1.6	RebO	50
7.03.2.1.7	Pyridine nucleotide oxidations	50
<b>7.03.2.2</b>	<b>Carbon–Oxygen Oxidation</b>	51
7.03.2.2.1	Glucose oxidase	51
7.03.2.2.2	Flavocytochrome <i>b</i> <sub>2</sub>	52
7.03.2.2.3	Lactate monooxygenase	54
7.03.2.2.4	L-lactate oxidase	55
<b>7.03.2.3</b>	<b>Carbon–Sulfur Oxidation</b>	55
7.03.2.3.1	Prenylcysteine lyase	55
<b>7.03.3</b>	<b>Oxidation and Reduction of Carbon–Carbon Bonds</b>	56
7.03.3.1	Old Yellow Enzyme	56
7.03.3.2	Acyl-CoA Dehydrogenases and Oxidases	58
7.03.3.3	Dihydroorotate Dehydrogenase, Dihydropyrimidine Dehydrogenase, and Dihydrouridine Synthase	60
7.03.3.3.1	Dihydroorotate dehydrogenase	60
7.03.3.3.2	Dihydropyrimidine dehydrogenase	62
7.03.3.3.3	Dihydrouridine synthase	63
7.03.3.4	Succinate Dehydrogenase and Fumarate Reductase	63
7.03.3.5	UDP- <i>N</i> -acetylenolpyruvylglucosamine reductase	64
<b>7.03.4</b>	<b>Thiol/Disulfide Chemistry</b>	65
7.03.4.1	Glutathione Reductase	65
7.03.4.2	Thioredoxin Reductase	67
7.03.4.3	Lipoamide Dehydrogenase	68
7.03.4.4	Sulfhydryl Oxidases	70
7.03.4.5	Mercuric Reductase	70
7.03.4.6	NADH Peroxidase	72
<b>7.03.5</b>	<b>Electron Transfer Reactions</b>	72
7.03.5.1	Electron-Transferring Flavoproteins	73
7.03.5.2	Phthalate Dioxygenase Reductase	74
7.03.5.3	Cytochromes P-450 Reductase	75
<b>7.03.6</b>	<b>Oxygen Reactions</b>	76
7.03.6.1	Reaction of Free Flavins with O <sub>2</sub>	77
7.03.6.2	Oxidases	78
7.03.6.3	Hydroxylases	79
7.03.6.4	<i>p</i> -Hydroxybenzoate Hydroxylase	79
7.03.6.5	MICAL	81
7.03.6.6	Tryptophan-7-Halogenase	82
7.03.6.7	ActVA and ActVB	83

7.03.6.8	Flavin-Containing Monooxygenase	84
7.03.6.9	Baeyer–Villiger Monooxygenases	85
7.03.6.10	Bacterial Luciferase	86
<b>7.03.7</b>	<b>Nonredox Reactions</b>	87
7.03.7.1	Alkyldihydroxyacetonephosphate Synthase	87
7.03.7.2	UDP-Galactopyranose Mutase	88
7.03.7.3	Phosphopantothenoylcysteine Decarboxylase	90
7.03.7.4	Chorismate Synthase	90
7.03.7.5	DNA Photolyase	94
7.03.7.6	Oxynitrilase	94
7.03.7.7	Acetohydroxyacid Synthase	95
7.03.7.8	Isopentenyl Diphosphate Isomerase	96
<b>7.03.8</b>	<b>Complex Flavoenzymes</b>	96
7.03.8.1	Glutamate Synthase	96
7.03.8.2	Aclacinomycin Oxidoreductase	98
7.03.8.3	Berberine Bridge Enzyme	99
7.03.8.4	2-Aminobenzyl-CoA Monooxygenase/Reductase	100
<b>7.03.9</b>	<b>Summary</b>	101
<b>References</b>		103

### 7.03.1 Introduction

Hundreds, perhaps thousands, of flavoenzymes have been identified and characterized.<sup>1</sup> Current estimates are that ~1–3% of bacterial and eukaryotic genomes encode proteins that bind flavin.<sup>2</sup> Flavins catalyze a wide variety of chemical transformations, and thus are involved in almost all aspects of biology: energy production, natural product synthesis, biodegradation, chromatin remodeling, DNA repair, apoptosis, protein folding, xenobiotic detoxification, and neural development. Flavoenzymes play a key role in metabolism, shuffling electrons between two-electron reactants (e.g., NADPH, fumarate, lactate) and one-electron reactants such as the metal centers and quinones of the respiratory chain. Many flavoenzymes are involved in natural product biosynthesis, such as the recently discovered RebF and RebH, which catalyze halogenation reactions leading to the antibiotic rebeccamycin.<sup>3</sup> A recently discovered flavoenzyme, lysine-specific demethylase 1 (LSD1), demethylates histone lysines.<sup>4</sup> MICAL (*m*olecule *i*nteracting with *CasL*), a multidomain protein, transmits signals that guide axonal growth. The essential N-terminal domain of this protein is a flavin-dependent monooxygenase.<sup>5</sup>

Flavins are derivatives of riboflavin (vitamin B<sub>2</sub>) and consist of a tricyclic isoalloxazine moiety derivatized with a ribityl chain at N10 (**Figure 1**). The two common enzymatically active flavins are flavin mononucleotide (FMN), which is riboflavin phosphorylated at the 5'-OH of the ribityl chain, and flavin adenine dinucleotide (FAD), which is the condensation product of FMN and AMP. The ribityl chain and its 5'-OH modification is generally a handle for binding by proteins, but occasionally an –OH group participates in catalysis.<sup>6</sup> In most cases, flavoenzymes are nondissociable noncovalent complexes of apoprotein and flavin, with dissociation constants in the nanomolar range or lower. Therefore, under most circumstances, the flavin can be thought of as part of the enzyme. In contrast to the usual noncovalent complexes, there are many cases where the flavin is covalently attached to the enzyme. Covalent attachment occurs at either the 8- $\alpha$  (methyl) or the 6-position of the isoalloxazine, and sometimes at both sites (see **Figure 1**), through either of the imidazole nitrogens of histidine, the sulfur of cysteine, or the phenolic oxygen of tyrosine. The need for covalent attachment is not fully understood, since there are examples of enzymes (e.g., cholesterol oxidases from different species) that use covalently and noncovalently bound flavins to accomplish the same chemistry.<sup>7</sup> In some cases, the redox potential of the flavin is increased markedly (~100 mV) by covalent attachment.<sup>8</sup> Covalent attachment could protect the flavin from deleterious modifications. For instance, covalent attachment of trimethylamine dehydrogenase at C6 of the FAD prevents modification to the 6-hydroxy derivative, which has been proposed to

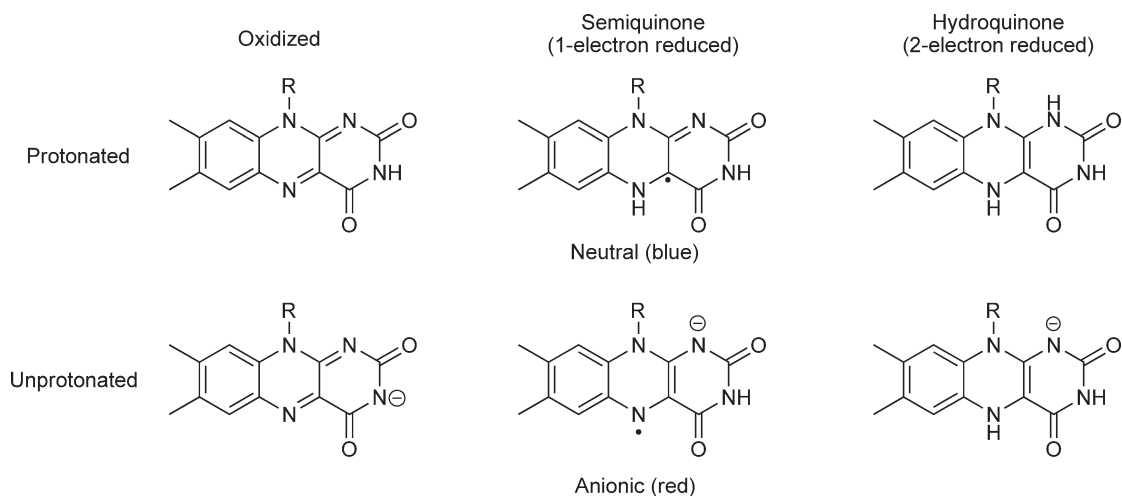


**Figure 1** Flavin structure and nomenclature.

occur by the reaction with either hydroxide or  $O_2$ .<sup>9,10</sup> The other extreme of binding is also known – sometimes, flavins bind so weakly to enzymes that they act as substrates rather than prosthetic groups. For example, bacterial luciferase uses FMN as a substrate after it is reduced by pyridine nucleotide in a reaction catalyzed by a separate enzyme.

The chemically active portion of flavins is the heterocyclic isoalloxazine ring system, which can have three stable oxidation states: fully oxidized, semiquinone (1-electron reduced), and hydroquinone (2-electron reduced) (**Figure 2**). The oxidized isoalloxazine ring system is planar in solution and in most protein structures, but is distorted in a few protein structures, doubtlessly influencing its ability to accept one or two electrons. This electron-deficient heterocycle is susceptible to nucleophilic attack at N5, C4a, and C6. Attack at C6 is not involved in catalysis – indeed, it can be an enzyme-inactivating side reaction – but is important in the covalent flavinylation during the posttranslational maturation of a few flavoenzymes. Oxidized isoalloxazines are highly polarizable, so the electron distribution – and, therefore, reactivity – is influenced greatly by interactions with proteins. The  $pK_a$  of N3, which is  $\sim 10$  in aqueous solution, may become quite low in an enzyme, but is not often suggested to be a major determinant of flavin reactivity.

The single-electron reduced state, the semiquinone, is thermodynamically unstable in aqueous solution, with only  $\sim 5\%$  present in equimolar mixtures of fully oxidized and fully reduced flavins.<sup>11</sup> However, protein interactions alter the stability of the semiquinone markedly – some stabilizing it, others suppressing it.<sup>12</sup> The



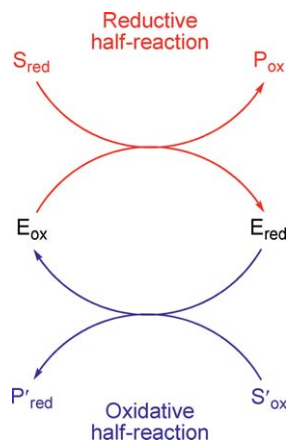
**Figure 2** Chemical states of flavin.

unpaired electron is delocalized over the highly conjugated isoalloxazine.<sup>13,14</sup> High spin density is observed by electron paramagnetic resonance (EPR) spectroscopy at C4a for the neutral (blue) radical. When deprotonated, high spin density is on N5 of the red anion. The  $pK_a$  of N5 is  $\sim 8.5$  in aqueous solution and can vary markedly in an active site, so either protonation state can be important in enzyme reactions. Semiquinones react by single-electron oxidations or reductions, and can also couple to other radicals.

The two-electron reduced, or hydroquinone, isoalloxazine is pale yellow. It is an electron-rich heterocycle and, when planar, is antiaromatic according to Hückel's rule. The ring system of the hydroquinone in some small-molecule structures and some protein structures is bent by as much as  $30^\circ$  along the N5–N10 axis,<sup>15</sup> presumably to relieve the antiaromaticity. However, the majority of hydroquinones bound to proteins do not deviate from planarity much more than oxidized isoalloxazines. This is likely to be influenced by protonation of N1, whose  $pK_a$  is  $\sim 6.7$  in aqueous solution. Quantum calculations show that neutral hydroquinone adopts butterfly conformation but the anion is planar.<sup>16–18</sup> A survey of crystal structures agrees with this correlation – anionic hydroquinones are generally nearly planar, while the few instances of the butterfly conformation belonged to neutral hydroquinones.<sup>19</sup> Hydroquinones react as single-electron donors, as hydride donors, or as nucleophiles at N5 or C4a.

Redox potentials indicate the ability of the flavin to accept electrons. Therefore, this important property provides information on reactivity toward redox substrates and also is a gauge of the electrophilicity or nucleophilicity of the isoalloxazine. The redox potential, like other properties, can be tuned by the protein. The two-electron redox potential of free flavin at pH 7 is  $-207$  mV, while the single-electron potentials are  $-314$  mV for the oxidized/semiquinone couple and  $-124$  mV for the semiquinone/hydroquinone couple.<sup>20</sup> All potentials in this chapter are relative to the standard hydrogen electrode. One- and two-electron redox potentials of flavoprotein span a remarkable range: from  $-495$  mV for the semiquinone/reduced couple in *Azotobacter vinelandii* flavodoxin<sup>21</sup> to  $+153$  mV for the oxidized/semiquinone couple of electron-transfer flavoprotein from *Methylophilus methylotrophus*.<sup>22</sup> The highly conjugated isoalloxazine heterocycle is very polarizable, causing its electron distribution to be altered markedly by nearby charges, hydrogen bonds, pi-stacking with aromatics, or other van der Waals interactions. Interactions that lower electron density in the conjugated diimine of the oxidized isoalloxazine will increase the redox potential. Interactions between the hydroquinone and the protein influence the electron distribution of the electron-rich enediamine, modulating the reactivity. Stabilizing the negative charge of the anionic reduced flavin increases the redox potential.

The catalytic cycles of flavoenzymes can usually be divided into reductive and oxidative half-reactions (Scheme 1). In the reductive half-reaction, the oxidized flavoprotein is reduced by the first substrate, resulting in reduced flavoenzyme. In the oxidative half-reaction, a second substrate oxidizes the reduced flavoprotein (Scheme 1), usually after the product of the reductive half-reaction dissociates, so flavoenzymes frequently have ping-pong kinetic mechanisms. Even when the first product remains bound during the oxidative



**Scheme 1** Typical catalytic cycle of flavoproteins.

half-reaction, the reductive half-reaction is often so exergonic as to be practically irreversible, giving ping-pong steady-state patterns.

Because the flavin prosthetic group is essentially part of the enzyme and does not dissociate during catalysis, the chemistry of the enzyme can be monitored by watching the flavin itself. Convenient absorbance and fluorescence signals allow the catalytic cycle and thermodynamic properties of most flavoenzymes to be studied and understood in detail, especially by using stopped-flow techniques. Very different absorbance spectra are associated with each chemical state of flavin, allowing the flavin to serve as an excellent reporter on reactions. In addition, the binding of ligands often causes perturbations of the absorbance spectra even though no chemical reaction occurs. Charge-transfer interactions are very common between flavoproteins and substrates, products, ligands, or even protein side chains, and give rise to new absorbance bands, usually at wavelengths longer than 500 nm. These interactions occur when the orbitals of an electron-rich molecule overlap with the empty orbitals of an electron-poor molecule. Oxidized flavin, being electron-deficient, can be a charge-transfer acceptor, while reduced flavin, being electron-rich, can be a donor. Free flavins are also inherently fluorescent, and this fluorescence can be enhanced or quenched by proteins. Often, ligand binding or the formation of intermediates causes changes in flavoprotein fluorescence, and sometimes reactions without absorbance changes can be readily observed by fluorescence.

Several reaction mechanisms are available for reductive half-reactions. Hydrides are accepted by N5 of the isoalloxazine ring system. Single-electron donors reduce flavins by transferring electrons through space or through protein bonds, with no known optimal orientation between the flavin and the donor. Flavin reduction can go through covalent intermediates as well. Nucleophilic attack by thiolates at the C4a-position forms adducts whose sulfur can be attacked by a second nucleophilic thiolate, heterolytically cleaving the C–S bond, giving reduced flavin and a disulfide. Nucleophilic attack at N5, leading to the formation of an intermediate adduct, has also been proposed as a mechanism for flavin reduction. A variety of enzymes form N5 adducts, but these are generally with unnatural substrates such as sulfite, cyanide, or nitronates and not with natural substrates.

The same mechanisms – operating in reverse – are available for oxidative half-reactions. Flavin oxidation occurs via hydride abstraction from N5, single-electron transfers through space or through protein, through a C4a adduct formed by nucleophilic attack of the hydroquinone on a disulfide, or through an N5 adduct after attack on a substrate. In addition to these reversals of reductive half-reactions, many flavoenzymes are oxidized by molecular oxygen. Some enzymes – the oxidases – use O<sub>2</sub> simply as a final electron acceptor, forming hydrogen peroxide. Other flavoproteins have evolved to trap reduced molecular oxygen as a flavin adduct, forming flavin C4a-hydroperoxides. These reactive intermediates are used to hydroxylate organic compounds.

The mechanisms for reductive and oxidative half-reactions described above are mixed and matched in enzymes, resulting in a wide diversity of known flavoenzymes (**Table 1**). However, not all conceivable combinations have been observed. Interestingly, oxidation by molecular oxygen is the exception; oxidases are known for each type of reductive half-reaction. For most flavoenzymes, one half-reaction is biologically important and the other regenerates the enzyme. For instance, the biological function of lactate oxidase is to convert lactate to pyruvate, which is accomplished by the reductive half-reaction. The oxidative half-reaction with molecular oxygen regenerates the oxidized flavin. Conversely, in *p*-hydroxybenzoate hydroxylase (PHBH), the oxidative half-reaction is biologically important – the reduced enzyme activates molecular oxygen as a flavin hydroperoxide and hydroxylates *p*-hydroxybenzoate (pOHB) to 3,4-dihydroxybenzoate. The reductive half-reaction with the general cellular reductant NADPH serves only to generate the reduced enzyme, poisoning it for hydroxylation.

In this review, the simple flavoenzymes will be discussed in groups according to the chemistry of a half-reaction – usually the biologically important half-reaction. This is different from the usual classifications based on the net reactions catalyzed or structural motifs. In addition to these enzymes, there are several flavoproteins that do not fit easily into any chemical group. These enzymes will be discussed at the end and include examples where more than one type of chemical conversion occurs, the enzyme contains additional prosthetic groups such as metal centers, or there is no evidence for the flavin being involved chemically in the reaction.

**Table 1** Mixed and matched half-reactions

		Oxidative Half-Reactions				
		H-Transfer	Disulfide Reduction	e <sup>-</sup> -Transfer	Oxidase	Hydroxylase
Reductive Half-Reactions		Dihydroorotate dehydrogenase	Glutathione reductase	Acyl-CoA dehydrogenase	D-amino acid oxidase	p-Hydroxybenzoate hydroxylase
		Lipoamide dehydrogenase	No examples known	Flavocytochrome c-sulfide dehydrogenase	Sulfhydryl oxidase	No examples known
		Glutamate synthase	No examples known	Electron-transferring flavoprotein	Xanthine oxidase	No examples known

### 7.03.2 Oxidation of Carbon-Heteroatom Bonds

Many flavoenzymes catalyze the oxidation of carbon-heteroatom bonds, where the heteroatom is usually nitrogen or oxygen. In these reactions, amines or alcohols are oxidized by the flavin. When nitrogen is the heteroatom, the resulting imine is hydrolyzed nonenzymatically. When the heteroatom is oxygen, a stable carbonyl is formed.

Several properties of these enzymes are shared irrespective of the identity of the flavin (FMN or FAD). Many react with sulfite to form N5 flavin adducts, which is readily detectable spectroscopically. In addition, they stabilize the benzoquinoid resonance structure of 8-mercaptoflavin when it is substituted for the natural prosthetic group, producing a characteristic intense absorbance band at 600 nm.<sup>23</sup> They also stabilize semiquinone. Since the  $pK_a$  of the stabilized semiquinone is often low, the anionic rather than the neutral (blue) semiquinone is detected. However, there are a few cases where the  $pK_a$  is experimentally accessible and both anionic and neutral semiquinones are observed. These general features are the consequence of a positive charge (contributed by either a protein residue or the positive end of a helix dipole) at the N1-C2=O locus of the flavin.

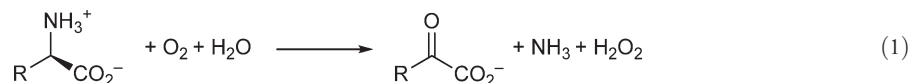
The chemical mechanism of flavin reduction in these enzymes has been debated for years. To date, there is no consensus; however, most proposed mechanisms envision N5 as the reactive position of the flavin. The mechanisms that have been considered include a direct hydride transfer from the substrate  $\alpha$ -carbon to the flavin and a family of mechanisms in which deprotonation of the  $\alpha$ -carbon forms a carbanionic intermediate. Three mechanisms of reaction of the hypothetical carbanion have been proposed. It could reduce the flavin by direct single-electron transfers. Alternatively, it may form a covalent adduct, either after a single-electron transfer and radical coupling, or by direct nucleophilic attack. The adduct would then eliminate reduced flavin.



### 7.03.2.1 Carbon–Nitrogen Oxidation

#### 7.03.2.1.1 D-Amino acid oxidase

The second flavoenzyme to be discovered and one of the most studied is D-amino acid oxidase (DAAO). This enzyme catalyzes the oxidation of D- $\alpha$ -amino acid to the corresponding  $\alpha$ -imino acid in the reductive half-reaction. The imino acid is then hydrolyzed nonenzymatically to the corresponding keto acid (Equation (1)).



In the oxidative half-reaction, the reduced enzyme reacts with molecular oxygen to regenerate oxidized FAD. Here we focus on the presumed biologically important reductive half-reaction, in which a carbon–nitrogen bond is oxidized.

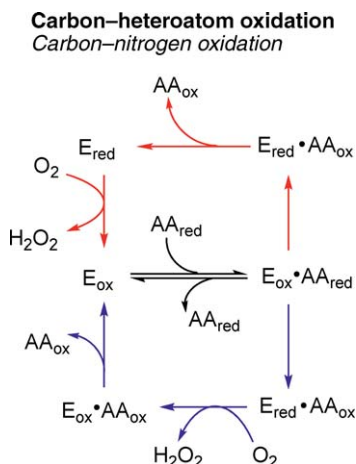
DAAO was once thought to be present only in eukaryotic cells, but within the past decade, DAAOs have been found in several bacterial genomes including *Mycobacterium tuberculosis*<sup>24</sup> and *Streptomyces avermitilis*.<sup>25</sup> The need to oxidize D-amino acids is not obvious since amino acids used for protein synthesis are L-isomers. In microorganisms, DAAO allows exogenous D-amino acids to be used as growth substrates providing necessary carbon and nitrogen.<sup>26</sup> The enzyme also protects organisms from D-amino acid-containing antibiotics.<sup>27</sup> In eukaryotes, DAAO plays different roles. DAAO is postulated to detoxify D-amino acids in mammalian liver and kidney cells, which accumulate from racemization of L-amino acids during aging.<sup>27</sup> Recently, DAAO has been shown to play a role in signaling in the brain by regulating the levels of D-serine. D-serine, produced by a specific racemase in the brain, is involved in the regulation of N-methyl-D-aspartate receptors, either as the free amino acid or in neuroactive peptides.<sup>28–31</sup>

The first DAAO studied mechanistically was from pig kidney (pkDAAO); many kinetic and mechanistic studies have been performed on this enzyme. More recently, yeast DAAOs from *Rhodotorula gracilis* (RgDAAO) and *Trigonopsis variabilis* (TvDAAO) have also been studied. Each has different substrate specificities. The best substrate for pkDAAO is D-proline, followed by hydrophobic and neutral amino acids. Positively charged amino acids are bad substrates, while negatively charged D-amino acids are not substrates.<sup>32–34</sup> In contrast, methionine and valine are the best substrates for RgDAAO.<sup>35</sup>

DAAOs also differ in their multimeric forms. The active form of pkDAAO is a monomer; however, at high concentrations *in vitro*, the enzyme will oligomerize.<sup>36</sup> The yeast enzymes function as homodimers in cells.<sup>27</sup> The homodimer formed by RgDAAO is very stable; it does not dissociate into monomers when diluted.<sup>37,38</sup> TvDAAO, similar to pkDAAO, forms higher-order oligomers at high enzyme concentrations.<sup>39</sup>

Before structures were available, the electrostatic environment of the active site of DAAO was probed spectrally using flavin analogues. The affinity of FAD for DAAO from different organisms varies widely. RgDAAO binds FAD ( $K_d = 2 \times 10^{-8} \text{ mol l}^{-1}$ )<sup>40</sup> an order of magnitude tighter than pkDAAO ( $K_d = 5 \times 10^{-7} \text{ mol l}^{-1}$ ).<sup>41</sup> The relatively weak binding of FAD to pkDAAO has facilitated its replacement with flavin analogues. pkDAAO reconstituted with 8-mercapto FAD has an absorbance spectrum of the quinonoid resonance form, with a maximum at 600 nm.<sup>23</sup> This was interpreted as indicating the presence of a positive charge near the N1–C2=O locus of the isoalloxazine.<sup>42</sup> Consistent with this interpretation, pkDAAO stabilizes the red anionic semiquinone form of the flavin, and both pkDAAO and RgDAAO form N5 sulfite adducts. Crystal structures of DAAOs show that instead of a positive charge, the positive end of a helix dipole points at the N1–C2=O locus of FAD.<sup>43,44</sup>

The reactions of several amino acids have been studied by both steady-state and stopped-flow methods. The overall kinetic mechanism depends on the substrate and the origin of the enzyme (Scheme 2). In all cases, the oxidized enzyme binds the substrate, and chemistry occurs in this complex, resulting in reduced flavin and the imine product. Afterward, catalysis can proceed by two pathways. In RgDAAO, which has a true ping-pong mechanism, the reduced flavin is reoxidized by molecular oxygen only after product release<sup>45</sup> (Scheme 2 (red)). This ping-pong mechanism is also seen with pkDAAO when using basic amino acids as substrates. When neutral amino acids are used, product release is slower than the reaction of the complex with oxygen,<sup>46,47</sup> so that H<sub>2</sub>O<sub>2</sub> is formed before imine dissociation, making the mechanism sequential rather than ping-pong (Scheme 2



**Scheme 2** D-amino acid oxidase kinetic mechanisms.

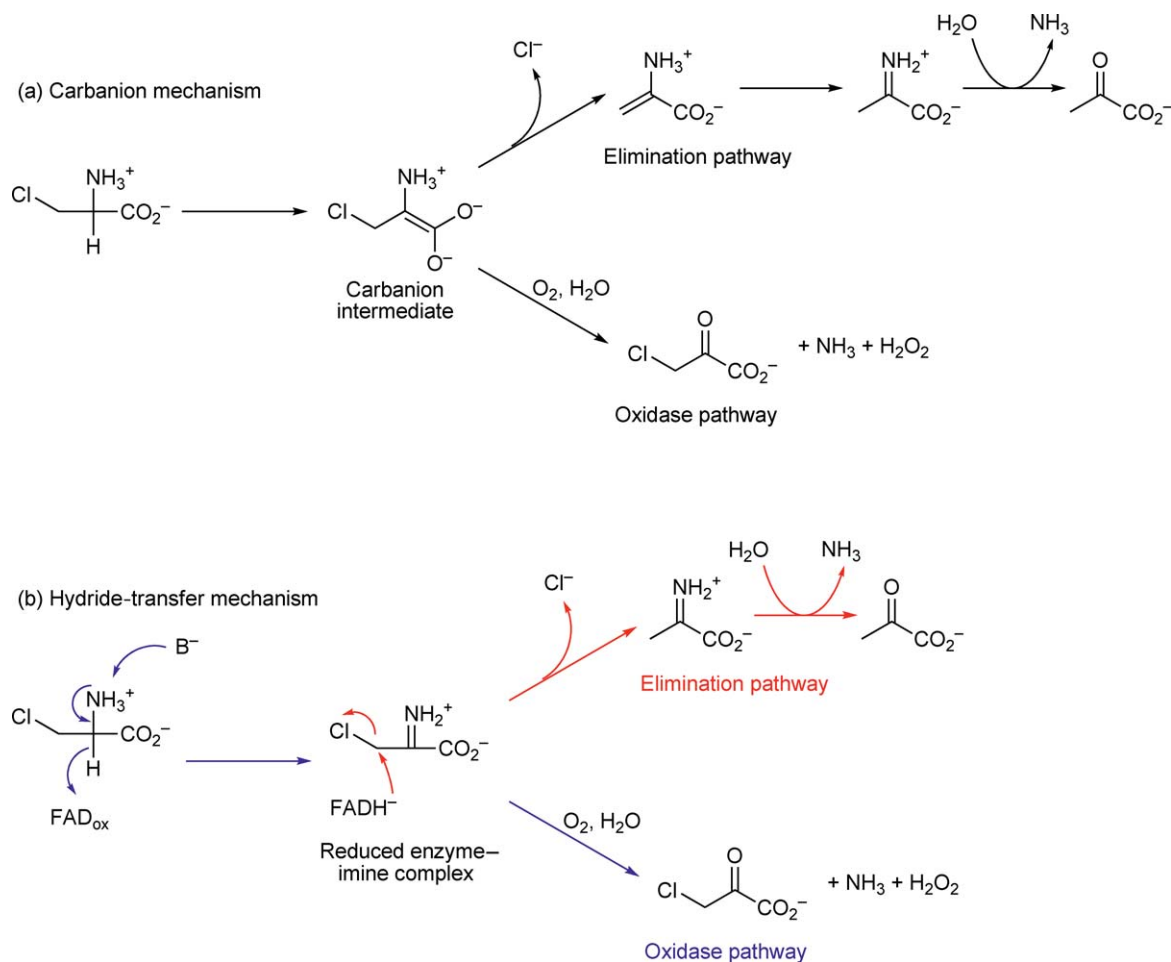
(blue)). However, in most cases, flavin reduction by amino acids is thermodynamically irreversible, resulting in parallel double-reciprocal plots despite the ternary complex.

Crystal structures rationalize the kinetic mechanisms of DAAOs. They show that RgDAAO does not contain the ‘lid’ that covers the active site of pkDAAO and is proposed to regulate product dissociation.<sup>48</sup> In RgDAAO, the side chain of Tyr238 might play the same role.<sup>44</sup> Stopped-flow studies have shown flavin reduction without intermediates and, in cases where the imine product dissociates slowly enough, an imine-reduced enzyme charge-transfer complex.<sup>45,47</sup>

The chemical mechanism of flavin reduction has been studied in many ways, including labeling experiments tracing the fate of the  $\alpha$ -hydrogen of substrates. Determining whether the  $\alpha$ -hydrogen is transferred directly to natural flavin is not experimentally feasible because proton exchange from N5 of reduced flavin is much faster than the techniques that could detect the label. To get around this, the flavin of DAAO has been replaced with 5-carba-5-deaza FAD. The isotope of  $\alpha$ -labeled amino acids is transferred almost quantitatively to the 5-position of the flavin analogue<sup>49</sup> and the flavin is reduced on its *re* face.<sup>50</sup> These experiments constrain any proposed mechanisms to those that ultimately transfer the  $\alpha$ -hydrogen of the substrate to N5 of the flavin of FAD.

In the early 1970s, Walsh investigated the reactions of  $\beta$ -chloro- $\alpha$ -amino acids as substrates.<sup>51,52</sup> When  $\beta$ -chloroalanine was used, two products were observed: chloropyruvate, which is expected in analogy to other amino acids, and pyruvate, suggesting the occurrence of an intermediate that eliminates chloride (**Scheme 3(a)**). In addition, the isotope effects were the same on the formation of chloropyruvate and pyruvate when  $\alpha$ -<sup>2</sup>H-chloroalanine was used as a substrate,<sup>51</sup> consistent with the C–H bond cleavage occurring in a common intermediate. These results suggested that an active site base abstracts the  $\alpha$ -proton from the amino acid substrate, forming a carbanion (**Scheme 4(a)**). This carbanion then reduces the flavin. The mechanism of flavin reduction by the putative carbanion was not addressed in these studies, but both radical and nucleophilic mechanisms have been proposed.<sup>53,54</sup> None of the proposed intermediates have been observed in stopped-flow experiments, so the rate-determining step of the sequence would have to be carbanion formation.

As an alternative to deprotonating the substrate  $\alpha$ -carbon, the oxidation of D-amino acids could occur by transfer of the  $\alpha$ -hydrogen as a hydride directly to N5 of the flavin without an intermediate, forming the imine product and reduced flavin<sup>49</sup> (**Scheme 4(b)**). Crystal structures support a hydride mechanism instead of a carbanion mechanism. Most strikingly, no structures have an active site base, a requirement for the carbanion mechanism.<sup>43,44,48,55,56</sup> If the substrate binds in a conformation similar to the product in the pkDAAO–imine–Trp complex, then the substrate would bind with its  $\alpha$ -hydrogen pointing toward N5 of FAD.<sup>48</sup> These structures are major evidence for the hydride-transfer mechanism. By studying a series of *p*-substituted phenylglycines with TvDAAO, it was concluded that the transition state is symmetrical, with little or no charge buildup, and the  $\alpha$ -hydrogen is cleaved in concert with imine product formation.<sup>57</sup> Kinetic isotope effects have been used to support a direct hydride-transfer mechanism. When C–H bond cleavage is completely rate limiting, both a deuterium effect and an <sup>15</sup>N isotope effect are observed, indicating that both C–H bond cleavage and C–N double-bond formation occur in a single transition state.<sup>58,59</sup>



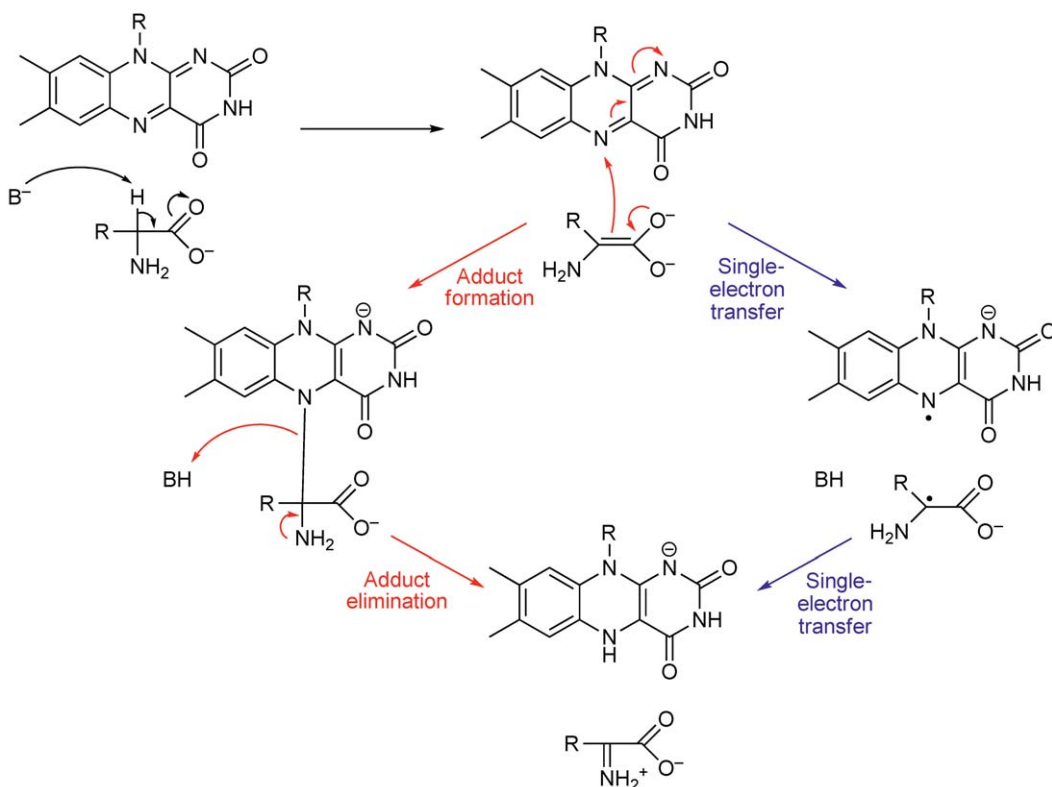
The quandary of chloride elimination can be reconciled with the hydride-transfer mechanism by invoking nucleophilic attack of a hydride from N5 of reduced flavin on the  $\beta$ -carbon, displacing chloride in an  $S_N2$ -like mechanism<sup>60</sup> (**Scheme 3(b)**). This mechanism predicts that the hydrogen removed from the  $\alpha$ -carbon of the  $\beta$ -chloroamino acid will be found on the  $\beta$ -carbon of the keto acid product. This has been observed for both  $\beta$ -chloroalanine and  $\beta$ -chloroaminobutyrate, although the transfer is not stoichiometric.<sup>51,52</sup> This mechanism also predicts that at low oxygen concentrations, the amount of chloride produced will increase; this is observed experimentally.<sup>51</sup>

### 7.03.2.1.2 Monoamine oxidase

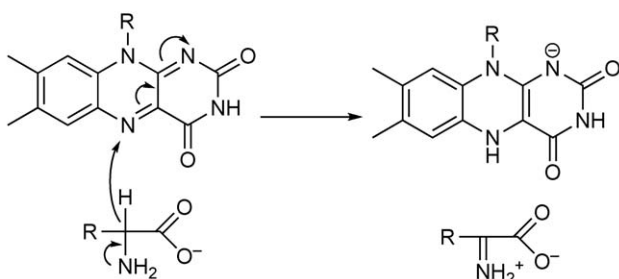
Monoamine oxidase (MAO) is another flavoenzyme that catalyzes the oxidation of carbon–nitrogen bonds. MAO has been studied extensively due to its physiological importance in the catabolism of amine neurotransmitters, such as norepinephrine, serotonin, and dopamine. There are two isozymes, MAO A and MAO B. The roles of both MAO A and MAO B are well documented in age-dependent neurodegenerative diseases and MAO inhibitors have been used to treat Parkinson's disease and depression.<sup>61</sup>

In MAO, the flavin is covalently attached to the enzyme at the 8- $\alpha$ -position via an active site cysteine residue.<sup>62</sup> MAO catalyzes the oxidation of several primary, secondary, and tertiary alkyl and arylalkyl amines to the corresponding imines, which are subsequently hydrolyzed nonenzymatically to the corresponding aldehydes or ketones.<sup>63–66</sup> MAO A and MAO B are 70% identical in sequence but vary in substrate specificity.<sup>67</sup>

## (a) Carbanion mechanism



## (b) Hydride-transfer mechanism



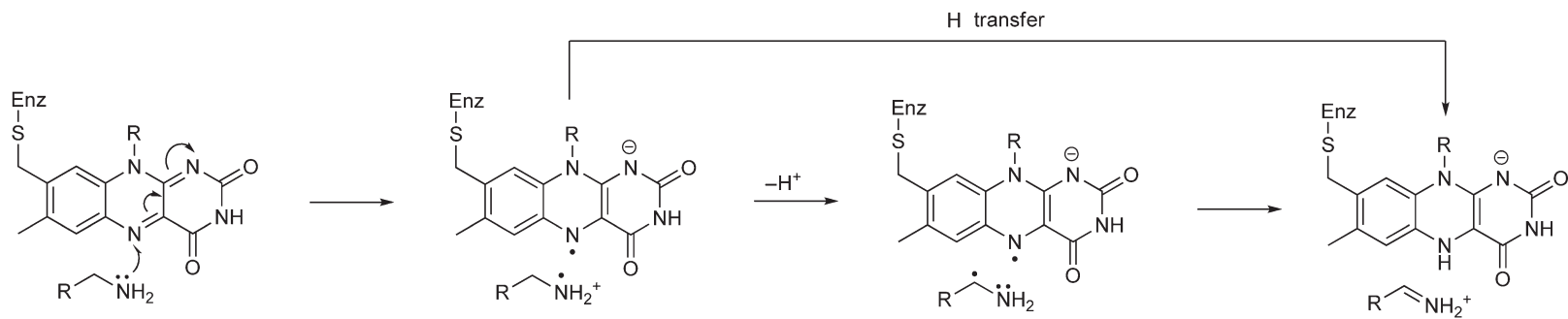
**Scheme 4** Alternative mechanisms of amino acid oxidation by D-amino acid oxidase.

Crystal structures have been solved for both human isozymes and show that the size of the active site cavity differs between the two enzymes,<sup>68–71</sup> resulting in the differing substrate specificities.

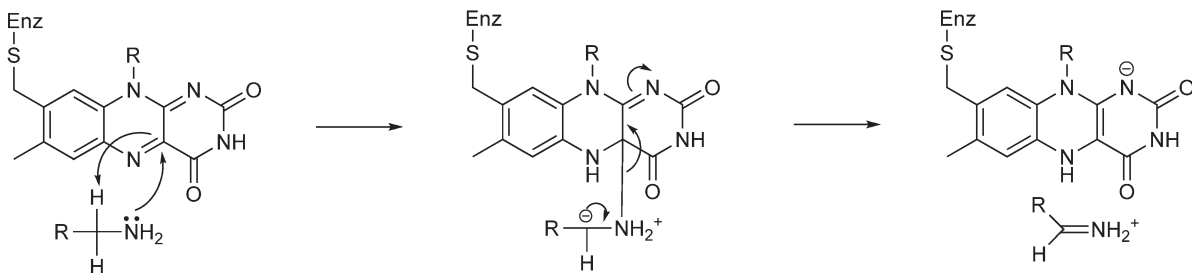
The kinetic mechanism of MAO is similar to that of DAAO – it depends on the substrate. In all cases, the amine binds to the enzyme and is oxidized. In some cases, the product dissociates before the enzyme is reoxidized by molecular oxygen, giving a ping-pong mechanism. In other cases, product dissociation is slow and the reduced enzyme–product complex reacts with molecular oxygen, forming the oxidized enzyme–product complex, followed by product dissociation.<sup>72</sup>

The chemical mechanism of the reductive half-reaction (the oxidation of the amine) is still controversial even after intense study. Silverman *et al.* proposed a radical mechanism after studying enzyme inactivation by mechanism-based inhibitors, several of which contain cyclopropyl groups or other moieties that rapidly rearrange when radicals are generated on amino substituents.<sup>73</sup> The first step of the proposed mechanism is a single-electron transfer from the amine nitrogen to the flavin (**Scheme 5(a)**). This can be followed by loss of a proton forming a carbon-centered radical that transfers a second electron to the flavin. Alternatively, following

(a) Radical mechanism



(b) Nucleophilic mechanism



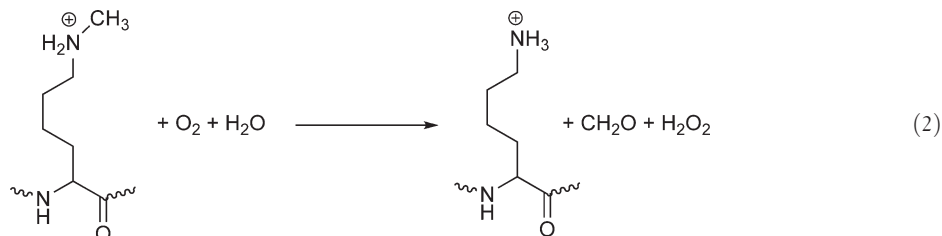
**Scheme 5** Alternative mechanisms of amine oxidation by monoamine oxidase.

the initial single-electron transfer step, a hydrogen atom could be transferred to the flavin, bypassing the formation of a carbon-centered radical. MAO is inactivated by aminomethylcubane, consistent with the formation of a carbon-centered radical.<sup>74</sup> However, stopped-flow experiments show no spectral evidence for a flavin semiquinone radical, which is required for this mechanism.<sup>72,75</sup> In addition, there is no EPR evidence of radicals during steady-state turnover.<sup>76</sup> It is still possible that these radical intermediates are present during catalysis, however, only transiently and at low concentrations (below detection limits) for kinetic reasons. Recently, spectroscopic evidence of a protein radical in partially reduced MAO A was reported. The tyrosyl radical develops following single-electron transfer to MAO A from dithionite and is present in equilibrium with the flavin semiquinone.<sup>77</sup> However, the detailed catalytic competence of this radical pair was not established.

As an alternative to radical mechanisms, the amine could attack the flavin as a nucleophile, forming a C4a adduct, which collapses to form the imine product and reduced flavin.<sup>54</sup> This mechanism is supported by model chemistry. The reaction of 2-phenylcyclopropylamine with 3,5-dimethylflavin produces a flavin adduct, which decays to the ring-opened aldehyde.<sup>78</sup> Using a series of ring-substituted benzylamines as substrates, Edmondson and coworkers concluded that the kinetics of oxidation by both MAO A and MAO B were consistent with a nucleophilic mechanism.<sup>79,80</sup> Large deuterium isotope effects were observed with benzylamines. To account for this, it was proposed that C–H bond cleavage and adduct formation occur in concert, with N5 of the flavin acting as a base<sup>79</sup> (Scheme 5(b)). The crystal structures of both MAO A and MAO B show two important tyrosyl residues, approximately perpendicular to the flavin, and have been studied by site-directed mutagenesis. These residues have been proposed to increase the nucleophilicity of the substrate amine by providing an environment that favors polarization of the substrate amine lone pair, consistent with the nucleophilic mechanism.<sup>81</sup>

#### 7.03.2.1.3 Lysine-specific demethylase 1

Histone modifications, such as phosphorylation, acetylation, and methylation, regulate gene expression.<sup>82–84</sup> Histone lysine methylation was considered a static modification, until relatively recently, when histone demethylases were discovered.<sup>4,85</sup> One such enzyme is LSD1, which belongs to the amine oxidase superfamily. LSD1 catalyzes the demethylation of mono- and disubstituted Lys 4 of histone H3<sup>4,86,87</sup> in a reductive half-reaction very similar to those catalyzed by DAO and MAO by oxidizing the amino group of the methylated lysine to the corresponding imino product, which hydrolyzes nonenzymatically to formaldehyde (Equation (2)).

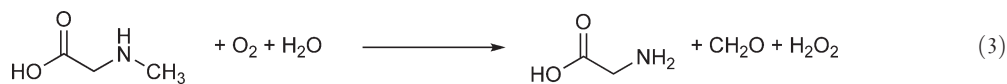


The activity of LSD1 is affected by other modifications on a histone H3 peptide, such as phosphorylation of Ser10, which abolishes LSD1 activity.<sup>88,89</sup> LSD1 can be inactivated by some of the same suicide inhibitors that inactivate MAO, such as *trans*-2-phenylcyclopropylamine (2-PCPA);<sup>90</sup> 2-PCPA shows little selectivity for MAO A and MAO B over LSD1, 2.4-fold and 16-fold higher, respectively.<sup>91</sup> The reaction of suicide inhibitors designed to inhibit MAO with LSD1 (or other amine oxidases) could be a cause of adverse side effects experienced with some drugs. Inhibitors have been designed specifically for LSD1 inactivation by creating peptide substrate analogues. One such analogue, in which the nitrogen atom of Lys 4 was derivatized with a propargyl group, acts as a suicide inhibitor.<sup>92</sup> By biotinylating this inactivator, LSD1 could be purified from nuclear extracts.<sup>93</sup> This could serve as a proteomic tool for studying epigenetics.

#### 7.03.2.1.4 Monomeric sarcosine oxidase

Monomeric sarcosine oxidase (MSOX) is a flavoenzyme that catalyzes the oxidative demethylation of sarcosine (*N*-methylglycine) to glycine (Equation (3)).





The reductive half-reaction is again similar to that of DAAO and MAO and includes the oxidation of sarcosine to the corresponding imino product, which is hydrolyzed to form glycine and formaldehyde. Sarcosine, a common soil metabolite, can be used as the sole source of carbon and energy for many microorganisms.<sup>94</sup>

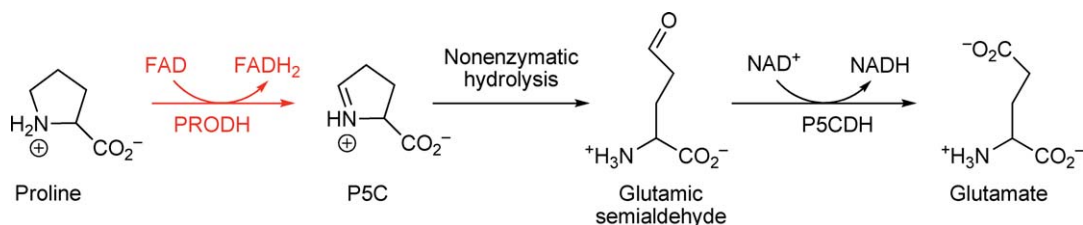
The flavin in MSOX is covalently attached to the protein at the 8- $\alpha$ -position through the sulfur of an active site cysteine residue.<sup>95</sup> In this instance, a dual role has been proposed for covalent flavin attachment – it both prevents loss of the weakly bound oxidized flavin and modulates the redox potential.<sup>96</sup>

Along with sarcosine, MSOX from *Bacillus* sp. B-0618 can use *N*-methyl-L-alanine, *N*-ethylglycine, and L-proline as substrates. In the steady state, the reduced enzyme/product complex is oxidized by O<sub>2</sub> prior to product dissociation.<sup>97</sup> The carboxylate of MSOX substrates is essential for binding, while the amino group is not.<sup>98</sup> The MSOX/sarcosine complex is characterized by a large charge-transfer absorbance at long wavelengths ( $\lambda_{\text{max}} \sim 520$  nm), suggesting that the pK<sub>a</sub> of the enzyme-bound substrate must be significantly lower than that of the free amino acid, since sarcosine can only act as a charge-transfer donor in its anionic form.<sup>99</sup>

The mechanism of flavin reduction in MSOX is still unknown. Several possible reductive half-reactions could be imagined. The reaction could occur by single-electron transfers similar to the reaction proposed for MAO. Transient kinetic studies of MSOX have shown no spectral evidence for a flavin semiquinone during the anaerobic reductive half-reaction,<sup>97,99</sup> but this does not rule out this intermediate. MSOX is inactivated by *N*-cyclopropylglycine, a mechanism-based inhibitor that can rapidly rearrange and modify the enzyme following a single-electron transfer. However, the rate of this inactivation is very slow, calling into question the single-electron transfer mechanism with the natural substrate.<sup>100,101</sup> The reaction could proceed through nucleophilic addition of the substrate amine at the C4a-position of the flavin forming an enzyme–substrate covalent intermediate. To probe the possibility that reduction occurs via a nucleophilic attack at C4a, the reaction of thioglycolate was examined. Thiols are known to form C4a adducts,<sup>102</sup> but MSOX and thioglycolate (HSCH<sub>2</sub>CO<sub>2</sub><sup>-</sup>) do not form a covalent flavin adduct. In fact, thioglycolate binds to MSOX and donates a single electron, forming flavin semiquinone.<sup>97</sup> The reductive half-reaction of MSOX could occur through hydride transfer to N5 of the oxidized flavin. Reaction at N5 was considered by studying a model reaction – that of sulfite, which probes the electrophilicity of N5 in flavoenzymes. MSOX makes an N5 sulfite adduct, but there is a significant kinetic barrier against sulfite addition.<sup>98</sup> ThiO, a glycine oxidase involved in thiamine biosynthesis, is structurally homologous to MSOX and DAAO. Interestingly, ThiO is proposed to use a hydride-transfer mechanism based on kinetic and structural studies.<sup>103</sup>

### 7.03.2.1.5 Proline dehydrogenases

Proline is oxidized in the biosynthesis of glutamate by a pair of related enzymes, proline dehydrogenase (PRODH) and  $\Delta^1$ -pyrroline-5-carboxylate dehydrogenase (P5CDH). PRODH is an FAD-containing enzyme that catalyzes the oxidation of proline to  $\Delta^1$ -pyrroline-5-carboxylate (P5C), which then hydrolyzes nonenzymatically to glutamic semialdehyde. The glutamic semialdehyde is further oxidized to glutamate in an NAD<sup>+</sup>-dependent reaction catalyzed by P5CDH<sup>104,105</sup> (Scheme 6). In eukaryotes and some bacteria, these two enzymes are independent proteins.<sup>106–110</sup> However, in some bacteria, such as *Escherichia coli* and *Salmonella typhimurium*, the enzymes are fused into a bifunctional enzyme known as proline utilization A (PutA).<sup>111</sup> Kinetic



**Scheme 6** Proline oxidation pathway.

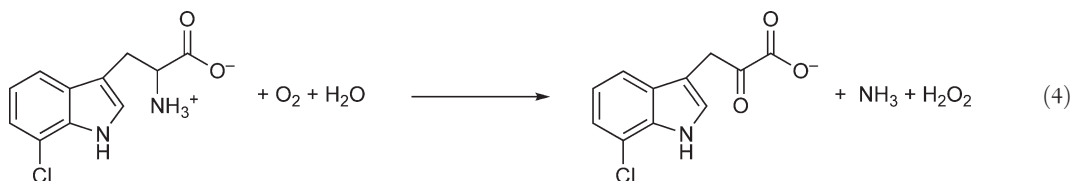
analysis of PutA from *S. typhimurium* suggests that an intermediate, P5C or possibly glutamic semialdehyde, is channeled between the two active sites via a tunnel.<sup>112</sup>

The PRODH monofunctional enzyme and the PRODH domain of PutA have  $(\beta\alpha)_8$ -barrel catalytic cores. FAD is in markedly different conformations in the monofunctional enzyme and the PRODH domain of PutA.<sup>110,113,114</sup> Modeling the substrate proline into the active site of the PRODH domain of PutA suggests that the hydrogen at the 5-position of proline points toward N5 of the flavin, consistent with a hydride-transfer mechanism.<sup>113</sup>

In addition to being an enzyme, PutA is also a transcription factor. The N-terminus of PutA from *E. coli* contains a ribbon-helix-helix motif, which enables the enzyme to bind DNA.<sup>115,116</sup> PutA acts as a transcriptional repressor of the *putA* and *putP* genes and its function is controlled by the redox state of the PRODH domain. Reduction of the FAD prosthetic group induces PutA to bind to the membrane, which disrupts the PutA-DNA complex and activates *put* gene expression.<sup>117-119</sup> Interestingly, the 2'-OH ribityl group of FAD acts as a molecular switch in PutA as it toggles between two unique positions. In the oxidized enzyme, the 2'-OH hydrogen bonds to an Arg residue on the  $8\alpha$ -helix. However, in the reduced enzyme, the 2'-OH hydrogen bonds to a backbone nitrogen between  $\beta$ -strand 5 and the  $\alpha 5a$ -helix.<sup>120</sup> This is an excellent example of the ribityl side chain playing an important function in a flavoenzyme.

#### 7.03.2.1.6 *RebO*

Rebeccamycin is a natural product that inhibits DNA topoisomerase I<sup>121</sup> and has been studied as a potential anticancer agent because of the importance of DNA topoisomerase I in cell growth and proliferation. Rebeccamycin analogues are being used in clinical trials for the treatment of neoplastic tumors,<sup>122</sup> renal cell cancer,<sup>123</sup> and leukemia,<sup>124</sup> *rebO* is one of the 11 genes in the rebeccamycin biosynthetic cluster. The protein product, RebO, is an L-amino acid oxidase that contains a noncovalently bound FAD.<sup>125</sup> RebO catalyzes the oxidation of 7-chloro-L-tryptophan in an early step in rebeccamycin biosynthesis (Equation (4)).



RebO also catalyzes the oxidation of 1-methyl-L-tryptophan and 5-fluoro-L-tryptophan, although at much slower rates.<sup>125</sup> The ability to oxidize other substrates suggests that the existing biosynthetic cluster could be used to create new antibiotics.

#### 7.03.2.1.7 *Pyridine nucleotide oxidations*

The reactions of pyridine nucleotides with flavoproteins are so widespread that they merit separate consideration. NADH and NADPH are general cellular reducing agents ( $E_{m7} = \sim -320$  mV), driving biologically important oxidative half-reactions to completion. Model studies show that reduction is by hydride transfer rather than through radicals.<sup>126,127</sup> Flavoenzymes are absolutely stereospecific for the diastereotopic hydride to be transferred from NAD(P)H, with the unusual exception of flavin-dependent thymidylate synthase, which preferentially transfers the *pro-R* hydride but also transfers the *pro-S* hydride to a lesser extent.<sup>128</sup> Dihyronicotinamides reduce flavins nonenzymatically by first forming weak complexes,<sup>129</sup> with dissociation constants of model compounds near  $0.1 \text{ mol l}^{-1}$ . Reduction in the complex is as rapid as in many enzymes, ranging from  $\sim 1$  to  $\sim 200 \text{ s}^{-1}$ , depending on redox potentials. Tethering the isoalloxazine moiety to a dihyronicotinamide<sup>130</sup> speeds up the reactions up to  $\sim 350 \text{ s}^{-1}$ , rivaling the rates of many NAD(P)H/flavoenzyme reactions and suggesting that proteins accelerate the reaction of NAD(P)H and flavins by simply bringing the two reactants together – specific catalytic groups are unnecessary. This is born out in the diversity of active site structures that cause rapid pyridine nucleotide-flavin reactions. However, quantum mechanical hydrogen tunneling promoted by protein dynamics has been proposed to be a key component in flavin-dihyronicotinamide reactions, suggesting a mechanical role for the protein structure.<sup>131</sup>

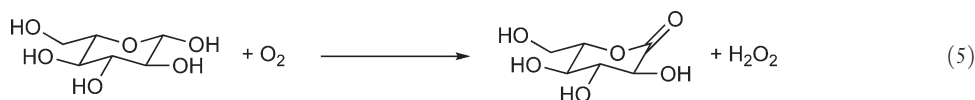
Model dihydronicotinamide-oxidized flavin complexes and nicotinamide–flavin hydroquinone complexes exhibit charge-transfer absorbance,<sup>130</sup> mimicking the behavior of many enzymes. The  $pK_a$  of N1 of the dihydroisoalloxazine in a tethered nicotinamide model decreased by 1.7 units as a result of contact with the positive charge of the nicotinamide moiety, and the two-electron potential became more positive by 116 mV. Similar large effects are likely to occur in enzymes but have not been explored extensively. Sometimes, the orbital overlap reported by charge-transfer absorbance is proposed to be an inherent part of the mechanism of hydride transfer. However, many flavoenzymes react rapidly with pyridine nucleotides without exhibiting observable charge-transfer absorbance. Charge-transfer complexes could be present in the reaction pathway but do not accumulate for kinetic reasons, making them unobservable. Studies on an enzyme related to ferredoxin reductase contradict this explanation.<sup>132</sup> At high pH, N3 of the flavin deprotonates (evident from large changes in the absorbance spectrum) and the more electron-rich oxidized isoalloxazine no longer forms charge-transfer complexes with NADPH as it does at lower pH. Nonetheless, reduction is rapid, and crystal structures show that pyridine nucleotides bind in the same orientation at high and low pH values, strongly suggesting that the orbital overlap responsible for charge-transfer absorbance is not a part of the hydride-transfer mechanism.

The minimum requirement for promoting hydride transfer between flavins and pyridine nucleotides is to bring C4 of the dihydronicotinamide and N5 of the flavin together in van der Waals contact; the relative orientations of the two rings can vary markedly. It is often assumed that the nicotinamide and isoalloxazine rings stack parallel to each other, but this is not necessary, either for reaction or charge-transfer interaction. For instance, there is a  $30^\circ$  angle between the rings in the NADP<sup>+</sup> complex of ferredoxin–NADP<sup>+</sup> reductase.<sup>133</sup> Interestingly, in many structures, the nicotinamide moiety is blocked from contact with the flavin by a protein side chain, or the nicotinamide is not visible in the electron density, suggesting high mobility; conformational changes to attain the reactive configuration appear to be very common. A variety of pyridine nucleotide conformations are observed in flavoenzyme complexes, ranging from extended conformations bound to Rossman folds to conformations in which the adenine and nicotinamide rings stack.<sup>134</sup> The structures of pyridine nucleotides bound to aromatic hydroxylases have remained largely elusive, but data collected on a mutant of PHBH suggest that the adenine ring of NADPH stacks with the *si* face of the isoalloxazine and the dihydronicotinamide rings stack with the *re* face.<sup>135</sup> Many flavoenzymes that bind FAD and pyridine nucleotides have been grouped into the ‘two nucleotides binding domains’ superfamily.<sup>136</sup> An analysis of 1600 structures and sequences in this superfamily suggests that the structural constraints imposed by bringing the two nucleotides together in a reactive complex constrain the locations allowed for the site of the other half-reaction, which is often at the juncture with another protein domain or subunit.

### 7.03.2.2 Carbon–Oxygen Oxidation

#### 7.03.2.2.1 Glucose oxidase

Glucose oxidase (GOX) oxidizes  $\beta$ -D-glucose to D-glucono- $\delta$ -lactone (Equation (5)) with the concomitant reduction of the enzyme-bound flavin.



The reduced enzyme is oxidized by molecular oxygen yielding hydrogen peroxide and oxidized flavoprotein. Several microorganisms contain GOXs; however, the enzyme from *Aspergillus niger* is the most thoroughly studied. GOX is a homodimer, with each monomer binding a single FAD molecule.<sup>137</sup> GOX is a glycoprotein (both *N*- and *O*-glycosylated) with high mannose content, which accounts for 10–16% of the total molecular weight.<sup>138–141</sup> Removing the carbohydrates does not have a significant effect on catalytic activity or stability.<sup>140,142</sup> However, in a detailed study using kinetic isotope effects, it was determined that less glycosylation in GOX results in a greater degree of quantum mechanical tunneling during the reduction reaction.<sup>143</sup>

The structure of ligand-free GOX has been solved.<sup>144</sup> FAD is found at the bottom of a deep pocket formed partly from the dimer interface. Kinetic studies have shown that glucose binds extremely weakly, if at all, to

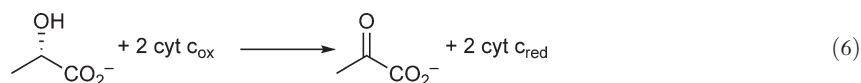
GOX.<sup>145</sup> Accordingly, the structure shows very few protein interactions available for substrate binding.<sup>144</sup> By replacing the flavin with 8-hydroxy-5-carba-5-deaza FAD and determining the stereochemistry of the reaction, it was shown that glucose reacts on the *re* face of the flavin.<sup>146</sup> Modeling the substrate in the active site with the 1-hydrogen of glucose near the *re* face of the flavin N5 places the 1-hydroxyl of glucose near His516, which is ideally positioned for deprotonation.<sup>147</sup> Mutating this His to an Ala decreases the activity of GOX four orders of magnitude, consistent with it being an important active site base.<sup>147</sup> In the crystal structure, the oxidized flavin is held in a bent conformation by hydrogen bonds at the pyrimidine end and nonpolar contacts at the xylene end.<sup>144</sup>

In the steady state, glucose is oxidized in a ping-pong mechanism. Stopped-flow studies on the reductive half-reaction using glucose as a substrate show that the observed pseudo-first-order rate constant never saturates up to the limit of experimentally attainable concentrations. Similar results, although with much lower rate constants, were obtained using mannose, xylose, and galactose as substrates.<sup>145</sup> Thus, there is no kinetic evidence of an enzyme–substrate complex with glucose in GOX. It is interesting to note that while the binding of these sugars cannot be detected in the reductive half-reaction, these same substrates have readily attainable  $K_m$  values in the steady state, nicely illustrating the fact that  $K_m$  is not a dissociation constant. Labeling the 1-position with deuterium causes a large primary isotope effect and a limiting rate constant in the reductive half-reaction is observed, indicating that an enzyme–substrate complex does in fact occur in GOX.<sup>148</sup>

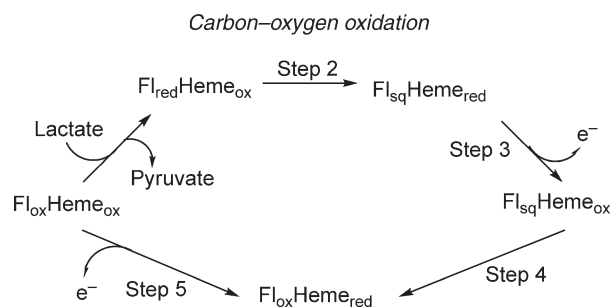
Interestingly, GOX oxidizes nitroalkanes.<sup>149</sup> The reaction proceeds through at least two significant pathways, leading to a mixture of products from nitroethane that include the major products nitrite and acetaldehyde, and the minor products nitrate and dinitroethane. During turnover, both neutral and anionic semiquinones are present, indicating a radical mechanism for nitroalkane oxidation. These findings are not necessarily relevant to physiological substrates, since both flavins and nitroalkanes can react by both radical and anionic mechanisms, but they illustrate the mechanistic plasticity of flavoenzymes.

### 7.03.2.2 Flavocytochrome $b_2$

Flavocytochrome  $b_2$  ( $Fb_2$ ) catalyzes the oxidation of L-lactate to pyruvate (Equation (6)).



$Fb_2$  is a member of a family of homologous flavoproteins that catalyze the oxidation of  $\alpha$ -hydroxy acids. It is located in the intermembrane space of yeast mitochondria and provides pyruvate for the Krebs cycle as well as participates in a short electron-transfer chain involving cytochrome  $c$  and cytochrome oxidase, making it an important respiratory enzyme. Unlike many other flavoproteins that oxidize  $\alpha$ -hydroxy acids,  $Fb_2$  has very poor reactivity toward oxygen.  $Fb_2$  is a homotetramer with each monomer containing both an FMN and a heme  $b_2$ .<sup>150–152</sup> The catalytic cycle has five redox steps (Scheme 7). In the first step, lactate is oxidized to pyruvate as the flavin of  $Fb_2$  is fully reduced to the hydroquinone. The reduced flavin transfers a single electron to the heme  $b_2$ , which then transfers the electron to cytochrome  $c$ . Another single-electron transfer results in oxidized flavin. Under saturating concentrations of ferricytochrome  $c$  and L-lactate, the turnover number is  $104 \text{ s}^{-1}$  (mol of lactate per mol of enzyme) at  $25^\circ\text{C}$ .<sup>153</sup> The oxidation of lactate (step 1) occurs at  $600 \text{ s}^{-1}$  at  $25^\circ\text{C}$ .<sup>153</sup> The transfer of an electron from the hydroquinone to heme  $b_2$  (step 2) has been studied using



**Scheme 7** Catalytic cycle of flavocytochrome  $b_2$ .

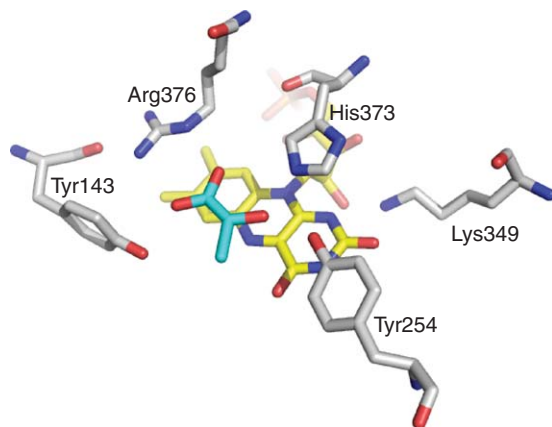
stopped-flow techniques and is very fast, estimated to be  $>1500\text{ s}^{-1}$  at  $25\text{ }^{\circ}\text{C}$ .<sup>154–158</sup> The cytochrome *c* reduction steps (steps 3 and 5) have been studied in detail and do not affect the rate of turnover.<sup>159</sup> The rate-limiting step in the reaction of  $Fb_2$  is the intramolecular electron transfer from the flavin semiquinone to the heme  $b_2$  (step 4); rapid-freeze quenched EPR showed about 75% of the enzyme in the semiquinone form during turnover.<sup>160</sup>

The mechanism of flavin reduction in  $Fb_2$  has been a topic of debate for decades. Soon after it was discovered that DAAO could catalyze the elimination of  $\text{Cl}^-$  from chlorinated substrates, very similar experiments were conducted with many  $\alpha$ -hydroxy acid oxidizing enzymes, including  $Fb_2$ . However, HCl elimination was not observed when using 3-Cl-lactate as a substrate.<sup>161</sup> Studies using 3-Br-pyruvate as an electron acceptor for lactate oxidation were interpreted as evidence of a carbanion mechanism of flavin reduction.<sup>162,163</sup>

Unlike DAAO, the crystal structure of  $Fb_2$  does not clarify the mechanism of flavin reduction. The monomer of  $Fb_2$  has a small N-terminal cytochrome domain and a larger C-terminal  $\alpha/\beta$  barrel containing FMN, connected by a short linker. The carboxylate of the product makes a salt bridge and a hydrogen bond to an Arg and a Tyr (Figure 3), residues also present in DAAO. Unlike in DAAO, the product in the crystal structure interacts with both Tyr254 and His373, either of which could be the active site base needed in a carbanion mechanism.<sup>164</sup> The substrate, lactate, has been modeled into the structure in two different conformations.<sup>165,166</sup> In one conformation, His373 is positioned to remove the  $\alpha$ -hydrogen from lactate to form a carbanion intermediate, while Tyr254 forms a hydrogen bond with the lactate hydroxyl. In the other conformation, His373 is positioned to abstract the substrate hydroxyl proton as a hydride is transferred to N5 of the flavin.

Mutations that broadened substrate tolerance were used to probe the mechanism of FMN reduction.  $Fb_2$  can oxidize (*S*)-2-hydroxyacids other than *L*-lactate; however, bulky substrates, such as mandelate, are poor.<sup>167</sup> The crystal structure shows that Ala198 and Leu230 interact with the methyl group of pyruvate, suggesting that a steric clash with these residues causes mandelate to be a poor substrate for  $Fb_2$ .<sup>164</sup> Rational mutagenesis was used to change the substrate specificity. Mandelate is the preferred substrate of the Ala198Gly/Leu230Ala double mutant. The crystal structure of the Leu230Ala single mutant with phenylglyoxylate (the product of mandelate oxidation) suggests that the binding mode of mandelate is incompatible with a carbanion mechanism, and instead exhibits the binding mode expected for a hydride-transfer mechanism,<sup>168</sup> indicating that at least in this mutant  $Fb_2$ , a hydride-transfer mechanism is utilized.

Primary isotope effects were measured using  $[^2\text{H}]$ lactate to study C–H bond cleavage and solvent isotope effects were used to study O–H bond cleavage. Isotope effects for the reaction of a mutant enzyme, Tyr254Phe, showed that the C–H and O–H bonds break in a single step, suggesting concerted hydride transfer and deprotonation in lactate oxidation.<sup>169</sup> Similar isotope effect experiments with wild-type  $Fb_2$  showed that the C–H and O–H bonds are not cleaved in the same step,<sup>170</sup> and were initially thought to provide evidence for a carbanion mechanism. However, the inverse solvent isotope effect and viscosity effect on  $V/K$ <sup>170</sup> could be interpreted as hydroxyl deprotonation being coupled to a conformational change that precedes hydride transfer to the flavin.<sup>59</sup> The reaction of 3-Br-pyruvate was also explained in terms of a hydride-transfer mechanism, similar to the mechanism proposed for DAAO.<sup>59</sup>

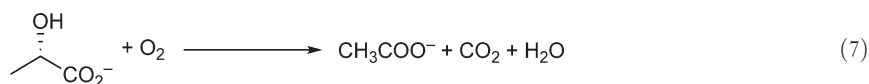


**Figure 3** Active site of yeast flavocytochrome  $b_2$  with pyruvate (cyan).



### 7.03.2.2.3 Lactate monooxygenase

L-lactate monooxygenase (LMO) catalyzes the reaction of L-lactate with  $O_2$ , producing acetate,  $CO_2$ , and  $H_2O$  (Equation (7)).



Though the enzyme is not a true oxidase, it catalyzes an oxidase-like reaction. The first step in catalysis is the oxidation of lactate to pyruvate. The two-electron reduced flavin then reacts with molecular oxygen producing  $H_2O_2$ . Because dissociation of pyruvate from the enzyme is a slow process, the  $H_2O_2$  formed at the active site by the oxidation of FMN decarboxylates pyruvate, forming the final products acetate,  $H_2O$ , and  $CO_2$ , which then dissociate. This oxidation/decarboxylation reaction can be uncoupled. In that instance, pyruvate dissociates before the oxygen reacts, producing pyruvate and  $H_2O_2$ . However, normal turnover in wild-type LMO is completely coupled. The uncoupled pathway can be mimicked by reducing the flavin with lactate and waiting for pyruvate dissociation before adding  $O_2$ .<sup>171</sup> Because of the oxidase-like mechanism, LMO has sometimes been referred to as lactate oxidase in the literature. However, it should be kept in mind that a true lactate oxidase exists (discussed later).

LMO from *Mycobacterium smegmatis* is the best-studied example of this enzyme. The enzyme is an octamer of identical subunits, with one FMN per monomer. Even though no structural information has been obtained for LMO, much can be inferred from the high similarity of LMO to other  $\alpha$ -hydroxy acid oxidizing flavoenzymes. LMO exhibits many common properties of flavoprotein oxidases. It makes an N5 sulfite adduct<sup>172,173</sup> and stabilizes anionic semiquinone<sup>12</sup> and the benzoquinoid form of 8-mercapto-FMN,<sup>42</sup> all of which are explained by the presence of a positive charge in the protein at the N1-C2=O locus. Sequence alignments show that Lys266 provides this positive charge in LMO. This was confirmed by site-directed mutagenesis.<sup>174</sup>

Redox potentials have been measured for both single-electron transfers in LMO. The oxidized enzyme-semiquinone couple has a redox potential of  $-67$  mV while the semiquinone-reduced enzyme couple has a redox potential of  $-231$  mV.<sup>175</sup> Pyruvate binds much more tightly to the semiquinone form of the enzyme than either the oxidized or reduced form, causing a considerable change in the redox potentials. In the presence of saturating pyruvate, the redox potential of the oxidized enzyme-semiquinone couple is  $+80$  mV while that of semiquinone-reduced enzyme couple is  $-370$  mV. Pyruvate dissociation from the semiquinone form of LMO is extremely slow and this pyruvate/semiquinone complex is essentially nonreactive with  $O_2$ . The rate of reoxidation observed is the rate of pyruvate dissociation.<sup>176</sup>

The kinetics of LMO has been extensively studied. In a reversal of what is observed in most flavoenzymes, the oxidized flavin in LMO is nonfluorescent, while the reduced flavin is fluorescent giving a useful signal for stopped-flow studies. The flavin is reduced by lactate at  $230$  s<sup>-1</sup> at pH 7.0 and 25 °C. The reduced enzyme-pyruvate complex formed exhibits long-wavelength charge-transfer absorbance and is nonfluorescent. Finally, at a rate too slow to be catalytically relevant, the fluorescence of free reduced enzyme appears as pyruvate dissociates. Thus, oxygen reacts with the reduced enzyme-pyruvate complex.<sup>171</sup>

LMO has each residue implicated in binding and catalysis in glycolate oxidase, and  $Fb_2$ .<sup>177</sup> In LMO, His290 is the proposed active base required in the carbanion mechanism. Tyr44 and Arg293 are proposed to interact with the substrate carboxylate. Tyr152 hydrogen bonds to the hydroxyl of the substrate and Lys266, mentioned earlier, is the positive charge at the N1-C2=O locus of FMN. The importance of each of these residues was demonstrated in an elegant series of stopped-flow experiments conducted with mutant enzymes.<sup>174,178,179</sup> Oxalate is a tight-binding inhibitor of wild-type LMO. The logarithms of the reduction rate constants for the wild-type and mutant enzymes were linearly correlated with the logarithms of the dissociation constants of oxalate, with a unit slope. Thus, enzymes that bound oxalate tightly also reduced the flavin quickly. This was taken to show that oxalate resembles the carbanion intermediate in the reductive half-reaction. However, oxalate could instead mimic a lactate alkoxide, an intermediate predicted in the hydride-transfer mechanism.<sup>59</sup>



### 7.03.2.2.4 L-lactate oxidase

L-lactate oxidase (LOX) catalyzes the oxidation of L-lactate to pyruvate with the concomitant reduction of the enzyme-bound FMN. The reduced enzyme is oxidized by molecular oxygen resulting in oxidized enzyme and H<sub>2</sub>O<sub>2</sub>. LOX has substantial sequence homology to other  $\alpha$ -hydroxy acid oxidizing flavoproteins, containing all the residues thought to be important for catalysis in this class of enzymes.<sup>180</sup>

As with most oxidases, LOX forms a stable semiquinone upon one-electron reduction. However, the pK<sub>a</sub> of the flavin semiquinone is 6.0 in LOX, making the neutral semiquinone experimentally accessible. As in LMO, pyruvate binds to the semiquinone form of LOX more tightly than either the oxidized or fully reduced form. The binding of pyruvate decreases the pK<sub>a</sub> of the semiquinone, so that neutral semiquinone is no longer observed.<sup>180</sup> The crystal structure of LOX shows that Lys241 is the positive charge located at the N1–C2=O locus responsible for this stabilization.<sup>181</sup> LOX also stabilizes a flavin N5-sulfite adduct, although sulfite binds about 10-fold less tightly to LOX as it does to LMO.<sup>180</sup> These findings suggest that LOX is less effective at stabilizing a negative charge at its active site than LMO.

The catalytic mechanisms of LOX and LMO are very similar, with one key exception. The rate of pyruvate dissociation is much faster in LOX than in LMO. LOX catalyzes a true ping-pong mechanism, which can be divided into two distinct half-reactions. In the reductive half-reaction, LOX is reduced as lactate is oxidized to pyruvate.<sup>180</sup> Flavin reduction occurs at 105 s<sup>-1</sup> at 4 °C; at 25 °C, the reaction is too fast to be measured. A fourfold primary isotope effect is observed on this step when using  $\alpha$ -<sup>2</sup>H-lactate as a substrate. The product of lactate oxidation is the reduced enzyme/pyruvate complex, which exhibits long-wavelength charge-transfer absorbance. Product dissociation from this complex occurs at 35 s<sup>-1</sup> about 7000-fold faster than pyruvate dissociation in LMO. In the oxidative half-reaction,<sup>180</sup> the free reduced enzyme, formed due to the quick release of pyruvate, is oxidized by molecular oxygen with a second-order rate constant of  $1.3 \times 10^6 \text{ mol}^{-1} \text{ l s}^{-1}$ .

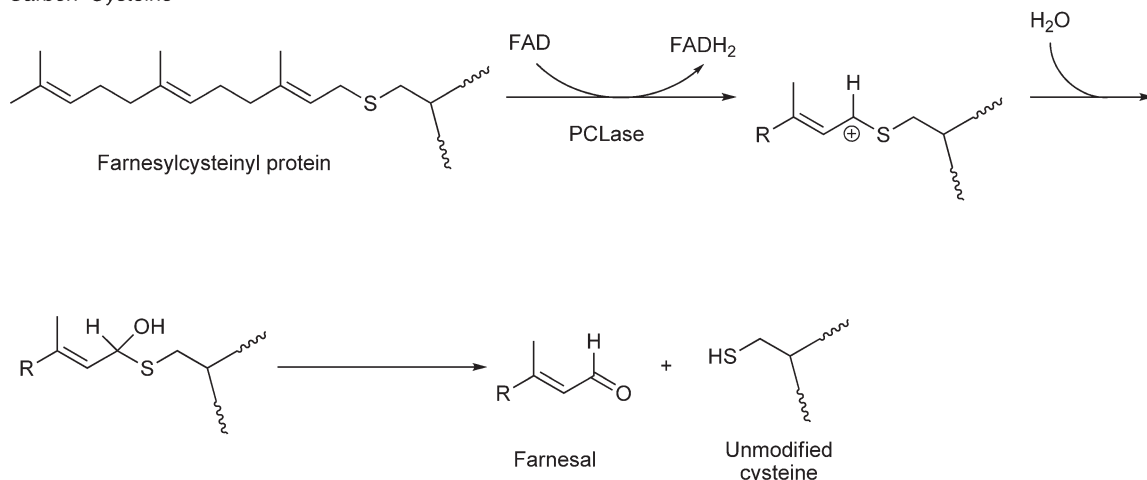
An elegant set of experiments examining the structure–function activity relationship for LOX were completed by substituting the native FMN with 12 FMN derivatives with various substitutions at the 6- and 8-positions of flavin.<sup>182</sup> The reduction rate constants and rate constants for sulfite adduct formation increased as the redox potential of the enzymes increased until a limiting value was obtained. Below the break point, the relationships of reduction rate constants and rate constants for sulfite adduct formation with the two-electron redox potential indicated significant negative charge development during the reaction, consistent with either a hydride-transfer or carbanion mechanism. The lack of variation with redox potentials above the break point indicates a change in rate-determining step, but the origin of this effect is not known.

### 7.03.2.3 Carbon–Sulfur Oxidation

#### 7.03.2.3.1 Prenylcysteine lyase

Proteins are subject to several different modifications *in vivo*. One such modification is prenylation, in which a 15-carbon farnesyl or a 20-carbon geranylgeranyl isoprenoid is attached to the C-terminus of specific proteins.<sup>183,184</sup> Usually, these proteins contain a so-called CaaX motif, which is defined as a cysteine residue fourth from the C-terminus of the protein to which the isoprenoid group is attached via a thioester linkage.<sup>185</sup> Prenylcysteine lyase (PCLase) cleaves prenylcysteines to form free cysteine and an isoprenoid aldehyde.<sup>186,187</sup> PCLase contains a noncovalently bound FAD.<sup>187</sup> In the proposed mechanism,<sup>187</sup> a hydride is transferred from C1 of the isoprenoid moiety of farnesylcysteine to FAD (**Scheme 8**), forming a sulfur-stabilized carbocation and reduced flavin. A hemithioacetal intermediate is formed by nucleophilic attack of a water molecule on the carbocation. This intermediate collapses to the isoprenoid aldehyde and free cysteine. The reduced flavin reacts with molecular oxygen to produce hydrogen peroxide and regenerates the oxidized enzyme. Only a little mechanistic work has been done, and not by directly observing the reactions of the flavin. Farnesylcysteine deuterated at either the *pro-S* or *pro-R* position of C1 established that the *pro-S* hydrogen is transferred during the reaction.<sup>188</sup> The turnover number of PCLase is 0.008 s<sup>-1</sup> at pH 7.4 and 25 °C.<sup>188</sup> While this is slow compared to other carbon–heteroatom oxidases, it is similar to the turnover numbers observed for the prenyltransferases, which attach the isoprenoid moiety to the cysteine.<sup>189,190</sup>

## Carbon–Cysteine



**Scheme 8** Proposed mechanism of prenylcysteine lyase.

### 7.03.3 Oxidation and Reduction of Carbon–Carbon Bonds

Several flavoproteins catalyze the oxidation or reduction of carbon–carbon bonds. Often, the substrates for these enzymes are  $\alpha,\beta$ -unsaturated carbonyl compounds. The mechanism used to catalyze these types of reactions involves hydride transfers to or from N5 of the flavin. In addition, an active site residue is involved in the acid/base chemistry needed for the reaction to occur. Several of these enzymes are involved in cellular energy metabolism.

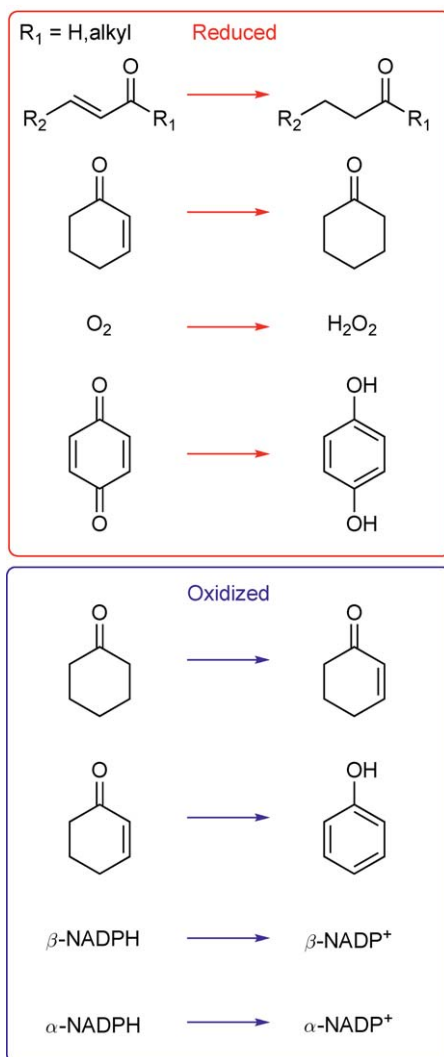
#### 7.03.3.1 Old Yellow Enzyme

Old yellow enzyme (OYE) was the first flavoenzyme to be discovered and purified. It was isolated from yeast in 1933 and has proven to be a diverse catalyst (**Scheme 9**). The physiological role of OYE is still uncertain, although it has been suggested to be involved in the oxidative stress response.<sup>191,192</sup> The enzyme has been detected in several species of yeast, and OYE-like enzymes – mostly of unknown function – are found in a multitude of bacteria. The most thoroughly studied OYEs are from *Saccharomyces cerevisiae* and *Saccharomyces carlsbergensis*. The enzyme is a dimer of 45 kDa subunits, with each monomer binding a single FMN. Brewer's bottom yeast has two OYE isozymes. Thus, when purified from yeast, OYE is a mixture of homodimers and heterodimers of different isozymes. Mixtures of homogeneous preparations of the cloned isozymes stored in the cold exchange subunits, showing that the dimers slowly dissociate spontaneously.<sup>193</sup>

OYE binds many different ligands including inorganic anions,<sup>193</sup> phenols,<sup>194</sup> and steroids.<sup>195,196</sup> OYE isolated from brewer's bottom yeast is green<sup>197</sup> due to charge-transfer absorbance from the complex of *p*-hydroxybenzaldehyde with OYE.<sup>198</sup> Investigations of the charge-transfer absorbance of OYE complexes played an important role in understanding this common spectral phenomenon. It was shown that the maximal wavelength of the charge-transfer absorbance produced by the binding of a series of *p*-substituted phenols to OYE correlates with the Hammett *para* constant of the substituent<sup>194</sup> and with the one-electron redox potential of the enzyme-bound flavin.<sup>199</sup> The crystal structures of OYE complexed with *p*-iodophenol and *p*-hydroxybenzaldehyde clearly show the parallel stacking of the ligand and the flavin.<sup>200</sup> The ligand binding site is mostly hydrophobic, but His191 and Asn194 hydrogen bond to the phenolate oxygen and stabilize the negative charge while Tyr375 hydrogen bonds to the carbonyl of *p*-hydroxybenzaldehyde.

The physiological reaction catalyzed by OYE has been a long-standing mystery. OYE is often assayed as an NADPH oxidase. However, its reaction with molecular oxygen is quite slow for an oxidase ( $10^2$ – $10^3$  mol<sup>-1</sup> l s<sup>-1</sup>), casting doubt on the relevance of O<sub>2</sub> as a substrate. Curiously, both  $\alpha$ - and  $\beta$ -NADPH rapidly reduce OYE, with the unnatural  $\alpha$ -isomer being a better substrate.<sup>201</sup> The crystal structure of OYE complexed with an unreactive NADPH analogue shows that C4 of the nicotinamide moiety lies over N5 of the flavin, as required

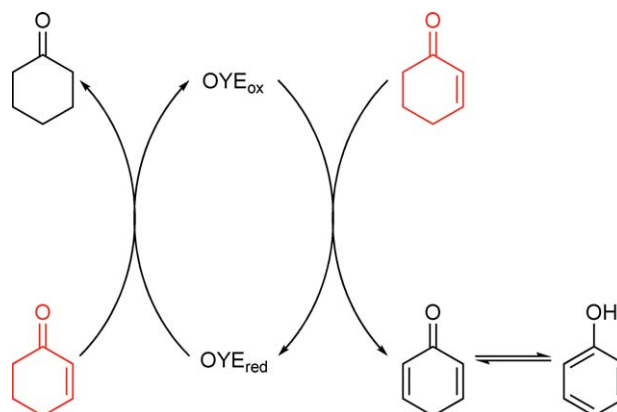
## Carbon-carbon oxidation



**Scheme 9** Various substrates of old yellow enzyme.

in a reactive complex.<sup>200</sup> Interestingly, the adenosine portion of the NADPH analogue is disordered in the structure, suggesting that it does not play a role in binding.

OYE catalyzes the NADPH-dependent reduction of quinones,<sup>202</sup> many  $\alpha,\beta$ -unsaturated carbonyl compounds,<sup>193,196</sup> conjugated nitroalkenes, and nitrate esters.  $\alpha,\beta$ -Unsaturated ketones or aldehydes are the only enones to be reduced; esters, carboxylates, and amides are not. Interestingly, 1,2-cyclohexadione is readily reduced by OYE, while 1,3-cyclohexadione is reduced poorly. Presumably, enol (or enolate) tautomers – both  $\alpha,\beta$ -unsaturated carbonyls – are the reactive species of these diketones. 1,4-Cyclohexadione, a compound incapable of forming an  $\alpha,\beta$ -unsaturated enol tautomer, is not a substrate.  $\alpha,\beta$ -Unsaturated carbonyl compounds are presumably reduced by hydride transfer from the flavin N5 to the  $\beta$ -carbon of the enone, producing either an enolate intermediate or a transition state with enolate character if  $\alpha$ -protonation is concerted with hydride transfer. The crystal structure<sup>200</sup> shows that Tyr196 is positioned to serve as an acid in the protonation of the  $\alpha$ -carbon. Interestingly, when  $\alpha,\beta$ -unsaturated ketones or aldehydes are used as substrates, the hydride derived from NADPH is transferred to the  $\beta$ -carbon of the product, even though the reaction is ping-pong and free reduced flavin is an intermediate during turnover. Therefore, proton exchange with solvent at N5 of the reduced flavin is much slower than oxidation of the flavin by good  $\alpha,\beta$ -unsaturated substrates.



**Scheme 10** Disproportionation reaction of cyclohexenone catalyzed by old yellow enzyme.

The mechanism discussed above allows OYE to catalyze the dismutation of six-membered cyclic  $\alpha,\beta$ -unsaturated ketones<sup>196</sup> to the corresponding phenols and saturated ketones (**Scheme 10**). Oxidized OYE is reduced by cyclohexenone, producing reduced flavin and cyclohexadienone, the latter being the *keto* tautomer of the thermodynamically favored phenol. The reduced enzyme is then oxidized by cyclohexenone, producing cyclohexanone.

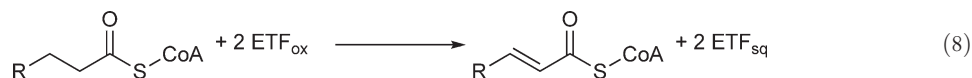
OYE reduces conjugated nitroolefins by hydride transfer to the  $\beta$ -carbon, followed by protonation of the  $\alpha$ -carbon of the nitronate by Tyr196.<sup>203</sup> The negatively charged nitronate oxygens are well positioned to form hydrogen bonds with the active site histidine and asparagine. The Tyr196Phe mutant enzyme reduces nitroolefins as rapidly as wild-type OYE, but releases the kinetically stable *aci*-nitro tautomer of the product. Wild-type OYE actually catalyzes the protonation of the  $\alpha$ -carbon of preformed nitronates introduced into solutions of low pH, or the deprotonation of nitroalkanes at pH values above their  $pK_a$ . Tyr196 and His191 are critical residues in this nonredox reaction.

OYE and many homologues occur widely in nature. The physiological function of most is unknown, but they are often isolated from bacteria that metabolize pollutants containing nitro compounds such as explosives.<sup>204,205</sup> Glycerol trinitrate<sup>206</sup> and trinitrotoluene<sup>207</sup> are substrates. Replacement of the natural FMN with 5-carba-5-deaza FMN slowed the reduction of glycerol trinitrate by 4–5 orders of magnitude, suggesting a mechanism involving radicals. Reduction of trinitrotoluene forms a Meisenheimer complex as a first step, consistent with hydride transfer to the aromatic ring.

### 7.03.3.2 Acyl-CoA Dehydrogenases and Oxidases

Acyl-CoA dehydrogenases (ACADs), flavoproteins found in the mitochondria, are involved in  $\beta$ -oxidation of fatty acids. Currently, there are five types of ACADs classified according to substrate specificity. Short-chain ACAD (SCAD), medium-chain ACAD (MCAD), long-chain ACAD (LCAD), and very long-chain ACAD (VLCAD) are most active with 4-, 8-, 14-, and 16-carbon substrates, respectively.<sup>208–210</sup> More recently, acyl-CoA dehydrogenase 9 (ACAD-9) was discovered.<sup>211</sup> This enzyme is most active with unsaturated long-chain acyl-CoAs.<sup>212</sup> Substrates whose acyl chain length is too long for a particular ACAD are excluded sterically by the length of the hydrophobic tunnel the alkyl group must occupy, while short-chain substrates bind weakly to ACADs specific for longer substrates because these enzymes do not interact extensively with the CoA portion of their cognate substrates; the short acyl groups do not extend deeply enough into the hydrophobic tunnel to provide significant binding energy. Mechanistically similar enzymes involved in amino acid metabolism are also members of the ACAD family, including isobutyryl-CoA dehydrogenase, isovaleryl-CoA dehydrogenase, and glutaryl-CoA dehydrogenase. Genetic deficiencies in members of the ACAD family are responsible for a number of medical disorders.<sup>213</sup>

ACADs catalyze the oxidation of acyl-CoA thioesters, forming the corresponding enoyl-CoA ester and reduced flavin<sup>214</sup> (Equation (8)).

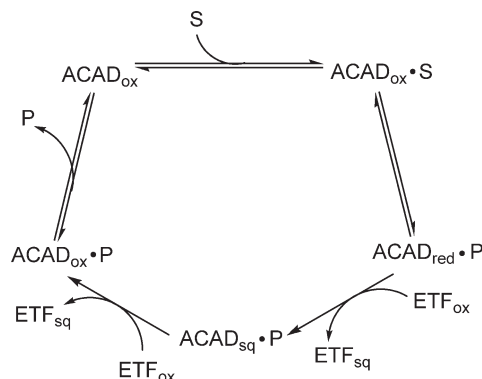


The electrons on reduced flavin are transferred to another flavoprotein, electron-transferring flavoprotein (ETF), which then transfers the reducing equivalents to membrane-bound ETF dehydrogenase, which reduces ubiquinone, a substrate of the respiratory chain.<sup>215</sup> Most ACADs are dimers of dimers, with 45 kDa subunits each containing an FAD.<sup>210</sup> However, VLCAD and ACAD-9 are homodimers of around 70 kDa subunits instead of tetramers. In addition to an N-terminus that shares sequence and structural homology to all ACADs, VLCAD and ACAD-9 contain an additional ~180 residues at the C-terminus,<sup>211,216</sup> which show ~14% sequence identity to residues in the N-terminus, suggesting that the residues probably originated from a partial gene duplication.

Acyl-CoA oxidases (ACOs) are very similar to ACADs. ACOs are found in the peroxisomes. They contain one FAD per subunit and are members of the same superfamily as ACADs.<sup>217</sup> The reductive half-reaction (oxidation of the acyl-CoA thioester) is very similar to that of ACADs. However, reduced ACOs react with molecular oxygen rapidly to produce H<sub>2</sub>O<sub>2</sub> and regenerate the oxidized flavin, rather than reacting with ETF. ACOs are homodimers and possess an additional C-terminal domain not present in the tetrameric ACADs.<sup>218</sup> In addition, the FAD in ACOs is more solvent exposed than in ACADs, presumably enhancing oxygen reactivity.

Ligand binding by ACADs is often accompanied by the appearance of intense charge-transfer absorbance. In fact, these enzymes are often purified as green complexes due to tightly bound CoA-persulfide that serves as a charge-transfer donor to the oxidized flavin.<sup>219</sup> Many ligands that have a negative charge or electron-rich functionality at the  $\alpha/\beta$  position will form charge-transfer bands with oxidized ACADs. The crystal structure of MCAD complexed with 3-thiooctanoyl-CoA shows the ligand stacked above the flavin in the proper orientation for charge-transfer absorbance.<sup>220</sup> In addition, electron-poor ligands, such as enoyl-CoA, have charge-transfer absorbance with the electron-rich reduced enzyme making the reduced enzyme-product complex blue-green.<sup>221,222</sup> Usually, acyl-CoA substrates bind quite tightly to all redox states of the enzyme.<sup>223–228</sup> Rapid reaction studies have shown that binding is a two-step process,<sup>229–232</sup> in which an initial enzyme-ligand complex forms, followed by its isomerization to a charge-transfer complex. The isomerization of the initial complex might represent the dehydration of the acyl-binding crevice, which is occupied by water molecules in the free enzyme.

Free ACADs stabilize the neutral flavin semiquinone when the enzyme is artificially reduced by single-electron donors; however, binding of the product enoyl-CoA esters shifts the pK<sub>a</sub> of the semiquinone to 7.3 to generate the anionic form.<sup>226</sup> The two-electron redox potential of free acyl-CoA substrates (−40 mV)<sup>233</sup> is not low enough to produce free reduced enzyme (−145 mV for MCAD).<sup>227</sup> However, ligands bind much more tightly to the reduced enzyme, shifting the redox equilibrium to promote flavin reduction. For example, the dissociation constant of oct-2-enoyl-CoA from oxidized enzyme is 200 nmol l<sup>−1</sup>, but with reduced enzyme it is 13 pmol l<sup>−1</sup>.<sup>229</sup> Therefore, it is thermodynamically unfavorable for the reduced enzyme to release product. Instead, turnover occurs by sequential one-electron oxidations by ETF while the product enoyl-CoA ester is bound<sup>234</sup> (Scheme 11).



**Scheme 11** Catalytic cycle of acyl-CoA dehydrogenases.

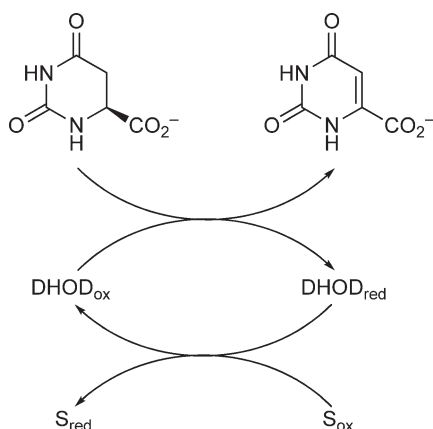
The chemical mechanism of  $\alpha,\beta$ -dehydrogenation has been thoroughly studied.<sup>235</sup> The 2-*pro-R* and 3-*pro-R* hydrogens are oxidized in the substrate,<sup>236,237</sup> leading to a *trans* enoyl thioester. Deuterium isotope effects on flavin reduction for butyryl-CoA labeled at the  $\alpha$ - or  $\beta$ -position or both were used to probe the order of the two C–H bond cleavages. The isotope effect observed with the doubly labeled substrate is roughly equal to the product of the isotope effects observed with the singly labeled substrates, indicating that both C–H bonds break in a single transition state.<sup>238,239</sup> Therefore, the mechanism involves proton abstraction from the substrate  $\alpha$ -carbon via an active site base concerted with hydride transfer from the substrate  $\beta$ -carbon to N5 of the flavin. Crystal structures of MCAD complexed with various ligands show that the  $\beta$ -carbon of the substrate lies over N5 of the flavin and the  $\alpha$ -carbon is in contact with the carboxylate of Glu376, the active site base.<sup>220,229,240</sup> The carbonyl forms hydrogen bonds with the backbone amide of Glu376 and the 2'-hydroxyl of the flavin ribityl chain. These hydrogen bonds stabilize the partial negative charge that develops on the carbonyl oxygen during the reaction, similar to the oxyanion hole in serine proteases. Replacing the natural flavin with 2'-deoxy FAD reduces activity by one to ten million-fold, demonstrating the importance of stabilizing the charge buildup on the carbonyl oxygen.<sup>6,241</sup> This charge delocalization lowers the  $pK_a$  of the  $\alpha$ -carbon by  $\sim 11$  units in the natural enzyme, and is accompanied by an increase of the  $pK_a$  of the active site glutamate to  $\sim 8$ – $9$ , promoting the necessary substrate deprotonation.

### 7.03.3.3 Dihydroorotate Dehydrogenase, Dihydropyrimidine Dehydrogenase, and Dihydrouridine Synthase

Dihydroorotate dehydrogenase (DHOD), dihydropyrimidine dehydrogenase (DPD), and dihydrouridine synthase (DUS), flavoenzymes involved in various stages of pyrimidine metabolism, catalyze very similar C–C oxidation/reduction reactions. These enzymes share common active site residues and structural homology. They have all been implicated as potential drug targets due to their important physiological roles.

#### 7.03.3.3.1 Dihydroorotate dehydrogenase

DHODs are FMN-containing enzymes that convert dihydroorotate (DHO) to orotate (OA) in the only redox step in the *de novo* synthesis of pyrimidines (Scheme 12). DHODs have been grouped into two classes based on sequence.<sup>242</sup> Class 2 DHODs are membrane-bound monomers that are reoxidized by ubiquinone, coupling pyrimidine biosynthesis to the respiratory chain.<sup>243</sup> On the other hand, Class 1 DHODs are cytosolic proteins that have been further divided into two subclasses. Class 1A DHODs are homodimers that are reoxidized by fumarate.<sup>244</sup> Class 1B DHODs are  $\alpha_2\beta_2$  heterotetramers with an FMN-containing subunit very similar to Class 1A enzymes and a second subunit that contains an iron–sulfur cluster and FAD, allowing Class 1B DHODs to be reoxidized by NAD.<sup>245</sup>



**Scheme 12** Catalytic cycle of dihydroorotate dehydrogenase.

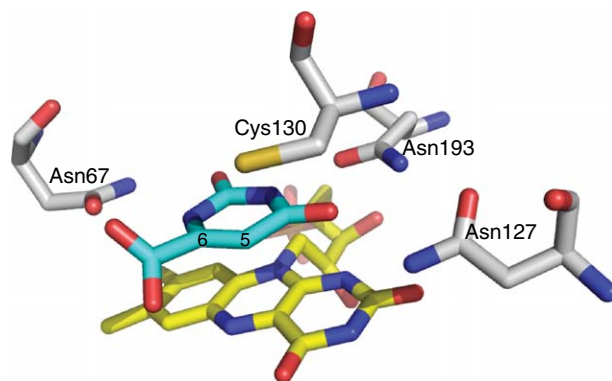


Both classes of DHODs are interesting drug targets. Currently, several Class 2 DHODs are being investigated as drug targets including the enzymes from humans,<sup>246</sup> *Candida albicans* (a pathogenic yeast species),<sup>247</sup> and *Plasmodium falciparum* (the malaria parasite).<sup>248–250</sup> Human DHOD is targeted for the treatment of arthritis by the drug leflunomide, whose active metabolite is a potent inhibitor.<sup>251</sup> Crystal structures have shown that Class 2 DHODs have two binding sites on opposite sides of the flavin, one for DHO oxidation and the other for quinone binding and electron transfer.<sup>248,252–254</sup> Inhibitors of Class 2 DHODs are directed at the so-called species-selective inhibitor site, which is near and possibly overlapping the quinone-binding site.<sup>250</sup> The sequence variability of this site has allowed species-specific inhibitors to be created. In Class 1A DHODs, both DHO and fumarate (the oxidizing substrate) bind to the same site,<sup>255,256</sup> whose structure is very similar to the pyrimidine-binding site of Class 2 enzymes. Nonetheless, small molecules that bind specifically to Class 1A DHODs have been discovered.<sup>257,258</sup> The underlying reason for the specificity of these compounds is not understood, especially since the pyrimidine-binding sites in both classes of enzymes appear to be nearly identical. However, these compounds demonstrate the possibility of creating inhibitors specific for Class 1A enzymes. Genetic studies have shown that DHOD is essential for the survival of *Trypanosoma cruzi*,<sup>259</sup> validating the idea that inhibitors specific for Class 1A DHODs could treat trypanosomal diseases such as Chagas' disease and African sleeping sickness.

In the oxidation of DHO, the C5 *pro-S* hydrogen<sup>260</sup> is removed as a proton by an active site base, while the C6 hydrogen is transferred to N5 of the isoalloxazine ring of the flavin as a hydride. Structures of product complexes of all DHODs show that C6 of OA is in van der Waals contact with N5 of the flavin and that C5 of OA is positioned correctly for proton abstraction by the active site base (Figure 4), either serine in Class 2 enzymes or cysteine in Class 1 enzymes.<sup>252,254,255,261,262</sup>

The two C–H bonds of DHO could break at the same time in a concerted mechanism or they could break sequentially in stepwise mechanisms. This mechanistic question was addressed by determining kinetic isotope effects using DHO labeled at the 5- or 6-position or both. Isotope effects determined in the steady state for the Class 2 DHOD from bovine liver<sup>263</sup> and the Class 1B enzyme from *Clostridium oroticum*<sup>264</sup> were consistent with a concerted mechanism, while isotope effects for the Class 1A enzyme from *Critibidia fasciculata*<sup>260</sup> were consistent with a stepwise mechanism. Similar experiments to probe the order of bond breaking were conducted by directly observing the isotope effect on flavin reduction in anaerobic stopped-flow experiments. A concerted mechanism was found for the Class 1A enzyme from *Lactococcus lactis*,<sup>265</sup> while the isotope effects observed with the Class 2 enzymes from humans and *E. coli* are consistent with a stepwise mechanism or a concerted mechanism with tunneling.<sup>266</sup> The conclusions from the reductive half-reaction contradict those from steady-state kinetics, presumably due to the complexities of interpreting steady-state measurements. Alternatively, the mechanism of DHO oxidation could be species specific.

The pH dependence of kinetics of DHODs has been studied both by steady-state and transient methods. By watching flavin reduction directly in anaerobic stopped-flow experiments, a  $pK_a$  of  $\sim 8.3$  controlling reduction was observed for the Class 1A DHOD from *L. lactis*,<sup>265</sup> while in the Class 2 enzymes from humans and *E. coli*,<sup>266,267</sup> the  $pK_a$  observed was  $\sim 9$ . The  $pK_a$  observed for *L. lactis* DHOD was attributed to ionization of the

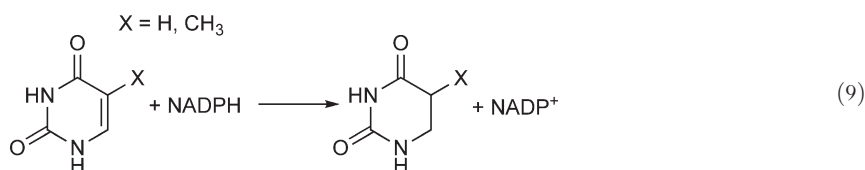


**Figure 4** Active site of dihydroorotate dehydrogenase from *Lactococcus lactis* with orotate (cyan).

active site base, Cys130. The  $pK_a$  observed in the Class 2 DHODs was a bit of a mystery. The structures of Class 2 DHODs show nothing in the active site that would lower the  $pK_a$  of the serine so dramatically<sup>252,254</sup> from  $\sim 13.6$  for an unperturbed residue.<sup>268</sup> Kinetic studies suggest that this is a consequence of the stepwise mechanism of DHO oxidation, rather than a true thermodynamic  $pK_a$ . The deuterium isotope effect for label at the 5-position of DHO is lost when experiments are performed at pH values above the observed  $pK_a$ . Above pH  $\sim 9$ , deprotonation becomes too fast to contribute to the rate of reduction, and the rate is determined only by hydride transfer. Conversely, the pH dependence of steady-state turnover rates by the Class 2 enzyme from bovine liver,<sup>263</sup> the Class 1A enzyme from yeast,<sup>269</sup> and the Class 1B enzyme from *C. oroticum*<sup>264</sup> all show a  $pK_a$  value  $\sim 7$ , which was ascribed to the active site base needed to deprotonate C5 of DHO. The kinetic complexity of steady-state measurements could account for the observed differences in pH dependencies.

### 7.03.3.3.2 Dihydropyrimidine dehydrogenase

DPD catalyzes the NADPH-dependent reduction of uracil and thymine to the corresponding dihydropyrimidines (Equation (9)) in the first step in the breakdown of pyrimidines.



DPD has become an important drug target due to its role in the degradation of 5-fluorouracil (5FU). 5FU is a commonly prescribed chemotherapeutic used for the treatment of colorectal, breast, and head/neck cancer. It blocks DNA synthesis by inhibiting thymidylate synthase.<sup>270</sup> However, much of the administered dose of 5FU is rapidly degraded by DPD,<sup>271</sup> decreasing the efficiency of the drug. In addition, the fluorinated degradation products cause multiple side effects in the nervous system.<sup>272</sup> Therefore, DPD has emerged as an interesting drug target to increase the therapeutic potential of 5FU.

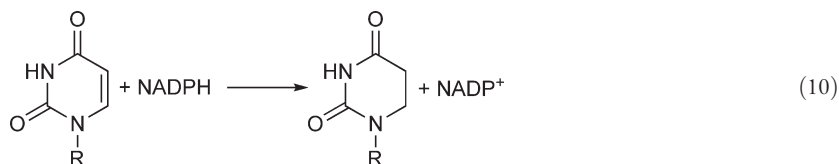
Porcine DPD is a homodimer of  $\sim 110$  kDa subunits.<sup>273</sup> Each monomer contains one FMN, one FAD, and four [4Fe-4S] clusters.<sup>274–276</sup> The crystal structure shows that each monomer contains five domains: an N-terminal Fe-S cluster domain, an FAD-binding domain, an NADPH-binding domain, an FMN/pyrimidine-binding domain, and a C-terminal Fe-S domain.<sup>277</sup> The FMN/pyrimidine-binding domain is a  $\alpha_8\beta_8$  barrel resembling the Class 1A DHOD from *L. lactis*, with which it shares 61% sequence similarity. Three of the four Fe-S clusters are coordinated by cysteine thiols; the second Fe-S cluster in the N-terminal domain is coordinated by three cysteine thiols and the side chain of Glu156. The two flavins are located on opposite ends of the protein and are connected via the C-terminal Fe-S domain from the same monomer and the N-terminal Fe-S domain from the other monomer, indicating that DPD is active only as a dimer.

Steady-state analysis of DPD<sup>278,279</sup> in addition to crystal structures<sup>277,280</sup> has led to the proposal of a two-site ping-pong mechanism, in which NADPH binds and reduces FAD at one end of the protein. The reducing equivalents are passed to the FMN at the other end of the protein via the Fe-S clusters. The reduced FMN reacts with the pyrimidine forming dihydropyrimidine and oxidized FMN. Here, we will focus on the pyrimidine reaction.

The reduction of the pyrimidine to dihydropyrimidine is the reverse of the oxidation reaction carried out by DHODs. The structure of the FMN/pyrimidine-binding site is very similar to the structure of *L. lactis* DHODs. Three Asn residues form hydrogen bonds with the nitrogens and carbonyls of the pyrimidine analogous to DHODs. DPD has an active site cysteine proposed to act in acid/base chemistry similar to Class 1 DHODs.<sup>277,280,281</sup> When mutated to alanine, only 1% of the wild-type activity was retained,<sup>278</sup> indicating the importance of this residue in catalysis. Secondary tritium isotope effects using 5-<sup>3</sup>H-uracil were determined in both H<sub>2</sub>O and D<sub>2</sub>O; an inverse isotope effect was observed in H<sub>2</sub>O and the value became more inverse in D<sub>2</sub>O.<sup>274</sup> This was taken as evidence of a stepwise mechanism in which hydride transfer to C6 is followed by protonation at C5.

### 7.03.3.3.3 Dihydrouridine synthase

Dihydrouridine is an abundant modified nucleoside found in tRNA<sup>282</sup> predominantly on the D-loop. DUSs, flavin-containing proteins that catalyze the conversion of uridine to dihydrouridine on tRNA, have been identified<sup>283,284</sup> (Equation (10)).

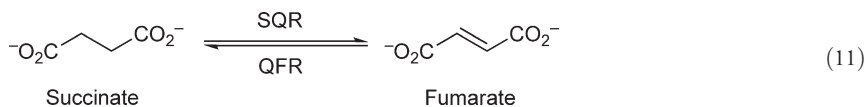


The physiological role of dihydrouridine, and, therefore, DUS, remains unknown, but it occurs throughout biology. Human DUS2 is upregulated in pulmonary carcinogenesis<sup>285</sup> and inhibits the kinase activity of PKR, an interferon-induced protein kinase involved in the regulation of antiviral innate immunity.<sup>286</sup>

Little mechanistic work has been published on these enzymes to date. The four DUS enzymes in *S. cerevisiae* each show distinct site specificity on tRNA, with each enzyme reducing either one or two specific uridines in tRNA.<sup>287</sup> An *in vivo* DUS-complementation assay has been developed to screen for residues important in catalysis in which an *E. coli* strain that has all the DUS genes knocked out is complemented with a plasmid-borne DUS and the dihydrouridine content of the tRNA is then determined.<sup>288</sup> This approach was used to identify a cysteine residue that is essential for DUS activity. Mapping this cysteine onto the crystal structure of a DUS from *Thermatoga maritima*<sup>289</sup> shows that it is positioned in the active site to act as a catalytic acid. Presumably, the mechanism of uridine reduction involves hydride transfer from N5 of the flavin to C6 of uridine and protonation of C5 of uridine by an active site acid. The actual substrate of DUSs has not been identified. The enzymes reduce *in vitro*-transcribed tRNAs several orders of magnitude slower than naturally modified tRNA.<sup>290</sup> The critical modification for DUS activity has not yet been identified.

### 7.03.3.4 Succinate Dehydrogenase and Fumarate Reductase

Succinate:quinone oxidoreductases (SQORs) refer collectively to succinate:quinone reductases (SQRs) and quinol:fumarate reductases (QFRs) (Equation (11)).



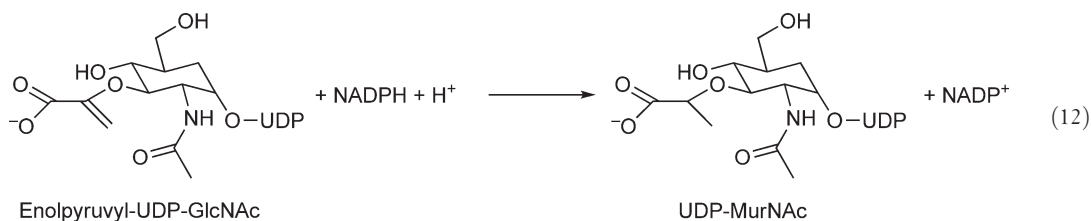
SQRs, also referred to as succinate dehydrogenases and Complex II, catalyze the oxidation of succinate to fumarate. SQRs are found in aerobic organisms and link the Krebs cycle and aerobic respiration.<sup>291</sup> QFRs catalyze the reverse reaction, the reduction of fumarate to succinate. QFRs are found in organisms capable of anaerobic respiration using fumarate as the terminal electron acceptor.<sup>292,293</sup> SQORs are generally multisubunit complexes containing two soluble domains and either one or two transmembrane domains anchoring the complex to the membrane. The soluble domains, a flavin-binding subunit and an Fe-S cluster subunit, of SQRs and QFRs share sequence and structural homology, while much variation is seen in the membrane-anchoring domains. Over the past decades, many in-depth reviews on SQORs structure and mechanism have been published.<sup>294–298</sup>

More recently, a soluble homologue of QFR has been identified<sup>299,300</sup> in *Shewanella*. This enzyme is a soluble, periplasmic, tetraheme flavocytochrome *c*<sub>3</sub> (Fcc<sub>3</sub>). The flavin-binding domain of Fcc<sub>3</sub> shares high sequence homology with the flavoprotein subunit of membrane-bound SQORs.<sup>300</sup> Several crystal structures of Fcc<sub>3</sub><sup>301–305</sup> have led to a proposal in which fumarate binds to the reduced enzyme; the C1 carboxylate twists out of the plane of the molecule and is held by steric constraints from two methionine residues and a hydrogen bond to a histidine. The C4 carboxylate on the other end of the molecule is bound in a very positively charged environment and is hydrogen bonded to a histidine and an arginine. This environment causes polarization of the C2–C3 bond, with positive charge buildup at C2 of fumarate, facilitating hydride transfer from N5 of the reduced FAD to the *si* face of the substrate. The structures show Arg402 positioned close to C3 of fumarate

where it could act as an active site acid, and this was verified by mutagenesis.<sup>304,305</sup> In addition, Arg402 is a member of a proton-transfer pathway completely conserved throughout SQORs that is responsible for reprotonation of the active site acid following fumarate reduction.<sup>306</sup> Given the sequence and structural homology of membrane-bound QFRs with Fcc<sub>3</sub>, the mechanism described here is likely a general mechanism of fumarate reduction by these enzymes.

### 7.03.3.5 UDP-*N*-acetylenolpyruvylglucosamine reductase

An enzyme instrumental in bacterial cell wall biosynthesis is UDP-*N*-acetylenolpyruvylglucosamine reductase (MurB). MurB catalyzes the NADPH-dependent reduction of enolpyruvyl-UDP-*N*-acetylglucosamine (enolpyruvyl-UDP-GlcNAc) to the corresponding *D*-lactyl compound UDP-*N*-acetylmuramic acid (UDP-MurNAc) (Equation (12)).

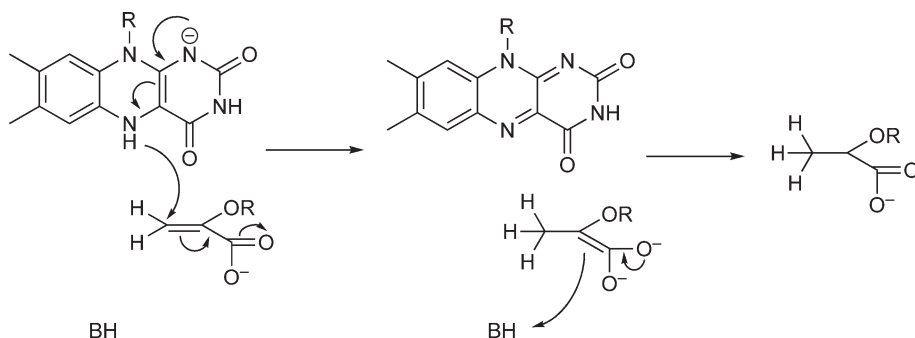


The monomeric protein is 38 kDa and contains one FAD.<sup>307</sup> With the increased prevalence of antibiotic-resistant bacteria, MurB is a possible drug target due to its critical role in cell wall biosynthesis.<sup>308</sup> MurB is an essential gene in both *E. coli* and *Bacillus subtilis*,<sup>309,310</sup> further validating it as a suitable drug target.

The enzyme from *E. coli* is the most thoroughly studied; however, recently, work has begun on enzymes from *Streptococcus pneumoniae*<sup>311</sup> and *Staphylococcus aureus*.<sup>312</sup> The catalytic mechanism of MurB is ping-pong with a turnover number of 62 s<sup>-1</sup> at pH 8.0 and 37 °C for the enzyme from *E. coli*.<sup>307,313</sup> NADPH binds to the oxidized enzyme and reduces the flavin with its *pro-S* hydrogen.<sup>307</sup> NADP<sup>+</sup> then dissociates and enolpyruvyl-UDP-GlcNAc binds to the reduced enzyme and is reduced to UDP-MurNAc. Binding of each substrate appears to be mutually exclusive and is possible in both redox states of the enzyme leading to substrate inhibition at high substrate concentrations.<sup>313</sup>

The oxidative half-reaction (reduction of the UDP-sugar substrate) occurs in two steps (Scheme 13).<sup>307,314</sup> First, a hydride is transferred from N5 of the flavin to C3 of the enolpyruvyl moiety of the UDP-sugar substrate forming an enediol intermediate. Monovalent cations activate the enzyme during turnover by coordinating with the negatively charged enolate oxygen. The intermediate is then protonated at C2, forming the product UDP-MurNAc. The hydrogen at C3 of the product is derived from NADPH, while the hydrogen at C2 is derived from the solvent.<sup>307</sup> Exchange at N5 of the reduced flavin is much slower than binding and reduction of enolpyruvyl-UDP-GlcNAc, but the active site acid responsible for protonating C2 rapidly exchanges with bulk solvent.

MurB can also reduce enolbutyryl-UDP-GlcNAc to UDP-methyl-*N*-acetylmuramic acid.<sup>314</sup> The crystal structures of MurB complexed with enolpyruvyl-UDP-GlcNAc and (*E*)-enolbutyryl-UDP-GlcNAc show that all



**Scheme 13** Oxidative half-reaction of MurB.

residues thought to be involved in hydride transfer and protonation are arranged *anti* relative to the enol bond.<sup>314,315</sup> The stereochemical outcome of the reaction was probed using (*E*)-enolbutyryl-UDP-GlcNAc as a substrate with NADPD in D<sub>2</sub>O. The (2*R*,3*R*)-dideutero product was formed. In addition to catalyzing the reduction reaction, MurB also catalyzes the isomerization of the (*E*)-enolbutyryl-UDP-GlcNAc into the (*Z*)-isomer. This was taken as evidence of a C2 carbanion/enol species that is present long enough to rotate around the C2–C3 single bond.

Several crystal structures of the enzyme with different carbohydrate substrates bound have been solved.<sup>314,315</sup> The substrate vinylic moiety is parallel to the plane of the isoalloxazine, with C3 of the substrate lying under N5 of the FAD. Glu325 and Arg159 are in close proximity to the carboxylate of the substrate and could perhaps stabilize the enolic intermediate. Ser229 is positioned in the plane below the C2–C3 bond and was proposed to be the acid that protonates C2. This was validated by mutagenesis; the Ser229Ala mutant shows a 10<sup>7</sup>-fold decrease in reduction of enolpyruvyl-UDP-GlcNAc when compared to wild type.<sup>316</sup>

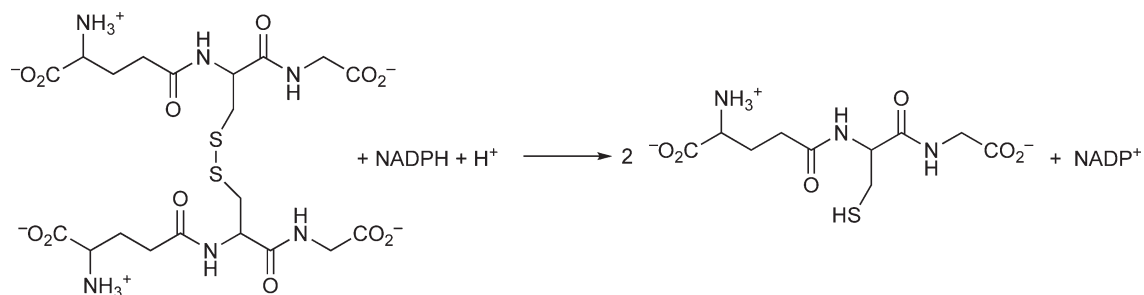
### 7.03.4 Thiol/Disulfide Chemistry

The reduction of disulfides by pyridine nucleotides is catalyzed by a family of flavoproteins that are well studied both mechanistically and structurally. These enzymes generally contain an FAD and a pair of cysteines capable of forming a disulfide. They are usually dimers, with each active site composed of residues contributed by each monomer. Members of this family all use the same basic catalytic mechanism, which can be divided into two half-reactions. In the reductive half-reaction, the oxidized flavin is reduced by the *pro-S* hydride of NAD(P)H; however, the reduced flavin is rarely observed experimentally. Instead, its enediamine attacks the enzyme disulfide, forming a flavin C4a-cysteine adduct and expelling a thiol (**Scheme 14(a)**). Thiol formation is assisted by an active site acid. Although the flavin–cysteine adduct is not usually observed experimentally, special circumstances such as alkylation<sup>317</sup> or removal by mutagenesis<sup>318</sup> of one of the thiols permit its observation. Subsequent elimination of thiolate from the flavin adduct rapidly forms oxidized flavin. Therefore, reduction by NAD(P)H leaves the flavin oxidized and the enzymatic disulfide reduced to a thiol–thiolate pair, yielding the two-electron reduced state of the enzyme (EH<sub>2</sub>). The thiolate near C4a forms a charge-transfer complex with the oxidized flavin, giving a characteristic absorbance from 500 to 600 nm. In most enzymes, EH<sub>2</sub> can be reduced again by NAD(P)H because the flavin is oxidized; however, the resulting four-electron reduced state (EH<sub>4</sub>) is not involved in steady-state turnover for most enzymes.

The oxidative half-reaction involves a series of thiol–disulfide interchanges between the thiols produced in the reductive half-reaction and the oxidizing substrate (**Scheme 14(b)**). First, the interchange thiol (the thiol not involved in charge-transfer absorbance with the flavin) attacks the disulfide of the substrate, expelling a leaving group, which is protonated by an active site acid. The enzyme–substrate mixed disulfide is then attacked by the charge-transfer thiolate, reforming the enzyme disulfide.

#### 7.03.4.1 Glutathione Reductase

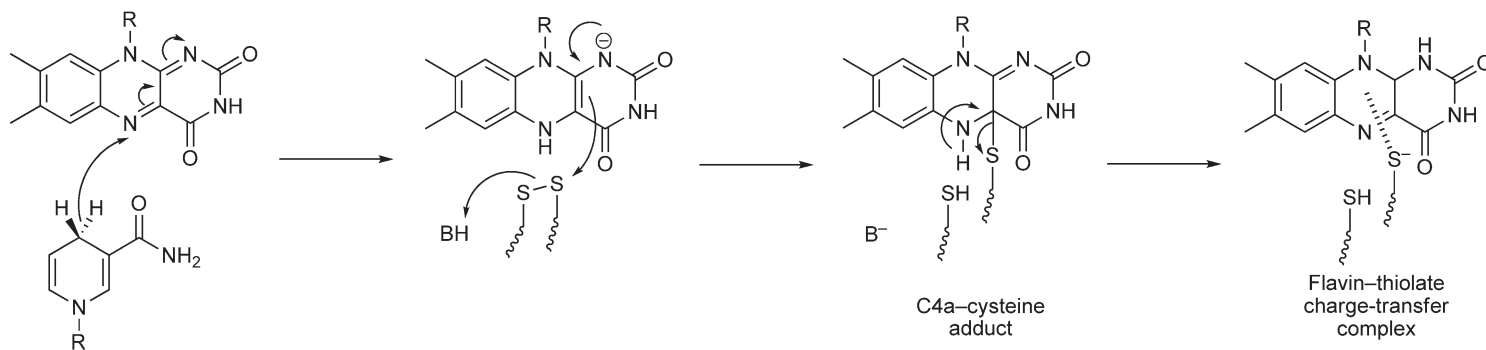
Glutathione reductase (GR) is a homodimer containing one FAD per monomer that catalyzes the reduction of glutathione disulfide (GSSG) to glutathione (GSH) (Equation (13)).



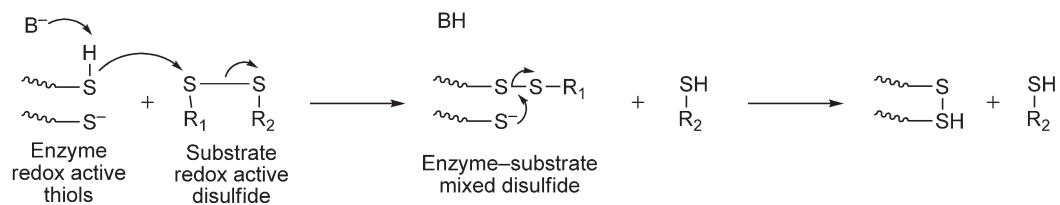
(13)

## Thiol–disulfide chemistry

(a) Reductive half-reaction



(b) Oxidative half-reaction



**Scheme 14** General mechanism used by flavoprotein disulfide oxidoreductases.



GSSG and GSH play key roles in cellular redox homeostasis. GSH helps detoxify reactive oxygen species by donating reducing equivalents to glutathione peroxidase and detoxifies electrophilic xenobiotics with glutathione *S*-transferase.<sup>319</sup> GR is a widely occurring enzyme and has been studied from several sources including *Plasmodium falciparum*,<sup>320</sup> *Anabaena* PCC 7120 (a cyanobacterium),<sup>321</sup> *Enterococcus faecalis*,<sup>322</sup> and most thoroughly from human erythrocytes<sup>323</sup> and *E. coli*.<sup>324</sup> The enzyme from *P. falciparum* is being investigated as a drug target for the treatment of malaria<sup>320,325–327</sup> due to its vital role in protecting the parasite from reactive oxygen species.

Crystals structures<sup>328–331</sup> show the enzymatic cysteines near the *si* face of the flavin. In the human erythrocyte enzyme, Cys63 is close to the flavin C4a and is responsible for the flavin–thiolate charge-transfer absorbance. Cys58, the other member of the active site disulfide, is found at the base of the GSH-binding domain and is solvent accessible. Electron flow is from NADPH to the *re* face of the flavin to the disulfides at the *si* face of the flavin to the disulfide in GSH.<sup>332</sup>

In the reductive half-reaction, NADPH binds to the oxidized enzyme, forming charge-transfer absorbance within the dead-time of stopped-flow experiments.<sup>333</sup> The first kinetic phase observed is composed of three chemical events: hydride transfer to N5 of flavin, nucleophilic attack of the reduced flavin on the enzymatic disulfide forming a flavin C4a–cysteine adduct, and elimination of thiolate from the adduct forming the oxidized flavin–thiolate charge-transfer complex (see **Scheme 14(a)**). These steps are not kinetically resolved, so that the oxidized flavin–NADPH charge-transfer complex appears to convert directly to the oxidized flavin–thiolate charge-transfer complex. The chemical steps have been resolved in a mutant *E. coli* enzyme in which the active site histidine is mutated to an alanine.<sup>334</sup> This mutant rapidly forms an NADPH–oxidized enzyme charge-transfer complex similar to the wild-type enzyme, followed by formation of a reduced flavin–NADP<sup>+</sup> charge-transfer complex, but the oxidation of the flavin by the enzyme disulfide is much slower, allowing the formation of reduced flavin to be observed. These experiments also indicated that the active site histidine is responsible for protonating the leaving interchange thiol, a common feature in this class of flavoenzymes.

The reduction of GSSG in the oxidative half-reaction is accomplished by binding, attack of the interchange thiol, formation of an enzyme–glutathione mixed disulfide, and expulsion of one molecule of GSH (see **Scheme 14(b)**). The active site histidine on the *si* face of the flavin protonates the leaving GSH, recapitulating its role in the reductive half-reaction. The mixed disulfide is then attacked by the charge-transfer thiol, reforming the enzymatic disulfide and releasing the second molecule of GSH.<sup>334</sup> Evidence for the mixed disulfide comes from crystallography,<sup>332,335</sup> chemical modification,<sup>336</sup> spectroscopy,<sup>337</sup> and kinetics.<sup>320</sup>

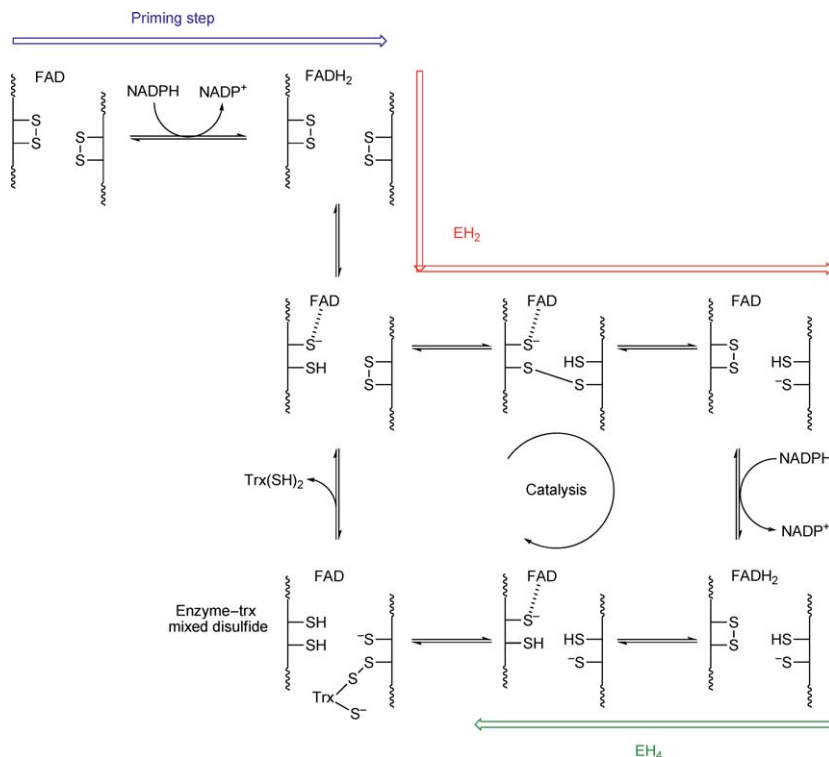
### 7.03.4.2 Thioredoxin Reductase

Thioredoxin is a small (12 kDa) protein containing a dithiol that is involved in many cellular processes including antioxidant defense, DNA synthesis and repair, redox signaling, and apoptosis.<sup>338,339</sup> Thioredoxin reductase (TrxR) reduces the disulfide of thioredoxin to a dithiol in an NADPH-dependent reaction. TrxR is essential for the survival of *P. falciparum*,<sup>340,341</sup> the parasite that causes malaria, and therefore, inhibitors specific for the enzyme are being developed as potential malaria treatments.<sup>342</sup> The human enzyme is being investigated as a potential chemotherapeutic drug target<sup>343</sup> since TrxR is elevated in tumor cells due to both the need for higher amounts of deoxyribonucleotides and for detoxification of reactive oxygen species.<sup>344,345</sup>

TrxR is a dimer with one molecule of FAD per monomer. There are two types of TrxRs, low molecular weight TrxR (low  $M_r$  TrxR) with ~35 kDa subunits and high molecular weight TrxR (high  $M_r$  TrxR) with ~55 kDa subunits.<sup>344</sup> Both types of TrxRs use the same basic chemistry as all members of the FAD–disulfide oxidoreductase family; however, the details of the reaction are quite different.

The prototype of low  $M_r$  TrxRs is the enzyme from *E. coli* (see Williams<sup>346</sup> for review). The mechanism of this enzyme involves large conformational changes of the protein. In one conformation, the redox active enzymatic disulfide is near FAD where it can be attacked by the reduced flavin, forming the oxidized flavin–thiolate charge transfer. The two protein domains then rotate ~66° into a conformation moving the dithiol pair to the hydrophilic surface of the enzyme where it can reduce the disulfide of thioredoxin. In this conformation, NADPH can also bind and reduce the flavin.<sup>347</sup>

The high  $M_r$  TrxRs use a different strategy to accomplish the same chemistry. The enzymes from human and *Drosophila melanogaster* are the best-studied examples. In addition to the usual enzymatic disulfide near the flavin, these enzymes have a second redox active disulfide or a selenosulfide pair at their C-terminus.<sup>344</sup> In the



**Scheme 15** Kinetic mechanism of high molecular weight thioredoxin reductases.

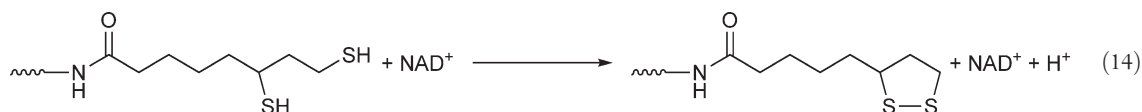
enzyme from humans, the C-terminal redox disulfide is formed by cysteine 497 and selenocysteine 498, while in the enzyme from *D. melanogaster*, it is formed by cysteine 489 and cysteine 490. In either case, the C-terminal redox active disulfide/selenosulfide is required for reduction of thioredoxin.<sup>348,349</sup> This extra disulfide or selenosulfide shuttles reducing equivalents away from the flavin in a bucket brigade fashion, thus obviating the need for large conformational changes to allow the reaction with thioredoxin.

Kinetic investigation of the *D. melanogaster* enzyme showed that reduction of the enzyme-bound flavin to form  $\text{EH}_2$  (Scheme 15)<sup>348</sup> is a priming step, but the enzyme cycles between  $\text{EH}_2$  and  $\text{EH}_4$  during turnover. The initial chemistry is identical to the reaction of GR. However, in TrxR, the interchange thiol of the thiolate-oxidized flavin complex attacks the C-terminal disulfide forming an N-terminal/C-terminal mixed disulfide. This complex still has thiolate-oxidized flavin charge-transfer absorbance. Finally, the charge-transfer thiolate attacks the mixed disulfide, regenerating the N-terminal disulfide. Another molecule of NADPH reduces the enzyme-bound flavin, beginning a very similar cascade of thiol chemistry resulting in an enzyme with two thiol/thiolate pairs, one at the C-terminus and one at the N-terminus. The thiolate of the C-terminal thiol/thiolate pair attacks the disulfide in thioredoxin, forming a C-terminal/thioredoxin mixed disulfide. The mixed disulfide is then attacked by the other C-terminal thiol, releasing reduced thioredoxin and reforming the C-terminal enzymatic disulfide. The enzyme is in the  $\text{EH}_2$  state after release of thioredoxin. Consequently, the first turnover of the purified enzyme requires two molecules of NADPH, while subsequent turnovers require only one molecule. An active site histidine aids in the dithiol/disulfide chemistry during catalysis, whose importance in both the reductive and oxidative half-reactions has been demonstrated by mutagenesis.<sup>350</sup>

### 7.03.4.3 Lipoamide Dehydrogenase

Lipoamide dehydrogenase is a dimeric enzyme with one FAD per monomer, found as part of ketoacid dehydrogenase multienzyme complexes; however, it can be resolved from the other components in pure,

active form.<sup>351</sup> When the enzyme is resolved from the complex, it can use (*R*)-lipoamide as a substrate in place of the reduced *N*-lysylipoamide of the transacylase component of the multienzyme complex (Equation (14)).



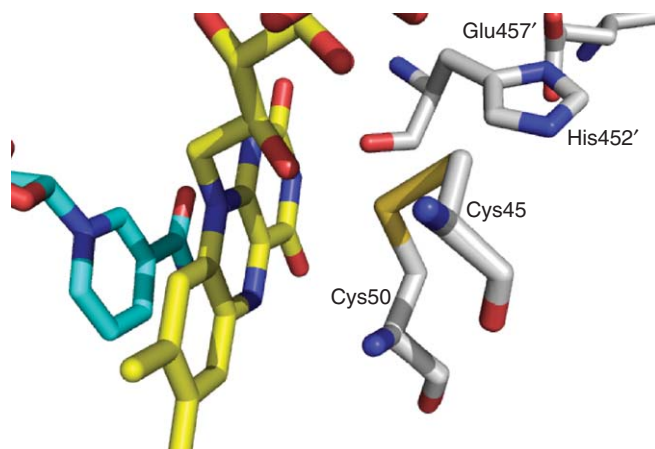
The enzyme's physiological role in energy production dictates that the reaction proceeds in a direction opposite to that of GR, that is, it normally reduces  $\text{NAD}^+$  to NADH.

Several crystal structures of the enzyme from various species have been solved.<sup>352–355</sup> Lipoamide dehydrogenase is structurally homologous to GR. The enzymatic disulfide is found on the *si* face of the flavin along with a histidine that is involved in acid/base chemistry (Figure 5). An isoleucine on the *re* face of the flavin, which is a tyrosine in GR, causes lipoamide dehydrogenase to be fluorescent, while GR is not.<sup>356</sup>

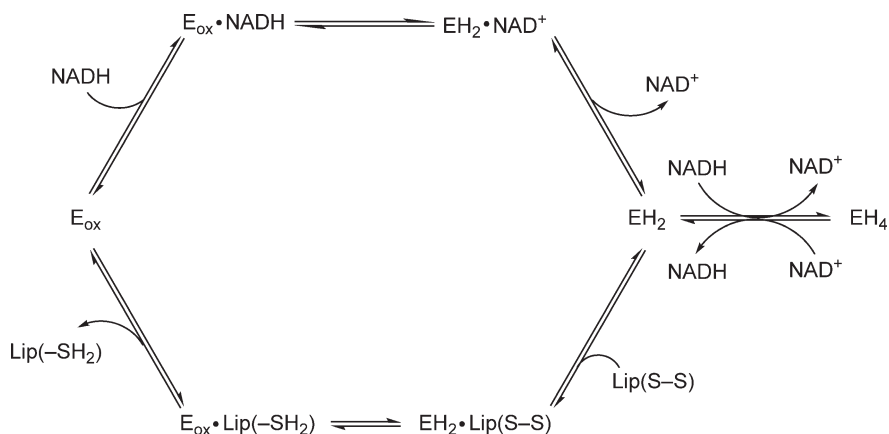
The chemistry of lipoamide dehydrogenase is essentially the same as that of GR. Lipoamide dehydrogenase uses a ping-pong mechanism, which can be divided into two half-reactions, each completely reversible.<sup>357</sup> The reduced flavin intermediate was observed in the reductive half-reaction for the first time in an unmodified enzyme with lipoamide dehydrogenase from *Mycobacterium tuberculosis*,<sup>358</sup> while the thiol-flavin C4a adduct was directly observed in pig heart enzyme whose interchange thiol was alkylated.<sup>317</sup>

The pig heart enzyme is very fast in the physiological direction ( $\text{NAD}^+$  to NADH) with a turnover number of  $550 \text{ s}^{-1}$  at pH 7.6 and  $25^\circ\text{C}$ . In the opposite direction ( $\text{NADH}$  to  $\text{NAD}^+$ ),  $\text{NAD}^+$  has an activating effect because excess NADH will over-reduce the enzyme to  $\text{EH}_4$ , which is not competent for catalysis (Scheme 16).  $\text{NAD}^+$  oxidizes this dead-end complex, increasing the amount of  $\text{EH}_2$ .<sup>359</sup> The enzyme from *E. coli* is especially susceptible to inhibition from over-reduction by NADH because the redox potentials of  $\text{EH}_2$  and  $\text{EH}_4$  are closer than in enzymes from other sources.<sup>358,360,361</sup> As a consequence of closer potentials, the  $\text{EH}_2$  state of the *E. coli* enzyme exists as a rapidly equilibrating mixture of three species: the reduced flavin-disulfide, the oxidized flavin–thiolate charge-transfer, and the oxidized flavin–thiol species.<sup>360</sup>

The behavior of lipoamide dehydrogenase in the intact pyruvate dehydrogenase complex has been investigated.<sup>362</sup> The multienzyme complex consists of 24 pyruvate decarboxylase subunits, 24 dihydrolipoyl transacylase subunits, and 12 lipoamide dehydrogenase subunits. The complex catalyzes a series of reactions beginning with pyruvate decarboxylation, followed by transfer of an acetyl moiety to CoA, and finally the oxidation of dihydrolipoamide by  $\text{NAD}^+$ . Stopped-flow studies showed that lipoamide dehydrogenase in the complex behaves largely the same as it does when studied alone.



**Figure 5** Active site of human lipoamide dehydrogenase complexed with NADH (cyan).



**Scheme 16** Catalytic cycle of lipoamide dehydrogenase.

#### 7.03.4.4 Sulfhydryl Oxidases

Quiescin-sulfhydryl oxidases (QSOXs) catalyze the oxidation of cysteines to disulfides in proteins. Unlike other enzymes of the flavin-dependent disulfide oxidoreductase family, these enzymes do not use a pyridine nucleotide, but instead use molecular oxygen as an electron acceptor. They are found in all metazoans, higher plants, and several protozoan parasites.<sup>363</sup> The protein is a fusion of one or more N-terminal thioredoxin domains, which contain one cysteine disulfide, and a C-terminal ERV1/ALR domain, which contributes two disulfides.<sup>364</sup> The enzyme contains a highly variable C-terminus ending in a single transmembrane-spanning peptide. The C-terminal fragment produced by chymotrypsin treatment contains FAD and retains weak sulfhydryl oxidase activity with DTT, but is unreactive with protein substrates, indicating that the N-terminal domain interacts with the protein substrates.<sup>365</sup> Consistent with this, site-directed mutagenesis of the human enzyme also indicates that the redox active disulfide in the thioredoxin domain is necessary for protein oxidation.<sup>366</sup>

A series of disulfide exchanges occurs in the reductive half-reaction of QSOX. A dithiol substrate such as a denatured protein is oxidized by the disulfide of the thioredoxin domain. The initial substrate–enzyme mixed disulfide likely occurs with the N-terminal solvent-exposed cysteine (Cys80 in the avian enzyme).<sup>363</sup> The reducing equivalents are then passed to the disulfides in the C-terminal ERV1/ALR domain. The roles of the two disulfides in this domain are unclear. Regardless, the electrons are eventually passed to the oxidized flavin, which then reacts with molecular oxygen.<sup>367</sup>

Most substrates are oxidized by QSOX with a  $k_{\text{cat}}$  of  $\sim 1000 \text{ min}^{-1}$ , so that substrate selectivity, determined by  $k_{\text{cat}}/K_m$ , is dominated by  $K_m$ .<sup>367</sup> By this measure, QSOX is much better at oxidizing protein substrates, such as denatured RNase, than small molecules, such as GSH, due to the much smaller  $K_m$  values obtained with reduced proteins.<sup>364,368,369</sup> Buried thiols in folded proteins are unreactive with QSOX,<sup>369</sup> indicating that a protein must be denatured for QSOX to insert disulfides. When denatured RNase is oxidized by QSOX, the correct disulfides are not introduced.<sup>370</sup> However, when protein disulfide isomerase (PDI) and GSH are added to the reaction, active RNase is quickly formed, indicating that the correct disulfide bonds have been introduced<sup>369</sup> and implying that QSOX and PDI work together to form correct disulfide bonds *in vivo*.

#### 7.03.4.5 Mercuric Reductase

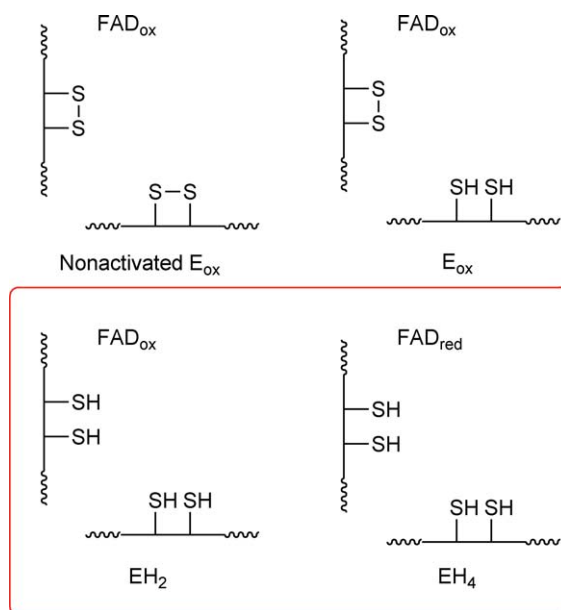
Many bacteria are resistant to inorganic and organic mercury compounds. Mercuric reductase (MerA) is a key enzyme in the mercury detoxification pathway. MerA catalyzes the NADPH-dependent reduction of  $\text{Hg}^{2+}$  to its volatile, uncharged, elemental state ( $\text{Hg}^0$ ). MerA is a cytosolic protein<sup>370</sup> that is homologous to GR,<sup>371,372</sup> but also has a short C-terminal extension and a long N-terminal extension not found in GR.<sup>373</sup> MerA contains three pairs of cysteines: one in the C-terminal extension, one in the N-terminal extension, and one in the GR-like region of the protein. The N-terminal domain binds one molecule of mercury and delivers it to the catalytic core of the protein, made up of the GR-like portion and the C-terminal extension, where it is reduced.<sup>374</sup> The disulfide from

the GR-like portion of the enzyme is termed the inner disulfide, while the other is called the C-terminal disulfide. Active site residues are contributed by both subunits of the dimer, so that the inner cysteines and the C-terminal cysteines are on separate polypeptide chains,<sup>375</sup> a common feature in the flavin-dependent disulfide oxidoreductase family. The crystal structure<sup>375</sup> shows that the cysteines are located on the *si* face of the flavin, the binding site for  $\text{Hg}^{2+}$ . NADPH binds on the *re* face of the flavin, but a tyrosine blocks direct approach, as in GR. The histidine involved in the acid/base chemistry of dithiol exchange in GR is a tyrosine in MerA.

Since the catalytic core of MerA contains four active site thiols, there are four possible redox states of the enzyme (**Scheme 17**):  $\text{EH}_4$ , in which both disulfides and the flavin are reduced;  $\text{EH}_2$ , in which both disulfides are reduced but the flavin is oxidized;  $\text{E}_{\text{ox}}$ , where the C-terminal disulfide is reduced and the flavin is oxidized; and nonactivated  $\text{E}_{\text{ox}}$ , which has two active site disulfides and an oxidized flavin. Nonactivated  $\text{E}_{\text{ox}}$  is not active in assays, but can be converted to an active form of the enzyme by incubation with NADPH.<sup>376,377</sup> Titrations of MerA show that four electrons are required to reach the  $\text{EH}_2$  state of the enzyme.<sup>376</sup>  $\text{EH}_2$  can bind  $\text{Hg}^{2+}$ , but the reaction is not completed until another molecule of NADPH is consumed,<sup>378</sup> showing that the enzyme cycles between  $\text{EH}_2$  and  $\text{EH}_4$  during turnover.

The reduction of  $\text{E}_{\text{ox}}$  to  $\text{EH}_2$  is analogous to the reduction of GR. An oxidized flavin-NADPH charge-transfer complex forms rapidly, followed by formation of the oxidized flavin-thiolate charge-transfer complex at  $117 \text{ s}^{-1}$  at pH 7.3 and  $4^\circ\text{C}$ . NADP<sup>+</sup> dissociation and binding of another NADPH follow<sup>376</sup> with a rate constant of  $5.5 \text{ s}^{-1}$ . During turnover,  $\text{Hg}^{2+}$  binds to  $\text{EH}_2$ , the flavin is reduced by NADPH, and the resulting reduced flavin attacks the charge-transfer sulfur, which is ligated to mercury. The S-Hg bond breaks, producing  $\text{Hg}^0$  and a flavin C4a-thiol adduct. The adduct eliminates thiolate, forming an oxidized flavin-thiolate charge-transfer complex. Mutation of all the active site cysteines to alanine except for the charge-transfer cysteine quantitatively formed the flavin C4a-thiol adduct when the mutant enzyme was titrated with NADP<sup>+</sup>, suggesting that this adduct is a reaction intermediate.<sup>318</sup>

The roles of each pair of active site dithiols were examined using site-directed mutagenesis. When  $\text{HgBr}_2$  was used as a substrate, identical rates were observed for wild-type and a mutant enzyme in which both C-terminal cysteines were mutated to alanine,<sup>379,380</sup> showing that the C-terminal thiols are not required for formation of a reducible  $\text{Hg}^{2+}$  complex, indicating that only the inner thiols bind  $\text{Hg}^{2+}$ . The C-terminal cysteine thiols are proposed to provide a ligand exchange pathway that would remove ligands such as  $\text{CN}^-$  before  $\text{Hg}^{2+}$  reaches the inner cysteine thiols, thus enhancing reactivity by avoiding buildup of negative charge at the active site.<sup>381</sup>

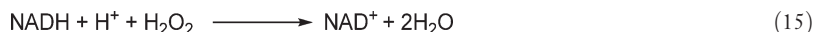


Catalytic forms

**Scheme 17** Four enzyme forms of mercuric reductase.

## 7.03.4.6 NADH Peroxidase

NADH peroxidase reduces peroxides (Equation (15)).



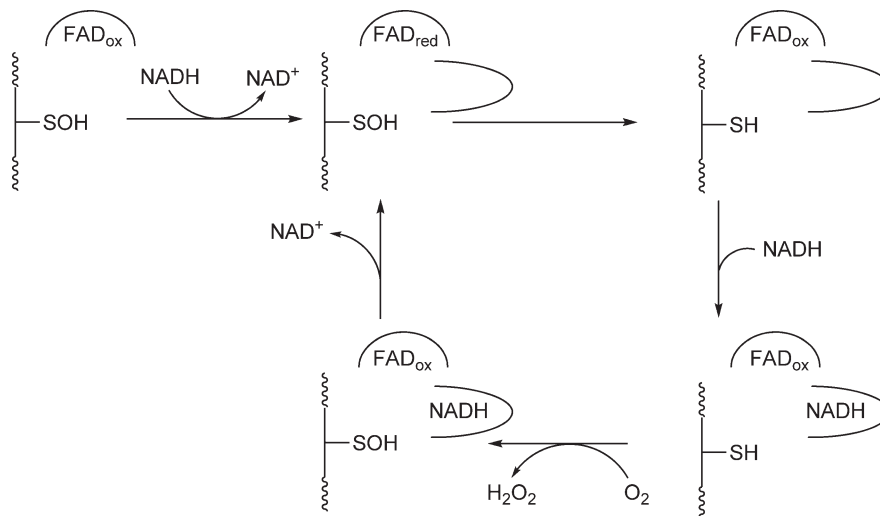
NADH peroxidase is structurally homologous to GR. The charge-transfer thiolate in GR is structurally equivalent to the redox active cysteine in NADH peroxidase;<sup>382</sup> however, no interchange thiol is present in NADH peroxidase. The reactivity of the equivalent active site cysteines with  $\text{H}_2\text{O}_2$  is drastically different, with the cysteine in NADH peroxidase reacting much faster than that in GR.<sup>383,384</sup>

During catalysis, NADH peroxidase cycles between the thiol and the sulfenic acid at the active site cysteine residue.<sup>385</sup> The enzyme is generally in the sulfenic acid form and must undergo a priming step before turnover can begin. The sulfenic acid is first reduced to the thiol by NADH via the flavin (**Scheme 18**). The enzyme then binds another molecule of NADH while the active site cysteine is a thiol. This complex reacts with  $\text{H}_2\text{O}_2$  reforming the sulfenic acid. NADH reduces the flavin, which in turn reduces the sulfenic acid to a thiol. The slowest step in turnover is FAD reduction with a rate constant of  $23 \text{ s}^{-1}$  at pH 7.0 and  $5^\circ\text{C}$ . The enzyme is slowly inactivated by  $\text{H}_2\text{O}_2$  due to the irreversible oxidation of the sulfenic acid to the sulfonate. An active site histidine adjacent to the cysteine residue is proposed to protect the sulfenic acid from over-oxidation. This histidine remains unprotonated throughout the catalytic cycle due to the formation of a hydrogen bond with an arginine.<sup>383</sup>

## 7.03.5 Electron Transfer Reactions

Flavins are unique coenzymes that are able to catalyze both one- and two-electron transfers. Because of this, many flavoproteins are involved in transferring electrons between other proteins. Often, flavoproteins are reduced by two-electron donors, such as pyridine nucleotides, and then pass those electrons one at a time to a single-electron acceptor, such as an iron-sulfur cluster in another protein. Conversely, some enzymes accept single electrons from reduced enzymes. In either case, the flavoenzymes are transferring single electrons; thus, flavin semiquinone is frequently stabilized and observed during turnover.

Single-electron transfer from a flavin to another redox active group, such as an iron-sulfur cluster or another flavin, requires no specific orientation of the two groups participating in the electron-transfer event. In fact, there are examples of electron transfer occurring between two redox groups that are far apart, but generally the groups are close to one another. The intrinsic rate of electron transfer is often very fast; however, the observed rate constant is usually much slower than the presumed intrinsic rate constant. This has led to the idea that



**Scheme 18** Catalytic cycle of NADH peroxidase.



electron transfer is gated. Often, electron transfers are proposed to be conformationally gated, controlled by a conformation change of the enzyme.

### 7.03.5.1 Electron-Transferring Flavoproteins

ETFs are soluble  $\alpha/\beta$ -heterodimeric enzymes containing FAD that are responsible for transferring electrons from dehydrogenases to the membrane-bound respiratory chain.<sup>386,387</sup> The enzymes that donate electrons to ETFs vary among organisms; in humans, many flavin-containing dehydrogenases involved in fatty acid  $\beta$ -oxidation, amino acid oxidation, and choline metabolism pass electrons to ETFs.<sup>388,389</sup> Mutations in ETF cause glutaric acidemia type II, also called multiple acyl-CoA dehydrogenase dysfunctional disease,<sup>390</sup> an autosomal recessively inherited disease of varying severity.<sup>391</sup>

ETFs have been divided into three groups based on sequence homology and function; examples from two of the groups have been studied in detail. Group I ETFs are typically found in mammals and also in some bacteria. Mammalian ETFs accept electrons from many mitochondrial matrix flavoprotein dehydrogenases including ACADs, glutaryl-CoA dehydrogenase, and sarcosine dehydrogenase,<sup>388,389</sup> and then pass these electrons to membrane-bound ETF-ubiquinone oxidoreductase.<sup>392,393</sup> Group II ETFs are found in nitrogen-fixing bacteria and diazotrophic bacteria.<sup>394</sup> A well-studied example of a Group II ETF comes from the bacterium *Methylobacillus methylotrophicus* strain W3A1, which oxidizes trimethylamine dehydrogenase, an iron-sulfur-containing flavoprotein.<sup>395,396</sup>

Structures of the Group I ETF from humans<sup>387</sup> and *Paracoccus denitrificans*<sup>397</sup> and the group II ETF from *M. methylotrophicus*<sup>395</sup> have been solved, but the structures of the group III ETF have not been solved. Each crystal structure contains three domains. Domains I and III form a shallow bowl upon which domain II, the flavin-binding domain, rests. Domain III contains a deeply buried AMP molecule, which is thought to play a structural role. Domain II is connected to the other domains by flexible linkers. The overall folds of the three proteins are very similar, despite low sequence homology between groups. The major difference in structures is in domain II. The flavin-binding domain of the group II ETF is rotated  $\sim 40^\circ$  relative to the group I ETFs. However, small-angle X-ray solution scattering (SAXS) has shown that the solvent envelopes of ETFs from the different groups are nearly identical and do not vary with flavin oxidation state.<sup>388</sup> These results were taken as evidence of multiple discrete conformations of domain II that arise from rotations between  $30$  and  $50^\circ$  with respect to domains I and III.

Structures of complexes of ETF with partner proteins have been solved, including the *M. methylotrophicus* group II ETF complex with trimethylamine dehydrogenase<sup>395</sup> and the human group I ETF complex with medium-chain acyl-CoA dehydrogenase.<sup>398</sup> Interestingly, in both cases, no density was observed for the flavin-binding domain of ETF, indicating that it is very mobile. Both structures revealed that the interaction site between the two proteins was a loop located in domain III, rather than in domain II, as might be expected. ETF recognizes a hydrophobic patch of the partner protein with this loop, leaving the flavin-binding domain of ETF free to sample various conformations, some of which allow electron transfer between the two proteins. When the recognition loop of the *M. methylotrophicus* ETF is removed by proteolysis, the resulting protein reacts with dithionite similar to the wild-type ETF; however, it cannot accept electrons from trimethylamine dehydrogenase,<sup>395</sup> indicating the importance of the recognition loop in protein-protein interactions.

The anionic semiquinone of ETF is very stable, so that the hydroquinone is normally not accessible. FAD is bound to domain II of ETF in a bent conformation due to a hydrogen bond between the 4'-OH of the ribityl side chain and N1 of the isoalloxazine ring. This interaction is proposed to stabilize the semiquinone of the semiquinone-hydroquinone redox couple. The oxidized flavin-semiquinone redox couple in the ETF from *M. methylotrophicus* is  $+153$  mV,<sup>22</sup> suggesting that the anionic semiquinone is stabilized. Mutagenesis of a serine residue that hydrogen bonds to N5 of the isoalloxazine ring system has shown the importance of this interaction in both stabilizing the semiquinone and establishing the high redox potential observed for the oxidized flavin-semiquinone couple.<sup>399,400</sup> The redox potential of the semiquinone-hydroquinone couple is less than  $-250$  mV. The redox potentials increase when ETF binds trimethylamine dehydrogenase<sup>401</sup> due to a conformational change in the flavin-binding domain.

ETF is the oxidizing substrate for the oxidation of trimethylamine catalyzed by trimethylamine dehydrogenase. The dehydrogenase has a flavin and an iron-sulfur cluster, allowing it to be reduced by up to three electrons. Trimethylamine is a two-electron reductant, while ETF is a one-electron oxidant. Therefore, depending on the

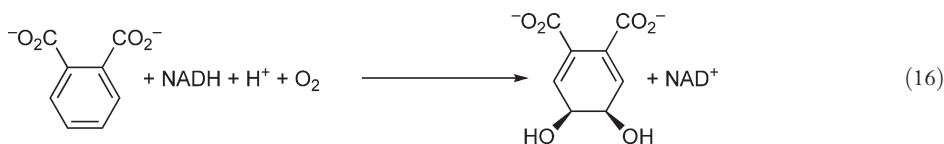
rates of reduction and oxidation, the dehydrogenase could cycle between the fully oxidized and two-electron reduced state, or the one-electron and three-electron reduced states, complicating steady-state turnover.<sup>397,402</sup>

Stopped-flow studies on electron transfer between two-electron reduced trimethylamine dehydrogenase and ETF are complicated by internal electron-transfer steps between the flavin and iron–sulfur center of trimethylamine dehydrogenase;<sup>403</sup> therefore, a modified enzyme in which the FMN prosthetic group was derivatized by phenylhydrazine, making it redox inert, was used.<sup>404</sup> This modified enzyme, reduced by one electron by dithionite, reacts with oxidized ETF<sup>402</sup> at 5 °C and pH 7.0 with a bimolecular rate constant of  $1.44 \times 10^6 \text{ mol}^{-1} \text{ s}^{-1}$ . Simulations of the kinetic scheme including binding of the two proteins, reversible conformational changes of the flavin-binding domain, and irreversible electron transfer have shown that if the rate constant for electron transfer is higher than  $10^3 \text{ s}^{-1}$ , a linear concentration dependence between the observed rate constant and ETF concentration will be observed.<sup>402</sup> This model predicts that complex formation is the rate-limiting step at accessible enzyme concentrations and that both the conformational changes and electron transfer occur very quickly.

Medium-chain acyl-CoA dehydrogenase also donates electrons to ETF and this reaction has been studied kinetically as well. Medium-chain acyl-CoA dehydrogenase is reduced by two electrons by an acyl-CoA thioester (see Section 7.03.2). The electrons are then passed one at a time to ETF reforming oxidized dehydrogenase. The product of the reductive half-reaction, enoyl-CoA, is not released until the enzyme has been fully oxidized.<sup>234</sup>

### 7.03.5.2 Phthalate Dioxygenase Reductase

Phthalate is oxidized to a *cis*-dihydrodiol by *Burkholderia cepacia* DB01 in a reaction requiring molecular oxygen and NADH (Equation (16)).



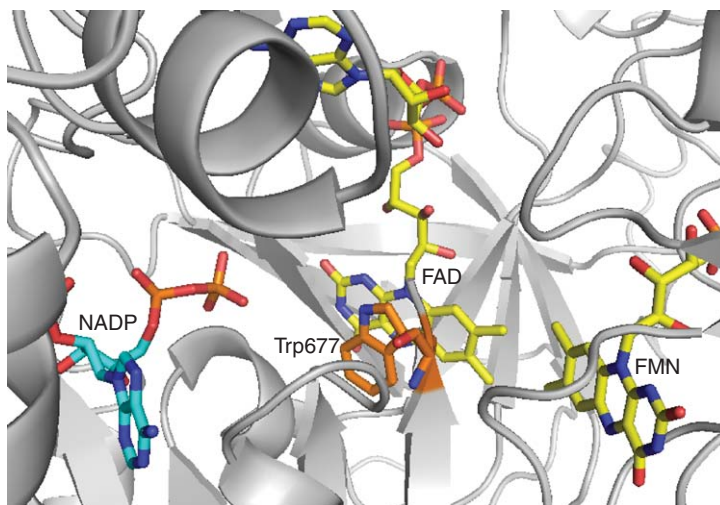
This reaction is catalyzed by a two-component dioxygenase system consisting of phthalate dioxygenase reductase (PDR), a 36 kDa protein containing both FMN and a [2Fe–2S] center, and phthalate dioxygenase (PDO), a multimeric protein with each 50 kDa subunit containing one Rieske [2Fe–2S] center and one mononuclear ferrous iron.<sup>405</sup> The iron–sulfur flavoprotein PDR is used to transfer electrons from NADH (a two-electron donor) to its partner protein PDO (a one-electron acceptor).

PDR has three domains: an N-terminal flavin-binding domain, a central pyridine nucleotide-binding domain, and a C-terminal [2Fe–2S]-containing domain. The N-terminal domain is a  $\beta$ -barrel with FMN bound to the outside of the barrel adjacent to the NADH-binding site. However, the nicotinamide ring does not contact the flavin in any of the complexes.<sup>406,407</sup> When the nicotinamide ring is visible in structures, a phenylalanine intervenes between the flavin and the nicotinamide ring. This residue must move to allow the nicotinamide ring and the flavin to react.

When PDR is titrated with NADH or dithionite, two stages of reduction are observed. First, the iron–sulfur center and FMN react simultaneously to form a reduced iron–sulfur center and a neutral flavin semiquinone. Further titration reduces the semiquinone to the hydroquinone. During the reduction, a maximum semiquinone concentration of  $\sim 80\%$  of the total enzyme concentration is reached. The redox potential of the [2Fe–2S] center and that of the oxidized flavin–semiquinone couple are the same,  $-174 \text{ mV}$ . The semiquinone–hydroquinone couple is well resolved from this at  $-287 \text{ mV}$ .<sup>408</sup> These midpoint potentials favor spontaneous electron transfer from NADH to FMN to [2Fe–2S].

The kinetics of reduction of PDR by NADH has been studied in stopped-flow experiments.<sup>408</sup> At pH 8.0 and 4 °C, PDR binds NADH very rapidly, within the dead-time of the stopped-flow instrument, with a  $K_d$  of  $50 \mu\text{mol l}^{-1}$ . The first observable kinetic phase is the formation of an oxidized flavin–NADH charge-transfer complex with long-wavelength absorbance extending beyond 800 nm, occurring at  $116 \text{ s}^{-1}$ . Next, hydride transfer occurs at  $70 \text{ s}^{-1}$ , forming a reduced flavin–NAD<sup>+</sup> charge-transfer complex. A deuterium isotope effect of  $\sim 7$  was observed on this phase when using NADD as a substrate. The next observable phase ( $35 \text{ s}^{-1}$ ) involves





**Figure 6** Cytochromes P-450 reductase complexed with NADP (cyan). Trp677 (orange) is stacked against FAD and must move for hydride transfer from NADPH.

Temperature-jump experiments with two-electron reduced CPR were used to investigate internal electron transfer between the flavins.<sup>422</sup> Two-electron reduced CPR is a 50/50 mixture of two species: enzyme in which both flavins are in the semiquinone form and enzyme in which the FAD is oxidized while the FMN is fully reduced. Increasing the temperature shifts the equilibrium between these two states allowing electron transfer from FAD to FMN to be observed. Two reaction phases are observed during the experiment. The fast phase ( $\sim 2000 \text{ s}^{-1}$ ) represents a conformational change around the FAD. The second phase ( $\sim 55 \text{ s}^{-1}$ ) represents electron transfer. This observed rate constant is much slower than the estimated intrinsic rate constant of electron transfer ( $10^{10} \text{ s}^{-1}$ ) based on the distance between the two flavins.<sup>421</sup> The rate of electron transfer is proposed to be conformationally gated.<sup>422,423</sup>

Another Class II P-450 redox system that has been extensively studied is the cytosolic flavocytochrome P-450 BM3 from *Bacillus megaterium*. BM3 is the fusion of a soluble P-450 domain with CPR.<sup>424,425</sup> The FAD of BM3 is reduced by NADPH; the electrons are transferred to FMN and then finally to the substrate-bound P-450 domain.<sup>414</sup> BM3 is the fastest reported P-450 monooxygenase,<sup>426</sup> with the rate of hydride transfer from NADPH to FAD and the rate of electron transfer from FMN to heme several-fold above the mammalian P-450 redox systems.<sup>427</sup> The structure of the full-length protein has yet to be solved, but the structure of the FAD- and NADPH-binding domain has been determined.<sup>428</sup> This domain closely resembles rat liver CPR and contains several conserved residues implicated in NADPH binding and flavin reduction.

### 7.03.6 Oxygen Reactions

Life on Earth arose under reducing conditions, but within a billion years or so the evolution of photosynthetic organisms filled the atmosphere with  $\text{O}_2$ , and oxygen became a player in biochemistry. Molecular oxygen has unique chemical properties. It is a small hydrophobic molecule, capable of rapidly diffusing into protein matrices through the crevices that open in the course of normal dynamic fluctuations.<sup>429</sup> The reduction of  $\text{O}_2$  is extremely favorable thermodynamically, but nonetheless occurs very slowly with most biochemicals because the electronic ground state of  $\text{O}_2$  is the triplet (two unpaired electrons), while the electronic ground state of most biochemicals is the singlet (all electron spins paired). Spin must be conserved in chemical reactions, making a single-electron transfer an obligatory first step in the oxidation of an organic molecule by  $\text{O}_2$ . Most organic radicals are extremely unstable so that their formation is highly disfavored, affording kinetic stability even in the presence of this thermodynamically powerful oxidant.

Flavins are different. Their ability to form stable semiquinones has been mentioned earlier. Even in an environment (such as water or many proteins) that disfavors semiquinone formation, the flavin radical is often

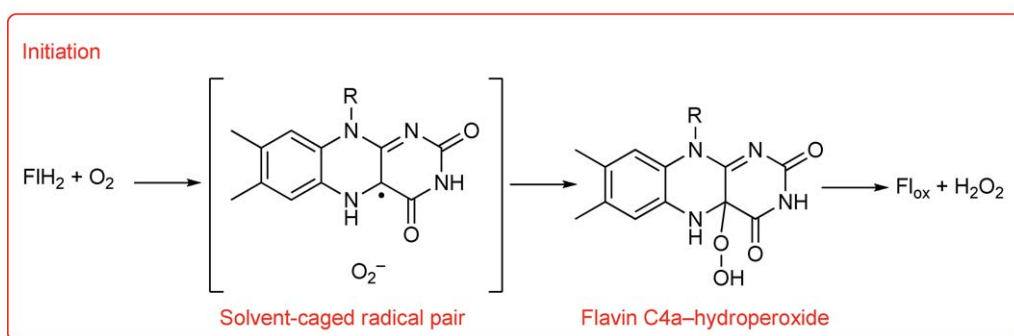
stable enough so that the exergonic single-electron transfer to  $O_2$ , giving  $O_2^-$ , drives the reaction. Thus, flavins in the hydroquinone state are generally oxidized by  $O_2$  much faster than typical organic molecules and at rates with biochemical significance. The presence of  $O_2$  in ecosystems, therefore, created new selective pressures on preexisting flavoenzymes and also new opportunities for expanding chemical repertoires. Some flavoenzymes evolved by suppressing reactivity with  $O_2$ , while others evolved with enhanced reactivity with  $O_2$ . Oxygen is used by oxidases as a substrate for the disposal of electrons generated in metabolism, driving oxidations to completion. The  $H_2O_2$  produced is detoxified by other enzymes, such as catalase, or might be used as a signal. Hydroxylases harness the reactivity of reduced oxygen intermediates to hydroxylate organic or (less frequently) inorganic substrates. This section presents the basic chemistry of flavin–oxygen reactivity; some ways by which proteins might enhance, inhibit, or otherwise control this reactivity; and some examples of oxidases and hydroxylases.

### 7.03.6.1 Reaction of Free Flavins with $O_2$

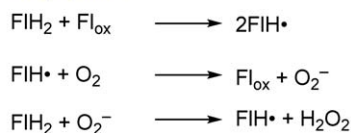
Reduced free flavins – FMN, FAD, riboflavin, or artificial models – react directly with  $O_2$  in aqueous solution rather slowly; for instance, tetraacetyl riboflavin reacts with a bimolecular rate constant<sup>430</sup> of  $250 \text{ mol}^{-1} \text{ s}^{-1}$ . Spin conservation requires the first step to be an electron-transfer reaction, forming a solvent-caged radical pair; this reaction is rate determining. Several routes may be imagined for the decomposition of the superoxide–semiquinone intermediate. Considerations of diffusion and proton-transfer rates, redox potentials, and  $pK_a$  values led to the conclusion that the radical pair forms the C–O bond in a C4a–peroxide, which is protonated to yield the corresponding hydroperoxide.<sup>431</sup> Flavin hydroperoxides eliminate  $H_2O_2$  rapidly in aqueous solution unless N5 is alkylated, and even then are unstable in protic solvents. Therefore, the putative hydroperoxide intermediate is never observed when aqueous flavin hydroquinones react with  $O_2$ . Aqueous flavin hydroperoxide generated by pulse radiolysis<sup>432</sup> eliminates  $H_2O_2$  with a rate constant of  $260 \text{ s}^{-1}$ , much higher than the maximum possible pseudo-first-order rate constant ( $\sim 0.25\text{--}0.5 \text{ s}^{-1}$ ) for its formation by the bimolecular reaction of  $O_2$ .

The bulk of free reduced flavin does not go through a hydroperoxide intermediate when reacting with  $O_2$  (Scheme 20). Instead, the reaction is autocatalytic.<sup>430</sup> Oxidized flavin, initially formed through the slow

#### Oxygen reactions Free flavins



#### Propagation



**Scheme 20** Oxidation of free reduced flavin by molecular oxygen.



hydroperoxide pathway, reacts rapidly with the remaining flavin hydroquinone to form semiquinone. Semiquinone reacts rapidly with  $O_2$ , forming oxidized flavin and superoxide;  $O_2^-$  reacts rapidly with flavin hydroquinone, forming  $H_2O_2$  and semiquinone, which propagates the chain. Because of the autocatalytic chain reactions, there is a distinct lag in the production of oxidized flavin, and this lag is lengthened by including superoxide dismutase.

These nonenzymatic reactions define the scope of possibilities available to flavoproteins. The protein environment provides a wide array of reactivities for the isoalloxazine. At one extreme, flavoprotein hydroxylases react rapidly with  $O_2$  and form C4a-hydroperoxides, which are often readily observed. Oxidases react quickly but rarely with observable intermediates. At the other extreme, many flavoenzymes, such as dehydrogenases, react with  $O_2$  sluggishly – some even slower than free reduced flavins – and in these cases usually produce superoxide as the initial oxidation product.

### 7.03.6.2 Oxidases

Flavoprotein oxidases react rapidly with  $O_2$ , with bimolecular rate constants ranging from  $\sim 10^4 \text{ mol}^{-1} \text{ s}^{-1}$  to more than  $\sim 10^6 \text{ mol}^{-1} \text{ s}^{-1}$ . Generally, the reductive half-reaction accomplishes the biologically important chemistry and  $H_2O_2$  can be regarded as a means to discard electrons; a notable exception is the macrophage NADPH oxidase, which produces  $H_2O_2$  to combat infections. The kinetic mechanisms of oxidases depend on competing rates. Reduced oxidases often react with  $O_2$  after the product of the reductive half-reaction dissociates, but in many instances  $O_2$  reacts faster than the product dissociates, so that the kinetic mechanism is not always ping-pong. For oxidases with broad substrate specificity, the product dissociation rate, and therefore the kinetic mechanism, can vary with the substrate (e.g., DAAO). With some enzyme–substrate combinations (e.g., monoamine oxidase B and benzylamine<sup>433</sup>), product dissociation and substrate association are rapid enough compared to the reaction of  $O_2$  that it is the reduced enzyme–substrate complex that reacts with  $O_2$ .

Until recently, there had been no indication in oxidases of an intermediate prior to forming oxidized enzyme and  $H_2O_2$ . In stopped-flow experiments with pyranose-2-oxidase, a C4a-hydroperoxide has been unambiguously detected.<sup>434</sup> This is the only instance this intermediate has been detected in flavoprotein oxidases – its generality is not yet clear. If this intermediate or a semiquinone is on the pathway of flavin oxidation in other oxidases, then the rate constant for intermediate decay must be very high. Interestingly, an attempt to synthesize the flavin hydroperoxide on GOX was successful.<sup>435</sup> A mixture of the enzyme semiquinone and superoxide was rapidly generated by pulse radiolysis, with radical recombination yielding the hydroperoxide, identified by its characteristic absorbance spectrum. However, the intermediate generated in this way was not kinetically competent. It eliminated  $H_2O_2$  with a rate constant of  $350 \text{ s}^{-1}$ , far too slow to be consistent with stopped-flow experiments. High-resolution crystal structure of DAAO<sup>56</sup> detected unidentified electron density over the face of the isoalloxazine, which might be consistent with a trapped hydroperoxide, but these results are ambiguous. Thus, it is not proven whether all oxidases form a fleeting hydroperoxide or simply reduce the initial superoxide intermediate very rapidly.

Although it is clear that proteins control the reactivity of oxygen with reduced flavins, it is unclear how.<sup>436</sup> Key concepts explaining the marked decrease in oxygen reactivity caused by ligand binding to ACADS are generally applicable to other flavoprotein/ $O_2$  reactions, including oxidases.<sup>437</sup> Biophysical studies, such as the quenching of fluorescence, show that  $O_2$  has access even to deeply buried protein residues,<sup>429</sup> so that reduced flavin is unlikely to be sterically protected by a protein. Conversely, reactivity is unlikely to be stimulated greatly by access channels for  $O_2$ . Instead, reactivity can be controlled by influencing the energetics of the obligate semiquinone–superoxide intermediate. A single-electron transfer to oxygen creates an anion whose stability will be very sensitive to polarity, local charges, hydrogen bonds, etc. According to the Hammond postulate, the transition state ought to resemble the diradical intermediate, so the polarity of the environment provided by the protein should influence the rate of reaction. Ligand binding to ACADS expels water, making the active site very hydrophobic and, presumably, inhospitable to the formation of superoxide, causing the low oxygen reactivity.<sup>437</sup> Conversely, oxidases could provide the polar environment that would enhance  $O_2$  reactivity.



This provided the underpinnings for investigations on oxygen reactivity of GOX. A combination of steady-state kinetics,  $^{18}\text{O}_2$  and solvent isotope effects, pH dependencies, viscosity effects, and site-directed mutagenesis points to the importance of His516 in accelerating the reaction of oxygen.<sup>438</sup> At very high pH values, when the histidine is not positively charged, or with the His516Asn mutant, the enzyme is oxidized at nearly the same rate as free flavin. The positive charge of protonated His516 decreases the activation barrier for the rate-determining formation of the superoxide–semiquinone intermediate. Other oxidases may use the same strategy,<sup>436</sup> including type II cholesterol oxidase, which has an Arg near the flavin. Interestingly, although the  $\text{O}_2$  reaction of most site-directed mutants of oxidases is perturbed only by small amounts, the mutation of Lys256 of MSOX to a neutral amino acid lowered the rate constant by a factor of  $\sim 8000$ , indicating the importance of the positive charge near the flavin in  $\text{O}_2$  reactivity.<sup>439</sup> Some amine oxidases generate what might be an oxygen-activating positive charge by the reductive half-reaction and then react with  $\text{O}_2$  before the iminium product dissociates, accelerating flavin oxidation about an order of magnitude over the rate of free reduced enzyme. However, counterexamples are known in which iminium complexes react slower with  $\text{O}_2$  than the free enzyme (e.g., cytokinin dehydrogenase), as well as examples in which noncationic ligands enhance flavin reactivity (e.g., LMO). Some oxidases (and hydroxylases) contain anion-binding pockets near the isoalloxazine, which might stabilize superoxide as it forms. However, such pockets are not present in all oxidases. In fact, no single structural feature – positive charge, anion-binding pocket, etc. – is common to all oxidases, suggesting that there are many ways to promote oxygen reactivity.

### 7.03.6.3 Hydroxylases

A large number of flavoenzymes react with  $\text{O}_2$  rapidly to form flavin C4a-hydroperoxide intermediates, which are used in much the same way a synthetic chemist uses organic hydroperoxides. The terminal, or distal, oxygen of the flavin hydroperoxide acts as an electrophile or a nucleophile, depending on the enzyme; the hydroperoxide of some enzymes reacts in either fashion, depending on the substrate. Interestingly, while organic hydroperoxides are widely used by synthetic chemists in the epoxidation of olefins, examples are rare in flavin enzymology. Styrene monooxygenase, a two-component system isolated from bacteria, is being studied for use in synthetic chemistry.<sup>440,441</sup> Squalene monooxygenase, a membrane-bound enzyme, is involved in cholesterol biosynthesis.<sup>442</sup> The mechanism of neither enzyme has been studied in detail. Hydroxylations occur widely in biology. They are important in the degradation of phenolic compounds derived from lignin and, therefore, account for a significant proportion of the carbon flux in the biosphere. Hydroxylations are also important routes for drug and xenobiotic metabolism in mammals, and are being discovered in signal-transduction and antibiotic synthesis pathways.

Hydroxylation of the substrate – the biologically important reaction – occurs in the oxidative half-reaction. Catalytic cycles are organized around the flavin hydroperoxide: synthesizing it *in situ* from reduced flavin and  $\text{O}_2$ , using the hydroperoxide to hydroxylate the substrate while protecting it from solvent to prevent elimination of  $\text{H}_2\text{O}_2$ , and then eliminating water after O–O bond cleavage. Sometimes the high reactivity of the initial hydroxylation product (often an unstable intermediate) is intercepted for nonredox reactions. Hydroxylases reduce the flavin prosthetic group with pyridine nucleotide in the reductive half-reaction. The very different requirements for promoting hydride transfer, hydroperoxide formation and stabilization, and substrate hydroxylation have led to remarkable control features in hydroxylases, including dramatic conformational changes, which spread the chemistry to different regions of the protein. Other hydroxylases accomplish the varied chemistry by using a separate reductase component and transferring the reduced flavin to the hydroxylase component.<sup>443</sup> In such cases, the flavin is actually a recycled substrate rather than a prosthetic group. New two-component reductase/monooxygenase systems are rapidly being discovered and are probably very common in nature. Both nucleophilic and electrophilic flavin hydroperoxides are formed by one- and two-component hydroxylases.

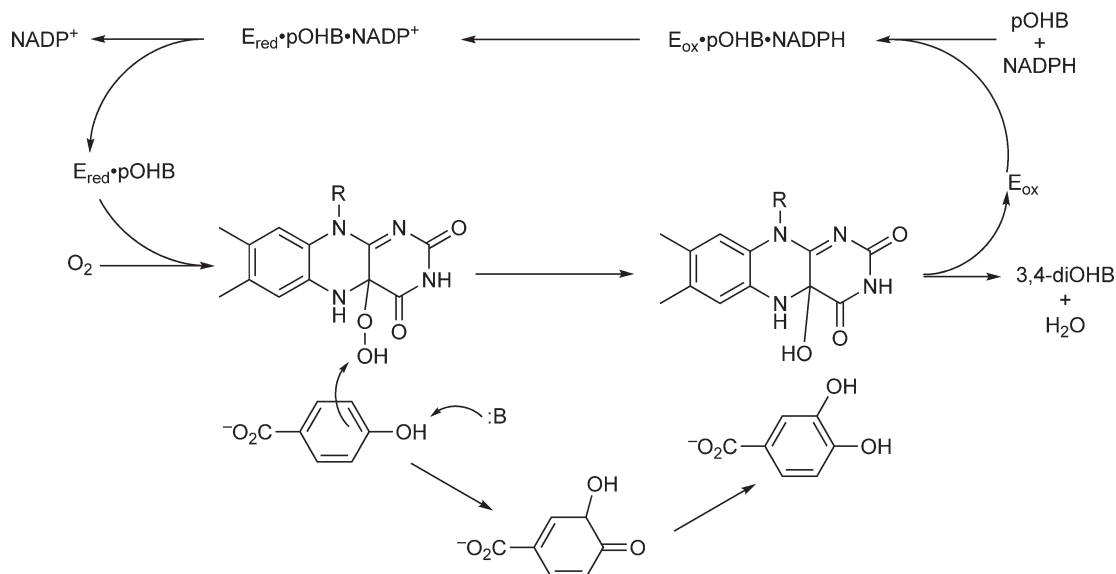
### 7.03.6.4 *p*-Hydroxybenzoate Hydroxylase

The most extensively studied flavoprotein monooxygenase is PHBH,<sup>444</sup> a homodimeric FAD-containing enzyme that hydroxylates *p*OHB using NADPH and  $\text{O}_2$ . PHBH is a bacterial enzyme, and is one of the many hydroxylases and dioxygenases catabolizing phenols liberated in the degradation of lignin. Carbon

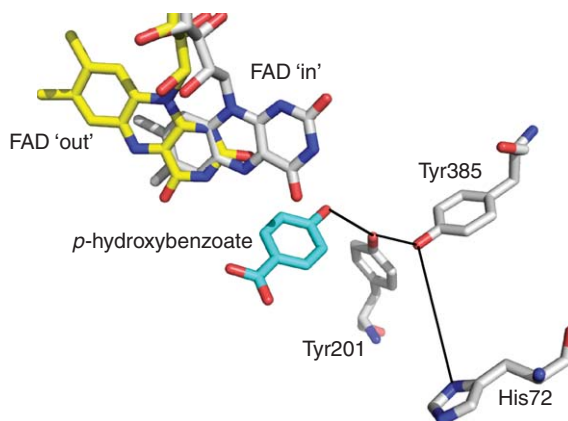
derived in such oxidations is funneled into metabolism through the Krebs cycle. PHBH, especially from *Pseudomonas fluorescens*, has been a model system for flavoprotein hydroxylases. The vast amount of mechanistic and structural work on PHBH provides a detailed insight into the regulation of a complex catalytic cycle and the chemistry of oxygen transfer, framing issues pertaining to all hydroxylase reactions.

In the reductive half-reaction, pOHB and NADPH bind in random order to the oxidized enzyme, NADPH reduces FAD, and  $\text{NADP}^+$  dissociates<sup>445</sup> (Scheme 21). Although pOHB does not react until the oxidative half-reaction, its presence accelerates reduction by about five orders of magnitude. PHBH detects pOHB by a proton-transferring network of hydrogen bonds that runs through the interior of the protein.<sup>446</sup> This network generates the phenolate of pOHB, whose negative charge triggers a large pendulum-like movement of the isoalloxazine from its largely buried 'in' conformation to a more exposed 'out' conformation where it reacts with NADPH<sup>135,447,448</sup> (Figure 7). Hydride transfer generates  $\text{FADH}^-$ , which is drawn back to the 'in' conformation and  $\text{NADP}^+$  dissociates.<sup>449</sup> Once the flavin is reduced, the exchange rate of pOHB from the active site plummets many orders of magnitude, effectively trapping the aromatic substrate in the active site for the

### Hydroxylases



**Scheme 21** Catalytic cycle of *p*-hydroxybenzoate hydroxylase.



**Figure 7** Active site of *p*-hydroxybenzoate hydroxylase showing two conformations of the flavin and the proton-transfer network. Internal waters that are part of the network are not shown.

oxidative half-reaction.<sup>450</sup> Thus, coupling reduction to the generation of the phenolate of pOHB tests for the presence of the correct substrate prior to committing NADPH, and may prevent the hydroxylation of *p*-aminobenzoate, a vital metabolite and excellent analogue of pOHB.<sup>451</sup>

The flavin hydroperoxide is formed from the rapid reaction ( $>10^5 \text{ mol}^{-1} \text{ s}^{-1}$  at 4 °C) of O<sub>2</sub> with the reduced enzyme–pOHB complex.<sup>450</sup> This intermediate is stable enough to be observed in stopped-flow experiments, but quickly hydroxylates pOHB. When using some poor substrates or mutated enzymes, the hydroperoxide eliminates H<sub>2</sub>O<sub>2</sub> without oxidizing the aromatic ligand. In the absence of ligands, the reduced enzyme reacts with O<sub>2</sub> 10-fold slower and the flavin hydroperoxide eliminates H<sub>2</sub>O<sub>2</sub> too quickly to be detected. Oxygen is transferred to pOHB from the hydroperoxide in an electrophilic aromatic substitution *ortho* to the electron-donating phenolic oxygen, forming 3,4-dihydroxybenzoate. Hydroxylating C3 of pOHB must transiently generate a nonaromatic product, but this is not observed, presumably because it rapidly tautomerizes to the aromatic product. However, PHBH will hydroxylate other activated benzoates such as 2,4-dihydroxybenzoate and *p*-aminobenzoate, and the nonaromatic intermediate is observed.

Deprotonating the phenolic oxygen of pOHB is critical for the hydroxylation reaction – the initial nonaromatic ketone produced by hydroxylation would have a prohibitively low p*K*<sub>a</sub>. pOHB could be deprotonated to phenolate (enhancing nucleophilicity) prior to attacking the electrophile, or it could be deprotonated during the reaction. However, pOHB bound to PHBH is buried from solvent, so the necessary proton transfers require the hydrogen bond network<sup>446</sup> mentioned above for its role in controlling flavin reduction. This network of two tyrosines, two water molecules, and a histidine extends through the protein interior where it communicates with the external solvent ~12 Å from the active site (see [Figure 7](#)). The network also lowers the phenolic p*K*<sub>a</sub> of pOHB by about two units, activating it as a nucleophile. Disrupting the network by mutagenesis prevents hydroxylation and the flavin hydroperoxide eliminates H<sub>2</sub>O<sub>2</sub>, while mutations that increase the p*K*<sub>a</sub> of pOHB but leave the hydrogen bond network intact slow, but do not prevent, hydroxylation.<sup>446,452</sup>

Attack of the aromatic nucleophile on the flavin hydroperoxide expels the anionic flavin C4a-oxide, which is protonated to the flavin C4a-hydroxide. The low p*K*<sub>a</sub> of the flavin hydroxide, a result of the electron-withdrawing isoalloxazine moiety, makes the C4a-oxide an excellent leaving group in heterolytic O–O cleavage and is responsible for the effectiveness of the flavin hydroperoxide as an electrophile. A Brønsted analysis of hydroxylation of pOHB was accomplished by replacing the natural FAD of PHBH with a series of flavins substituted at the 8-position.<sup>453</sup> Rate constants for hydroxylation increased as the calculated p*K*<sub>a</sub> of the flavin C4a-hydroxide decreased, giving a β-value of –0.41, indicating the development of negative charge in the transition state.

After oxygen transfer, the hydroxyflavin eliminates water, regenerating oxidized flavin, and the product dissociates from the enzyme. Although these events are observed as a single kinetic phase, it is likely that product release occurs first because the hydroxyflavin form of the enzyme can be trapped at very high aromatic substrate concentrations. This is the origin of inhibition by substrate in PHBH, a phenomenon common among the aromatic hydroxylases. Structures show no clear path for ligand exchange, so conformational changes are necessary and likely to involve flavin movement.

Many of the features of catalysis by PHBH – flavin movement, substrate activation, stimulation of reduction by aromatic ligand – are common in all aromatic hydroxylases, although details often vary and generally are not known as well. The long internal hydrogen bond network appears to be unique to PHBH; other hydroxylases have different mechanisms for deprotonating their substrates and controlling the conformation of their flavins, and these are not necessarily linked as in PHBH. However, the movement of the flavin between an ‘in’ and ‘out’ conformation appears to be common to all one-component aromatic hydroxylases. It resolves the conflicting chemical requirements of exposing the isoalloxazine for reaction with pyridine nucleotide and sequestering it to prevent elimination of H<sub>2</sub>O<sub>2</sub>. It is tempting to speculate (so I will) that the ‘in’ and ‘out’ flavin conformations of one-component aromatic hydroxylases arose from the fusion of the reductase and monooxygenase from a two-component system, with the conformational change representing a much abbreviated (and more efficient) transfer of flavin between active sites.

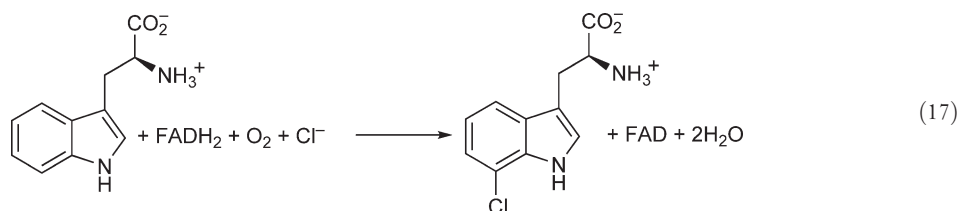
### 7.03.6.5 MICAL

While the majority of aromatic hydroxylases are bacterial catabolic enzymes, several have been discovered in mammals, playing a role in ubiquinone biosynthesis, neurotransmitter chemistry, and cell differentiation.<sup>5</sup>

MICAL is a complex signal-transduction enzyme that causes actin depolymerization, collapsing axon growth cones in developing neurons, thereby shaping the connectivity of the nervous system. The FAD of MICAL is critical for the correct development of the nervous system of fruit flies. Little is known about MICALs. They are activated by membrane receptors called plexins, which in turn respond to signals from proteins called semaphorins. MICALs are large enzymes,  $\sim 1000$  residues long, composed of five domains. The largest, N-terminal, domain is highly homologous to PHBH and other aromatic hydroxylases, while the other domains are homologous to various signal-transduction domains. Structures of the N-terminal domain strongly resemble bacterial hydroxylases.<sup>454,455</sup> When oxidized, the isoalloxazine of the N-terminal MICAL domain adopts an 'out' conformation, while the reduced isoalloxazine adopts an 'in' conformation. Even when the flavin is 'in', the isoalloxazine is unusually accessible to solvent, suggesting that some other macromolecular component – either one of the four domains absent from the N-terminal construct or a binding partner – binds and blocks access to the flavin. Consistent with this, the N-terminal domain is an active NADPH oxidase,<sup>454</sup> which is very unusual for an aromatic hydroxylase. Experiments with truncated MICAL constructs indicate that the C-terminal 290 amino acids are autoinhibitory.<sup>456</sup> No substrate for hydroxylation has yet been identified. While it has been speculated that MICAL transmits its signal by producing active oxygen species ( $\text{H}_2\text{O}_2$  or  $\text{O}_2^-$ ), this would be unprecedented behavior for an aromatic hydroxylase scaffold. Another signaling protein, CRMP, has been identified as a binding partner of the N-terminal domain of MICAL. Interestingly, its binding appears to prevent  $\text{H}_2\text{O}_2$  production, suggesting that CRMP might be the substrate of MICAL or might be a carrier of a small-molecule substrate. The latter possibility is suggested by the similarity of CRMP to enzymes of the dihydropyrimidinase family.

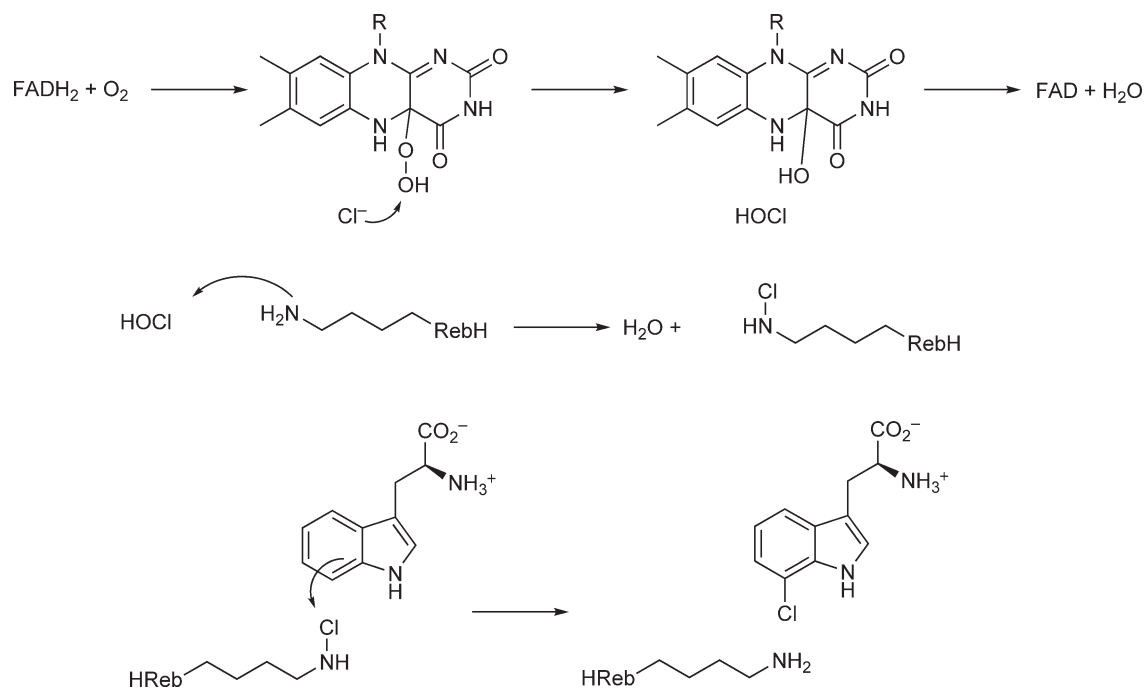
### 7.03.6.6 Tryptophan-7-Halogenase

Hydroxylases can synthesize unstable intermediates because of the very strong oxygen-donating ability of flavin hydroperoxides. These intermediates can then perform normally inaccessible reactions. One example of this is RebH. RebH chlorinates Trp at the 7-position (Equation (17)) in the synthesis of the antibiotic rebeccamycin.



Related halogenases appear to be involved in the synthesis of other antibiotics such as chloramphenicol and vancomycin.<sup>457</sup> RebH is the monooxygenase component of a two-component hydroxylase system. It uses reduced FAD produced by a reductase,  $\text{Cl}^-$ , Trp, and  $\text{O}_2$  as substrates in an interesting variation of standard flavin hydroperoxide chemistry. RebH uses its hydroperoxide to hydroxylate  $\text{Cl}^-$ , forming HOCl, a powerful chlorinating agent (Scheme 22). The reaction of  $\text{O}_2$  involves the usual flavin hydroperoxide and flavin hydroxide intermediates, detected by absorbance and fluorescence changes in stopped-flow experiments.<sup>458</sup> These experiments also showed that the reaction of reduced flavin was not influenced by the presence of Trp. Reduced flavin binds rapidly, and this initial complex can either react rapidly with  $\text{O}_2$  to form the flavin hydroperoxide or isomerize reversibly to an unreactive form that slowly converts to a reduced flavin–enzyme complex that oxidizes without observable intermediates. Chemical quenching experiments showed that 7-chlorotryptophan forms  $\sim 3$ -fold more slowly than oxidized FAD forms, suggesting that these processes are independent.

The fate of HOCl is tightly controlled by the protein. Structures of PrnA, a related halogenase,<sup>459</sup> and later of RebH itself,<sup>460</sup> show that there is a tunnel  $\sim 10$  Å long between the flavin and the substrate Trp-binding site. Rather than allowing HOCl to diffuse through this tunnel to the substrate, it chlorinates a lysine side chain near the middle of the tunnel<sup>461</sup> (Scheme 22). This lysyl chloramine is the penultimate chlorinating agent,



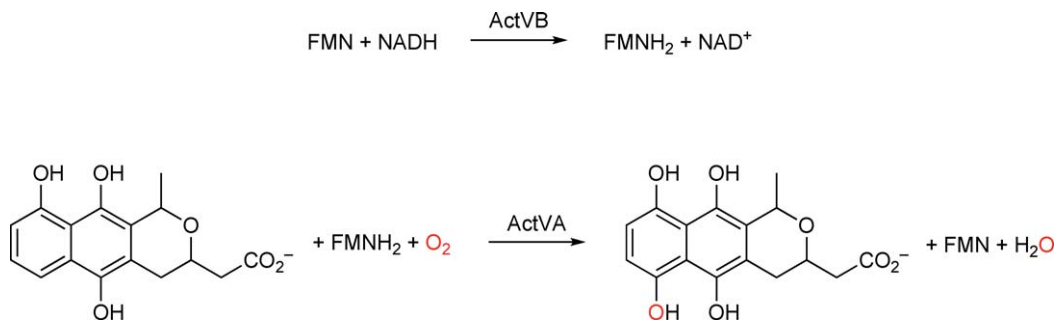
**Scheme 22** Mechanism of 7-chlorotryptophan halogenase.

delivering the halogen to Trp in the rate-determining step of the RebH reaction. In the absence of Trp, the lysyl chloramine has a half-life of 63 h at 4 °C.

### 7.03.6.7 ActVA and ActVB

Many two-component hydroxylase systems, consisting of separate flavin reductase and monooxygenase enzymes, have been discovered in bacteria. Two-component systems are a diverse group; their flavin reductases are often simple proteins but some have a flavin prosthetic group; some use NADH, others, NADPH; some specifically reduce either FMN or FAD, some have no preference; some monooxygenases catalyze electrophilic hydroxylations, while others catalyze nucleophilic hydroxylations.<sup>443</sup> However, they all share the common challenge of transferring reduced flavin from the reductase to the monooxygenase. Data are sparse on whether flavin transfer occurs by diffusion of the flavin through solution from the reductase to the monooxygenase, or by direct transfer between active sites in a transient protein complex.

A particularly well-studied example where the question of flavin transfer has been addressed is the ActVA–ActVB system from *Streptomyces coelicolor*, which catalyzes a step in the biosynthesis of the antibiotic actinorhodin.<sup>462</sup> The substrate of the reaction is the dihydroquinone form of dihydrokalafungin, a trihydroxynaphthalene that is hydroxylated *para* to an activating phenolic oxygen by ActVA, the monooxygenase component (**Scheme 23**). The oxidation of reduced FMN was studied in the absence and presence of dihydrokalafungin. Without the aromatic substrate, reduced flavin binds rapidly to the monooxygenase and reacts with O<sub>2</sub> with a rate constant of  $3 \times 10^4 \text{ mol}^{-1} \text{ s}^{-1}$  at 4 °C. The flavin hydroperoxide formed by this reaction eliminates H<sub>2</sub>O<sub>2</sub> slowly, with a half-life of ~20 min. Identical yields of hydroperoxide were obtained whether reduced flavin was mixed with aerobic solutions of monooxygenase or the preformed reduced flavin–monooxygenase complex was mixed with aerobic buffer, showing that the association of reduced flavin with the monooxygenase is much faster than the nonenzymatic reaction of O<sub>2</sub> with reduced FMN. Dihydrokalafungin has almost no effect on the rate of hydroperoxide formation. This is followed by substrate hydroxylation ( $1.5 \text{ s}^{-1}$ ), elimination of water from the flavin hydroxide ( $0.16 \text{ s}^{-1}$ ), and the nonenzymatic oxidation of the tetrahydroxynaphthalene hydroxylation product to a quinone by O<sub>2</sub>.

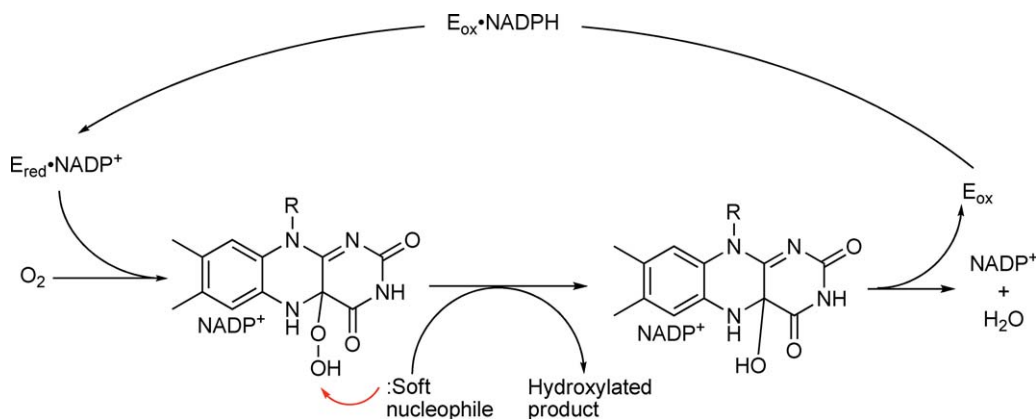


**Scheme 23** Reaction catalyzed by ActVA/ActVB.

When reduced FMN is delivered from the reductase component, ActVB, to the monooxygenase in oxidative half-reactions, hydroperoxide forms at a rate identical to the rate of FMN dissociation from the reductase, measured independently with trapping agents. Thus, there is no kinetic evidence that flavin passes directly from the reductase to the monooxygenase, nor does the poorly competing reaction of free reduced FMN with  $\text{O}_2$  require it. When  $\text{NAD}^+$  is present, it binds tightly to the reductase–reduced FMN complex, forming a charge-transfer complex. When this complex is used as a source of FMN in oxidative half-reactions – a situation most closely mimicking physiology – the charge-transfer complex disappears simultaneously with the formation of the flavin hydroperoxide, and this rate approaches zero as the concentration of  $\text{NAD}^+$  increases, showing that the rate of dissociation of pyridine nucleotide determines the rate of flavin delivery and hydroperoxide formation.

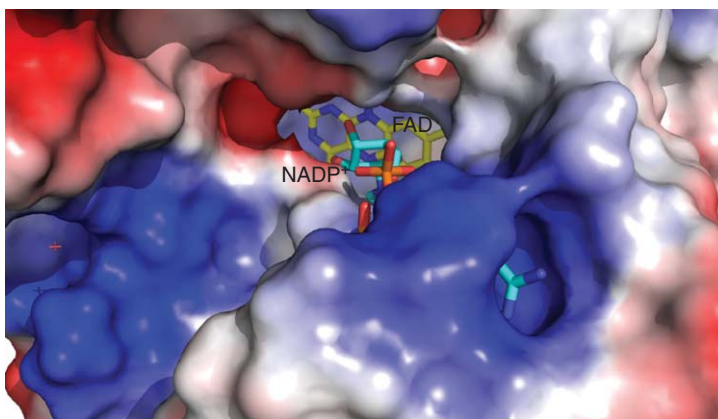
### 7.03.6.8 Flavin-Containing Monooxygenase

An FAD-containing enzyme discovered in liver microsomes, the flavin-containing monooxygenase (FMO), detoxifies many of the xenobiotics encountered in animals and is very important in drug metabolism.<sup>463</sup> FMO homologues are widespread in nature.<sup>464,465</sup> In some microbes, FMOs appear to be involved in nutrient assimilation or protein folding. The pathway for the synthesis of the plant hormone auxin uses FMOs. FMOs use  $\text{NADPH}$  and  $\text{O}_2$  to hydroxylate substrates that are soft nucleophiles (**Scheme 24**). The mammalian enzyme has a very broad substrate tolerance. Hydroxylation occurs on nitrogen, sulfur, selenium, phosphorus, or iodine atoms. In several instances, the initial hydroxylation product is unstable and reacts further independently of the FMO. For instance, thiols are oxidized to sulfenic acids, which spontaneously form disulfides by the reaction with thiols. Nonpolar compounds are better substrates than polar compounds. The hydroperoxide of FMO is not strictly limited to reacting as an electrophile; alkyl boronates are also oxidized at the electrophilic boron atom, presumably by a nucleophilic hydroperoxide.



**Scheme 24** Catalytic cycle of flavin-containing monooxygenases.



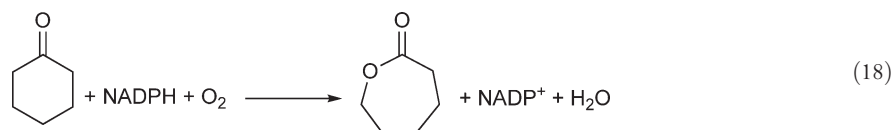


**Figure 8** Active site of flavin-containing monooxygenase with NADP bound blocking access to N5.

FMOs are not homologues of the aromatic hydroxylases – they apparently evolved independently to harness  $O_2$ . Consequently, there are large differences between the catalytic cycles of FMOs and aromatic hydroxylases,<sup>466</sup> but both classes of hydroxylases regulate turnover to minimize possible wasteful and toxic NAD(P)H oxidase activity. The aromatic hydroxylases do this by preventing flavin reduction when the aromatic substrate is absent and then greatly slowing substrate dissociation after flavin reduction. FMOs and related enzymes react rapidly with NADPH regardless of whether the nucleophilic substrate is present.<sup>467</sup> However,  $NADP^+$  remains tightly bound after the reduction reaction and the complex reacts with  $O_2$  (though not extremely rapidly;  $\sim 10^3 \text{ mol}^{-1} \text{ s}^{-1}$ ), forming the hydroperoxide.<sup>468</sup> In the absence of a nucleophilic substrate, this complex is very stable – the half-life for  $H_2O_2$  elimination is  $\sim 2 \text{ h}$  at  $4^\circ\text{C}$ . Early in the study of the mechanism of FMO,  $NADP^+$  was suggested to shield N5 of FAD (Figure 8), slowing  $H_2O_2$  elimination;<sup>469</sup> this prediction was confirmed some two decades later with the crystal structures of two microbial FMOs.<sup>470,471</sup> Substrate hydroxylation forms the C4a-hydroxyflavin, which regenerates the free oxidized enzyme upon  $NADP^+$  dissociation and elimination of water in the rate-determining step of catalysis. Thus, *in vivo*, FMO could be expected to be in either the relatively stable hydroperoxyflavin form, ready for an encounter with a substrate, or the hydroxyflavin form as it recovers from an encounter.

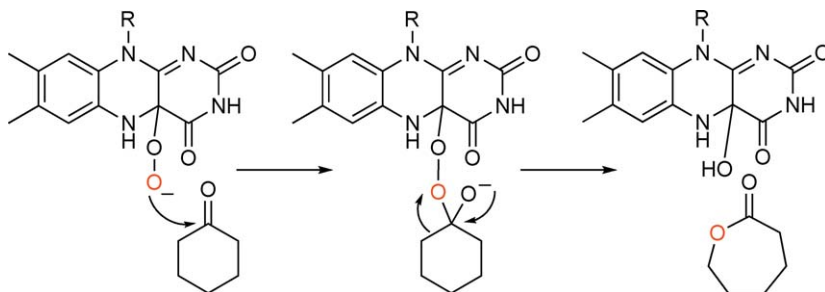
### 7.03.6.9 Baeyer–Villiger Monooxygenases

Flavin hydroperoxides, acting as nucleophiles, can oxidize ketones to esters in Baeyer–Villiger reactions, analogous to peracids in organic synthesis (Equation (18)).



Many FAD-dependent enzymes catalyzing these reactions are found in bacteria and fungi, either in pathways allowing them to metabolize hydrocarbons or in pathways that synthesize secondary metabolites such as antibiotics.<sup>472–474</sup> There has been long-standing interest in using these enzymes for enantioselective chemical syntheses, either as purified enzymes or expressed in engineered microbes for bioreactors.

The mechanisms of a few Baeyer–Villiger monooxygenases have been studied in detail. The catalytic cycles of cyclohexanone monooxygenase from *Acinetobacter* NCIB 9871 and phenylacetone monooxygenase from *Thermobifida fusca* are similar to each other and to that of FMO.<sup>475–477</sup> As with FMO, wasteful and toxic NADPH oxidase activity is minimized by stabilizing the flavin hydroperoxide intermediate rather than inhibiting flavin reduction in the absence of substrate. NADPH reduces FAD in the reductive half-reaction and  $NADP^+$  remains bound, shielding the hydroperoxide. The subsequent reaction with  $O_2$  is much more rapid than that of FMO, with rate constants close to  $10^6 \text{ mol}^{-1} \text{ s}^{-1}$  or higher. Transient kinetic studies strongly



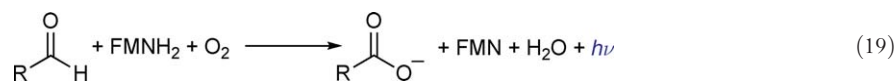
**Scheme 25** Mechanism of cyclohexanone monooxygenase.

suggest that the intermediate observed immediately after the reaction of reduced cyclohexanone monooxygenase with  $O_2$  is the flavin C4a-peroxide,<sup>476</sup> and, in the absence of substrate, the peroxide is protonated to the hydroperoxide over the course of a second at pH 7.2 and 4 °C. The isomerization was observed by a 22 nm shift in the absorbance spectrum whose extent was controlled by a  $pK_a$  of 8.4, consistent with the assignment of the ionizing group as the hydroperoxide. The putative flavin peroxide was significantly more reactive toward cyclohexanone than the putative hydroperoxide, consistent with the role of the flavin peroxide as a nucleophile. The crystal structure of phenylacetone monooxygenase shows that there is a conserved arginine in a position that could stabilize the flavin peroxide anion;<sup>478</sup> mutations of this residue eliminated the ability to form product, although the (hydro)peroxide still formed.

Once formed, the flavin peroxide attacks the ketone of the substrate, forming the so-called Criegee adduct (**Scheme 25**). Alkyl group migration displaces the flavin oxide leaving group, which can eliminate water after protonation to the flavin hydroxide.  $NADP^+$  remains tightly bound throughout the oxidative half-reaction and shields the (hydro)peroxide from solvent. The last step of the half-reaction is the dissociation of the oxidized pyridine nucleotide, which is the slowest step in the catalytic cycle.

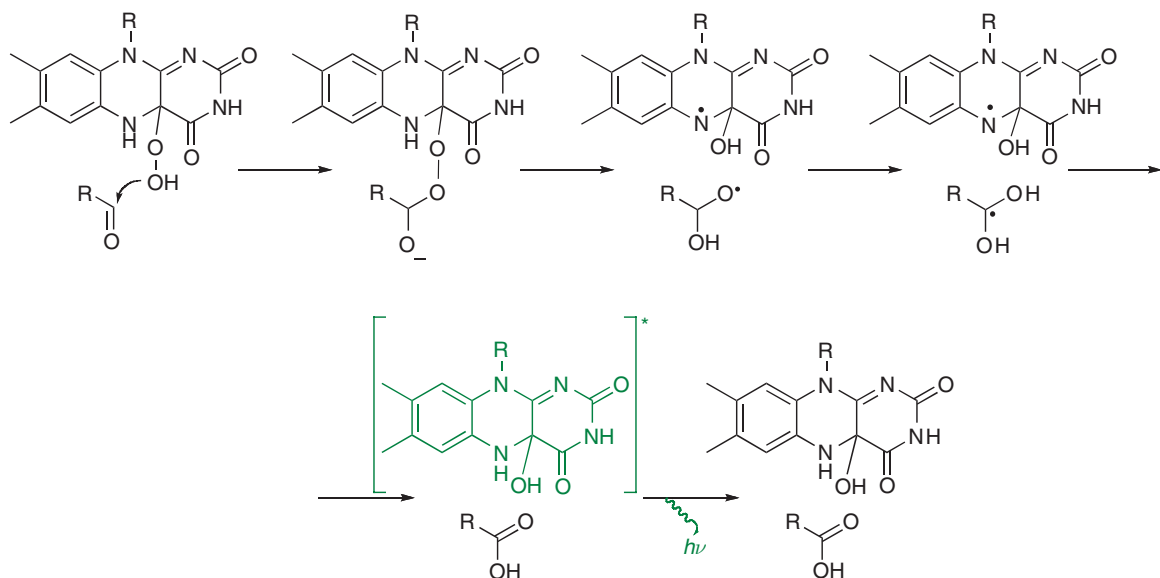
### 7.03.6.10 Bacterial Luciferase

Bioluminescence in the sea is produced by bacterial luciferase, a two-component flavin hydroxylase system.<sup>479</sup> Luciferase oxidizes long-chain aldehydes to the corresponding fatty acids using  $O_2$  and reduced FMN as substrates (Equation (19)).



The biologically important product of the reaction, however, is light. Chromophoric accessory proteins that are expressed with luciferase modulate the color of emitted light. Other proteins including flavin reductases and enzymes that recycle the fatty acid back to the aldehyde are also expressed.<sup>480,481</sup>

Because it is a two-component hydroxylase, reduced flavin is a substrate of luciferase. After it binds, it reacts with  $O_2$ , forming a hydroperoxide, and then aldehyde binds; binding of aldehyde to the free enzyme prevents flavin binding and luminescence. The flavin hydroperoxide bound to luciferase is very stable and may be trapped for many hours by adding long-chain alcohols instead of aldehydes. The hydroperoxide intermediate was prepared in this way using  $^{13}C$ -labeled FMN;<sup>482</sup> its NMR spectrum was critically important in confirming that the product of flavin- $O_2$  reactions is actually the C4a-hydroperoxide. The enzyme-bound hydroperoxide reacts with aldehyde substrates forming an adduct, which ultimately breaks down to fatty acid and hydroxyflavin. The net reaction appears to be similar to a Baeyer-Villiger reaction with hydride derived from the aldehyde as the migratory group. However, the hydroxyflavin produced in this reaction is in the excited state<sup>483</sup> and either emits light directly or transfers energy to an accessory protein – a simple Baeyer-Villiger-like mechanism would not produce the excited state. Instead, a mechanism involving homolytic O–O cleavage followed by electron transfer – termed CIEEL (chemically induced electron exchange luminescence) – has been proposed (**Scheme 26**) and is supported by linear free-energy relationships using flavins substituted at the 8-position as substrates.<sup>484</sup>



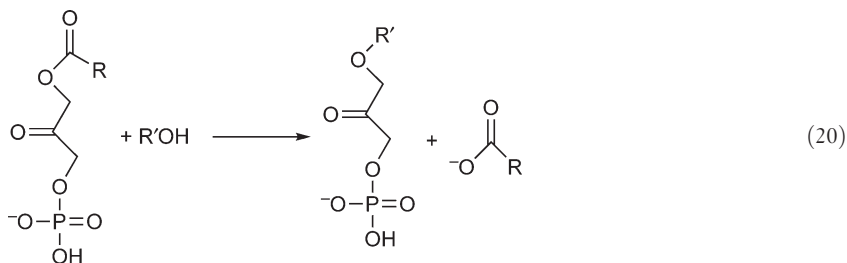
**Scheme 26** CIEEL mechanism proposed for bacterial luciferase.

### 7.03.7 Nonredox Reactions

Some flavoproteins do not catalyze redox transformations.<sup>485</sup> In some, there is no evidence that the flavin plays any role in catalysis, but instead plays a purely structural role. In other cases, it is clear that the flavin is involved in catalysis, but exactly how is often unclear. Since there is no net oxidation or reduction of substrates, the common strategy of studying half-reactions is not generally applicable to these enzymes. Many of these enzymes require reduced flavin, which can oxidize during purification, leading to loss of activity and impeding study of these enzymes. The enzymes discussed below have been grouped together because they all catalyze reactions that have no net redox change – not because they are related or use similar mechanisms. Examples are discussed below that include elimination reactions and photochemical carbon–carbon bond cleavage.

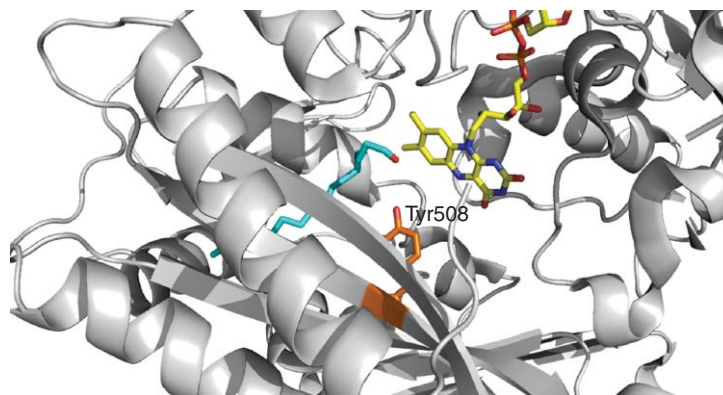
#### 7.03.7.1 Alkyldihydroxyacetonephosphate Synthase

Ether phospholipids are important cell membrane biomolecules involved in signal transduction, neuroplasticity, tumor growth, and membrane fluidity and trafficking.<sup>486</sup> Alkyldihydroxyacetonephosphate synthase (ADPS) is an FAD-containing protein that catalyzes the formation of the ether bond of ether phospholipids.<sup>487–489</sup> Alkyldihydroxyacetonephosphate (alkyl-DHAP) is formed by exchange of the acyl moiety of acyl-DHAP with a fatty alcohol (Equation (20)).



Even though there is no net redox change in the overall reaction, FAD is essential for catalysis.<sup>490</sup>

A truncated ADPS from *Dictyostelium discoideum* was recently crystallized and shown to be a dimer<sup>486</sup> in accordance with gel filtration experiments. The protein contains three domains: an FAD-binding domain, a cap domain, and an N-terminal domain. The structure shows a long tunnel that is covered by a disordered loop at



**Figure 9** Alkylidihydroxyacetonephosphate synthase complexed with a C<sub>16</sub> fatty alcohol (cyan).

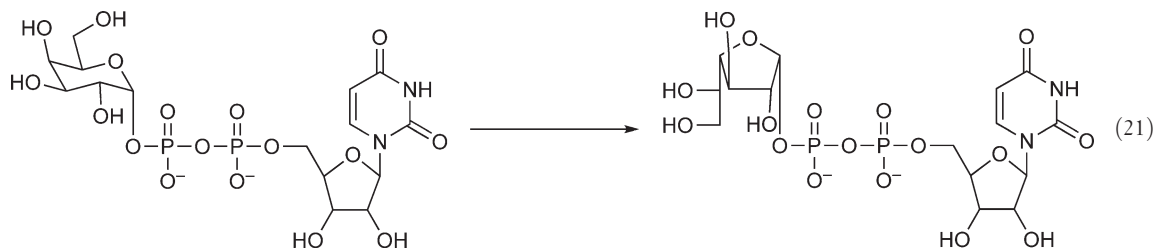
one end. The isoalloxazine of FAD is at this opening with its *si* face exposed to solvent due to the disordered loop (**Figure 9**). Even though no ligand was added during the crystallization procedure, density was observed in the active site consistent with a C<sub>16</sub> fatty alcohol. The fatty alcohol binds in a narrow 18-Å-long tunnel. The disordered loop near the flavin contains three histidine residues and is proposed to shield the active site from solvent when acyl-DHAP is bound due to interactions with the negatively charged phosphate moiety of acyl-DHAP. Access to the other end of the tunnel is controlled by the so-called gating helix.

The enzyme associates with the peroxisomal membrane.<sup>489</sup> Several positive residues create an electropositive patch on the surface of the protein<sup>486</sup> that could bind with the negatively charged heads of membrane lipids. The release of fatty acid product and the entrance of the fatty alcohol substrate is proposed to occur at the protein–membrane interface because residues in the gating helix contribute to the membrane-binding positive patch.

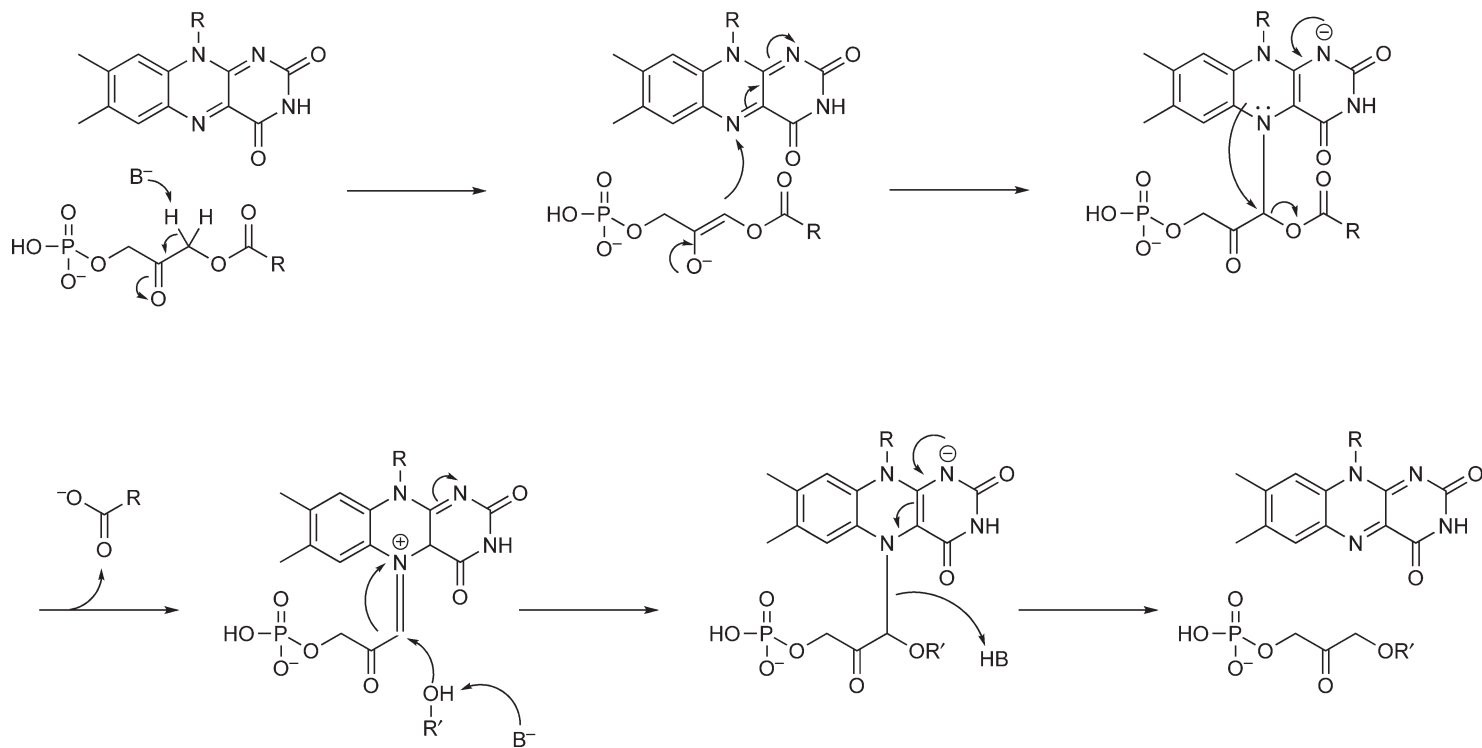
Steady-state analysis<sup>490,491</sup> shows that ADPS uses a ping-pong mechanism with a turnover number of  $\sim 1 \text{ s}^{-1}$ . It is proposed that an active site base deprotonates the C1'-position of acyl-DHAP forming an enolate intermediate (**Scheme 27**). The *pro-R* hydrogen of acyl-DHAP was shown to exchange with solvent.<sup>492,493</sup> A tyrosine, whose importance was demonstrated by site-directed mutagenesis, is positioned correctly to act as the base. The enolate attacks the flavin forming a covalent N5 adduct and the fatty acid product, which is released from the tunnel by movement of the gating helix.<sup>486</sup> This reaction would cause a loss of flavin absorbance, as was observed when the enzyme was mixed with acyl-DHAP.<sup>490</sup> The fatty alcohol then enters the tunnel with its OH group attacking the flavin adduct, forming alkyl-DHAP and oxidized flavin, followed by product release. This mechanism is consistent with the findings that the fatty acid product contains both oxygens of the acyl-DHAP ester bond and the ether oxygen comes from the alcohol.<sup>492,493</sup>

### 7.03.7.2 UDP-Galactopyranose Mutase

Galactofuranose (Gal<sub>f</sub>)-containing oligosaccharides are part of the cell walls of bacteria, but are not found in humans. The first step of Gal<sub>f</sub> incorporation into oligosaccharides is catalyzed by UDP-galactopyranose mutase (UGM). UGM is an FAD-containing enzyme that catalyzes the conversion of UDP-galactopyranose (UDP-Gal<sub>p</sub>) to UDP-Gal<sub>f</sub> (Equation (21)).<sup>494,495</sup>



**Nonredox reactions**



**Scheme 27** Proposed mechanism of alkyldihydroxyacetonephosphate synthase.

Gal $\beta$ -containing oligosaccharides are essential for the survival and infectivity of *Mycobacterium*, making UGM a drug target for the treatment of tuberculosis.<sup>496,497</sup> This isomerization reaction does not involve a net redox change; however, reduced flavin is essential for catalytic activity.<sup>498</sup>

Several structures of enzymes from various species have been solved<sup>498,499</sup> and are nearly identical, with the only major difference being the position of a mobile loop located near the active site. The redox state of the flavin has no significant effect on the structure. The protein can be divided into three domains. Domain I is the FAD-binding domain. The active site is located in a cleft formed between domains I and II. Domain III seals one end of this cleft. The *re* face of FAD is open to the cleft, which is lined with conserved residues. Site-directed mutagenesis of residues in this cleft indicates that it is the site of UDP-Gal $\beta$  binding.<sup>498,500</sup>

The mechanism of UGM is of great interest and is still not completely resolved. Both a single-electron transfer mechanism and a nucleophilic mechanism have been proposed (**Scheme 28**). Both mechanisms involve the formation of a substrate–flavin N5 adduct, which has been trapped during turnover.<sup>501</sup> This could arise due to direct nucleophilic attack of the reduced flavin on the sugar substrate. Alternatively, this adduct could arise from one-electron transfer from the reduced flavin to an oxocarbenium ion generated by elimination of UDP, followed by radical recombination of the flavin semiquinone and hexose radical.<sup>502</sup> An oxocarbenium ion is a proposed intermediate based on positional isotope exchange experiments<sup>503</sup> and studies with substrate analogues,<sup>504</sup> ruling out an S<sub>N</sub>2-like mechanism. However, a radical mechanism or an S<sub>N</sub>1-like mechanism is still possible.

### 7.03.7.3 Phosphopantothenoylcysteine Decarboxylase

Phosphopantothenoylcysteine decarboxylase (PPC-DC) is an essential flavoenzyme in the CoA biosynthetic pathway that catalyzes the decarboxylation of phosphopantothenoylcysteine (PPC).<sup>505,506</sup> In many prokaryotes, PPC-DC is fused with PPC synthetase to form a bifunctional enzyme, while in most eukaryotes it is a monofunctional enzyme.<sup>507</sup>

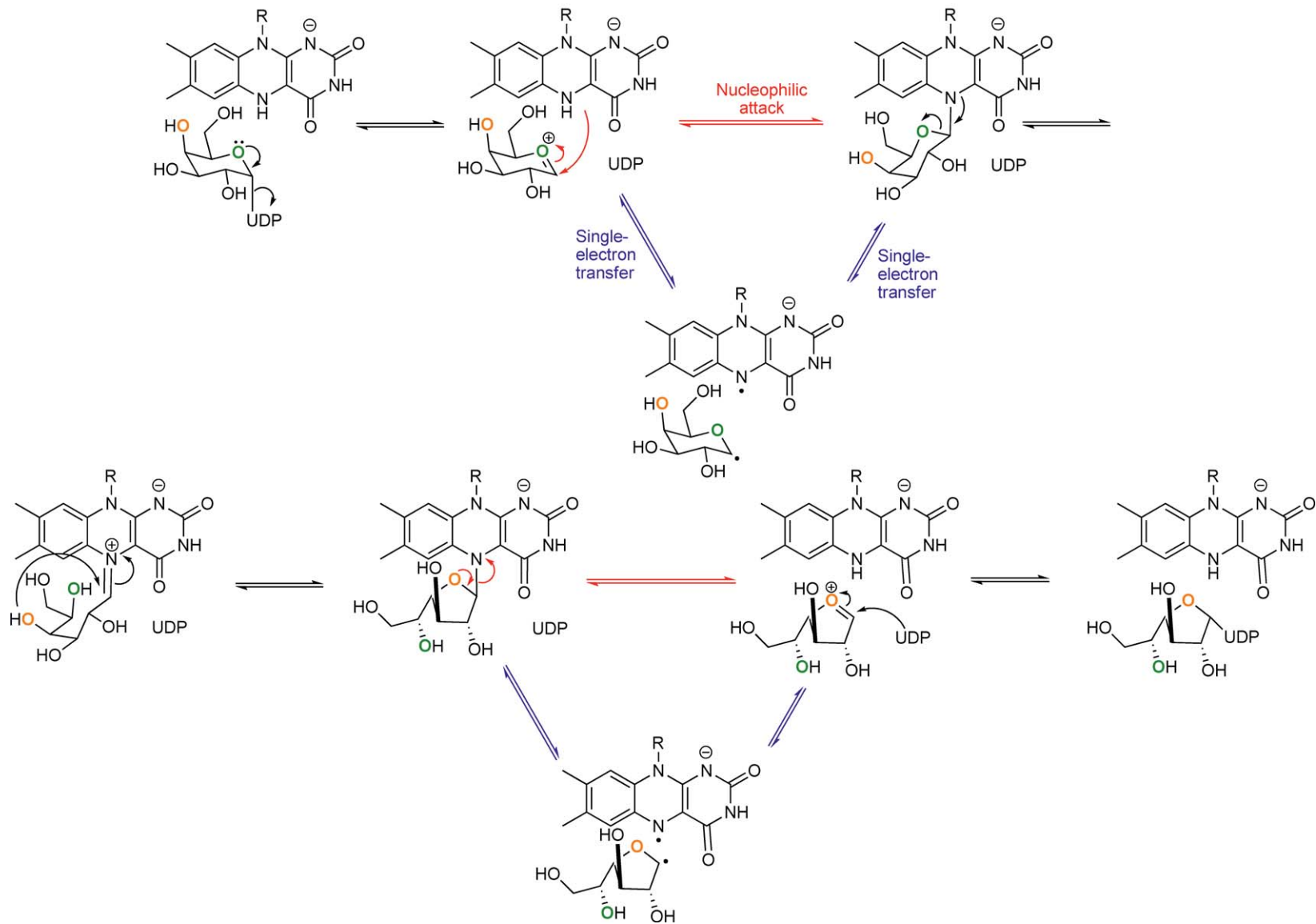
In the proposed mechanism of decarboxylation, the thiol group of PPC is oxidized by FMN to a thioaldehyde via a sulfur–flavin C4a adduct (**Scheme 29**). Flavin–thiolate charge-transfer absorbance is observed when the oxidized enzyme is mixed with the substrate,<sup>508</sup> a common feature of flavoenzymes catalyzing thiol redox chemistry via C4a adducts. The thioaldehyde delocalizes the negative charge that develops during the decarboxylation reaction. The resulting enethiolate is reduced by the reduced flavin forming the decarboxylated product and oxidized flavin. In accordance with the proposed mechanism, retention of stereochemistry at the C $_{\alpha}$ -position was observed and the *pro-R* hydrogen on C $_{\beta}$  of the cysteine moiety was reversibly removed during the reaction, while the hydrogen at C $_{\alpha}$  was not.<sup>509</sup>

A cysteine residue is proposed to be the active site acid needed to protonate the enethiolate intermediate as it is reduced. When the cysteine is mutated to a serine, the enzyme can no longer form the decarboxylated product.<sup>510</sup> The mutant enzyme catalyzes the reductive half-reaction to form the enethiolate intermediate; however, without the active site acid, the enzyme cannot complete the oxidative half-reaction. Soaking a crystal of the mutant with the substrate produced the enethiolate intermediate bound at the active site,<sup>511</sup> further indicating the importance of the cysteine residue in the oxidative half-reaction.

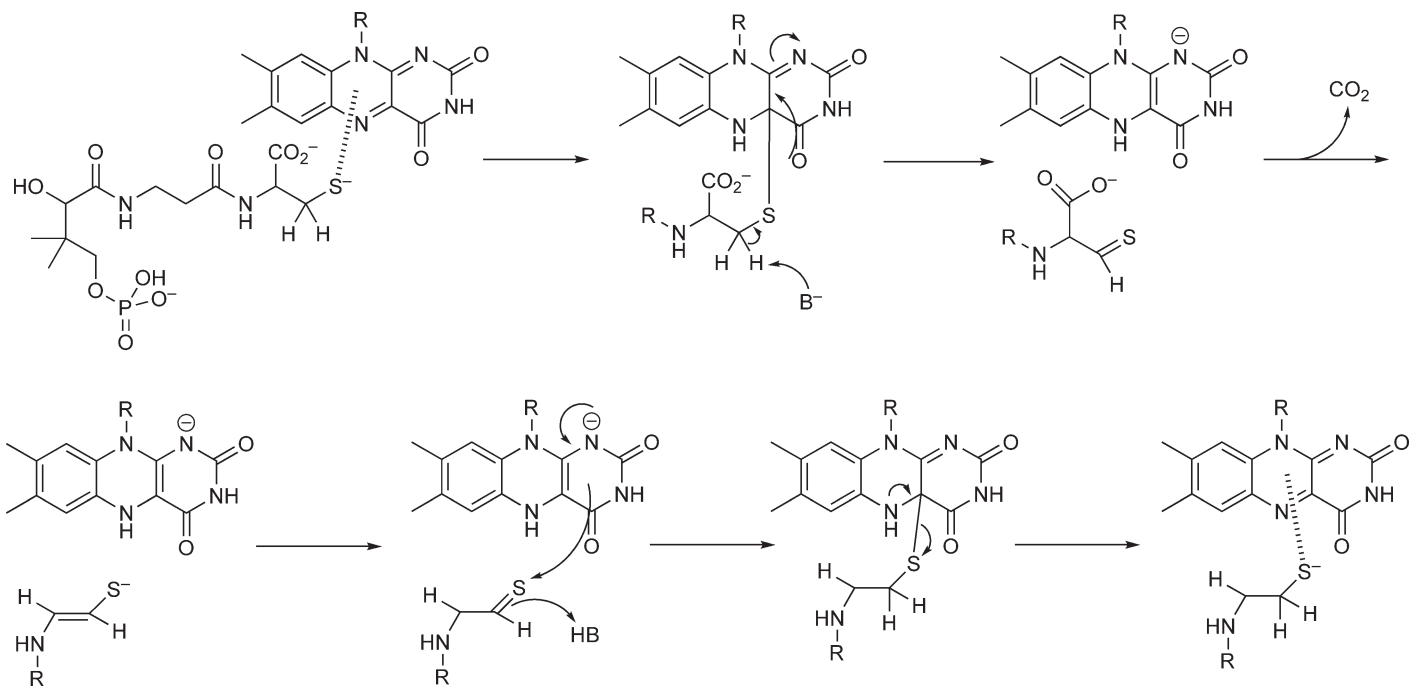
### 7.03.7.4 Chorismate Synthase

Chorismate synthase (CS) catalyzes the formation of chorismate, the last step in the shikimate pathway. Chorismate is a branch-point metabolite used for the synthesis of aromatic amino acids, *p*-aminobenzoic acid, folate, and other cyclic metabolites such as ubiquinone. The shikimate pathway is found only in plants, fungi, and bacteria, making the enzymes of the pathway potential targets for herbicides, antifungals, and antibiotics.

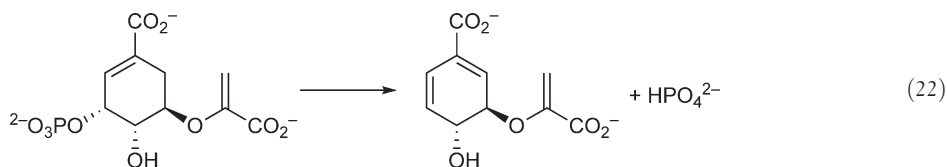




**Scheme 28** Proposed mechanisms of UDP-galactopyranose mutase.



**Scheme 29** Proposed mechanism of 4'-phosphopantothencysteine decarboxylase.

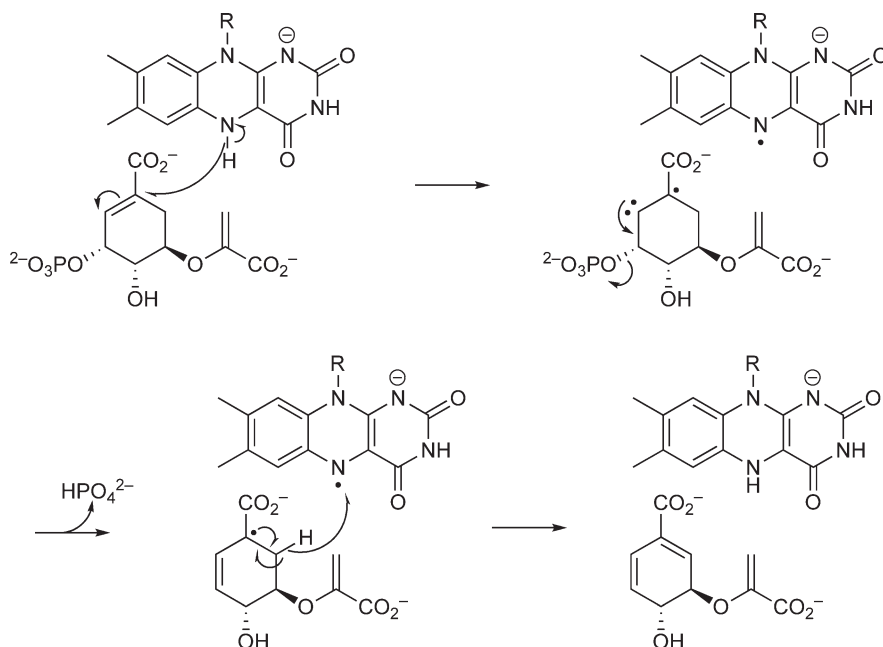


CS catalyzes the 1,4-*anti*-elimination of the 3-phosphate group and the C6 *pro-R* hydrogen from 5-enolpyruvylshikimate-3-phosphate (EPSP) (Equation (22)) in a reaction that requires reduced FMN.<sup>512,513</sup>

The enzymes have been divided into two classes. Bifunctional enzymes have the ability to reduce oxidized FMN via NADPH, while monofunctional enzymes do not. It is unclear how monofunctional enzymes acquire reduced FMN for catalysis. Regardless of the mode of acquiring reduced flavin, all enzymes use similar mechanisms for the elimination reaction.<sup>514</sup> CS has a rather weak affinity for oxidized FMN, which is increased dramatically by the presence of the substrate.<sup>515</sup> The crystal structure shows that the binding of EPSP blocks the exit of FMN, and in addition the structural changes upon substrate binding allow the flavin to make more interactions with the protein,<sup>516</sup> explaining the observed increase in affinity. The structure also explains the ordered binding, flavin followed by EPSP, observed experimentally.<sup>517</sup>

The proposed mechanism of phosphate elimination in CS involves radicals. After binding reduced flavin and EPSP, a single electron is transferred from reduced flavin to the substrate double bond allowing the elimination of phosphate (**Scheme 30**). An active site histidine is proposed to protonate the leaving phosphate. Mutagenesis has shown the importance of this residue for catalysis.<sup>514</sup> The neutral flavin semiquinone then oxidizes the intermediate substrate radical. The structure of the ternary complex with flavin and EPSP shows that N5 of the flavin is directly under C6 of the substrate, implying a direct role of N5 in the reaction.<sup>516</sup>

Radical flavin species have not been detected during turnover of CS; therefore, these species must be short lived. However, a radical flavin species was observed using the EPSP analogue (6*R*)-6-fluoro-EPSP.<sup>518</sup> Phosphate is eliminated from this analogue; however, due to the replacement of the C6 *pro-R* hydrogen with fluorine, the reaction cannot proceed further, resulting in the formation of a substrate-derived radical and neutral flavin semiquinone.



**Scheme 30** Proposed mechanism of chorismate synthase.

### 7.03.7.5 DNA Photolyase

Two main UV lesions that form on DNA are cyclobutane pyrimidine dimers and 6-4 photoadducts. 6-4 photolyase uses reduced flavin to correct 6-4 photoadducts.<sup>519,520</sup> DNA photolyases are flavoenzymes that photochemically reverse pyrimidine dimers formed in DNA by UV irradiation. DNA photolyases are monomeric proteins<sup>521</sup> that have been divided into two classes based on sequence.<sup>522-524</sup> Class 1 enzymes are found in many microbes, while Class 2 enzymes have been found in higher organisms including goldfish<sup>525</sup> and *Drosophila melanogaster*.<sup>522,526</sup> Each monomer of the enzyme contains not only a molecule of FAD, but also a second chromophore that functions as a light-harvesting antenna.<sup>527</sup> This second chromophore is generally either the pterin 5,10-methenyltetrahydropteroylpolyglutamate or the flavin derivative 7,8-dimethyl-8-hydroxy-5-carba-5-deazariboflavin.

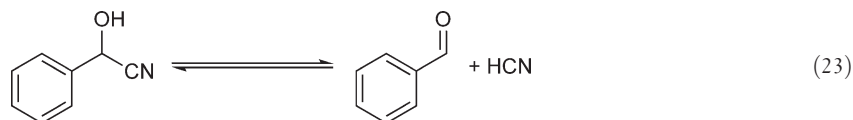
Structures show that FAD is bent in a U-shape with the adenosine moiety near the isoalloxazine ring system. It is deeply buried in the protein with direct access to solvent limited to a cavity leading from the edge of the adenine moiety to the enzyme surface.<sup>528-530</sup> The cavity is thought to be the avenue of approach for the pyrimidine dimer of the substrate. The negatively charged phosphate backbone of DNA is thought to bind a band of positive potential that runs along the surface of the protein around the cavity entrance. The cavity itself is asymmetric, with one side consisting of hydrophobic residues and the other of polar groups, which is thought to aid binding of the asymmetric pyrimidine dimer. The cavity is wide enough to accommodate the dimer only if it is rotated out of the helical DNA structure.

The flavin of DNA photolyase can be in any oxidation state, and the enzyme is often purified as an air-stable neutral semiquinone; however, only the fully reduced form is active.<sup>531-534</sup> The presence of the second chromophore is not required for catalysis.<sup>535,536</sup> Its function is rather to gather light in regions where the reduced flavin absorbs very weakly and transfer the excitation energy to reduced flavin.<sup>535,537-539</sup> The energy transfer between the chromophores is extremely efficient,<sup>540</sup> with rate constants on the order of  $10^{10} \text{ s}^{-1}$ .

The proposed chemical mechanism starts with generation of the excited state of the reduced flavin, followed by transfer of a single electron to the pyrimidine dimer (**Scheme 31**). The ketyl radical formed is presumed to decompose in steps to a pyrimidine-pyrimidine anion radical pair, which transfers a single electron to the flavin semiquinone, regenerating reduced flavin. The repaired DNA is then released.<sup>541</sup>

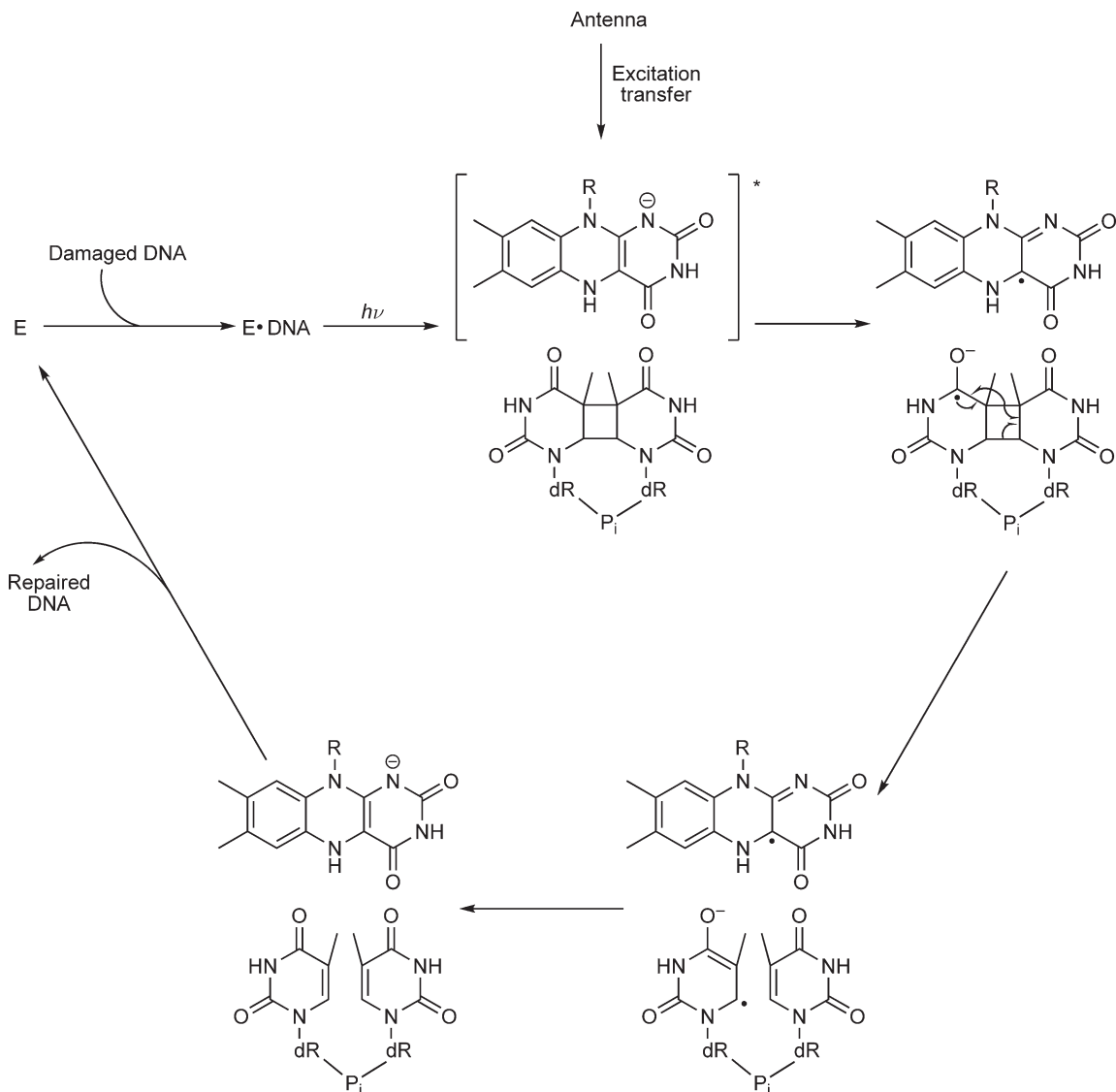
### 7.03.7.6 Oxynitrilase

A number of plants are capable of producing HCN by the breakdown of cyanogenic glycosides. Oxynitrilase catalyzes the release of HCN from (*R*)-mandelonitrile (Equation (23)).



Oxynitrilase isolated from almonds contains one FAD molecule per monomer, while the enzyme isolated from other sources does not contain a flavin.<sup>542</sup> The FAD-containing enzyme from almonds catalyzes both the decomposition of mandelonitrile ( $k_{\text{cat}} = 630 \text{ s}^{-1}$ ,  $25^\circ\text{C}$ ) and the reverse reaction, formation of mandelonitrile ( $k_{\text{cat}} = 1700 \text{ s}^{-1}$ ,  $25^\circ\text{C}$ ).<sup>543</sup>

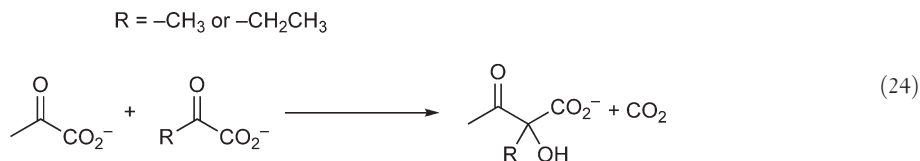
Flavin-dependent oxynitrilase has many properties common to flavoprotein oxidases. The enzyme binds sulfite in an N5 adduct, and one-electron reduction produces anionic semiquinone.<sup>544</sup> The sulfite adduct and one- or two-electron reduced enzyme are inactive, as is apoenzyme, suggesting that the flavin is involved in catalysis. When the flavin is replaced by 5-carba-5-deaza FAD, a low level of activity is retained.<sup>544</sup> However, when the enzyme containing the artificial flavin is exposed to  $\text{H}_2\text{O}_2$ , the isoalloxazine ring system is partially degraded to a redox-inactive heterocyclic system.<sup>545</sup> After the formation of this redox-inactive flavin derivative, enzyme activity actually increases dramatically, suggesting that the flavin plays a structural role. It is possible that the redox state of the flavin serves a regulatory role, since enzyme activity decreases when the natural flavin is reduced; however, this has yet to be proven.



**Scheme 31** Proposed mechanism of DNA repair by DNA photolyase.

### 7.03.7.7 Acetohydroxyacid Synthase

Acetohydroxyacid synthase (AHAS), formerly referred to as acetolactate synthase, is involved in the biosynthesis of branched-chain amino acids in plants and many microorganisms.<sup>546</sup> AHAS catalyzes the condensation of two molecules of pyruvate to form acetolactate and  $\text{CO}_2$ , or the condensation of one molecule of pyruvate with one molecule of  $\alpha$ -ketobutyrate to form  $\alpha$ -aceto- $\alpha$ -hydroxybutyrate and  $\text{CO}_2$  (Equation (24)).



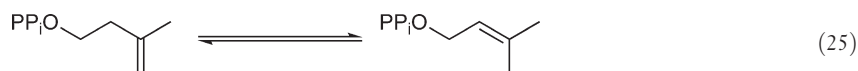
The enzyme has two subunits: a catalytic subunit and a regulatory subunit. There are three prosthetic groups, thiamine diphosphate, a divalent metal ion, and FAD, present in the catalytic subunit. The regulatory subunit

greatly enhances the reactivity of the catalytic subunit and is necessary for feedback inhibition by branched-chain amino acids.

FAD is necessary for enzyme activity and structures show that it is in the active site; however, it is not believed to be involved in catalysis.<sup>546</sup> When the flavin was replaced with several analogues, modest effects on activity were observed at most,<sup>547</sup> suggesting that the flavin plays a structural role. Interestingly, AHAS shares significant sequence homology with pyruvate oxidase, an enzyme that uses its flavin for catalysis. It is thought that the flavin in AHAS is an evolutionary relic left from pyruvate decarboxylase.<sup>548</sup> Therefore, the mechanism proposed for AHAS does not involve FAD. Instead, a thiamine–pyruvate adduct is formed and decarboxylated, followed by condensation with the second  $\alpha$ -keto acid substrate. The flavin has been shown to protect against alkylation of a cysteine residue by bromopyruvate, leading to the idea that FAD shields the anionic substrate–thiamine adduct from solvent during catalysis.<sup>547</sup> A vestigial FAD is also present in the thiamine-dependent enzyme YerE, which is involved in the biosynthesis of yersiniose.<sup>549</sup>

### 7.03.7.8 Isopentenyl Diphosphate Isomerase

Isopentenyl diphosphate (IPP) isomerase catalyzes the conversion of IPP to dimethylallyl diphosphate (DMAPP), an early step in isoprenoid metabolism (Equation (25)).



DMAPP is a precursor of many isoprenoid compounds including carotenoids, sterols, and ubiquinones.<sup>550</sup> IPP isomerase is an essential enzyme in organisms that use the mevalonate pathway to synthesize isoprenoid units,<sup>551</sup> making the enzymes from *S. pneumoniae* and *S. aureus* interesting drug targets.

Two isozymes of IPP isomerase have been identified. Type I IPP isomerase, an enzyme containing two divalent metal ions ( $\text{Zn}^{2+}$  and  $\text{Mg}^{2+}$ ), has been studied for decades.<sup>552,553</sup> More recently, type II IPP isomerases have been identified.<sup>554</sup> These enzymes require  $\text{Mg}^{2+}$  and reduced FMN to be active,<sup>554–556</sup> even though the overall reaction has no redox change.

The mechanism of type II IPP isomerase is currently not known. Two mechanisms have been proposed. Epoxy and fluorinated substrate analogues have been used to probe the mechanism of type II IPP isomerase. These compounds were irreversible inhibitors of the enzyme forming N5 covalent adducts with the reduced flavin.<sup>557–559</sup> These studies were taken as evidence of a protonation/deprotonation mechanism in which protonation of the double bond of IPP is followed by deprotonation of the carbocation intermediate to form DMAPP, similar to the mechanism of type I IPP isomerase. Conversely, a radical mechanism has been proposed. Type II IPP isomerase stabilizes neutral semiquinone during reduction.<sup>555,556,560</sup> In addition, the enzyme is inactive when the natural flavin is replaced by 5-carba-5-deaza flavin, a flavin analogue incapable of single-electron transfers.<sup>555,560</sup> More work is required to determine the mechanism of these enzymes and the role of the flavin in catalysis.

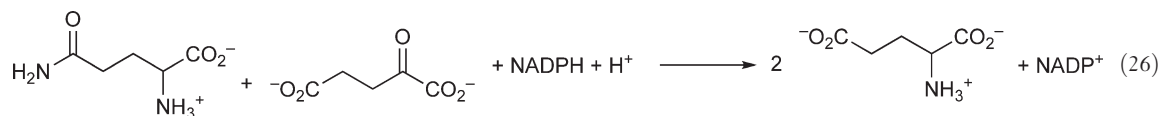
## 7.03.8 Complex Flavoenzymes

As defined in this review, complex flavoproteins catalyze more than one type of reaction used to classify enzymes. Some of these enzymes use a single active site to catalyze different chemical conversions, whereas some contain multiple active sites and other prosthetic groups, such as iron–sulfur clusters. In any case, the study of these enzymes is often more complicated than the simple examples discussed earlier.

### 7.03.8.1 Glutamate Synthase

Glutamate synthase (GltS) catalyzes the conversion of 2-oxoglutarate into L-glutamate with L-glutamine serving as the nitrogen source for the reaction (Equation (26)).





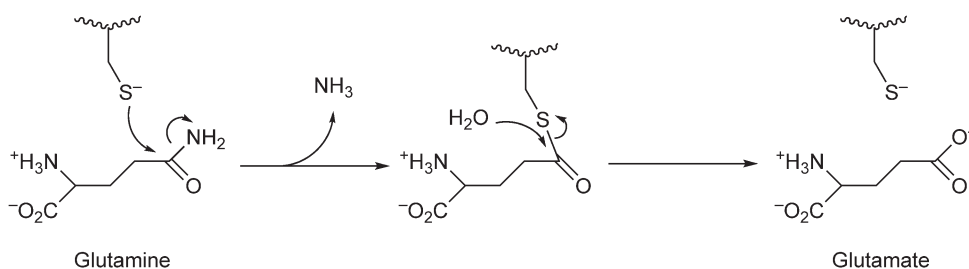
The reducing equivalents required for the reaction originate from either a pyridine nucleotide or ferredoxin. GltS along with glutamine synthetase forms a nitrogen assimilation pathway of great importance in plants and microorganisms.<sup>561–563</sup> This pathway produces glutamine and glutamate, which serve as nitrogen donors for the biosynthesis of many compounds including other amino acids, nucleotides, chlorophylls, polyamines, and alkaloids.<sup>564,565</sup>

GltSs have been divided into three classes based on sequences and biochemical properties:<sup>566</sup> ferredoxin-dependent GltS (Fd-GltS), NADPH-dependent GltS (NADPH-GltS), and NADH-dependent GltS (NADH-GltS). All GltSs are iron–sulfur flavoproteins. Fd-GltS is found in higher plants, cyanobacteria, and algae, and is composed of a single polypeptide chain of ~165 kDa that contains FMN and an iron–sulfur cluster. The enzyme acquires its reducing equivalents from ferredoxin. NADPH-GltS is found mostly in bacteria. This enzyme is a dimer composed of an  $\alpha$  and a  $\beta$  subunit and contains FMN, FAD, and three iron–sulfur clusters. The large  $\alpha$  subunit (~160 kDa) catalyzes the reductive synthesis of glutamate. The smaller  $\beta$  subunit (~52 kDa) is an FAD-dependent NADPH oxidoreductase that transfers reducing equivalents from NADPH to FMN on the  $\alpha$  subunit. NADH-GltS is not well studied. These enzymes are found in fungi, lower animals, and nongreen tissues of plants like seeds and roots. NADH-GltS is composed of a single polypeptide chain of ~200 kDa that is proposed to have formed from gene fusion of the  $\alpha$  and  $\beta$  subunits of NADPH-GltS; however, NADH-GltS is highly specific for NADH.<sup>567</sup>

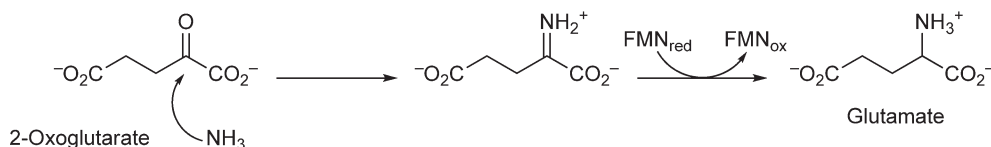
The mechanism of glutamate synthesis is thought to be the same in all GltSs, with the only difference being the mechanism of reducing FMN. The enzyme must first catalyze the hydrolysis of glutamine forming glutamate and ammonia. The ammonia produced by this reaction attacks 2-oxoglutarate and is reduced to form a second molecule of glutamate. These reactions occur in separate domains of the protein, with the two active sites over 30 Å apart,<sup>566–568</sup> a common feature of amidotransferases. The hydrolysis of glutamine occurs in the amidotransferase domain when an active site cysteine attacks the amide of glutamine, forming a  $\gamma$ -glutamyl thioester intermediate (**Scheme 32(a)**). The thioester is then hydrolyzed, resulting in glutamate and the free cysteine. The tetrahedral intermediates that form during acylation and hydrolysis of the thioester

### Complex reaction

#### (a) Glutaminase reaction



#### (b) Synthase reaction



**Scheme 32** Two reactions catalyzed by glutamate synthase.

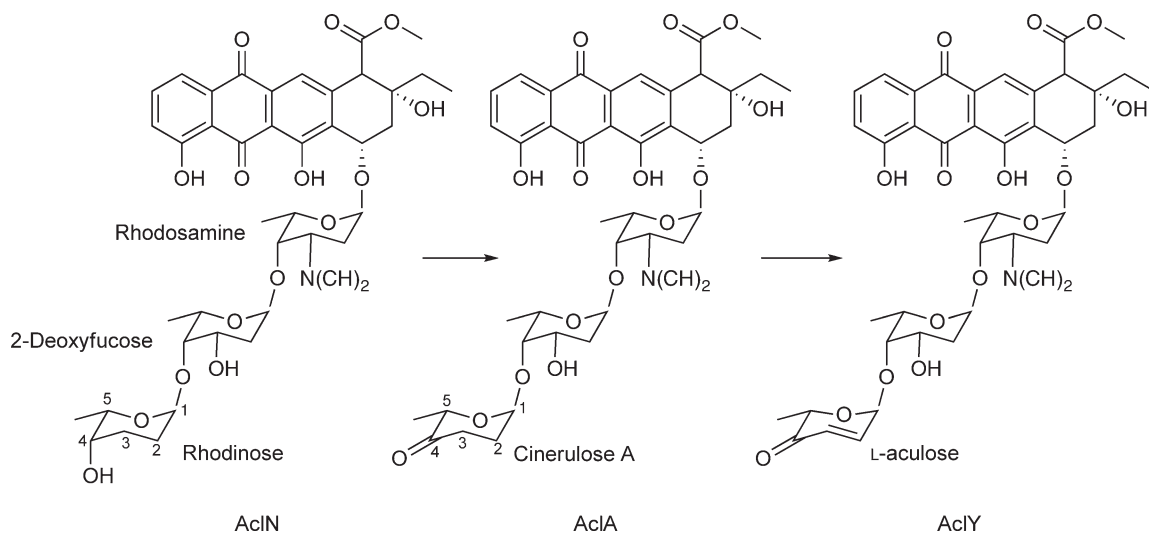
intermediate are stabilized by an oxyanion hole formed by the side-chain nitrogen of an asparagine and a backbone nitrogen. The ammonia formed then travels through a tunnel to the active site of the FMN-binding domain and attacks C2 of 2-oxoglutarate forming 2-iminoglutarate (**Scheme 32(b)**) in a reaction proposed to be assisted by polarization of the carbonyl group by the protein and precise delivery of the ammonia at the exit of the tunnel. Direct hydride transfer from N5 of reduced flavin to C2 of 2-iminoglutarate reduces the intermediate to form glutamate. Rapid reduction of the intermediate by the reduced flavin has been proposed to drive the reaction toward glutamate formation.<sup>568</sup>

The NADPH-GltS from *Azospirillum brasilense* has a turnover number of  $\sim 60 \text{ s}^{-1}$  at pH 8.0 and 25 °C using NADPH, 2-oxoglutarate, and glutamine.<sup>568</sup> Ammonia can replace glutamine as the nitrogen donor in NADPH-GltS, but the ammonia-dependent reaction is much slower than the physiological reaction, though the rate increases with pH.<sup>569</sup> The enzyme binds glutamine in the absence of 2-oxoglutarate and NADPH;<sup>569</sup> however, glutamine hydrolysis is observed only when all substrates are present in the NADPH-GltS from *A. brasilense*.<sup>570</sup> During turnover with all three substrates present, all the ammonia produced from glutamine hydrolysis is added to 2-oxoglutarate, with none being released into the solvent.<sup>570</sup> Using <sup>14</sup>C-labeled glutamine and 2-oxoglutarate, no glutaminase activity (hydrolysis of glutamine to produce glutamate and ammonia) uncoupled from glutamate synthesis was observed in the holoenzyme. However, the isolated  $\alpha$  subunit does exhibit glutaminase activity.<sup>571</sup> Thus, the two reactions in separate domains of the protein are tightly coupled. It should be noted that glutaminase activity has been reported in other bacterial GltSs.<sup>572</sup>

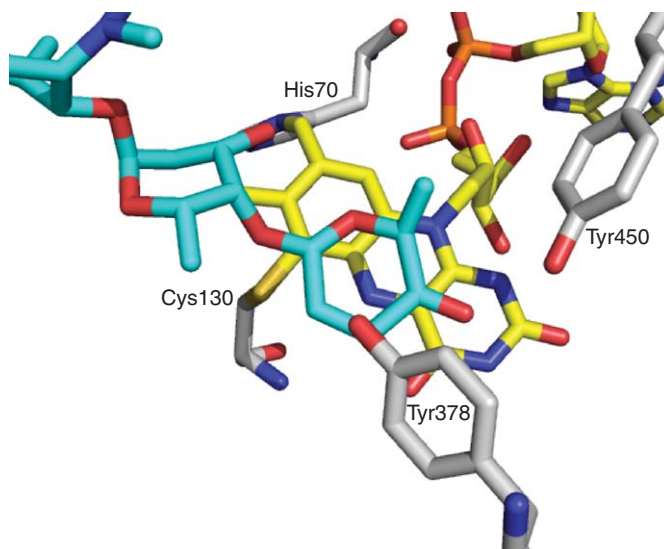
NADPH-GltS and Fd-GltS also catalyze the reverse reaction – glutamate oxidation forming ammonia and 2-oxoglutarate.<sup>571,573</sup> This reaction is sufficiently slow to be observed using a diode-array spectrophotometer under anaerobic conditions.<sup>573</sup> Reduction of the flavin is observed following addition of glutamate, indicating that glutamate binds to the synthase site of the enzyme and that FMN is directly involved in the interconversion of glutamate and 2-iminoglutarate.

### 7.03.8.2 Aclacinomycin Oxidoreductase

Aclacinomycins (Acls) are aromatic polyketides produced by several strains of *Streptomyces* and are potential chemotherapeutic agents.<sup>574,575</sup> Acls are composed of a trisaccharide moiety attached to aklavinone. The first sugar residue of the trisaccharide moiety is rhodosamine, the second is 2-deoxyfucose, while the third varies depending on the specific Acl. In AclN, the third sugar residue is rhodinosose, while it is cinerulose A in AclA and L-aculose in AclY.<sup>576</sup> Aclacinomycin oxidoreductase (AknOx) is an FAD-containing enzyme that catalyzes two chemical conversions of the terminal sugar of Acls<sup>577</sup> (Equation (27)).



(27)



**Figure 10** Active site of acIcinomycin oxidoreductase with the product, AcIY (cyan).

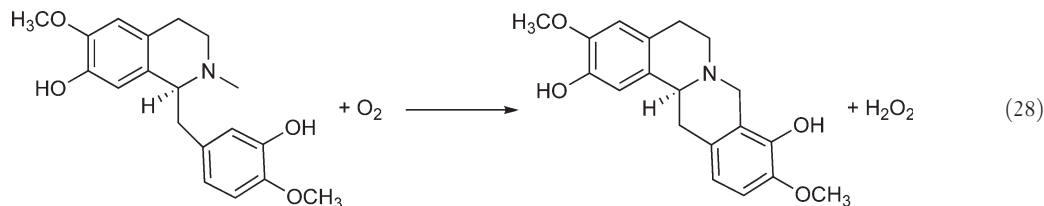
First, AcIN is converted to AcIA and then to AcIY.

The structure of the AcIY–enzyme complex shows that the enzyme contains a single active site.<sup>578</sup> The enzyme can be divided into two domains: an FAD-binding domain and a substrate-binding domain. The product is bound in a pocket extending from the surface to deep inside the enzyme, with the terminal sugar moiety near the isoalloxazine ring system of the flavin (Figure 10). Interestingly, the structure shows that the flavin is bicovalently bound to the enzyme by a histidine at the 8- $\alpha$ -position of FAD and by a cysteine at the 6-position of FAD. Sequence alignment of AknOx homologues shows that the histidine and cysteine residues are conserved suggesting that bicovalent flavinylation is a common feature of these enzymes.

AknOx catalyzes two consecutive steps in the biosynthesis of AcIY: the conversion of the terminal sugar residue of AcIN to cinerulose A and the desaturation of cinerulose A to L-aculose. Both these reactions occur in the same active site. AcIN, the first substrate of the enzyme, was modeled into the active site based on the structure of the AcIY–enzyme complex.<sup>578</sup> Based on this model, it is proposed that an active site tyrosine (Tyr450) deprotonates the C4 hydroxyl of rhodinose and the C4 proton is transferred as a hydride to N5 of the flavin resulting in AcIA, in which the terminal sugar residue is cinerulose A. Mutation of the tyrosine to phenylalanine results in a severely impaired enzyme, indicating the importance of this residue. The reduced flavin is reoxidized by molecular oxygen. Another active site tyrosine (Tyr378) is proposed to deprotonate C3 of cinerulose A. Hydride transfer from C2 to the flavin results in the formation of a double bond between C2 and C3 of the sugar giving the product AcIY. Mutation of the second proposed active base results in complete loss of activity in the second reaction, although AcIN is still converted to AcIA.<sup>578</sup> Again, the reduced flavin is oxidized by molecular oxygen. Thus, AknOx catalyzes both C–O and C–C oxidation in the same active site using different active site residues.

### 7.03.8.3 Berberine Bridge Enzyme

Alkaloids, such as benzophenanthridines, are potentially useful as pharmaceuticals.<sup>579</sup> Because of this, there has been much interest in understanding the biosynthesis of these compounds. A common metabolite, (*S*)-reticuline, is involved in the biosynthesis of many alkaloids. Berberine bridge enzyme (BBE) catalyzes the conversion of (*S*)-reticuline to (*S*)-scoulerine in an FAD-dependent reaction (Equation (28)).



The reaction involves the oxidation of the *N*-methyl group of (*S*)-reticulin. However, the hypothetical methylene iminium ion is not formed;<sup>580,581</sup> the enzyme catalyzes the concerted intramolecular attack of the phenolic moiety of the substrate on the *N*-methyl group as it is oxidized, forming a C–C bond.<sup>582</sup> The reduced flavin is oxidized by molecular oxygen.

The flavin is bicovalently bound to BBE at the 8- $\alpha$ -position by a histidine and at the 6-position by a cysteine. BBE has some common properties of flavoprotein oxidases.<sup>583</sup> BBE stabilizes anionic semiquinone and reacts with sulfite, forming an N5 adduct. These properties indicate a positive charge in the active site that can stabilize negative charge at the N1–C2=O locus of the flavin.

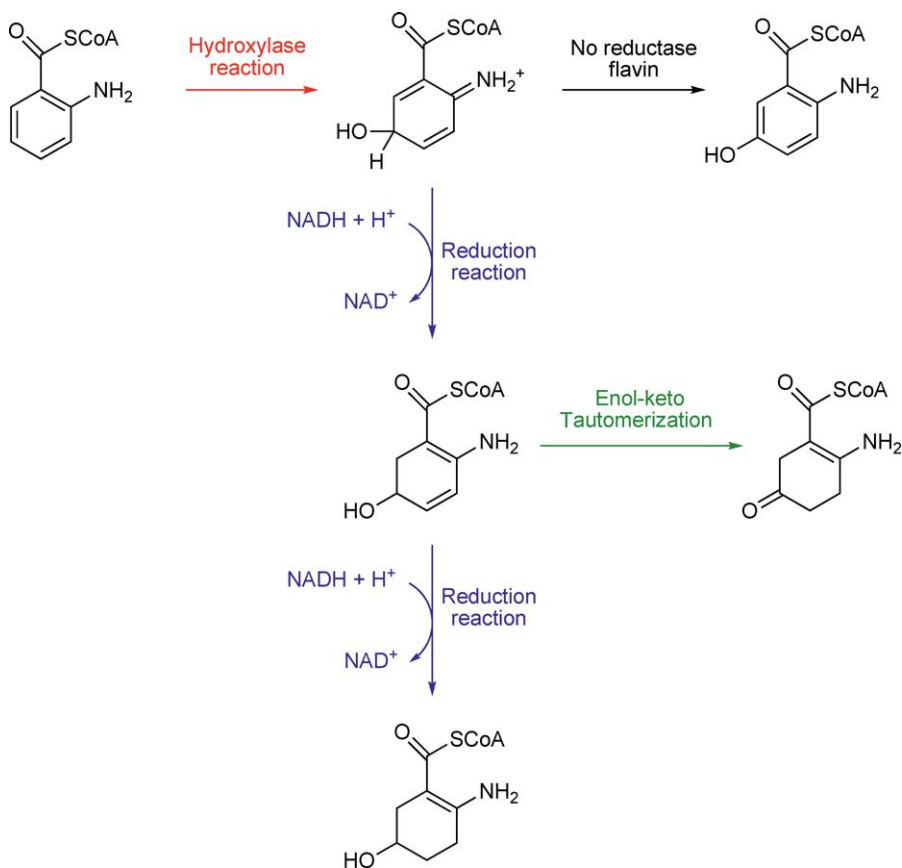
BBE has a turnover number of  $\sim 8 \text{ s}^{-1}$  at pH 9 and 37 °C.<sup>583</sup> Under these conditions, flavin reduction<sup>584</sup> occurs at  $\sim 100 \text{ s}^{-1}$ , and turnover is limited by the reaction of reduced enzyme with molecular oxygen. When the cysteine responsible for flavinylation at the 6-position is mutated, the turnover number is decreased to  $0.48 \text{ s}^{-1}$  under identical conditions<sup>584</sup> because flavin reduction is rate limiting. The redox potential is greatly affected by flavinylation at the 6-position as well. The redox potential of the mutant enzyme is +53 mV, compared to +132 mV in the wild-type enzyme,<sup>584</sup> indicating that flavinylation at the 6-position is responsible for increasing the redox potential. This increase in the redox potential is responsible for the differences in the rate of flavin reduction observed.<sup>584</sup>

#### 7.03.8.4 2-Aminobenzyl-CoA Monooxygenase/Reductase

2-Aminobenzyl-CoA monooxygenase/reductase (AbzCoA-M/R) is a flavoprotein involved in anthranilate degradation.<sup>585</sup> The enzyme is a homodimer with one FAD per monomer, but apparently only one active site per dimer.<sup>586</sup> The enzyme catalyzes both a hydroxylation and a reduction using 2-aminobenzyl-CoA, NADH, and molecular oxygen as substrates, with one flavin being responsible for each reaction. The enzyme first catalyzes the hydroxylation of the aromatic ring of 2-aminobenzyl-CoA *para* to the activating amine. The hydroxylation product is then reduced by the second flavin. The reduction product can be further reduced<sup>587</sup> complicating analysis.

The two flavins present in AbzCoA-M/R are not equivalent. The flavin responsible for the reductase activity is bound weakly compared to the other and can dissociate during purification.<sup>586</sup> The flavins also differ in their reactivity. In the absence of 2-aminobenzyl-CoA, only one flavin reduces rapidly ( $\sim 70 \text{ s}^{-1}$ ) when the enzyme is mixed with NADH. However, when CoA substrate is present, both flavins reduce rapidly ( $\sim 70$  and  $\sim 2 \text{ s}^{-1}$ ), indicating that the presence of substrate stimulates reduction of the hydroxylase flavin. In addition, the reduced hydroxylase flavin reacts quickly with molecular oxygen, while the reductase flavin does not.<sup>588</sup> While the two flavins are quite different kinetically, they are indistinguishable thermodynamically. Only a single flavin redox potential can be measured for the enzyme, indicating that the flavins have identical potentials. The redox potential of the free enzyme is  $-219 \text{ mV}$ , while in the presence of 2-aminobenzyl-CoA it is  $-202 \text{ mV}$ .<sup>589</sup>

The reduced enzyme–substrate complex reacts with molecular oxygen forming a flavin hydroperoxide, which hydroxylates 2-aminobenzyl-CoA at the 5-position at  $80 \text{ s}^{-1}$  (Scheme 33).<sup>588</sup> In the absence of the reductase flavin, an aromatic product is obtained. However, in its presence, the hydroxylated intermediate is reduced rapidly and undergoes tautomerization to give the product 2-amino-5-oxocyclohex-1-enecarboxyl-CoA.



**Scheme 33** Reactions of 2-aminoacyl-CoA monoxygenase/reductase.

### 7.03.9 Summary

The few examples considered in this chapter demonstrate the wide range of chemistry available to flavins, and the large number of enzyme catalysts possible by the combination of the various half-reactions. The enzymes in nature have evolved to accelerate the chemistry of their half-reactions and suppress unwanted side reactions, enabling flavoenzymes to participate in most areas of biochemistry.

#### Abbreviations

<b>2-PCPA</b>	<i>trans</i> -2-phenylcyclopropylamine
<b>5FU</b>	5-fluorouracil
<b>ABzCoA-M/R</b>	2-aminobenzyl-CoA monoxygenase/reductase
<b>ACAD-9</b>	acyl-CoA dehydrogenase 9
<b>ACADs</b>	acyl-CoA dehydrogenases
<b>Acis</b>	aclacinomycins
<b>ACOs</b>	acyl-CoA oxidases
<b>ADPS</b>	alkyldihydroxyacetonephosphate synthase
<b>AHAS</b>	acetohydroxyacid synthase
<b>AknOx</b>	aclacinomycin oxidoreductase
<b>alkyl-DHAP</b>	alkyldihydroxyacetonephosphate
<b>BBE</b>	berberine bridge enzyme

<b>CPR</b>	cytochrome P-450 reductase
<b>CS</b>	chorismate synthase
<b>DAAO</b>	D-amino acid oxidase
<b>DHO</b>	dihydroorotate
<b>DHOD</b>	dihydroorotate dehydrogenase
<b>DMAPP</b>	dimethylallyl diphosphate
<b>DPD</b>	dihydropyrimidine dehydrogenase
<b>DUS</b>	dihydrouridine synthase
<b>enolpyruvyl-UDP-GlcNAc</b>	enolpyruvyl-UDP- <i>N</i> -acetylglucosamine
<b>EPR</b>	electron paramagnetic resonance
<b>EPSP</b>	5-enolpyruvylshikimate-3-phosphate
<b>ETF</b>	electron-transferring flavoprotein
<b>FAD</b>	flavin adenine dinucleotide
<b>Fb<sub>2</sub></b>	flavocytochrome <i>b</i> <sub>2</sub>
<b>Fcc<sub>3</sub></b>	flavocytochrome <i>c</i> <sub>3</sub>
<b>Fd-GltS</b>	ferredoxin-dependent GltS
<b>FMN</b>	flavin mononucleotide
<b>FMO</b>	flavin-containing monooxygenase
<b>GltS</b>	glutamate synthase
<b>GOX</b>	glucose oxidase
<b>GR</b>	glutathione reductase
<b>GSH</b>	glutathione
<b>GSSG</b>	glutathione disulfide
<b>high <i>M<sub>r</sub></i>, TxrR</b>	high molecular weight TxrR
<b>IPP</b>	isopentenyl diphosphate
<b>K<sub>d</sub></b>	dissociation constant
<b>LCAD</b>	long-chain ACAD
<b>LMO</b>	L-lactate monooxygenase
<b>low <i>M<sub>r</sub></i>, TxrR</b>	low molecular weight TxrR
<b>LOX</b>	L-lactate oxidase
<b>LSD1</b>	lysine-specific demethylase 1
<b>M</b>	molar
<b>MAO</b>	monoamine oxidase
<b>MCAD</b>	medium-chain ACAD
<b>MerA</b>	mercuric reductase
<b>MICAL</b>	molecule interacting with CasL
<b>min</b>	minute
<b>MSOX</b>	monomeric sarcosine oxidase
<b>MurB</b>	UDP- <i>N</i> -acetylenolpyruvylglucosamine reductase
<b>NADH-GltS</b>	NADH-dependent GltS
<b>NADPH-GltS</b>	NADPH-dependent GltS
<b>OA</b>	orotate
<b>OYE</b>	old yellow enzyme
<b>P5C</b>	Δ <sup>1</sup> -pyrroline-5-carboxylate
<b>P5CDH</b>	Δ <sup>1</sup> -pyrroline-5-carboxylate dehydrogenase
<b>PCLase</b>	prenylcysteine lyase
<b>PDI</b>	protein disulfide isomerase
<b>PDO</b>	phthalate dioxygenase
<b>PDR</b>	phthalate dioxygenase reductase
<b>PHBH</b>	<i>p</i> -hydroxybenzoate hydroxylase
<b>pkDAAO</b>	pig kidney DAAO



<b>pOHB</b>	<i>p</i> -hydroxybenzoate
<b>PPC</b>	phosphopantothencycysteine
<b>PPC-DC</b>	phosphopantothencycysteine decarboxylase
<b>PRODH</b>	proline dehydrogenase
<b>PutA</b>	proline utilization A
<b>QFRs</b>	quinol:fumarate reductases
<b>QSOX</b>	quiescin-sulfhydryl oxidases
<b>RgDAAO</b>	<i>Rhodotorula gracilis</i> DAAO
<b>SCAD</b>	short-chain ACAD
<b>SQORs</b>	succinate:quinone oxidoreductases
<b>SQRs</b>	succinate:quinone reductases
<b>TrxR</b>	thioredoxin reductase
<b>TvDAAO</b>	<i>Trigonopsis variabilis</i> DAAO
<b>UDP-Galf</b>	UDP-galactofuranose
<b>UDP-Galp</b>	UDP-galactopyranose
<b>UDP-MurNAc</b>	UDP- <i>N</i> -acetylmuramic acid
<b>UGM</b>	UDP-galactopyranose mutase
<b>VLCAD</b>	very long-chain ACAD

## References

1. B. A. Palfey; V. Massey, Flavin-Dependent Enzymes. In *Comprehensive Biological Catalysis, Volume III/Radical Reactions and Oxidation/Reduction*; M. Sinnott, Ed.; Academic Press: London and San Diego, 1998; Chapter 29, pp 83–154.
2. G. A. Reid, Flavins, Flavoproteins, and Flavoproteomics. In *Flavin and Flavoproteins*; S. Chapman, R. Perham, N. Scrutton, Eds.; Rudolf Weber: Berlin, 2002; pp 3–10.
3. E. Yeh; S. Garneau; C. T. Walsh, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3960–3965.
4. Y. Shi; F. Lan; C. Matson; P. Mulligan; J. R. Whetstone; P. A. Cole; R. A. Casero; Y. Shi, *Cell* **2004**, *119*, 941–953.
5. J. R. Terman; T. Mao; R. J. Pasterkamp; H. H. Yu; A. L. Kolodkin, *Cell* **2002**, *109*, 887–900.
6. S. Engst; P. Vock; P. Wang; J.-J. P. Kim; S. Ghisla, *Biochemistry* **1999**, *38*, 257–267.
7. D. E. Edmondson; P. Newton-Vinson, *Antioxid. Redox Signal.* **2001**, *3*, 789–806.
8. M. Ghanem; G. Gadda, *Biochemistry* **2006**, *45*, 3437–3447.
9. M. Mewies; J. Basran; L. C. Packman; R. Hille; N. S. Scrutton, *Biochemistry* **1997**, *36*, 7162–7168.
10. X. Lu; D. Nikolic; D. J. Mitchell; R. B. van Breemen; J. A. Merfelder; R. Hille; R. B. Silverman, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4129–4132.
11. Q. H. Gibson; V. Massey; N. M. Atherton, *Biochem. J.* **1962**, *85*, 369–383.
12. V. Massey; G. Palmer, *Biochemistry* **1966**, *5*, 3181–3189.
13. A. Ehrenberg; F. Müller; P. Hemmerich, *Eur. J. Biochem.* **1967**, *2*, 286–293.
14. F. Müller; P. Hemmerich; A. Ehrenberg; G. Palmer; V. Massey, *Eur. J. Biochem.* **1970**, *14*, 185–196.
15. P. Kierkegaard; R. Norrestam; P.-E. Werner; I. Csöreg; M. von Glehn; R. Karlsson; M. Leijonmark; O. Rönnquist; B. Stensland; O. Tillberg; L. Torbjörnsson, X-Ray Structure Investigation of Flavin Derivatives. In *Flavins and Flavoproteins*; H. Kamin, Ed.; University Park Press: Baltimore, 1971; pp 1–22.
16. Y.-J. Zheng; R. L. Ornstein, *J. Am. Chem. Soc.* **1996**, *118*, 9402–9408.
17. C. J. Rizzo, *Antioxid. Redox Signal.* **2001**, *3*, 737–746.
18. J. Rodríguez-Otero; E. Martínez-Núñez; A. Peña-Gallego; S. A. Vázquez, *J. Org. Chem.* **2002**, *67*, 6347–6352.
19. B. W. Lennon; C. H. Williams; M. L. Ludwig, *Prot. Sci.* **1999**, *8*, 2366–2379.
20. R. F. Anderson, *Biochim. Biophys. Acta* **1983**, *722*, 158–162.
21. S. G. Mayhew; G. Tollin, General Properties of Flavodoxins. In *Chemistry and Biochemistry of Flavoenzymes VIII*; F. Müller, Ed.; CRC Press: Boca Raton, FL, 1992; pp 389–426.
22. F. Talfournier; A. W. Munro; J. Basran; M. J. Sutcliffe; S. Daff; S. K. Chapman; N. S. Scrutton, *J. Biol. Chem.* **2001**, *276*, 20190–20196.
23. S. Ghisla; V. Massey, *Biochem. J.* **1986**, *239*, 1–12.
24. S. T. Cole; R. Brosch; J. Parkhill; T. Garnier; C. Churcher; D. Harris; S. V. Gordon; K. Eiglmeier; S. Gas; C. E. Barry, III; F. Tekaia; K. Badcock; D. Basham; D. Brown; T. Chillingworth; R. Connor; R. Davies; K. Devlin; T. Feltham; S. Gentles; N. Hamlin; S. Holroyd; T. Hornsby; K. Jagels; A. Krogh; J. McLean; S. Moule; L. Murphy; K. Oliver; J. Osborne; M. A. Quail; M.-A. Rajandream; J. Rogers; S. Rutter; K. Seeger; J. Skelton; R. Squares; S. Squares; J. E. Sulston; K. Taylor; S. Whitehead; B. G. Barrell, *Nature* **1998**, *393*, 537.
25. S. Omura; H. Ikeda; J. Ishikawa; A. Hanamoto; C. Takahashi; M. Shinose; Y. Takahashi; H. Horikawa; H. Nakazawa; T. Osonoe; H. Kikuchi; T. Shiba; Y. Sakaki; M. Hattori, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12215–12220.

26. H. A. Krebs, *Biochem. J.* **1935**, *29*, 1620–1644.
27. V. I. Tishkov; S. Khoronenkova, *Biochemistry (Moscow)* **2005**, *70*, 40–54.
28. M. J. Schell; M. E. Molliver; S. H. Snyder, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3948–3952.
29. H. Wolosker; K. N. Sheth; M. Takahashi; J. P. Mothet; R. O. Brady, Jr; C. D. Ferris; S. H. Snyder, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 721–725.
30. K. Hamase; T. Inoue; A. Morikawa; R. Konno; K. Zaitso, *Anal. Biochem.* **2001**, *298*, 253–258.
31. T. Haradahira; T. Okouchi; J. Maeda; M.-R. Zhang; T. Nishikawa; R. Konno; K. Suzuki; T. Suhara, *Synapse* **2003**, *50*, 130–136.
32. A. D'Aniello; A. Vetere; L. Petrucelli, *Comp. Biochem. Physiol. B* **1993**, *105*, 731–734.
33. M. Dixon; K. Kleppe, *Biochim. Biophys. Acta* **1965**, *96*, 357–365.
34. H. Scannone; D. Wellner; A. Novogronsky, *Biochemistry* **1964**, *11*, 1742–1745.
35. M. Gabler; M. Hensel; L. Fischer, *Enzyme Microb. Technol.* **2000**, *27*, 605–611.
36. B. Curti; S. Ronchi; M. Pilone Simonetta, D- and L-Amino Acid Oxidases. In *Chemistry and Biochemistry of Flavoenzymes*; F. Müller, Ed.; CRC Press: Boca Raton, FL, 1992; Vol. 3, pp 69–94.
37. M. S. Pilone; L. Pollegioni; P. Casalin; B. Curti; S. Ronchi, *Eur. J. Biochem.* **1989**, *180*, 199–204.
38. L. Piubelli; L. Caldinelli; G. Molla; M. S. Pilone; L. Pollegioni, *FEBS Lett.* **2002**, *526*, 43–48.
39. T. Schröder; J. R. Andreesen, *Arch. Microbiol.* **1996**, *165*, 465–471.
40. P. Casalin; L. Pollegioni; B. Curti; M. S. Pilone, *Eur. J. Biochem.* **1991**, *197*, 513–517.
41. V. Massey; B. Curti; H. Ganther, *J. Biol. Chem.* **1966**, *241*, 2347–2357.
42. V. Massey; S. Ghisla; E. G. Moore, *J. Biol. Chem.* **1979**, *254*, 9640–9650.
43. A. Mattevi; M. A. Vanoni; F. Todone; M. Rizzi; A. Teplyakov; A. Coda; M. Bolognesi; B. Curti, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 7496–7501.
44. L. Pollegioni; K. Diederichs; G. Molla; S. Umhau; W. Welte; S. Ghisla; M. S. Pilone, *J. Mol. Biol.* **2002**, *324*, 535–546.
45. L. Pollegioni; B. Langkau; W. Tischer; S. Ghisla; M. S. Pilone, *J. Biol. Chem.* **1993**, *268*, 13850–13857.
46. D. J. Porter; J. G. Voet; H. J. Bright, *J. Biol. Chem.* **1973**, *248*, 4400–4416.
47. D. J. Porter; J. G. Voet; H. J. Bright, *J. Biol. Chem.* **1977**, *252*, 4464–4473.
48. F. Todone; M. A. Vanoni; A. Mozzarelli; M. Bolognesi; A. Coda; B. Curti; A. Mattevi, *Biochemistry* **1997**, *36*, 5853–5860.
49. L. B. Hersh; M. Schuman-Jorns, *J. Biol. Chem.* **1975**, *250*, 8728–8734.
50. D. J. Manstein; V. Massey; S. Ghisla; E. F. Pai, *Biochemistry* **1988**, *27*, 2300–2305.
51. C. T. Walsh; A. Schonbrunn; R. H. Abeles, *J. Biol. Chem.* **1971**, *246*, 6855–6866.
52. C. T. Walsh; E. Krodel; V. Massey; R. H. Abeles, *J. Biol. Chem.* **1973**, *248*, 1946–1955.
53. T. W. Chan; T. C. Bruice, *Biochemistry* **1978**, *17*, 4784–4793.
54. L. E. Brown; G. E. Hamilton, *J. Am. Chem. Soc.* **1970**, *92*, 7225–7227.
55. H. Mizutani; I. Miyahara; K. Hirotsu; Y. Nishina; K. Shiga; C. Setoyama; R. Miura, *J. Biochem. (Tokyo)* **1996**, *120*, 14–17.
56. S. Umhau; L. Pollegioni; G. Molla; K. Diederichs; W. Welte; M. S. Pilone; S. Ghisla, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12463–12468.
57. L. Pollegioni; W. Blodig; S. Ghisla, *J. Biol. Chem.* **1997**, *272*, 4924–4934.
58. K. A. Kurtz; M. A. Rishavy; W. W. Cleland; P. F. Fitzpatrick, *J. Am. Chem. Soc.* **2000**, *122*, 12896–12897.
59. P. F. Fitzpatrick, *Bioorg. Chem.* **2004**, *32*, 125–139.
60. P. F. Fitzpatrick, *Acc. Chem. Res.* **2001**, *34*, 299–307.
61. A. M. Cesura, *Prog. Drug Res.* **1992**, *38*, 171–297.
62. E. B. Kearney; J. I. Slach; W. H. Walker; R. Seng; T. P. Singer, *Biochem. Biophys. Res. Commun.* **1971**, *42*, 490.
63. H. Blaschko, *Pharmacol. Rev.* **1952**, *4*, 415–458.
64. L. De Colibus; M. Li; C. Binda; A. Lustig; D. E. Edmondson; A. Mattevi, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12684–12689.
65. G. A. Akkes; E. V. Heegaard, *J. Biol. Chem.* **1943**, *147*, 487–503.
66. C. Guffroy; C. J. Fowler; M. J. Strolin Benedetti, *J. Pharm. Pharmacol.* **1982**, *35*, 416–420.
67. R. B. Yamasaki; R. B. Silverman, *Biochemistry* **1985**, *24*, 6543–6550.
68. T. P. Singer; R. R. Ramsay, *FASEB J.* **1995**, *9*, 605–610.
69. C. Binda; P. Newton-Vinson; F. Hubálek; D. E. Edmondson; A. Mattevi, *Nat. Struct. Biol.* **2002**, *9*, 22–26.
70. C. Binda; M. Li; F. Hubálek; N. Restelli; D. E. Edmondson; A. Mattevi, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9750–9755.
71. J. Ma; M. Yoshimura; E. Yamashita; A. Nakagawa; A. Ito; T. Tsukihara, *J. Mol. Biol.* **2004**, *338*, 103–114.
72. M. Husain; D. E. Edmondson; T. P. Singer, *Biochemistry* **1982**, *21*, 595–600.
73. R. B. Silverman, *Acc. Chem. Res.* **1995**, *28*, 335–342.
74. R. B. Silverman; J. P. Zhou; P. E. Eaton, *J. Am. Chem. Soc.* **1993**, *115*, 8841–8842.
75. J. R. Miller; D. E. Edmondson, *J. Am. Chem. Soc.* **1995**, *117*, 7830–7831.
76. A. Tan; M. D. Glantz; L. H. Peitte; K. T. Yasunobu, *Biochem. Biophys. Res. Commun.* **1983**, *117*, 517–523.
77. S. E. J. Rigby; R. M. G. Hynson; R. R. Ramsay; A. W. Munro; N. S. Scrutton, *J. Biol. Chem.* **2005**, *280*, 4627–4631.
78. J. M. Kim; S. E. Hoegy; P. S. Mariano, *J. Am. Chem. Soc.* **1995**, *117*, 100–105.
79. J. R. Miller; D. E. Edmondson, *Biochemistry* **1999**, *38*, 13670–13683.
80. M. C. Walker; D. E. Edmondson, *Biochemistry* **1994**, *33*, 7088.
81. M. Li; C. Binda; A. Mattevi; D. E. Edmondson, *Biochemistry* **2006**, *45*, 4775–4784.
82. S. L. Schreiber; B. E. Bernstein, *Cell* **2002**, *111*, 771–778.
83. T. Jenuwein; C. D. Allis, *Science* **2001**, *293*, 1074–1080.
84. C. Martin; Y. Zhang, *Nat. Rev. Mol. Cell. Biol.* **2005**, *6*, 838–849.
85. Y. Shi; J. R. Whetstone, *Mol. Cell.* **2007**, *25*, 1–14.
86. F. Forneris; C. Binda; M. A. Vanoni; A. Mattevi; E. Battaglioli, *FEBS Lett.* **2005**, *579*, 2203–2207.
87. R. Anand; R. Mamorstein, *J. Biol. Chem.* **2007**, *282*, 35425–35429.
88. F. Forneris; C. Binda; M. A. Vanoni; E. Battaglioli; A. Mattevi, *J. Biol. Chem.* **2005**, *280*, 41360–41365.
89. F. Forneris; C. Binda; A. Dall'Aglio; M. W. Fraaije; E. Battaglioli; A. Mattevi, *J. Biol. Chem.* **2006**, *281*, 35289–35295.
90. M. G. Lee; C. Wynder; D. M. Schmidt; D. G. McCafferty; R. Shiekhata, *Chem. Biol.* **2006**, *13*, 563–567.

91. D. M. Schmidt; D. G. McCafferty, *Biochemistry* **2007**, *46*, 4408–4416.
92. J. C. Culhane; L. M. Szewczuk; X. Liu; G. Da; R. Marmorstein; P. A. Cole, *J. Am. Chem. Soc.* **2006**, *128*, 4536–4537.
93. L. M. Szewczuk; J. C. Culhane; M. Yang; A. Majumdar; H. Yu; P. A. Cole, *Biochemistry* **2007**, *46*, 6892–6902.
94. K. Kvalnes-Krick; M. S. Jorns, Role of the Covalent and Noncovalent Flavins in Sarcosine Oxidase. In *Chemistry and Biochemistry of Flavoenzymes VII*; F. Müller, Ed.; CRC Press: Boca Raton, FL, 1991; pp 425–435.
95. M. A. Wagner; P. Khanna; M. S. Jorns, *J. Biol. Chem.* **1999**, *270*, 18252–18259.
96. A. Hassan-Adballah; G. Zhao; M. S. Jorns, *Biochemistry* **2006**, *45*, 9454–9462.
97. M. A. Wagner; M. S. Jorns, *Biochemistry* **2000**, *39*, 8825–8829.
98. M. A. Wagner; P. Trickey; Z. Chen; F. S. Mathews; M. S. Jorns, *Biochemistry* **2000**, *39*, 8813–8824.
99. G. Zhao; M. S. Jorns, *Biochemistry* **2006**, *45*, 5985–5992.
100. G. Zhao; F. A. Franklin; M. S. Jorns, *Biochemistry* **2000**, *39*, 14341–14347.
101. Z. Chen; G. Zhao; S. Martinovic; M. S. Jorns; F. S. Mathews, *Biochemistry* **2005**, *44*, 15444–15450.
102. I. Yokoe; T. C. Bruice, *J. Am. Chem. Soc.* **1975**, *97*, 450–451.
103. E. C. Settembre; P. C. Dorestein; J. Park; A. M. Augustine; T. P. Begley; S. E. Ealick, *Biochemistry* **2003**, *42*, 2971–2981.
104. E. Adams; L. Frank, *Annu. Rev. Biochem.* **1980**, *49*, 1005–1061.
105. J. M. Phang, *Curr. Top. Cell Reg.* **1985**, *25*, 92–132.
106. R. Menzel; J. Roth, *J. Biol. Chem.* **1981**, *256*, 9762–9766.
107. E. Brown; J. M. Wood, *J. Biol. Chem.* **1992**, *267*, 13086–13092.
108. A. M. Muro-Pastor; P. Ostrovsky; S. Maloy, *J. Bacteriol.* **1995**, *179*, 2788–2791.
109. D. F. Becker; E. A. Thomas, *Biochemistry* **2001**, *40*, 4714–4722.
110. T. A. White; K. Navasona; D. F. Becker; J. J. Tanner, *J. Biol. Chem.* **2007**, *282*, 14316–14327.
111. R. C. Scarpulla; R. L. Soffer, *J. Biol. Chem.* **1978**, *253*, 5997–6001.
112. M. W. Surber; S. Maloy, *Arch. Biochem. Biophys.* **1998**, *354*, 281–287.
113. Y. H. Lee; S. Nadaraja; D. Gu; D. F. Becker; J. J. Tanner, *Nat. Struct. Biol.* **2003**, *10*, 109–114.
114. M. Zhang; T. A. White; J. P. Schuermann; B. A. Baban; D. F. Becker; J. J. Tanner, *Biochemistry* **2004**, *43*, 12539–12548.
115. D. Gu; Y. Zhou; V. Kallhoff; B. Baban; J. J. Tanner; D. F. Becker, *J. Biol. Chem.* **2004**, *279*, 31171–31176.
116. J. D. Larson; J. L. Jenkins; J. P. Schuermann; Y. Zhou; D. F. Becker; J. J. Tanner, *Protein Sci.* **2006**, *15*, 2630–2641.
117. J. Wood, *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 373–377.
118. E. D. Brown; J. M. Wood, *J. Biol. Chem.* **1993**, *268*, 8972–8979.
119. W. Zhang; Y. Zhou; D. F. Becker, *Biochemistry* **2004**, *43*, 13165–13174.
120. Z. Zhang; M. Zhang; W. Zhu; Y. Zhou; S. Wanduragala; D. Rewinkel; J. J. Tanner; D. F. Becker, *Biochemistry* **2007**, *46*, 483–491.
121. P. Moreau; F. Anizon; M. Sancelme; M. Prudhomme; C. Mailly; D. Severe; J.-F. Riou; D. Fabbro; T. Meyer; A.-M. Aubertin, *J. Med. Chem.* **1999**, *42*, 584–592.
122. M. W. Saif; R. B. Diassia, *Clin. Colorectal Cancer* **2005**, *5*, 27–36.
123. M. Hussian; U. Vaishampayan; L. K. Heilbrun; V. Jain; P. M. LoRusso; P. Ivy; L. Flaherty, *Invest. New Drugs* **2003**, *21*, 465–471.
124. M. Facompre; J.-F. Goossens; C. Bailey, *Biochem. Pharmacol.* **2001**, *61*, 299–310.
125. M. Nishimura; Y. D. Akhmedov; K. Strzalka; T. Akazawa, *Arch. Biochem. Biophys.* **1983**, *222*, 397–402.
126. G. Blankenhorn, *Eur. J. Biochem.* **1976**, *67*, 67–80.
127. M. F. Powell; W. H. Wong; T. C. Bruice, *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4604–4608.
128. N. Agrawal; S. A. Lesley; P. Kuhn; A. Kohen, *Biochemistry* **2004**, *43*, 10295–10301.
129. G. Blankenhorn, *Eur. J. Biochem.* **1975**, *14*, 3172–3176.
130. G. Blankenhorn, *Eur. J. Biochem.* **1975**, *50*, 351–356.
131. J. Basran; R. J. Harris; M. J. Sutcliffe; N. S. Scrutton, *J. Biol. Chem.* **2003**, *278*, 43973–43982.
132. A. Pennati; G. Zanetti; A. Aliverti; A. Gadda, *Biochemistry* **2008**, *47*, 3418–3425.
133. Z. Deng; A. Aliverti; G. Zanetti; A. K. Arakaki; J. Ottado; E. G. Orellano; N. B. Calcaterra; E. A. Ceccarelli; N. Carrillo; P. A. Karplus, *Nat. Struct. Biol.* **1999**, *6*, 847–853.
134. R. H. H. Van den Heuvel; A. H. Westphal; A. J. R. Heck; M. A. Walsh; S. Rovida; W. J. H. van Berkel; A. Mattevi, *J. Biol. Chem.* **2004**, *279*, 12860–12867.
135. J. Wang; M. Ortiz-Maldonado; B. Entsch; V. Massey; D. P. Ballou; D. L. Gatti, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 608–613.
136. S. Ojha; E. C. Meng; P. C. Babbitt, *PLoS Comp. Biol.* **2007**, *3*, 1268–1280.
137. J. H. Pazur; K. Kleppe, *Biochemistry* **1964**, *3*, 2951–2959.
138. J. H. Pazur; K. Kleppe; A. Cepure, *Arch. Biochem. Biophys.* **1965**, *111*, 351–357.
139. S. Hayashi; S. Nakamura, *Biochim. Biophys. Acta* **1981**, *657*, 40–51.
140. K. Takegawa; K. Fukiwara; S. Iwahara; K. Yamamoto; T. Tochikura, *Biochem. Cell Biol.* **1989**, *67*, 460–464.
141. K. Takegawa; K. Fukiwara; S. Iwahara; K. Yamamoto; T. Tochikura, *Agric. Biol. Chem.* **1991**, *55*, 883–884.
142. H. M. Kalisz; R. D. Schmid; H. J. Hecht; D. Schomburg, *Biochim. Biophys. Acta* **1991**, *1080*, 138–142.
143. A. Kohen; T. Jonsson; J. P. Klinman, *Biochemistry* **1997**, *36*, 2603–2611.
144. H. J. Hecht; H. M. Kalisz; J. Hendle; R. D. Schmid; D. Schomburg, *J. Mol. Biol.* **1993**, *229*, 153–172.
145. Q. H. Gibson; B. E. P. Swoboda; V. Massey, *J. Biol. Chem.* **1964**, *239*, 3927–3934.
146. D. J. Manstein; E. F. Pai; S. Schopfer; V. Massey, *Biochemistry* **1986**, *25*, 6807–6816.
147. G. Wohlfart; S. Witt; J. Hendle; D. Schomburg; H. M. Kalisz; H.-J. Hecht, *Acta Crystallogr.* **1999**, *D55*, 969–977.
148. H. J. Bright; Q. H. Gibson, *J. Biol. Chem.* **1967**, *242*, 994–1003.
149. D. J. Porter; H. J. Bright, *J. Biol. Chem.* **1977**, *252*, 4361–4370.
150. C. Jaccq; E. Lederer, *Eur. J. Biochem.* **1974**, *41*, 311–320.
151. A. Baudras, *Biochem. Biophys. Res. Commun.* **1962**, *7*, 310–314.
152. B. Guiard; O. Groudinsky; F. Lederer, *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 2539–2543.
153. C. S. Miles; N. Rouviere-Fourmy; F. Lederer; F. S. Mathews; G. A. Reid; F. S. Chapman, *Biochem. J.* **1992**, *285*, 187–192.
154. C. Capeillere-Blandin, *Eur. J. Biochem.* **1975**, *56*, 91–101.
155. C. Capeillere-Blandin; R. C. Bray; M. Iwatsubo; F. Labeyrie, *Eur. J. Biochem.* **1975**, *54*, 549–566.

156. S. K. Chapman; G. A. Reid; S. Daff; R. E. Sharp; P. White; F. D. C. Manson; F. Lederer, *Biochem. Soc. Trans.* **1994**, *22*, 713–718.
157. D. Pompon; *Eur. J. Biochem.* **1980**, *106*, 151–159.
158. D. Pompon; M. Iwatsubo; F. Lederer, *Eur. J. Biochem.* **1980**, *104*, 479–488.
159. S. Daff; R. E. Sharp; D. M. Short; C. Bell; P. White; F. D. C. Manson; S. K. Chapman, *Biochemistry* **1996**, *35*, 6351–6357.
160. S. Daff; W. J. Ingledew; G. A. Reid; S. K. Chapman, *Biochemistry* **1996**, *35*, 6345–6350.
161. F. Lederer, *Eur. J. Biochem.* **1974**, *46*, 393–399.
162. P. Urban; P. M. Alliel; F. Lederer, *Eur. J. Biochem.* **1983**, *134*, 275–281.
163. P. Urban; F. Lederer, *J. Biol. Chem.* **1985**, *260*, 11115–11122.
164. Z.-X. Xia; F. S. Mathews, *J. Mol. Biol.* **1990**, *212*, 837–863.
165. J. Dubios; S. K. Chaomen; F. S. Mathews; G. A. Reid; F. Lederer, *Biochemistry* **1990**, *29*, 6396–6400.
166. M. Gondry; J. Dubios; M. Terrier; F. Lederer, *Eur. J. Biochem.* **2001**, *268*, 4918–4927.
167. R. Sinclair; G. A. Reid; S. K. Chapman, *Biochem. J.* **1998**, *333*, 117–120.
168. C. G. Mowat; A. Wehenkel; A. J. Green; M. D. Walkinshaw; G. A. Reid; S. K. Chaoman, *Biochemistry* **2004**, *43*, 9519–9526.
169. P. Sobrado; P. F. Fitzpatrick, *Biochemistry* **2003**, *42*, 15208–15214.
170. P. Sobrado; S. C. Daubner; P. F. Fitzpatrick, *Biochemistry* **2001**, *40*, 994–1001.
171. O. Lockridge; V. Massey; P. A. Sullivan, *J. Biol. Chem.* **1972**, *247*, 8097–8106.
172. V. Massey; F. Müller; R. Feldberg; M. Schuman; P. A. Sullivan; L. G. Howell; S. G. Mayhew; R. G. Matthews; G. P. Foust, *J. Biol. Chem.* **1969**, *244*, 3999–4006.
173. F. Müller; V. Massey, *J. Biol. Chem.* **1969**, *244*, 4007–4016.
174. U. Müh; V. Massey; C. H. Williams, Jr., *J. Biol. Chem.* **1994**, *269*, 7982–7988.
175. M. Stankovich; B. Fox, *Biochemistry* **1983**, *22*, 4466–4472.
176. Y. S. Choong; V. Massey, *J. Biol. Chem.* **1980**, *255*, 8672–8677.
177. S. Ghisla; V. Massey, L-Lactate Oxidase. In *Chemistry and Biochemistry of Flavoenzymes*; F. Müller, Ed.; CRC Press: Boca Raton, 1991; Vol. II, pp 243–289.
178. U. Müh; V. Massey; C. H. Williams, Jr., *J. Biol. Chem.* **1994**, *269*, 7989–7993.
179. U. Müh; V. Massey; C. H. Williams, Jr., *J. Biol. Chem.* **1994**, *269*, 7994–8000.
180. K. Maeda-Yorita; K. Aki; H. Sagai; H. Misaki; V. Massey, *Biochimie* **1995**, *77*, 631–642.
181. Y. Urnena; K. Yorita; T. Matsuoka; A. Kita; K. Fukui; Y. Morimoto, *Biochem. Biophys. Res. Commun.* **2006**, *350*, 249–256.
182. K. Yorita; H. Misaki; B. A. Palfey; V. Massey, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2480–2485.
183. F. L. Zhang; P. J. Casey, *Annu. Rev. Biochem.* **1996**, *65*, 241–269.
184. R. R. Rando, *Biochim. Biophys. Acta* **1996**, *1300*, 5–16.
185. H. W. Fu; P. J. Casey, *Recent Prog. Horm. Res.* **1999**, *54*, 315–343.
186. L. Zhang; W. R. Tschantz; P. J. Casey, *J. Biol. Chem.* **1997**, *272*, 23354–23359.
187. W. R. Tschantz; J. A. Digits; H. Pyun; R. M. Coates; P. J. Casey, *J. Biol. Chem.* **2001**, *276*, 2321–2324.
188. J. A. Digits; H. Pyun; R. M. Coates; P. J. Casey, *J. Biol. Chem.* **2002**, *277*, 41086–41093.
189. Y. Reiss; J. L. Goldstein; M. C. Seabra; P. J. Casey; M. S. Brown, *Cell* **1990**, *62*, 81–88.
190. D. L. Pompliano; E. Rands; M. D. Schaber; S. D. Mosser; N. J. Anothony; J. B. Gibbs, *Biochemistry* **1990**, *31*, 3800–3807.
191. B. K. Haarer; D. C. Amberg, *Mol. Biol. Cell.* **2004**, *15*, 4522–4531.
192. T. B. Fitzpatrick; N. Amrhein; P. Macheroux, *J. Biol. Chem.* **2003**, *278*, 19891–19897.
193. K. Stott; K. Saito; D. Thiele; V. Massey, *J. Biol. Chem.* **1993**, *268*, 6097–6106.
194. A. S. Abramovitz; V. Massey, *J. Biol. Chem.* **1976**, *251*, 5327–5336.
195. V. Massey; The Enigma of Old Yellow Enzyme II. In *Flavins and Flavoproteins 1993*; K. Yagi, Ed.; Walter de Gruyter: Berlin, 1994; pp 371–380.
196. A. D. N. Vaz; S. Chakraborty; V. Massey, *Biochemistry* **1995**, *34*, 4246–4256.
197. R. G. Matthews; V. Massey, *J. Biol. Chem.* **1969**, *244*, 1779–1786.
198. R. G. Matthews; V. Massey, *J. Biol. Chem.* **1975**, *250*, 9294–9298.
199. R. C. Stewart; V. Massey, *J. Biol. Chem.* **1985**, *260*, 13639–13647.
200. K. M. Fox; P. A. Karplus, *Structure* **1994**, *2*, 1089–1105.
201. V. Massey; L. M. Schopfer, *J. Biol. Chem.* **1986**, *261*, 1215–1222.
202. L. M. Schopfer; V. Massey, Old Yellow Enzyme. In *A Study of Enzymes*; S. A. Kuby, Ed.; CRC Press: Cleveland, OH, 1991; pp 247–269.
203. Y. Meah; V. Massey, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10733–10738.
204. R. E. Williams; N. C. Bruce, *Microbiology* **2002**, *148*, 1607–1614.
205. J. L. Ramos; M. M. González-Pérez; A. Caballero; P. van Dillewijn, *Curr. Opin. Biotech.* **2005**, *16*, 275–281.
206. Y. Meah; B. J. Brown; S. Chakraborty; V. Massey, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8560–8565.
207. H. Khan; R. J. Harris; T. Barna; D. H. Craig; N. C. Bruce; A. W. Munro; P. C. E. Moody; N. S. Scrutton, *J. Biol. Chem.* **2002**, *277*, 21906–21912.
208. T. Aoyama; M. Souri; S. Ushikubo; T. Kamijo; S. Yamaguchi; R. I. Kelley; W. J. Rhead; K. Uteake; K. Tanakn; T. Hasimoto, *J. Clin. Invest.* **1995**, *95*, 2465–2473.
209. Y. Ikeda; K. Okamura-Ikeda; K. Tanaka, *J. Biol. Chem.* **1985**, *260*, 1311–1325.
210. K. Izai; Y. Uchida; T. Orii; S. Yamamoto; T. Hasimoto, *J. Biol. Chem.* **1992**, *267*, 1027–1033.
211. J. Zhang; W. Zhang; D. Zou; G. Chen; T. Wan; M. Zhang; X. Cao, *Biochem. Biophys. Res. Commun.* **2002**, *297*, 1033–1042.
212. R. Ensenuer; M. He; J. M. Willard; E. S. Goetxman; T. J. Corydon; B. N. Vandahl; A. W. Mohsen; G. Isaya; J. Voxkley, *J. Biol. Chem.* **2005**, *280*, 32309–32316.
213. N. Gregersen; P. Bross; B. S. Andersen, *Eur. J. Biochem.* **2004**, *271*, 470–782.
214. F. L. Crane; H. Beinert, *J. Biol. Chem.* **1956**, *218*, 717–731.
215. F. J. Ruzicka; H. Beinert, *J. Biol. Chem.* **1977**, *252*, 8440–8445.
216. M. Souri; T. Aoyama; G. Hoganson; T. Hasimoto, *FEBS Lett.* **1998**, *426*, 187–190.



217. Y. Matsubara; Y. Indo; E. Naito; H. Ozasa; R. Glassburg; J. Vockley; Y. Ikeda; J. Kraus; K. Tankaka, *J. Biol. Chem.* **1989**, *264*, 16321–16331.
218. Y. Nakajima; I. Miyahara; K. Hirotsu; Y. Nishina; K. Shiga; C. Setoyama; H. Tamaoki; R. Miura, *J. Biochem.* **2002**, *131*, 365–374.
219. G. Williams; P. C. Engel; J. P. Mizzer; C. Thorpe; C. Massey, *J. Biol. Chem.* **1982**, *257*, 4314–4320.
220. A. Satoh; Y. Nakajima; I. Miyahara; K. Hirotsu; T. Tanaka; Y. Nishina; K. Shiga; H. Tamaokie; C. Setoyama; R. Miura, *J. Biochem. (Tokyo)* **2003**, *134*, 297–304.
221. P. C. Engel, Acyl-Coenzyme A Dehydrogenases. In *Chemistry and Biochemistry of Flavoenzymes*; F. Muller, Ed.; CRC Press: Boca Raton, FL, 1992; pp 597–655.
222. E. P. Steyn-Parve; H. Beinert, *J. Biol. Chem.* **1958**, *233*, 853–861.
223. H. Beinert, Acyl-Coenzyme A Dehydrogenases. In *Enzymes*; P. D. Boyer, H. Lardy, K. Myrback, Eds.; Academic Press: New York, NY, 1963; pp 447–466.
224. R. C. Trievel; R. Wang; V. E. Anderson; C. Thorpe, *Biochemistry* **1995**, *34*, 8597–8605.
225. P. J. Powell; S.-M. Lau; C. Thorpe, *Biochemistry* **1987**, *26*, 3704–3710.
226. J. P. Mizzer; C. Thorpe, *Biochemistry* **1981**, *19*, 4965–4970.
227. B. D. Johnson; G. J. Mancini-Samuelsen; M. T. Stankovich, *Biochemistry* **1995**, *34*, 7047–7055.
228. N. R. Kumar; D. K. Srivastava, *Biochemistry* **1995**, *34*, 9434–9443.
229. C. Thrope; J. J. Kim, *P. FASEB J.* **1995**, *9*, 718–725.
230. N. R. Kumar; D. K. Srivastava, *Biochemistry* **1994**, *33*, 8833–8841.
231. J. K. Johnson; Z.-X. Wang; D. K. Srivastava, *Biochemistry* **1992**, *31*, 10564–10575.
232. I. Rudik; S. Ghisla; C. Thorpe, *Biochemistry* **1998**, *37*, 8437–8445.
233. N. D. Lenn; M. T. Stankovich; H.-W. Liu, *Biochemistry* **1990**, *29*, 3709–3715.
234. R. J. Gorelick; L. M. Schopfer; D. P. Ballou; V. Massey; C. Thorpe, *Biochemistry* **1985**, *24*, 6830–6839.
235. S. Ghisla; C. Thorpe, *Eur. J. Biochem.* **2004**, *271*, 494–508.
236. J. F. Biellmann; C. G. Hirth, *FEBS Lett.* **1970**, *9*, 335–336.
237. A. Kawaguchi; S. Tsubotani; Y. Seyama; T. Yamakawa; T. Osumi; T. Hasmitot; T. Kikuchi; M. Ando; S. Okuda, *J. Biochem. (Tokyo)* **1980**, *88*, 1481–1486.
238. B. Pohl; T. Raichle; S. Ghisla, *Eur. J. Biochem.* **1986**, *160*, 109–115.
239. L. M. Schopfer; V. Massey; S. Ghisla; C. Thorpe, *Biochemistry* **1988**, *27*, 6599–6611.
240. J.-J. P. Kim; J. Wu, Three Dimensional Structure of Medium-Chain Acyl-CoA Dehydrogenase. In *Chemistry and Biochemistry of Flavoenzymes*; F. Müller, Ed.; CRC Press: Boca Raton, FL, 1992; Vol. III, pp 299–307.
241. S. Ghisla; S. Engst; P. Vock; V. Kieweg; P. Bross; A. Nandy; I. Rasched; A. W. Strauss, Mechanism of  $\alpha,\beta$ -Dehydrogenation by Acyl-CoA Dehydrogenases. In *Flavins and Flavoproteins 1993*; K. Yagi, Ed.; Walter de Gruyter: Berlin, 1994; pp 283–292.
242. O. Björnberg; P. Rowland; S. Larsen; K. F. Jensen, *Biochemistry* **1997**, *36*, 16197–16205.
243. M. E. Jones, *Annu. Rev. Biochem.* **1980**, *49*, 253–279.
244. M. Nagy; F. Lacroute; D. Thomas, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8966–8970.
245. F. S. Nielsen; P. S. Andersen; K. F. Jensen, *J. Biol. Chem.* **1996**, *271*, 29359–29365.
246. R. Baumgartner; M. Walloschek; M. Kralik; A. Gotschlich; S. Tasler; J. Miles; J. Leban, *J. Med. Chem.* **2006**, *49*, 1239–1247.
247. E. Zameitat; Z. Gojkovic; W. Knecht; J. Piskur; M. Löffler, *FEBS Lett.* **2006**, *273*, 3183–3191.
248. D. E. Hurt; J. Widom; J. Clardy, *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **2006**, *62*, 312–323.
249. T. Heikkia; C. Ramsey; M. Davies; C. Galtier; A. M. Stead; W. A. P. Johnson; C. W. G. Fishwick; A. N. Boa; G. A. McConkey, *J. Med. Chem.* **2007**, *50*, 186–191.
250. N. A. Malmquist; R. Gujjar; P. K. Rathod; M. A. Phillips, *Biochemistry* **2008**, *47*, 2466–2475.
251. M. L. Herrmann; R. Schleyerbach; B. J. Kirschbaum, *Innopharmacology* **2000**, *47*, 273–289.
252. S. Liu; E. A. Neidhardt; T. H. Grossman; T. Ocain; J. Clardy, *Structure* **2000**, *8*, 25–33.
253. J. Baldwin; A. M. Farajallah; N. A. Malmquist; P. K. Rathod; M. A. Phillips, *J. Biol. Chem.* **2002**, *277*, 41827–41834.
254. S. Norager; K. F. Jensen; O. Björnberg; S. Larsen, *Structure* **2002**, *10*, 1211–1223.
255. P. Rowland; O. Björnberg; F. S. Nielsen; K. F. Jensen; S. Larsen, *Protein Sci.* **1998**, *7*, 1269–1279.
256. D. K. Inaoka; K. Sakamoto; H. Shimizu; T. Shiba; G. Kurisu; T. Nara; T. Aoki; K. Kita; S. Harada, *Biochemistry* **2008**, *47*, 10881–10891.
257. B. A. Palfey; O. Björnberg; K. F. Jensen, *J. Med. Chem.* **2001**, *44*, 2861–2864.
258. A. E. Wolfe; M. Thymark; S. G. Gattis; R. L. Fagan; Y. C. Hu; E. Johansson; S. Arent; S. Larsen; B. A. Palfey, *Biochemistry* **2007**, *46*, 5741–5753.
259. T. Annoura; T. Nara; T. Makiuchi; T. Hashimoto; T. Aoki, *J. Mol. Evol.* **2005**, *60*, 113–127.
260. R. Pascal; C. T. Walsh, *Biochemistry* **1984**, *23*, 2745–2752.
261. P. Rowland; S. Norager; K. F. Jensen; S. Larsen, *Structure* **2000**, *8*, 1227–1238.
262. T. L. Arakaki; F. S. Buckner; J. R. Gillespie; N. A. Malmquist; M. A. Phillips; O. Kalyuzhnyi; J. R. Luft; G. T. DeTitta; C. L. M. J. Verlinde; W. C. Van Voorhis; W. G. Hol; E. A. Merritt, *Mol. Microbiol.* **2008**, *68*, 37–50.
263. V. Hines; M. Johnston, *Biochemistry* **1989**, *28*, 1227–1234.
264. A. Argyrou; M. W. Washabaugh; C. M. Pickart, *Biochemistry* **2000**, *39*, 10373–10384.
265. R. L. Fagan; K. F. Jensen; O. Björnberg; B. A. Palfey, *Biochemistry* **2007**, *46*, 4028–4036.
266. R. L. Fagan; M. N. Nelson; P. M. Pagano; B. A. Palfey, *Biochemistry* **2006**, *45*, 14926–14932.
267. B. A. Palfey; O. Björnberg; K. F. Jensen, *Biochemistry* **2001**, *40*, 4381–4390.
268. T. C. Bruice; T. H. Fife; J. J. Bruno; N. E. Brandon, *Biochemistry* **1962**, *1*, 7–12.
269. D. B. Jordan; J. J. Bisaha; M. A. Picolleli, *Arch. Biochem. Biophys.* **2000**, *378*, 84–92.
270. D. Papamichael, *Stem Cells* **2000**, *18*, 166–185.
271. G. Milano; M.-C. Etienne, *Anticancer Res.* **1994**, *14*, 2295–2297.
272. W. E. Hull; R. E. Port; R. Herrmann; B. Britsch; W. Kunz, *Cancer Res.* **1988**, *48*, 1680–1688.
273. B. Podschun; K. D. Wahler; K. D. Schackerz, *Eur. J. Biochem.* **1989**, *185*, 219–224.
274. K. Rosenbaum; K. Jahnke; K. D. Schanckerz; P. F. Cook, *Biochemistry* **1998**, *37*, 9156–9159.

275. H. Yokota; P. Fernandez-Salguero; H. Furuya; K. Lin; O. W. McBride; B. Podschun; K. D. Schanckerz; F. J. Gonzalez, *J. Biol. Chem.* **1994**, 269, 23192–23196.
276. W. R. Hagen; M. A. Vanoni; K. D. Rosenbaum; K. D. Schanckerz, *Eur. J. Biol.* **2000**, 267, 3640–3646.
277. D. Dobritzsch; G. Schneider; K. D. Schanckerz; Y. Lindqvist, *EMBO J.* **2001**, 20, 650–660.
278. K. Rosenbaum; K. Jahnke; B. Curti; W. R. Hagen; K. D. Schackerz; M. A. Vanoni, *Biochemistry* **1998**, 37, 17589–17609.
279. B. Podschun; P. F. Cook; K. D. Schanckerz, *J. Biol. Chem.* **1990**, 265, 12966–12972.
280. D. Dobritzsch; S. Ricagno; G. Schneider; K. D. Schanckerz; Y. Lindqvist, *J. Biol. Chem.* **2002**, 277, 13155–13166.
281. B. Podschun; K. Jahnke; K. D. Schanckerz; P. F. Cook, *J. Biol. Chem.* **1993**, 268, 3407–3413.
282. M. Sprinzl; C. Horn; M. Brown; A. Ioudovitch; S. Steinberg, *Nucleic Acid Res.* **1998**, 26, 148–153.
283. F. Xing; M. R. Martzen; E. M. Phizicky, *RNA* **2002**, 8, 370–381.
284. A. C. Bishop; J. Xu; R. C. Johnson; P. Schimmel; V. de Crecy-Lagard, *J. Biol. Chem.* **2002**, 277, 25090–25095.
285. T. Kato; Y. Daigo; S. Hayama; N. Ishikawa; T. Yamabuki; T. Ito; M. Miyamoto; S. Kondo; Y. Nakamura, *Cancer Res.* **2005**, 65, 5638–5646.
286. M. Mittelsadt; A. Frump; T. Khuu; V. Fowlkes; I. Handy; C. V. Patel; R. C. Patel, *Nucleic Acid Res.* **2008**, 36, 998–1008.
287. F. Xing; S. L. Hiley; T. R. Hughes; E. M. Phizicky, *J. Biol. Chem.* **2004**, 279, 17850–17860.
288. D. F. Savage; V. de Crecy-Lagard; A. C. Bishop, *FEBS Lett.* **2006**, 580, 5198–5202.
289. F. Park; K. Gajiwala; B. Noland; L. Wu; D. He; J. Molinari; K. Loomis; B. Pagarigan; P. Kearins; J. Christopher; T. Peat; J. Badger; J. Hendle; J. Lin; S. Buchanan, *Proteins* **2004**, 55, 772–774.
290. L. W. Rider; M. B. Ottosen; S. G. Gattis; B. A. Palfey, *J. Biol. Chem.* **2009**, 284, 10324–10333.
291. M. Saraste, *Science* **1999**, 283, 1488–1493.
292. A. Kröger, *Biochim. Biophys. Acta* **1978**, 505, 129–145.
293. C. R. D. Lancaster, *FEBS Lett.* **2003**, 555, 21–28.
294. G. Cecchini; I. Schroder; R. P. Gunsalus; E. Maklasina, *Biochim. Biophys. Acta* **2002**, 1553, 140–157.
295. T. M. Iverson; C. Luna-Chavez; I. Schroder; G. Cecchini; D. C. Rees, *Curr. Opin. Struct. Biol.* **2000**, 10, 448–455.
296. C. Hägerhäll, *Biochim. Biophys. Acta* **1997**, 1320, 107–141.
297. T. P. Singer; M. K. Johnson, *FEBS Lett.* **1985**, 190, 189–198.
298. S. L. Pealing; M. R. Cheesman; G. A. Reid; A. J. Thomas; F. B. Ward; S. K. Chapman, *Biochemistry* **1995**, 34, 6153–6158.
299. S. L. Pealing; A. C. Black; F. D. Manson; F. B. Ward; S. K. Chapman; G. A. Reid, *Biochemistry* **1992**, 31, 12132–12140.
300. P. Taylor; S. L. Pealing; G. A. Reid; S. K. Chapman; M. D. Walkinshaw, *Nat. Struct. Biol.* **1999**, 6, 1108–1115.
301. A. Kröger; V. Geisler; E. Lemma; F. Theis; R. Lenger, *Arch. Microbiol.* **1992**, 158, 311–314.
302. D. Leys; A. S. Tsapin; K. H. Nealon; T. E. Meyer; M. A. Cusanovich; J. J. Van Beeumen, *Nat. Struct. Biol.* **1999**, 6, 1113–1117.
303. V. Bamford; P. S. Dobbin; D. J. Richardson; M. Hemmings, *Nat. Struct. Biol.* **1999**, 6, 1104–1107.
304. M. K. Doherty; S. L. Pealing; C. S. Miles; R. Moysey; P. Taylor; M. D. Walkinshaw; G. A. Reid; S. K. Chapman, *Biochemistry* **2000**, 39, 10695–10701.
305. C. G. Mowat; R. Moysey; C. S. Miles; D. Leys; M. K. Doherty; P. Taylor; M. D. Walkinshaw; G. A. Reid; S. K. Chapman, *Biochemistry* **2001**, 40, 12292–12298.
306. K. L. Pankhurst; C. G. Mowat; E. L. Rothery; J. M. Hudson; A. K. Jones; C. S. Miles; M. D. Walkinshaw; F. A. Armstrong; G. A. Reid; S. K. Chapman, *J. Biol. Chem.* **2006**, 281, 20589–20597.
307. T. E. Benson; J. L. Marquardt; A. C. Marquardt; F. A. Etkorn; C. T. Walsh, *Biochemistry* **1993**, 32, 2024–2030.
308. C. J. Andres; J. J. Bronson; S. V. D'Andrea; M. S. Deshpande; P. J. Palf; K. A. Grant-Young; W. E. Harte; H.-T. Ho; P. F. Misco; J. G. Robertson; D. Stock; Y. Sun; A. W. Walsh, *Bioorg. Med. Chem. Lett.* **2000**, 10, 715–717.
309. M. J. Pucci; L. F. Discotto; T. J. Dougherty, *J. Bacteriol.* **1992**, 174, 1690–1693.
310. S. L. Rowland; J. Errington; R. G. Wake, *Gene* **1995**, 164, 113–116.
311. D. R. Sylvester; E. Alavarez; K. Ratnam; H. Kallender; N. G. Wallis, *Biochem. J.* **2001**, 355, 431–435.
312. S. Nishida; K. Kurokawa; M. Matsuo; K. Sakamoto; K. Ueno; K. Kita; K. Sekimizu, *J. Biol. Chem.* **2006**, 281, 1714–1724.
313. A. M. Dhalla; J. Yanchunas; H. T. Ho; P. J. Falk; J. Villafranca; J. G. Robertson, *Biochemistry* **1995**, 34, 5390–5402.
314. W. J. Lees; T. E. Benson; J. M. Hogle; C. T. Walsh, *Biochemistry* **1996**, 35, 1342–1351.
315. T. E. Benson; C. T. Walsh; J. M. Hogle, *Biochemistry* **1997**, 36, 806–811.
316. T. E. Benson; C. T. Walsh; V. Massey, *Biochemistry* **1997**, 36, 796–805.
317. C. Thorpe; C. H. Willaims, Jr., *Biochemistry* **1981**, 20, 1507–1513.
318. S. M. Miller; V. Massey; D. P. Ballou; C. H. Williams, Jr.; M. D. Distefano; M. J. Moore; C. T. Walsh, *Biochemistry* **1990**, 29, 2831–2841.
319. H. F. Gilbert, *Adv. Enzymol.* **1990**, 63, 69–172.
320. C. C. Bohme; L. D. Arscott; K. Becker; R. H. Schirmer; C. H. Williams, Jr., *J. Biol. Chem.* **2000**, 275, 37317–37323.
321. U. H. Danielson; F. Jiang; L. O. Hansson; B. Mannervik, *Biochemistry* **1999**, 38, 9258–9263.
322. M. P. Patel; J. Marcinkiewicz; J. S. Blanchard, *FEMS Microbiol. Lett.* **1998**, 166, 155–163.
323. E. Scott; I. W. Duncan; V. Eksrand, *J. Biol. Chem.* **1963**, 238, 3928–3933.
324. A. Bashir; R. N. Perham; N. S. Scrutton; A. Berry, *Biochem. J.* **1995**, 312, 527–533.
325. E. Davioud-Charvet; S. Delarue; C. Biot; B. Schwobel; C. C. Bohme; A. Mussigbrodt; L. Maes; C. Sergheraert; P. Grellier; R. H. Schirmer; K. A. Becker, *J. Med. Chem.* **2001**, 44, 4268–4276.
326. G. N. Sarma; S. N. Savvides; K. Becker; M. Schirmer; R. H. Schirmer; P. A. Karplus, *J. Mol. Biol.* **2003**, 328, 893–907.
327. W. Friebolin; B. Jannack; N. Wenzel; J. Furrer; T. Oeser; C. P. Sanchez; M. Lanzer; V. Yardley; K. Becker; E. Davioud-Charvet, *J. Med. Chem.* **2008**, 53, 1260–1277.
328. P. A. Karplus; G. E. Schulz, *J. Mol. Biol.* **1989**, 210, 163–180.
329. P. A. Karplus; G. E. Schulz, *J. Mol. Biol.* **1987**, 195, 701–729.
330. E. F. Pai; P. A. Karplus; G. E. Schulz, *Biochemistry* **1988**, 27, 4465–4474.
331. P. R. Mittl; G. E. Schulz, *Protein Sci.* **1994**, 3, 799–809.
332. E. F. Pai; G. E. Schulz, *J. Biol. Chem.* **1983**, 258, 1752–1757.
333. P. W. Huber; K. G. Brandt, *Biochemistry* **1980**, 19, 4568–4575.



334. P. Rietveld; L. D. Arscott; A. Berry; N. S. Scrutton; M. P. Deonarain; R. N. Perham; C. H. Williams, Jr., *Biochemistry* **1994**, *33*, 13888–13895.
335. P. A. Karplus; E. F. Pai; G. E. Schulz, *Eur. J. Biochem.* **1989**, *178*, 693–703.
336. K. K. Wong; M. A. Vanoni; J. S. Blanchard, *Biochemistry* **1988**, *27*, 7091–7096.
337. L. D. Arscott; D. M. Veine; C. H. Williams, Jr., *Biochemistry* **2000**, *39*, 4711–4721.
338. E. S. J. Arner; A. Holmgren, *Eur. J. Biol.* **2000**, *277*, 11521–11526.
339. W. H. Watson; X. Yang; Y. E. Choi; D. P. Jones; J. P. Kehrer, *Toxicol. Sci.* **2004**, *78*, 3–14.
340. K. Becker; S. Gromer; R. H. Schirmer; S. Müller, *Eur. J. Biol.* **2000**, *267*, 6118–6125.
341. Z. Krnajsiki; T. W. Gilberger; R. D. Walter; A. F. Cowman; S. Müller, *J. Biol. Chem.* **2002**, *277*, 25970–25975.
342. E. Davioud-Charvet; M. J. McLeish; D. M. Veine; D. Giegel; L. D. Arscott; A. D. Andricopulo; K. Becker; S. Müller; R. H. Schirmer; C. H. Williams, Jr.; G. L. Kenyon, *Biochemistry* **2003**, *42*, 13319–13330.
343. S. Gromer; S. Urig; K. Becker, *Med. Res. Rev.* **2004**, *24*, 40–89.
344. C. H. Williams, Jr.; L. D. Arscott; S. Müller; B. W. Lennon; M. L. Ludwig; P. F. Wang; D. M. Veine; K. Becker; R. H. Schirmer, *Eur. J. Biol.* **2000**, *267*, 6110–6117.
345. T. Tamura; T. C. Stadtman, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 1006–1011.
346. C. H. Williams, Jr., *FASEB J.* **1995**, *9*, 1267–1276.
347. B. W. Lennon; C. H. Williams, Jr.; M. L. Ludwig, *Science* **2000**, *289*, 1190–1194.
348. H. Bauer; V. Massey; L. D. Arscott; R. H. Schirmer; D. P. Ballou; C. H. Williams, Jr., *J. Biol. Chem.* **2003**, *278*, 33020–33028.
349. L. Zhong; A. Holmgren, *J. Biol. Chem.* **2000**, *275*, 18121–18128.
350. H. H. Huang; L. D. Arscott; D. P. Ballou; C. H. Williams, Jr., *Biochemistry* **2008**, *47*, 1721–1731.
351. V. Massey, *Biochim. Biophys. Acta* **1960**, *38*, 447–460.
352. A. Mattevi; A. J. Schierbeek; W. G. Hol, *J. Mol. Biol.* **1991**, *220*, 975–994.
353. A. Mattevi; G. Obmolova; J. R. Sokatch; C. Betzel; W. G. Hol, *Proteins* **1992**, *13*, 336–351.
354. A. Mattevi; G. Obmolova; K. H. Kalk; W. J. van Berkel; W. G. Hol, *J. Mol. Biol.* **1992**, *230*, 1200–1215.
355. K. R. Rajashankar; R. Bryk; R. Kniewel; J. A. Buglino; C. F. Nathan; C. D. Lima, *J. Biol. Chem.* **2005**, *280*, 33977–33983.
356. K. Maeda-Yordita; G. C. Russel; J. R. Guest; V. Massey; C. H. Williams, Jr., *Biochemistry* **1991**, *30*, 11788–11795.
357. V. Massey; Q. H. Gibson; C. Veegar, *Biochem. J.* **1960**, *77*, 341–451.
358. A. Arygyrou; J. S. Blanchard; B. A. Palfey, *Biochemistry* **2002**, *41*, 14580–14590.
359. R. G. Matthews; D. P. Ballou; C. H. Williams, Jr., *J. Biol. Chem.* **1979**, *254*, 4974–4981.
360. K. D. Wilkinson; C. H. Williams, Jr., *J. Biol. Chem.* **1979**, *254*, 852–862.
361. R. G. Matthews; C. H. Williams, Jr., *J. Biol. Chem.* **1976**, *251*, 3956–3964.
362. S. K. Akiyama; G. G. Hammes, *Biochemistry* **1981**, *20*, 1491–1497.
363. D. L. Coppock; C. Thorpe, *Antioxid. Redox Signal.* **2006**, *8*, 300–311.
364. K. L. Hooper; N. M. Glynn; J. Burnside; D. L. Coppock; C. Thorpe, *J. Biol. Chem.* **1999**, *274*, 31759–31762.
365. S. Rajce; C. Thorpe, *Biochemistry* **2003**, *42*, 4560–4568.
366. L. Ellgaard; L. W. Ruddock, *EMBO Rep.* **2005**, *6*, 28–32.
367. E. J. Heckler; P. C. Rancy; V. K. Kodali; C. Thorpe, *Biochim. Biophys. Acta* **2008**, *1783*, 567–577.
368. K. L. Hooper; B. Joneja; H. B. White, III; C. Thorpe, *J. Biol. Chem.* **1996**, *271*, 30510–30516.
369. K. L. Hooper; S. S. Sheasley; H. F. Gilbert; C. Thorpe, *J. Biol. Chem.* **1999**, *274*, 22147–22150.
370. A. O. Summers; L. I. Sugarman, *J. Bacteriol.* **1974**, *119*, 242–249.
371. B. Fox; C. T. Walsh, *J. Biol. Chem.* **1982**, *257*, 2498–2503.
372. B. Fox; C. T. Walsh, *Biochemistry* **1983**, *22*, 4082–4088.
373. N. L. Brown; S. J. Ford; R. D. Pridmore; D. C. Fritzinger, *Biochemistry* **1983**, *22*, 4089–4095.
374. R. L. Ledwidge; B. Patel; A. Dong; D. Fiedler; M. Falkowski; J. Zelikova; A. O. Summers; E. F. Pai; S. M. Miller, *Biochemistry* **2005**, *44*, 11402–11416.
375. N. Schiering; W. Kabsch; M. J. Moore; M. D. Distefano; C. T. Walsh; E. F. Pai, *Nature* **1991**, *352*, 168–172.
376. S. M. Miller; M. J. Moore; V. Massey; C. H. Williams, Jr.; M. D. Distefano; D. P. Ballou; C. T. Walsh, *Biochemistry* **1989**, *28*, 1194–1205.
377. A. Sandstorm; S. Lindskog, *Eur. J. Biol.* **1987**, *164*, 243–249.
378. S. M. Miller; D. P. Ballou; V. Massey; C. H. Williams, Jr.; C. T. Walsh, *J. Biol. Chem.* **1986**, *261*, 8081–8084.
379. S. Engst; S. M. Miller, *Biochemistry* **1998**, *37*, 11496–11507.
380. S. Engst; S. M. Miller, *Biochemistry* **1999**, *38*, 853–854.
381. S. Engst; S. M. Miller, *Biochemistry* **1999**, *38*, 3519–3529.
382. T. Stehle; S. A. Ahmed; A. Claiborne; G. E. Schulz, *J. Mol. Biol.* **1991**, *221*, 325–355.
383. E. J. Crane; D. Parsonage; A. Claiborne, *Biochemistry* **1996**, *35*, 2380–2387.
384. H. Miller; A. Claiborne, *J. Biol. Chem.* **1991**, *266*, 19342–19350.
385. H. S. Toogood; D. Leys; N. S. Scrutton, *FEBS J.* **2007**, *274*, 5481–5504.
386. D. L. Roberts; F. E. Frerman; J. J. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14355–14360.
387. K. K. Chohan; M. Jones; J. G. Grossmann; F. E. Frerman; N. S. Scrutton; M. J. Sutcliffe, *J. Biol. Chem.* **2001**, *276*, 34142–34147.
388. F. E. Frerman, *Biochem. Soc. Trans.* **1988**, *16*, 416–418.
389. E. J. Crane; D. Parsonage; L. B. Poole; A. Claiborne, *Biochemistry* **1995**, *34*, 14114–14124.
390. R. K. Olsen; B. S. Andersen; E. Christensen; P. Bross; F. Skovby; N. Gregersen, *Hum. Mutat.* **2003**, *22*, 12–23.
391. A. Curcoy; R. K. Olsen; A. Ribes; V. Trenchs; M. A. Vilaseca; J. Campistol; J. H. Osorio; B. S. Andersen; N. Gregersen, *Mol. Genet. Metab.* **2003**, *78*, 247–249.
392. J. D. Beckmann; F. E. Frefman, *Biochemistry* **1985**, *16*, 3913–3921.
393. J. Zhang; F. E. Frerman; J. J. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16212–16217.
394. J. D. Scott; R. A. Ludwig, *Microbiology* **2004**, *150*, 117–126.
395. D. Leys; J. Basran; F. Talfournier; M. J. Sutcliffe; N. S. Scrutton, *Nat. Struct. Biol.* **2003**, *10*, 219–225.
396. V. L. Davidson; M. Husain; J. W. Neher, *J. Bacteriol.* **1986**, *166*, 812–817.

397. D. L. Roberts; D. Salazar; J. P. Fulmer; F. E. Freman; J. J. Kim, *Biochemistry* **1999**, *38*, 1977–1989.
398. H. S. Toogood; A. van Thiel; J. Basran; M. J. Sutcliffe; N. S. Scrutton; D. Leys, *J. Biol. Chem.* **2004**, *279*, 32904–32912.
399. K. Y. Yang; R. P. Swenson, *Biochemistry* **2007**, *46*, 2289–2297.
400. K. Y. Yang; R. P. Swenson, *Biochemistry* **2007**, *46*, 2298–2305.
401. M. H. Jang; N. S. Scrutton; R. Hille, *J. Biol. Chem.* **2000**, *275*, 12546–12552.
402. J. Basran; K. K. Chohan; M. J. Sutcliffe; N. S. Scrutton, *Biochemistry* **2000**, *39*, 9188–9200.
403. L. Huang; R. J. Rohlfis; R. Hille, *J. Biol. Chem.* **1995**, *270*, 23958–23965.
404. J. Nagy; W. C. Kenney; T. P. Singer, *J. Biol. Chem.* **1979**, *254*, 2684–2688.
405. C. J. Batie; E. LaHaie; D. P. Ballou, *J. Biol. Chem.* **1987**, *262*, 1510–1528.
406. C. C. Correll; C. J. Batie; D. P. Ballou; M. L. Ludwig, *Science* **1992**, *258*, 1604–1610.
407. C. C. Correll; D. L. Gatti; M. L. Ludwig, Oxidized, Reduced, and Liganded States of the Iron-Sulfur Flavoprotein, Phthalate Dioxygenase Reductase. In *Flavin and Flavoproteins 1994*; K. Yagi, Ed.; Walter de Gruyter and Co: Berlin, 1994; pp 649–658.
408. G. Gassner; L. Wang; C. Batie; D. P. Ballou, *Biochemistry* **1994**, *33*, 12184–12193.
409. G. Gassner; D. P. Ballou, *Biochemistry* **1995**, *34*, 13460–13471.
410. G. Gassner; M. L. Ludwig; D. L. Gatti; C. C. Correll; D. P. Ballou, *FASEB J.* **1995**, *9*, 1411–1418.
411. M. Tarasev; D. P. Ballou, *Biochemistry* **2005**, *44*, 6197–6207.
412. M. Tarasev; D. P. Ballou, *Biochemistry* **2004**, *43*, 12799–12808.
413. O. Pylypenko; I. Schlichting, *Annu. Rev. Biochem.* **2004**, *73*, 991–1018.
414. A. J. Warman; O. Roitel; R. Neeli; H. M. Girvan; H. E. Seward; S. A. Murray; K. J. McLean; M. G. Joyce; H. Toogood; R. A. Holt; D. Leys; N. S. Scrutton; A. W. Munro, *Biochem. Soc. Trans.* **2005**, *33*, 747–753.
415. H. M. Girvan; T. N. Waltham; R. Neeli; H. F. Collins; K. J. McLean; N. S. Scrutton; D. Leys; A. W. Munro, *Biochem. Soc. Trans.* **2006**, *34*, 1173–1177.
416. T. Iyanagi; H. S. Mason, *Biochemistry* **1973**, *12*, 2297–2308.
417. T. Iyanagi; N. Makino; H. S. Mason, *Biochemistry* **1974**, *13*, 1701–1710.
418. J. L. Vermilion; M. J. Coon, *J. Biol. Chem.* **1978**, *253*, 2694–2704.
419. J. L. Vermilion; D. P. Ballou; V. Massey; M. J. Coon, *J. Biol. Chem.* **1981**, *256*, 266–277.
420. B. S. Masters; M. H. Bilimoria; H. Kamin; Q. H. Gibson; *J. Biol. Chem.* **1965**, *240*, 4081–4088.
421. M. Wang; D. L. Roberts; R. Paschke; T. M. Shea; B. S. Masters; J. J. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8411–8416.
422. A. Gutierrez; M. Paine; C. R. Wolf; N. S. Scrutton; G. C. Roberts, *Biochemistry* **2002**, *41*, 4626–4637.
423. A. Gutierrez; A. W. Munro; A. Grunau; C. R. Wolf; N. S. Scrutton; G. C. Roberts, *Eur. J. Biochem.* **2003**, *270*, 2612–2621.
424. R. T. Ruettinger; L. P. Wen; A. J. Fulco, *J. Biol. Chem.* **1989**, *264*, 10987–10995.
425. A. W. Munro; D. G. Leys; K. J. McLean; K. R. Marshall; T. W. Ost; S. Daff; C. S. Miles; S. K. Chapman; D. A. Lysek; C. C. Moser; C. C. Page; P. L. Dutton, *Trends Biochem. Sci.* **2002**, *27*, 250–257.
426. M. A. Noble; C. S. Miles; S. K. Chapman; D. A. Lysek; A. C. MacKay; G. A. Reid; R. P. Hanzlik; A. W. Munro, *Biochem. J.* **1999**, *339*, 371–379.
427. A. W. Munro; S. Daff; J. R. Coggins; J. G. Lindsay; S. K. Chapman, *Eur. J. Biochem.* **1996**, *239*, 403–409.
428. P. A. Hubbard; A. L. Shen; R. Paschke; C. B. Kasper; J. J. Kim, *J. Biol. Chem.* **2001**, *276*, 29163–29170.
429. J. R. Lakowicz; G. Weber, *Biochemistry* **1973**, *12*, 4171–4179.
430. V. Massey, *J. Biol. Chem.* **1994**, *269*, 22459–22462.
431. T. C. Bruice, *Isr. J. Chem.* **1984**, *24*, 54–61.
432. R. F. Anderson, Flavin-Oxygen Complex Formed on the Reaction of Superoxide Ions with Flavosemiquinone Radicals. In *Flavins and Flavoproteins*; V. Massey, C. H. Williams, Eds.; Elsevier North-Holland: New York, 1982; pp 278–283.
433. R. R. Ramsay; S. C. Koerber; T. P. Singer, *Biochemistry* **1987**, *26*, 3045–3050.
434. J. Sucharitakul; M. Prongjit; D. Haltrich; P. Chaiyen, *Biochemistry* **2008**, *47*, 8485–8490.
435. V. Massey; L. M. Schopfer; R. F. Anderson, Structural Determinants of the Oxygen Reactivity of Different Classes of Flavoproteins. In *Oxidases and Related Redox Systems*; T. E. King, H. S. Mason, M. Morrison, Eds.; Alan Liss: New York, 1988; pp 147–166.
436. A. Mattevi, *Trends Biochem. Sci.* **2006**, *31*, 276–283.
437. R. Wang; C. Thorpe, *Biochemistry* **1991**, *30*, 7895–7901.
438. J. P. Klinman, *Acc. Chem. Res.* **2007**, *40*, 325–333.
439. G. Zhao; R. C. Bruckner; M. S. Jorns, *Biochemistry* **2008**, *47*, 9124–9135.
440. K. Otto; K. Hofstetter; M. Röthlisberger; B. Witholt; A. Schmid, *J. Bacteriol.* **2004**, *186*, 5292–5302.
441. A. Kantz; F. Chin; N. Nallamotheu; T. Nguyen; G. T. Gassner, *Arch. Biochem. Biophys.* **2005**, *442*, 102–116.
442. B. P. Laden; Y. Tang; T. D. Porter, *Arch. Biochem. Biophys.* **2000**, *374*, 381–388.
443. D. P. Ballou; B. Entsch; L. J. Cole, *Biochem. Biophys. Res. Commun.* **2005**, *338*, 590–598.
444. B. Entsch; L. J. Cole; D. P. Ballou, *Arch. Biochem. Biophys.* **2005**, *433*, 297–311.
445. M. Husain; V. Massey, *J. Biol. Chem.* **1979**, *254*, 6657–6666.
446. B. A. Palfey; B. Entsch; D. P. Ballou; V. Massey, *Biochemistry* **1994**, *33*, 1545–1554.
447. D. L. Gatti; B. A. Palfey; M. S. Lah; B. Entsch; V. Massey; D. P. Ballou; M. L. Ludwig, *Science* **1994**, *266*, 110–114.
448. B. A. Palfey; R. Basu; K. K. Frederick; B. Entsch; D. P. Ballou, *Biochemistry* **2002**, *41*, 8438–8446.
449. K. K. Frederick; B. A. Palfey, *Biochemistry* **2005**, *44*, 13304–13314.
450. B. Entsch; V. Massey; D. P. Ballou, *J. Biol. Chem.* **1976**, *251*, 2550–2563.
451. B. A. Palfey; G. R. Moran; B. Entsch; D. P. Ballou; V. Massey, *Biochemistry* **1999**, *38*, 1153–1158.
452. B. Entsch; B. A. Palfey; D. P. Ballou; V. Massey, *J. Biol. Chem.* **1991**, *266*, 17341–17349.
453. M. Ortiz-Maldonado; D. P. Ballou; V. Massey, *Biochemistry* **1999**, *38*, 8124–8137.
454. M. Nadella; M. A. Bianchetti; S. B. Gabelli; J. Barrila; L. M. Amzel, *Proc. Natl. Acad. U.S.A.* **2005**, *102*, 16830–16835.
455. C. Siebold; N. Berrow; T. S. Walter; K. Harlos; R. J. Owens; D. I. Stuart; J. R. Terman; A. L. Kolodkin; R. J. Pasterkamp; E. Y. Jones, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16836–16841.
456. E. F. Schmidt; S.-O. Shim; S. M. Strittmatter, *J. Neurosci.* **2008**, *28*, 2287–2297.

457. K.-H. Van Pée; E. P. Patallo, *Appl. Microbiol. Biotechnol.* **2006**, *70*, 631–641.
458. E. Yeh; L. J. Cole; E. W. Barr; J. M. Bollinger; D. P. Ballou; C. T. Walsh, *Biochemistry* **2006**, *45*, 7904–7912.
459. C. Dong; S. Flecks; S. Uverschuch; C. Haupt; K.-H. van Pée; J. H. Naismith, *Science* **2005**, *309*, 2216–2219.
460. E. Bitto; Y. Huang; C. A. Bingman; S. Singh; J. S. Thorson; G. N. Phillips, *Proteins* **2008**, *70*, 289–293.
461. E. Yeh; L. C. Blasiak; A. Koglin; C. L. Drennan; C. T. Walsh, *Biochemistry* **2007**, *46*, 1284–1292.
462. J. Valton; C. Mathevon; M. Fontecave; V. Nivière; D. P. Ballou, *J. Biol. Chem.* **2008**, *283*, 10287–10296.
463. S. K. Krueger; D. E. Williams, *Pharm. Ther.* **2005**, *106*, 357–387.
464. H. S. Choi; J. K. Kim; E. H. Cho; Y. C. Kim; J. I. Kim; S. W. Kim, *Biochem. Biophys. Res. Commun.* **2003**, *306*, 930–936.
465. N. Schlaich, *Trends Pharmacol. Sci.* **2007**, *12*, 1360–1385.
466. N. B. Beaty; D. P. Ballou, *J. Biol. Chem.* **1980**, *255*, 3817–3819.
467. N. B. Beaty; D. P. Ballou, *J. Biol. Chem.* **1981**, *256*, 4611–4618.
468. N. B. Beaty; D. P. Ballou, *J. Biol. Chem.* **1981**, *256*, 4619–4625.
469. K. C. Jones; D. P. Ballou, *J. Biol. Chem.* **1986**, *261*, 2553–2559.
470. A. Alfieri; E. Malito; R. Orru; M. W. Fraaije; A. Mattevi, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6572–6577.
471. S. Eswaramoorthy; J. B. Bonanno; S. K. Burley; S. Swaminathan, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9832–9837.
472. M. W. Fraaije; N. M. Kamerbeek; W. J. H. van Berkel; D. B. Janssen, *FEBS Lett.* **2002**, *518*, 43–47.
473. M. Gibson; M. Nur-e-alam; F. Lipata; M. A. Oliveira; J. Rohr, *J. Am. Chem. Soc.* **2005**, *127*, 17594–17595.
474. H. Iwaki; S. Wang; S. Grosse; H. Bergerson; A. Nagahashi; J. Lertvorachon; J. Yang; Y. Konishi; Y. Hasegawa; P. C. K. Lau, *Appl. Environ. Microbiol.* **2006**, *72*, 2707–2720.
475. C. C. Ryerson; D. P. Ballou; C. T. Walsh, *Biochemistry* **1982**, *21*, 2644–2655.
476. D. Sheng; D. P. Ballou; V. Massey, *Biochemistry* **2001**, *40*, 11156–11167.
477. D. E. Torres Pazmiño; B.-J. Baas; D. B. Janssen; M. W. Fraaije, *Biochemistry* **2008**, *47*, 4082–4093.
478. E. Malito; A. Alfieri; M. W. Fraaije; A. Mattevi, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13157–13162.
479. S.-C. Tu, *Photochem. Photobiol. Sci.* **2008**, *7*, 183–188.
480. S. Nijvipakul; J. Wongratana; C. Suadee; B. Entsch; D. P. Ballou; P. Chaiyen, *J. Bacteriol.* **2008**, *190*, 1531–1538.
481. J. Vervoort; F. Müller; J. Lee; W. A. M. van der Berg; C. T. Moonen, *Biochemistry* **1986**, *25*, 8067–8075.
482. B. Lei; Q. Ding; S.-C. Tu, *Biochemistry* **2004**, *43*, 15975–15982.
483. J. W. Eckstein; J. W. Hastings; S. Ghisla, *Biochemistry* **1993**, *32*, 404–411.
484. E. J. Nelson; H. S. Tunsjø; P. M. Fidopiastis; H. Sørum; E. G. Ruby, *Appl. Environ. Microbiol.* **2007**, *73*, 1825–1833.
485. S. Bornemann, *Nat. Prod. Rep.* **2002**, *19*, 761–772.
486. A. Razeto; F. Mattioli; E. Carpanelli; A. Aliverti; V. Pandini; A. Coda; A. Mattevi, *Structure* **2007**, *15*, 683–692.
487. A. K. Hajra, *Biochem. Biophys. Res. Commun.* **1970**, *39*, 1037–1044.
488. R. L. Wykle; C. Plantadosi; F. Snyder, *J. Biol. Chem.* **1972**, *247*, 2944–2948.
489. E. C. de Vet; H. van den Bosch, *Cell Biochem. Biophys.* **2000**, *32*, 117–121.
490. E. C. de Vet; Y. H. Hilkes; M. W. Fraaije; H. van den Bosch, *J. Biol. Chem.* **2000**, *275*, 6276–6283.
491. A. J. Brown; F. Snyder, *J. Biol. Chem.* **1982**, *257*, 8835–8839.
492. S. J. Friedberg; S. T. Weintraub; M. R. Singer; R. C. Greene, *J. Biol. Chem.* **1983**, *258*, 136–142.
493. A. J. Brown; G. L. Glish; E. H. McBay; F. Snyder, *Biochemistry* **1985**, *24*, 8012–8016.
494. G. Stevenson; B. Neal; D. Liu; M. Hobbs; N. H. Packer; M. Batley; J. W. Redmond; L. Linquist; P. J. Reeves, *J. Bacteriol.* **1994**, *176*, 4144–4156.
495. P. M. Nassau; S. L. Martin; R. E. Brown; A. Weston; D. Monsey; M. R. McNeil; K. Duncan, *J. Bacteriol.* **1996**, *178*, 1047–1052.
496. K. Duncan, *Curr. Pharm. Des.* **2004**, *10*, 3185–3194.
497. M. S. Scherman; K. A. Winans; R. J. Stern; V. Jones; C. R. Bertozzi; M. R. McNeil, *Antimicrob. Agents Chemother.* **2003**, *47*, 378–382.
498. D. A. R. Sanders; A. G. Staines; S. A. McMahon; M. R. McNeil; C. Whitfield; J. H. Naismith, *Nat. Struct. Biol.* **2001**, *8*, 858–863.
499. K. Beis; V. Srikanthasan; H. Liu; S. W. B. Fullerton; V. Bamford; D. A. R. Sanders; C. Whitfield; M. R. McNeil; J. H. Naismith; *J. Mol. Biol.* **2005**, *348*, 971–982.
500. J. M. Chad; K. P. Sarathy; T. D. Gruber; E. Addala; L. L. Kiessling; D. A. R. Sanders, *Biochemistry* **2007**, *46*, 6723–6732.
501. M. Soltero-Higgin; E. E. Carlson; T. D. Gruber; L. L. Kiessling, *Nat. Struct. Mol. Biol.* **2004**, *11*, 539–543.
502. Z. Huang; Q. Zhang; H. Liu, *Bioorg. Chem.* **2003**, *31*, 494–502.
503. J. N. Barlow; M. E. Girvin; J. S. Blanchard, *J. Am. Chem. Soc.* **1999**, *121*, 6968–6969.
504. K. Itoh; Z. Huang; H. Liu, *Org. Lett.* **2007**, *9*, 879–882.
505. E. Strauss; C. Kinsland; Y. Ge; F. W. McLafferty; T. P. Begley, *J. Biol. Chem.* **2001**, *276*, 12513–12516.
506. T. Kupke; P. Hernandez-Acosta; S. Steinbacher; F. A. Culiñez-Macia, *J. Biol. Chem.* **2001**, *276*, 19190–19196.
507. E. Strauss; H. Zhai; L. A. Brand; F. W. McLafferty; T. P. Begley, *Biochemistry* **2004**, *43*, 15520–15533.
508. E. Strauss; T. P. Begley, *J. Am. Chem. Soc.* **2001**, *123*, 6449–6450.
509. E. Strauss; T. P. Begley, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 339–342.
510. P. Hernandez-Acosta; D. G. Schmid; G. Jung; F. A. Culiñez-Macia; T. Kupke, *J. Biol. Chem.* **2002**, *277*, 20490–20498.
511. S. Steinbacher; P. Hernandez-Acosta; B. Bieseler; M. Blaesse; R. Huber; F. A. Culiñez-Macia; T. Kupke, *J. Mol. Biol.* **2003**, *327*, 192–202.
512. H. G. Floss; D. K. Onderka; M. Carroll, *J. Biol. Chem.* **1972**, *247*, 736–744.
513. H. Morell; M. J. Clark; P. F. Knowles; D. B. Sprinson, *J. Biol. Chem.* **1967**, *242*, 82–90.
514. K. Kitzing; S. Auweter; N. Amrhein; P. Macheroux, *J. Biol. Chem.* **2004**, *279*, 9451–9461.
515. P. Macheroux; E. Schonbrunn; D. I. Svergun; V. V. Volkov; M. H. J. Koch; S. Bornemann; R. N. F. Thorneley, *Biochem. J.* **1998**, *335*, 319–327.
516. J. Maclean; S. Ali, *Structure* **2003**, *11*, 1499–1511.
517. P. Macheroux; J. Peterson; S. Bornemann; D. J. Lowe; R. N. F. Thorneley, *Biochemistry* **1996**, *35*, 1643–1652.
518. S. Bornemann; M. K. Ramjee; S. Balasubramanian; C. Abell; J. R. Coggins; D. J. Lowe; R. N. F. Thorneley, *J. Biol. Chem.* **1995**, *270*, 22811–22815.

519. T. Carell; L. T. Burgdorf; L. M. Kundu; M. Cichon, *Curr. Opin. Chem. Biol.* **2001**, *5*, 491–498.
520. M. K. Cichon; S. Arnold; T. Carell, *Angew. Chem. Int. Ed.* **2002**, *41*, 767–770.
521. G. B. Sancar, *Mutat. Res.* **1990**, *236*, 147–160.
522. A. Yasui; A. P. M. Eker; S. Yasuhira; H. Yajima; T. Kobayashi; M. Oikawa; A. Oikawa, *EMBO J.* **1994**, *13*, 6143–6151.
523. S. Kanai; R. Kikuno; H. Toh; H. Ryo; T. Todo, *J. Mol. Evol.* **1997**, *45*, 535–548.
524. K. Hitomi; K. Okamoto; H. Daiyasu; H. Miyashita; S. Iwai; H. Toh; M. Ishiura; T. Todo, *Nucleic Acids Res.* **2000**, *28*, 2353–2362.
525. S. Yasuhira; A. Yasui, *J. Biol. Chem.* **1992**, *267*, 25644–25647.
526. T. Todo; H. Ryo; H. Takemori; H. Toh; T. Nomura; S. Kondo, *Mutat. Res.* **1994**, *315*, 213–228.
527. A. Sancar, *Biochemistry* **1994**, *33*, 2–9.
528. H.-W. Park; S.-T. Kim; A. Sancar; J. Deisenhofer, *Science* **1995**, *268*, 1866–1872.
529. T. Tamada; K. Kitadokoro; Y. Higuchi; K. Inaka; A. Yasui; P. E. de Ruiter; A. P. M. Eker; K. Miki, *Nat. Struct. Biol.* **1997**, *4*, 887–891.
530. H. Komori; R. Masui; S. Kuramitsu; S. Yokoyama; T. Shibata; Y. Inoue; K. Miki, *J. Am. Chem. Soc.* **1999**, *121*, 5127–5134.
531. G. B. Sancar; M. S. Jorns; D. J. Payne; D. J. Fluke; C. S. Rupert; A. Sancar, *J. Biol. Chem.* **1987**, *262*, 492–498.
532. P. F. Heelis; A. Sancar, *Biochemistry* **1986**, *25*, 8163–8166.
533. M. S. Jorns; E. T. Baldwin; G. B. Sancar; A. Sancar, *J. Biol. Chem.* **1987**, *262*, 486–491.
534. G. Payne; P. F. Heelis; B. R. Rohrs; A. Sancar, *Biochemistry* **1987**, *26*, 7121–7127.
535. P. F. Heelis; G. Payne; A. Sancar, *Biochemistry* **1987**, *26*, 4634–4640.
536. M. S. Jorns; B. Wang; S. P. Jordan, *Biochemistry* **1987**, *26*, 6810–6818.
537. G. Payne; A. Sancar, *Biochemistry* **1990**, *29*, 7715–7727.
538. M. S. Jorns; B. Wang; S. P. Jordan; L. P. Chanderkar, *Biochemistry* **1990**, *29*, 552–561.
539. R. S. A. Lipman; M. S. Jorns, *Biochemistry* **1992**, *31*, 786–791.
540. S. T. Kim; P. F. Heelis; A. Sancar, *Biochemistry* **1992**, *45*, 11244–11248.
541. S. Weber, *Biochim. Biophys. Acta* **2005**, *1707*, 1–23.
542. H. Griengl; H. Schwab; M. Fechter, *Trends Biotechnol.* **2000**, *18*, 252–256.
543. M. S. Jorns, *Biochim. Biophys. Acta* **1980**, *613*, 203–209.
544. M. S. Jorns, *J. Biol. Chem.* **1979**, *254*, 12145–12152.
545. D. Vargo; A. Pokora; S. W. Wang; M. S. Jorns, *J. Biol. Chem.* **1981**, *256*, 6027–6033.
546. R. G. Duggleby; J. A. McCourt; L. W. Guddat, *Plant Physiol. Biochem.* **2008**, *46*, 309–324.
547. J. Schloss; L. Ciskanik; E. Pai; C. Thorpe, Acetolactate Synthase: A Deviant Flavoprotein. In *Flavins and Flavoproteins 1990*; B. Curti, S. Ronchi, G. Zanetti, Eds.; Walter de Gruyter: Berlin, 1991; pp 907–914.
548. R. G. Duggleby, *Acc. Chem. Res.* **2006**, *39*, 550–557.
549. H. Chen; Z. Guo; H.-W. Liu, *J. Am. Chem. Soc.* **1998**, *120*, 11796–11797.
550. H. V. Thulasiram; H. K. Erickson; C. D. Poulter, *Science* **2007**, *316*, 73–76.
551. T. Kuzuyama; H. Seto, *Nat. Prod. Rep.* **2003**, *20*, 171–183.
552. B. W. Agronoff; H. Eggerer; J. Henning; F. Lynen, *J. Am. Chem. Soc.* **1959**, *81*, 1254–1255.
553. J. W. Cornforth; G. Popjak, *Methods Enzymol.* **1969**, *15*, 359–371.
554. K. Kaneda; T. Kuzuyama; M. Takagi; Y. Hayakawa; H. Seto, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 932–937.
555. H. Hemmi; Y. Ikeda; S. Yamashita; T. Nakayama; T. Nishino, *Biophys. Res. Commun.* **2004**, *322*, 905–910.
556. S. C. Rothman; T. R. Helm; C. D. Poulter, *Biochemistry* **2007**, *46*, 5437–5445.
557. S. C. Rothman; J. B. Johnston; S. Lee; J. R. Walker; C. D. Poulter, *J. Am. Chem. Soc.* **2007**, *130*, 4906–4913.
558. J. R. Walker; S. C. Rothman; C. D. Poulter, *J. Org. Chem.* **2008**, *73*, 726–729.
559. J. de Rudyck; J. Pouyez; S. C. Rothman; D. Poulter; J. Wouters, *Biochemistry* **2008**, *47*, 9051–9053.
560. W. Kittleman; C. J. Thibodeaux; Y. Liu; H. Zhang; H.-W. Liu, *Biochemistry* **2007**, *46*, 8401–8413.
561. D. W. Tempest; J. L. Meers; C. M. Brown, *Biochem. J.* **1970**, *117*, 405–407.
562. P. J. Lea; S. A. Robinson; G. R. Stewart, The Enzymology and Metabolism of Glutamine, Glutamate, and Asparagine. In *The Biochemistry of Plants*; B. J. Milfin, P. J. Lea, Eds.; Academic Press: New York, 1990; Vol. 16, pp 121–159.
563. S. J. Temple; C. P. Vance; J. S. Gantt, *Trends Plant Sci.* **1998**, *3*, 51–56.
564. P. J. Lea; R. J. Ireland, Nitrogen Metabolism in Higher Plants. In *Plant Amino Acids: Biochemistry and Biotechnology*; B. K. Singh, Ed.; Marcel Dekker: New York, 1999; pp 1–47.
565. L. J. Rietzer, Ammonia Assimilation and the Biosynthesis of Glutamine, Glutamate, Aspartate, Asparagine, L-Alanine, and D-Alanine. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*; F. C. Neidhart, Ed.; ASM Press: Washington, DC, 1996; Vol. 1, pp 391–407.
566. M. A. Vanoni; B. Curti, *Cell. Mol. Life Sci.* **1999**, *55*, 617–638.
567. R. H. H. van den Heuvel; B. Curti; M. A. Vanoni; A. Mattevi, *Cell. Mol. Life Sci.* **2004**, *61*, 669–681.
568. M. A. Vanoni; B. Curti, *Arch. Biochem. Biophys.* **2005**, *433*, 193–211.
569. M. A. Vanoni; P. Accornero; G. Carrera; B. Curti, *Arch. Biochem. Biophys.* **1994**, *309*, 222–230.
570. M. A. Vanoni; D. E. Edmondson; M. Rescigno; G. Zanetti; B. Curti, *Biochemistry* **1991**, *30*, 11478–11484.
571. M. A. Vanoni; F. Fischer; E. Ravasio; S. Verzotti; D. E. Edmondson; W. R. Hagen; G. Zanetti; B. Curti, *Biochemistry* **1998**, *37*, 1828–1838.
572. P. Mäntsälä; H. Zalkin, *J. Biol. Chem.* **1970**, *251*, 3294–3299.
573. S. Ravasio; L. Dossena; E. Martin-Figueroa; F. J. Florencio; A. Mattevi; P. Morandi; B. Curti; M. A. Vanoni, *Biochemistry* **2002**, *41*, 8120–8133.
574. J. Kantola; T. Kunnari; A. Hautala; J. Hakala; K. Ylihonko; P. Mäntsälä, *Microbiology* **2000**, *146*, 155–163.
575. I. Fugii; Y. Ebizuka, *Chem. Rev.* **1997**, *97*, 2511–2524.
576. K. Ylihonko; J. Hakala; J. Niemi; J. Lundell; P. Mäntsälä, *Microbiology* **1994**, *140*, 1359–1365.
577. A. Yoshimoto; T. Ogasawara; I. Kitamura; T. Oki; T. Takeuchi; H. Umezawa, *J. Antibiot. (Tokyo)* **1979**, *32*, 472–481.
578. I. Alexeev; A. Sultana; P. Mäntsälä; J. Niemi; G. Schneider, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6170–6175.
579. R. H. Cheney, *Q. J. Crude Drug Res.* **1964**, *3*, 413–416.

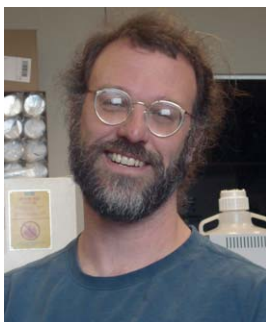


580. D. H. R. Barton; R. H. Hesse; G. W. Kirby, *Proc. Chem. Soc. (London)* **1963**, 1, 267–268.
581. T. M. Kutchan; H. Dittirch, *J. Biol. Chem.* **1995**, 270, 24475–24481.
582. A. Winkler; A. Lyskoski; S. Riedl; M. Puhl; T. M. Kutchan; P. Macheroux; K. Gruber, *Nat. Chem. Biol.* **2008**, 4, 739–741.
583. A. Winkler; F. Hartner; T. M. Kutchan; A. Glieder; P. Macheroux, *J. Biol. Chem.* **2006**, 281, 21276–21285.
584. A. Winkler; T. M. Kutchan; P. Macheroux, *J. Biol. Chem.* **2007**, 282, 24437–24443.
585. R. Buder; G. Fuchs, *Eur. J. Biochem.* **1989**, 185, 629–635.
586. B. Langkau; P. Vock; V. Massey; G. Fuchs; S. Ghilsa, *Eur. J. Biochem.* **1995**, 230, 676–685.
587. B. Langkau; S. Ghilsa; R. Buder; K. Ziegler; G. Fuchs, *Eur. J. Biochem.* **1990**, 191, 365–371.
588. B. Langkau; S. Ghilsa, *Eur. J. Biochem.* **1995**, 230, 686–697.
589. B. Langkau, Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften, Universität Konstanz, 1993; p 32.

### Biographical Sketches



Rebecca L. Fagan did her undergraduate work at Indiana State University in Terre Haute, IN, where she obtained a B.S. in chemistry. She then came to the University of Michigan to pursue her graduate studies in biological chemistry. She joined Dr. Bruce A. Palfey's laboratory in 2005, began her work on the differences in the mechanism of dihydroorotate oxidation by Class 1A and Class 2 dihydroorotate dehydrogenases, and received her Ph.D. in 2009. She is now a postdoctoral researcher at the University of Iowa.



Bruce A. Palfey did his undergraduate work at Pennsylvania State University where he majored in biochemistry. After few years as a technician, he earned an M.S. in chemistry at Drexel University and then a Ph.D. in biological chemistry at the University of Michigan working jointly in the laboratories of Professors David Ballou and Vincent Massey. After postdoctoral work in the laboratories of Vincent Massey and David Ballou, he became a lecturer at the University of Michigan and, in 2003, an Assistant Professor. Research in his laboratory is focused on flavoenzymes in pyrimidine metabolism.

## 7.04 Biotechnology of Riboflavin Production

Hans-Peter Hohmann, DSM Nutritional Products, Kaiseraugst, Switzerland

Klaus-Peter Stahmann, Fachhochschule Lausitz, Senftenberg, Germany

© 2010 Elsevier Ltd. All rights reserved.

---

7.04.1	Scope	115
7.04.2	Introduction	116
7.04.3	Industrial Production	117
7.04.3.1	Commercial Aspects	117
7.04.3.2	Chemical Riboflavin Synthesis	117
7.04.3.3	Microbial Riboflavin Production	119
7.04.4	Riboflavin Biosynthesis	120
7.04.4.1	Riboflavin Biosynthesis in Bacteria and Fungi	120
7.04.4.2	Riboflavin Biosynthetic Genes	120
7.04.4.2.1	<i>Bacillus subtilis</i> and other bacteria	120
7.04.4.2.2	Fungi	120
7.04.4.3	Regulation of <i>rib</i> Gene Expression	122
7.04.4.3.1	Bacteria	122
7.04.4.3.2	Fungi	123
7.04.5	Construction of Riboflavin Production Strains	123
7.04.5.1	<i>Bacillus subtilis</i> and <i>Ashbya gossypii</i> as Preferred Host Strains	123
7.04.5.2	Riboflavin Pathway Engineering	124
7.04.5.2.1	<i>Bacillus subtilis</i> production strains: <i>rib</i> genes driven by their natural promoters	124
7.04.5.2.2	<i>Bacillus subtilis</i> production strains: <i>rib</i> gene expression driven by heterologous promoters	125
7.04.5.2.3	<i>Corynebacterium ammoniagenes</i>	126
7.04.5.2.4	<i>Ashbya gossypii</i>	127
7.04.5.3	Metabolic Pathway Engineering to Improve Riboflavin Precursor Supply	127
7.04.5.3.1	Purine and guanosine pathway, nitrogen metabolism in <i>Bacillus subtilis</i>	127
7.04.5.3.2	Purine and glycine metabolism in <i>Ashbya gossypii</i>	128
7.04.5.3.3	Pentose phosphate pathway	128
7.04.5.3.4	Modifications in the central metabolism of <i>Bacillus subtilis</i> production strains	129
7.04.5.3.5	Isocitrate lyase and isocitrate dehydrogenase in <i>Ashbya gossypii</i> production strains	130
7.04.5.4	Riboflavin Transmembrane Transport	130
7.04.5.5	Energetic Efficiency of the <i>Bacillus subtilis</i> Respiratory Chain	131
7.04.5.6	Classical Strain Improvement and Whole Genome Sequencing	131
7.04.6	Process Development	132
7.04.6.1	Glucose and Biotin-Limited Fed-Batch Fermentation Process for <i>Bacillus subtilis</i> Production Strains	132
7.04.6.2	<i>Ashbya gossypii</i> Fermentation Process with Plant Oils as Carbon and Energy Source	133
7.04.6.3	Product Isolation and Purification	133
7.04.7	Conclusions	135
References		136

---

### 7.04.1 Scope

Microbial riboflavin (vitamin B<sub>2</sub>) production processes converting commodity fermentation substrates like monomeric or oligomeric carbohydrates or vegetable oils to riboflavin are the focus of the present review. The current developmental status of the riboflavin production strains derived from metabolic pathway engineering

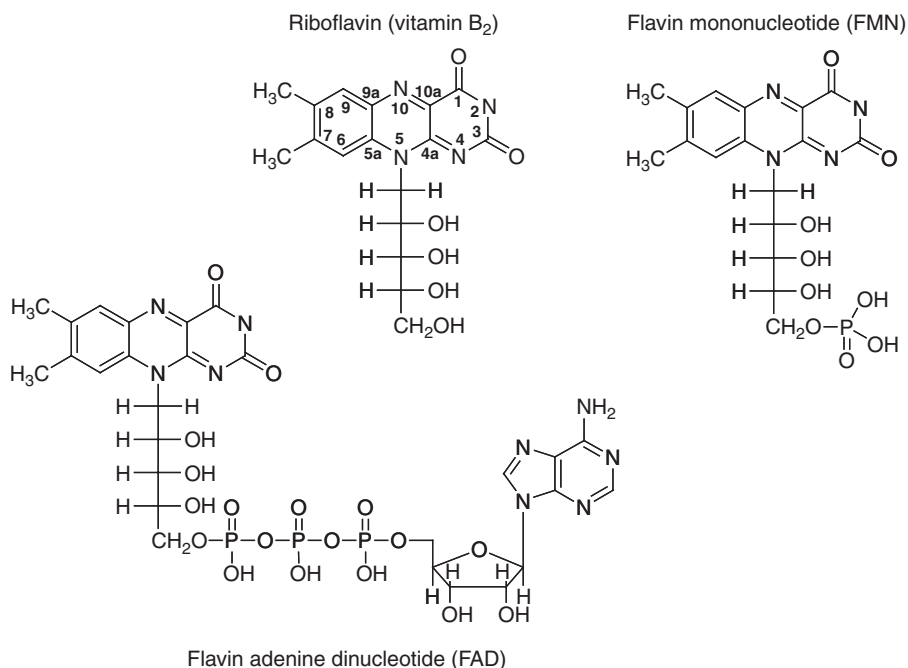


programs, frequently supported by classical strain improvement campaigns, is described in detail. The topics covered in this chapter are (1) the riboflavin biosynthetic genes and the regulation of their expression in the microorganisms relevant as production hosts, (2) the engineering of the riboflavin pathways in these organisms to attain massive *rib* gene overexpression, (3) modifications in the biosynthetic routes to GTP and ribulose-5-phosphate, the two immediate riboflavin precursors, (4) modifications in the central metabolism of the host strains to optimize metabolic building block supply, (5) riboflavin import and export, and (6) energetic aspects. Fermentation process development and product isolation and purification are also addressed. The chemical production processes providing riboflavin for decades, but now almost abandoned, are outlined. The layout of riboflavin biosynthesis in microorganisms and the fascinating enzymes constituting the riboflavin biosynthetic pathway are only briefly touched. References to extensive reviews are provided.

The present review is derived from publicly available academic and patent literature. Occasionally, unpublished observations are included where appropriate. The review incorporates mainly literature published in the English language during the last 10–15 years. This includes important contributions from Russian scientists. More recent literature from Chinese authors are incorporated as well. It is recognized that a vast body of Chinese literature particularly on fungal riboflavin production strains exists. Certainly interesting aspects of microbial riboflavin production are communicated therein, but unfortunately, because of language barriers, this literature could not be included in the present review.

## 7.04.2 Introduction

Around 1880 a yellow, fluorescing pigment was isolated from whey by the English chemist A. W. Blyth. Later, the pigment was isolated from other biological sources as well and designated as lactochrom, ovaflavin, lactoflavin, and others. Kuhn, Weygand, and Karrer determined the structure (Figure 1) of the yellow pigment and proved it by chemical synthesis in 1933 and 1934. It became clear that all of the yellow compounds mentioned above were identical, for which the common name riboflavin was chosen. Riboflavin is biosynthesized in plants and in many microorganisms. Vegetables and milk are major sources of the vitamin in human



**Figure 1** Chemical structures of riboflavin (vitamin B<sub>2</sub>) and the flavin coenzymes flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD).

nutrition. One cup of milk per day is sufficient to cover the recommended dosage of 2 mg per day for humans. Ruminants can derive vitamin B<sub>2</sub> from their intestinal flora.

The first enzyme that was identified in the early 1930s to use a flavin compound as prosthetic group was Warburg's old yellow enzyme OYE from yeast. In fact, the basic biochemical concept of a holoenzyme consisting from a proteinaceous apoenzyme and a relative small organic compound, the coenzyme, in the reaction center of the holoenzyme was developed from studies with OYE. Also in the 1930s two other flavoproteins were identified, D-amino acid oxidase and liponamid dehydrogenase. Theorell found out that the flavin associated to OYE was not riboflavin itself but flavin adenine mononucleotide (FMN) synthesized from the biological precursor riboflavin (Figure 1) by phosphorylation. Later it became clear that the coenzymes of all flavoproteins are either FMN or flavin adenine dinucleotide (FAD). The flavin nucleotides are noncovalently or covalently bound to their apoenzyme. Covalent linkage occurs via cysteinyl, histidyl, or tyrosyl residues of the apoprotein and the positions 6 or 8a of the flavin. With some remarkable (see below) exceptions, flavoenzymes are redox catalysts capable of both one- and two-electron transfer processes. They play a pivotal role in coupling the two-electron oxidation of most organic substrates to the one-electron transfers of the respiratory chain. Flavoproteins, such as L-lactate monooxygenase or D-amino acid oxidase, catalyze the oxidation of  $\alpha$ -hydroxy acids and  $\alpha$ -amino acids via a carbanion mechanism. Flavoprotein disulfide reductases such as lipoyl dehydrogenase and glutathione reductase involve a hydride transfer mechanism from NAD(P)H and a flavin C4a-cysteinyl adduct. Flavoprotein monooxygenases, like *p*-hydroxybenzoate-hydroxylase, bind molecular oxygen to the reduced form of their flavin coenzyme to form a peroxide, which is responsible for the oxygenation of the substrate. There are various other important flavoreductases, dehydrogenases, and electron transferases including acyl-CoA dehydrogenases involved in the oxidation of fatty acids, succinate dehydrogenases, and NADH dehydrogenases, cytochrome P-450 reductases, and aforementioned OYE known to reduce  $\alpha$ ,  $\beta$ -unsaturated carbonyl compounds *in vitro*. There are pertinent reviews<sup>1,2</sup> for a comprehensive overview on the chemistry and biochemistry of flavoenzymes.

Besides its function as a precursor of FAD and FMN riboflavin is a pigment. In *A. gossypii* riboflavin was shown to protect the hyaline ascospores against UV-light.<sup>3</sup> In photoreceptors of migratory birds FAD works as a pigment and a redox system to visualize the geomagnetic field.

### 7.04.3 Industrial Production

#### 7.04.3.1 Commercial Aspects

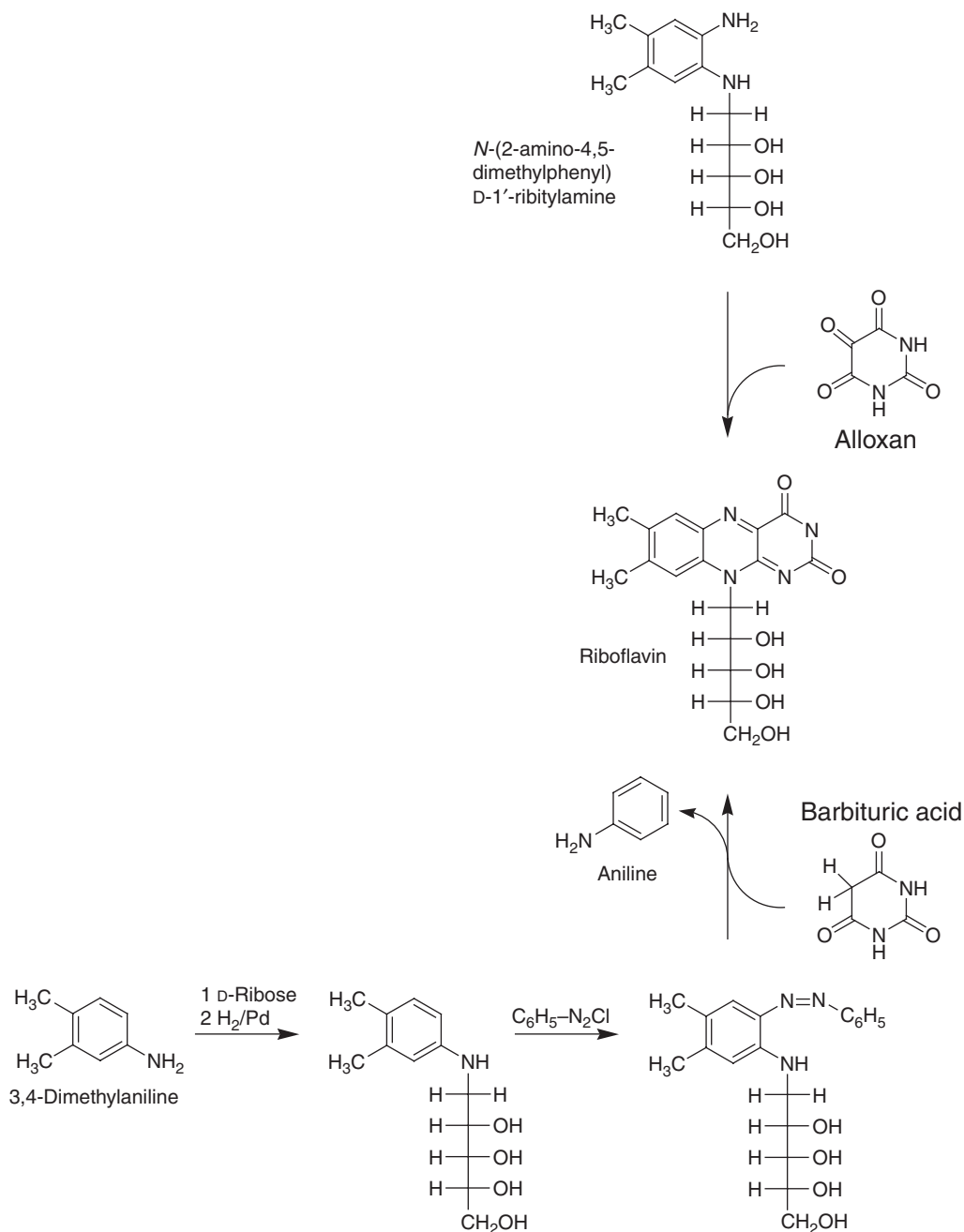
At present, over 3000 tons of riboflavin are industrially produced each year. About 70% of this material is used as feed additive in the form of free-flowing, spray-dried granules or microgranules. The remaining 30% are required for the fortification of foods like breakfast cereals, pastas, sauces, processed cheese, fruit drinks, vitamin-enriched milk products, baby formulas, and clinical infusions.

During the 1990s riboflavin market prices were quite constant at 45 and 55 EUR per kg for feed and food applications, respectively. During the last 8 years the prices dropped to less than 20 EUR per kg, primarily due to overcapacities in the market caused by the appearance of Chinese riboflavin producers. However, recently, the riboflavin market sensed supply shortages and the price of the various riboflavin products became highly volatile.

Owing to fierce price competition, the riboflavin market is consolidated to a high degree with only a few producers remaining. The market leader is DSM Nutritional Products Ltd. from Switzerland producing the vitamin in a plant in Southern Germany. BASF AG from Germany moved their riboflavin production facilities to South Korea some years ago. The main Chinese producer is Hubei Guangji Pharmaceutical Co., Ltd. from Hubei Province. Shanghai Desano Vitamin Co., Ltd. discontinued operation of their chemical production plant and might come up with a fermentation plant soon.

#### 7.04.3.2 Chemical Riboflavin Synthesis

The first chemical synthesis of riboflavin was accomplished by Kuhn and Karrer in 1934 and 1935.<sup>4,5</sup> Starting from 2-(ethoxycarbonylamino)-4,5-dimethylaniline (Karrer) or nitroxylidine (Kuhn) and D-ribose the *N*-(2-amino-4,5-dimethylphenyl)-D-1'-ribitylamine was synthesized, which was reacted with alloxane to obtain riboflavin (Figure 2). In the technical production process 3,4-dimethylaniline reacted with D-ribose to form



**Figure 2** Chemical synthesis routes to riboflavin. In the upper part of the figure the laboratory route via *N*-(2-amino-4,5-dimethylphenyl)-*D*-1'-ribitylamine followed by Kuhn and Karrer is depicted. The lower part shows the technical route devised by Tishler using 3,4-dimethylaniline, *D*-ribose, benzenediazonium chloride, and barbituric acid. During the reaction of the azo compound with barbituric acid aniline is produced in molar stoichiometry.

a Schiff base. After reduction of the Schiff base the resulting ribityl aniline derivative was diazotized and reacted with barbituric acid to riboflavin.<sup>6</sup> 3,4-Dimethylaniline and barbituric acid are cheap and easily available starting compounds. *D*-Ribose, originally synthesized from *D*-glucose by alkaline oxidation to afford arabonic acid, epimerization, and reduction, was later obtained in a highly efficient microbial process also starting from *D*-glucose.<sup>7</sup> As production organisms *Bacillus* transketolase knockout mutants (see Section

7.04.5.3.3) were used. Commercial riboflavin production with the technical demanding, multistep chemical process outlined above was optimized over decades and delivered the vitamin at >60% yield.<sup>8</sup>

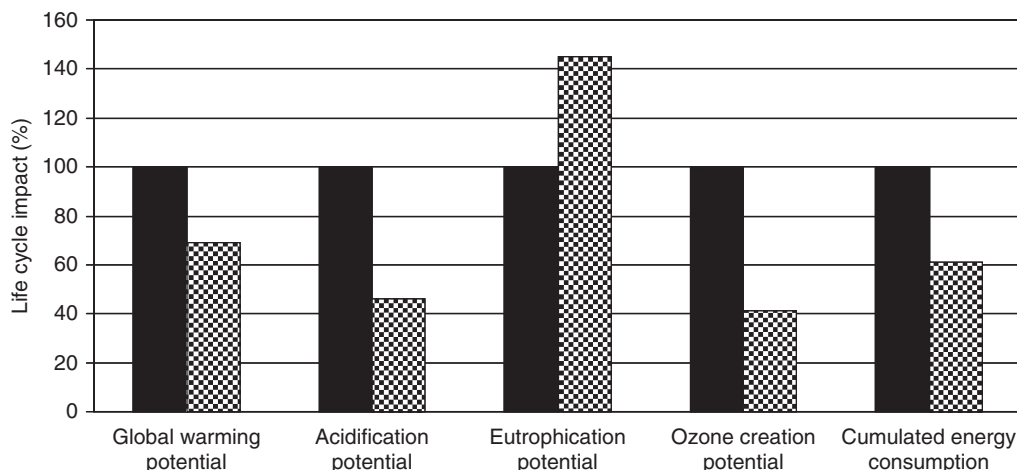
### 7.04.3.3 Microbial Riboflavin Production

The first microbial riboflavin production processes developed in the 1940s used *Eremothecium ashbyi* and *A. gossypii*, two natural riboflavin-overproducing yeast species, as production strains.<sup>9,10</sup> The US companies Commercial Solvents, Grain Processing Corporation, and Premier Malt Products built yeast-based commercial riboflavin production plants in the 1960s, but closed them after several years of operation because they could not compete with the chemical plants.<sup>11</sup>

A riboflavin production process based on *Candida famata*, another natural riboflavin-overproducing fungus growing as yeast, was developed and used for commercial riboflavin production by Coors Brewing Company, USA. Later, Archer Daniels Midland Company, USA, acquired the process and produced riboflavin from hydrolyzed corn starch. Meanwhile, the process is not used at the industrial scale anymore.

Merck Sharp and Dohme Ltd., USA, developed an *A. gossypii*-based riboflavin process in the early 1970s, which was later purchased by BASF AG, Germany, and developed further. BASF, with a long and impressive history in chemical manufacturing, implemented the *Ashbya* process at industrial scale and used it in parallel with the traditional chemical process for several years. During this period the microbial process was improved continuously with the consequence that finally the chemical production was shut down and the microbial plant was enlarged to provide sufficient production capacity. Roche Vitamins Inc. from Switzerland, now DSM Nutritional Products, the riboflavin market leader then and now, followed suit by switching entirely to a *B. subtilis*-based microbial riboflavin process in the year 2000.

The one-step microbial processes using highly effective production strains are economically superior to the chemical production technologies because of lower technical complexity and reduced consumption of raw materials, mainly renewable carbon sources like vegetable oils or carbohydrates. From this it is obvious that also under ecologic aspects the microbial riboflavin processes are superior. On behalf of the German Federal Environmental Protection Agency a cradle to grave life cycle comparison of five environmental impact categories (global warming, acidification, ozone creation, energy consumption, and eutrophication) of the traditional chemical and both, the *B. subtilis* and *A. gossypii*-based microbial production process, showed (Figure 3) that with the exception of eutrophication all other categories were clearly in favor of the microbial processes.<sup>12</sup> The study concluded that the microbial riboflavin production processes improved the life cycle performance of riboflavin production and offer more sustainable production processes.



**Figure 3** Results of a cradle to grave life cycle comparison of five environmental impact categories of the traditional chemical (solid bars) and both the *Bacillus subtilis* and *Ashbya gossypii*-based microbial riboflavin production processes (checker board bars).

## 7.04.4 Riboflavin Biosynthesis

### 7.04.4.1 Riboflavin Biosynthesis in Bacteria and Fungi

Riboflavin biosynthesis in *B. subtilis* (Figure 4) follows the common prokaryotic route starting from guanosine triphosphate and ribulose-5-phosphate in a 1:2 stoichiometric ratio, respectively. The hydrolytic opening of the imidazole ring of GTP is followed by (1) deamination of the resulting pyrimidinone ring, (2) reduction of the ribosyl side chain, and (3) dephosphorylation of the resulting ribityl side chain. 6,7-Dimethyl-8-ribityl lumazine (DMRL), the direct biosynthetic precursor of riboflavin, is synthesized by adding the 4-carbon moiety 3,4-dihydroxy-2-butanone 4-phosphate (DHBP) derived from ribulose-5-phosphate to the pyrimidinedione. Finally, two molecules of DMRL are converted in a dismutation reaction to riboflavin and the pyrimidinedione intermediate of the pathway. In the fungal riboflavin biosynthetic pathway the ribosyl side chain is first reduced and then the imidazole ring is deaminated. Otherwise, the bacteria and fungi pathways follow the same outline. Chemically plausible and experimentally supported reaction mechanisms were derived for all biosynthetic reaction steps including the intriguing conversion of ribulose-5-phosphate to DHBP and the unique dismutation reaction of DMRL to afford riboflavin.<sup>13,14</sup> High-resolution structures of the riboflavin biosynthetic enzymes are available. Important contributions to the elucidation of riboflavin biosynthesis and the enzymatic reactions catalyzed by the riboflavin biosynthetic enzymes came from G. W. E. Plaut in the 1950s and 1960s and from A. Bacher since the 1970s. For a more detailed insight into riboflavin biosynthesis and the biocatalytic mechanisms of the involved enzymes the reader is referred to excellent comprehensive reviews.<sup>15–17</sup> From all of the water-soluble vitamins the biosynthetic pathway toward riboflavin is studied to the greatest detail. It might not be just a coincidence that so far only in the case of riboflavin a microbial process following the natural biosynthetic pathway could replace an established chemical production route.

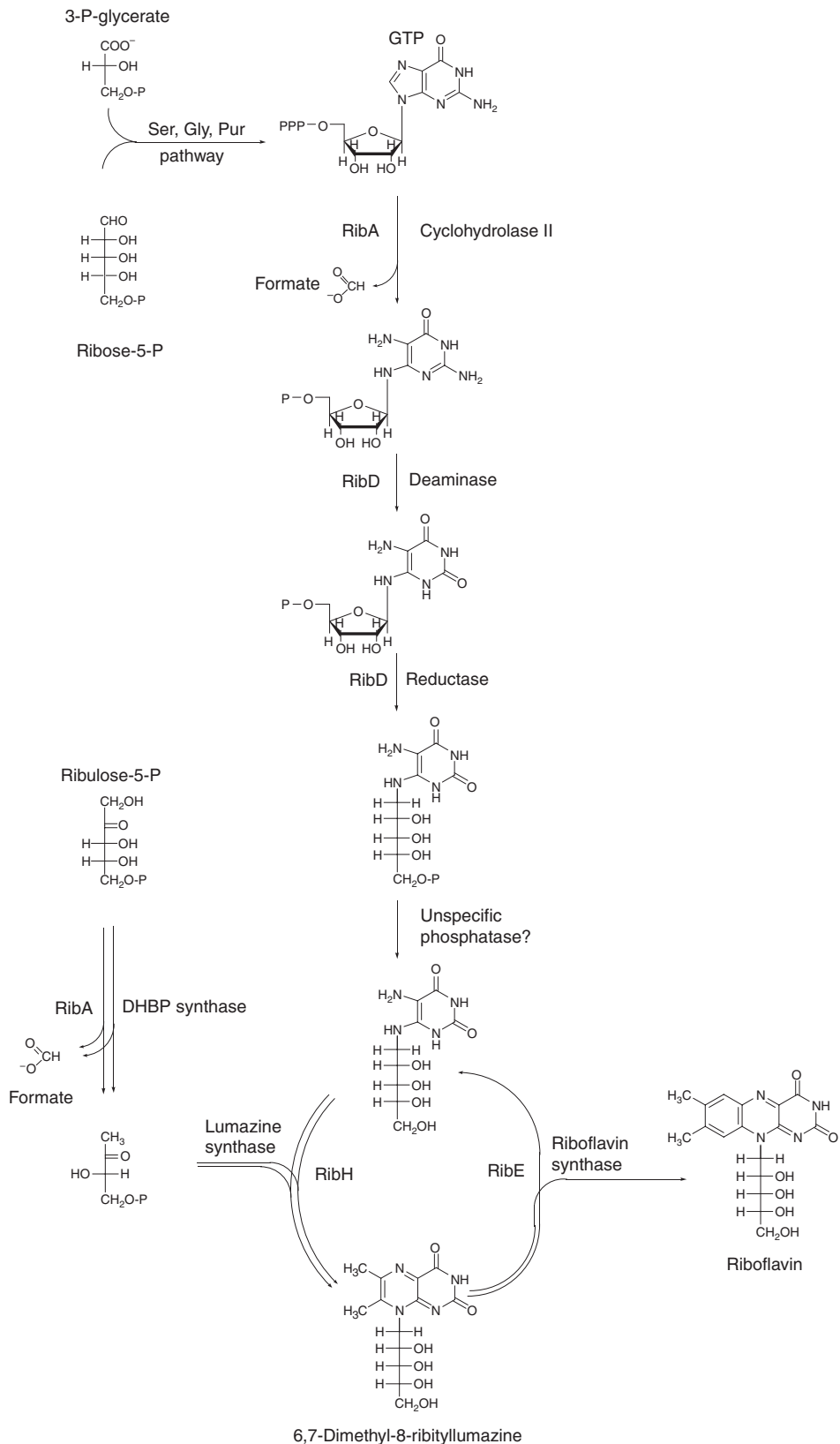
### 7.04.4.2 Riboflavin Biosynthetic Genes

#### 7.04.4.2.1 *Bacillus subtilis* and other bacteria

The *Bacillus subtilis* *rib* genes were identified and cloned by plasmid library screening with a guess-a-mer probe synthesized according to peptide sequences from purified *B. subtilis* riboflavin synthase.<sup>18</sup> Independently, the *B. subtilis* *rib* genes were isolated by functional complementation of riboflavin auxotrophic *E. coli* mutants.<sup>19</sup> In *B. subtilis* the six enzymatic activities constituting the riboflavin biosynthetic pathway are associated with four proteins (Figure 4) (1) RibD (deaminase and reductase affording 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione phosphate), (2) RibE (riboflavin synthase), (3) RibA (DHBP synthase and cyclohydrolase II), and (4) RibH (DMRL synthase). The latter synthase has a remarkable structure consisting from 60 monomers arranged into 12 pentamers, which form a capsid of icosahedral 532 symmetry (similar to capsids of viruses). The inner core of the capsid contains a riboflavin synthase homotrimer. The corresponding genes are clustered together with a fifth gene of unknown function (*ribT*) as an operon at map position 2428 kb of the *B. subtilis* chromosome (Figure 5). Dephosphorylation of the pyrimidinedione phosphate, which is a prerequisite for the pyrimidinedione to become a substrate for the DMRL synthase, is presumably achieved by the action of one or several phosphatases with broad substrate specificity. Engineering work to increase the phosphatase activity has not been reported. Obviously, the various microbial host strains used for riboflavin production express sufficient activity of these phosphatases to allow high production performance. Genomic *rib* gene arrangements as in *B. subtilis* are found, for example, in *Bacillus licheniformis*,<sup>20</sup> *Lactococcus lactis*,<sup>21</sup> and *Corynebacterium glutamicum*<sup>22</sup> at chromosomal positions 2406, 1026, and 1691 kb, respectively. In contrast to the *rib* gene arrangement in these Gram-positive bacterial genomes the *rib* genes of Gram-negative *Escherichia coli* are scattered over the whole genome<sup>23</sup> with *ribD* and *ribH* at 434 kb and *ribE* at 1741 kb. Cyclohydrolase II and DHBP synthase activity are carried by two different proteins, RibA and RibB, the genes of which are located at 1337 and 3182 kb of the *E. coli* chromosome.

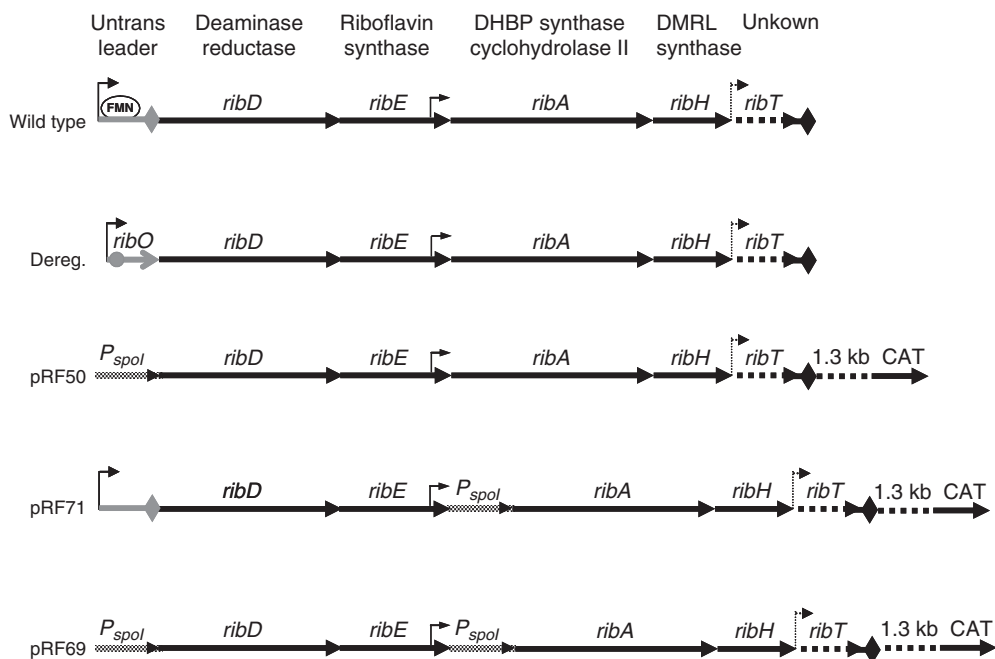
#### 7.04.4.2.2 Fungi

*Saccharomyces cerevisiae* played an important role in the genetic analysis of riboflavin biosynthesis. Mating analysis between riboflavin auxotrophic mutants obtained by random mutagenesis and screening of haploid



**Figure 4** Biosynthesis of riboflavin in bacteria. In the fungal route, the ribosyl side chain is first reduced and then the imidazole ring is deaminated.





**Figure 5** Structures of the *Bacillus subtilis* *rib* operon and various *rib* expression cassettes.

strains revealed the existence of six complementation groups.<sup>24</sup> The six *RIB* genes from *S. cerevisiae*, *RIB1* to *RIB5* and *RIB7*, were then cloned by functional complementation of the respective auxotrophic mutants.<sup>25</sup> The *RIB* genes of *A. gossypii*, which were crucial for genetic engineering of *A. gossypii* riboflavin production strains, could not be identified based on the *S. cerevisiae* *RIB* sequence information. But functional expression of *A. gossypii* cDNA and genomic DNA libraries in the *S. cerevisiae* riboflavin auxotrophs finally allowed isolation and cloning of the *A. gossypii* *RIB* genes.<sup>26</sup> The *C. famata* *RIB* genes were identified by functional complementation of *C. famata* *RIB* mutants. The genes were found to be functional in *Pichia guilliermondii* auxotrophs as well.<sup>27</sup>

### 7.04.4.3 Regulation of *rib* Gene Expression

#### 7.04.4.3.1 Bacteria

In contrast to several fungi riboflavin overproduction and riboflavin export is not observed from wild-type bacteria explaining why microbial riboflavin processes were based originally on fungal production strains. However, *B. subtilis* mutants excreting the vitamin are easily obtained by selection for resistance against riboflavin analogues, for example, roseoflavin.<sup>28,29</sup> Northern blot analysis showed the presence of a 4.3 kb *rib* transcript comprising *ribD*, *ribE*, *ribA*, *ribH*, and *ribT* in RNA preparations of a riboflavin-excreting strains.<sup>30</sup> A small transcript originating from the 5' untranslated leader region of the *B. subtilis* *rib* operon (Figure 5) was observed in addition to the 4.3 kb *rib* transcript in RNA preparations from wild-type *B. subtilis*.<sup>31</sup> Cultivation of the latter strain in the presence of riboflavin only led to the accumulation of the small transcript (DSM, unpublished). These observations together with the presence of a stem loop structure at the 3' end of the *rib* leader sequence suggested that *rib* gene expression in *B. subtilis* involves a transcription attenuation mechanism. Two classes of riboflavin-excreting mutants, designated *ribO* and *ribC* mutants, were distinguished.<sup>32</sup> Mutants of the *ribO* type carry point mutations in the *rib* leader sequence.<sup>33</sup> Riboflavin-excreting mutants were also obtained from *Lactobacillus plantarum*, *Leuconostoc mesenteroides*, and *Propionibacterium freudenreichii* cultures after selection for roseoflavin resistance.<sup>34</sup> *RibO* type mutations were confirmed in the *rib* leader sequences of the *Lb. plantarum* and *Lc. mesenteroides* mutants. The chromosomal lesion in *ribC* mutants, which were isolated at a much rarer frequency than *ribO* mutants, were mapped at 1737.1 kb of the *B. subtilis* chromosome. Surprisingly, *ribC* encoded a flavokinase/FAD synthase.<sup>35–37</sup> The flavokinase of a *ribC* mutant had only about 1% activity compared to the wild-type enzyme. The replacement of the wild type by the mutant

enzyme caused reduced intracellular concentrations of the coenzyme FMN, which is synthesized from riboflavin by a RibC-catalyzed phosphorylation step.<sup>37</sup> Since FMN, but not riboflavin is the effector molecule of the *rib* regulatory system, *ribC* mutants are deregulated for riboflavin biosynthesis and excrete up to 30 mg l<sup>-1</sup> riboflavin in overnight shake flask cultures. From the plethora of riboflavin-deregulated *B. subtilis* mutants isolated over decades in various laboratories, only cis-acting *ribO* and trans-acting *ribC* type mutants were identified, but never a mutant that indicated the existence of a *rib* regulator protein. Hence, repression of *rib* gene expression in *B. subtilis* by FMN does not most likely involve a mediator protein.<sup>37–39</sup> In fact, it has been shown *in vitro* that premature transcription attenuation can occur by direct binding of FMN to the RFN element at the 5' end of the nascent wild-type *rib* mRNA.<sup>40,41</sup> Nascent *rib* mRNA from *ribO* mutants failed to bind to FMN and transcription attenuation was prevented.

#### 7.04.4.3.2 Fungi

Overproduction of riboflavin in glucose-cultivated *A. gossypii* was found to be linked to the formation of ascospores. Extracellular supplementation with cAMP or nonmetabolizable derivatives of cAMP arrested the fungus in the state of a vegetative mycelium without riboflavin overproduction indicating that a cAMP-dependent signaling cascade controlled riboflavin overproduction.<sup>42</sup> While *A. gossypii* did not overproduce riboflavin during continuous cultivation at constant dilution rates, dilution rate downshifts associated with nutritional stress resulted in sporulation of the culture and riboflavin overproduction. Increased *RIB3*, *RIB4*, and *RIB5*, but not *RIB2* and *RIB7* mRNA levels were detected upon the nutritional downshift. Concerning *RIB1* the results were inconsistent.<sup>43</sup> After entering the stationary growth phase the DHBP synthase activity increased 50-fold due to increased transcription initiation as revealed by promoter–reporter fusion experiments.<sup>42</sup> While the molecular details underlying control of *rib* gene expression in *A. gossypii* are still to be elucidated a mechanism involving glucose repression could be excluded.<sup>42</sup>

The initiation of *RIB* gene expression in *A. gossypii* correlates with sporulation of the fungus suggesting a common regulatory mechanism for asci development and the induction of riboflavin overproduction. Riboflavin might contribute to both integrity and dissemination of *A. gossypii* spores. The natural habitats of *A. gossypii*, a weak plant pathogen, are plant tissues like cotton balls (*Gossypium hirsutum*), which become yellowish probably due to pigmentation with riboflavin when infected by the fungus. The pigments might protect the hyaline *A. gossypii* spores against UV–light-induced damage. *A. gossypii* is unable to penetrate the epidermis of plant tissue by itself. Fungal dissemination therefore depends on insects with piercing–sucking mouthparts like *Dysdercus nigrofasciatus* taking up and distributing the spores during feeding at the plants. The riboflavin pigmentation of *A. gossypii* spores in the cotton balls might attract the insects preferably to the infected plants.

### 7.04.5 Construction of Riboflavin Production Strains

#### 7.04.5.1 *Bacillus subtilis* and *Ashbya gossypii* as Preferred Host Strains

*Bacillus subtilis* was the organism of choice to develop a bacterial route to the commercial production of riboflavin for several reasons. First, *Bacillus* species, including *B. subtilis*, are preferred host organisms for recombinant protein production and classified as GRAS (generally regarded as safe). *Bacillus natto* with a genome almost identical to *B. subtilis* is used in the fermentation of steamed soybeans to produce the popular Japanese food product, itohiki-natto, which is also obtainable in Western countries. *B. natto* is not removed from the fermentation product, and natto too is not pasteurized prior to human consumption. Thus, the safe use of *B. natto* in human nutrition is well documented. Second, classical mutagenesis had been used previously to develop strains of various *Bacillus* species that were capable of producing significant quantities of inosine or guanosine (20–40 g l<sup>-1</sup>).<sup>44–47</sup> This indicated that *Bacillus* species could be engineered to efficiently supply GTP, one of the two immediate riboflavin biosynthetic precursors, for high riboflavin production. Third, a vast knowledge on the genetics and physiology of *B. subtilis* is available,<sup>48,49</sup> which is of great importance for a successful strain engineering program.

The *A. gossypii* wild-type strain isolated by Ashby and Novel in 1926 is already a significant natural overproducer of riboflavin accumulating 100 mg g<sup>-1</sup> biomass under suited cultivation conditions.<sup>42</sup> The identification of improved production strains is possible by simple visual inspection of colonies obtained by

cultivation of a randomly mutagenized *A. gossypii* population on agar plates. Since haploid monokaryotic spores can be easily harvested, classical mutants can be obtained conveniently at high throughput. A fully developed genetic tool box is available including plasmid transformation, replacement mutagenesis with dominant marker genes, restriction enzyme-mediated chromosomal integration, and transposon mutagenesis.<sup>50</sup> Site-specific, double crossover chromosomal integration is possible with less than 50 bp homology regions flanking the DNA to be integrated at the 5' and 3' end.<sup>51</sup> Looping out of selection markers, for example, the geneticin resistance gene, or non-*Asbyya* genes is possible via short repetitive DNA sequences, which allow the repeated use of the same dominant marker for consecutive chromosomal DNA integrations.

Because introns are rare in the *A. gossypii* genome, functional complementation of *Saccharomyces cerevisiae* mutants with *A. gossypii* genomic libraries is possible. The entire *A. gossypii* genome sequence is meanwhile publicly available.<sup>52</sup> *A. gossypii*, although a filamentous fungus, can grow relatively fast with a maximum specific growth rate  $\mu_m = 0.6 \text{ h}^{-1}$  in a minimal medium. The fungus is resistant against shear stress and can grow to high cell density, a prerequisite for large-scale industrial fermentations.

## 7.04.5.2 Riboflavin Pathway Engineering

### 7.04.5.2.1 *Bacillus subtilis* production strains: rib genes driven by their natural promoters

*Bacillus subtilis* VNIIGenetika 304/pMX45 from the Russian Institute for Genetics and Selection of Industrial Microorganisms, Moscow, was the first riboflavin production strain and maybe the first production strain for a small organic molecule at all, that was obtained by a genetic engineering program. The Russian patent application SU3599355 claiming the strain and a fermentation process for riboflavin was filed with priority date 2 June 1983. The host strain VNIIGenetika 304 carried a resistance marker against 8-azaguanine and a *ribC* mutation to deregulate the purine and the riboflavin biosynthetic pathway, respectively. The plasmid pMX45 present in VNIIGenetika 304 contained a 10 kbp *EcoRI* fragment from the chromosome of a *B. subtilis* *ribO* mutant comprising the entire *rib* operon as deduced from complementation studies<sup>53</sup> and an erythromycin selection marker. With VNIIGenetika 304/pMX45 4.5 g l<sup>-1</sup> riboflavin were produced from a total of 100 g l<sup>-1</sup> saccharose supplied at fermentation start and 12 h after fermentation start during a 25 h process in a small stirred tank reactor.<sup>54</sup>

The co-occurrence of a chromosomal and episomal copy of the *rib* operon in *B. subtilis* strains provided with pMX45 gave rise to plasmid instability, which could be prevented by introducing a *recE4* mutation into the host strain genome.<sup>55</sup> However, the *recE4* mutation resulted in decreased growth rates and prevented further selection of genetically modified strains because of the low survival rate of *recE4* mutants. As an attempt to breed stable production strains, the *rib* operon comprising the four *rib* genes preceded by the deregulated *rib* leader of a *B. amyloliquefaciens* roseoflavin-resistant mutant was introduced on a replicative plasmid into a *B. subtilis* host strain<sup>55</sup> reasoning that the heterologous *rib* operon on the plasmid would reduce the recombination frequency between chromosomal and episomal DNA sequences. In fact, in *B. subtilis* Marburg 168 *trpC2* the plasmid was stably passed on. In a riboflavin production strain background, however, the plasmid turned to be unstable, which was explained with an 'overdosing' of *rib* gene expression. Genetically stable riboflavin production strains were finally obtained using integrative vector constructs.<sup>56</sup> First, the plasmid pES8A comprising a 6 kbp *EcoRI* fragment from a *B. amyloliquefaciens* *ribO* mutant encoding the entire *rib* operon and a chloramphenicol resistance marker was constructed. Plasmid pES8A, which contained an *E. coli*, but lacking a *B. subtilis* replication origin, was then linearized at a unique *BamHI* site, ligated *in vitro* with random *BamHI* fragments of *B. subtilis* DNA and the ligation mix was used to transform a riboflavin auxotrophic *B. subtilis* recipient strain. Circular DNA molecules in the ligation mix containing *B. subtilis* *BamHI* fragments were targeted via the *BamHI* fragments in a Campbell integration event into various loci of the recipient host. The resulting transformants, which were obtained in reasonable frequency, overproduced riboflavin, but only a small percentage of these could stably transmit the overproducing phenotype in the absence of antibiotic selection. The high frequency of unstable strains was explained by the formation of direct repeats during the single crossover transformation event. From one of the stable transformants containing the transgenes in close proximity to the *pur* operon, a phage lysate was prepared and used to transduce a deregulated riboflavin-producing *B. subtilis* host strain. The resulting strain Y32 produced 3 g l<sup>-1</sup> riboflavin during a 72 h shake flask fermentation. Like Y32 the riboflavin production strain GM41 pMX45<sup>57</sup> harbours in its chromosome a deregulated *B. amyloliquefaciens* *rib* operon and the plasmid pMX45 (see above, this section). In addition, GM41

pMX45 was selected for its ability to grow with 1 g/l glycerophosphate as sole carbon source in the presence of 1 g/l glyoxylate. Under small scale fed-batch fermentation conditions the strain accumulated 21 g/l riboflavin during 70 h. It is presumed that riboflavin-producing companies from China employ *B. subtilis* production strains with a genetic make up similar to that of GM41 pMX45.

Precise copy number dosing of the deregulated *rib* operon in a *B. subtilis* host strain for optimal riboflavin production was recently emphasized by a research group from the Tianjin University, Tianjin, China.<sup>58</sup> Seven to eight copies gave best results. Chromosomal integration via double crossover proved superior over single crossover, presumably because of stability reasons.<sup>59</sup>

#### 7.04.5.2.2 *Bacillus subtilis* production strains: rib gene expression driven by heterologous promoters

After the *B. subtilis* *rib* operon was sequenced,<sup>18,19</sup> the correct functional assignment of the open reading frames within the operon and the identification of the relevant transcriptional and translational signals became possible. Before that only a tentative and partly even incorrect structure of the *rib* operon based on mapping data was available. A more precise genetic engineering attempt to construct riboflavin production strains was then made.<sup>18,60</sup> The endogenous upstream promoter and the leader sequence with the repressible *rib* operator were replaced by a strong constitutive promoter derived from the *Bacillus* bacteriophage *SPO1* (Figure 5). The modified *rib* operon was then ligated into a standard cloning vector in close proximity to a chloramphenicol resistance cassette affording plasmid pRF50. In analogy pRF71 with the *SPO1* promoter element between *ribE* and *ribA* and a hybrid of pRF50 and pRF71 with two *SPO1* elements designated pRF69 were constructed. For control purposes, pRF40 was obtained with a 6.5 kb DNA fragment comprising the entire wild-type *B. subtilis* *rib* operon linked to its natural transcription regulatory elements. The plasmids were integrated via single crossover recombination into the native *rib* locus of a *B. subtilis* RB50 recipient strain provided with a derepressing *ribC* mutation and resistance markers against purine analogues. In the transformants the chloramphenicol resistance gene was flanked 5' and 3' by extended repetitive DNA sequences, the native *rib* operon and the modified operon, respectively. The repetitive genomic structures gave rise to individuals within the bacterial population with multiple copies of the *rib* operon and of the chloramphenicol resistance gene.<sup>61</sup> The increased chloramphenicol resistance (60–80 mg l<sup>-1</sup> chloramphenicol) of bacteria with multiple copies of the *cat* gene provided a convenient selection tool for individuals with amplified structures (in average around 10 copies).

Northern blot analysis revealed *rib* mRNAs with expected sizes of 4.2 and 2.3 kb in RNA preparation from strains provided with pRF40 and pRF71, respectively.<sup>60</sup> Surprisingly, pRF69 gave rise to *rib* mRNA species below 2 kb from the 5' end of the *rib* operon, which did not comprise RibA. Obviously, the absence of the *rib* leader had detrimental effects on the accumulation of unmitigated *rib* mRNA.

Riboflavin production processes<sup>18,60</sup> with the RB50 strains containing the engineered and amplified *rib* operons were carried out for 48 h under carbon-limited fed-batch conditions. Starch hydrolysate was supplied as the main carbon source. Ammonia served as pH titrant and nitrogen source. The cultures were aerated to keep a 15% oxygen saturation level. During the inoculum preparation the antibiotic selection pressure was maintained, whereas the main fermentations were run in the absence of antibiotics. The copy number of the engineered *rib* operons did not dwindle as revealed by Southern analysis (DSM, unpublished). With production strains provided with pRF40, pRF50, and pRF71, 7.4, 11, and 10.5 g l<sup>-1</sup> riboflavin were obtained, respectively. The results indicated that the functional linkage of a strong promoter element to the *rib* genes significantly increased the riboflavin production capabilities of *B. subtilis*. Comparing the fermentation results of the various strains with the results of the Northern analysis revealed that overexpression of either the upstream genes *ribD* and *ribE* or the downstream genes *ribA* or *ribH* boosted riboflavin production. Accordingly, *SPO1*-driven expression of all *rib* genes in pRF69 led to a further increase in production reaching 14 g l<sup>-1</sup> with a product yield of about 7% (w/w) on the carbon source.

Cultivation of RB50::[pRF69]*n* with fractional <sup>13</sup>C-labeled glucose as carbon source and metabolic flux balancing indicated that riboflavin formation in this strain was still limited by the fluxes through the terminal biosynthetic rather than the central carbon metabolic pathways.<sup>62</sup> To debottleneck the riboflavin biosynthetic pathway further, a similar approach as described above was applied to introduce a second P<sub>SPO1</sub>-driven and amplifiable *rib* expression cassette linked to a tetracycline resistance marker at the *bpr* locus (map position 1598.6 kb) affording strain RB50::[pRF69]*n*::[pRF93]*m*.<sup>60</sup> Proteome analysis of an extract of amplified

RB50::[pRF69]*n*::[pRF93]*m* (*n* around 10, *m* around 5) revealed that each of the Rib proteins, which were barely detectable in the untransformed parent strain, accumulated to 3–5% of the total soluble protein content (DSM, unpublished data). The *in vitro* activities of the riboflavin biosynthetic enzymes are in the range of 200 nmol mg<sup>-1</sup> min<sup>-1</sup> and below.<sup>17</sup> The rather low *in vitro* and presumably *in vivo* enzymatic activities makes massive Rib protein overexpression mandatory to obtain production strains with a high specific productivity.

Several attempts to simultaneously overexpress all the genes of the riboflavin operon to still higher levels than those reached in RB50::[pRF69]*n*::[pRF93]*m* failed suggesting that one or several riboflavin biosynthetic enzymes might exert a negative effect on the cell when expressed above a critical level. Therefore, the effect on riboflavin production of double crossover integrations at the *sacB* locus (map position 3535.0 kb) of single copies of the individual *rib* genes transcribed from the medium strong *B. subtilis* *veg* promoter were evaluated. Additional expression of *ribA* encoding the protein that carries both pathway entrance enzymes, cyclohydrolase II and DHBP synthase, further increased riboflavin productivity by 25%.<sup>63</sup> The additional *ribA* copy in strain VB2XL1 resulted in a three- to sixfold increase in the activities of the RibA-associated enzymes suggesting that the P<sub>veg</sub> driven *ribA* gene is much more efficiently expressed than the *ribA* genes from both of the engineered *rib* operons. This could result from more efficient transcription or translation, improved mRNA stability, or a combination thereof. The moderate productivity and yield increase upon a several-fold increase in cyclohydrolase II and DHBP synthase activity measured in cell extracts of VB2XL1 indicated that in this strain one of the remaining riboflavin biosynthetic enzymes or limitations in the precursor supply pathways restrained product formation.

Recently, after a detailed study of the *rib* leader sequence as a regulatory and mRNA stabilizing element a new generation of highly effective riboflavin production strains became available provided with a single copy of a carefully deregulated *rib* operon driven by a *SPOI* promoter element at the 5' end of the operon (DSM, patent application submitted). These strains are free of antibiotic markers and fulfill the legal requirements for self-cloning strains (EC Council directive 98/81), which is of some significance to market fermentative riboflavin in Europe. *B. subtilis* production strains similar to the ones described above are employed by DSM Nutritional Products of Switzerland for commercial riboflavin production.

A *B. subtilis* riboflavin production strain has been claimed recently in a Chinese patent application from the Tianjin University, Tianjin.<sup>64</sup> The *B. subtilis* riboflavin operon without its natural upstream promoter and regulatory sequences was synthesized by PCR, linked at the 3' end to a kanamycin resistance cassette and introduced as a second *rib* copy via Campbell integration into the *rib* locus of a roseoflavin and purine analogue-resistant *B. subtilis* host strain. The promoter element driving the additional *rib* operon was not disclosed. A protoplast fusion procedure without any counterselection was applied to combine the presumed genetic diversity in a population of roseoflavin and purine analogue-resistant mutants. With the resulting strain F4-RH33 small-scale aerobic fed-batch fermentation runs with molasses sucrose as the main carbon source afforded up to 12 g l<sup>-1</sup> riboflavin at a yield of 5.8% (w/w) on sucrose.<sup>65</sup>

#### 7.04.5.2.3 *Corynebacterium ammoniagenes*

The *rib* operon of *Corynebacterium ammoniagenes* is structured as the *rib* operons of *B. subtilis* or *C. glutamicum*. Purine analogue-resistant *C. ammoniagenes* mutants, which were transformed with plasmids encoding the *C. ammoniagenes* *rib* operon, proved efficient riboflavin production strains at Kyowa Hakko Kogyo Co, Japan.<sup>66,67</sup> Strain KY13628, which was provided with plasmid pFM67 encoding the *rib* operon driven by the strong *C. ammoniagenes* promoter P54-6, accumulated 5 g l<sup>-1</sup> riboflavin during a 72 h aerobic fermentation run at 32 °C with ammonia as pH titrant and nonlimiting glucose supply. The activities of *ribA*-encoded cyclohydrolase II and *ribE*-encoded riboflavin synthase in protein extracts of the strain were 0.011 and 2.8 mU mg<sup>-1</sup>, respectively. With an optimized ribosomal binding site in front of *ribA* the activity of cyclohydrolase II was increased threefold and 15.3 g l<sup>-1</sup> riboflavin were produced at 4.8% (w/w) yield on glucose. Riboflavin overproduction did not depend on a *ribC* type mutation indicating that *rib* gene expression in *C. ammoniagenes* is independent from the intracellular FMN levels. The mechanism regulating the expression of the *Corynebacteria* *rib* genes might in fact be different from the regulation in *B. subtilis* as *C. glutamicum* showed no FMN-binding RFN regulatory element associated with its *rib* operon. Nevertheless, *Corynebacteria* might make use of FMN-responsive RFN elements, since an RFN signature was identified in front of the *C. glutamicum* *pnuX* gene encoding a riboflavin transporter protein.<sup>68,69</sup>



#### 7.04.5.2.4 *Ashbya gossypii*

To introduce additional copies of the *RIB3* (DHBP synthase), *RIB4* (DMRL synthase), and *RIB5* (riboflavin synthase) genes into the chromosome of *A. gossypii*<sup>70</sup> a geneticin resistance cassette preceded by the *A. gossypii* *TEF* promoter<sup>71</sup> and the *RIB3*, *RIB5*, and *RIB4* genes, the latter three genes preceded by their original promoters, were cloned back to back into a conventional cloning vector. A 6.3 kb *SpeI* fragment comprising the resistance marker and the *RIB* genes together with their promoter sequences was generated from the plasmid and introduced into the genome of the itaconate-resistant mutant Ita-GS-01<sup>72</sup> by restriction-enzyme-mediated integration (REMI). One of the resulting transformants designated Ita-GS-01.17 accumulated in shake flask cultures almost 250 mg l<sup>-1</sup> riboflavin, whereas the parent strain afforded only around 100 mg l<sup>-1</sup>.<sup>73</sup> Introduction of additional copies of the *RIB1* (cyclohydrolase II), *RIB2* (deaminase), *RIB4* (DMRL synthase), and *RIB7* (reductase) genes into the genome of *A. gossypii* LU21 again applying REMI also resulted in a significant increase in riboflavin production.<sup>74</sup> Flanking the geneticin resistance gene with two TEF promoter elements at the 5' and 3' end facilitated a looping out of the resistance marker via the repetitive structure in the transformants. The resulting marker-free strains might fulfill the legal requirements for self-cloning strains according to EC Council directive 98/81. Despite these two examples of successful application of REMI, the method has several disadvantages, first of all the risk of disruption of loci<sup>74</sup> involved in product formation. Site-specific targeted insertion of exogenous DNA by double crossover into the *A. gossypii* is possible.<sup>51</sup> The low frequency of site-directed double crossover insertion events might ask for more screening efforts, but will provide strains with a better-defined genetic structure.

#### 7.04.5.3 Metabolic Pathway Engineering to Improve Riboflavin Precursor Supply

##### 7.04.5.3.1 Purine and guanosine pathway, nitrogen metabolism in *Bacillus subtilis*

Since GTP is one of the two precursor compounds for the converging riboflavin biosynthetic pathway it is obvious that a high flux through the purine pathway is a prerequisite for superior riboflavin production. Efficient inosine and guanosine production strains of both, *Bacillus* and *Corynebacterium* origin, were obtained in the 1960s and 1970s by selecting for purine analogue-resistant mutants.<sup>47</sup> Typical purine analogues used were azaguanine, decoyinine, and psicofuranine. The mutations leading to the deregulated metabolic flux through the purine pathway have not been cloned and sequenced so far, but it is assumed that they interfere with adenine-imposed repression of the *pur* operon (in azaguanine-resistant mutants) or inhibition of GMP synthase (decoyinine- and psicofuranine-resistant mutants).<sup>47</sup> Since purine accumulation triggering the shutdown of purine *de novo* biosynthesis is not to be expected in strains overexpressing the riboflavin biosynthetic enzymes including GTP-consuming cyclohydrolase II, the necessity to interfere with purine-induced repression of the purine pathway is not immediately evident. Nevertheless, all of the *Bacillus* and *Corynebacterium* strains genetically engineered for enhanced *rib* gene expression were also provided with one or several of the resistance markers against the purine analogues azaguanine, 6-mercaptoguanine, decoyinine, or combinations thereof. Mutations deregulating the purine biosynthetic pathway seem necessary for efficient riboflavin production at least in the case of *Bacillus* host strains, as only roseoflavin-resistant *B. subtilis* strains provided with decoyinine and azaguanine resistance markers formed intensely fluorescing colonies indicating riboflavin overproduction.<sup>60</sup> Molecular engineering of the purine pathway, although successfully applied to enhance guanosine production,<sup>75</sup> has not been reported for breeding improved riboflavin production strain.

The nitrogen atoms N1 and N10 of riboflavin are recruited from L-glutamine via the phosphoribosylpyrophosphate (PRPP) amidotransferase and FGAM synthase reactions of the purine metabolic pathway imposing an elevated glutamine demand on riboflavin-overproducing strains. This might explain an inverse correlation of the decreasing glutamate synthase (GOGAT, glutamine oxoglutarate aminotransferase) activity in cell extracts of the *B. subtilis* RosR, *B. subtilis* strain 41, *B. subtilis* strain 24, and *B. subtilis* strain AS5, and the increasing riboflavin productivity in this strain lineage.<sup>76</sup> Glutamate synthase is essential for L-glutamate synthesis in *B. subtilis* starting from L-glutamine and 2-oxoglutarate to produce two molecules of L-glutamate. Reduced glutamate synthase activity might result in elevated intracellular L-glutamine levels enhancing riboflavin production.



#### 7.04.5.3.2 Purine and glycine metabolism in *Ashbya gossypii*

As in many other microorganisms the phosphoribosylamine synthesis in *A. gossypii* catalyzed by PRPP amidotransferase, the committing step in the purine biosynthesis pathway, is negatively regulated at the transcriptional level by adenine. In addition, ATP and GTP exert a strong inhibitory effect on PRPP amidotransferase activity. For deregulated expression of PRPP amidotransferase in *A. gossypii* the PRPP amidotransferase-encoding *AgADE4* gene was functionally linked to the strong constitutive promoter of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene. Furthermore, a feedback-resistant PRPP amidotransferase mutant (asp310val, lys333glu, ala417trp) was constructed by site-directed mutagenesis. Replacing wild-type *AgADE4* by the mutant gene under the control of the GAPDH promoter in wild-type *A. gossypii* ATCC 10895 increased riboflavin production from 28 to 228 mg l<sup>-1</sup> after 72 h shake flask cultivation.<sup>77</sup>

Glycine supplementation had a positive effect on riboflavin overproduction with *A. gossypii*<sup>78</sup> presumably by stimulating *de novo* biosynthesis of the riboflavin precursor GMP. The glycine analogue aminomethyl phosphonic acid did not increase, but reduced riboflavin production<sup>79</sup> suggesting that glycine did not stimulate riboflavin production via a regulatory function, but alleviated as a substrate a possible glycine shortage.

L-Threonine supplementation of an *A. gossypii* strain overexpressing the threonine aldolase *GLY1* gene increased riboflavin production ninefold. The strong enhancement of riboflavin production exceeding riboflavin stimulation by glycine was attributed to an almost quantitative uptake of L-threonine and its intracellular conversion into glycine. This became evident by glycine excretion of an L-threonine-supplemented culture.<sup>80</sup> Elevated supply with glycine has also been held responsible for the enhanced riboflavin overproduction in *SHM2*-disrupted *A. gossypii* mutants. *SHM2* encodes the cytosolic serine hydroxymethyltransferase of *A. gossypii* catalyzing glycine to serine interconversion.<sup>81</sup> *In vivo* <sup>13</sup>C-labeling experiments confirmed that this cytosolic serine hydroxymethyltransferase was a major glycine-consuming enzyme during the riboflavin production phase. *SHM2*-disrupted *A. gossypii* mutants are not glycine auxotroph, obviously due to the presence of other genes in the *A. gossypii* genome, which encode glycine producing enzymes, that is, mitochondrial serine hydroxymethyltransferase, threonine aldolase, and a glyoxylate aminotransferase.

#### 7.04.5.3.3 Pentose phosphate pathway

Guanosine triphosphate and ribulose-5-phosphate are recruited in a 1:2 stoichiometric ratio by GTP cyclohydrolase II and DHBP synthase, respectively, for riboflavin biosynthesis. Since at substrate saturation the activity of *B. subtilis* DHBP is twice the activity of *B. subtilis* cyclohydrolase II (DSM, unpublished observations) and since both enzymatic activities are associated with the same bifunctional protein encoded by *ribA*, the balanced formation of the pyrimidinedione and the dihydroxybutanone intermediates is ensured. However, the  $K_{0.5}$  constant of DHBP synthase ( $\sim 1$  mmol l<sup>-1</sup>) is about 100-fold higher than the  $K_{0.5}$  constant of GTP cyclohydrolase II imposing the risk of excessive synthesis of the pyrimidinone and pyrimidinedione intermediates in case of reduced intracellular concentrations of pentose phosphate pathway intermediates. This can be expected, for instance, in glucose-limited fed-batch fermentations, which are frequently used in industrial applications. The pyrimidinone and pyrimidinedione intermediates are highly reactive, oxidative compounds, which can do serious damage on the bacteria.

In mutants of various *Bacillus* strains defective for transketolase, a key enzyme of the pentose phosphate pathway, elevated intracellular C5 carbon sugar pools are reached up to a level that exceeds the physiological requirements of the bacteria and leads to excretion of excess ribose into the fermentation broth.<sup>82,83</sup> The *B. subtilis* chromosome comprises only one transketolase-encoding gene and lacks *edd*-encoding gluconate-6-P dehydratase, a key enzyme of the Entner–Doudoroff pathway. Therefore, transketolase mutants cannot synthesize erythrose-4-P leading to an auxotrophy for shikimate or aromatic amino acids and are unable to grow on carbon sources assimilated via the C5 sugar phosphate shunt. The ribose-excreting capability of transketolase-defective *Bacillus* mutants is exploited for industrial production of ribose, which was used for instance in the chemical riboflavin production process (see Section 7.04.3.2). Because of the importance of a sufficiently high intracellular ribulose-5-phosphate concentration to ensure balanced precursor supply for riboflavin biosynthesis, the effects of mutations in the transketolase-encoding *tkt* gene on the riboflavin production performance have been evaluated in *B. subtilis* and *Corynebacteria*. Improved riboflavin production was reported with the *B. subtilis* transketolase mutant 24.<sup>84</sup> The mutant was auxotroph for shikimic acid and could not grow with gluconate as sole carbon source in accordance with the expected phenotype of a transketolase knockout mutant.

*Corynebacterium glutamicum* transketolase mutants, which contained a plasmid encoding the *C. ammoniagenes* rib operon, demonstrated superior riboflavin production over transketolase wild-type host strains provided with the same plasmid. The experiments were carried out in test tubes reaching 41 and 108 mg l<sup>-1</sup> riboflavin with the transketolase wild type and mutant strain, respectively.<sup>85</sup> It is not clear to which extent these results are relevant for industrial high-intensity fed-batch fermentation runs.

In the *B. subtilis* RB50::[pRF69]*n* strain (see Section 7.04.5.2.2), riboflavin production was negatively affected by a transketolase knock-out mutation.<sup>86</sup> However, RB50::[pRF69]*n* strains expressing impaired, but not completely inactive transketolases showed increased riboflavin production performance. The *tkt* mutants encoding partially inactivated transketolases were generated by saturation mutagenesis at amino acid position R357, corresponding to R359 of the intensively studied *S. cerevisiae* transketolase orthologue.<sup>87</sup> This position is critical for substrate binding. Amino acid replacements at this position in the yeast enzyme resulted in mutant transketolases with considerable enzyme activity, but increased  $K_{0.5}$  values for the donor and in particular for the acceptor substrate. *B. subtilis* transketolase mutants were selected by their ability to rescue *B. subtilis tkt* deletion mutants from shikimic acid auxotrophy and to allow them to grow on gluconate albeit at reduced growth rates compared to a *B. subtilis* strain containing wild-type *tkt*. For example, strain BS3534 expressing the R357A transketolase mutant showed a 43% riboflavin yield increase on glucose compared to parent RB50::[pRF69]*n* in glucose-limited fed-batch fermentation runs. It is assumed that similar to the R557 mutations in the yeast orthologue the R359A mutation in the *B. subtilis* enzyme leads to increased  $K_{0.5}$  values for the donor and acceptor substrates xylulose-5-P and ribose-5-phosphate, respectively, resulting in elevated intracellular concentrations of these sugar phosphates. The intracellular ribulose-5-phosphate concentration should be increased as well, since the three C5 sugar phosphates are easily converted into each other via epimerase and isomerase reactions. The important riboflavin building block ribulose-5-phosphate and the purine building block ribose-5-phosphate, thus, become more readily available.

As an alternative approach to convey increased metabolic flux toward C5 carbon sugar phosphates glucose dehydrogenase GDH, a *B. subtilis* enzyme whose expression is naturally limited to the developing forespore, was put under the expression control of the constitutive *cdd* promoter P43 and inserted into the *amyE* locus of a *B. subtilis* riboflavin production strain.<sup>88</sup> Expression of *glcU* encoding a glucose importer and forming an operon with *gdb* in wild-type *B. subtilis* was not modified. GDH oxidizes intracellular glucose to gluconate and might open a second oxidative route toward pentose phosphates in addition to *zwf*-encoded glucose-6-phosphate dehydrogenase. This alternative route depends on the availability of nonphosphorylated intracellular glucose and requires gluconate kinase to be released from catabolite repression. In shake flask cultures with glucose as carbon source a *B. subtilis* riboflavin production strain constitutively expressing GDH activity demonstrated improved riboflavin production performance.

#### 7.04.5.3.4 Modifications in the central metabolism of *Bacillus subtilis* production strains

In *B. subtilis*, the genes encoding gluconeogenic glyceraldehyde-3-phosphate dehydrogenase (GapB) and phosphoenolpyruvate carboxykinase (PckA) are repressed by glucose. However, glucose repression is not exerted by CcpA, the major transcriptional regulator for catabolite repression in *B. subtilis*, but by CcpN.<sup>89</sup> CcpN knock-out mutants of the aforementioned RB50::[pRF69]*n* riboflavin production strain (see Section 7.04.5.2.2) demonstrated increased riboflavin yield and productivity in deep-well plate cultivations with glucose, slowly metabolizable raffinose, or gluconeogenic malate as carbon source.<sup>90</sup> A plausible explanation for the positive *ccpN* knock-out effect on riboflavin production might be the increased gluconeogenic flux as a result of deregulated *gapB* and *pckA* expression in the *ccpN* knockout mutant, possibly making more substrate available for product formation by reducing the net catabolic flux toward the tricarboxylate cycle.

Disruption of the phosphotransacetylase-encoding *pta* gene in combination with overexpression of the acetolactate synthase-encoding *als* gene in a *B. subtilis* riboflavin production strain led to increased riboflavin production (4.5% w/w yield on glucose compared 3.0% yield without the *pta* and *als* modifications) during batch cultivation with glucose as carbon source.<sup>91</sup> Elevated citrate synthase and isocitrate dehydrogenase activities suggested an increased metabolic flux through the tricarboxylate cycle, which could explain the 5.8-fold elevated ATP-to-ADP ratio in the *pta*, *als* mutant. The high concentration of ATP was held responsible for the observed increase in riboflavin production.

Various other attempts, in addition to the one mentioned above, to redirect the central metabolic flux toward the riboflavin building blocks were made including expression modification of glycolysis,<sup>92</sup> TCA cycle, or pentose phosphate pathway enzymes. At the level of intracellular fluxes, the to-date most comprehensive 'fluxome' analysis with a plethora of *B. subtilis* knock-out mutants revealed a robust flux distribution, which indicates a surprisingly high degree of network rigidity of *B. subtilis*.<sup>93</sup> This result might explain why, with the remarkable exception of the impaired transketolase modification leading to improved riboflavin production under industrially relevant conditions, all of these attempts were met with limited success. Possibly the application of multiple omics methods<sup>94</sup> may increase the success rate.

#### **7.04.5.3.5 Isocitrate lyase and isocitrate dehydrogenase in *Ashbya gossypii* production strains**

*Ashbya gossypii* mutants resistant to isocitrate lyase inhibitors, for example, itaconate, showed increased riboflavin production.<sup>72</sup> Isocitrate lyase together with malate synthase constitute the anaplerotic glyoxylate shunt providing malate from acetyl-CoA, which is derived, for example, from  $\beta$ -oxidation of fatty acids. Both enzymes are induced upon glucose depletion<sup>95</sup> and assimilation of the intracellular reserve lipids.<sup>96</sup> The mutation leading to itaconate resistance might increase the capacity of the glyoxylate shunt favoring anaplerosis. In fact, in a protein extract of one of the itaconate-resistant mutants, which was characterized in more detail, increased isocitrate lyase activity was determined, and 3,4-dihydroxy-2-butanone 4-phosphate synthase activity was also significantly increased.<sup>72</sup> The genomic lesion in the mutant leading to itaconate resistance could not be localized within the isocitrate lyase gene or its upstream region. Obviously, the itaconate resistance resulted from a more pleiotropic mutation, affecting the glyoxylate shunt as well as the riboflavin pathway.

The enzymes of the glyoxylate shunt are localized in the peroxisomes of *A. gossypii*.<sup>97</sup> It was anticipated that interference with the peroxisomal, NADP-dependent isocitrate dehydrogenase reaction should benefit growth and riboflavin production, since more isocitrate would become available for the isocitrate lyase reaction due to interference with the competing dehydrogenase reaction. However, knockout of the peroxisomal isocitrate dehydrogenase led to a reduced production of riboflavin. Five-fold overexpression of the enzyme using the constitutive *TEF* promoter led to a more than 50% increase in riboflavin production. It was presumed that overexpression of the peroxisomal, NADP-dependent isocitrate dehydrogenase improved the intracellular levels of NADPH, which might be limiting in lipid-assimilating *A. gossypii*.<sup>98</sup>

#### **7.04.5.4 Riboflavin Transmembrane Transport**

A still unresolved question in the engineering of riboflavin-overproducing *B. subtilis* strains is how the product is excreted from the producing microorganisms. A long-term candidate for a *B. subtilis* riboflavin transporter was a putative membrane protein encoded by *ypaa*.<sup>99</sup> Besides the *rib* operon *ypaa* was identified as the only other gene deregulated in *B. subtilis* *ribC* mutants<sup>100</sup> in line with the presence of an FMN-responsive RFN element in the *ypaa* leader sequence.<sup>38</sup> Meanwhile the physiological function of Ypaa, now renamed RibU, as a riboflavin importer spanning the bacterial membrane with five hydrophobic  $\alpha$ -helices has been confirmed. A proton symporter mode of action has been proposed.<sup>69</sup> In contrast, for the *ribU* orthologue of *Lactococcus lactis* a facilitated diffusion mechanism was suggested.<sup>101</sup> RibU is not required for riboflavin excretion from *B. subtilis* production strains since a *ribU* knock-out mutation had no effect on the performance of the strains (DSM, unpublished).

*Bacillus subtilis* can excrete large amounts of riboflavin without modifying or engineering an exporter function indicating that a membrane protein, maybe one of the many ABC transporters encoded by the *B. subtilis* genome with still unknown function, is available with the right specificity and sufficient capacity to remove excess riboflavin from the bacterial interior. Alternatively, the planar riboflavin molecule might diffuse through the lipid bilayer of the bacterial membrane and accumulate in the cultivation medium spontaneously. The low solubility of riboflavin in aqueous media of about  $0.5 \text{ g l}^{-1}$  ( $\sim 1.3 \text{ mmol l}^{-1}$ ) might facilitate a passive diffusion process.

In *A. gossypii* kinetic studies of riboflavin import and excretion using riboflavin, auxotrophic and overproducing *A. gossypii* mutants indicated the presence of active riboflavin transport mechanisms in the fungus.<sup>102</sup> Inhibition of riboflavin uptake by FMN or FAD led to an increase in the apparent riboflavin efflux in the early

production phase indicating the presence of a separate excretion carrier. Excretion rates in the mid-production phase were determined to be  $2.5 \text{ nmol min}^{-1} \text{ g}^{-1}$  dry weight for wild-type cells and more than 20-fold increased to  $66.7 \text{ nmol min}^{-1} \text{ g}^{-1}$  dw for an overproducing mutant.

Riboflavin-overproducing *A. gossypii* strains excrete the vitamin not only into the cultivation medium, but a significant amount is transported into the vacuole and is retained there. In contrast to *S. cerevisiae* or *N. crassa* disruption of *VMA1* encoding the vacuolar ATPase subunit A did not interfere with the viability of *A. gossypii*, but *VMA1* deletion mutants were devoid of any detectable vacuolar riboflavin and completely excreted the vitamin into the culture broth.<sup>103</sup>

The *MCH5* gene of *S. cerevisiae* encodes a transporter protein involved in riboflavin uptake by a facilitated diffusion mechanism.<sup>104</sup> Evidence for the functional assignment came from growth assays, protein localization, and uptake assays. Expression of *MCH5* in a heterologous system, *Schizosaccharomyces pombe*, induced riboflavin uptake across the *S. pombe* plasma membrane. However, the data do not formally exclude the possibility that MCH5p is a positive regulator of riboflavin transport. The function of *A. gossypii* ADL156Cp (176/443, i.e., 39% identities with MCH5) has not been reported.

A dedicated riboflavin exporter protein might be the multidrug transporter breast cancer resistance protein BCRP1, which is strongly induced in the mammary gland of mice during pregnancy and lactation.<sup>105</sup> In *Bcrp1* (−/−) mice, secretion of riboflavin into the milk was still detectable but reduced to less than 2% compared to wild type. It was concluded that BCRP1 is responsible for pumping riboflavin into the milk, thus supplying the pups with the vitamin.

#### 7.04.5.5 Energetic Efficiency of the *Bacillus subtilis* Respiratory Chain

In addition to the central metabolic and the product-specific biosynthetic pathways, the maintenance metabolism of a host strain requires special attention in a developmental program for high-performance production strains. Maintenance metabolism is defined as the carbon source and oxygen consumption of a microbial culture at an extrapolated growth rate of zero corrected for the requirements for product formation. The maintenance demand, interpreted as the energetic requirements of an organism to remain in a viable and active state, is an intrinsic property of an organism and becomes a critical process determinant in industrial fed-batch fermentations with slow-growing cells at high density. A particular problem of *B. subtilis* as a production host strain is its comparatively high maintenance energy demand.<sup>106</sup> As most other bacteria *B. subtilis* possesses a branched respiratory chain consisting of both quinol and cytochrome *C* terminal oxidases. Depending on the composition of the respiratory chain, between one and four protons may be translocated over the membrane per electron reacting with oxygen. The fraction of free energy released by the oxidation of electron carriers like NADH/H<sup>+</sup> and FADH<sub>2</sub> that is captured by building up a proton gradient and drives ATP synthesis, is dependent on the distribution of the electron flux through the various branches of the respiratory chain. Knockout of the quinol cytochrome *bd* branch (encoded by *cydAB*) in the riboflavin production strain RB50::[pRF69]*n* (see Section 7.04.5.2.2) reduced the maintenance demand by about 40% from 0.67 to 0.39 mmol glucose per gram biomass per hour.<sup>107</sup> Since the putative Yth and the cytochrome *caa3* oxidases are of minor importance in this strain, the most likely explanation for this reduction is the translocation of two protons per transported electron via the remaining cytochrome *aa3* oxidase, instead of only one proton via the cytochrome *bd* oxidase. The maintenance demand reduction significantly improves the yield of biomass and riboflavin in fed-batch cultures of *cyd* knockout *B. subtilis* production strains.<sup>108</sup>

The beneficial effect of a *cyd* knockout mutation on reducing maintenance demand and riboflavin production was also demonstrated with the *B. subtilis* riboflavin production strain PK. In this strain background the startlingly high maintenance coefficient of 1.9 mmol glucose per gram biomass per hour was reported to be reduced by 26% along with a 30% higher riboflavin production rate.<sup>109</sup>

#### 7.04.5.6 Classical Strain Improvement and Whole Genome Sequencing

In addition to metabolic engineering of the specific riboflavin pathway and its immediate supply routes, that is, purine and pentose phosphate pathways, improving of general host properties that relate to the performance under industrial process conditions is mandatory for commercially viable production processes. Industrial

process conditions typically include low growth rates, high process intensity, high material flow rates, cultivation in the more inhomogeneous environment of large-scale bioreactors compared to laboratory-scale reactors, and high stirrer tip speed in aerated cultures. Much of the contemporary industrial practise to address general host properties still exploits the traditional approach of random, mutagen-forced mutagenesis and screening or selection for strains with improved traits. Microtiter plate-based screening devices, robot-assisted sampling, and advanced analytical tools, for instance high-throughput NMR, have become available and allow high-throughput screening of reasonably large numbers of mutated bacterial clones.

The aforementioned highly engineered DSM *B. subtilis* production strain RB50::[pRF69]*m*::[pRF93]*n* (see Section 7.04.5.2.2) was subjected to several rounds of random mutagenesis and selection before being used as the strain of choice for industrial riboflavin production (DSM, unpublished). Mutants resistant against threonine or proline analogues were selected from a mutagenized culture of the *B. subtilis* riboflavin production strain AS5 by researchers from the CJ Corporation of Korea.<sup>110</sup> Riboflavin titers of 22.4 and 26.5 g l<sup>-1</sup> in 60–70 h fed-batch fermentation runs were obtained with the parent strain AS5 and the analogue-resistant strains, respectively.

While still indispensable for successful industrial strain development programs, the drawback of brute-force random mutagenesis is the accumulation of irrelevant or even counterproductive mutations that often result in crippled production hosts.<sup>111,112</sup> Low-cost whole genome resequencing techniques developed recently by companies such as NimbleGen, Affymetrix, 454 Life Sciences, or Solexa<sup>113–115</sup> offer the possibility to relate SNPs (single nucleotide polymorphisms) in strains obtained from classical mutagenesis programs to their respective phenotype. The relatively small size of bacterial genomes ensures the technical and economical feasibility of SNP analysis for production strain development. The technique allows to distinguish between mutations beneficial for metabolite overproduction and collateral damage from the mutagenesis procedure facilitating the rational reconstruction of production strains without the otherwise accumulating collateral traits. Furthermore, SNP analysis in production strain lineages can provide a link between classical strain improvement and metabolic engineering, because the technique facilitates the identification of genes, which are not directly involved in product formation and due to the tremendous complexity of the dynamic interactions in cellular systems are difficult to single out with some certainty as promising engineering targets by mere biochemical or physiological considerations. Once identified these nonobvious target genes can be subjected to the whole arsenal of advanced genetic engineering techniques to modify their expression or the characteristics of the encoded gene products. Such methods are starting to find their ways into industrial strain engineering programs, exemplified already for L-lysine production strain improvement.<sup>116,117</sup>

## 7.04.6 Process Development

### 7.04.6.1 Glucose and Biotin-Limited Fed-Batch Fermentation Process for *Bacillus subtilis* Production Strains

Carbon source-limited fed-batch protocols are frequently applied in industrial fermentation processes. During an initial batch phase all substrates required for microbial growth are available in excess and the inoculated microorganisms can grow at their maximal possible rate sometimes enhanced by the presence of complex media components. After complete consumption, the carbon source is pumped into the fermenter at a rate that subjects the microbial growth rate to the rate of substrate supply. The fermentation process developed for cultivation of prototrophic *B. subtilis* riboflavin production strains like RB50::[pRF69]*n*<sup>60</sup> and derivatives of this strain follows the common strategy of carbon source-limited fermentation processes. During the initial batch phase, mainly biomass and only minor amounts of riboflavin are produced. Riboflavin is synthesized and excreted into the culture broth at low growth rates under the strictly glucose-limited conditions of the feeding phase. The *Bacillus*-based fed-batch process for commercial riboflavin production could be optimized using model-predictive control based on artificial neural networks. The optimum feed profile for increasing the product yield and the amount of riboflavin at the time of harvesting was adjusted continuously and applied to the process. Both the total amount of riboflavin produced and the product yield increased by more than 10%.<sup>118</sup>



Glucose-limited continuous fermentation runs revealed a dichotomy between riboflavin productivity, that is, space–time yield versus product yield on the consumed carbon source (DSM, unpublished), an observation not uncommon in microbial production processes. To facilitate a riboflavin process of high yield on the carbon source and high productivity, a decoupled fermentation regime was developed, in which the rate of biomass and riboflavin production is controlled by different fermentation substrates. Controlling the growth rate of amino acid auxotrophic derivatives of *B. subtilis* production strains through limiting supply of amino acids had a negative effect on riboflavin productivity. The reason could be general detrimental effects on protein synthesis due to amino acid shortage or induction of the *relA*-mediated stringent response caused by unloaded tRNAs. Biotin-limited growth of the biotin auxotrophic derivative RB50::[pRF69]*n* Bio<sup>-</sup> of the riboflavin production strain RB50::[pRF69]*n* proved to be a suited strategy.<sup>119</sup> In chemostat fermentations with RB50::[pRF69]*n* Bio<sup>-</sup> glucose was the growth-limiting fermentation substrate if the feed medium contained 0.5 µg biotin per gram glucose. Under these conditions biomass and riboflavin were produced with a 29.4 and 3.04% yield on glucose, respectively. The specific riboflavin productivity was 15.5 mg riboflavin per g biomass per h. Upon reduction of the biotin-to-glucose ratio to 0.15 µg g<sup>-1</sup> the biomass yield on glucose was reduced to 17.0% indicating that biotin had a growth-limiting effect. The riboflavin yield on glucose was increased to 4.05%. The specific riboflavin productivity more than doubled to 35.5 mg riboflavin per g biomass per hour.

#### 7.04.6.2 *Ashbya gossypii* Fermentation Process with Plant Oils as Carbon and Energy Source

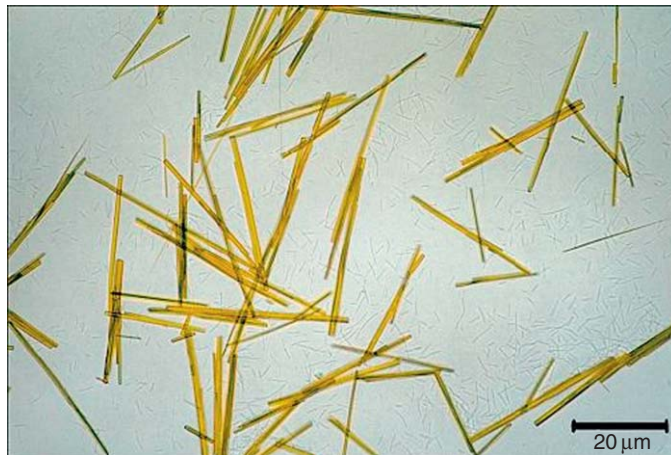
*Ashbya gossypii* riboflavin production processes utilize plant oils as carbon and energy source. Since the oils are present in the fermentation broth as osmotically inactive emulsified lipid droplets and rigid fat particles, a high nutrient load is possible. *Ashbya gossypii* secretes a 35 kDa lipase<sup>120</sup> hydrolyzing the triglycerides in the culture medium into free fatty acids, which are then ingested by the fungus. The lipase is inactivated within minutes presumably due to aggregation at the lipid/water interface. Furthermore, elevated concentrations of free fatty acids interfere with lipase secretion.<sup>120</sup> During the *A. gossypii* fermentation process with plant oils the carbon dioxide development and the dissolved oxygen concentration in the fermentation broth oscillate in counter-phase. Addition of a stable *Candida* lipase or free fatty acids to the fermentation broth interfered with the oscillation and kept the carbon dioxide development and the dissolved oxygen at their maximal and minimal levels, respectively, for several hours.<sup>121</sup> These observations indicate a periodic change in the availability of the carbon source during the fermentation run caused by the periodic inhibition of the lipase secretion by elevated concentrations of free fatty acids. Obviously, the fermentation does not reach a steady-state equilibrium with a lipase activity in the broth that produces just the amount of fatty acids that is taken up by the fungus.

#### 7.04.6.3 Product Isolation and Purification

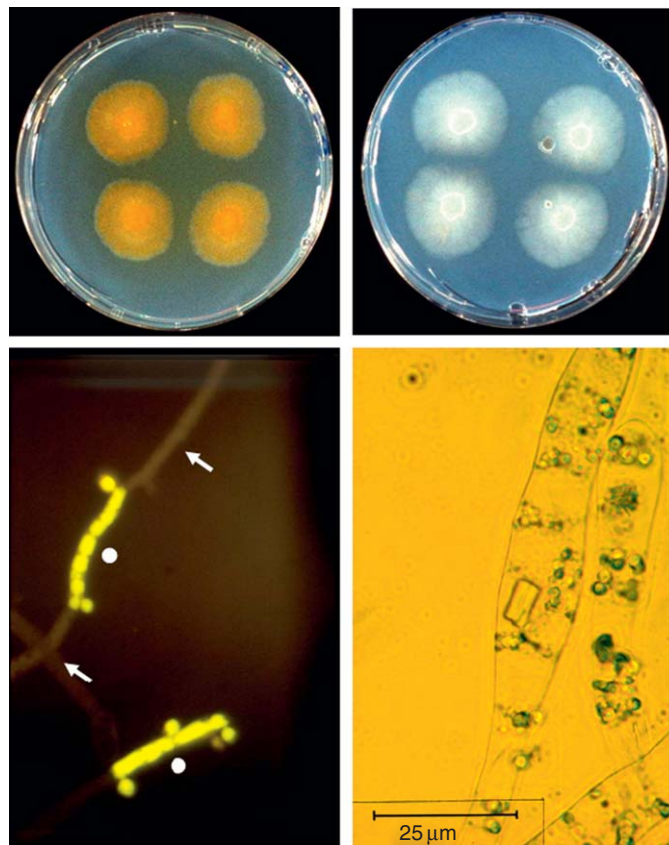
Since riboflavin is rather insoluble at neutral pH in aqueous solutions the vitamin accumulates in the fermentation broth as needle-shaped crystals (**Figure 6**). At the end of the fermentation run with the *B. subtilis* production strain RB50::[pRF69]*n*::[pRF93]*m* the fermentation broth containing riboflavin crystals and biomass was pasteurized and subjected to differential centrifugation to harvest and partially purify the crystals. After a washing step with 0.2 mol l<sup>-1</sup> hot mineral acid, a feed quality product of >96% purity was obtained, which was devoid of any recombinant DNA as revealed by PCR analysis.<sup>122</sup> A food and pharma product (>98% purity) was obtained by recrystallization of riboflavin from hot concentrated acid. The feed-grade riboflavin material contained small amounts of amino acids and amino sugars and the biosynthetic riboflavin precursor dimethyl-ribityl-lumazine. Other minor impurities were identified as riboflavin degradation products resulting from the purification procedure, which were also present in riboflavin obtained by chemical synthesis.

Riboflavin-overproducing *A. gossypii* store significant amounts of the product as intracellular crystals (**Figure 7**). A heating step after completion of the main fermentation run with *A. gossypii* induces self-lysis and liberates riboflavin from the biomass.<sup>123</sup> In addition, heating of the fermentation broth and slow cooling over several hours promote growth of the riboflavin crystals in the broth facilitating the separation of the crystals from the biomass by decantation.<sup>124</sup> Further purification is attained by resuspension of the





**Figure 6** Bright yellow needle-like riboflavin crystals in the culture medium of *Bacillus subtilis* production strain RB50::[pRF69]n::[pRF93]m. The *B. subtilis* cells are seen as small rods in the background.



**Figure 7** Growing disks of riboflavin-overproducing *Ashbya gossypii* mycelium on the surface of agar plates (top left). Chlorophenolthio-cAMP ( $5 \text{ mmol l}^{-1}$ ) interferes with riboflavin overproduction (top right). Riboflavin accumulating (white dot) and nonaccumulating cells (white arrow) within *Ashbya gossypii* hyphae (bottom left). Riboflavin crystals within *A. gossypii* vacuoles (bottom right).

riboflavin-containing precipitate in diluted aqueous acids, heating, and decantation.<sup>125</sup> Since *A. gossypii* production strains with increased copy numbers of the *RIB* genes are classified self-clones (see Section 7.04.5.2.4) remains of the production organism in the feed application product are tolerated.

### 7.04.7 Conclusions

Riboflavin is unique in the sense that for its industrial production biosynthetic processes have almost completely replaced the well-established and efficient chemical processes. Microbial production of the vitamin C precursor compound 2-keto-L-gulonic acid (2KGA) from D-sorbitol comes close, but still two chemical steps, catalytic D-glucose hydrogenation to D-sorbitol and costly lactonization of 2KGA, are required to afford vitamin C. Furthermore, vitamin C synthesis via 2KGA is an artificial reaction schema, as the two natural vitamin C biosynthetic pathways in plants and animals follow different routes. A microbial production process for pantothenic acid biosynthesized in a mechanistically quite simple and well-understood pathway might be close to industrial application, since highly effective *B. subtilis*-based production strains were constructed recently.<sup>126</sup> Biosynthetic processes to other water-soluble vitamins like thiamine, pyridoxol, biotin, or folic acid are far from being economically feasible according to the published academic and patent literature. The biocatalytic mechanisms toward these compounds involve particular and uncommon mechanisms of carbon bond rearrangements (hydroxymethylpyrimidine moiety of thiamine), heterocycle formation (pyridoxol), or sulfur incorporation (biotin). Once the molecular details of these pathways are elucidated to a level already reached for the riboflavin pathway involving serious biochemical studies but also holistic approaches provided by the various 'omics' methods construction of efficient production strains for all of the water-soluble vitamins should become a promising endeavor.

The success story of fermentative riboflavin production might not have come to an end. Considering the maximum theoretical riboflavin yield on glucose in the absence of growth and maintenance requirements of about 55% w/w<sup>127</sup> there might still be room for further strain and process improvement. Product yields exceeding 50% on the carbon source, the key parameter for economic success of bioprocesses particular in times of persistently rising carbon source quotes, are possible as the microbial production process for L-lysine involving a multistep, certainly not trivial biosynthetic pathway demonstrates.

It can be expected that the yield improvement might result from the application of new technological developments, like whole genome sequencing and the 'omics' methods. The impact of SNP analysis in production strain lineages for target gene identification has been discussed above (see Section 7.04.5.6). The available 'omics' methods for *B. subtilis* including different formats of mRNA expression microarrays<sup>100,128</sup> and proteomics technologies, which allow identification and quantification of several hundred *B. subtilis* proteins<sup>129,130</sup> are well advanced. Metabolomics platforms detect several hundred *B. subtilis* metabolites,<sup>131,132</sup> but a large number of them remains to be identified.

Novel target genes for the improvement of lysine production with *C. glutamicum* MH20-22B, a standard laboratory lysine strain of comparatively low productivity, could be identified by global gene expression analysis.<sup>133</sup> Whether fully developed industrial production strains can be further improved by applying similar methods has not become public so far. Despite the availability of highly developed technologies for holistic determination of intracellular concentrations (mRNAs, proteins, metabolites) a true success story of 'omics' applications in industrial biotechnology has not been published yet. Possibly, these technologies are still not precise enough to guide the subtle fine-tuning of the metabolic network of a production strain that has already reached an advanced stage of development. Furthermore, the immense data sets resulting from 'omics' experimentation await a truly integrated mathematical/computational framework, such as, for example, network-embedded thermodynamic analysis,<sup>134</sup> to capture the complex and interdependent component interactions and processes that define a living cell. The principles and methods of systems biology could become a key for the successful application of data-driven approaches of strain development.<sup>135,136</sup>

Improved production strains might also result from the systematic removal of nonessential DNA from the genome of production host strains. From a total of 4101 genes of *B. subtilis* only 271 genes were found to be essential for growth in a standard laboratory rich medium at 37 °C.<sup>137</sup> A few hundred more genes might be required to facilitate bacterial growth in a mineral salt medium with a sugar carbon source. Reducing the chromosomal complexity will save the metabolite and energy resources to maintain and express the deleted genetic information. Furthermore, deletion of autolysin genes within prophage elements frequently encountered in bacterial genomes will eliminate the risk of cell lysis during industrial fermentations due to autolysin

activation. Finally, the behavior of production strains derived from minimal genome host strains should be become more predictable, which is an important prerequisite for rational strain design. A recent report demonstrated the increased capability to produce recombinant extracellular digestive enzymes in the *Bacillus subtilis* strain MBG874 depleted of 874 kb (20%) of genomic DNA encoding among others the prophages and prophage-like regions *SP $\beta$* , *PBSX*, *skin*, and *pro1-7*, and the polyketide and plipastatin synthesis operons *pks* and *pps*.<sup>138</sup>

Two quite different microbial production systems are currently being used for industrial riboflavin production, one based on *A. gossypii* using vegetable oil as major fermentation substrate, the other using engineered *B. subtilis* strains, which are fed with carbohydrates. Both systems coexist in the highly competitive environment of the riboflavin market, but which of them will finally prevail will depend on many factors. Raw material price development for carbohydrates and vegetable oils will be an important extrinsic factor. Intrinsically, it will be decisive which of the two microbial systems can be engineered further such that the good part of the metabolic flux can be diverted away from biomass and carbon dioxide production toward riboflavin. The microorganism accomplishing this ambitious task more effectively will be the riboflavin production strain of the future.

## References

1. F. Müller, Ed., *Chemistry and Biochemistry of Flavoenzymes*. CRC Press: Boca Raton, FL, 1992.
2. B. Palfey; V. Massey, Flavin-dependent enzymes. In *Comprehensive Biochemical Catalysis*; M. Sinnott, Ed.; Academic Press: New York, 1998; pp 83–154.
3. K. P. Stahmann; H. N. Arst, Jr.; A. Althöfer; J. L. Revuelta; N. Monschau; C. Schlüpen; C. Gätgens; A. Wiesenburg; T. Schlösser, *Environ. Microbiol.* **2001**, 3 (9), 545–550.
4. R. Kuhn; K. Reinemund, *et al.*, *Ber. Dtsch. Chem. Ges.* **1934**, 67, 1460–1463.
5. P. Karrer; K. Schöpp, *et al.*, *Helv. Chim. Acta* **1935**, 18 (1), 426–429.
6. M. Tishler; K. Pfister, *et al.*, *J. Am. Chem. Soc.* **1947**, 69, 1487–1492.
7. K.-I. Sasajima; M. Doi, *et al.*, Verfahren zur Herstellung von D-Ribose. DE2454931, Takeda Chemical Industries Ltd., 1975.
8. R. Kurth; J. Paust, *et al.*, *Vitamins*; VCH Verlagsgesellschaft: Weinheim, 1996; Chapter 7.
9. L. Wickerham; M. Flickinger, *et al.*, *Arch. Biochem.* **1946**, 9, 95–98.
10. D. Perlman, Microbial Process for Riboflavin Production. In *Microbial Technology, Microbial Processes*, 2nd ed.; H. Pepler, D. Perlman, Eds.; Academic Press: New York, 1979; Vol. 1, pp 521–527.
11. B. Lago; L. Kaplan, *Adv. Biotechnol.* **1981**, 3, 241–246.
12. K. Hoppenheidt; W. Mücke, *et al.*, Reducing environmental load of chemical engineering processes and chemical products by biotechnological substitutes. *Summary of the Final Report. Project UFOPLAN 202 66 326*; Bayerisches Institut für Angewandte Umweltforschung und -technik GmbH: Augsburg, Germany, 2004.
13. R. Volk; A. Bacher, *J. Biol. Chem.* **1991**, 266 (31), 20610–20618.
14. M. Fischer; W. Romisch, *et al.*, *Biochem. Soc. Trans.* **2005**, 33 (4), 780–784.
15. A. Bacher, Biosynthesis of Flavins. In *Chemistry and Biochemistry of Flavoenzymes*; F. Muller, Ed.; CRC Press: Boca Raton, FL, 1990; Vol. 1, pp 215–259.
16. A. Bacher; S. Eberhardt, *et al.*, *Annu. Rev. Nutr.* **2000**, 20, 153–167.
17. M. Fischer; A. Bacher, *Nat. Prod. Rep.* **2005**, 22 (3), 324–350.
18. J. Perkins; J. Pero, *et al.*, Riboflavin Overproducing Strains of Bacteria. EP0405370, Hoffmann La Roche, 1991.
19. V. Mironov; A. Kraev, *et al.*, *Dokl. Akad. Nauk. SSSR* **1989**, 305 (2), 482–487.
20. B. Veith; C. Herzberg, *et al.*, *J. Mol. Microbiol. Biotechnol.* **2004**, 7 (4), 204–211.
21. A. Bolotin; P. Wincker, *et al.*, *Genome. Res.* **2001**, 11 (5), 731–753.
22. J. Kalinowski; B. Bathe, *et al.*, *J. Biotechnol.* **2003**, 104 (1–3), 5–25.
23. F. R. Blattner; G. Plunkett 3rd, *et al.*, *Science* **1997**, 277 (5331), 1453–1474.
24. O. Oltmanns; A. Bacher, *J. Bacteriol.* **1972**, 110 (3), 818–822.
25. D. Revuelta; G. Santos, *et al.*, Riboflavin Synthesis in Yeast. WO9411515, BASF AG, 1994.
26. D. Revuelta; S. Buitrago, *et al.*, Riboflavin Synthesis in Fungi. WO9526406, BASF AG, 1995.
27. K. Dmytruk; C. Abbas, *et al.*, *Ukr. Biokhim. Zh.* **2004**, 76 (1), 78–87.
28. A. I. Stepanov; A. Kukanova, *et al.*, *Genetika* **1977**, 13 (3), 490–495.
29. K. Matsui; H. C. Wang, *et al.*, *Agric. Biol. Chem.* **1982**, 46 (8), 2003–2008.
30. V. N. Mironov; A. S. Kraev, *et al.*, *Mol. Gen. Genet.* **1994**, 242 (2), 201–208.
31. A. Sorokin; E. Zumstein, *et al.*, *Mol. Microbiol.* **1993**, 10 (2), 385–395.
32. S. E. Bresler; E. I. Cherepenko, *et al.*, *Genetika* **1971**, 7 (11), 1466–1470.
33. Y. V. Kil; V. N. Mironov, *et al.*, *Mol. Gen. Genet.* **1992**, 233 (3), 483–486.
34. C. M. Burgess; E. J. Smid, *et al.*, *Microb. Cell. Fact.* **2006**, 5, 24.
35. D. Coquard; M. Huecas, *et al.*, *Mol. Gen. Genet.* **1997**, 254 (1), 81–84.
36. I. Gusarov; R. Kreneva, *et al.*, *Mol. Biol. (Mosk)* **1997**, 31, 820–825.
37. M. Mack; A. P. G. M. van Loon, *et al.*, *J. Bacteriol.* **1998**, 180 (4), 950–955.
38. M. S. Gelfand; A. A. Mironov, *et al.*, *Trends Genet.* **1999**, 15 (11), 439–442.

39. J. B. Perkins; J. Pero, Biosynthesis of Riboflavin, Biotin, Folic acid, and Cobalamin: From Genes to Cells. In *Bacillus subtilis and Its Closest Relatives*; A. L. Sonenshein, J. A. Hoch, R. Losick, Eds.; ASM Press: Washington, DC, 2002, pp 271–286.
40. A. S. Mironov; I. Gusarov, *et al.*, *Cell* **2002**, *111* (5), 747–756.
41. W. C. Winkler; S. Cohen-Chalamish, R. R. Breaker, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (25), 15908–15913.
42. T. Schlosser; G. Schmidt, *et al.*, *Microbiology* **2001**, *147* (Pt. 12), 3377–3386.
43. T. Schlosser; A. Wiesenburg, *et al.*, *Appl. Microbiol. Biotechnol.* **2007**, *76* (3), 569–578.
44. I. Nogami; M. Katsumata, *et al.*, Method for the Production of Inosine. US3616206, 1965.
45. V. Debabov, The Industrial Use of Bacilli. In *The Molecular Biology of the Bacilli*; D. A. Dubnau, Ed.; Academic Press: New York, 1982, pp 331–370.
46. Y. Sumino; K. Sono, *et al.*, Producing Nucleosides. US4578336, 1982.
47. I. Shio, Production of Primary Metabolites. In *Bacillus subtilis: Molecular Biology and Industrial Application*; B. Maruo, H. Yoshikawa, Eds.; Kodansha Ltd. and Elsevier Science Publishers: Tokyo, Japan, 1989.
48. A. Sonenshein; J. Hoch, *et al.*, Eds., *Bacillus subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics*; American Society for Microbiology: Washington, DC, 1993.
49. A. L. Sonenshein; J. A. Hoch, *et al.*, Eds., *Bacillus subtilis and Its Relatives: From Genes to Cells*. American Society for Microbiology: Washington, DC, 2001.
50. M. C. Wright; P. Philippsen, *Gene* **1991**, *109* (1), 99–105.
51. J. Wendland; Y. Ayad-Durieux, *et al.*, *Gene* **2000**, *242* (1–2), 381–391.
52. A. Gattiker; R. Rischatsch, *et al.*, *BMC Genomics* **2007**, *8*, 9.
53. L. I. Panina; I. V. Iomantas, *et al.*, *Genetika* **1983**, *19* (1), 174–176.
54. A. I. Stepanov; V. G. Zhdanov, *et al.*, Riboflavin Preparation. FR2546907, Inst Genetiki i Selektcii, 1984.
55. N. V. Stoinova; E. G. Abalakina, *et al.*, *Biotehnologiya* **1996**, *11*, 1–5.
56. N. V. Stoinova; E. G. Abalakina, *et al.*, *Biotehnologiya* **1996**, *11*, 7–10.
57. A. S. M. Mironov; K. A. Valentinova, *et al.*, Method for Producing Riboflavin. WO2004046347, State Res Inst of Genetics, 2004.
58. T. Chen; X. Chen, *et al.*, *Trans. Tianjin Univ.* **2005**, *11* (1), 1–5.
59. T. Chen; J. Y. Wang, *et al.*, *J. Wuxi Univ. Light Ind.* **2005**, *24* (1), 6–10.
60. J. B. Perkins; A. Sloma, *et al.*, *J. Ind. Microbiol. Biotechnol.* **1999**, *22* (1), 8–18.
61. A. M. Albertini; A. Galizzi, *J. Bacteriol.* **1985**, *162* (3), 1203–1211.
62. U. Sauer; V. Hatzimanikatis, *et al.*, *Nat. Biotechnol.* **1997**, *15* (5), 448–452.
63. M. Hümbelin; V. Griesser, *et al.*, *J. Ind. Microbiol. Biotechnol.* **1999**, *22* (1), 1–7.
64. X. Zhao; T. Chen, *et al.*, A Genetic-Engineering Bacterial Strain Producing Riboflavin and Its Construction Method. CN1891813 A, Tianjin University, 2007.
65. X. Zhao; T. Chen, *et al.*, Riboflavin-Produced Engineering Strain and Its Method for Producing Riboflavin. CN1891814 A, Tianjin University, 2007.
66. Koizumi; S. S. Teshiba, *J. Ferment. Bioeng.* **1998**, *86*, 130–133.
67. S. Koizumi; Y. Yonetani, *et al.*, *Appl. Microbiol. Biotechnol.* **2000**, *53* (6), 674–679.
68. A. G. Vitreschak, D. A. Rodionov, *et al.*, *Nucleic Acids Res.* **2002**, *30* (14), 3141–3151.
69. C. Vogl; S. Grill, *et al.*, *J. Bacteriol.* **2007**, *189* (20), 7367–7375.
70. H. Althöfer; H. Seuberger, *et al.*, Genetic Method for Producing Riboflavin. WO 9961623, BASF AG, 1999.
71. S. Steiner; P. Philippsen, *Mol. Gen. Genet.* **1994**, *242* (3), 263–271.
72. G. Schmidt; K. Stahmann, *et al.*, *Microbiology* **1996**, *142*, 419–426.
73. H. Althöfer; D. J. Revuelta, Genetic Strain Optimization for Improving the Production of Riboflavin. WO03048367 A1. DE, BASF AG, 2003.
74. S. Casas-Flores; T. Rosales-Saavedra, *et al.*, *Methods. Mol. Biol.* **2004**, *267*, 315–325.
75. K. Miyagawa; H. Kimura, *et al.*, *Bio/Technology* **1986**, *4*, 225–228.
76. V. N. Gershanovich; T. N. Bol'shakova, *et al.*, *Mol. Gen. Mikrobiol. Virusol.* **2005**, *3* (3), 29–34.
77. A. Jimenez; M. A. Santos, *et al.*, *Appl. Environ. Microbiol.* **2005**, *71* (10), 5743–5751.
78. A. L. Demain, *Annu. Rev. Microbiol.* **1972**, *26*, 369–388.
79. N. Monschau; K. P. Stahmann, *et al.*, Fungi of the Genus *Ashbya* for the Production of Riboflavin. EP0930367, 1999.
80. N. Monschau; H. Sahm, *et al.*, *Appl. Environ. Microbiol.* **1998**, *64* (11), 4283–4290.
81. C. Schlupen; M. A. Santos, *et al.*, *Biochem. J.* **2003**, *369* (Pt. 2), 263–273.
82. K. Sasajima; M. Yoneda, *Biotechnol. Genet. Eng. Rev.* **1984**, *2*, 175–213.
83. P. De Wulf; E. J. Vandamme, *Appl. Microbiol. Biotechnol.* **1997**, *48*, 141–148.
84. V. N. Gershanovich; A. Kukanova, *et al.*, *Mol. Gen. Mikrobiol. Virusol.* **2000**, *3*, 3–7.
85. M. Ikeda; K. Okamoto, *et al.*, Method for Producing Metabolites Biologically Synthesized via Phosphoribosyl Pyrophosphate. US6258554, Kyowa Hakko, 2001.
86. M. Lehmann; H. P. Hohmann, *et al.*, Modified Transketolase and Use Thereof. EP1957640, DSM, 2008.
87. U. Nilsson; L. Meshalkina, *et al.*, *J. Biol. Chem.* **1997**, *272* (3), 1864–1869.
88. Y. Zhu; X. Chen, *et al.*, *Biotechnol. Lett.* **2006**, *28* (20), 1667–1672.
89. P. Servant; D. Le Coq, *et al.*, *Mol. Microbiol.* **2005**, *55* (5), 1435–1451.
90. H. P. Hohmann; S. Aymerich, *et al.*, Novel Genes Involved in Gluconeogenesis. EP1831249, 2007.
91. Y. Zhu; X. Chen, *et al.*, *FEMS Microbiol. Lett.* **2007**, *266* (2), 224–230.
92. N. Zamboni; H. Maaheimo, *et al.*, *Metab. Eng.* **2004**, *6* (4), 277–284.
93. E. Fischer; U. Sauer, *Nat. Genet.* **2005**, *37* (6), 636–640.
94. N. Zamboni; E. Fischer, *et al.*, *Biotechnol. Bioeng.* **2005**, *89* (2), 219–232.
95. I. Maeting; G. Schmidt, *et al.*, *FEBS Lett.* **1999**, *444* (1), 15–21.
96. K. P. Stahmann; C. Kupp, *et al.*, *Appl. Microbiol. Biotechnol.* **1994**, *42* (1), 121–127.
97. I. Maeting; G. Schmidt, *et al.*, *J. Mol. Catal.* **2000**, *10*, 335–343.
98. H. Althöfer; O. Zelder, *et al.*, Monocellular or Multicellular Organisms for the Production of Riboflavin. WO0111052, 2001.



99. R. A. Kreneva; M. Gelfand, *et al.*, *Russ. J. Genet.* **2000**, 36, 972–974.
100. J. M. Lee; S. Zhang, *et al.*, *J. Bacteriol.* **2001**, 183 (24), 7371–7380.
101. C. M. Burgess; D. J. Slotboom, *et al.*, *J. Bacteriol.* **2006**, 188 (8), 2752–2760.
102. C. Forster; J. L. Revuelta, *et al.*, *Appl. Microbiol. Biotechnol.* **2001**, 55 (1), 85–89.
103. C. Forster; M. A. Santos, *et al.*, *J. Biol. Chem.* **1999**, 274 (14), 9442–9448.
104. P. Reihl; J. Stolz, *J. Biol. Chem.* **2005**, 280 (48), 39809–39817.
105. A. E. van Herwaarden; E. Wagenaar, *et al.*, *Mol. Cell. Biol.* **2007**, 27 (4), 1247–1253.
106. U. Sauer; V. Hatzimanikatis, *et al.*, *Appl. Environ. Microbiol.* **1996**, 62 (10), 3687–3696.
107. N. Zamboni; N. Mouncey, *et al.*, *Metab. Eng.* **2003**, 5 (1), 49–55.
108. H. P. Hohmann; N. J. Mouncey, *et al.*, Fermentation Process. EP1481064, DSM, 2004.
109. X. J. Li; T. Chen, *et al.*, *Appl. Microbiol. Biotechnol.* **2006**, 73 (2), 374–383.
110. H. Choi; J. G. Han, *et al.*, Microorganism Producing Riboflavin and Production Methods of Riboflavin Using Thereof, CJ CORP. KR2003051237, 2003.
111. U. Sauer, *Adv. Biochem. Eng. Biotechnol.* **2001**, 73, 129–169.
112. J. L. Adrio; A. L. Demain, *FEMS Microbiol. Rev.* **2006**, 30 (2), 187–214.
113. T. J. Albert; G. Dailidiene, *et al.*, *Nature. Meth.* **2005**, 2 (12), 951–953.
114. S. T. Bennett; C. Barnes, *et al.*, *Pharmacogenomics* **2005**, 6 (4), 373–382.
115. M. Margulies; M. Egholm, *et al.*, *Nature* **2005**, 437 (7057), 376–380.
116. J. Ohnishi; S. Mitsuhashi, *et al.*, *Appl. Microbiol. Biotechnol.* **2002**, 58 (2), 217–223.
117. M. Ikeda; J. Ohnishi, *et al.*, *J. Ind. Microbiol. Biotechnol.* **2006**, 33 (7), 610–615.
118. K. Kovarova-Kovar; S. Gehlen, *et al.*, *J. Biotechnol.* **2000**, 79 (1), 39–52.
119. H. P. Hohmann; N. J. Mouncey, Process for Producing a Fermentation Product. EP1186664, DSM, 2002.
120. K. P. Stahmann; T. Bodecker, *et al.*, *Eur. J. Biochem.* **1997**, 244 (1), 220–225.
121. K. P. Stahmann, Untersuchungen zum Stoffwechsel des Riboflavin-bildenden Pilzes *Ashbya gossypii*. Universität Düsseldorf, 1999.
122. W. Bretzel; W. Schurter, *et al.*, *J. Ind. Microbiol. Biotechnol.* **1999**, 22 (1), 19–26.
123. R. Kurth, Process for the Enhancement of Riboflavin Levels in Spray-Dried Riboflavinfermentation Extracts. EP0487985, Germany, BASF AG, 1992.
124. T. Faust; J. Meyer, *et al.*, Process for the Separation of Riboflavin from a Fermentation Suspension. EP0438767, 1991.
125. J. Grimmer; H. Kiefer, *et al.*, Process for the Purification of Riboflavin Obtained by Fermentation. EP0464582, Germany, BASF AG, 1992.
126. R. Yocum; T. Patterson, *et al.*, Microorganisms and Processes for Enhanced Production of Pantothenate. EP1390519, BASF AG, 2004.
127. M. Dauner; U. Sauer, *Biotechnol. Bioeng.* **2001**, 76 (2), 132–143.
128. B. Jürgen; S. Tobisch, *et al.*, *Biotechnol. Bioeng.* **2005**, 92 (3), 277–298.
129. M. Hecker; U. Volker, *Proteomics* **2004**, 4 (12), 3727–3750.
130. U. Völker; M. Hecker, *Cell. Microbiol.* **2005**, 7 (8), 1077–1085.
131. T. Soga; Y. Ohashi, *et al.*, *J. Proteome. Res.* **2003**, 2 (5), 488–494.
132. M. M. Koek; B. Muilwijk, *et al.*, *Anal. Chem.* **2006**, 78 (4), 1272–1281.
133. G. Sindelar; V. F. Wendisch, *Appl. Microbiol. Biotechnol.* **2007**, 76 (3), 677–689.
134. A. Kummel; S. Panke, *et al.*, *Mol. Syst. Biol.* **2006**, 2, 0034.
135. H. Kitano, *Science* **2002**, 295 (5560), 1662–1664.
136. L. Alberghina; H. Westerhoff, Eds., *Systems Biology Definitions and Perspectives*; Topics in Current Genetics; Springer: Berlin/Heidelberg, 2005.
137. K. Kobayashi; S. D. Ehrlich, *et al.*, Essential bacillus subtilis genes. *Proc. Natl. Acad. Sci. USA* **2003**, 100 (8), 4678–4683.
138. T. Morimoto; R. Kadoya, *et al.*, *DNA Res.* **2008**, 15 (2), 73–81.

## Biographical Sketches



Hans-Peter Hohmann studied chemistry and biochemistry at the Universities of Marburg and Heidelberg, both Germany, between 1976 and 1982. He was a Ph.D. student until 1987 at the Max Planck Institute for Biochemistry, Munich, with Professor Gerisch, working on the biosynthesis of the CsA cell adhesion protein of the social amoeba *Dictyostelium discoideum*. Dr. Hohmann was a research scientist from 1987 to 1991 at Hoffmann-La Roche Inc., Basel, Switzerland, studying cell surface receptors for immunoreactive cytokines including tumor necrosis factor and their intracellular signaling. He moved to the vitamins research department of Roche, first in Basel, then in New Jersey, USA, working on microbial production strain engineering. After his return to Switzerland in 2001 and the takeover of Roche Vitamins Ltd. by Royal DSM N.V. in 2003, he was employed as a principal scientist in the biotechnological research department of DSM Nutritional Products of Switzerland.



Klaus-Peter Stahmann is a professor of applied microbiology at Fachhochschule Lausitz in Senftenberg, Germany. He studied biology at Universität Kaiserslautern (1983–88) focusing on the isolation of antibiotics from fungi in the laboratory of Timm Anke. Afterward he moved to the Research Center Juelich working on the synthesis and degradation of  $\beta$ -glucan by an ascomycete at the Institut für Biotechnologie headed by Hermann Sahm. In 1992 he submitted his doctoral thesis at Heinrich-Heine-Universität Düsseldorf (Germany). Subsequently he founded a research group investigating the metabolism of *Aschbya gosypii* and *Saccharomyces cerevisiae* in Juelich. In 1999 Professor Stahmann became habilitated at the Heinrich-Heine-Universität Düsseldorf, Germany, and received the *venia legendi* for microbiology. In 2002 he moved to Senftenberg where he is in charge of the biotechnology department.



## 7.05 The Use of Subsystems to Encode Biosynthesis of Vitamins and Cofactors

**Andrei L. Osterman**, Burnham Institute for Medical Research, La Jolla, CA, USA

**Ross Overbeek**, Fellowship for Interpretation of Genomes, Burr Ridge, IL, USA

**Dmitry A. Rodionov**, Burnham Institute for Medical Research, La Jolla, CA, USA

© 2010 Elsevier Ltd. All rights reserved.

---

7.05.1	The Goal	141
7.05.2	More Precisely, <i>What Is a Subsystem?</i>	143
7.05.3	How Are Subsystems Built?	150
7.05.4	What Is Revealed by the Construction of Subsystems?	154
7.05.5	The Project to Annotate 1000 Genomes	155
7.05.6	Summary	155
References		156

---

### 7.05.1 The Goal

Our ability to acquire and annotate complete genomes has accelerated rapidly since the first completely sequenced genome became available in 1995.<sup>1</sup> Prior to that point in time, our understanding of the biochemical networks that sustain life were largely limited to a relatively few *model organisms*. This is largely true even for the most universal and the best-studied metabolic pathways, such as biosynthesis of the ubiquitous cofactors and their precursors (vitamins) covered in this volume. Hundreds of complete genomes are currently available for analysis, and the availability of thousands is just a few years away. This wealth of data for the first time opened an opportunity to practically assess the famous statement of Jacques Monod: “What is true for *E. coli* is true for the elephant.” The expectation was that the identification of genes associated with pathways of interest (e.g., biosynthesis of cofactors) would allow us to establish the presence or absence and to reconstruct the details of these pathways across multiple diverse species. Many research groups who understood the power of comparative genome analysis for projecting the knowledge of genes and pathways from model organisms to others have asked the question “How can we organize the data from thousands of genomes to support analysis of specific areas of metabolism?” The initial efforts to address this question led to the establishment of metabolic reconstruction technology based on comparative analysis.<sup>2,3</sup> The development of the first genomic integrations connecting genes with formally encoded biochemical reactions and pathways have undoubtedly benefited the research community.<sup>4–10</sup> In this chapter, we briefly describe a technology for implementing one particular style of organization termed *the subsystems-based approach*.<sup>11</sup> The proven utility of this approach is illustrated here by examples from vitamin and cofactor biosynthesis. We discuss the encoding, projection, assessment of variation, and prediction of novel aspects of representative pathways. Although we largely limit the discussion of specific examples to one subsystem, ‘Biosynthesis of Riboflavin’ (vitamin B<sub>2</sub>) and the derived cofactors, flavin mononucleotide (FMN) and dinucleotide (FAD), the same approach is applicable to many other cofactors described in this volume.

One goal of subsystems-based annotation is to offer researchers interested in a single biochemical process a ‘summary’ of exactly how the process is implemented in each of the sequenced genomes. This summary attempts to clarify the different variations one sees in the basic process, which genes play roles in the process, and which open questions remain. We believe that a well-organized collection of the genomic data becomes a framework to support research into the remaining open questions. A growing collection of subsystems capturing a substantial fraction of the Core Metabolic Machinery projected across hundreds of completely sequenced genomes, as well as the tools for its further expansion and curation is provided in The SEED database.<sup>11</sup> A status of The SEED subsystems (that are also available through the National Medical Pathogen Data Resource, NMPDR,<sup>12</sup>) encoding the biosynthesis of several major cofactors are listed in [Table 1](#).

**Table 1** Examples of vitamin and cofactor biosynthesis subsystems in The SEED database

#	Vitamin	Cofactors	Subsystem in SEED	Roles <sup>a</sup>	Genomes <sup>b</sup>	Genes <sup>c</sup>	Rare roles <sup>d</sup>	Core roles <sup>e</sup>
1	H	Biotin	Biotin biosynthesis	17	415	2780	6	8
2	B <sub>5</sub>	CoA	Coenzyme A biosynthesis	18	720	6701	5	10
3	B <sub>12</sub>	Cobalamine	Coenzyme B <sub>12</sub> biosynthesis	61	201	4585	16	22
4		PQQ	Coenzyme PQQ synthesis	6	65	345	0	6
5	B <sub>9</sub> /B <sub>11</sub>	Folates	Folate biosynthesis	28	593	8231	7	13
6		Heme	Heme and siroheme biosynthesis	15	596	6474	0	11
7		Lipoate	Lipoic acid metabolism	3	655	1639	0	3
8	B <sub>3</sub>	NAD(P)	NAD and NADP cofactor biosynthesis	32	747	7156	15	9
9	B <sub>6</sub>	PLP, PMP	Pyridoxin (vitamin B <sub>6</sub> ) biosynthesis	9	578	2818	0	5
10	B <sub>2</sub>	FMN, FAD	Riboflavin, FMN, and FAD metabolism	17	412	3492	7	8
11	B <sub>1</sub>	Thiamin-PP	Thiamin biosynthesis	32	310	2971	11	6
12	K	Ubiquinone	Ubiquinone biosynthesis	15	490	4583	6	8
				253		51 775	29%	43%

<sup>a</sup> The number of distinct functional roles (isofunctional protein families) in a subsystem.

<sup>b</sup> The number of genomes that contain an operational variant of a given subsystem.

<sup>c</sup> The total number of genes associated with a set of functional roles in a subsystem (from all genomes counted in the column 'genomes').

<sup>d</sup> Roles that are present in <10% genomes within a subsystem.

<sup>e</sup> Roles that are present in >50% genomes within a subsystem.

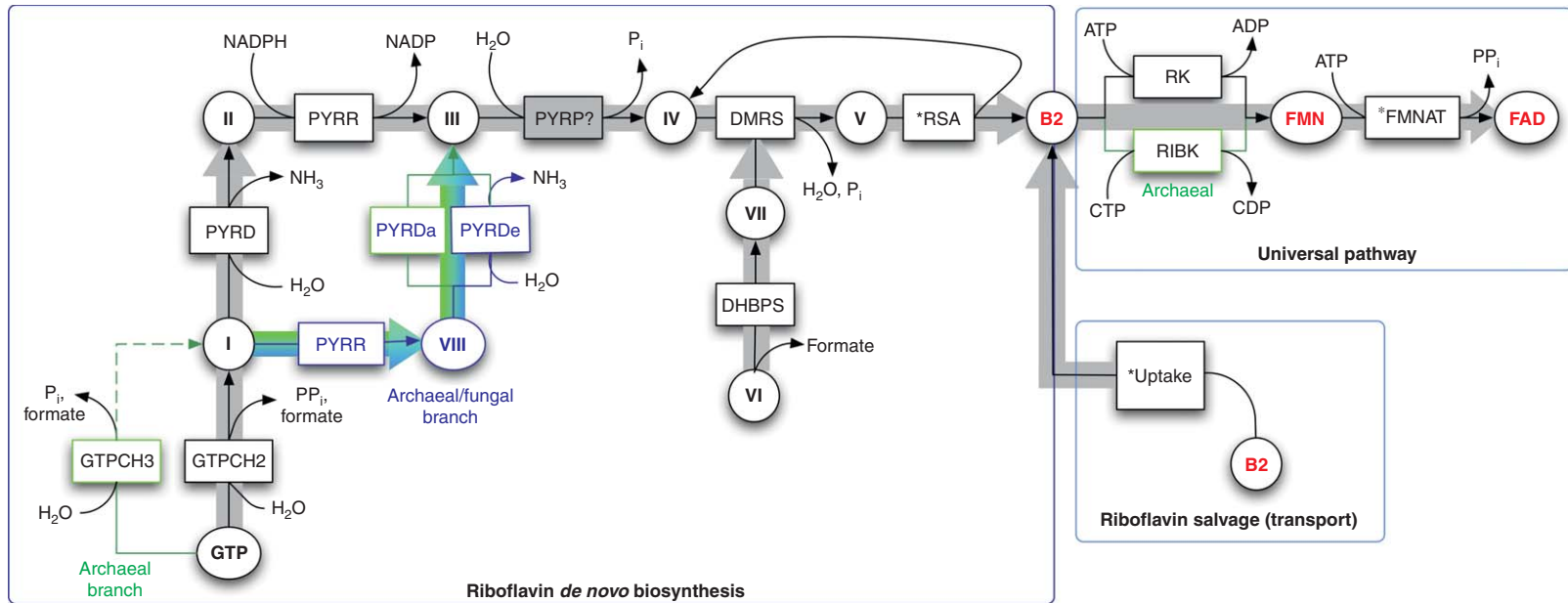
### 7.05.2 More Precisely, *What Is a Subsystem?*

Subsystem-based annotation seeks to organize the genomic data relating to a small, well-defined set of 'functional roles' that make up a pathway. For metabolic pathways these collections of functional roles include mostly enzymes, sometimes enhanced with transporters and transcriptional regulators. We refer to this set of functional roles as a 'subsystem'. Each functional role is typically associated with a set of homologous genes (members of a single protein family) that implement this role in specific organisms. We create a 'populated subsystem' as a spreadsheet in which the columns represent functional roles, and the rows represent specific genomes. Each cell in the spreadsheet contains the genes that encode proteins that implement the functional role in a specific genome.

A number of factors determines the scope, phylogenetic coverage, accuracy, and level of completion of each subsystem. Of course, the extent of the experimental evidence, the curator's depth of knowledge in the respective area of metabolism, and the stage of the curator's analysis (usually reflected in the notes attached to every subsystem) constitute the main factors. Overall, the small collection of 12 subsystems shown in **Table 1** includes >50 000 annotated genes spanning ~750 analyzed diverse genomes (among them ~90% are bacterial, with a relatively small fraction of archaeal and only about a dozen of representative eukaryotic genomes). Briefly looking at **Table 1**, one may notice substantial variations in the number of functional roles (from three in lipoate metabolism up to 61 in coenzyme B<sub>12</sub> biosynthesis, with a more typical size being ~15–25 roles), as well as in the number of genomes containing an operational variant of each pathway (from 65 to over 700). Functional roles included in each subsystem may be split into two main categories depending on their occurrence in the genomes. On average ~40% of the functional roles constitute the core of a subsystem, and they are rather universal (present in >1/2 of the analyzed genomes), whereas ~30% of the included roles correspond to less ubiquitous and species-specific aspects of the pathway (present in <1/10 of all genomes). A subsystem core often includes the most conserved and universal enzymes, whereas transcriptional regulators, uptake transporters, and rare alternative forms of enzymes frequently constitute a labile periphery of the subsystem.

To illustrate the concept of a subsystem, we will consider the subnetwork of biochemical transformations that convert GTP and ribulose-5-phosphate first into vitamin B<sub>2</sub>, and then into FMN and FAD cofactors. This subsystem was chosen due to a combination of reasons: (1) the high level of biochemical understanding (as reflected in Chapter 7.02); (2) the ubiquitous and essential nature of flavin cofactors in the three kingdoms of life; (3) a relatively simple topology, which includes the *de novo* biosynthesis of B<sub>2</sub> (replaced by salvage in some species) followed by its two-step conversion into FMN and FAD cofactors; and (4) broad conservation of most biochemical reactions and enzymes combined with some interesting variations between species that allow us to illustrate the application of comparative genomic techniques.

**Figure 1** provides a simplified subsystem diagram that schematically shows major intermediary metabolites (depicted by ovals with abbreviations or Roman numerals I through VII) and enzymes (shown as rectangles with abbreviations) known to catalyze the respective reactions (shown by arrows) in at least some of the characterized species. Similar diagrams are broadly used by some of the pathway-oriented databases (such as KEGG<sup>5</sup>) to capture all possible reactions and pathways within a subnetwork, whereas other resources (e.g., MetaCyc<sup>6</sup>) prefer to display organism-specific pathway diagrams. As in KEGG pathway maps, subsystem diagrams in The SEED database provide the ability to highlight the functional roles present in any selected organism. This allows the user to make a preliminary assessment of which of the possible fluxes (shown by thick gray arrows in **Figure 1**) or metabolic scenarios<sup>13</sup> are present in the organism of interest. This depiction of the genes from specific organisms can be used to clearly reveal incomplete functional variants of pathways<sup>14</sup> containing gaps (*missing genes*) and inconsistencies (*out-of-context genes*) that reflect incomplete knowledge or annotation errors. Although some of the revealed problems may be reconciled by similarity-based annotation techniques (e.g., by finding a gene candidate with a lower homology score or by the detailed analysis of gene grouping in a family of paralogues), others may not be effectively addressed without application of additional genome context analysis techniques. Application of these techniques, primarily clustering of functionally related genes on the prokaryotic chromosome,<sup>15</sup> analysis of protein fusion events,<sup>16</sup> co-occurrence profiles<sup>17</sup> and shared regulatory sites,<sup>18</sup> substantially improves the quality and consistency of genomic annotations. Such analysis can lead to accurate prediction of novel gene candidates and other conjectures about pathways that can be further tested by focused experiments (see Osterman and Overbeek<sup>19</sup>



Symbol	Compound name
I	2,5-Diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate
II	5-Amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate
III	5-Amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate
IV	5-Amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione
V	6,7-Dimethyl-8-ribityl-lumazine
VI	Ribulose 5-phosphate
VII	L-3,4-Dihydroxy-2-butanone 4-phosphate

**Figure 1** A subsystem diagram of the biosynthesis of riboflavin, FMN, and FAD. Enzymes are indicated by rectangles with abbreviations as in [Table 3](#). Pathway intermediates and products are shown in circles by standard abbreviations (GTP, FMN, FAD, and B<sub>2</sub> for riboflavin) or roman numerals enlisted in the inset. Major fluxes are outlined by thick arrows (gray, for bacterial and universal routes and colored, for the archaeal/fungal branch). Subsets of roles marked by "\*" combine alternative (nonorthologous) forms of genes (protein families) that play equivalent roles in the subsystem.

for the overview). Some examples of such bioinformatics predictions related to the B<sub>2</sub>/FMN/FAD metabolism will be illustrated below, while additional examples in the metabolism of other cofactors are listed in [Table 2](#).

Functional roles that constitute the analyzed subsystem, as well as the respective reactions (as captured in the KEGG database), are provided in [Table 3](#). These represent key elements in the encoding of any subsystem. In addition to the list of the core functional roles, this table includes all known cases of the so-called nonorthologous gene displacement, which amounts to the existence of two or more alternative forms of the same role encoded by distinct protein families. For example, two alternative forms of pyrimidine deaminase (EC 3.5.4.26) are known, the bacterial enzyme PYRD (represented by the N-terminal domain of the *ribD* gene product), and the eukaryotic enzyme PYRDe (represented by the gene *YOL066C* of *Saccharomyces cerevisiae*). As none of these two enzymes is present in Archaea, the third, typically archaeal, form (PYRDa, represented by the gene *MJ0699* of *Metbanocaldococcus jannaschii*) was predicted by bioinformatics techniques (as described below). When this manuscript was in preparation for publication, the authors became aware of the groundbreaking discovery by Zahra Mashhadi and R. White (Virginia Tech) of the last missing gene in the archaeal FAD biosynthesis. The gene MJ1179 from *M. jannaschii* was experimentally proven to encode an archaeal-specific version of FADS (Zahra Mashhadi, personal communication). Remarkably, this gene shows a substantial level of chromosomal clustering with other genes of the pathway. For example, in *Sulfolobus solfataricus* and several other archaeal genomes it is located next to Riboflavin kinase, whereas in *Archaeoglobus fulgidus* and a few other genomes, next to Riboflavin synthase.

[Table 4](#) shows the condensed form of the populated subsystem spreadsheet showing only a set of representative genomes. Normally, such a spreadsheet would contain gene IDs in the cells, but for the purpose of this discussion, we have simplified it by reflecting only the presence or absence of genes implementing the corresponding functional roles in each genome. Shading is used to convey clustering on the chromosome, which provides important supporting evidence (in addition to sequence similarity) for gene assignment in pathways.

Thus, in *Bacillus subtilis*, where this pathway was studied in most detail,<sup>56</sup> all of the six functional roles involved in B<sub>2</sub> *de novo* biosynthesis are encoded within a single chromosomal cluster (see [Figure 2](#)). An additional level of functional coupling of the involved genes is manifested at the level of domain fusion. Three bi-functional multidomain enzymes are present in the subsystem variant implemented in *B. subtilis*. Two of them are involved in *de novo* B<sub>2</sub> biosynthesis: the fusion protein encoded by the gene *ribBA* containing two enzymatic domains, 3,4-dihydroxy-2-butanone 4-phosphate synthase (DHBPS) and GTP cyclohydrolase II (GTPCH2), and the bi-functional enzyme pyrimidine deaminase/pyrimidine reductase (PYRD/PYRR) encoded by the gene *ribD*. Another bifunctional enzyme with two enzymatic domains, riboflavin kinase/FMN adenylyltransferase (RK/FMNAT), is involved in the consecutive conversion of B<sub>2</sub> into FMN and FAD. The *ribC* gene encoding the fusion RK/FMNAT enzyme is located on the chromosome remotely from the *ribDEBAH* operon (see [Table 4](#) and [Figure 2](#)).

The membrane transporter RibU is a nonenzymatic component of the B<sub>2</sub>/FMN/FAD subsystem, which functions in the uptake of vitamin B<sub>2</sub>. Although it is located separately on the *B. subtilis* chromosome, its functional coupling with the pathway is implemented through shared regulatory sites in the upstream regions of the gene *ribU* and the *ribDEBAH* operon. These *cis*-regulatory RNA elements (the so-called *RFN*-elements) were first discovered by comparative genomics<sup>20,22</sup> and then attributed to a class of metabolite-sensing riboswitches that control gene expression through formation of alternating leader mRNA structures depending on the presence of particular metabolites.<sup>57</sup> For example, the riboswitch associated with *rib*-genes is controlled by FMN, thus providing a negative feedback loop for the control of cofactor biosynthesis and B<sub>2</sub> uptake.<sup>26</sup> Two other riboswitches involved in the control of vitamin/cofactor biosynthesis, the thiamine pyrophosphate riboswitch (*THI*-element) and the adenosylcobalamin riboswitch (*B<sub>12</sub>*-element), were discovered by comparative genomic analysis and experimentally characterized.<sup>58</sup> Identification of these well-conserved regulatory elements upstream of various genes provides a strong indication of their involvement with respective pathways, thus contributing to functional gene annotation. In the case of the analyzed subsystem, the identification of the FMN riboswitch upstream of the *ribU* gene (previously *ypaA* in *B. subtilis*) led to the prediction of its role in B<sub>2</sub> uptake,<sup>20</sup> which was later experimentally confirmed.<sup>24</sup> A recent study revealed a novel mechanism of this uptake that includes association of the B<sub>2</sub>-specific permease component RibU with the energizing cassette composed of the two cytoplasmic ATPse subunits (genes *ybaD* and *ybaE* in *B. subtilis*) and the anchor membrane subunit (gene *ybaF*).<sup>21</sup> The energizing cassettes of this type (termed ECF for energy-coupling factor) are

**Table 2** Impact of comparative genomics in expanding the knowledge of vitamin/cofactor biosynthesis

Prediction based on comparative genomics	Lineages <sup>a</sup>	Reference(s)	Experimental verification	Reference(s)
<i>Riboflavin, FMN, and FAD metabolism</i>				
RibU/EcfAA'T – riboflavin transporter	Firmicutes	20–22	Genetics and uptake assays in <i>Bacillus subtilis</i> , <i>Lactococcus lactis</i>	21, 23, 24
PnuX – riboflavin transporter	Actinobacteria	22	Cloned from <i>Corynebacterium glutamicum</i> , uptake assay	24
RFN riboswitch – regulation	[B]	22	<i>In vitro</i> assays	25, 26
MJ0699 – NOD of pyrimidine deaminase (EC 3.5.4.26)	[A]	14		
<i>NAD and NADP cofactor biosynthesis</i>				
NadD – nicotinate mononucleotide adenylyltransferase (EC 2.7.7.18)	[B]	27	Cloned from <i>Escherichia coli</i> , enzyme assay	28
NadR, C-domain – Ribosylnicotinamide kinase (EC 2.7.1.22)	$\gamma$ -Proteobacteria	29	Cloned from <i>Haemophilus influenzae</i> , enzyme assay	29
KynB – bacterial form of Kynurenine formamidase (EC 3.5.1.9)	[B]	30	Cloned from <i>Pseudomonas aeruginosa</i> , enzyme assay	30
Kynurenine pathway of quinolinate synthesis in bacteria	[B]	31	Cloned from <i>Cytophaga hutchinsonii</i> , assays, complementation	31
NadE' – NMN synthetase, alternative route of NAD synthesis	<i>Francisella</i>	32	Cloned from <i>Francisella tularensis</i> , enzyme assay	33
NiaP – niacin transporter	Firmicutes	33	Mutant phenotype in <i>B. subtilis</i>	33
NiaX/EcfAA'T – niacin transporter	Firmicutes	21, 33		
NiaR – transcriptional regulator, regulon	Firmicutes, Thermotogales	33, 34	Cloned from <i>B. subtilis</i> and <i>Thermotoga maritima</i> assay	33, 34
NrtR – transcriptional regulator, regulon	[B]	35	Cloned from <i>Shewanella</i> and <i>Synechocystis</i> , assay	35
<i>Coenzyme A biosynthesis</i>				
NOD of pantetheine-phosphate adenylyltransferase (EC 2.7.7.3)	[E] and [A]	36	Cloned from <i>Homo sapiens</i> , assay, pathway reconstitution	36
NOD of bacterial PanK (EC 2.7.1.33) by the eukaryotic form	Staphylococci	37	Cloning, expression, enzyme assay,	38
MJ0969 – NOD of PanK (EC 2.7.1.33)	[A]	11		
PanT/EcfAA'T – pantothenate transporter	Firmicutes	21		
<i>Thiamin biosynthesis</i>				
THI riboswitch – regulation	[B], [A]	39	<i>in vitro</i> assays	40
ThiT/EcfAA'T – thiamine uptake	Firmicutes	21	Cloned from <i>Lactobacillus casei</i> , binding and uptake assays	41
TenA – enzyme in thiamine metabolism	[B], [A]	39	Cloned from <i>B. subtilis</i> , enzyme assay	42



Thi4 – enzyme in thiazole biosynthesis	Thermotogales, [A]	39	Cloned from <i>T. maritima</i> , complementation	43
ThiN – NOD of ThiE	Thermotogales, [A]	39	Cloned from <i>T. maritima</i> , enzyme assay	43
ThiXYZ – transporter of HMP precursor	[B]	39	Cloned from <i>B. halodurans</i> , binding assay	44
YkoEDC – transporter of HMP precursor	[B]	39	Mutant phenotype in <i>B. subtilis</i>	45
CytX – transporter of HMP precursor	[B], [A]	39		
PnuT – thiamine transporter	Bacteroidetes	39		
OMR – outer membrane thiamine transporter	Proteobacteria	39		
<i>Biotin biosynthesis</i>				
BioY/BioMN – biotin transporter	[B], [A]	46	Cloned from <i>Rhodobacter capsulatus</i> , uptake assay	47
BioR – transcription factor, regulon	$\alpha$ -Proteobacteria	48		
BioQ – transcription factor, regulon	Actinobacteria	49	Cloned from <i>C. glutamicum</i> , assay <sup>b</sup>	
BioG – NOD of BioH	Bacteroidetes	46		
BioK – NOD of BioH	Cyanobacteria	46		
<i>Coenzyme B<sub>12</sub> biosynthesis</i>				
B <sub>12</sub> riboswitch – regulation	[B], [A]	50, 51	<i>In vitro</i> assays	52
CobZ – NOD of CobC	[A]	50	Cloned from <i>M. mazei</i> , complementation	53
CobY – NOD of CobU	[A]	50	Mutant phenotype in <i>Halobacterium</i> sp.	54
PduX – threonine kinase	[B]	50	Cloned from <i>Salmonella</i> , enzyme assay	55
CbtAB, CbtC, HupE, CbtD, CbtE, CbtF, CbtG – Co transporters	[B]	50		
ChIID – components of Co chelatase CobN	[B]	50		
CblT, CblS – components of alpha-ribazole-5P synthesis	Firmicutes	50		
<i>Lipoic acid metabolism</i>				
LipT – lipoate transporter	Phytoplasma	21		
<i>Folate biosynthesis</i>				
FolT/EcfAA'T – folate transporter	Firmicutes	21	Cloned from <i>L. casei</i> , binding and uptake assays	41

<sup>a</sup> Groups of species that contain predicted genes are shown as [B], [A], or [E] (for various representatives of Bacteria, Archaea, or Eukaryota) or by more specific phylogenetic groups.

<sup>b</sup> A. Tauch, personal communication

This table contains a list of selected functional predictions published by one (or more) coauthors of this chapter (references in column 3). It includes only those cases where the straightforward homology-based analysis would not be sufficient to assign a gene a specific function. Functional predictions were made by combining metabolic reconstruction with genome context analysis techniques such as clustering on the chromosome, protein fusion, and the analysis of shared regulatory sites. For many examples, supporting experimental evidence data have been published by the same or other authors (references in column 5).

NOD, nonorthologous gene displacement.

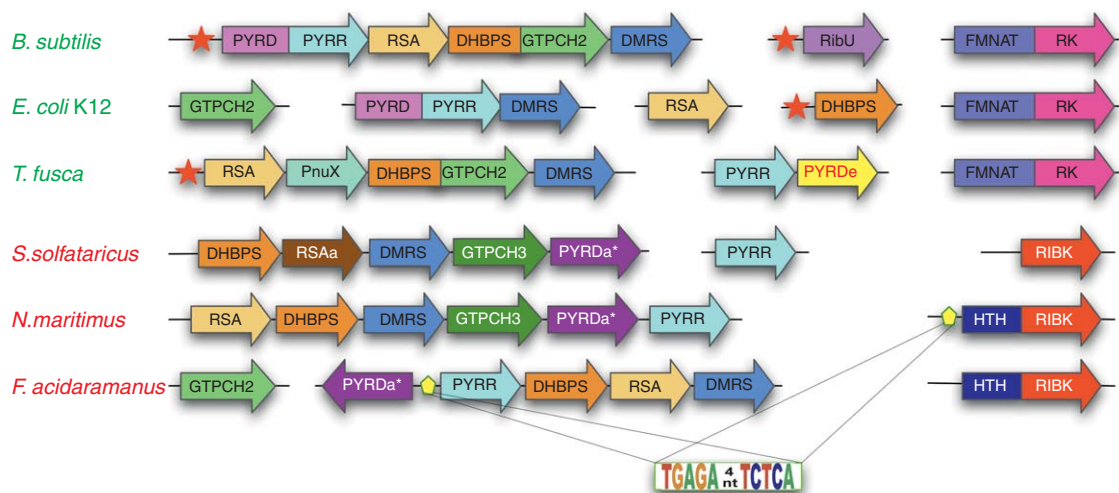
**Table 3** Functional roles included in riboflavin, FMN, and FAD biosynthesis subsystems

Abbreviations	Subset <sup>a</sup>	Functional role	Reaction <sup>b</sup>	Gene name <sup>c</sup>
GTPCH2	*GPTCH	GTP cyclohydrolase II (EC 3.5.4.25)	$GTP + 3H_2O = I + \text{formate} + PP_i$	<i>ribA</i>
GPTCH3		GTP cyclohydrolase III (EC 3.5.4.29)	$GTP + 3H_2O = I + \text{formate} + 2P_i$	
PYRD	*PYRD	Pyrimidine deaminase (EC 3.5.4.26)	$I + H_2O = II + NH_3$	<i>ribD</i>
PYRDe		Pyrimidine deaminase eukaryotic (EC 3.5.4.26)	$II + H_2O = III + NH_3$	
PYRDa		Pyrimidine deaminase archaeal predicted (EC 3.5.4.26)	$II + H_2O = III + NH_3$	
PYRR		Pyrimidine reductase (EC 1.1.1.193)	$II + NADP^+ = III + NADPH + H^+$ $(I + NADP^+ = II + NADPH + H^+)$	<i>ribD</i>
PYRP		5-Amino-6-(5'-phosphoribitylamino)uracil phosphatase	$III + H_2O = IV + P_i$	?
DHBPS		3,4-Dihydroxy-2-butanone 4-phosphate synthase	$VI = VII + \text{formate}$	<i>ribB</i>
DMRS		6,7-Dimethyl-8-ribityllumazine synthase (EC 2.5.1.9)	$IV + VII = V + H_2O + P_i$	<i>ribH</i>
RSA	*RSA	Riboflavin synthase (EC 2.5.1.9)	$V = B_2 + IV$	<i>ribE</i>
RSAa		Riboflavin synthase archaeal (EC 2.5.1.9)		
RibU	*Uptake	RibU component of riboflavin ECF transporter		<i>ribU (ypaA)</i>
PnuX		Riboflavin transporter PnuX		
RK	*RK	Riboflavin kinase (EC 2.7.1.26)	$B_2 + ATP = FMN + ADP$	<i>ribC</i>
RIBK		CTP-dependent archaeal riboflavin kinase		
FMNAT	*FMNAT	FMN adenyltransferase (EC 2.7.7.2)	$FMN + ATP = FAD + PP_i$	<i>ribC</i>
FMNAT2		FMN adenyltransferase, type 2 eukaryotic (EC 2.7.7.2)		

<sup>a</sup>Subsets of roles marked by "\*" combine alternative (nonorthologous) forms of genes (protein families) that play equivalent roles in the subsystem.

<sup>b</sup>Symbols (and abbreviations) of compounds are as in [Figure 1](#).

<sup>c</sup>Gene names are provided as in *Bacillus subtilis* (note that in some other species such as *Escherichia coli* some of these genes were historically named differently).



**Figure 2** Examples of clustering on the chromosome and coregulation of genes involved in riboflavin, FMN, and FAD biosynthesis in prokaryotes. Thick arrows of matching color correspond to orthologous genes (or functional domains) encoding functional roles that are marked by the same abbreviations as in **Table 3** and **Figure 1**. Chromosomal loci containing respective genes are shown for three representative bacterial genomes (names in green) and three archaeal genomes (names in red) (see **Table 3** for more information). Typically, bacterial and universal forms of genes are shown by light colors and abbreviations in black, whereas typically archaeal form are marked by dark background and white abbreviations help to reveal cases of nonorthologous replacement. The only case of nonorthologous replacement in bacteria by the eukaryotic form of PYRD is outlined by the red font. The alternative form of PYRD in Archaea predicted based on chromosomal clustering and distant homology with some characterized families of hydrolases is marked by asterisk (PYRDa\*). The presence of the HTH domain fused with RIBK enzymatic domains in some archaeal genomes suggests that this domain may play a role in the transcriptional regulation. Predicted position of DNA binding sites for this regulator are shown by yellow pentagons and the predicted sequence is illustrated by the Web-logo. Instances of RFN (FMN-responding riboswitch regulating the expression of respective genes and operons) in bacterial genomes are shown by a red star.

conserved in most Firmicutes, where they are shared by two or more solute-binding permease components. For example, in *B. subtilis* the EcfAA'T cassette is probably shared between the binding components for riboflavin (RibU), biotin (BioY), and thiamine (ThiT). Genes encoding the shared EcfAA'T cassettes always form operons but they are neither clustered on the chromosome, nor coregulated with solute-specific components (such as *ribU*, *bioY*, or *thiT*). Therefore, their inclusion in the subsystems describing respective metabolic pathways is optional (e.g., these genes have not been included in the B<sub>2</sub>/FMN/FAD subsystem and they are not shown in **Tables 3** and **4** or **Figure 2**).

In contrast to solute-binding proteins of ABC transporters, the respective components of the ECF modular transporters are integral membrane proteins.<sup>21</sup> Remarkably, most of these structurally divergent components appear to be involved with the uptake of vitamins. In addition to the three vitamins mentioned above, members of the ECF class identified in other genera (e.g., *Streptococci*, *Lactobacilli*, and *Mycoplasma*) include previously uncharacterized transporters of folate (FolT), niacin (NiaX), pantothenate (PanT), and lipoate (LipT). For all these solute-binding transporters, the specificity was inferred using genome context analysis (e.g., chromosomal clustering with genes from the respective metabolic pathways) and some of them have been experimentally verified (see **Table 2** for specific examples).

It is important to emphasize that operational variants in other metabolic pathways (including biosynthesis of cofactors listed in **Table 1**) are, in many cases, also associated with missing gene problems. Those may emerge due to errors in annotation, incomplete knowledge or, more often, nonorthologous gene displacements within subsets of species.<sup>59</sup> The subsystems approach offers a framework to characterize and project operational pathway variants (or metabolic scenarios) even if they contain gaps and to clarify open questions including identification of gene candidates for missing genes (see **Table 2** and further discussion for some examples). In the following section, we continue to use the example of the B<sub>2</sub>/FMN/FAD subsystem to illustrate how subsystems are built and expanded to capture the entire set of relevant functional roles and operational variants over hundreds of sequenced genomes.

### 7.05.3 How Are Subsystems Built?

The first step in constructing a subsystem is to assemble the relevant literature. In the case of our example, this has been done in Chapter 7.02 of this volume. On the basis of this primary literature and, in some cases, with the help of some web-based pathway-oriented resources (such as KEGG and MetaCyc), an initial description of the subsystem is composed. This description amounts to a list of relevant functional roles, including alternative forms when they are known to exist (as illustrated by **Table 3**). Additional associated information (captured in The SEED subsystems) includes links to encoded reactions, GO terms, PubMed citations and curator's notes. A subsystem diagram conveying the biochemical transformations and other interactions of functional roles within the subsystem (e.g., complexes, alternative roles, etc.) is constructed and displayed (as illustrated in **Figure 1**). At this preparatory stage, the information about the most common variants is gathered and captured in the form of annotated genes (usually from organisms in which the steps were actually determined and verified).

The subsystem can then be populated by considering the details in a few key genomes. Upon inclusion of the analyzed genome(s) as rows to the spreadsheet, the cells corresponding to the functional roles (columns in the spreadsheet) are automatically populated with the IDs of annotated genes (as shown in simplified form in **Table 4**). A *subsystems editor* was implemented to facilitate this process and to support integration of the main *spreadsheet* with the subsystem *description* and *diagram*. In addition, it also supports visual shading of the spreadsheet to indicate genes that occur in close proximity on the chromosome. **Table 4** shows instances of fusion proteins by matching numbers. For example, the symbol <1> is printed in both columns, GTPCH2 and DHBPS, for those genomes that contain genes encoding a bifunctional fusion protein (such as *ribBA* in *B. subtilis*).

A carefully constructed initial subsystem will designate the key variations captured in a subsystem by curator-assigned variant codes (examples are shown in **Table 4**). Subsystem variants in general correspond to identical or similar gene patterns reflecting the presence and absence of respective functional roles in a genome. Subsystem variants may be defined and classified using different approaches depending on the subsystem complexity, the extent of variation between species and, most importantly, the style of reasoning and research scope of a curator. For example, a semi-automated approach to variant classification based on the application of graph theory was described.<sup>14</sup> Here we illustrate the approach that tends to classify the variants of the B<sub>2</sub>/FMN/FAD subsystem using two hierarchical levels of granularity. The first (coarse) classification level distinguishes only two major functional variants. Variant 'A' (the first capitalized symbol in the variant codes shown in the respective column of the **Table 4**) refers to the organisms that have a version of a complete *de novo* pathway of B<sub>2</sub> synthesis. This variant is found in most bacteria (570 of 680 analyzed genomes), Archaea (43 of 50 genomes), fungi (*S. cerevisiae* and five other analyzed fungal genomes), and plants (represented by *Arabidopsis thaliana*). The second variant 'B' includes all species that lack *de novo* B<sub>2</sub> biosynthesis but have a downstream pathway converting the salvaged vitamin into FMN and FAD cofactors. This variant is present in 79 bacterial species (many of them are Gram-positive pathogens), five Archaea and all nine of the analyzed metazoan genomes. Finally, 26 bacterial genomes (mostly intracellular bacterial pathogens such as *Rickettsia* spp.) and one archaeal genome (*Nanoarchaeum equitans*) have lost all components of the subsystem. Most probably, these species developed the ability to salvage the intact flavin cofactors from their eukaryotic hosts. Examples of all functional variants in all three kingdoms are provided in **Table 4**.

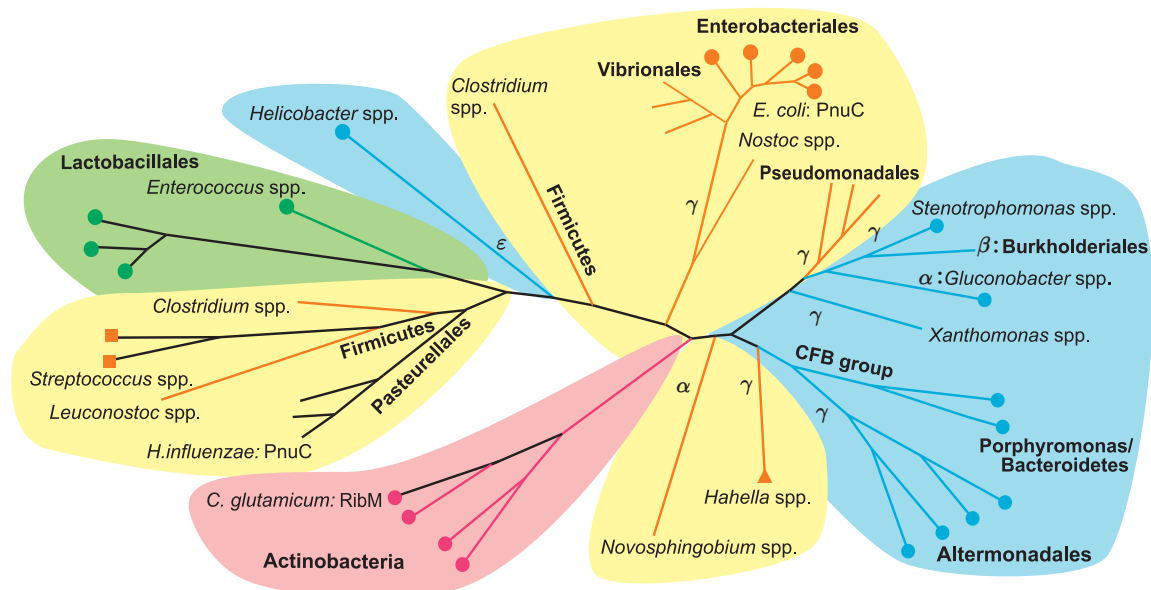
These examples also provide an illustration of a second level of classification, which is more granular and can potentially reflect any number of variations within otherwise similar pathways. For example, the variant marked 'A.b2' in **Table 4** differs from the basic 'A.b1' variant by the absence of one functional role, a RibU transporter. The absence of this transporter may have rather dramatic physiological consequences. For example, *E. coli* (variant A.b2) is unable to salvage vitamin B<sub>2</sub> leading to the observed essentiality of all B<sub>2</sub> biosynthetic genes even when B<sub>2</sub> is present in the medium.<sup>37</sup> This is in contrast with *B. subtilis*, where any gene of the *rib* operon can be inactivated producing viable B<sub>2</sub> auxotrophs. It is important to mention that the absence of a *ribU* homologue in a genome does not necessarily mean that an organism cannot uptake B<sub>2</sub>, as the same role might be potentially implemented by another nonorthologous gene. Nonorthologous replacements are particularly common among various classes of transporters.<sup>60</sup> At least one case of such replacement in B<sub>2</sub> uptake was predicted for *Thermobifida fusca* (see **Table 4** and **Figure 2**) and several other Actinobacteria by bioinformatics techniques.<sup>22</sup> This prediction was based on genome context analysis that revealed the presence of a putative transporter gene colocalized and/or coregulated (by FMN riboswitch) with the riboflavin biosynthesis genes in these species.

**Table 4** Variations in riboflavin, FMN, FAD biosynthesis subsystems (examples)

Genome	Variant	From GTP to riboflavin							From riboflavin to FMN and FAD									
		*GTPCH		*PYRD			PYRR	DHBPS	DMRS	*RSA		*UPTAKE		*RK	*FMNAT			
		B	A	B	A <sup>a</sup>	E	U	U	U	B	A	RibU	PnuX	B/E	A	B	E	A
<i>Bacteria</i>																		
<i>Bacillus subtilis</i>	A.b1	<1>		<2>			<2>	<1>	+	+	+			<3>		<3>		
<i>Synechocystis</i> sp.	A.b2	<1>		<2>			<2>	<1>	+	+	+			<3>		<3>		
<i>Escherichia coli</i> K12	A.b2	+		<2>			<2>	+	+	+				<3>		<3>		
<i>Buchnera aphidicola</i>	A.b2	+		+			+	+	+	+				<3>		<3>		
<i>Thermobifida fusca</i>	A.b3	<1>				+	+	<1>	+	+		+		<3>		<3>		
<i>Streptococcus pyogenes</i>	B.b1												+		<3>		<3>	
<i>Rickettsia prowazekii</i>	-1																	
<i>Archaea</i>																		
<i>Methanocaldococcus jannaschii</i>	A.a1		+		+		+	+	+		+					+		?
<i>Sulfolobus solfataricus</i>	A.a1		+		+		+	+	+		+					+		?
<i>Ferroplasma acidarmanus</i>	A.a2	+			+		+	+	+		+					+		?
<i>Nitrosopumilus maritimus</i>	A.a3		+		+		+	+	+		+					+		?
<i>Pyrococcus furiosus</i>	A.a4	+		+			+	+	+		+					+		?
<i>Pyrococcus horikoshii</i>	B.a1															+		?
<i>Nanoarchaeum equitans</i>	-1																	
<i>Eukarya</i>																		
<i>Saccharomyces cerevisiae</i>	A.e1	+				+	+	+	+		+				+			+
<i>Arabidopsis thaliana</i>	A.e2	<1>		<2>			<2>	<1>	+		+				+			+
<i>Homo sapiens</i>	B.e1														+			+

<sup>a</sup> A predicted archaeal form of PYRD, which has not been experimentally verified.

Upper row – abbreviations of functional roles in the subsystem (as in [Table 3](#)). The subsets roles that are associated with essentially the same biochemical transformations but exist in two or more alternative forms (encoded by nonorthologous genes) are marked by asterisks (\*). Abbreviations in the next row show the predominant phylogenetic occurrence of the respective forms: 'B' – mostly bacterial form; 'A' – archaeal; 'E' – eukaryotic; 'U' – universal. In the cells, '+' – indicate the presence of the respective gene (blank – absence). Fusion proteins composed of two functional domains are shown by matching numbers, for example, <1> corresponds to a fusion of GTPCH and DHBPS enzymatic domains. '?' notation points to a case of a missing gene in the subsystem (in Archaea the gene encoding FMNAT is still unknown, although the existence of this enzyme is anticipated). Matching background shading reflects proximity of the respective genes on the chromosome.



**Figure 3** Phylogenetic tree and genome context analysis of the PnuC family of transporters. The PnuC protein family includes the known transporters of nicotinamide ribose (PnuC) and the predicted transporters of riboflavin (PnuX), thiamine, (PnuT) and deoxynucleosides (PnuN). The PnuC transporters from *Escherichia coli* and *Haemophilus influenzae*, and PnuX from *Corynebacterium glutamicum* (renamed RibM) have been experimentally characterized. Chromosomal clustering with genes implicated in the corresponding metabolic pathways is indicated by color of the respective branch on the tree: NAD salvage (orange), thiamine salvage (blue), and nucleoside salvage (green). Predicted regulatory elements found in upstream regions of candidate transporter genes are indicated by color symbols at the ends of the branches: NadR operator (orange circle), NiaR operator (orange square), NrtR operator (orange triangle), *THI* riboswitch (blue circle), FMN riboswitch (red circle), and NrdR operator (green circle). Taxonomic groups are indicated in bold.

The predicted novel B<sub>2</sub> transporters were termed PnuX to reflect upon their similarity with PnuC transporters that are involved in salvage of nicotinamide ribose, a precursor of NAD cofactor (see Chapter 7.08). The predicted B<sub>2</sub>-specificity of the representative PnuX transporter from *Corynebacterium glutamicum* (re-named RibM) was experimentally confirmed<sup>24</sup> (see Table 2). Remarkably, the PnuC family of membrane transporters contains several specific subgroups with different candidate specificities (see phylogenetic tree in Figure 3), including nicotinamide ribose (PnuC), riboflavin (PnuX), thiamine (PnuT), and deoxynucleosides (PnuN) that were predicted by simultaneous phylogenetic analysis, chromosomal clusters, and shared regulatory sites.<sup>60</sup> This combined bioinformatics approach is of particular importance for accurate annotation of members of large families that contain groups of paralogues with related but distinct functions, most notably transporters and transcriptional regulators. For example, nearly all members of the PnuC family annotate in most databases as nicotinamide ribose transporters (involved in NAD metabolism) because of an indiscriminate homology-based projection.

A version of the subsystem that contains PnuX transporter instead of RibU was captured as the A.b3 variant (Table 4). There are reasons to expect the existence of at least one more (yet unknown) alternative form of the B<sub>2</sub> transporter in a group of bacteria that implement subsystem variant B, which is equivalent to B<sub>2</sub> auxotrophy (20 species, mostly pathogenic *Mycoplasma* and *Spirochetes*). However, the majority of bacterial species (57 in our set) that belong to variant B contain a copy of *ribU* gene in the genome as illustrated in Table 4 by the example of *Streptococcus pyogenes* (variant B.b1).

Additional variations illustrated in Table 4 and Figure 2 that are not reflected in variant codes include (1) the degree of clustering on the chromosome (e.g., all of the *rib* genes are scattered on the chromosome of *Synechocystis* sp.); (2) presence or absence of fusion events (as in the extreme case of *Buchnera aphidicola*, in which both common fusion proteins DHBPS/GTPCH2 and PYRD/PYRS are replaced by the four respective monofunctional components), and (3) nonorthologous replacement of enzymes (as in the case of *T. fusca*, where the typically bacterial PYRD enzyme is replaced by the functionally equivalent PYRDe enzyme,



characteristic of the fungi (see below)). The tentative functional assignment of this gene, a probable subject of horizontal gene transfer (of all analyzed bacterial genomes it occurs only in seven species of Actinobacteria), is supported by its chromosomal clustering with PYRR gene (see the shaded regions in **Table 4** and an illustration of gene clusters in **Figure 2**).

Once the kernel of the subsystem, displaying instances of the common variants, has been constructed, new genomes resembling the initial representatives can be rapidly added. Occasionally, new problems are detected and substantial manual effort is required to determine the most probable genes needed to populate the growing spreadsheet. Eventually, a state is reached in which the majority of newly sequenced genomes can be added with minimal effort. Tools to increase the productivity of the analyst by semi-automating the process have been developed and are in constant use.<sup>61</sup>

Most of the difficulties are associated with projecting genes and pathways over a substantial phylogenetic distance, especially when projecting from rather deeply studied Bacteria to Archaea and Eukaryota. In addition to generally weaker homology scores (creating particularly difficult problems in the case of protein families with many paralogues implementing distinct functions) and more frequent cases of nonorthologous gene replacements, challenges associated with such long-range projection include alternative biochemical transformations. Even in a relatively conserved B<sub>2</sub>/FMN/FAD biosynthesis subsystem, the key differences between the typically bacterial and archaeal versions include

1. nonorthologous replacements of PYRD, RSA, and RK enzymes;
2. an alternative GTPCH3 enzyme in Archaea appears to generate a different intermediate product that undergoes further dephosphorylation to (I) by an unknown phosphatase;<sup>62</sup>
3. the order of the deamination and reduction steps in Archaea (and fungi) is opposite to the order of these two reactions in the bacterial version of the pathway (**Figure 1**);
4. the enzyme for the last step of the pathway, conversion of FMN into FAD, in Archaea is unknown, as archaeal genomes do not contain homologues of either bacterial or eukaryotic FMNAT enzymes.

However, despite these challenges and remaining knowledge gaps, the reconstruction and analysis of subsystem variants in archaeal species can be accomplished using the same general approach as described above. As illustrated in **Table 4**, the coarse-level classification of operational subsystem variants in Archaea includes the same two major lifestyles, the complete *de novo* pathway (A) and B<sub>2</sub>-auxotrophy (B). A more granular classification distinguishes several cases of nonorthologous gene replacements of typically archaeal enzymes by their bacterial-like counterparts. The examples shown in **Table 4** include the replacement of GTPCH3 by GTPCH2 (as in *Ferroplasma acidarmanus*, variant A.a2) or RSAa by RSA (as in *Nitrosopumilus maritimus*, variant A.a2). An extreme case of the horizontal transfer of the entire *rib* operon is illustrated by the example of *Pyrococcus furiosus* (variant A.a4). Remarkably, the closely related *Pyrococcus horikoshii* lacks the entire B<sub>2</sub> *de novo* synthesis (variant B.a1), suggesting that the hypothesized transfer of the bacterial *rib* operon to *P. furiosus* occurred after speciation within the *Pyrococcus* branch. Interestingly, the typically archaeal version of the subsystem, in contrast with its bacterial equivalent, does not contain a single fusion protein.

On the other hand, the existence of conserved chromosomal clusters (operons) of functionally coupled genes in archaeal genomes plays an important role in achieving high accuracy in gene assignment, pathway reconstruction, and identification of candidates for missing genes. Thus, a candidate for a gene encoding PYRDa, which is considered missing, was proposed based on its frequent chromosomal clustering with other genes in the pathway (see **Figure 2**). Orthologues of this gene candidate (*M70699* in *M. jannaschii*) are conserved in all archaeal genomes except those that either lack the B<sub>2</sub> *de novo* biosynthesis (variant B) or have the bacterial version of this pathway (variant A.a4). PYRDa proteins belong to a large superfamily of metallo-dependent hydrolases providing a general class function consistent with its predicted deaminase activity. This functional prediction remains to be tested.

Although the prediction of transcriptional regulators in Archaea is generally more difficult than in bacteria, at least one regulatory mechanism can be proposed for some of the archaeal genomes that contain a version of the RIBK protein fused with the N-terminal helix-turn-helix (HTH) domain. HTH domains play a role in DNA binding in many families of transcription factors. In some cases, these transcription factors include enzymatic domains that are involved in the same pathways as other members of respective regulons. Among the known examples in vitamin/cofactor metabolism are BirA (in biotin metabolism, see Chapter 7.08) and NadR (in NAD

metabolism, see Chapter 7.06). Based on genomic evidence, we hypothesized that HTH/RIBK proteins may play a role in the regulation of FMN/FAD synthesis through interaction with cognate binding sites in the upstream regions of some of the *rib* genes. Using a comparative genomic approach (for review see Rodionov<sup>49</sup>) we tentatively identified the consensus sequence (TGAGAnnnnTCTCA) and several instances of possible regulatory sites for this predicted regulator (see **Figure 2**). This conjecture is to be experimentally tested.

Projection of the metabolic subsystems toward eukaryotic genomes is generally the most challenging task. Among the key limiting factors are (1) substantial differences at all levels, from genes to pathways, (2) relatively limited biochemical knowledge (beyond one or two model species), (3) a very limited number of available genomes, (4) problems associated with gene identification and (5) inability to use clustering of genes on the chromosome as an evidence of their functional coupling. Nevertheless, the homology-based projection from yeast and prokaryotes often provides sufficient evidence for the reconstruction of the core metabolic pathways in various eukaryotes. The B<sub>2</sub> *de novo* biosynthesis that was characterized in detail in yeast (variant A.e1 in **Table 4**) combines some features of the bacterial (e.g., GTPCPH2) and archaeal (the alternative order PYRD and PYRR steps) versions. Nonorthologous gene replacements include the alternative forms of PYRD and FMNAT. Remarkably, plants (represented by *A. thaliana* in **Table 4**) combine a typically bacterial version of B<sub>2</sub> synthesis with the presence of the characteristic eukaryotic-like downstream enzymes (RK and FMNAT2). All animals appear to lack the *de novo* synthesis of B<sub>2</sub> being fully dependent on the salvage of dietary vitamin (variant B.e1), which is typical for most other vitamins listed in **Table 1** (e.g., pantothenate, thiamine, cobalamine, etc.). Among the few exceptions is the biosynthesis of NAD(P) cofactor, where both pathways, *de novo* biosynthesis and salvage of niacin are present in vertebrates (see Chapter 7.08).

The key principles of constructing and populating a subsystem illustrated above by the example of B<sub>2</sub>/FMN/FAD biosynthesis, are applicable to other vitamins and cofactors, and, more generally, to any other group of metabolic pathways. However, it would be just to admit that quite often, these subsystems are far less straightforward, less conserved, supported by much less functional coupling evidence, contain more variations even between phylogenetically close species, have more noncommitted enzymes (those that are involved in several pathways), include unresolved paralogues, and they reflect knowledge gaps (missing genes, unknown reactions) that create impediments for recognition and classification of operational variants. The initial construction and expansion of such subsystems requires a substantial effort, which may nevertheless be successful and rewarding, especially with the help of experts in the relevant areas of metabolism.

### 7.05.4 What Is Revealed by the Construction of Subsystems?

There are several reasons why the effort to construct subsystems is worth expending. First, it reveals *missing genes*, genes that are yet to be identified, but that we believe must exist in the genome to implement a specific functional role. Some examples in B<sub>2</sub>/FMN/FAD biosynthesis (the missing PYRD<sub>a</sub> and a novel PnuX transporter) are briefly described above. Additional examples in other vitamin/cofactor-related subsystems are listed in **Table 2**. In addition to clustering on the chromosome and shared regulatory sites, identification of fusion events (originally termed the Rosetta Stone approach<sup>16</sup>) in some cases provides evidence of functional coupling leading to prediction of previously unknown genes. This is of particular importance for functional prediction of novel gene families in eukaryotes that lack operon-like chromosomal clusters. For example, a typically eukaryotic form of phosphopantetheine adenylyltransferase (EC 2.7.7.3), one of the key enzymes in coenzyme A biosynthesis (see Chapter 7.20), was originally identified due to its fusion with the dephospho-CoA kinase (EC 2.7.1.24) domain in mammals.<sup>36</sup> Although some genes are missing in even the most advanced subsystems (e.g., PYRP in B<sub>2</sub> synthesis in all species or FMNAT in the archaeal version of FAD synthesis), it seems probable that they will eventually be identified, allowing us to fill-in the remaining gaps in the subsystem spreadsheet. The systematic identification of missing genes has been a central use of the existing subsystem collection.

The process of organizing genomic data in a form that supports comparative analysis has set the stage for systematic analysis of regulation. One important use of a subsystem like that shown in our example has been to support characterization of coregulated genes – through analysis of chromosomal gene clusters, fusion events, expression data, and through identification of transcription regulatory sites. The details and implications of this approach are described in a recent review,<sup>49</sup> and briefly illustrated by two examples in B<sub>2</sub>/FMN/FAD

subsystem. Among the recent applications of this approach to the analysis of other vitamin/cofactor-related pathways is the identification of two novel regulons associated with NAD metabolism in a large number of diverse bacteria.<sup>35,33</sup>

In a number of subsystems, authors have expended substantial effort to construct alignments and phylogenetic trees to support reconstruction of the evolutionary events that led to variations within a subsystem.<sup>63–65</sup> The emerging set of tools to support such analysis will ultimately dramatically reduce the effort required to develop a historical overview. As a component of the overall analysis of a subsystem, it becomes common to characterize the overall evolutionary history, relating the estimate with existing instances of clustering on the chromosome, fused proteins, and nonorthologous gene displacements. A few examples listed above illustrate that the information gathered in subsystems provides a good starting point for expending the detailed evolutionary analysis toward biosynthesis of vitamins and cofactors such as B<sub>2</sub>/FMN/FAD.

### 7.05.5 The Project to Annotate 1000 Genomes

During late 2003, the Fellowship for Interpretation of Genomes (FIG) initiated a project to provide annotations for the first 1000 genomes. A number of cooperating research groups formulated a strategy for systematically and consistently annotating large numbers of genomes based on the subsystems technology. Rather than annotating genomes one-at-a-time, annotators selected specific subsystems and annotated them through the entire set of existing genomes. The benefits in terms of productivity of annotators, derives from the fact that they become familiar with the central complexities related to each subsystem. The integration of large numbers of genomes into an existing subsystem leads to better accuracy and consistency when compared with the more common approach of asking annotators to systematically work through genomes one-at-a-time (an approach that assures that annotations in most cases are being done by individuals with no special knowledge of the functional components being annotated).

The systematic creation and extension of subsystems has now led to a collection including 700–800 distinct subsystems containing genes from hundreds of genomes. The collection is constantly being extended, refined, and updated by teams from a number of collaborating institutions.

It has become a framework for careful manual annotation, but it has also become the key resource in supporting high-throughput semi-automated annotations. The RAST server,<sup>61</sup> which offers subsystem-based automated annotations, is now being used to annotate 150–200 genomes per month, and it has become clear that it is capable of annotating thousands of genomes consistently, significantly reducing the manual effort required to extend existing subsystems.

### 7.05.6 Summary

The construction of subsystems now offers one approach to systematically organize the growing body of genomic data. It offers a technology for exposing issues that can only be resolved in the wet lab, and then supports propagation of experimental advances to the entire collection of relevant genomes. As we move into an era that will involve thousands of genomes, the integration of biochemical and genomic understanding offered by this approach will rapidly grow in significance. It is worth emphasizing that each of the nontrivial subsystems we have considered has exposed variations. These variations are explored systematically in a continuing effort to attain a consistent picture of the molecular mechanisms supported by experimental confirmation. It is our belief that the vast majority of the major variations within these subsystems be exposed by studying a growing set of diverse genomes. Within the next few years, it seems probable that exemplars of most significant variants will exist, and these will methodically be confirmed by experimental characterization. This implies that within just a few years it will be possible to construct a framework that accurately reflects the most significant variations, and that this resource will play a fundamental role in supporting analysis of the thousands of genomes that will rapidly become available during the next few decades.

## Acknowledgments

We thank our collaborators in ‘The Project to Annotate a Thousand Genomes’ for their help with the development of SEED tools, genomic annotations, and subsystem curation. The genomic analysis performed in this study was supported by the grant from the National Institute of Allergy and Infectious Diseases (NIAID) ‘Genomics of Coenzyme Metabolism in Bacterial Pathogens’ AI066244 to A. O. and ‘Bioinformatics Resource Centers for Biodefense and Emerging/Re-Emerging Infectious Diseases’ NIAID Contract Number: HHSN266200400042C (RO codirector).

### Abbreviations

<b>B<sub>2</sub></b>	riboflavin
<b>DHBPS</b>	3,4-dihydroxy-2-butanone 4-phosphate synthase
<b>DMRS</b>	6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9)
<b>FAD</b>	flavin adenine dinucleotide
<b>FMN</b>	flavin mononucleotide
<b>FMNAT</b>	FMN adenylyltransferase (EC 2.7.7.2)
<b>FMNAT2</b>	FMN adenylyltransferase, type 2 eukaryotic (EC 2.7.7.2)
<b>GPTCH3</b>	GTP cyclohydrolase III (EC 3.5.4.29)
<b>GTP</b>	guanosine triphosphate
<b>GTPCH2</b>	GTP cyclohydrolase II (EC 3.5.4.25)
<b>I</b>	2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 59-phosphate
<b>II</b>	5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione 59-phosphate
<b>II</b>	5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 59-phosphate
<b>IV</b>	5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione
<b>PYRD</b>	pyrimidine deaminase (EC 3.5.4.26)
<b>PYRDa</b>	pyrimidine deaminase archaeal predicted (EC 3.5.4.26)
<b>PYRDe</b>	pyrimidine deaminase eukaryotic (EC 3.5.4.26)
<b>PYRP</b>	5-amino-6-(59-phosphoribitylamino)uracil phosphatase
<b>PYRR</b>	pyrimidine reductase (EC 1.1.1.193)
<b>RIBK</b>	CTP-dependent archaeal riboflavin kinase
<b>RK</b>	riboflavin kinase (EC 2.7.1.26)
<b>RSA</b>	riboflavin synthase (EC 2.5.1.9)
<b>RSAA</b>	riboflavin synthase archaeal (EC 2.5.1.9)
<b>V</b>	6,7-dimethyl-8-ribityl-lumazine
<b>VI</b>	ribulose 5-phosphate
<b>VII</b>	L-3,4-dihydroxy-2-butanone 4-phosphate

## References

1. R. D. Fleischmann; M. D. Adams; O. White; R. A. Clayton; E. F. Kirkness; A. R. Kerlavage; C. J. Bult; J. F. Tomb; B. A. Dougherty; J. M. Merrick; K. Mckenney; G. Sutton; W. Fitzhugh; C. Fields; J. D. Gocayne; J. Scott; R. Shirley; L. I. Liu; A. Glodek; J. M. Kelley; J. F. Weidman; C. A. Phillips; T. Spriggs; E. Hedblom; M. D. Cotton; T. R. Utterback; M. C. Hanna; D. T. Nguyen; D. M. Saudek; R. C. Brandon; L. D. Fine; J. I. Fritchman; J. I. Fuhrmann; N. S. M. Geoghagen; C. I. Gnehm; L. A. McDonald; K. V. Small; C. M. Fraser; H. O. Smith; J. C. Venter, *Science* **1995**, 269, 496–512.
2. H. Bono; H. Ogata; S. Goto; M. Kanehisa, *Genome Res.* **1998**, 8, 203–210.
3. E. Selkov; N. Maltsev; G. J. Olsen; R. Overbeek; W. B. Whitman, *Gene* **1997**, 197, GC11–GC26.
4. D. H. Haft; J. D. Selengut; L. M. Brinkac; N. Zafar; O. White, *Bioinformatics* **2005**, 21, 293–306.
5. M. Kanehisa; S. Goto; S. Kawashima; A. Nakaya, *Nucleic Acids Res.* **2002**, 30, 42–46.
6. P. D. Karp; M. Riley; M. Saier; I. T. Paulsen; S. M. Paley; A. Pellegrini-Toole, *Nucleic Acids Res.* **2000**, 28, 56–59.
7. R. Overbeek; D. Devine; V. Vonstein, *Targets* **2003**, 2, 138–146.

8. R. Overbeek; N. Larsen; T. Walunas; M. D'Souza; G. Pusch; E. Selkov, Jr.; K. Liolios; V. Joukov; D. Kaznadzey; I. Anderson; A. Bhattacharyya; H. Burd; W. Gardner; P. Hanke; V. Kapatral; N. Mikhailova; O. Vasieva; A. Osterman; V. Vonstein; M. Fonstein; N. Ivanova; N. Kyrpides, *Nucleic Acids Res.* **2003**, *31*, 164–171.
9. R. Overbeek; N. Larsen; G. D. Pusch; M. D'Souza; E. Selkov, Jr.; N. Kyrpides; M. Fonstein; N. Maltsev; E. Selkov, *Nucleic Acids Res.* **2000**, *28*, 123–125.
10. E. Selkov; S. Basmanova; T. Gaasterland; I. Goryanin; Y. Gretchkin; N. Maltsev; V. Nenashev; R. Overbeek; E. Panyushkina; L. Pronevitch; E. Selkov, Jr.; I. Yunus, *Nucleic Acids Res.* **1996**, *24*, 26–28.
11. R. Overbeek; T. Begley; R. M. Butler; J. V. Choudhuri; H. Y. Chuang; M. Cohoon; V. de Crecy-Lagard; N. Diaz; T. Disz; R. Edwards; M. Fonstein; E. D. Frank; S. Gerdes; E. M. Glass; A. Goesmann; A. Hanson; D. Iwata-Reuyl; R. Jensen; N. Jamshidi; L. Krause; M. Kubal; N. Larsen; B. Linke; A. C. McHardy; F. Meyer; H. Neuweiger; G. Olsen; R. Olson; A. Osterman; V. Portnoy; G. D. Pusch; D. A. Rodionov; C. Ruckert; J. Steiner; R. Stevens; I. Thiele; O. Vassieva; Y. Ye; O. Zagnitko; V. Vonstein, *Nucleic Acids Res.* **2005**, *33*, 5691–5702.
12. L. K. McNeil; C. Reich; R. K. Aziz; D. Bartels; M. Cohoon; T. Disz; R. A. Edwards; S. Gerdes; K. Hwang; M. Kubal; G. R. Margaryan; F. Meyer; W. Mihalo; G. J. Olsen; R. Olson; A. Osterman; D. Paarmann; T. Paczian; B. Parrello; G. D. Pusch; D. A. Rodionov; X. Shi; O. Vassieva; V. Vonstein; O. Zagnitko; F. Xia; J. Zinner; R. Overbeek; R. Stevens, *Nucleic Acids Res.* **2007**, *35*, D347–D353.
13. M. DeJongh; K. Formsma; P. Boillot; J. Gould; M. Rycenga; A. Best, *BMC Bioinformatics* **2007**, *8*, 139.
14. Y. Ye; A. Osterman; R. Overbeek; A. Godzik, *Bioinformatics* **2005**, *21*, i1–i9.
15. R. Overbeek; M. Fonstein; M. D'Souza; G. D. Pusch; N. Maltsev, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2896–2901.
16. E. M. Marcotte; M. Pellegrini; H. L. Ng; D. W. Rice; T. O. Yeates; D. Eisenberg, *Science* **1999**, *285*, 751–753.
17. M. Pellegrini; E. M. Marcotte; M. J. Thompson; D. Eisenberg; T. O. Yeates, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4285–4288.
18. M. S. Gelfand; P. S. Novichkov; E. S. Novichkova; A. A. Mironov, *Brief Bioinform.* **2000**, *1*, 357–371.
19. A. Osterman; R. Overbeek, *Curr. Opin. Chem. Biol.* **2003**, *7*, 238–251.
20. M. S. Gelfand; A. A. Mironov; J. Jomantas; Y. I. Kozlov; D. A. Perumov, *Trends Genet.* **1999**, *15*, 439–442.
21. D. A. Rodionov; P. Hebbeln; A. Eudes; J. ter Beek; I. A. Rodionova; G. B. Erkens; D. J. Slotboom; M. S. Gelfand; A. L. Osterman; A. D. Hanson; T. Eitinger, *J. Bacteriol.* **2009**, *190*, 42–51.
22. A. G. Vitreschak; D. A. Rodionov; A. A. Mironov; M. S. Gelfand, *Nucleic Acids Res.* **2002**, *30*, 3141–3151.
23. C. M. Burgess; D. J. Slotboom; E. R. Geertsma; R. H. Duurkens; B. Poolman; D. van Sinderen, *J. Bacteriol.* **2006**, *188*, 2752–2760.
24. C. Vogl; S. Grill; O. Schilling; J. Stulke; M. Mack; J. Stolz, *J. Bacteriol.* **2007**, *189*, 7367–7375.
25. A. S. Mironov; I. Gusarov; R. Rafikov; L. E. Lopez; K. Shatalin; R. A. Krenova; D. A. Perumov; E. Nudler, *Cell* **2002**, *111*, 747–756.
26. W. C. Winkler; S. Cohen-Chalamish; R. R. Breaker, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15908–15913.
27. A. L. Osterman; T. P. Begley, *Prog. Drug Res.* **2007**, *64*, 131, 133–170 64.
28. R. A. Mehl; C. Kinsland; T. P. Begley, *J. Bacteriol.* **2000**, *182*, 4372–4374.
29. O. V. Kurnasov; B. M. Polanuyer; S. Ananta; R. Sloutsky; A. Tam; S. Y. Gerdes; A. L. Osterman, *J. Bacteriol.* **2002**, *184*, 6906–6917.
30. O. Kurnasov; L. Jablonski; B. Polanuyer; P. Dorrestein; T. Begley; A. Osterman, *FEMS Microbiol. Lett.* **2003**, *227*, 219–227.
31. O. Kurnasov; V. Goral; K. Colabroy; S. Gerdes; S. Anantha; A. Osterman; T. P. Begley, *Chem. Biol.* **2003**, *10*, 1195–1204.
32. N. Huang; L. Sorci; X. Zhang; C. A. Brautigam; X. Li; N. Raffaelli; G. Magni; N. V. Grishin; A. L. Osterman; H. Zhang, *Structure* **2008**, *16*, 196–209.
33. D. A. Rodionov; X. Li; I. A. Rodionova; C. Yang; L. Sorci; E. Dervyn; D. Martynowski; H. Zhang; M. S. Gelfand; A. L. Osterman, *Nucleic Acids Res.* **2008**, *36*, 2032–2046.
34. P. Rossolillo; I. Marinoni; E. Galli; A. Colosimo; A. M. Albertini, *J. Bacteriol.* **2005**, *187*, 7155–7160.
35. D. A. Rodionov; J. De Ingeniis; C. Mancini; F. Cimadamore; H. Zhang; A. L. Osterman; N. Raffaelli, *Nucleic. Acids Res.* **2008**, *36*, 2047–2059.
36. M. Daugherty; B. Polanuyer; M. Farrell; M. Scholle; A. Lykidis; V. de Crecy-Lagard; A. Osterman, *J. Biol. Chem.* **2002**, *277*, 21431–21439.
37. S. Y. Gerdes; M. D. Scholle; M. D'Souza; A. Bernal; M. V. Baev; M. Farrell; O. V. Kurnasov; M. D. Daugherty; F. Mseeh; B. M. Polanuyer; J. W. Campbell; S. Anantha; K. Y. Shatalin; S. A. Chowdhury; M. Y. Fonstein; A. L. Osterman, *J. Bacteriol.* **2002**, *184*, 4555–4572.
38. A. E. Choudhry; T. L. Mandichak; J. P. Broskey; R. W. Egolf; C. Kinsland; T. P. Begley; M. A. Seefeld; T. W. Ku; J. R. Brown; M. Zalacain; K. Ratnam, *Antimicrob. Agents. Chemother.* **2003**, *47*, 2051–2055.
39. D. A. Rodionov; A. G. Vitreschak; A. A. Mironov; M. S. Gelfand, *J. Biol. Chem.* **2002**, *277*, 48949–48959.
40. W. Winkler; A. Nahvi; R. R. Breaker, *Nature* **2002**, *419*, 952–956.
41. A. Eudes; G. B. Erkens; D. J. Slotboom; D. A. Rodionov; V. Naponelli; A. D. Hanson, *J. Bacteriol.* **2008**, *22*, 7591–7594.
42. A. V. Toms; A. L. Haas; J. H. Park; T. P. Begley; S. E. Ealick, *Biochemistry* **2005**, *44*, 2319–2329.
43. E. Morett; J. O. Korbelt; E. Rajan; G. Saab-Rincon; L. Olvera; M. Olvera; S. Schmidt; B. Snel; P. Bork, *Nat. Biotechnol.* **2003**, *21*, 790–795.
44. A. H. Jenkins; G. Schyns; S. Potot; G. Sun; T. P. Begley, *Nat. Chem. Biol.* **2007**, *3*, 492–497.
45. G. Schyns; S. Potot; Y. Geng; T. M. Barbosa; A. Henriques; J. B. Perkins, *J. Bacteriol.* **2005**, *187*, 8127–8136.
46. D. A. Rodionov; A. A. Mironov; M. S. Gelfand, *Genome Res.* **2002**, *12*, 1507–1516.
47. P. Hebbeln; D. A. Rodionov; A. Alfandega; T. Eitinger, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2909–2914.
48. D. A. Rodionov; M. S. Gelfand, *FEMS Microbiol. Lett.* **2006**, *255*, 102–107.
49. D. A. Rodionov, *Chem. Rev.* **2007**, *107*, 3467–3497.
50. D. A. Rodionov; A. G. Vitreschak; A. A. Mironov; M. S. Gelfand, *J. Biol. Chem.* **2003**, *278*, 41148–41159.
51. A. G. Vitreschak; D. A. Rodionov; A. A. Mironov; M. S. Gelfand, *RNA* **2003**, *9*, 1084–1097.
52. A. Nahvi; J. E. Barrick; R. R. Breaker, *Nucleic Acids Res.* **2004**, *32*, 143–150.
53. C. L. Zayas; J. D. Woodson; J. C. Escalante-Semerena, *J. Bacteriol.* **2006**, *188*, 2740–2743.
54. J. D. Woodson; R. F. Peck; M. P. Krebs; J. C. Escalante-Semerena, *J. Bacteriol.* **2003**, *185*, 311–316.



55. C. Fan; T. A. Bobik, *J. Biol. Chem.* **2008**, *283*, 11322–11329.
56. A. Bacher; S. Eberhardt; W. Eisenreich; M. Fischer; S. Herz; B. Illarionov; K. Kis; G. Richter, *Vitam. Horm.* **2001**, *61*, 1–49.
57. W. C. Winkler; R. R. Breaker, *Annu. Rev. Microbiol.* **2005**, *59*, 487–517.
58. A. G. Vitreschak; D. A. Rodionov; A. A. Mironov; M. S. Gelfand, *Trends Genet.* **2004**, *20*, 44–50.
59. M. Y. Galperin; E. V. Koonin, *Genetica* **1999**, *106*, 159–170.
60. M. S. Gelfand; D. A. Rodionov, *Phys. Life Rev.* **2008**, *5*, 22–49.
61. R. K. Aziz; D. Bartels; A. A. Best; M. DeJongh; T. Disz; R. A. Edwards; K. Formsma; S. Gerdes; E. M. Glass; M. Kubal; F. Meyer; G. J. Olsen; R. Olson; A. L. Osterman; R. A. Overbeek; L. K. McNeil; D. Paarmann; T. Paczian; B. Parrello; G. D. Pusch; C. Reich; R. Stevens; O. Vassieva; V. Vonstein; A. Wilke; O. Zagnitko, *BMC Genomics* **2008**, *9*, 75.
62. M. Graupner; H. Xu; R. H. White, *J. Bacteriol.* **2002**, *184*, 1952–1957.
63. C. A. Bonner; T. Disz; K. Hwang; J. Song; V. Vonstein; R. Overbeek; R. A. Jensen, *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 13–53; table of contents.
64. E. Merino; R. A. Jensen; C. Yanofsky, *Curr. Opin. Microbiol.* **2008**, *11*, 78–86.
65. G. Xie; N. O. Keyhani; C. A. Bonner; R. A. Jensen, *Mol. Biol. Rev.* **2003**, *67*, 303–342, table of contents.

### Biographical Sketches



Andrei L. Osterman received his Ph.D. in Biochemistry from the Moscow University in Russia. The main focus of his research, including postdoctoral studies at UT Southwestern in Dallas, was in the field of mechanistic and structural enzymology. In 1999, he joined Integrated Genomics, a start-up biotech firm in Chicago, to lead an effort on genomics-driven discovery of metabolic enzymes and pathways. He joined the Burnham Institute for Medical Research in La Jolla, CA as an associate professor in the Bioinformatics and Systems Biology Program in 2003. He is one of the founders of the Fellowship for Interpretation of Genomes (FIG), and he actively participates in the Project to Annotate 1000 Genomes launched by FIG in 2003. His research group combines bioinformatics and experimental techniques to elucidate the key metabolic subsystems, such as biogenesis of vitamins and cofactors, across hundreds of diverse species with completely sequenced genomes.



Ross Overbeek received a doctorate in computer science in 1972 from Pennsylvania State University. He taught at Northern Illinois University for 11 years (mathematics and



computer science). His research areas were computational logic and database systems. From about 1983 to 1998, he worked at Argonne National Laboratory, focusing on parallel computation and logic programming. Although a senior scientist at ANL, he was convinced that interest in high-performance computing was losing ground. At that critical moment, he met Carl Woese, who convinced him that the most important part of science during the next decades would be in the field of biology, and that it would be driven by comparative analysis based on a rapidly growing body of genomic sequence data. He collaborated with Woese and participated in the founding of the Ribosomal Database Project. He went on to participate in the analysis of *M. jannaschii* (the first archaeal genome). He was the lead architect of the PUMA and WIT systems at ANL before becoming a founder of Integrated Genomics (where he spent 1998–mid-2003). While at IG, he participated in the sequencing and analysis of over 50 genomes and led the bioinformatics effort. The most significant product was ERGO, a system to support comparative analysis. In mid-2003, he left IG to become a founding fellow of FIG (the Fellowship for Interpretation of Genomes). His efforts at FIG have centered on building a new system for comparative analysis, the SEED, which will be an open source and free for all. Since 2004, he has been co-PI of the National Microbial Pathogen Data Resource, a framework to support comparative analysis of pathogen genomes.



Dmitry A. Rodionov received his M.Sc. degree from the Moscow Engineering Physics Institute in 2000 and his Ph.D. degree in Molecular Biology from the State Research Institute for Genetics and Selection of Industrial Microorganisms (Moscow, Russia) in 2003. He was a postdoctoral fellow in the group of Prof. Mikhail S. Gelfand at the Institute for Information Transmission Problems of Russian Academy of Sciences (IITP RAS, Moscow, Russia), and in the group of Prof. Andrei Osterman at the Burnham Institute for Medical Research (BIMR, La Jolla, CA, USA). He participated in multiple collaborative projects as a visitor scientist at National Center for Biotechnology Information (Bethesda, MD, USA), Unit of Microbiology and Genetics (INSA-Lyon, France), Lawrence Berkeley National Laboratory (CA, USA), and Institute for Microbiology at Humboldt University (Berlin, Germany). He is currently a research assistant professor at BIMR (CA, USA). His research interests are different aspects of comparative genomic reconstruction of metabolic pathways and regulatory networks in prokaryotes including identification of transcription factor-binding sites and RNA regulatory elements, functional annotation of genes and identification of missing genes in metabolic pathways, phylogenetic analysis of protein families, and evolution of metabolic pathways and regulatory systems.

## 7.06 Biosynthesis of Biotin

Andrée Marquet, Université Pierre et Marie Curie, Paris, France

© 2010 Elsevier Ltd. All rights reserved.

7.06.1	Introduction	161
7.06.2	The Biosynthetic Pathway	161
7.06.2.1	The Different Steps	163
7.06.2.1.1	The origin of pimeloyl-CoA	163
7.06.2.1.2	8-Amino-7-oxopelargonic acid (AOP) synthase	165
7.06.2.1.3	7,8-Diaminopelargonic acid aminotransferase	169
7.06.2.1.4	Dethiobiotin synthetase	171
7.06.2.1.5	Biotin synthase	174
7.06.2.2	Regulation of the Pathway	177
References		178

### 7.06.1 Introduction

(+) Biotin, also called vitamin H (**Figure 1**), which belongs to the water-soluble group of vitamins, is produced by bacteria, plants, and a few fungi.

It plays a very important physiological role as the prosthetic group of several carboxylases. Biotin, always covalently bound to its partner enzymes serves as a CO<sub>2</sub> carrier between bicarbonate and the acceptor substrate. *N*-Carboxybiotin is first produced from HCO<sub>3</sub><sup>-</sup> and ATP, and this activated form of CO<sub>2</sub> is then transferred to the substrate. These reactions are essential in central metabolism, including fatty acid biosynthesis and gluconeogenesis, as well as in secondary metabolism.<sup>1-3</sup> It has also been recognized that biotin influences the expression of genes, through different pathways, the best documented one being the biotinylation of histones. This process appears to play a role in cell proliferation, gene silencing, and cellular response to DNA repair.<sup>3,4</sup>

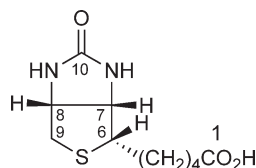
The biotin content of several cell types has been estimated. Besides biotin covalently bound to the carboxylases, the pool of free biotin is generally low. For instance, *Escherichia coli* can grow with only 100–200 molecules of biotin per cell.<sup>5</sup> In plants, the free biotin pool is higher, with a ratio free/bound of around 6 in pea leaf protoplasts, for example.<sup>6</sup> Anyway, the biotin requirement of most organisms is low. In mammals, some amount is synthesized by intestinal bacteria, food being the main source. Biotin present in food, covalently linked to the proteins, is liberated during the digestive process, the last enzyme of which is biotinidase. Biotin is transported in the blood, taken up by the cells where it can be recycled to biotinylate new proteins. In humans, no severe biotin deficiency has been observed, except in cases of some inherited ones resulting in a less efficient recycling.<sup>3,7,8</sup> Deficiencies related to the diet are rare. However, some cases have been observed in animals and biotin is often added to animal food. It is also used in fermentation broths and in cosmetics, where it proved to be beneficial.

Biotin is still obtained industrially by chemical synthesis, in spite of many efforts to produce it by fermentation.<sup>9</sup>

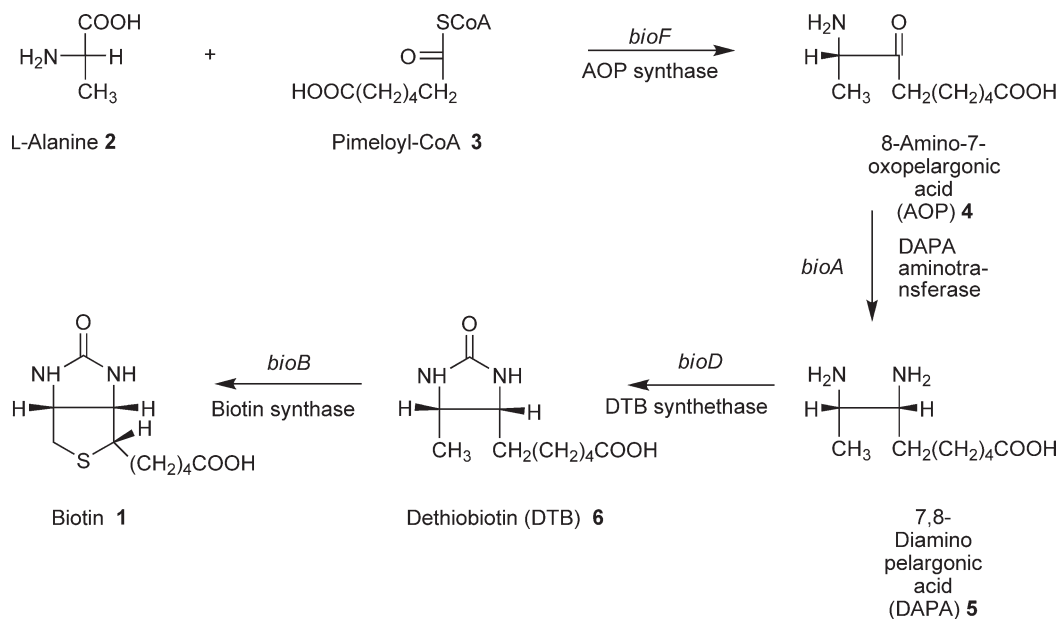
### 7.06.2 The Biosynthetic Pathway

The pathway for biotin biosynthesis elucidated in bacteria by application of the mutants technique was proposed by Eisenberg in 1968.<sup>10</sup> This pathway, which has never been questioned ever since,<sup>8</sup> is depicted in **Figure 2**. It has also been shown that the same pathway is present in plants.<sup>11</sup>

The genetics and enzymology of biotin biosynthesis have been the subject of a reinforced interest when several industrial companies undertook studies directed toward the production of biotin by fermentation, using recombinant strains. This has stimulated a great deal of work for the identification of the *bio* genes of many



**Figure 1** The structure of biotin.



**Figure 2** The biotin biosynthetic pathway.

microorganisms and the overexpression and characterization of these genes products. The *bio* genes have been first identified in *E. coli* and *Bacillus spbaericus*. In *E. coli*, they are organized in a divergent operon (*bioABFCD*). A rightward transcription unit includes *B*, *F*, *C*, and *D*, coding (**Figure 2**) for biotin synthase (BS), 8-amino-7-oxopelargonic acid (AOP) synthase, an unknown enzyme, and dethiobiotin (DTB) synthetase, respectively. *BioA*, corresponding to 7,8-diaminopelargonic acid (DAPA) aminotransferase is leftward transcribed. Another *bio* gene, *bioH* is located outside the operon. The biotin operon has been completely sequenced by Otsuka *et al.*<sup>12</sup> and *bioH* by O'Regan *et al.*<sup>13</sup> A different organization was found in *B. spbaericus*<sup>14</sup> since the *bio* genes are located in two separate operons *bioXWF* and *bioDAYB*. *BioF*, *D*, *A*, and *B* have the same function as in *E. coli*, and *W* encodes a pimeloyl-CoA synthetase (from pimelate). The genes products of *bioC* and *bioH* in *E. coli* and of *bioX* and *Y* in *B. spbaericus* are involved in the synthesis of pimeloyl-CoA (see Section 7.06.2.1.1).

Presently, the *bio* genes have been totally or partially characterized in several organisms.<sup>8,9</sup> This will not be compiled in this review where we shall focus mainly on the mechanistic problems. One can, however, point out that *A*, *B*, *F*, and *D* are universally found, with an important degree of homology, whereas several classes of organisms can be distinguished for the synthesis of pimeloyl-CoA.

The availability of the purified enzymes in large amounts has allowed mechanistic investigations and structural studies. The interest in the fundamental mechanistic enzymology of the pathway, which includes very intriguing reactions, was reinforced by the observation that biotin auxotrophy and biotin biosynthesis inhibitors have a lethal effect in plants. Thus, the enzymes involved in biotin biosynthesis represent targets for the discovery of new herbicides devoid of toxicity for animals,<sup>2,15</sup> and special attention has been paid to the design of inhibitors. Inhibitors of the pathway have similarly been reported to have antibiotic properties against many microorganisms *in vitro*; however, these activities are generally reversed by biotin and disappear *in vivo*

because of the presence of biotin. This seems to be less pronounced in plants, and also in a very important bacteria, *Mycobacterium tuberculosis*.<sup>16</sup>

Several extensive reviews on biotin biosynthesis have been published.<sup>8,10,17</sup> The early work will be summarized to the extent necessary for an independent reading, but we shall focus on the recent results.

## 7.06.2.1 The Different Steps

### 7.06.2.1.1 The origin of pimeloyl-CoA

If there is only one biosynthetic pathway for biotin, present in all organisms and rather easily established, starting from pimeloyl-CoA, the origin of this simple molecule has been discussed for a long time. It is just starting to be understood, and some questions are still pending. Several pathways to pimeloyl-CoA have been identified, depending on the organism considered.

**7.06.2.1.1(i) Conversion of pimelic acid into pimeloyl-CoA** Izumi *et al.*<sup>18</sup> have observed a pimeloyl-CoA synthesis from pimelate in various cell-free extracts, but the reaction has been completely characterized only later on in *B. sphaericus*. All activities required for the transformation of pimelate into AOP were encoded by the *bioXWF* cluster. BioF was identified as the AOP synthase, and initial genetic experiments suggested that bioW was encoding a pimeloyl-CoA synthase.<sup>14</sup> The enzyme has been purified from an *E. coli* strain overproducing this gene. From an assay containing pimelate, CoA, and ATP, pimeloyl-CoA and AMP were identified as products, and the reaction can be written as shown in **Figure 3**.

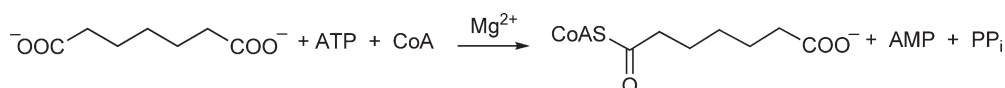
The enzyme is very specific since it does not accept alternate substrates for pimelate. No significantly similar protein was found in the NBRF-PIR library. The use of pimelate by *B. sphaericus* implies that the cells are permeable to it and *bioX* was suspected to be a permease. The uptake of [<sup>3</sup>H]pimelate was thus studied and it was observed that it was entering the cells by passive diffusion (no permease behavior, no energy dependence), a result that left open the function of *bioX*.<sup>19</sup>

A pimeloyl-CoA synthetase, the *pauA* gene product has also been isolated from *Pseudomonas mendocina* 35. Its specificity is not as high as that of *B. sphaericus*, the relative activity for adipic acid was 72% and for azelaic acid 18% of that for pimelic acid. The permeability of this organism toward pimelate has not been specifically examined, but as its choice was the result of a screening of different bacteria for growth on pimelic acid, permeability can be inferred.<sup>20</sup> In *Bacillus subtilis*, the pimeloyl-CoA synthetase activity has not been investigated *in vitro*, but in the biotin operon, a gene homologous to *bioW* is present.<sup>21</sup> In plants, labeled pimelate has been efficiently incorporated into biotin using cell cultures of lavender.<sup>6</sup>

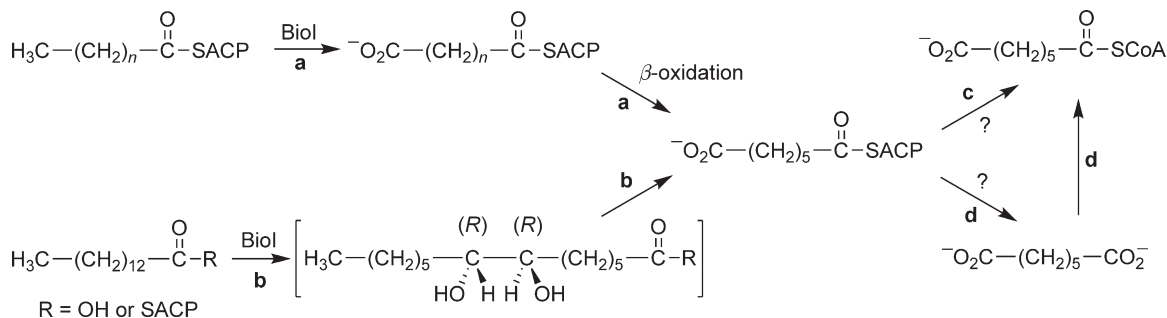
Thus, one can conclude that all these organisms can use pimelate as the precursor of biotin, via pimeloyl-CoA, but it just shifts the question to the origin of pimelate.

**7.06.2.1.1(ii) The origin of pimelate** In *B. subtilis*: The most complete study concerns *B. subtilis*. In the biotin operon, two additional open reading frames (ORFs) were found besides the ORFs homologous to *bioF*, *D*, *A*, and *B* of *E. coli* and *B. sphaericus*, and *bioW* of *B. sphaericus*. The first one, *bioI* was homologous to a number of bacterial cytochrome P-450.<sup>21</sup> *BioI* is able to complement *bioC* and *bioH* mutants of *E. coli*, indicating that it can produce pimeloyl-CoA in the absence of BioW, the pimeloyl-CoA synthetase. When *BioI* was overexpressed in *E. coli*, it copurified with the acyl carrier protein (ACP) of *E. coli*, covalently bound to different fatty acids. When this complex was incubated with a reducing system, an obligatory partner of cytochrome P-450, no oxidation product was detected, unless a saponification step was added. In that case, a single product, pimelate, was detected. It has hence been produced through pimeloyl ACP.<sup>22</sup>

*BioI* has now been purified and well characterized as a P-450<sup>23,24</sup> and its redox partners investigated. The unique ferredoxin gene *fer* in the genome of *B. subtilis* has been overexpressed and the Fer protein, a (4Fe-4S)



**Figure 3** The reaction catalyzed by pimeloyl-CoA synthase.



**Figure 4** Postulated mechanisms for pimelate synthesis in *Bacillus subtilis*.

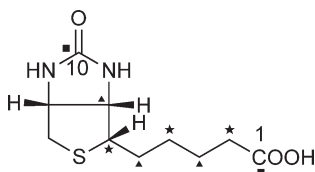
ferredoxin, has been fully characterized.<sup>25</sup> Two putative flavodoxines found at another position in the gene, part of the Fur regulon could replace Fer in situations of iron deficiencies,<sup>26</sup> but the reductase of *B. subtilis* has not yet been isolated. However, functional *in vitro* systems have been developed using ferredoxin/flavodoxin and reductases of other organisms. All these systems were efficient to reduce BioI and oxidize free fatty acids.

Two independent studies with free fatty acids have been reported (Figure 4). For Green *et al.*,<sup>23</sup> the first step is an  $\omega$ -hydroxylation that should take place on rather rare odd chain fatty acids to lead to pimelate by successive  $\beta$ -oxidations (Figure 4, a), whereas Cryle *et al.*<sup>27,28</sup> observe an in-chain hydroxylation that occurs with different even chain fatty acids. Indeed, the latter authors brought convincing evidence that pimelate was formed from myristic acid through a double hydroxylation at C-7 and C-8 [(*R,R*) 7,8-dihydroxytetradecanoic acid], followed by C-C bond cleavage (Figure 4, b). The mechanism is analogous to that of P-450<sub>scc</sub>, which cleaves the side chain of cholesterol. BioI is thus the second P-450 identified as being capable of such bond fission, and the first from a prokaryote.

This raises the question of the physiological substrates of BioI *in vivo*, free or ACP-bound fatty acids. It is tempting to consider that it is the last one. Cryle and De Voss<sup>28</sup> consider that the mechanism they have demonstrated starting with free fatty acids is valid for ACP-bound ones. If pimeloyl ACP is the product formed by BioI, the way it is converted into pimeloyl-CoA has still to be clarified. It can be a transesterification (Figure 4, c) or hydrolysis to free palmitate, which would then be converted to pimeloyl-CoA (Figure 4, d). If it is obtained according to Figure 4, c, the presence of BioW is difficult to rationalize. Of course, the presence of BioW explains why pimelate can be used as a precursor of biotin.

In *B. sphaericus*: The situation is quite different in *B. sphaericus*. No *biol* homolog has been identified in its genome, and there is till now no argument favoring the production of pimelate from higher fatty acids. It contains instead, as mentioned above, the *X* gene, belonging to the *bioXWF* cluster, which has not yet been definitely identified. However, it was recognized that the consensus phosphopantetheine attachment site found in several ACPs was present in *bioX*, and it was suggested that it could encode a specific ACP involved in pimeloyl-CoA synthesis, eventually a specific ACP for carrying a malonyl moiety, in a nonclassical fatty acid biosynthesis.<sup>29</sup> If it is the case, it means that in two closely related Gr(+) organisms, pimelate is formed by two completely different routes. In this hypothesis, the intermediacy of free pimelate remains open, as in the case of *B. subtilis*.

In *E. coli*: The conversion of pimelate into pimeloyl-CoA does not take place in *E. coli*. Preliminary labeling experiments with <sup>14</sup>C<sub>2</sub>, which ruled out the symmetrical pimelate as intermediate, were confirmed by further studies using <sup>13</sup>C-labeled acetate, with *E. coli* strains overexpressing the biotin operon.<sup>30,31</sup> Nuclear magnetic resonance (NMR) analysis of DTB or biotin revealed that carbons 3, 5, 7 and 2, 4, 6 were obtained from C-1 and C-2 of acetate, respectively (Figure 5). Sanyal *et al.*<sup>31</sup> confirmed using [1,2-<sup>13</sup>C]acetate that intact acetate units were incorporated, consistent with a classical fatty acid or polyketide pathway. The results of the two groups differ, however, on an important point. Sanyal *et al.* did not observe <sup>13</sup>C incorporation at C-1 or C-10 with 1- or 2-<sup>13</sup>C acetate and none, either with <sup>13</sup>CO<sub>2</sub>, whereas Ifuku *et al.* reported enrichment at C-1 with the labeled acetates, higher with [1-<sup>13</sup>C] than with [2-<sup>13</sup>C]acetate.



**Figure 5**  $^{13}\text{C}$ -Labeling of biotin isolated from feeding experiments with the following: ★,  $[2-^{13}\text{C}]$ acetate; ▲,  $[1-^{13}\text{C}]$ acetate;<sup>30,31</sup> and ■,  $[1-^{13}\text{C}]$ acetate or  $[2-^{13}\text{C}]$ acetate.<sup>30</sup>

The origin of the discrepancies between the two groups is not clear and an unambiguous experimental proof of the origin of C-1 is still awaited.<sup>8</sup> An interpretation assuming a malonyl-CoA as starter (**Figure 6**) seems, however, reasonable.

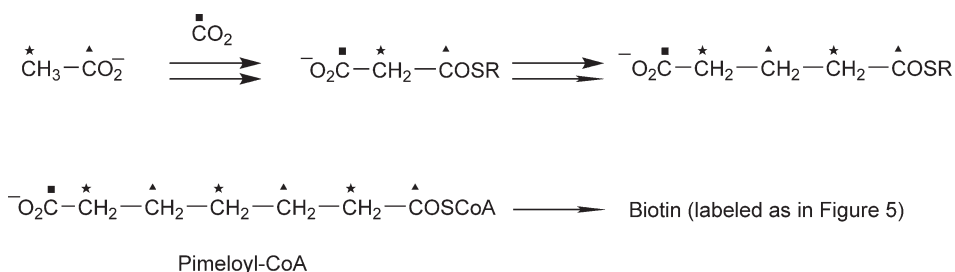
In *E. coli*, there are two genes, *bioC* and *bioH* that encode enzymes involved in pimeloyl-CoA synthesis, which have no homology with *bioX* and *bioW* of *B. sphaericus*, whereas sequences of all the other genes of the *bio* operon show an average of 50% identity. The BioH protein has now been well characterized. It is a carboxyl esterase, as predicted by sequence analysis, biochemical characterization,<sup>32</sup> and finally, resolution of its three-dimensional (3D) structure, which reveals clearly the classical catalytic triad Asp–His–Ser.<sup>33</sup> Since it has to produce pimeloyl-CoA without forming pimelate, one can speculate that it transfers the pimeloyl moiety from a pimeloyl ACP to CoA. A hypothesis is that BioC could be this specific ACP binding the malonyl moiety.<sup>29,31</sup> But as pointed above, there is no homology between *BioC* of *E. coli* and *bioX* of *B. sphaericus* although similar functions have been proposed for the corresponding proteins. In both cases, the condensing enzymes of the postulated nonclassical fatty acid synthase should not belong to the biotin operon.

**Table 1** summarizes the presence of the genes in the three cases discussed here.

Clearly, the different pathways to pimeloyl-CoA are not completely unraveled. Consideration of other organisms increases the complexity: BioH is not conserved in all organisms in which *bioC* is observed, where it is replaced by different genes.<sup>34</sup>

### 7.06.2.1.2 8-Amino-7-oxopelargonic acid (AOP) synthase

**7.06.2.1.2(i) Characterization of the enzyme** Early studies on the AOP synthase of *E. coli*<sup>10</sup> and *B. sphaericus*<sup>35</sup> have shown that the only requirement for the reaction, besides the two substrates, L-Ala and pimeloyl-CoA (**Figure 7**), was pyridoxal phosphate (PLP). Later on, AOP synthases of several organisms have been

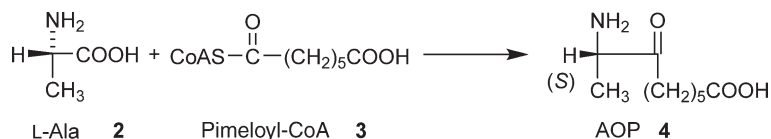


**Figure 6** Proposed pathway for the origin of pimeloyl-CoA in *Escherichia coli*.

**Table 1** Genes (or homologs) involved in the synthesis of pimeloyl-CoA in different organisms

	<i>BioW</i>	<i>BioI</i>	<i>BioC</i>	<i>BioH</i>
<i>Escherichia coli</i>	–	–	+	+
<i>Bacillus sphaericus</i>	+	–	–	–
<i>Bacillus subtilis</i>		+	–	–





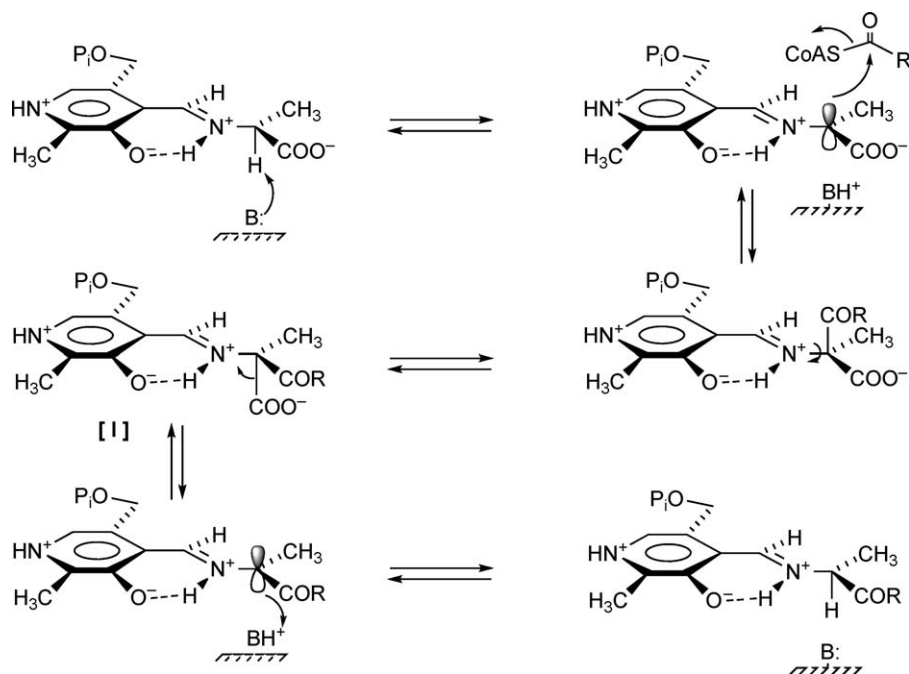
**Figure 7** The reaction catalyzed by AOP synthase.

cloned and the enzymes from *B. spbaericus*,<sup>36</sup> *E. coli*,<sup>37</sup> *M. tuberculosis*,<sup>38</sup> *Thermus thermophilus*,<sup>39</sup> and *Arabidopsis thaliana*<sup>40</sup> have been purified and characterized.

AOP synthase belongs to the small group of PLP enzymes that catalyze the condensation of an acyl-CoA with an  $\alpha$ -amino acid, leading to an  $\alpha$ -amino ketone. There are three other identified enzymes in this subfamily, referenced in the papers cited above: 5-amino levulinate synthase (ALS) that catalyzes the first step of the porphyrin ring of hemes; serine palmitoyl transferase (SPT), the first enzyme in the sphingolipids biosynthetic pathway; and 2-amino-3-oxobutyryl-CoA ligase (KBL), involved in threonine degradation. They all share strong sequence similarities.

Three AOP synthases have been mechanistically studied in detail, those of *B. spbaericus*, *E. coli*, and *M. tuberculosis*. Two plausible mechanisms, already discussed for ALS and SPT had to be considered. The one which is proposed, depicted in **Figure 8**,<sup>36</sup> starts with the abstraction of the C-2 proton on the Schiff base of alanine followed by acylation and decarboxylation. In the alternative one, the carbanion would be produced by decarboxylation of the Ala Schiff base.

These two mechanisms could be distinguished by exchange experiments. In the first one, the C-2 proton of alanine has to be lost in the product, whereas it should be retained if the carbanion is generated by decarboxylation. It has been shown in the cases of ALS and SPT that the C-2 proton was indeed lost. The same methodology was applied to AOP synthase, the reaction being followed by NMR.<sup>36</sup> When the enzyme was incubated with pimeloyl-CoA and either L-2- $^{[2}\text{H}$ ]Ala in  $\text{H}_2\text{O}$  or L-Ala in  $^{[2}\text{H}_2]\text{O}$ , it was observed that the C-2 hydrogen of alanine was lost, and that AOP had incorporated one  $^{[2}\text{H}$ ] from the solvent. As no exchange at



**Figure 8** Overall reaction mechanism proposed for AOP synthase<sup>36</sup> [R = (CH<sub>2</sub>)<sub>5</sub>COO<sup>-</sup>].

C-2 was observed in the residual alanine, it can be inferred that once the carbanion is formed, it reacts readily with the thioester. The deuterium kinetic isotope effect determined with  $[^2\text{H}]\text{Ala}$  was weak ( $^{D}V/^{H}V=1.3$ ), indicating that the proton abstraction was only partially rate determining.

The proposed stereochemistry represented in **Figure 8** relies on the fact that most PLP-dependent decarboxylations occur with retention and with the assumption that it is the same base that abstracts the C-2 proton and reprotonates the final intermediate. This last point is consistent with the observation that when L-Ala was incubated alone with the enzyme in  $[^2\text{H}_2]\text{O}$ , a slow exchange was observed (with a rate constant 10 times less than the catalytic constant) and with complete retention of configuration. The overall picture represented in **Figure 8** implies that the first condensation step occurs with inversion.

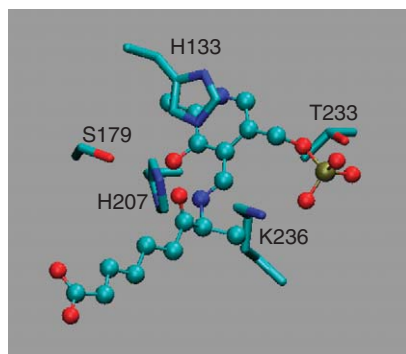
Webster *et al.*<sup>41</sup> have carried out a thorough spectroscopic and kinetic analysis of the reaction with the *E. coli* enzyme, showing the appearance and decay of different intermediates. A very strong argument in favor of the proposed mechanism was recently presented by Kerbach *et al.*<sup>42</sup> When alanine methyl ester was used instead of alanine, thus preventing the decarboxylation of intermediate [I], they observed indeed the accumulation of a new product that was characterized as the methyl ester of [I] by mass spectrometry.

The structures of the apo- and PLP-bound forms (**Figure 9**) of the *E. coli* enzyme, at 1.61 and 2.1 Å resolution, respectively, have been published.<sup>37</sup> The protein is a symmetrical homodimer, each subunit consisting of three domains. PLP is bound to Lys236. Lys236, which is the only conserved Lys, is located at the end of a deep cleft that allows access of the pantothenate arm of pimeloyl-CoA. It is postulated to be the base involved in the deprotonation of the Schiff base. Modeling experiments lead to the conclusion that the acyl-CoA moiety then has to approach on the opposite side,<sup>37</sup> consistent with the proposal made by Ploux and Marquet.<sup>36</sup> These structures reveal the amino acids of the active site presumably involved in the catalysis, residues that are conserved in the different AOP synthases, which have been sequenced and aligned.

The important catalytic residues are also conserved in other  $\alpha$ -acyl-CoA oxoamine synthases indicating that they have similar active site geometries in good agreement with the observed mechanistic analogy.<sup>37</sup>

The enzyme of *B. spbaericus* was also crystallized<sup>43</sup> and its 3D structure to 2.2 Å resolution is very similar to that of the *E. coli* one (Cambillau, Ploux, Marquet, unpublished results).

As mentioned above, two other AOP synthases have been studied, although less extensively: one from a thermophilic organism (*T. thermophilus*) and the second from an eukaryotic one, *A. thaliana*. Their mechanisms, evidenced by the spectroscopic characterization of the intermediates, are identical. There are some differences in the kinetic parameters (**Table 2**), and interestingly, in the substrate specificity. Analogs of pimeloyl-CoA were tested only with *A. thaliana* and *T. thermophilus*. Whereas all the compounds assayed were neither substrates nor inhibitors with the *A. thaliana* enzyme, with L-Ala as cosubstrate, the thermophilic one also accepts palmitoyl-CoA and acetyl-CoA, which are indeed better substrates than pimeloyl-CoA. Different amino acids were also tested with these two enzymes. The first one accepts Ser with pimeloyl-CoA (specific activity 35%/Ala), the second one Gly and Ser (specific activity 50%/Ala). Interestingly, with *T. thermophilus*, the best specific activity is obtained with the couple acetyl-CoA and glycine, that is, the substrates of KBL.



**Figure 9** Structure of the AOP-PLP external aldimine in the active site of AOP synthase.<sup>41</sup>

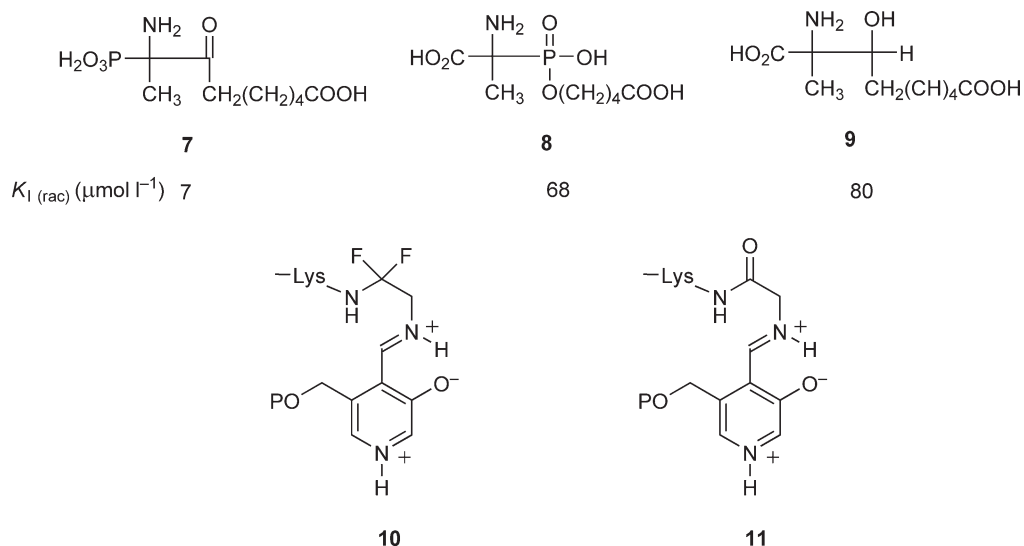
**Table 2** Steady-state kinetic constants of AOP synthase

	<i>L</i> -Alanine		<i>Pimeloyl</i> -CoA	
	$K_M$ (mmol l <sup>-1</sup> )	$k_{cat}/K_M$ (s <sup>-1</sup> mol l <sup>-1</sup> )	$K_M$ (μmol l <sup>-1</sup> )	$k_{cat}/K_M$ (s <sup>-1</sup> mol l <sup>-1</sup> )
<i>Bacillus sphaericus</i> <sup>36</sup>	2.0	150	1.5	2 × 10 <sup>5</sup>
<i>Escherichia coli</i> <sup>41</sup>	0.5	120	25	2400
<i>Mycobacterium tuberculosis</i> <sup>38</sup>	467	15	1.5	2600
<i>Thermus thermophilus</i> <sup>39</sup>	12	7	80	1730
<i>Arabidopsis thaliana</i> <sup>40</sup>	1400	78	1.6	62 500

D-Ala is not substrate of the *T. thermophilus* enzyme and was not considered in the study with *A. thaliana*. On the other hand, its behavior was carefully studied in the three other cases. Consistent results were obtained with *E. coli* and *B. sphaericus*. Ploux *et al.*<sup>44</sup> found that D-Ala formed the external aldimine but did not exchange its α-proton and did not give the first quinonoid intermediate, but was a competitive inhibitor of L-Ala (with a  $K_I$  of 0.6 mmol l<sup>-1</sup>, smaller than the  $K_M$  of L-Ala). Other experiments showed that the binding of D-Ala was 10 times weaker than that of L-Ala, but increased in the presence of pimeloyl-CoA to a value more compatible with the observed  $K_I$  value. It was suggested that the binding of pimeloyl-CoA induces a conformational change in the active site. This induced active site topology might be complementary to the reaction intermediate [I], which is supposed to have the opposite configuration of L-alanine. The carboxylate group of D-Ala might occupy the same position as that of the intermediate. The same conclusions were reached for the *E. coli* enzyme.<sup>41</sup>

A different situation is observed in the case of *M. tuberculosis*.<sup>38</sup> D-Ala forms the external aldimine, either, but also, slowly, the first quinonoid intermediate. It is claimed that this form reacts with pimeloyl-CoA to give D-AOP. However, the characteristics of the reference synthetic AOP are those of the racemic compound, as already discussed.<sup>45</sup> Thus, this reaction with D-Ala should be reinvestigated. It was also reported that D-Ala inhibits the reaction with L-Ala; however, the inhibition is no longer competitive, but of the linear mixed type, which would mean that the two enantiomers bind independently at different sites.

**7.06.2.1.2(ii) Search for inhibitors** Several phosphonic derivatives, designed as analogs of the reaction intermediates, have been tested.<sup>44</sup> Three of them, **7**, **8**, and **9** (Figure 10) were found to be competitive inhibitors of L-Ala with good affinities, revealed by the low values of  $K_I/K_M$  Ala. The best one is compound **7**, an analog of [I], which cannot decarboxylate, with a  $K_I$  of 7 μmol l<sup>-1</sup> (3.5 μmol l<sup>-1</sup> if we assume that only one

**Figure 10** Inhibitors of AOP synthase.

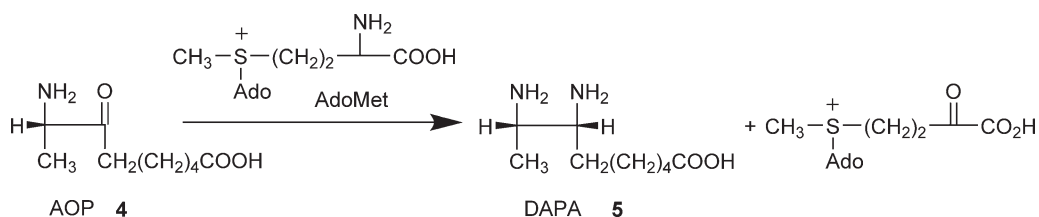
enantiomer is recognized). It is a slow binding inhibitor, competitive with L-Ala. The two others are transition state analogs of the condensation step with the acyl-CoA.

Alexeev *et al.*<sup>46</sup> have examined the inhibition by trifluoro alanine, according to the well-known ability of fluoro amino acids to inactivate PLP-dependent enzymes. They indeed observe an irreversible inactivation and established the structure of the adduct. When the crystals of the protein were soaked in a buffer containing trifluoro alanine, they observed the adduct **10** (Figure 10), whereas when the incubation was carried out in solution, the amide **11**, resulting from HF elimination, was obtained. The postulated mechanism implies a decarboxylation of the external aldimine in the first step. This is rather surprising since in the normal mechanism, decarboxylation occurs in the intermediate [I].

### 7.06.2.1.3 7,8-Diaminopelargonic acid aminotransferase

**7.06.2.1.3(i) Characterization of the enzyme** The reaction catalyzed by 7,8-diaminopelargonic acid aminotransferase (DAPA AT) (Figure 11) was first characterized in *E. coli* cell-free extracts by Pai,<sup>47</sup> who first determined the basic requirements for the reaction. The enzyme was then purified and further kinetically studied by Stoner and Eisenberg.<sup>48</sup> Two further studies confirmed these early results.<sup>45,49</sup> Izumi *et al.*<sup>50</sup> and Hotta *et al.*<sup>51</sup> did a similar work with the DAPA AT of *Brevibacterium divaricatum*. Recently, the enzyme of *M. tuberculosis* has been cloned and characterized.<sup>16</sup> The conclusion of all these studies are in good agreement: The reaction, like all transamination reactions, occurs in two steps: (1) formation of pyridoxamine by NH<sub>2</sub> transfer from the NH<sub>2</sub> donor to PLP and (2) transfer of NH<sub>2</sub> from pyridoxamine to the ketonic substrate. The enzyme follows a ping-pong bi-bi mechanism, like other transaminases and is inhibited by an excess of substrate. This enzyme presents, however, a unique feature, which is the use of AdoMet as NH<sub>2</sub> donor in the first half reaction. Breen *et al.*<sup>52</sup> have definitely identified the product obtained from AdoMet as 4-(methylthioadenosyl)-2-oxobutanoate (Figure 11). It is the only example where AdoMet is the NH<sub>2</sub> source in transamination reactions. It was, however, recently shown that another unusual amino donor is used by *B. subtilis*, namely, L-lysine.<sup>53</sup>

Kinetic parameters for the two half reactions are reported in Table 3. It has to be pointed out that there is some confusion in the literature regarding the enantiomeric purity of the AOP samples used in the assays. Mann *et al.* used racemic AOP, but they have shown unambiguously that only (*S*)-AOP is the substrate of DAPA AT.<sup>45</sup> (Both enantiomers support the growth of *Saccharomyces cerevisiae*, but it is probably due to the racemization of (*R*)-AOP *in vivo*.) It is claimed in the other studies that the AOP samples used as substrates had the natural 8-(*S*) configuration but they are very likely also racemic compounds, since they have been obtained under racemizing conditions,<sup>54</sup> as discussed by Mann *et al.*<sup>45</sup> Only in the study performed with the



**Figure 11** The reaction catalyzed by DAPA aminotransferase.

**Table 3** Steady-state kinetic constants of DAPA aminotransferase

Organism	$K_M$ AdoMet (mmol l <sup>-1</sup> )	$K_M$ AOP (rac) (μmol l <sup>-1</sup> )	$K_I$ AOP (μmol l <sup>-1</sup> )	$k_{cat}$ (s <sup>-1</sup> )
<i>Escherichia coli</i> <sup>48</sup>	0.2	1.2	25	0.1
<i>Escherichia coli</i> <sup>49</sup>	0.15	<2	25	0.013
<i>Escherichia coli</i> <sup>45</sup>	–	1	–	0.08
<i>Mycobacterium tuberculosis</i> <sup>16</sup>	0.78	3.8	14	60 <sup>-1</sup>
<i>Bacillus subtilis</i> <sup>53</sup>	2–25 <sup>a</sup>	1–5	–	–

<sup>a</sup> $K_M$  lys.

*B. subtilis* enzyme,<sup>53</sup> where AOP was obtained by fermentation, it could be the (*S*) enantiomer, provided it was not racemized during the isolation.

The recombinant protein of *E. coli* has been expressed, completely purified and crystallized, and its 3D structure was determined to 1.8 Å resolution by X-ray crystallography.<sup>55</sup> The enzyme is dimeric with two equal active sites. Each monomer contains two domains, residues 50–329 and residues 1–49 plus 330–429. PLP is bound to a cleft formed by domains of both subunits anchored to the conserved Lys274. The structure of the nonproductive ternary complex, enzyme–PLP–AOP has also been solved. As no structure of a complex containing SAM is available, a complete discussion of the residues involved in the catalytic steps is not yet possible. However, Eliot *et al.*<sup>49</sup> and Sandmark *et al.*<sup>56</sup> have studied different mutants and propose a safe working hypothesis.

**7.06.2.1.3(ii) Search for inhibitors** Inhibition of DAPA aminotransferase by various analogs of AdoMet and by 7-amino-8-oxo pelargonic acid, a regio isomer of the substrate, was studied but led to a complex pattern.<sup>48</sup> Nudelman *et al.* synthesized AOP analogs, which were not tested with the enzyme, but directly as herbicides with *A. thaliana*. Removing the methyl group of AOP, or replacing it by ethyl, isopropyl, or hydroxymethyl groups led to interesting activities.<sup>57</sup>

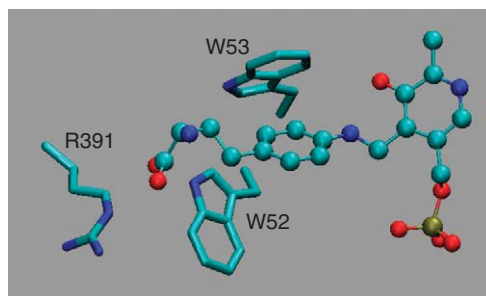
Another very interesting type of inhibition has been observed with amiclenomycin (ACM) (**12** in **Figure 13**). This natural product was isolated from cultures of different *Streptomyces* strains as a component of di- and tripeptides showing antibiotic properties.<sup>58–60</sup> These antibiotic properties were reversed by biotin and it was established that the target was DAPA aminotransferase.<sup>61</sup> Kinetic inhibition studies showed that ACM was probably recognized at the AOP binding site.<sup>51</sup> Inactivation was observed after preincubation but was found to be reversible, the activity being recovered after dialysis for 20 h. These interesting results justified a reexamination of the mechanism, which was undertaken in our group.

A total synthesis of ACM was necessary because the natural compound was no longer available, and because we speculated that the trans configuration proposed by Okami *et al.*<sup>59</sup> was ambiguous and needed confirmation. This synthetic approach led to the synthesis of both cis and trans isomers by unequivocal routes, with the conclusion that the natural compound was indeed the cis isomer.<sup>62</sup>

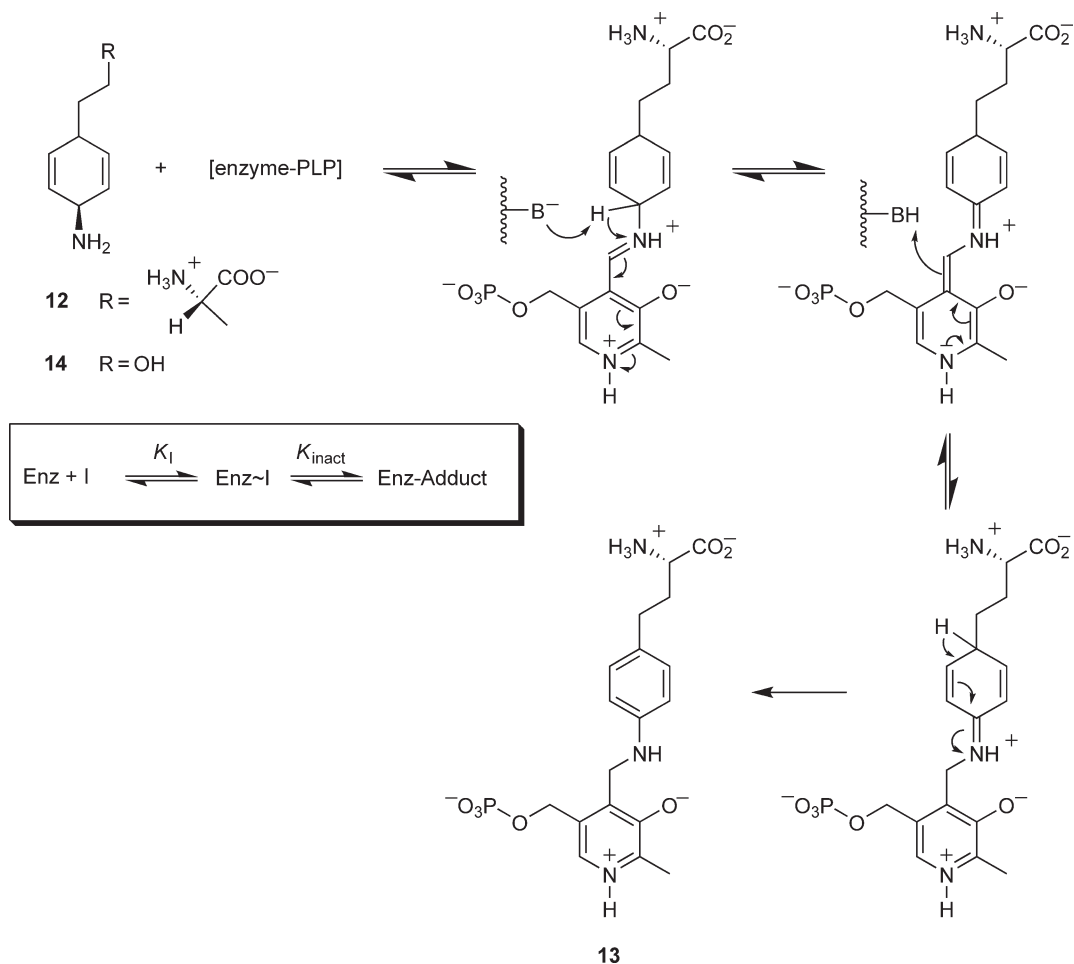
The inhibition studies were first performed with the *E. coli*, and later with the *M. tuberculosis* enzymes, since it was recently reported that *bioA*, the gene encoding for DAPA AT, was implicated in the long-term survival of mycobacteria. With the *E. coli* protein, a strong time-dependent inactivation was observed with the cis isomer, whereas the trans one was less potent. It was established that an aromatic PLP–inhibitor adduct was formed, the structure of which (**13**) was established by comparison with an authentic sample.<sup>45</sup> Although not covalently bound to the enzyme, its affinity makes the inactivation quasi-irreversible. The X-ray structure of the inactivated enzyme reveals that the aromatic moiety of **13** is located in the AOP binding site, constituted of aromatic residues<sup>63</sup> (**Figure 12**).

The proposed mechanism of inactivation is depicted in **Figure 13**. It is reminiscent of that proposed for the inactivation of  $\gamma$  aminobutyric acid transaminase by gabaculine<sup>64</sup> or cycloserine.<sup>65</sup>

The L- $\alpha$ -amino acid moiety of ACM is not absolutely required, since other aminocyclohexadienes with a D- $\alpha$ -amino acid or without functionality in the side chain are also efficient inhibitors. In all cases, the rate of inactivation is higher with the cis than with the trans isomers.<sup>45</sup> Very similar results were obtained with the *M. tuberculosis* enzyme that was inhibited by ACM and another analog (**14**) irreversibly, with the kinetic



**Figure 12** Binding of the amiclenomycin–PLP adduct to DAPA AT.<sup>63</sup>



**Figure 13** Mechanism of inactivation of DAPA AT by ampiclinomycin.<sup>45</sup>

**Table 4** Kinetic parameters for the inactivation of DAPA AT

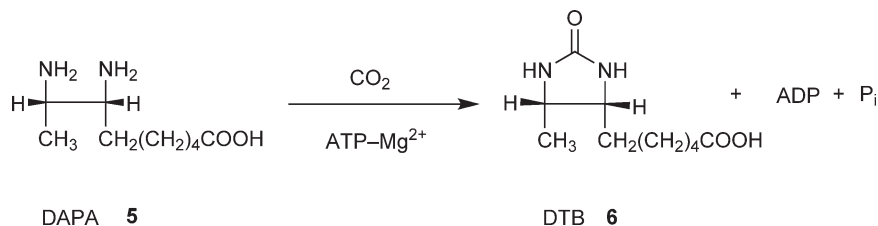
	Inhibitor	$K_I$ ( $\mu\text{mol l}^{-1}$ )	$k_{inact}$ ( $\text{min}^{-1}$ )
<i>Escherichia coli</i> <sup>45</sup>	12	2	0.4
<i>Mycobacterium tuberculosis</i> <sup>16</sup>	12	12	0.35
	14	20	0.56

parameters for inactivation given in **Table 4**. The observed partition ratio of about 1, which means that these compounds inactivate the enzyme every turnover, indicate that they can be very efficient inhibitors. They are also active *in vivo*.<sup>16</sup>

#### 7.06.2.1.4 Dethiobiotin synthetase

**7.06.2.1.4(i) Characterization of the enzyme** Important results on the reaction catalyzed by *E. coli* DTB synthetase (**Figure 14**) have been obtained by the group of Eisenberg, who established that the substrate for carboxylation was  $\text{CO}_2$  and not bicarbonate and that the reaction required 1 mol of ATP. They proposed on this basis a reasonable mechanism involving the formation of a carbamate, activated with ATP into a mixed

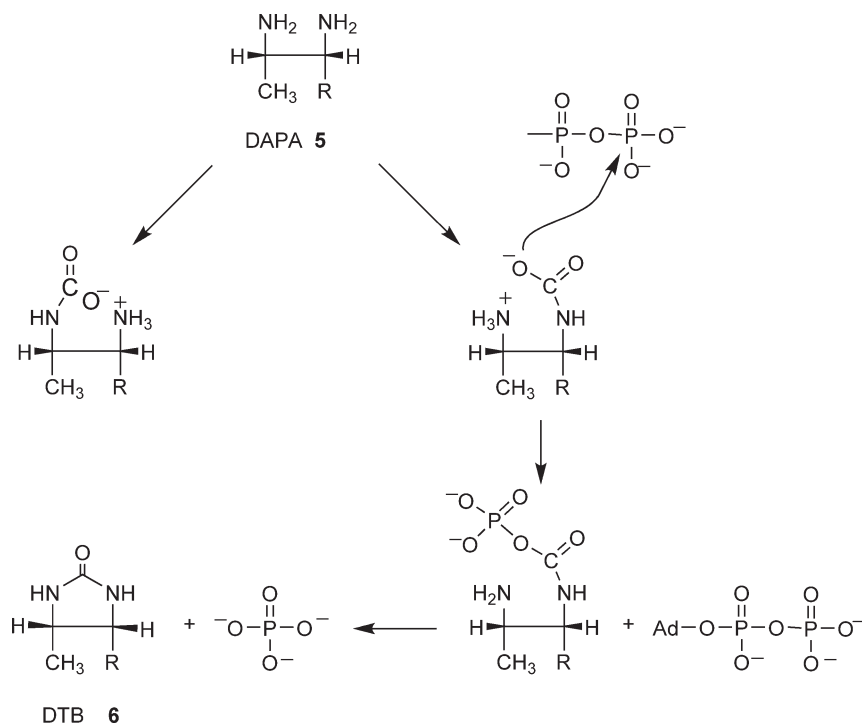




**Figure 14** The reaction catalyzed by DTB synthetase.

carbamic phosphoric anhydride, which cyclizes by nucleophilic attack of the neighboring amino group (**Figure 15**).

The recombinant *E. coli* enzyme has been purified and crystallized. Its 3D structure has been solved by two groups.<sup>66,67</sup> The enzyme is homodimeric and each subunit consists of a single globular  $\alpha/\beta$  domain. Although the typical motif of ATP and GTP binding proteins Gly-X-X-Gly-X-Gly-Lys-Thr/Ser is replaced in DTB synthetase by Gly-X-X-X-X-X-Gly-Lys-Thr, the overall topology of the central  $\beta$  sheets is similar. The ATP binding site has been inferred from crystallographic data of complexes with ADP or ATP analogs and divalent metal ions. DAPA binds at the interface and interacts with residues from both subunits. Detailed mechanistic studies performed by the two abovementioned groups have brought convincing experimental arguments in favor of the hypothesis of Eisenberg. The stoichiometry of the reaction was confirmed and experiments were carried out with  $\gamma$ -[<sup>18</sup>O<sub>3</sub>]- $\beta$ - $\gamma$ -[<sup>18</sup>O]ATP, the reaction products being analyzed by <sup>31</sup>P NMR.<sup>68</sup> [<sup>18</sup>O<sub>3</sub>]phosphate and  $\gamma$ -[<sup>18</sup>O]ADP were obtained proving a nucleophilic attack on the  $\gamma$  phosphate of ATP (attack on the  $\beta$  phosphate leading to an ADP mixed anhydride would lead to [<sup>18</sup>O<sub>4</sub>]phosphate and nonlabeled ADP) (**Figure 15**). A moderately stable intermediate with the properties expected for a mixed anhydride was later isolated by Gibson.<sup>69</sup>

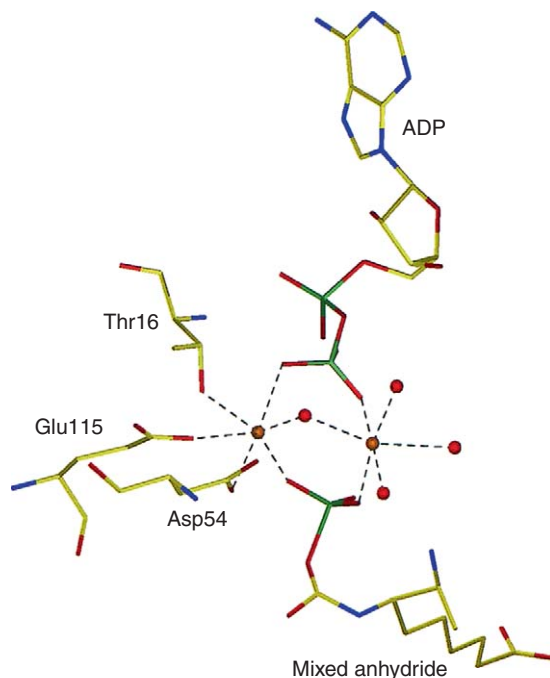


**Figure 15** Mechanism of dethiobiotin synthetase [R = (CH<sub>2</sub>)<sub>5</sub>COOH].<sup>68</sup>

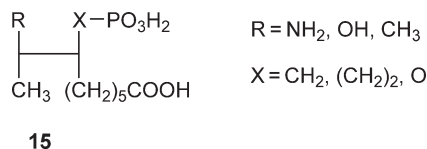
As shown in **Figure 15**, two pathways are of course susceptible to lead to the ureido ring, the carbamoylation taking place at N-7 or N-8. After trapping experiments, which were not entirely conclusive, a  $^{13}\text{C}$  NMR study established that carbamoylation takes place at N-7. When  $[7\text{-}^{15}\text{N}]\text{DAPA}$  and  $^{13}\text{CO}_2$  were incubated with an excess of enzyme, only the N-7 carbamate was observed, and pulse-chase experiments showed that this carbamate was kinetically competent.<sup>70</sup>

This result was fully confirmed by subsequent crystallographic studies.<sup>66,67</sup> Indeed, this reaction is one of the most precisely studied by X-ray with the structures of a large number of enzyme–substrates or enzyme–products complexes solved. The postulated intermediate, the N-7 carbamate, has been observed by soaking the crystals in a DAPA solution in the presence of  $\text{CO}_2$ , and it has been recognized in the structure of a ternary complex containing DAPA and an ATP analog in the presence of  $\text{Mn}^{2+}$  that one of the oxygens of the carbamate is well positioned for in-line attack of the  $\gamma$  phosphate of ATP, with inversion of configuration at the phosphorus atom.<sup>66</sup> A detailed mechanism has been proposed based on these structures, refined by the study of different mutants.<sup>71</sup> The unstable mixed anhydride has also been observed using kinetic crystallography, confirming the inversion of configuration at the phosphorous atom (**Figure 16**). Interestingly, besides the  $\text{Mg}^{2+}$  atom found in all ATP-using enzymes, a second  $\text{Mg}^{2+}$  is present in this structure, stabilizing the mixed anhydride.<sup>72</sup>

**7.06.2.1.4(ii) Search for inhibitors** Inhibitors of DTB synthetase were designed with the aim of obtaining novel herbicides. Alexeev *et al.*<sup>73</sup> synthesized a pyrimidine derivative designed to bind to the purine base of ATP subsite. X-ray crystallography proved that it was indeed the case. The compound was a competitive inhibitor, but with a too low affinity to be useful. Rendina *et al.*<sup>2</sup> have described the synthesis of 54 compounds, which can be classified into mimics of DAPA and mimics of DAPA carbamates. As no plant sequence has been reported, the structural and mechanistic data concerning the *E. coli* enzyme were used and the compounds were tested *in vitro* as substrates and inhibitors with the bacterial enzyme. It is beyond the scope of this review to analyze in detail this set of results. Some of the compounds are substrates but generally very poor ones, considering the relative values of  $V/K$  and they are also poor inhibitors. The best ones (**15**) (**Figure 17**), with  $K_I$  in the millimolar range at high ATP and micromolar range at high ATP, mimic the N-7 carbamates. However, none of these compounds showed significant herbicidal activity against *A. thaliana*.



**Figure 16** Structure of the complex of DTB synthetase with ADP and the mixed anhydride.<sup>72</sup>



**Figure 17** Inhibitors of DTB synthetase.

### 7.06.2.1.5 Biotin synthase

**7.06.2.1.5(i) Early studies** The conversion of DTB into biotin (**Figure 18**), which implies the formation of two carbon–sulfur bonds from two nonactivated carbons, has represented a very puzzling but fascinating mechanistic problem.<sup>8,74,75</sup> However, decisive progress has recently been made and, although several questions remain unanswered, the main features of the mechanism are now understood.

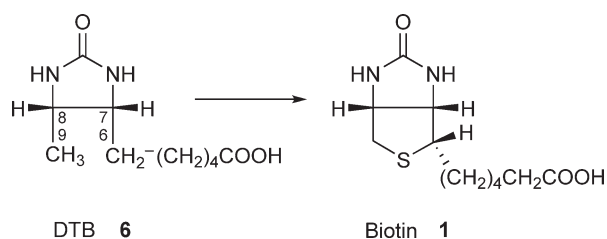
Early studies (described in detail by Marquet *et al.*<sup>8</sup>) consisted of *in vivo* bioconversion experiments, with labeled DTB, since all attempts to obtain *in vitro* activity had failed. The first experiments using regiospecifically labeled DTB samples were carried out independently by the groups of Parry, with *Aspergillus niger*, and Marquet, with *E. coli*. They were designed to check the two reasonable hypotheses (according to the mechanisms known at that time) for the activation of carbons 6 and 9, namely, introduction of a double bond or hydroxylation. Both groups showed that only two hydrogens of DTB were lost (one at C-9 and H<sub>5</sub> at C-6) ruling out the formation of double bonds. The possible intermediate alcohols 6(*R*)-, 6(*S*)-, and 9-hydroxydethiobiotins as well as 6(*R*), 9-dihydroxydethiobiotin were synthesized but they were not substrates of the reaction.

This led us to postulate a direct sulfur insertion. The mercapto compounds 6(*R*)-, 6(*S*)-, and 9-mercapto-dethiobiotins were also synthesized and tested for their ability to be converted into biotin *in vivo*. A very complex situation was observed with growing cells of *E. coli* or *B. sphaericus*, including desulfuration. However, it was found that 9-mercaptodethiobiotin (9-DTBSH) (**16** in **Figure 19**) was incorporated intact by the resting cells of *B. sphaericus*, <sup>34</sup>S- or <sup>35</sup>S-labeled 9-DTBSH gave biotin where 80% of the label was retained, whereas the two other thiols were not substrates.<sup>76</sup> We thus formulated the hypothesis that the first C–S bond was formed at C-9 and further results obtained *in vitro* fully confirmed this proposal.

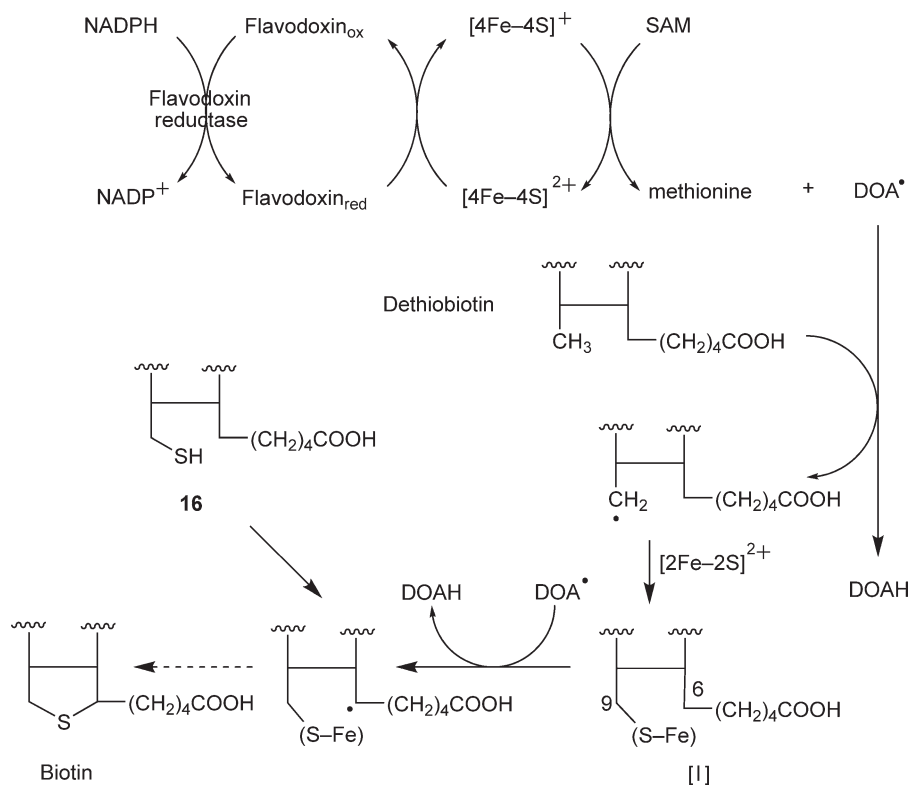
The mechanism of the formation of this C–S bond and the sulfur source remained of course completely unknown. The only experimental data concerning this mechanism were obtained by Marti,<sup>77</sup> who showed, starting with DTB bearing a chiral methyl group, that the reaction at C-9 occurred with racemization, a result consistent with a radical mechanism. The decisive breakthrough was done when it was shown that an active cell-free extract could be obtained by the addition of SAM (AdoMet or SAM)<sup>78</sup> followed by the observation that SAM was not the sulfur source.<sup>79</sup>

**7.06.2.1.5(ii) Characterization of the enzyme and mechanistic studies** Based on these results, dependence on SAM and the radical character of the reaction, it was postulated that BS could belong to the recently discovered ‘radical-SAM’ family of enzymes, and a mechanism could be proposed.<sup>80</sup> The presently accepted mechanism is shown in **Figure 19**, and we shall describe the most important experimental data that support it.

The *bioB* genes were sequenced, first in *E. coli*<sup>12</sup> and *B. sphaericus*,<sup>14</sup> and then in many other organisms. Comparison of the sequences revealed highly conserved motifs, a Cys triad CX3CX2C, an isolated C, and a



**Figure 18** The reaction catalyzed by biotin synthase.



**Figure 19** Biotin synthase mechanism.<sup>81</sup>

cluster YNHNL. The CX<sub>3</sub>CX<sub>2</sub>C motif has then been recognized as the signature of the ‘radical-SAM’ enzymes,<sup>82</sup> which are iron–sulfur proteins.

Biotin synthases from several organisms have been purified and characterized, the first one from *E. coli*.<sup>83</sup> They are indeed [Fe–S] proteins, isolated as homodimers containing one [2Fe–2S]<sup>2+</sup> center per monomer when purified aerobically. However, the elucidation of the nature of the [Fe–S] centers in the active enzyme has represented a very difficult problem, as will be discussed below.

For the reductive cleavage of SAM into methionine and the deoxyadenosyl radical (DOA<sup>•</sup>), an electron has to be transferred from a reduced [Fe–S] cluster. The reducing system consisting of NADPH, flavodoxine, and flavodoxine reductase has been identified in *E. coli*.<sup>84,85</sup> This allowed to set up an *in vitro* assay, which besides the obviously necessary components (BS, flavodoxin, flavodoxin reductase, DTB, SAM, and NADPH), requires also DTT, Fe<sup>2+</sup>, and a sulfide source. With such a system, the reaction was not catalytic, the turnover being at most 1. This bottleneck has since never been removed. The reducing system is still unknown in *B. sphaericus*, and in plants, adrenodoxin and adrenodoxin reductase<sup>86</sup> have been identified. The physiological electron transfer systems can be replaced *in vitro* by photoreduced deazaflavine.<sup>87</sup>

The DOA<sup>•</sup> then cleaves a C–H bond, here at C-9 of DTB, which implies that methionine and deoxyadenosine (DOAH) are formed in equimolar amounts. This is indeed the case.<sup>80</sup> The following question was the number of SAM equivalents (one or two) required for producing biotin. Although divergent results were reported, it is now accepted that two DOAs<sup>•</sup>, that is, two SAM molecules, one per C–H bond to be cleaved, are necessary,<sup>80,88,89</sup> a strong argument being that the cyclization of 9-DTBSH, which is also a substrate *in vitro* requires SAM. The observed DOAH/biotin ratio is often found greater than 2 and the excess has been attributed to the occurrence of an abortive process, DOA<sup>•</sup> being quenched by some component of the assay, very likely the protein.

The next important question was to know to which subfamily within the radical-SAM family BS belonged. When the SAM requirement for BS was discovered, only three similar enzymes were known. For

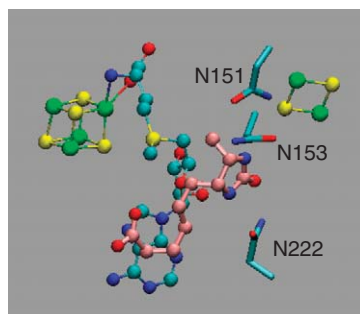
one of them, lysine 2,3-aminomutase (LAM), the  $\text{DOA}^\bullet$  reacts directly with the substrate to cleave a C–H bond, whereas for the other ones, pyruvate formate lyase (PFL) and anaerobic ribonucleotide reductase (aRNR), the enzyme is associated to an activase, where  $\text{DOA}^\bullet$  is generated;  $\text{DOA}^\bullet$  then generates a radical on the enzyme itself, which reacts with the substrate. Now a large number of radical-SAM enzymes have been identified and classified into these two subfamilies.<sup>75</sup> To answer this question in the case of BS, DTB labeled with deuterium at positions 6 and/or 9 was synthesized. The DOAH, which was isolated from the assays and analyzed by mass spectrometry, contained deuterium, thus proving a direct hydrogen transfer from the substrate to  $\text{DOA}^\bullet$ ,<sup>88</sup> a result consistent with the fact that no evidence for an activase was obtained. This deuterium transfer was however not complete, and no satisfying explanation could be proposed. It is now established that this is probably due to an isotope effect although previous studies had concluded that there was none.<sup>90</sup>

Thus the first part of the mechanism ending in the formation of the carbon radical at C-9 is rather well understood, but the sulfur source for the next step has raised long debates. Early studies showed that BS itself was the sulfur source and two hypotheses were considered: the sulfur could arise either from the [Fe–S] center or from a persulfide of unknown origin.<sup>87,91</sup> Reconstitution of the [Fe–S] center in the apoenzyme using  $^{34}\text{S}^{2-}$  or  $^{35}\text{S}^{2-}$  gave a protein with a labeled [Fe–S] center, which after purification produced labeled biotin in the *in vitro* assay.<sup>92,93</sup> This was a strong argument in favor of the first hypothesis.

At that time, the nature of the [Fe–S] clusters present in the active enzyme was unknown, and it has also been a very controversial point. After many independent studies, there is now more or less agreement between several groups, with the strong support from the X-ray structure,<sup>94</sup> that active BS contains two clusters, a  $[\text{4Fe–4S}]^{2+}$  common to all radical-SAM enzymes, which mediates the electron transfer, and a  $[\text{2Fe–2S}]^{2+}$  located at a different site.<sup>81,95–97</sup> It is reasonable to admit that the  $[\text{2Fe–2S}]^{2+}$  cluster is the sulfur source. The X-ray structure of the enzyme containing both clusters, SAM and DTB, revealed that the methyl group of DTB is properly oriented, at a compatible distance to this cluster (Figure 20).<sup>94</sup> Other arguments can be drawn from optical<sup>98</sup> and Mössbauer<sup>81,99</sup> spectroscopy studies, which have shown that during the formation of biotin, the amount of  $[\text{2Fe–2S}]^{2+}$  decreased, whereas the amount of  $[\text{4Fe–4S}]^{2+}$  remained constant. Recently, the selenium version of BS was investigated. When the  $[\text{2Fe–2S}]^{2+}$  enzyme was assayed in the presence of  $\text{Na}_2\text{Se}$ , instead of  $\text{Na}_2\text{S}$ , a mixture of biotin and selenobiotin was obtained, and it was observed by resonance Raman spectroscopy that an exchange had occurred during the reaction at the level of the cluster since a mixture of  $[\text{2Fe–2S}]^{2+}$  and  $[\text{2Fe–2Se}]^{2+}$  was formed. This exchange, which did not take place in the absence of reaction, reveals a perturbation of the cluster's structure during the catalytic event. This brings confirmation of the participation of the  $[\text{2Fe–2S}]$  center in the C–S bond formation.<sup>100</sup>

It is now possible to understand the requirements of the *in vitro* assay. The  $[\text{4Fe–4S}]^{2+}$  cluster, which is very oxygen sensitive, is lost during the purification and the aerobically isolated enzyme contains only the stable  $[\text{2Fe–2S}]^{2+}$  center. It is clear that the  $[\text{4Fe–4S}]^{2+}$  one has to be reconstituted for obtaining an active enzyme, by addition of iron and sulfide. In many protocols, cysteine has been added instead of sulfide. It is a source of sulfide due to the contamination of BS by a cysteine desulfurase.

However, the precise way by which the sulfur atom can be transferred from the cluster to the DTB radical is still not understood. One can postulate a direct reaction between the substrate radical and the  $[\text{2Fe–2S}]^{2+}$  center. On the other hand, Johnson and coworkers have proposed that the destruction of the cluster could



**Figure 20** Biotin synthase active site<sup>94</sup> (iron is green; sulfur is yellow; SAM carbons are blue; and DTB carbons are pink).

produce an intermediate persulfide on the enzyme, which would be the final sulfur donor reacting with the carbon radical.<sup>99</sup> This hypothesis cannot be completely excluded, but the *in vitro* assay is performed in a reductive medium, which should not favor the formation of a persulfide.

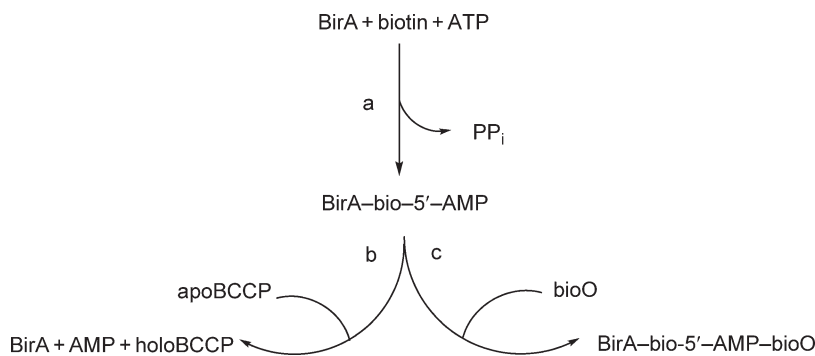
Thus, BS represents an interesting case of an enzyme that is also a substrate of the reaction. A simple explanation for the absence of turnover *in vitro* is that, once the [2Fe–2S] is destroyed by giving one sulfur atom, it remains bound to the protein, preventing the formation of a new cluster, the enzyme being inactivated after the catalytic event. Attempts to regenerate the cluster after turnover have been unsuccessful, only weak additional activity has been obtained.<sup>101</sup> When 9-DTBSH (16), which contains the sulfur atom of biotin was used as substrate, the same turnover of 1 was observed, with similar reaction rates. This experimental result does not invalidate the proposed mechanism, we assume that both substrates lead to biotin through the same intermediate, [I] in Figure 19.<sup>101</sup> This intermediate is probably cyclized without leaving the enzyme. Only recently has 16 been unambiguously characterized in the assay by mass spectrometry, at about 10% of the enzyme concentration. Interestingly, it is formed in stoichiometric amount with the Asn153Ser mutant of the enzyme. According to the X-ray structure, Asn153 is involved in the binding of the DTB substrate, and indeed it does not produce biotin and the intermediate can accumulate.<sup>89</sup>

The question of the *in vivo* situation was of course open. As most organisms require only tiny amounts of biotin, the hypothesis that BS could also be noncatalytic *in vivo* had to, and has been considered. It looks, however, more reasonable to expect a catalytic function. It is now well established that *in vivo* synthesis of [Fe–S] centers makes use of a very complex machinery, namely, the Isc or Suf systems in *E. coli*.<sup>102</sup> They include chaperone proteins, which may be necessary for repairing the cluster, or for the regeneration of a ‘native’ empty [2Fe–2S] site. A first answer to this puzzling question has been given in a recent paper that describes *in vivo* experiments.<sup>103</sup> The amount of biotin produced from DTB was determined by coexpression in *E. coli* of BS, biotin ligase, and a fragment of acetyl-CoA carboxylase to trap the biotin produced, followed by quantification of the biotinylated protein. A turnover of 20–60 equivalents of biotin has been observed, but a quantitative evaluation is difficult due to the fact that turnover renders the protein susceptible to proteolysis.

No inhibitors of BS have been developed. Actithiazic acid, a natural compound, remains the only one.<sup>104,105</sup>

### 7.06.2.2 Regulation of the Pathway

Numerous studies performed with *E. coli* have established that, in *E. coli*, biotin regulates very efficiently its biosynthetic pathway, with an absolute specificity, the biotin vitamers being inactive. As the topic has been largely reviewed,<sup>10,17,106</sup> it will be only summarized here. The regulation occurs at the transcriptional level and the biotin operon repressor (BirA) has been well characterized. This 33.5 kDa bifunctional protein is both an enzyme and a transcriptional regulator (Figure 21). It activates biotin into biotinyl-5'-AMP with ATP (reaction a) and transfers biotin on a specific lysine residue of the biotin accepting proteins (in *E. coli*, the biotin carboxyl carrier protein (BCCP), a subunit of acetyl-CoA carboxylase) (reaction b). When all the



**Figure 21** Schematic representation of the biotin regulatory system.<sup>108</sup>



apoprotein has been biotinylated, the noncovalent complex BirA-biotinyl-5'-AMP accumulates and binds to the *bio*-operator, *bioO* (reaction c), repressing transcription of the biotin biosynthetic genes. Thus, BirA synthesizes its own corepressor, a unique property among known DNA-binding proteins, as pointed out by Wilson *et al.*<sup>107</sup>

BirA has been completely purified and its 3D structure was established. The protein consists of three domains: the N-terminal domain containing the helix–turn–helix motif found in many DNA-binding proteins, is loosely connected to the rest of the protein, suggesting that important conformational changes could accompany DNA binding. On the other hand, the function of the carboxy-terminal domain is unknown. The central domain contains the active site, where biotin binds. A disordered loop including the consensus sequence (GXGXXG) associated with ATP binding is found close to the biotin site. This site is inferred from the crystal structure of a complex BirA–biotinyl lysine (biocytin) but no BirA–biotin or BirA–biotinyl-5'-AMP complex has been obtained. The crystallographic data favor a monomeric unit for apo BirA. Mutations that either affect the repressor function or the enzymatic activity are easily rationalized with the help of this 3D structure.<sup>107</sup>

Extensive solution studies have also been performed indicating that apo BirA is monomeric in solution. After binding of biotinyl-5'-adenylate, the complex dimerizes and binds cooperatively to the two operator half-sites.<sup>108</sup>

The regulation region of the *bio* operon has also been identified in *B. sphaericus*<sup>109</sup> and *B. subtilis*.<sup>21</sup>

## References

1. H. G. Wood; R. E. Barden, *Annu. Rev. Biochem.* **1977**, *46*, 385–413.
2. A. R. Rendina; W. S. Taylor; C. Gibson; G. Lorimer; D. Rayner; B. Lockett; K. Kranis; B. Wexler; D. Marcovici-Mizrahi; A. Nudelman; A. Nudelman; E. Masilii; H. Chi; Z. Wawrzak; J. Calabresse; W. Huang; J. Jia; G. Schneider; Y. Lindqvist; G. Yang, *Pest. Sci.* **1999**, *55*, 236–249.
3. F. Rebeillé; S. Ravanel; A. Marquet; R. R. Mendel; A. G. Smith; M. J. Warren, *Nat. Prod. Rep.* **2007**, *24*, 949–962.
4. J. Zempleni, *Annu. Rev. Nutr.* **2005**, *25*, 175–196.
5. J. E. Cronan, Jr., *J. Biol. Chem.* **2001**, *276*, 37355–37364.
6. P. Baldet; H. Gerbling; S. Axiotis; R. Douce, *Eur. J. Biochem.* **1993**, *217*, 479–485.
7. B. Wolf, *J. Nutr. Biochem.* **2005**, *16*, 441–445.
8. A. Marquet; B. Tse Sum Bui; D. Florentin, *Vitamins and Hormones*; Academic Press: New York, 2001; Vol. 61, pp 51–101.
9. W. R. Streit; P. Entcheva, *Appl. Microbiol. Biotechnol.* **2003**, *61*, 21–31.
10. M. A. Eisenberg, *Advances in Enzymology and Related Areas of Molecular Biology*; Interscience: New York, 1973; Vol. 38, pp 317–372.
11. C. Alban; D. Job; R. Douce, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2000**, *51*, 17–47.
12. A. J. Otsuka; M. R. Buoncristiani; P. K. Howard; J. Flamm; C. Johnson; R. Yamamoto; K. Uchida; C. Cook; J. Ruppert; J. Matsuzaki, *J. Biol. Chem.* **1988**, *263*, 19577–19585.
13. M. O'Regan; R. Gloeckler; S. Bernard; C. Ledoux; I. Ohsawa; Y. Lemoine, *Nucleic Acids Res.* **1989**, *17*, 8004.
14. R. Gloeckler; I. Ohsawa; D. Speck; C. Ledoux; S. Bernard; M. Zinsius; D. Villeval; T. Kisou; K. Kamogawa; Y. Lemoine, *Gene* **1990**, *87*, 63–70.
15. D. A. Patton; A. L. Schetter; L. H. Franzmann; K. Nelson; R. E. Ward; D. W. Meinke, *Plant Physiol.* **1998**, *116*, 935–946.
16. S. Mann; O. Ploux, *FEBS J.* **2006**, *273*, 4778–4789.
17. M. A. Eisenberg, *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*; American Society for Microbiology: Washington, DC, 1987; Vol. 2, pp 544–550.
18. Y. Izumi; H. Morita; Y. Tani; K. Ogata, *Agric. Biol. Chem.* **1974**, *38*, 2257–2262.
19. O. Ploux; P. Soularue; A. Marquet; R. Gloeckler; Y. Lemoine, *Biochem. J.* **1992**, *287*, 685–690.
20. A. Binieda; M. Fuhrmann; B. Lehner; C. Rey-Berthod; S. Frutiger-Hughes; G. Hughes; N. M. Shaw, *Biochem. J.* **1999**, *340*, 793–801.
21. S. Bower; J. B. Perkins; R. Rogers Yocum; C. L. Howitt; P. Rahaim; J. Pero, *J. Bacteriol.* **1996**, *178*, 4122–4130.
22. J. E. Stok; J. J. De Voss, *Arch. Biochem. Biophys.* **2000**, *384*, 351–360.
23. A. J. Green; S. L. Rivers; M. Cheesman; G. A. Reid; L. G. Quaroni; I. D. G. Macdonald; S. K. Chapman; A. W. Munro, *J. Biol. Inorg. Chem.* **2001**, *6*, 523–533.
24. R. J. Lawson; D. Leys; M. J. Sutcliffe; C. A. Kempf; M. R. Cheesman; S. J. Smith; J. Clarkson; W. E. Smith; I. Haq; J. B. Perkins; A. W. Munro, *Biochemistry* **2004**, *43*, 12410–12426.
25. A. J. Green; A. W. Munro; M. R. Cheesman; G. A. Reid; C. von Wachenfeldt; S. K. Chapman, *J. Inorg. Biochem.* **2003**, *93*, 92–99.
26. R. J. Lawson; C. von Wachenfeldt; I. Haq; J. Perkins; A. W. Munro, *Biochemistry* **2004**, *43*, 12390–12409.
27. M. J. Cryle; N. J. Matovic; J. J. De Voss, *Org. Lett.* **2003**, *5*, 3341–3344.
28. M. J. Cryle; J. J. De Voss, *Chem. Commun.* **2004**, 86–87.
29. Y. Lemoine; A. Wach; J. M. Jeltsch, *Mol. Microbiol.* **1996**, *19*, 645–647.
30. O. Ifuku; H. Miyaoka; N. Koga; J. Kishimoto; S. Haze; Y. Wachi; M. Kajiwara, *Eur. J. Biochem.* **1994**, *220*, 585–591.
31. I. Sanyal; S. L. Lee; D. H. Flint, *J. Am. Chem. Soc.* **1994**, *116*, 2637–2638.
32. N. H. Tomczyk; J. E. Nettleship; R. L. Baxter; H. J. Crichton; S. P. Webster; D. J. Campopiano, *FEBS Lett.* **2002**, *513*, 299–304.

33. R. Sanishvili; A. F. Yakunin; R. A. Laskowski; E. Evdokimova; A. Doberty-Kyrby; G. A. Lajoie; J. M. Thornton; C. H. Arrowsmith; A. Savchenko; A. Joachimiak; A. M. Edwards, *J. Biol. Chem.* **2003**, *278*, 26039–26045.
34. M. E. Webb; A. Marquet; R. R. Mendel; F. Rebeillé; A. G. Smith, *Nat. Prod. Rep.* **2007**, *24*, 988–1008.
35. Y. Izumi; H. Morita; Y. Tani; K. Ogata, *Agric. Biol. Chem.* **1973**, *37*, 1327–1333.
36. O. Ploux; A. Marquet, *Eur. J. Biochem.* **1996**, *236*, 301–308.
37. D. Alexeev; M. Alexeeva; R. L. Baxter; D. J. Campopiano; S. P. Webster; L. Sawyer, *J. Mol. Biol.* **1998**, *284*, 401–419.
38. V. M. Bohr; S. Dev; G. R. Vasanthakumar; P. Kumar; S. Sinha; A. Surolia, *J. Biol. Chem.* **2006**, *281*, 25076–25088.
39. T. Kubota; J. Shimono; C. Kanameda; Y. Izumi, *Biosci. Biotechnol. Biochem.* **2007**, *71*, 3033–3040.
40. V. Pinon; S. Ravel; R. Douce; C. Alban, *Plant Physiol.* **2005**, *139*, 1666–1676.
41. S. P. Webster; D. Alexeev; D. J. Campopiano; R. M. Watt; M. Alexeeva; L. Sawyer; R. L. Baxter, *Biochemistry* **2000**, *39*, 516–528.
42. O. Kerbach; D. J. Campopiano; R. L. Baxter, *Chem. Commun.* **2006**, 60–62.
43. S. Spinelli; O. Ploux; A. Marquet; C. Anguille; C. Jelsch; C. Cambillau; C. Martinez, *Acta Crystallogr. D* **1996**, *52*, 866–868.
44. O. Ploux; O. Breyne; S. Carillon; A. Marquet, *Eur. J. Biochem.* **1999**, *259*, 63–70.
45. S. Mann; D. Florentin; D. Lesage; T. Drujon; O. Ploux; A. Marquet, *Helv. Chim. Acta* **2003**, *86*, 3836–3850.
46. D. Alexeev; R. L. Baxter; D. J. Campopiano; O. Kerbach; L. Sawyer; N. Tomczyk; S. P. Webster, *Org. Biomol. Chem.* **2006**, *4*, 1209–1212.
47. C. H. Pai, *J. Bacteriol.* **1971**, *105*, 793–800.
48. G. L. Stoner; M. A. Eisenberg, *J. Biol. Chem.* **1975**, *250*, 4037–4043.
49. A. C. Eliot; J. Sandmark; G. Schneider; J. F. Kirsch, *Biochemistry* **2002**, *41*, 12582–12589.
50. Y. Izumi; K. Sato; Y. Tani; K. Ogata, *Agric. Biol. Chem.* **1973**, *37*, 2683–2684.
51. K. Hotta; T. Kitahara; O. Okami, *J. Antibiot.* **1975**, *28*, 222–228.
52. R. S. Breen; D. J. Campopiano; S. Webster; M. Brunton; R. Watt; R. L. Baxter, *Org. Biomol. Chem.* **2003**, *1*, 3498–3499.
53. S. W. Van Aesdell; J. B. Perkins; R. R. Yocum; L. Luan; C. L. Howitt; N. P. Chatterjee; J. G. Pero, *Biotechnol. Bioeng.* **2005**, *91*, 75–83.
54. A. Nudelman; A. Nudelman; D. Marcovici-Mizrahi; D. Flint, *Bioorg. Chem.* **1998**, *26*, 157–168.
55. H. Käck; J. Sandmark; K. Gibson; G. Schneider; Y. Lindqvist, *J. Mol. Biol.* **1999**, *291*, 857–876.
56. J. Sandmark; A. C. Eliot; K. Famm; G. Schneider; J. F. Kirsch, *Biochemistry* **2004**, *43*, 1213–1222.
57. A. Nudelman; D. Marcovici-Mizrahi; A. Nudelman; D. Flint; V. Wittenbach, *Tetrahedron* **2004**, *60*, 1731–1748.
58. K. H. Baggaley; B. Blessington; C. P. Falshaw; W. D. Ollis, *J. Chem. Soc. Chem. Commun.* **1969**, 101–102.
59. Y. Okami; T. Kitahara; M. Hamada; K. Naganawa; S. Kondo; K. Maeda; T. Takenchi; H. Umezawa, *J. Antibiot.* **1974**, *27*, 656–664.
60. A. Kern; V. Kabatek; G. Jung; R. G. Werner; M. Poetsch; H. Zähler, *Liebigs Ann. Chem.* **1985**, 877–892.
61. T. Kitahara; K. Hotta; M. Yoshida; J. Okami, *J. Antibiot.* **1975**, *28*, 215–220.
62. S. Mann; S. Carillon; O. Breyne; A. Marquet, *Chem. Eur. J.* **2002**, *8*, 439–450.
63. J. Sandmark; S. Mann; A. Marquet; G. Schneider, *J. Biol. Chem.* **2002**, *277*, 43352–43358.
64. R. R. Rando, *Biochemistry* **1977**, *16*, 4605–4610.
65. G. T. Olson; M. Fu; S. Lau; K. L. Rinehart; R. B. Silverman, *J. Am. Chem. Soc.* **1998**, *120*, 2256–2267.
66. W. Huang; J. Jia; K. J. Gibson; W. S. Taylor; A. R. Rendina; G. Schneider; Y. Lindqvist, *Biochemistry* **1995**, *34*, 10985–10995.
67. D. Alexeev; R. L. Baxter; O. Smekal; L. Sawyer, *Structure* **1995**, *3*, 1207–1215.
68. R. L. Baxter; H. C. Baxter, *J. Chem. Soc. Chem. Commun.* **1994**, 759–760.
69. K. J. Gibson, *Biochemistry* **1997**, *36*, 8474–8478.
70. K. J. Gibson; G. H. Lorimer; A. R. Rendina; W. S. Taylor; G. Cohen; A. A. Gatenby; W. G. Payne; D. C. Roe; B. A. Lockett; A. Nudelman; D. Marcovici; A. Nachum; B. A. Wexler; E. L. Marsilli; I. M. Turner, Sr.; L. D. Howe; C. E. Kalbach; H. Chi, *Biochemistry* **1995**, *34*, 10976–10984.
71. G. Yang; T. Sandalova; K. Lohman; Y. Lindqvist; A. R. Rendina, *Biochemistry* **1997**, *36*, 4751–4760.
72. H. Käck; K. J. Gibson; Y. Lindqvist; G. Schneider, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5495–5500.
73. D. Alexeev; R. L. Baxter; D. J. Campopiano; R. S. McAlpine; L. McIver; L. Sawyer, *Tetrahedron* **1998**, *54*, 15891–15898.
74. A. Marquet, *Curr. Opin. Chem. Biol.* **2001**, *5*, 541–549.
75. A. Marquet; B. Tse Sum Bui; A. G. Smith; M. J. Warren, *Nat. Prod. Rep.* **2007**, *24*, 1027–1040.
76. A. Marquet; F. Frappier; G. Guillerme; M. Azoulay; D. Florentin; J. C. Tabet, *J. Am. Chem. Soc.* **1993**, *115*, 2139–2145.
77. F. B. Marti, Zur Biosynthese Von Biotin: Einsatz Von Vorläufern Mit Chiraler Methyl Gruppe. Ph.D. Thesis, ETH, Zürich, 1983; No 7236.
78. O. Ifuku; J. Kishimoto; S. Haze; M. Yanagi; S. Fukushima, *Biosci. Biotechnol. Biochem.* **1992**, *56*, 1780–1785.
79. D. Florentin; B. Tse Sum Bui; A. Marquet; T. Ohshiro; Y. Izumi, *C. R. Acad. Sci. Paris, Life Sci.* **1994**, *317*, 485–488.
80. D. Guianvarc'h; D. Florentin; B. Tse Sum Bui; F. Nunzi; A. Marquet, *Biochem. Biophys. Res. Commun.* **1997**, *236*, 402–406.
81. B. Tse Sum Bui; R. Benda; V. Schünemann; D. Florentin; A. X. Trautwein; A. Marquet, *Biochemistry* **2003**, *42*, 8791–8798.
82. H. J. Sofia; G. Chen; B. G. Hetzler; J. F. Reyes-Spindola; N. E. Miller, *Nucleic Acids Res.* **2001**, *29*, 1097–1106.
83. I. Sanyal; G. Cohen; D. H. Flint, *Biochemistry* **1994**, *33*, 3625–3631.
84. O. Ifuku; N. Koga; S. Haze; J. Kishimoto; Y. Wachi, *Eur. J. Biochem.* **1994**, *224*, 173–178.
85. O. M. Birch; M. Fuhrmann; N. M. Shaw, *J. Biol. Chem.* **1995**, *270*, 19158–19165.
86. A. Picciocchi; R. Douce; C. Alban, *J. Biol. Chem.* **2003**, *278*, 24966–24975.
87. A. Méjean; B. Tse Sum Bui; D. Florentin; O. Ploux; Y. Izumi; A. Marquet, *Biochem. Biophys. Res. Commun.* **1995**, *217*, 1231–1237.
88. F. Escalettes; D. Florentin; B. Tse Sum Bui; D. Lesage; A. Marquet, *J. Am. Chem. Soc.* **1999**, *121*, 3571–3578.
89. A. M. Taylor; C. E. Farrar; J. T. Jarrett, *Biochemistry* **2008**, *47*, 9309–9317.
90. D. Florentin, unpublished results.
91. K. J. Gibson; D. A. Pelletier; I. M. Turner, Sr., *Biochem. Biophys. Res. Commun.* **1999**, *254*, 632–635.
92. B. Tse Sum Bui; D. Florentin; F. Fournier; O. Ploux; A. Méjean; A. Marquet, *FEBS Lett.* **1998**, *440*, 226–230.
93. B. Tse Sum Bui; F. Escalettes; G. Chottard; D. Florentin; A. Marquet, *Eur. J. Biochem.* **2000**, *267*, 2688–2694.
94. F. Berkovitch; Y. Nicolet; J. T. Wan; J. T. Jarrett; C. L. Drennan, *Science* **2004**, *303*, 76–79.

95. N. B. Ugulava; B. R. Gibney; J. T. Jarrett, *Biochemistry* **2001**, *40*, 8343–8351.
96. N. B. Ugulava; K. K. Surerus; J. T. Jarrett, *J. Am. Chem. Soc.* **2002**, *124*, 9050–9051.
97. M. M. Cosper; G. N. L. Jameson; H. L. Hernandez; C. Krebs; B. H. Huynh; M. K. Johnson, *Biochemistry* **2004**, *43*, 2007–2021.
98. N. B. Ugulava; C. J. Sacanell; J. T. Jarrett, *Biochemistry* **2001**, *40*, 8352–8358.
99. G. N. L. Jameson; M. M. Cosper; H. L. Hernandez; M. K. Johnson; B. H. Huynh, *Biochemistry* **2004**, *43*, 2022–2031.
100. B. Tse Sum Bui; T. A. Mattioli; D. Florentin; G. Bolbach; A. Marquet, *Biochemistry* **2006**, *45*, 3824–3834.
101. B. Tse Sum Bui; M. Lotierzo; F. Escalettes; D. Florentin; A. Marquet, *Biochemistry* **2004**, *43*, 16432–16441.
102. D. C. Johnson; D. R. Dean; A. D. Smith; M. K. Johnson, *Annu. Rev. Biochem.* **2005**, *74*, 247–281.
103. E. Choi-Rhee; J. E. Cronan, *Chem. Biol.* **2005**, *12*, 461–468.
104. K. Ogata; Y. Izumi; Y. Tani, *Agric. Biol. Chem.* **1970**, *34*, 1872.
105. M. A. Eisenberg; S. C. Hsiung, *Antimicrob. Agents Chemother.* **1982**, *21*, 5.
106. A. Chapman-Smith; J. E. Cronan, Jr., *Trends Biochem. Sci.* **1999**, *24*, 359–363.
107. K. P. Wilson; M. L. Schewchuk; R. G. Brennan; A. J. Otsuka; B. W. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9257–9261.
108. E. Eisenstein; D. Beckett, *Biochemistry* **1999**, *38*, 13077–13084.
109. D. Speck; I. Ohsawa; R. Gloeckler; M. Zinsius; S. Bernard; C. Ledoux; T. Kisou; T. Kamogawa; Y. Lemoine, *Gene* **1991**, *108*, 39–45.

### Biographical Sketch



Andrée Marquet is presently a professor emeritus at the Université Pierre et Marie Curie (Paris VI). She completed her Ph.D. under Jean Jacques at the Collège de France in Paris, studying reaction mechanisms, and then did her postdoctoral work with Duilio Arigoni, at ETH Zürich, on terpene biosynthesis. After a career as a research director at the Centre National de la Recherche Scientifique, she was appointed as professor of chemistry at Paris VI, where she taught bioorganic chemistry.

For several years, she studied chemical reaction mechanisms, especially those involving sulfoxide carbanions and their synthetic applications. Then, she turned to mechanistic enzymology. Her main interests concerned steroid isomerases and cytochrome P-450, vitamin K-dependent carboxylations, and biotin biosynthesis. She contributed to the mechanistic understanding of several enzymes of the pathway, namely, amino-oxopelargonate synthase, diaminopelargonate aminotransferase, and more importantly biotin synthase.

## 7.07 Lipoic Acid Biosynthesis and Enzymology

Elizabeth S. Billgren, Robert M. Cicchillo, Natasha M. Nesbitt, and Squire J. Booker,  
The Pennsylvania State University, University Park, PA, USA

© 2010 Elsevier Ltd. All rights reserved.

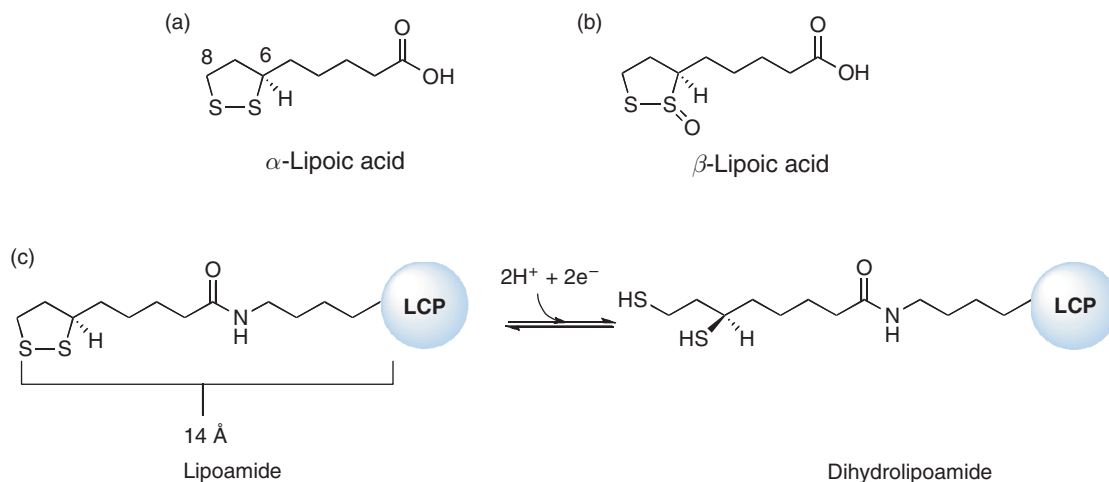
7.07.1	Introduction	181
7.07.1.1	Properties of Lipoic Acid	181
7.07.1.2	Properties of the Lipoyl Cofactor (Lipoamide)	183
7.07.1.3	Discovery of Lipoic Acid	183
7.07.2	The Biological Functions of Lipoamide	183
7.07.2.1	Discovery of Lipoamide	183
7.07.2.2	The $\alpha$ -Keto Acid Dehydrogenase Complexes	184
7.07.2.3	The Glycine Cleavage System	188
7.07.2.4	Lipoamide Dehydrogenase	189
7.07.3	Pathways for Construction of the Lipoyl Cofactor	190
7.07.4	Characterization of Enzymes Involved in Lipoic Acid Biosynthesis and Metabolism	193
7.07.4.1	Lipoate Protein Ligase A	193
7.07.4.2	Lipoate-Activating Enzyme	195
7.07.4.3	Lipoyltransferase	196
7.07.4.4	Octanoyl-[Acyl Carrier Protein]-Protein Transferase	196
7.07.4.5	Lipoyl Synthase	198
7.07.4.5.1	Metabolic feeding studies	198
7.07.4.5.2	Isolation and characterization	200
7.07.4.5.3	Mechanistic studies of LipA	203
7.07.4.6	Lipoamidase	206
References		208

### 7.07.1 Introduction

#### 7.07.1.1 Properties of Lipoic Acid

Lipoic acid (1,2-dithiolane-3-pentanoic acid or 6,8 thioctic acid) is an essential biomolecule that is present in organisms throughout all kingdoms of life.<sup>1,2</sup> It is an eight-carbon-containing straight-chain fatty acid with sulfur atoms inserted at C6 and C8, a configuration that allows it to form an intramolecular disulfide bond in its biologically relevant oxidized form (Figure 1).<sup>3</sup> Its most notable role in biology is as a central redox-related cofactor used by several multienzyme complexes that are involved in primary metabolism, and the reversible formation of this five-membered cyclic disulfide, termed a 1,2-dithiolane ring, is key to its cofactor-related activity.<sup>4</sup> The dithiolane ring is also responsible for the compound's pale yellow color and UV-visible (UV-vis) absorption maximum at 333 nm ( $\epsilon = 150 \text{ mol}^{-1} \text{ l cm}^{-1}$ ).<sup>5,6</sup> Lipoic acid contains one chiral center, at C6, and it is the *R* enantiomer that is produced naturally.<sup>7</sup> The appellation  $\alpha$ -lipoic acid is sometimes used to distinguish the compound from  $\beta$ -lipoic acid, which is the two-electron oxidized thiosulfinate derivative of lipoic acid (Figure 1).  $\beta$ -Lipoic acid is also found in nature, however less frequently, and has considerably less biological activity than  $\alpha$ -lipoic acid.<sup>8,9</sup> A detailed overview of the structure, intrinsic reactivity, and chemical syntheses of lipoic acid has been published,<sup>10</sup> and readers are encouraged to refer to this work for additional insight into the general properties of the molecule.

In addition to its role as a cofactor, lipoic acid is also gaining notoriety as a broad-spectrum antioxidant and regulator of certain cellular processes.<sup>11</sup> The reversible midpoint potential for its two-electron reduction to dihydrolipoic acid (DHLA) has been measured to be  $-0.325 \text{ V}$  at pH 7 and  $25^\circ\text{C}$ ,<sup>12</sup> and  $-0.294 \text{ V}$  at pH 6.8 and  $22^\circ\text{C}$ .<sup>13</sup> This redox potential is lower than that of other common cellular reducing agents such as the



**Figure 1** Structures of (a)  $\alpha$ -lipoic acid, (b)  $\beta$ -lipoic acid, and (c) lipoamide in its oxidized (*left*) and reduced (*right*) forms. The two-electron reduced form of lipoamide is designated dihydrolipoamide.

glutathione redox couple ( $-0.25$  V, pH 6.8, and  $30^\circ\text{C}$ )<sup>14</sup> and the thioredoxin redox couple ( $-0.27$  V, pH 7.0, and  $25^\circ\text{C}$ ),<sup>15,16</sup> and on par with the nicotinamide adenine dinucleotide ( $\text{NAD}^+$ )/ $\text{NADH}$  and  $\text{NADP}^+$ / $\text{NADPH}$  redox couples ( $-0.318$  and  $-0.317$  V, respectively).<sup>5</sup> Lipoic acid is also believed to play a role in the recycling of other antioxidants, such as glutathione and vitamins E (tocopherol) and C (ascorbic acid).<sup>17–19</sup> Its versatility as a broad-spectrum antioxidant is underscored by its water and lipid solubility, and its ability to cross the blood–brain barrier.<sup>20</sup> Although DHLA has the better reducing ability,  $\alpha$ -lipoic acid has been shown to scavenge a number of reactive oxygen species (ROS), including singlet oxygen – which is not scavenged by DHLA – and is a potent scavenger of hypochlorous acid.<sup>21,22</sup>

Lipoic acid has been indicated in the amelioration of the complications of a variety of neurological disorders, including diabetic neuropathy<sup>23,24</sup> and Alzheimer's disease,<sup>25</sup> because of its antioxidant properties. It was first introduced as a treatment for a number of illnesses – such as liver cirrhosis, mushroom poisoning, heavy metal intoxication, and diabetic polyneuropathy – in the late 1960s when a German physician discovered that individuals afflicted with such ailments had lower levels of lipoic acid.<sup>10</sup> The underlying pathophysiological factor associated with these ailments is oxidative stress, which originates from improper control of oxygen reduction and/or attenuated clearance of ROS.<sup>26,27</sup> The efficacy of lipoic acid in the treatment of type II diabetes appears to result from its ability to increase glucose uptake into skeletal muscle by stimulating the translocation of the GLUT4 receptor from the cytosol to the plasma membrane – a process that is regulated by insulin in healthy cells – as well as regulating its intrinsic activity.<sup>28</sup>

Lipoic acid has also been shown to regulate the activity of several cellular signaling pathways.<sup>11,29</sup> One of the best characterized is its effect on  $\text{NF-}\kappa\text{B}$ , a transcription factor that is triggered as part of the mammalian immune or inflammatory response, or as a result of stress from ROS.<sup>30</sup> Upon activation,  $\text{NF-}\kappa\text{B}$  translocates from the cytoplasm into the nucleus where it binds to specific sites in DNA and regulates transcription of genes that are typically involved in defense against pathogens, but which also play roles in inflammation and the pathogenesis of atherosclerosis.<sup>30,31</sup> It is known that the redox status of the cell also modulates  $\text{NF-}\kappa\text{B}$  activity,<sup>32</sup> and lipoic acid can attenuate its activation, most likely through reduction of a disulfide bond on the protein, similarly to the established effect of the redox protein thioredoxin.<sup>33</sup> This redox regulation of  $\text{NF-}\kappa\text{B}$  activity, as well as oxidative stress in general, is also important in infection by human immunodeficiency virus (HIV) and the resulting pathophysiology.<sup>34</sup> HIV-1 proviral DNA contains two binding sites for  $\text{NF-}\kappa\text{B}$ . Upon activation of the transcription factor and translocation to the nucleus, its binding to integrated HIV DNA initiates transcription of the HIV genome.<sup>35,36</sup> Studies have shown that  $\alpha$ -lipoic acid is a potent inhibitor of  $\text{NF-}\kappa\text{B}$  in human T cells, requiring a concentration of  $4\text{ mmol l}^{-1}$  for complete inhibition,<sup>37</sup> and can also inhibit HIV-1 replication *in vivo*.<sup>38</sup>



### 7.07.1.2 Properties of the Lipoyl Cofactor (Lipoamide)

The most established role for lipoic acid is as a central cofactor in several multienzyme complexes that are involved in energy metabolism and the catabolism of certain amino acids. These complexes include the pyruvate dehydrogenase complex (PDC), the  $\alpha$ -ketoglutarate dehydrogenase complex (KDC), the branched-chain  $\alpha$ -keto acid dehydrogenase complex (BCKDC), the acetoin dehydrogenase complex (ADC), and the glycine cleavage system (GCS).<sup>2,4,39–41</sup> In its capacity as a cofactor, lipoic acid is attached to the  $\epsilon$ -amino group in an amide linkage to a specific lysyl residue that is housed on a designated lipoyl carrier protein (LCP) (Figure 1). This linkage creates a 14 Å appendage that can access multiple active sites within these huge complexes to facilitate the direct transfer of intermediates during turnover. The key functional property of this prosthetic group is its ability to undergo redox chemistry, interchanging between the cyclic disulfide (lipoamide) and the reduced dithiol (dihydrolipoamide) forms. The disulfide form of the cofactor acts as an acceptor in the transfer of acetyl, succinyl, or methylamine groups from appropriate substrates, becoming two-electron reduced in the process.<sup>2,4,41</sup> The reduced form must then be reoxidized to lipoamide in order for subsequent cycles of catalysis to occur. This oxidation is catalyzed by the flavin-containing enzyme lipoamide dehydrogenase (LipDH), with transfer of electrons ultimately to  $\text{NAD}^+$  to produce NADH. In eukaryotes, the PDC, KDC, BCKDC, and GCS are found in the mitochondrion,<sup>4,41,42</sup> as are the enzymes involved in the biosynthesis and attachment of the cofactor to LCPs (*vide infra*). As will be detailed below, the ADC is found only in certain species of bacteria.

### 7.07.1.3 Discovery of Lipoic Acid

Lipoic acid was first identified in the late 1940s in an attempt to recognize factors that replaced the requirement for acetate in the growth medium of certain lactic acid bacteria.<sup>43</sup> In early studies, the cofactor had been deemed ‘acetate-replacing factor,’ because it aided bacterial growth of organisms such as *Lactobacillus acidophilus*, which could not survive under conditions devoid of acetate.<sup>44</sup> Other laboratories had demonstrated the necessity of lipoic acid for the oxidation of pyruvate to acetate and carbon dioxide, which led to another trivial name, ‘pyruvate oxidation factor.’<sup>45</sup> The combined efforts of Lester Reed and the Eli Lilly industrial group led to the first isolation of crystalline acetate-replacing/pyruvate oxidation factor in the early 1950s.<sup>3</sup> Distinct physical characteristics, such as solubility in hydrophobic solvents, acidity, and involvement in the formation of acetate through the oxidative decarboxylation of pyruvate, led to the naming of the compound as ‘ $\alpha$ -lipoic acid’.<sup>46</sup> Molecular and structural characterization of the yellow crystals followed. Initial examination of the cofactor produced positive results for sulfur, whereas a nitroprusside assay for free thiols was negative. Treatment of the compound with sodium cyanide, an agent that can reduce disulfide bonds, resulted in a positive nitroprusside test, indicating that  $\alpha$ -lipoic acid forms a disulfide bond.<sup>3,47</sup> The molecular formula,  $\text{C}_8\text{H}_{14}\text{O}_2\text{S}_2$ , was determined by elemental analysis and electrochemical titration experiments. A titratable group having a  $\text{p}K_a$  of 4.76 was also observed, corresponding to an aliphatic carboxyl group separated from any unsaturation or polar groups by at least two methylene groups. X-ray data along with infrared spectroscopy and polarographic studies revealed that lipoic acid behaved similar to an aliphatic chain containing a cyclic disulfide. Treatment of lipoic acid with Raney nickel, a desulfurating metal, yielded *n*-caprylic acid (octanoic acid). These characteristics of the compound were consistent with a cyclic disulfide that was derived from 4,8-, 5,8-, or 6,8-dimercapto-*n*-caprylic acid.<sup>3</sup> The chemical nature of the biologically active cofactor was confirmed by comparisons to a chemically synthesized thiocetic acid racemate.<sup>6,48</sup>  $\alpha$ -Lipoic acid has one chiral center, located at carbon six. Synthesis of the optical isomers, ( $\pm$ )- $\alpha$ -lipoic acid, was first achieved at Merck and facilitated the determination of the absolute configuration at this carbon by Mislow in 1956.<sup>7,49</sup>

## 7.07.2 The Biological Functions of Lipoamide

### 7.07.2.1 Discovery of Lipoamide

The physiological role of lipoic acid was first determined to be its involvement in the acetyl-coenzyme A (acetyl-CoA) and NAD-dependent oxidative decarboxylation of pyruvate and  $\alpha$ -keto acids. Lipoic acid was shown to be essential for the oxidation of pyruvate,  $\alpha$ -ketobutyrate,  $\beta$ -methyl- $\alpha$ -ketobutyrate, and



$\beta$ -methyl- $\alpha$ -ketovalerate by cells of *Streptococcus faecalis* – now known as *Enterococcus faecalis* – harvested from lipoic acid-deficient medium.<sup>43,50</sup> Enzyme systems capable of catalyzing these decarboxylation reactions were shown to contain significant amounts of protein-bound lipoic acid.<sup>13,50</sup> Release of bound lipoic acid from these proteins by lipoamidase (Lpa), the enzyme that hydrolyzes the amide bond between lipoic acid and the  $\epsilon$ -amino group of the target lysine group,<sup>51,52</sup> resulted in the loss of decarboxylation activity.<sup>51,53,54</sup>

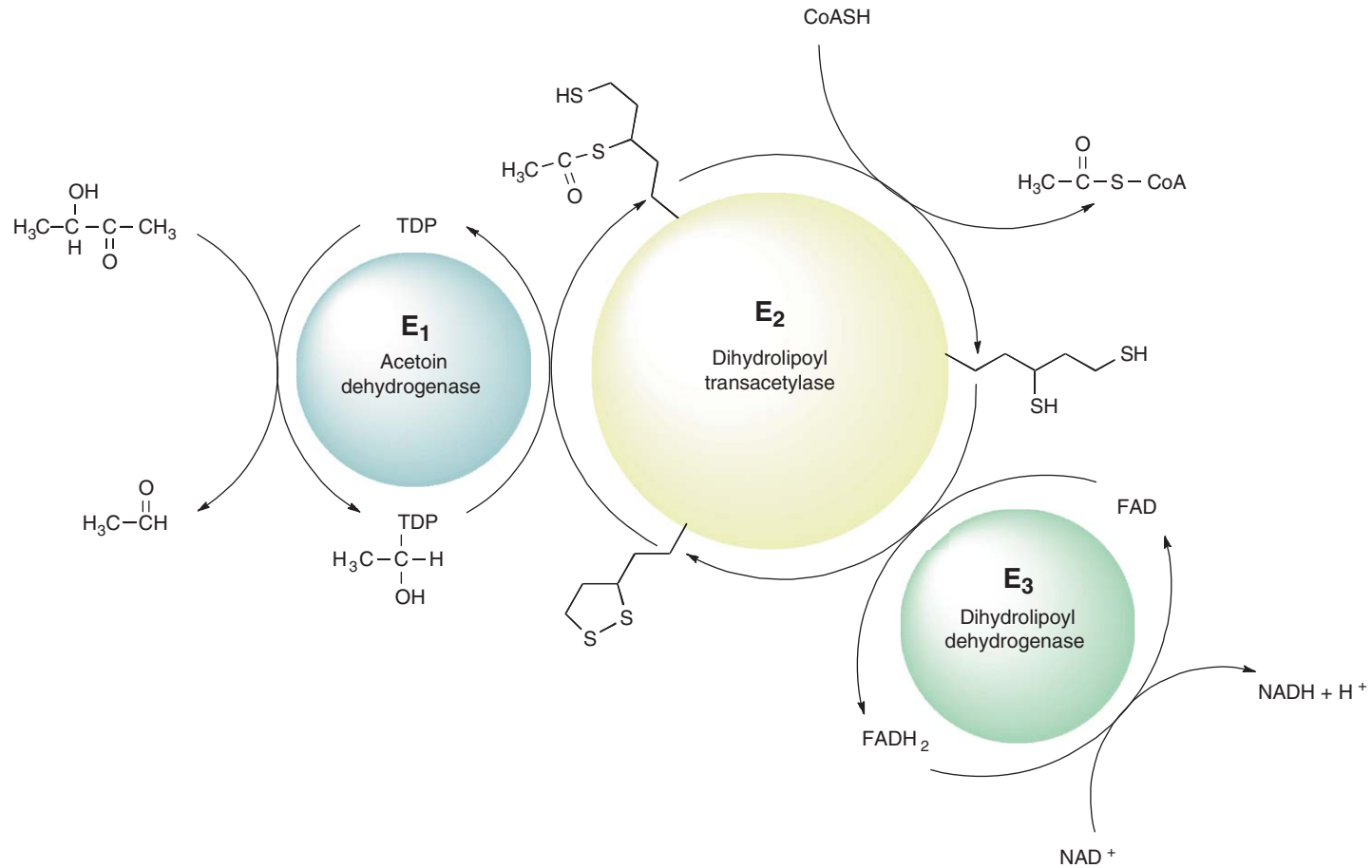
Because the presence of lipoic acid was required for these decarboxylation reactions, its association with these enzyme systems was investigated. The participation of lipoic acid in the pyruvate dehydrogenase and  $\alpha$ -KDCs was resolved in experiments in which *Escherichia coli* was cultured aerobically in the presence of <sup>35</sup>S-lipoic acid, allowing for incorporation of the radioactive compound into the complexes. Upon oxidation of the protein with performic acid and partial hydrolysis with hydrochloric acid, the hydrolysates contained a radioactive conjugate that was identified as  $N^{\epsilon}$ -(6,8-disulfooctanoyl)-L-lysine by comparing its migratory properties with that of authentic  $N^{\epsilon}$ -(6,8-disulfooctanoyl)-L-lysine using paper electrophoresis and paper chromatography with ninhydrin staining.<sup>55</sup> Acid hydrolysis of the P<sub>2</sub> protein – later referred to as the H-protein – of the GCS using 3 N *p*-toluenesulfonic acid along with the chemical reactions by thiol-modifying agents showed that this cofactor was also fundamental for the activity of the GCS.<sup>56</sup>

### 7.07.2.2 The $\alpha$ -Keto Acid Dehydrogenase Complexes

The role of lipoic acid in the pyruvate dehydrogenase and  $\alpha$ -KDCs has been firmly established.<sup>53,54,57,58</sup> These large (~5–14 MDa) multienzyme complexes contain multiple copies of a complex specific dehydrogenase (E<sub>1</sub>), a complex specific dihydrolipoyl transacetylase (E<sub>2</sub>), and a common flavin adenine dinucleotide (FAD)-dependent LipDH (E<sub>3</sub>).<sup>59–61</sup> The E<sub>2</sub> subunit serves as the core of the complexes, and its associated lipoyl group forms a ~14 Å flexible appendage that permits the lipoyl moiety to rotate among the catalytic sites of the three component enzymes of each complex.<sup>4</sup> The PDC serves a key role in the metabolism of glucose, catalyzing the conversion of pyruvate to acetyl-CoA, which can be further metabolized through the tricarboxylic acid (TCA) cycle (i.e., Krebs cycle or citric acid cycle) for additional energy generation, or be used for the biosynthesis of fatty acids.<sup>62</sup> The overall reaction of the PDC proceeds in five sequential steps (**Figure 2**). The thiamine diphosphate (TDP)-dependent pyruvate dehydrogenase (E<sub>1</sub>) catalyzes the decarboxylation of pyruvate, yielding CO<sub>2</sub>. The hydroxyethyl group formed is transferred to the lipoamide group located on the dihydrolipoyl transacetylase (E<sub>2</sub>) with concomitant reduction of the disulfide bond of the cofactor. The resulting thioester is cleaved by CoA, producing acetyl-CoA. For multiple turnovers to occur, the dihydrolipoamide group must be reoxidized. This occurs through disulfide interchange mediated by a reactive disulfide located on the FAD-dependent LipDH (E<sub>3</sub>). The reduced E<sub>3</sub> subunit in turn is reoxidized by NAD<sup>+</sup>. The reactions of the  $\alpha$ -KDC, which catalyzes the oxidative decarboxylation of  $\alpha$ -ketoglutarate, producing succinyl-CoA, CO<sub>2</sub>, and NADH in the citric acid cycle, resemble those of the PDC. The BCKDC catalyzes the oxidative decarboxylation of the  $\alpha$ -keto acids generated from branched-chain amino acids into their corresponding acyl-CoA derivatives by a mechanism akin to the reactions catalyzed by the PDC and KDC.

In certain Gram-negative bacteria that respire anaerobically (e.g., *Pelobacter carbinolicus*, *Alcaligenes eutrophus*, and *Clostridium magnum*), the degradation of acetoin (3-hydroxy-2-butanone) to acetaldehyde and acetyl-CoA is catalyzed by an ADC, which allows these bacteria to use acetoin as a sole carbon source.<sup>63,64</sup> In many respects the ADC is analogous to the  $\alpha$ -keto acid dehydrogenase complexes. It is also composed of multiple copies of E<sub>1</sub>, E<sub>2</sub>, and E<sub>3</sub> subunits. The E<sub>1</sub> subunit is a TDP-dependent  $\alpha_2\beta_2$  homotetramer and is encoded by the *acoA* and *acoB* genes, respectively. In contrast to the other lipoyl-dependent dehydrogenase complexes, the TDP cofactor in the E<sub>1</sub> subunit is not used to catalyze the decarboxylation of an  $\alpha$ -keto acid, but is used instead to catalyze the fragmentation of an  $\alpha$ -hydroxy ketone with concomitant formation of acetyl-TDP. The E<sub>2</sub> and E<sub>3</sub> subunits act in the same fashion as that described for the other lipoyl-dependent dehydrogenase complexes. Although a mechanistic analysis of the ADC has not yet been conducted, a reasonable working model for the reaction can be drawn from its similarity to those of the  $\alpha$ -keto acid dehydrogenase complexes (**Figure 3**). Concomitant with fragmentation of the  $\alpha$ -hydroxy ketone and the release of acetaldehyde, catalyzed by the E<sub>1</sub> subunit, the acetyl moiety of the resulting hydroxyethyl-TDP intermediate is transferred to the lipoyl cofactor on the E<sub>2</sub> subunit with concomitant reduction of the cofactor. The remaining steps proceed as described for the PDC, with the production of acetyl-CoA and reoxidation of the cofactor by the E<sub>3</sub> subunit (**Figure 3**).

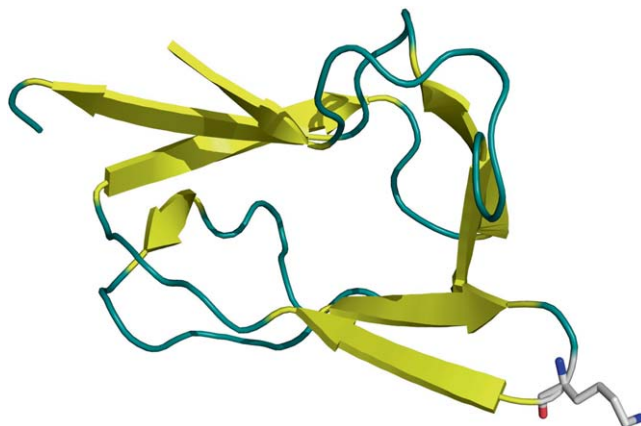




**Figure 3** Schematic representation of the reactions catalyzed by the acetoin dehydrogenase complex. The E<sub>1</sub> subunit (acetoin dehydrogenase) binds thiamine diphosphate (TDP); the E<sub>2</sub> subunit (dihydrolipoyl transacetylase) contains lipoamide; the E<sub>3</sub> subunit (dihydrolipoyl dehydrogenase) binds flavin adenine dinucleotide (FAD).

As discussed above, the E<sub>2</sub> subunit serves as the core of these multienzyme complexes. In *E. coli* it is composed of 24 identical polypeptides arranged in a geometry that approximates a cube, in which trimers occupy each of the vertices.<sup>4</sup> Each polypeptide bears at least one lipoyl domain, and most Gram-negative bacteria, including *E. coli*, contain two to three.<sup>41</sup> The Gram-negative bacterium *Zymomonas mobilis*, however, has only one, as do most Gram-positive bacteria.<sup>65</sup> The number of lipoyl domains does not seem to play a substantial role in the activity of the complex, because deletion of one or two of the domains in *E. coli* has very little effect on overall activity.<sup>66</sup> Each E<sub>2</sub> polypeptide also contains a peripheral subunit-binding domain, which is involved in the binding of the E<sub>1</sub> and/or E<sub>3</sub> subunits to the E<sub>2</sub> core. A large core-forming dihydrolipoyl transacetylase (catalytic) domain resides at the C-terminus.<sup>61,67,68</sup> These domains are linked by flexible polypeptide chains (~25–30 residues) rich in alanine, proline, and charged amino acids.<sup>69,70</sup> Eukaryotic PDC, as well as some from the Gram-positive bacteria such as *Bacillus stearothermophilus*, differ in overall structure. They are typically composed of 60 E<sub>2</sub> subunits arranged in a pentagonal dodecahedron, displaying icosahedral symmetry.<sup>68</sup> Mammalian PDC also bind two regulatory enzymes, a kinase and a phosphatase, which attach to the E<sub>2</sub> subunit core, as well as an E<sub>3</sub>-binding protein (E3BP), previously referred to as protein X, which acts as a recognition motif for LipDH.<sup>41</sup> The E3BP contains a single lipoyl domain, a peripheral subunit-binding domain, and a transacetylase catalytic domain; however, it is unable to catalyze transacetylation because of the absence of key catalytic residues.<sup>71,72</sup>

The principal role of lipoyl domains is to serve as carriers of reaction intermediates. Each contains ~80 amino acids, which includes one fully conserved lysyl residue that is the site of lipoylation. The three-dimensional structures of the lipoyl domains from the KDC of *E. coli* and *Azotobacter vinelandii* as well as that of the PDC from *E. coli* and *B. stearothermophilus* have been determined using nuclear magnetic resonance (NMR) spectroscopy, and are all quite similar.<sup>73–76</sup> They contain a flattened β-barrel composed of a well-defined hydrophobic core that is surrounded by two almost parallel four-stranded antiparallel β-sheets. The lipoyllysine is located at the tip of one β-sheet, whereas the N- and C-termini are in close proximity in a β-sheet on the other side of the domain (Figure 4). Movement of the target lysine to adjacent positions in the β-hairpin using site-directed mutagenesis resulted in no lipoylation, indicating that the location on the hairpin is crucial in proper substrate recognition.<sup>77</sup> Although the lipoyl domains from different multienzyme complexes are similar in size, amino acid composition, and overall structure, the domain from one type of complex cannot substitute for the corresponding domain from another.<sup>78</sup>

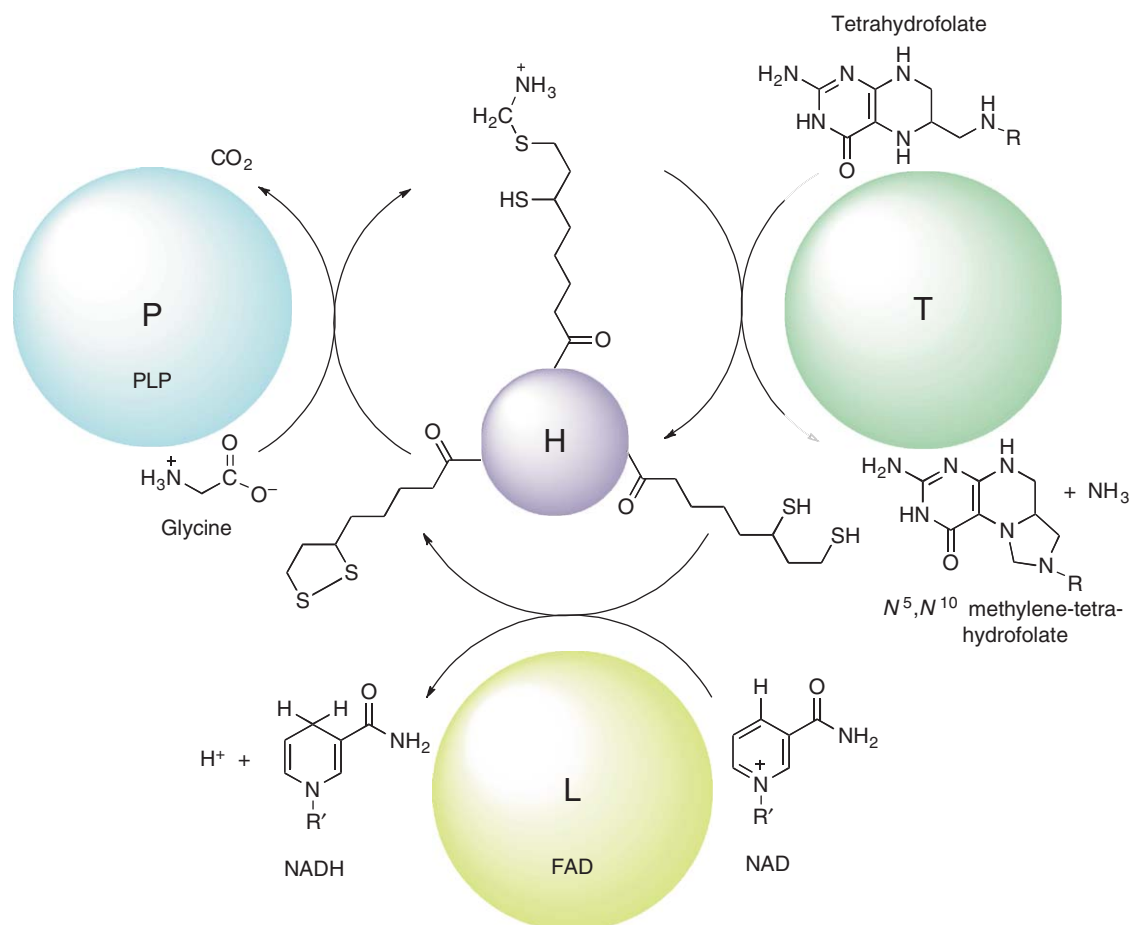


**Figure 4** Structure of the lipoyl domain of the E<sub>2</sub> subunit of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus* solved by NMR (PDB 1LAB). The lysyl residue that bears the lipoyl group (Lys42), found at the tip of a β-hairpin, is shown as sticks, and individual atoms are color-coded (gray, carbon; blue, nitrogen; red, oxygen). The structure was prepared using the PyMOL Molecular Graphics (<http://www.pymol.org>).

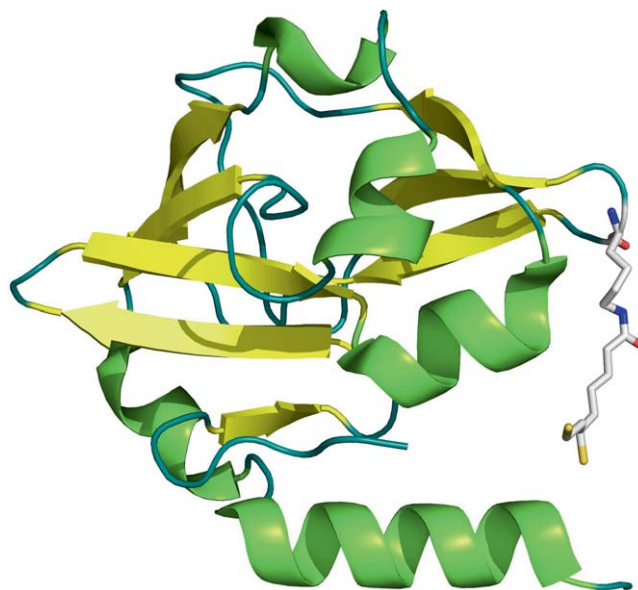
## 7.07.2.3 The Glycine Cleavage System

The GCS catalyzes the reversible oxidative decarboxylation of glycine to  $\text{CO}_2$ , ammonia, and  $N^5, N^{10}$ -methylene-tetrahydrofolate (THF).<sup>79</sup> Unlike the PDC and KDC, the GCS is not a tightly bound multienzyme complex, but a system of four independent proteins (P, H, T, and L) that loosely associate<sup>80</sup> (Figure 5). The homodimeric P-protein, which is equivalent to the  $E_1$  subunits of PDC and KDC, catalyzes the pyridoxal 5'-phosphate (PLP)-dependent decarboxylation of glycine, yielding  $\text{CO}_2$ .<sup>81–83</sup> The resulting aminomethyl group is transferred to the lipoyl-bearing aminomethyltransferase (H-protein) with concomitant reduction of the lipoyl group, forming an aminomethyl H-protein intermediate.<sup>56,84–87</sup> The T-protein catalyzes the formation of ammonia concomitant with the transfer of the methylene group from the lipoyl moiety of the H-protein to THF, affording  $N^5, N^{10}$ -methylene-THF.<sup>85,88</sup> The resulting dihydrolipoyl group is then reoxidized by LipDH (L-protein) as described for the PDC.<sup>89,90</sup> It has been shown that in pea leaf mitochondria the PDC and GCS share the same LipDH.<sup>91</sup>

Parallel to the  $E_2$  subunit of the PDC and KDC, the H-protein contains the catalytically relevant lipoyl pendant. The X-ray crystal structure of the H-protein from *Pisum sativum* containing a reduced lipoyl cofactor has been solved and resembles that of the lipoyl domains of the PDC and KDC.<sup>92</sup> Similarly to the lipoyl domains, the H-protein is composed of two antiparallel  $\beta$ -sheets, forming a 'sandwich' (Figure 6). One unique aspect, however, is that the N-terminal exposed loop of the lipoyl domains is replaced by a helix in the



**Figure 5** Schematic representation of the reactions catalyzed by the glycine cleavage system. The P-protein (acetoin dehydrogenase) binds pyridoxal 5'-phosphate (PLP); the T-protein (dihydrolipoyl transacetylase) contains lipoamide; the L-protein (dihydrolipoyl dehydrogenase) binds flavin adenine dinucleotide (FAD).



**Figure 6** Structure of the H-protein of the glycine cleavage system from *Pisum sativum* (PDB 1DXM). The lipoyllysine cofactor is in its reduced form, dihydrolipoamide, and is shown as sticks. Individual atoms are color-coded (gray, carbon; blue, nitrogen; red, oxygen). The structure was prepared using the PyMOL Molecular Graphics (<http://www.pymol.org>).

H-protein structure. The H-protein also contains two additional  $\beta$ -strands at the N-terminus, and a short helix at the C-terminus,<sup>67,92,93</sup> rendering it significantly larger than the lipoyl domains of the E<sub>2</sub> subunits of the PDC and KDC. The lipoyl group is covalently attached to a lysine residue that resides at the tip of a  $\beta$ -hairpin structure, and in its reduced form is exposed to solvent, allowing it to interact with the L-protein to be reoxidized. By contrast, the structure of H-protein containing a methylaminated lipoyl group showed the cofactor to be locked in a cavity in the protein.<sup>94,95</sup>

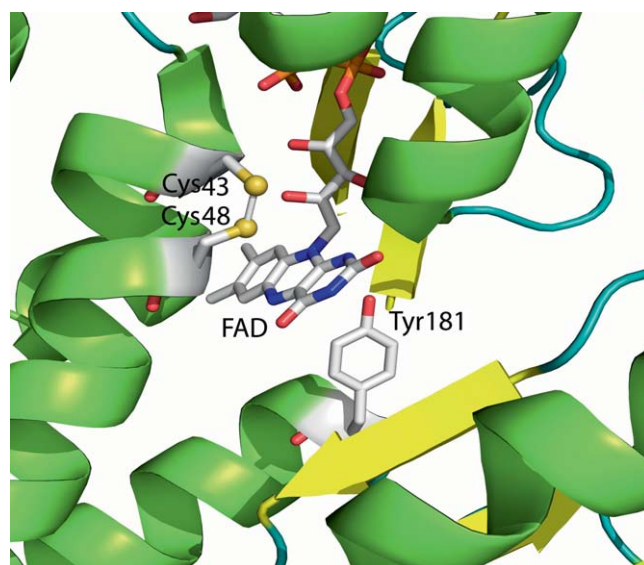
#### 7.07.2.4 Lipoamide Dehydrogenase

LipDH (L-protein or E<sub>3</sub>) belongs to a special class of flavin-containing proteins called the flavoprotein disulfide reductases, which also include, among others, glutathione reductase, trypanothione reductase, mercuric ion reductase, and mycothione reductase.<sup>96,97</sup> In each of these proteins, the flavin cofactor, FAD, acts in concert with two cysteines in the active site of the enzyme to catalyze oxidation or reduction of reduced or oxidized lipoic acid through disulfide interchange. Detailed mechanistic analyses of E<sub>3</sub> from pig heart indicate that catalysis takes place by a two-step process, as described in Equations (1) and (2). Excellent reviews describing the catalytic mechanism of LipDH and the experiments that support it have been published.<sup>96,98</sup> A few key details are mentioned herein to provide a thorough analysis of enzymes that use the lipoyl cofactor, act on the lipoyl cofactor, or participate in its construction.



Several structures of LipDHs have been determined by X-ray diffraction, including a structure from *Pseudomonas putida*, solved to 2.45 Å resolution.<sup>99</sup> The enzyme from *P. putida* consists of two identical subunits – as do all LipDHs – each of molecular mass 48 159 Da, and is only active as a dimer.<sup>96</sup> Each subunit contains four domains. The largest domain binds FAD and is in contact with the other three domains: an NAD-binding domain, a central domain, and an interface domain. The active site of one subunit of LipDH is shown in **Figure 7**. In addition to the FAD cofactor, it contains the disulfide bond formed by conserved cysteine residues 43 and 48, as well as Tyr181, which covers the NAD<sup>+</sup>-binding pocket in the absence of the pyridine cofactor,



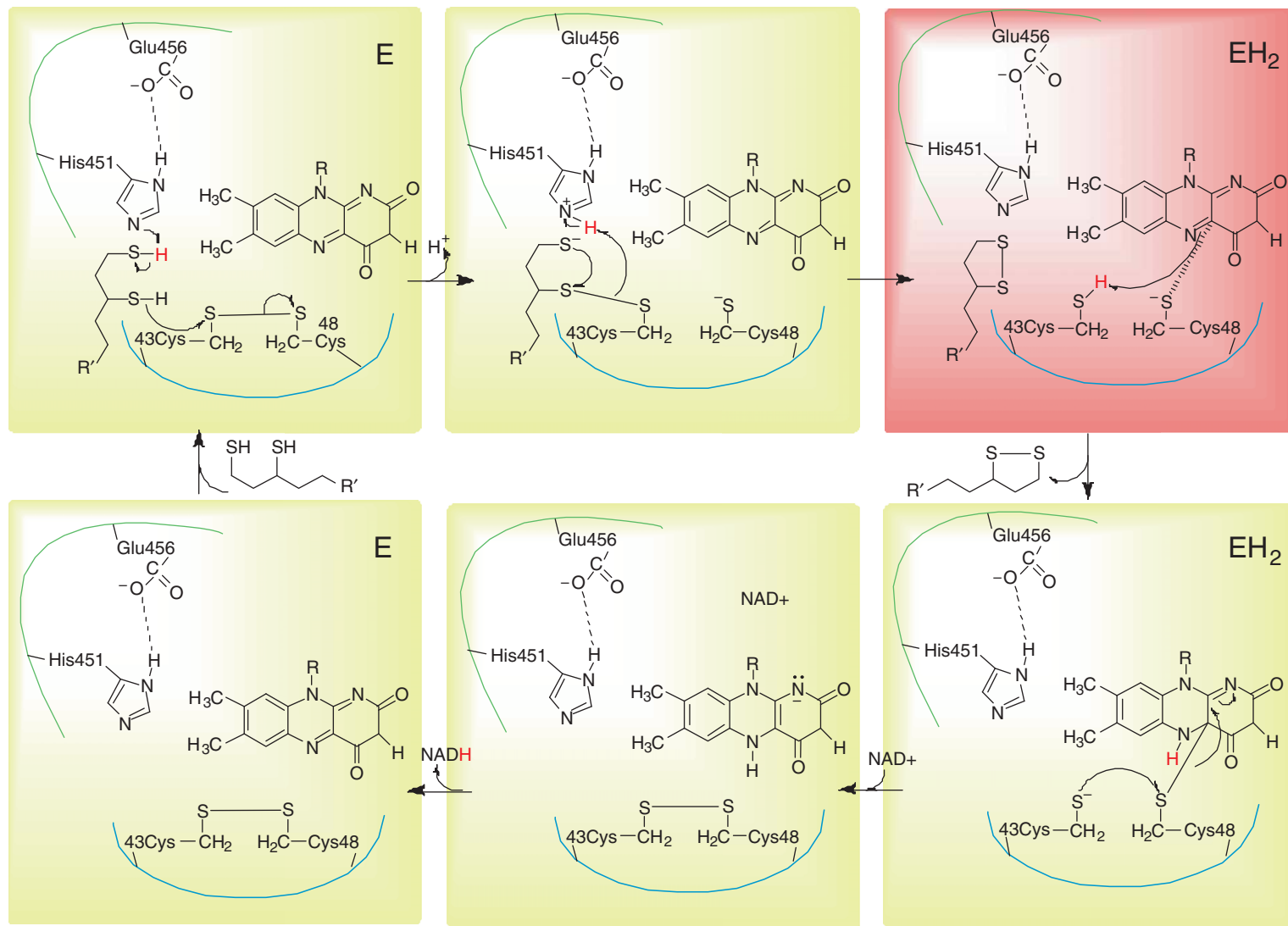


**Figure 7** Structure of the active site of lipoyl dehydrogenase from *Pseudomonas putida*. The key disulfide bond between Cys43 and Cys48 is shown in stick format, as is the flavin cofactor and conserved amino acid Tyr181. The structure was prepared using the PyMOL Molecular Graphics (<http://www.pymol.org>).

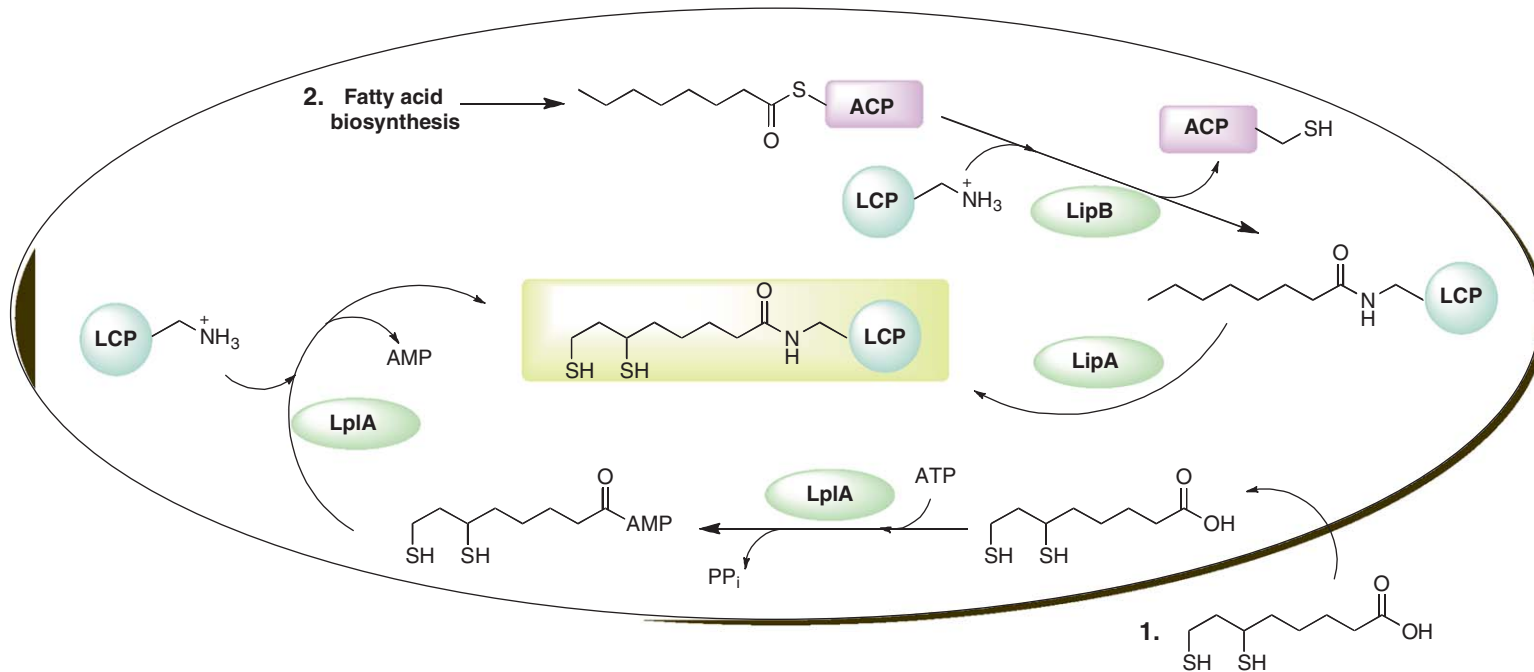
shielding the flavin from bulk solution. In the first half-reaction, written in the direction of substrate oxidation (Figure 8), DHLA reduces the Cys43/Cys48 disulfide bond of fully oxidized LipDH (denoted E), yielding the two-electron reduced state (denoted  $\text{EH}_2$ ). Disulfide interchange is mediated by the general acid/base properties of His451, which is located at the C-terminus on the opposite subunit. The  $\text{EH}_2$  form of the enzyme has a characteristic red color that derives from a charge transfer interaction between the oxidized flavin (the acceptor) and a Cys48 thiolate anion (the donor) (Figure 8). This thiolate anion initiates reduction of the flavin through a covalent mechanism that involves its attack on the flavin 4a carbon, and reduction is completed upon reformation of the Cys43/Cys48 disulfide bond. In the second half-reaction,  $\text{NAD}^+$  oxidizes the flavin by direct hydride transfer, affording NADH with concomitant return to the fully oxidized state of the protein. Pre-steady-state kinetic analyses of pig heart LipDH indicate that the slow step in the reaction is the reduction of the enzyme by dihydrolipoamide to generate the  $\text{EH}_2$  state. Reoxidation of this state by  $\text{NAD}^+$  takes place with a significantly greater rate constant, which has precluded observation of the state containing the reduced flavin ( $\text{E-FADH}^-$ ). In the reverse reaction, reduction of the  $\text{E-FAD}$  state of the enzyme by excess NADH is too fast to measure and affords the  $\text{EH}_2$  state with no detectable intermediates. Support for the reduced flavin as an intermediate in the reaction derives from studies on a variant form of LipDH generated by site-directed mutagenesis, in which His450 of the enzyme from *A. vinelandii* was changed to Ser. This variant displays only 0.5% of the activity of the wild-type enzyme, and in the reverse reaction is drastically compromised in its ability to effect reduction of oxidized lipoamide by the reduced flavin.<sup>100</sup> More recently, pre-steady-state kinetic analysis of a LipDH from *Mycobacterium tuberculosis* – an enzyme that is kinetically slower than previously characterized LipDHs – has allowed the reduced flavin intermediate to be observed in the wild-type enzyme.<sup>101</sup>

### 7.07.3 Pathways for Construction of the Lipoyl Cofactor

Very little lipoate exists in the cell as the free acid unless the cell is supplemented with substantial quantities of the molecule. Instead, it is found almost exclusively in its cofactor form, appended in an amide linkage to a specific lysyl residue on an LCP.<sup>1,54</sup> The pathways in *E. coli* by which LCPs are posttranslationally modified with lipoic acid have recently begun to be illuminated and then extended to other organisms through comparative bioinformatics, and biochemical and genetic approaches. *E. coli* and many other organisms use two distinct routes for the construction of the lipoyl cofactor (Figure 9).<sup>39</sup> One route, often referred to as the



**Figure 8** Reaction mechanism of lipoamide dehydrogenase. The panel in red displays the charge transfer interaction that affords the red color observed for this species.



**Figure 9** Pathways of lipoylation in *Escherichia coli*. (1) Exogenous pathway, wherein lipoic acid is obtained as a nutrient source. Lipoate-protein ligase A (LplA) catalyzes both the ATP-dependent activation of lipoic acid and its transfer to lipoate-carrying proteins (LCPs). The endogenous pathway (2) involves the *de novo* biosynthesis of the lipoyl cofactor as an offshoot of type II fatty acid biosynthesis, occurring on the acyl carrier protein (ACP). LipB catalyzes the transfer of an eight-carbon fatty acyl chain to the target lysyl residue of an appropriate LCP. LipA catalyzes the subsequent insertion of sulfur atoms into C-H bonds at C6 and C8.

exogenous pathway, involves the uptake of lipoic acid as a nutrient source and its attachment to LCPs in an ATP-dependent process. The reaction proceeds through a discrete lipoyl-AMP intermediate and takes place according to Equations (3) and (4).<sup>54,102</sup> In *E. coli*, both half-reactions are catalyzed by lipoyl-protein ligase A (LplA), a protein composed of a single polypeptide.<sup>103</sup> In various other organisms, including vertebrates, each half-reaction is catalyzed by two separate and noninteracting proteins. Lipoyl-activating enzyme (LAE) catalyzes the formation of lipoyl-AMP,<sup>104</sup> whereas lipoyltransferase (LT) catalyzes the transfer of the lipoyl group from lipoyl-AMP to an appropriate LCP.<sup>105</sup> The second route for the construction of the lipoyl cofactor involves its *de novo* biosynthesis through the type II fatty acid biosynthetic pathway.<sup>106,107</sup> In this pathway, two carbon units are added in succession to a 4'-phosphopantetheine prosthetic group that is attached covalently to a small acidic protein referred to as acyl carrier protein (ACP).<sup>108</sup> The 4'-phosphopantetheine cofactor contains a terminal sulfhydryl group to which the growing fatty acyl chain is appended in a thioester linkage. Upon formation of a fatty acyl chain containing eight carbons (i.e., octanoyl chain), octanoyl-[acyl carrier protein]-protein transferase (LipB), the first enzyme in the pathway that is committed to lipoic acid biosynthesis, catalyzes the transfer of the octanoyl chain to the target lysyl group of an appropriate LCP.<sup>109</sup> Lipoyl synthase (LipA), the second enzyme in the *de novo* pathway, then catalyzes sulfur insertion into C-H bonds at C6 and C8 of the appended octanoyl chain.<sup>109,110</sup>



## 7.07.4 Characterization of Enzymes Involved in Lipoic Acid Biosynthesis and Metabolism

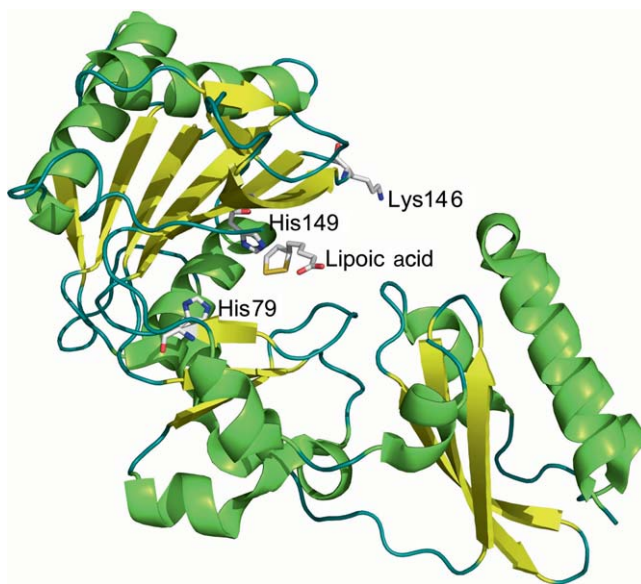
### 7.07.4.1 Lipoyl Protein Ligase A

Considerable biochemical evidence suggested that *E. coli* contained a lipoyl-activating system, which could activate free lipoic acid and catalyze its attachment to LCPs.<sup>102</sup> Early investigations of lipoyl-activating systems used *S. faecalis* strain 10C1 – which is lipoyl deficient – as the model organism. Seminal studies by Reed and his associates showed that the lipoyl-activating system in this organism required lipoic acid, ATP,  $\text{Mg}^{2+}$ , and two separable components, one of which was heat stable and one of which was heat labile. By contrast, the lipoic acid-activating components of *E. coli* could not be fractionated, suggesting that only one protein or a tightly bound complex of two or more proteins was responsible for catalyzing both activation and transfer of lipoic acid to LCPs. Other nucleoside triphosphates could substitute for ATP, with CTP, GTP, ITP, and UTP supporting 40, 25, 15, and 50%, respectively, of the activity supported by ATP.<sup>102</sup> The requirement for a nucleoside triphosphate could be alleviated by the use of lipoyl-AMP, providing evidence for the ping-pong mechanism depicted in Equations (3) and (4). The reversible nature of Equation (3) was further supported in isotope exchange experiments conducted in extracts of *A. vinelandii*, in which the disappearance of lipoyl-AMP was shown to be dependent on the addition of pyrophosphate, with concomitant incorporation of <sup>32</sup>P-labeled pyrophosphate into ATP.<sup>111</sup>

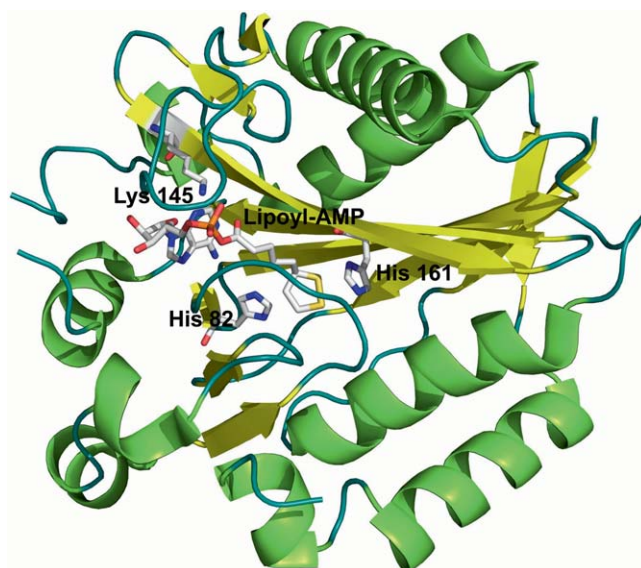
LplA from *E. coli* has enjoyed the most mechanistic interrogation of enzymes that catalyze this reaction.<sup>103</sup> The gene that encodes it, *lplA*, was identified in the early 1990s. Strains with null mutations in the *lplA* locus were able to transport lipoic acid successfully into the cell, but were unable to attach it to LCPs.<sup>112</sup> These strains were also resistant to the effects of seleno-lipoate, which had been shown to inhibit the growth of *E. coli*, because of an inability of the diselenide to be reductively acetylated during catalysis by the PDC, a key step in the dismutation of pyruvate.<sup>113</sup> Isolated *E. coli* LplA was shown to form at least one disulfide bond, which resulted in inactivation of the protein; however, this inactivation could be reversed by the addition of reducing agents. Kinetic studies revealed that the protein could catalyze attachment of dihydro-DL-lipoic acid, D-lipoic acid, octanoic acid, and methyl-lipoic acid with rates that were 80, 83, 3, and 73%, respectively, of the rate with DL-lipoic acid. The protein could also use L-lipoic acid as a substrate, although only the D-form is transferred to the PDC *in vivo*.<sup>103</sup> LplA has also been shown to catalyze attachment of 6-thiooctanoic acid and

6-selenooctanoic acid with rates that were 150 and 6%, respectively, of the rate in the presence of DL-lipoic acid. By contrast, 8-thiooctanoic acid and 8-selenooctanoic acid were not good substrates.

The structures of LplAs from *Thermoplasma acidophilum* and *E. coli* have been solved by X-ray crystallography. Shown in **Figures 10 and 11** are the structures of the *E. coli* enzyme complexed with lipoic acid<sup>114</sup> and the *T. acidophilum* enzyme complexed with lipoyl-AMP,<sup>115</sup> respectively. The *T. acidophilum* enzyme was crystallized in three forms in two separate studies: in the absence of bound substrate;<sup>116</sup> complexed with ATP;<sup>115</sup> and



**Figure 10** Structure of *Thermoplasma acidophilum* LplA modified from PDB file 2ART. LplA is complexed with lipoyl-AMP, shown in stick format, as are several conserved amino acids. The structure was prepared using the PyMOL Molecular Graphics (<http://www.pymol.org>).



**Figure 11** Structure of *Escherichia coli* LplA modified from PDB file 1X2H. LplA is shown in complex with lipoic acid (stick format). Also indicated in stick format are several key conserved residues. The structure was prepared using the PyMOL Molecular Graphics (<http://www.pymol.org>).



complexed with lipoyl-AMP.<sup>115</sup> Its structure is monomeric and is composed of a central core of two nearly orthogonal  $\beta$ -sheets with eight  $\alpha$ -helices surrounding them. Similarity searches yielded two other proteins with significant structural homology. The first was a putative LplA from *Streptococcus pneumoniae* whose primary structure was longer than that of the *T. acidophilum* enzyme by 67 residues at its C-terminus, but which shared 29% sequence identity. The second was the biotin holoenzyme synthase/bio repressor (BirA) from *E. coli*. This enzyme is responsible for ligation of the biotinyl appendage to target proteins and has a sequence identity of 13% with LplA from *E. coli*.

The structure of *T. acidophilum* LplA shows a bifurcated binding pocket, which binds lipoyl-AMP in a U-shaped conformation with the phosphate group and part of the ribose sugar accessible to solvent. The dithiolane ring of the lipoyl group is surrounded by hydrophobic residues and by the side chains of His81 and His161. The presence of two histidyl amino acids in the binding pocket of the dithiolane ring appears to be a common motif in enzymes that are involved in attachment or transfer of the cofactor to LCs. Lys145, conserved also in LipB and biotin ligases, is believed to play a key role in the mechanism of transfer, although detailed mechanistic studies have not yet been forthcoming. Sequence homology searches reveal the presence of three conserved sequence motifs in lipoate ligases: the first is (RRXXGGGXV(F/Y)HD), the second is (KhXGXa), and the third is (HXX(L/M)LXXX(D/N)LXXLXXHL). Note the use of the standard one-letter code for amino acids, in which X represents a nonconserved position and h represents any hydrophobic amino acid. All of these three conserved motifs contain key residues that are part of the lipoyl-AMP-binding pocket. Interestingly, the group that reported activity measurements on the *T. acidophilum* enzyme indicated that the isolated enzyme was unable to catalyze lipoylation of target proteins and suggested that it may function to catalyze only the activation of lipoic acid to lipoyl-AMP.<sup>116</sup>

The structure of *E. coli* LplA is similar to that of the *T. acidophilum* enzyme, although it is markedly larger.<sup>114</sup> It is tempting to speculate that this increased mass contains the necessary machinery to allow this protein to catalyze both steps in lipoic acid attachment, which *E. coli* LplA is known to do. The general structure is composed of two large domains, a C-terminal domain and an N-terminal domain. Lipoic acid is bound in a pocket located between these two domains, with the binding channel on the N-terminal domain. The N-terminal domain is composed of three  $\alpha$ -helices and one  $\beta$ -sheet. The lipoic acid-binding pocket is a hydrophobic cavity formed by Leu17, Phe24, Phe147, and Ala138, and aliphatic parts of the side chains of Glu21 and Ser72. Additionally, residues Ser72 in crystals in one orientation and Arg140 in crystals in a different orientation form a hydrogen bond with the C-terminus of lipoic acid. It was speculated that this weak hydrogen bonding and the shape of the cavity likely allowed LplA to bind both *R*(+)- and *S*(-)-lipoic acid, lipoic acid analogs, and octanoic acid. It must be mentioned that despite the presence of two conserved histidines in the lipoic acid-binding pocket, the orientation of the dithiolane ring of the bound lipoic acid with respect to the histidines is somewhat askew as compared with the structures of *T. acidophilum* LplA with lipoyl-AMP bound, suggesting that this orientation may not be catalytically relevant. The binding of both lipoic acid and ATP, which affords the catalytically relevant complex, might shift the orientation of the dithiolane ring.

#### 7.07.4.2 Lipoate-Activating Enzyme

Although lipoate-protein ligase activity was also detected in mammals, it was determined that the activation and attachment of the lipoyl group to LCs was not carried out by a single protein as it is in *E. coli*, but rather by two separate proteins as is observed in *S. faecalis*. The first, LAE, is responsible for the activation of free lipoic acid to lipoyl-AMP,<sup>117</sup> whereas the second, LT, is responsible for attachment of the lipoyl arm to the target lysyl residue on LCs.<sup>105</sup> Both enzymes have been purified and characterized to some extent. The prototype for mammalian LAEs is the enzyme from bovine liver, which was initially isolated directly from the natural source by conventional methods.<sup>104,117</sup> In addition to activating lipoic acid, it activated octanoic acid, 1,2-dithiolane-3-butyric acid, and 1,2-dithiolane-3-caproic acid, all of which contain a free carboxyl group. Activity assays showed the enzyme to display an absolute requirement for *R*-lipoate, the enantiomer found in nature, as well as MgCl<sub>2</sub>. Surprisingly, the specific activity of the enzyme was about 1000-fold higher when GTP replaced ATP as the requisite nucleoside triphosphate.<sup>104</sup> Subsequent cloning of the cDNA that encodes LAE from bovine liver showed that it contains 577 amino acids, the first 37 of which encoded a mitochondrial targeting sequence. Interestingly, sequence homology searches revealed it to be identical to xenobiotic-metabolizing/



medium chain fatty acid:CoA ligase-III, an enzyme that had been shown to catalyze formation of acyl-CoA and acyl-GMP derivatives of a number of different fatty acids.<sup>104</sup> Therefore, the specificity of attachment of lipoic acid to LCPs must derive from the second enzyme, LT, that functions in attachment.

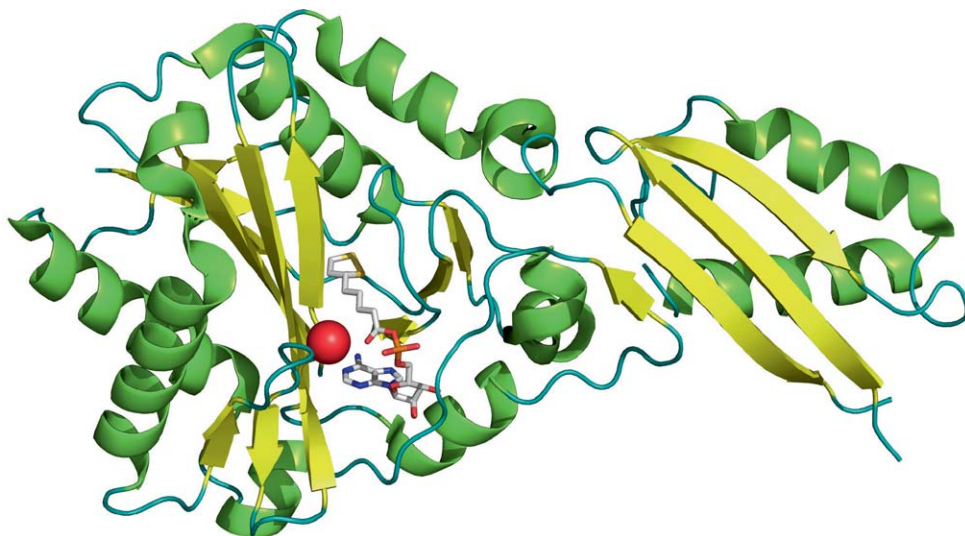
#### 7.07.4.3 Lipoyltransferase

As indicated above, LT catalyzes only the attachment of lipoic acid to LCPs using lipoyl-AMP, or presumably lipoyl-GMP, as the substrate. It does not activate lipoic acid, however.<sup>105</sup> The enzyme from bovine liver has been well characterized and exists as two isoforms, designated LT 1 (LipTI) and LT 2 (LipTII). The only difference in sequence between these two isoforms is the presence of an additional asparaginyl residue at the N-terminus of LipTII. It is hypothesized that both isoforms derive from the same gene product but are processed differently. The cDNA encoding LipTII revealed a protein that contained 373 amino acids, 26 of which were determined to be a mitochondrial targeting sequence. The amino acid sequence of the mature protein has an estimated molecular weight of 39 137 Da and 35% identity with LplA from *E. coli*.<sup>118</sup> The X-ray crystal structure of bovine LT in complex with lipoyl-AMP has been solved, and is shown in **Figure 12**.<sup>119</sup> Although the crystal was not soaked in lipoyl-AMP, the structure contained electron density consistent with its presence, indicating a high affinity of the substrate for the enzyme despite reported  $K_m$  values from steady-state kinetic studies of 13–16  $\mu\text{mol l}^{-1}$ . As described for the similar structure of LplA from *T. acidophilum*, the lipoyl-AMP moiety is bound in a U-shaped conformation, with Lys135, which corresponds to Lys145 in *T. acidophilum* LplA—oriented toward the carbonyl oxygen of lipoyl-AMP, suggesting that it might be important in charge stabilization upon formation of the tetrahedral intermediate during transfer.

Human LT has also been cloned and purified.<sup>120</sup> The gene encoding this protein is organized into four exons and three introns that span ~8 kb of DNA. The amino acid sequence of human LT was determined to have 88% identity with bovine LT I, and 31% identity with LplA from *E. coli*. The mRNA levels for the enzyme were analyzed in various tissues by Northern blot analysis. Tissues with the highest levels of the mRNA were skeletal muscle and heart, followed by kidney and pancreas, with lesser amounts of expression in the liver, brain, placenta, and lung.<sup>120</sup>

#### 7.07.4.4 Octanoyl-[Acyl Carrier Protein]-Protein Transferase

Under lipoate-deficient conditions, octanoylation of the lipoyl domain of the PDC has been observed, consistent with *in vitro* studies that indicate that octanoic acid can serve as an alternate substrate for



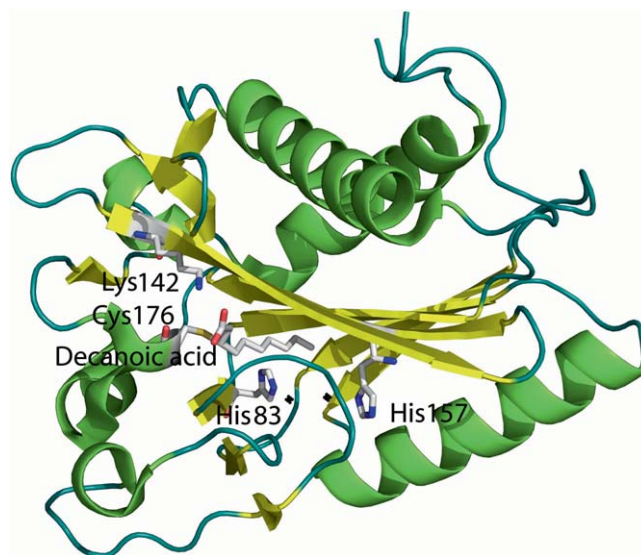
**Figure 12** Structure of bovine liver lipoyltransferase modified from PDB file 2E5A. The protein is in complex with lipoyl-AMP, shown in stick format. A magnesium ion in the binding pocket is shown as a red sphere. The structure was prepared using the PyMOL Molecular Graphics (<http://www.pymol.org>).

LplA.<sup>121,122</sup> *In vivo* labeling experiments with [<sup>35</sup>S]lipoic acid and [<sup>14</sup>C]octanoic acid using a strain of *E. coli* with null mutations in the *lplA* gene showed that lipoylation (or octanoylation) of LCPs occurs, albeit poorly, suggesting an alternate pathway for protein lipoylation.<sup>42,103,112,123</sup> This alternate pathway was shown to be mediated by the product of the *lipB* gene. In this pathway an octanoyl fatty acyl chain, derived from fatty acid biosynthesis as an appendage to the ACP, is transferred to apo LCPs by the action of octanoyl-[acyl carrier protein]-protein transferase, designated LipB. Upon formation of the octanoyl-LCP, LipA, the product of the *lipA* gene in *E. coli*, inserts two sulfur atoms into the octanoyl chain, thus forming the intact cofactor.<sup>107,124,125</sup>

It has been speculated that the exogenous pathway takes precedence over this endogenous pathway for the construction of the lipoyl cofactor.<sup>1,42</sup> *E. coli* LipB is a monomer of molecular mass 23 884 Da and is regulated at least at the level of translation, because initiation of translation is encoded by the rare UUG initiation codon rather than the typical AUG codon.<sup>126</sup> Consistent with this finding, the level of LipB in a culture of wild-type *E. coli* cells was found to be near the level of detection of an antibody that was used to visualize it, which is not uncommon for genes that encode proteins that perform rare forms of posttranslational modification.<sup>124</sup> Moreover, over-expression of the *lipB* gene in *E. coli* was found to significantly decrease the growth rate of the bacterium, suggesting that high concentrations of the protein are toxic to the organism.<sup>126</sup>

Both hexahistidine-tagged and native forms of the recombinant *E. coli* enzyme have been purified.<sup>124,127</sup> The activity of the native enzyme was assessed by monitoring its ability to catalyze transfer of the [1-<sup>14</sup>C]octanoyl group from [1-<sup>14</sup>C]octanoyl-ACP to apo-H-protein. The protein exhibited a turnover number of 0.2 s<sup>-1</sup>, and  $K_m$  values for octanoyl-ACP and apo-H-protein of 10.2 ± 4.4 μmol l<sup>-1</sup> and 13.2 ± 2.9 μmol l<sup>-1</sup>, respectively. *E. coli* LipB contains three cysteinyl residues, and treatment of the protein with 5,5'-dithiobis-(2-nitrobenzoic acid) as a function of time showed that one cysteine is modified with a rate constant that is ~50-fold greater than that of the other two.<sup>127</sup> An independent study by the Cronan laboratory showed that *E. coli* LipB catalyzes its reaction through a ping-pong mechanism, in which the octanoyl group from octanoyl-ACP is first transferred to Cys169 of LipB before it is transferred to the appropriate LCP.<sup>125</sup>

Detailed studies of the mechanism of catalysis by LipB have not been forthcoming; however, an X-ray structure of the enzyme from *M. tuberculosis* has allowed a glimpse of the active site of the protein, which shows interesting similarities to the active site of LplAs (Figure 13).<sup>128</sup> The most striking feature of the LipB structure from *M. tuberculosis*, however, is electron density that corresponds to decanoic acid covalently attached at C3 in



**Figure 13** Structure of octanoyl-[acyl carrier protein]-protein transferase (LipB) from *Mycobacterium tuberculosis*, adapted from PDB file 1W66. Key conserved amino acids are labeled and shown in stick format. The two black spheres represent water molecules that are within hydrogen bonding distance to the two indicated histidines. Decanoic acid is depicted as an adduct with Cys176, the residue that participates in covalent catalysis. The structure was prepared using the Pymol Molecular Graphics (<http://www.pymol.org>).

a thioether linkage to Cys176. This cysteine is equivalent to Cys169 of the *E. coli* enzyme, which is known to participate in covalent catalysis by acting as an intermediate carrier of the fatty acyl chain to be transferred. This finding was unexpected, although revealing, and was proposed to be an artifact of heterologous expression of the *M. tuberculosis* enzyme in *E. coli*, arising from a postulated Michael addition of the cysteine onto *cis*- or *trans*-2-decanoyl-ACP, which are intermediates in fatty acid biosynthesis in *E. coli* but not in *M. tuberculosis*, wherein small-chain fatty acids are synthesized using a type I fatty acid synthase. The structure also shows the presence of two histidines at the distal end of the fatty acyl chain (His83 and His157), which are conserved in LipB proteins and are also found in LplAs and LTs. In addition, Lys142, also conserved in LipB proteins and found in LplAs and LT (Lys135 in *T. acidophilum* LplA and Lys145 in bovine liver LT), is positioned  $\sim 3.4$  Å from the sulfur atom of Cys176, and has been proposed to function in concert with Cys176 as a Cys/Lys dyad.

Based on the structure of *M. tuberculosis* LipB as well as prior mechanistic studies, a mechanism for catalysis by LipB was proposed (Figure 14). The reaction is initiated by binding of octanoyl-ACP to LipB, upon which *trans*-thioesterification takes place to generate the covalent octanoyl-LipB intermediate. The appropriate LCP (denoted E<sub>2</sub>) binds subsequent to release of holo-ACP, whereupon amide formation takes place to afford the corresponding octanoyl-LCP. It is postulated that proton transfers to and from substrates and products are mediated, at least in part, by Lys142.<sup>128</sup>

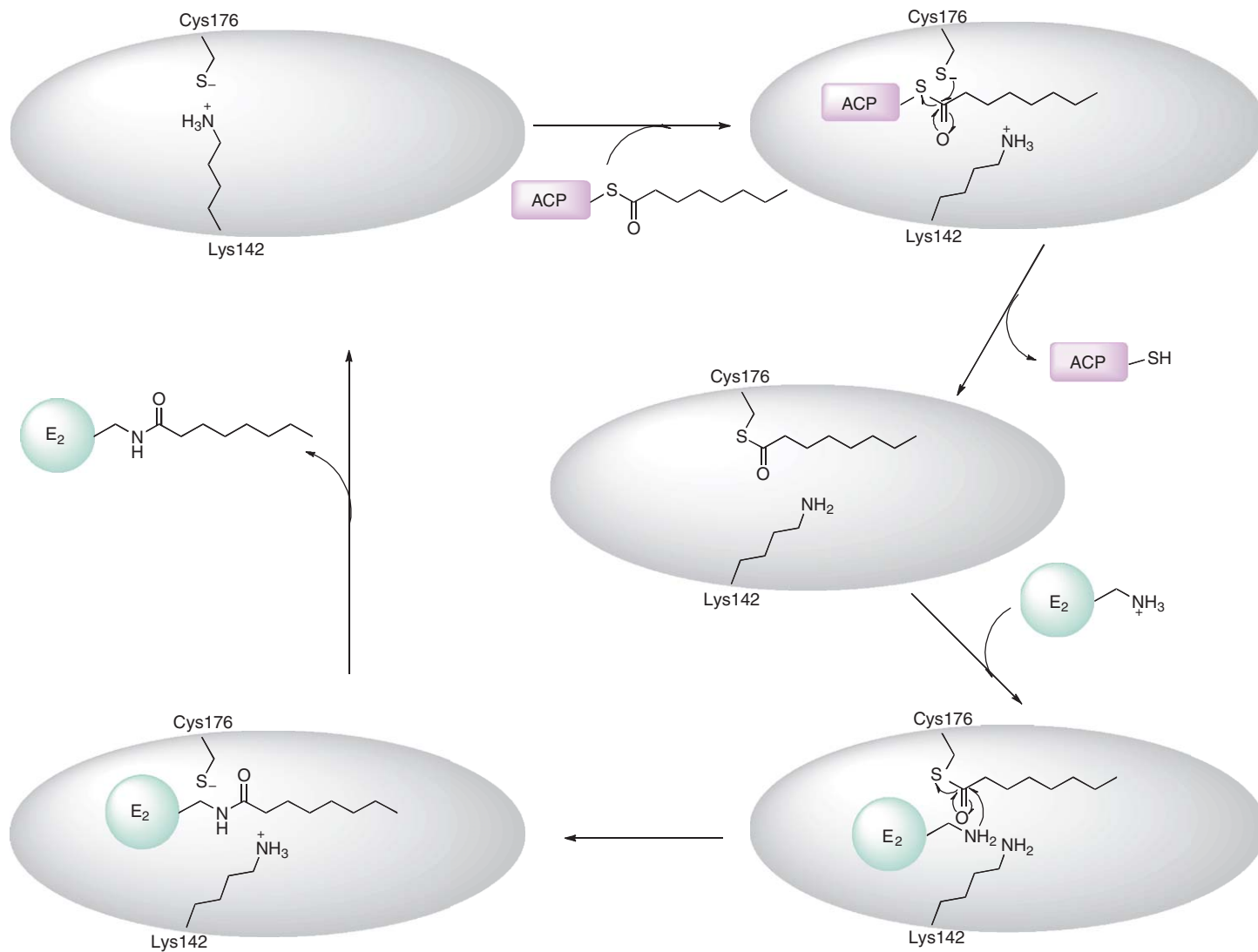
#### 7.07.4.5 Lipoyl Synthase

The identification of the pathways for the construction of the lipoyl cofactor resulted primarily from genetic studies in *E. coli*. A recent review, authored by the key contributor to the understanding of lipolic acid biosynthesis and attachment to LCPs, is available and worthy of attention.<sup>39</sup> In 1968, Herbert and Guest generated *E. coli* mutants that were auxotrophic for lipolic acid through chemical mutagenesis.<sup>129</sup> These mutants could grow anaerobically on glucose, but required glucose and lipolic acid for aerobic growth. The lipolic acid requirement could be circumvented, however, if the auxotrophs were cultured in the presence of acetate and succinate, given that the lipoyl-containing multienzyme complexes PDC and KDC catalyze formation of acetyl-CoA and succinyl-CoA, respectively. The use of these strains facilitated cloning of the *E. coli lipA* gene by complementation and showed that both auxotrophs had mutations in this gene. The finding that these two auxotrophs could be complemented with lipolic acid suggested that the *lipA* gene was involved in the *de novo* biosynthesis of the lipoyl cofactor rather than its attachment to LCPs.

##### 7.07.4.5.1 Metabolic feeding studies

LipA, the product of the *lipA* gene in *E. coli*, is the enzyme that catalyzes the insertion of sulfur atoms at C6 and C8 of an octanoyl fatty acyl chain attached to an LCP. This is an intriguing reaction, because it involves functionalization of a completely unactivated alkyl chain, suggesting the need for chemistry involving carbon-centered radicals. A great deal of what is known about this reaction has been gleaned from metabolic feeding studies that were initiated in the 1960s. A comprehensive review of these studies, authored by one of the major players in the field at that time, has been published and is worthy of reading.<sup>130</sup> Several key findings from these studies are described below. Early investigations into the biosynthesis of lipolic acid by Lester Reed showed that [1-<sup>14</sup>C]lipoate could be isolated from cultures of *E. coli* when the growth media was supplemented with [1-<sup>14</sup>C]octanoate. By contrast, supplementation with [1-<sup>14</sup>C]hexanoate or [1,6-<sup>14</sup>C]adipate did not give rise to radiolabeled lipolic acid.<sup>131</sup> The failure of these latter two compounds to be converted into lipolic acid suggested that octanoate is incorporated intact and that the incorporation of radioactivity does not stem from a pathway that uses building blocks derived from the catabolism of octanoate, perhaps by fatty acid  $\beta$ -oxidation.

Further independent studies provided specific mechanistic insight into the incorporation of sulfur atoms at C6 and C8 of octanoate. Substrates containing tritium at C5, C6, C7 (racemic mixtures), and C8 were coadministered with [1-<sup>14</sup>C]octanoate, and the ratio of <sup>3</sup>H/<sup>14</sup>C in isolated lipoate was used to analyze for hydrogen atom abstraction at a given position. Minimal loss of tritium was observed with substrates containing radiolabel at C5 and C7, suggesting that intermediate desaturation is an unlikely route to the formation of lipolic acid. Minimal loss of tritium at C8 was also observed and was attributed to a substantial C-<sup>3</sup>H kinetic isotope effect at this position. By contrast,  $\sim 50\%$  of the tritium content was found in isolated lipolic acid when [6(*R,S*)-<sup>3</sup>H]octanoate was administered to *E. coli*, consistent with stereoselective removal of hydrogen at this



**Figure 14** Proposed mechanism for the LipB reaction. ACP denotes acyl carrier protein.  $E_2$  denotes a generic lipoyl domain. Cys176 and Lys142 are proposed to exist in an ion pair in the resting state of the enzyme and to act as a Cys/Lys dyad. Adapted from Q. Ma; X. Zhao; A. N. Eddine; A. Geerlof; X. Li; J. E. Cronan; S. H. E. Kaufmann; M. Wilmanns, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8662.

position. The subsequent synthesis and use of stereospecifically labeled forms of  $[6\text{-}^3\text{H}]$ -,  $[6(S)\text{-}^3\text{H}]$ -, and  $[6(R)\text{-}^3\text{H}]$ octanoate showed that incorporation of sulfur at C6 occurs with removal of only the *proR* hydrogen, because only 10.9% of the initial tritium content remained in isolated lipoic acid when the *proR*-labeled compound was used, whereas 84% of the tritium remained when the *proS*-labeled compound was used. The stereochemical configuration at C6 of natural lipoic acid is also *R*; however, there is a change in priority of the substituents around C6 in lipoic acid as compared with octanoic acid according to the Cahn–Ingold–Prelog rules. Therefore, sulfur insertion occurs with inversion of configuration at this carbon.

Similar results were obtained in independent studies that were designed to probe the origins, stoichiometries, and stereochemistries of hydrogen atoms incorporated into fatty acids during their biosynthesis.<sup>132,133</sup> GC–MS analysis of lipoic acid extracted from *E. coli* B cells cultured in the presence of  $[\text{methyl-}^2\text{H}_3]$ acetate showed that it possessed a deuterium-labeling pattern that is consistent with its direct biosynthesis from a fatty acid. The additional finding that deuterium was still present at C6 of the isolated lipoic acid indicated that sulfur insertion at this position occurs with inversion of configuration, because acetate-derived hydrogens on even-numbered carbons in fatty acids are incorporated only in the *L*-configuration during fatty acid biosynthesis, whereas the C6 hydrogen in natural lipoic acid is known to be in the *D*-configuration. This finding also corroborates the earlier finding of Parry and Trainor that the *proR* hydrogen is removed from C6. Finally, when lipoic acid was extracted from *E. coli* cells cultured in the presence of  $[\text{U-}^2\text{H}_{15}]$ octanoate, 90.3% of it was found to be derived from the fatty acid and to contain only 13 deuterium atoms, consistent with the removal of only two during its synthesis.

To account for the change in stereochemistry, it was suggested that the conversion of octanoic acid to lipoic acid might occur through intermediate hydroxylation with the retention of configuration, followed by an  $\text{S}_{\text{N}}2$  displacement of an activated form of the hydroxyl group with a sulfur nucleophile.<sup>132</sup> To address this possibility,  $[6(R,S)\text{-}^2\text{H}_1]$ -6-hydroxyoctanoate,  $[8\text{-}^2\text{H}_2]$ -8-hydroxyoctanoate, and  $[8\text{-}^2\text{H}_2](\pm)$ -6,8-dihydroxyoctanoate were synthesized and tested for their ability to be converted into lipoic acid *in vivo*. None of these compounds were used to make lipoic acid,<sup>134</sup> consistent with the known ability of anaerobically cultured *E. coli* to biosynthesize lipoic acid;<sup>132</sup> however, the corresponding monothiolated species,  $[8\text{-}^2\text{H}_2]$ -8-mercaptooctanoic acid and  $[6(R,S)\text{-}^2\text{H}_1]$ -6-mercaptooctanoic acid, were converted into lipoic acid with varying efficiencies.<sup>134</sup> The observation that conversion of  $[8\text{-}^2\text{H}_2]$ -8-mercaptooctanoic into lipoic acid was approximately sixfold greater than that for the labeled 6-mercaptooctanoic acid led to the suggestion that sulfur insertion takes place at C6 prior to C8.<sup>134</sup>

Further studies in Gary Ashley's laboratory to probe the mechanism of sulfur insertion in the biosynthesis of lipoic acid made use of two strains of *E. coli*, W1485-*lip2* and JRG33-*lip9*, that were known to be deficient in lipoic acid biosynthesis because of a mutation in the *lipA* gene.<sup>135</sup> Strain W1485-*lip2* contains a single mutation that converts Ser308 to Phe, whereas strain JRG33-*lip9* contains a single mutation that converts Glu196 to Lys. Each auxotroph was cultured in the presence of octanoic acid, 6-mercaptooctanoic acid, or 8-mercaptooctanoic acid, and growth was assessed as a function of time. No growth was observed with the auxotroph either in the absence of supplementation or when cultures were supplemented with octanoic acid or 6-mercaptooctanoic acid at concentrations up to  $250\ \mu\text{mol l}^{-1}$ . Growth was apparent when cultures were supplemented with 8-mercaptooctanoic acid, with W1485-*lip2* affording a 20-fold higher response than JRG33-*lip9*. Nevertheless, this growth response required micromolar concentrations of substrate to achieve a similar response as cultures supplemented with lipoic acid at nanomolar levels. It was argued that strain W1485-*lip2* must contain a mutation that primarily affects sulfur insertion at C8, whereas strain JRG33-*lip9* contains a mutation that affords a seriously damaged protein that can catalyze neither of the two sulfur insertion steps.

A somewhat similar, but yet perplexing, result was obtained in John Cronan's laboratory, in which an *E. coli* *lipA* auxotroph, generated by transposon mutagenesis, was also shown to grow in the presence of 8-mercaptooctanoic acid or 6-mercaptooctanoic acid.<sup>42</sup> Since insertion of a transposon into the coding region of a protein would be expected to completely ablate a protein's catalytic ability, this result suggests that another protein in the cell is capable of catalyzing sulfur insertion at C6 of 8-mercaptooctanoic acid or a derivative thereof.

#### 7.07.4.5.2 Isolation and characterization

Two *E. coli* strains that were auxotrophic for lipoic acid (initially called *lip*<sup>-</sup> mutants) were used to map the *lip* locus on the *E. coli* chromosome. These strains were shown to lack 2-oxoacid dehydrogenase activity, and required supplementation with lipoic acid or both acetate and succinate for growth.<sup>129</sup> More recent studies showed that overproduction of E<sub>2</sub> of the PDC in either of the mutant strains afforded proteins that were

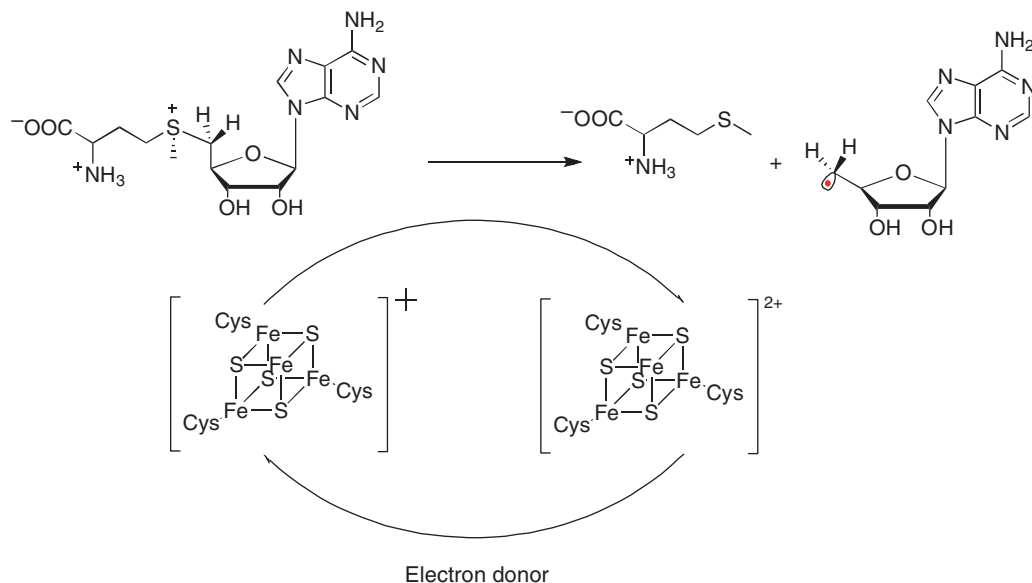


exclusively octanoylated.<sup>121</sup> The *E. coli lipA* gene was independently cloned by complementation and sequenced by the Ashley<sup>136</sup> and Cronan<sup>42</sup> laboratories. The report by Reed and Cronan<sup>42</sup> corrected a misassignment of the actual start codon in the earlier report of Hayden *et al.*,<sup>136</sup> indicating that the mature protein contains 321 amino acids and has a molecular mass of 36 072 Da.

The primary structure of LipA showed it to contain a CXXXCXXC motif, which at that time was known to exist only in a few proteins, and most notably in biotin synthase (BioB), an enzyme that catalyzes a similar reaction.<sup>42,136</sup> In fact, it was noted that the central region of the LipA and BioB amino acid sequences contained ~36% identity.<sup>136</sup> It is now known that the CXXXCXXC motif is a signature sequence for a large class of enzymes called the radical *S*-adenosyl-L-methionine (SAM) superfamily, which uses SAM as a precursor to a 5'-deoxyadenosyl 5'-radical (5'-dA $\cdot$ ), a common and requisite intermediate in catalysis.<sup>137-140</sup> The role of the 5'-dA $\cdot$  in each enzyme system is to abstract a hydrogen atom from the substrate, which either initiates catalysis or creates a stable protein radical cofactor. The formation of a 5'-dA $\cdot$  from SAM takes place through a reductive cleavage of the molecule, affording L-methionine as a by-product (Figure 15). The electron is supplied directly by a [4Fe-4S]<sup>+</sup> cluster that is ligated by the cysteines in the CXXXCXXC motif. The cluster receives electrons from the flavodoxin/flavodoxin reductase system in *E. coli*, which ultimately derives from NADPH. Artificial electron donors, such as sodium dithionite or 5-deazaflavin in combination with light, can also function to reduce the [4Fe-4S]<sup>2+</sup> cluster in *in vitro* experiments.<sup>137-139</sup>

Initial expression of the *lipA* gene yielded a protein product that was found primarily as insoluble aggregates. The aggregates could be resolubilized by treatment with guanidine-HCl, affording protein that was judged to be ~80% pure by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).<sup>42</sup> The isolated protein was ~36 kDa in size, as predicted, but was unable to catalyze conversion of octanoylated E<sub>2</sub> of the KDC into lipoylated E<sub>2</sub>.<sup>42</sup> Octanoylated LCPs were a presumed substrate, since *E. coli lipA* auxotrophs were known to synthesize LCPs that were exclusively octanoylated when their growth media were not supplemented with exogenous lipoic acid.<sup>121</sup>

Subsequent isolations of the enzyme by two additional laboratories afforded protein that was essentially homogeneous, and that contained significant amounts of iron and sulfide.<sup>141,142</sup> In one instance, chemical analysis of the as-isolated protein showed the presence of 1.8 mol of iron per polypeptide; however, analysis by UV-vis and resonance Raman spectroscopies suggested the presence of [4Fe-4S] clusters. Since the purified



**Figure 15** Cleavage of *S*-adenosyl-L-methionine by enzymes within the radical SAM superfamily. The reductive cleavage reaction generates a 5'-deoxyadenosyl 5'-radical and L-methionine, a spectator in the reaction. The physiological electron donor is the flavodoxin/flavodoxin reductase reducing system, with electrons deriving ultimately from NADPH. Artificial electron donors such as sodium dithionite and 5-deazaflavin plus light can also effect reduction in *in vitro* activity determinations.



protein was present in both monomeric and dimeric forms, it was hypothesized that a [4Fe-4S] cluster was formed at the interface of two LipA monomers, each containing a [2Fe-2S] cluster.<sup>141</sup> In the other instance, anaerobic chemical reconstitution of LipA that was solubilized and purified from inclusion bodies afforded a protein with UV-vis and electron paramagnetic resonance (EPR) spectra that were consistent with the presence of [4Fe-4S] clusters.<sup>142</sup>

In neither of the above studies was the isolated protein shown to catalyze the formation of lipic acid. Indeed, at this time, confusion about the identity of the substrate for the reaction was a major issue. Octanoic acid was initially hypothesized to be the substrate, because experimental evidence indicated that it is converted intact into lipic acid *in vivo*. However, null mutations in the *lplA* gene, which encodes the protein that would catalyze attachment of lipic acid to LCPs if it were synthesized from octanoic acid, did not result in growth defects unless either the *lipA* or *lipB* genes also contained null mutations, indicating a second, redundant, pathway for biosynthesis of the lipoyl cofactor in *E. coli*. Because octanoic acid was known to be synthesized *de novo* by type II fatty acid biosynthesis, it followed that the substrate might be an intermediate in the fatty acid biosynthetic pathway. Since octanoyl-E<sub>2</sub> domains had been shown not to be converted into lipoyl-E<sub>2</sub> domains, it was assumed that the substrate for LipA was octanoyl-ACP, and that the substrate for LipB was lipoyl-ACP rather than octanoyl-ACP.

The first evidence of *in vitro* turnover was reported in a study from the Cronan and Marletta laboratories using both native LipA and a form containing a C-terminal hexahistidine tag.<sup>143</sup> Purification of the hexahistidine-tagged protein from the soluble fraction of crude lysate was performed under anaerobic conditions, and analysis of the resulting as-isolated protein, which was dark brown in color, revealed the presence of  $3.4 \pm 0.4$  atoms of iron and  $4.8 \pm 0.8$  atoms of sulfide per monomer. Subsequent EPR analysis of the dithionite-reduced protein again showed the presence of [4Fe-4S]<sup>+</sup> clusters, although spin quantification indicated less than 5% of the reduced form.

A somewhat complex *in vitro* assay for quantification of turnover was developed. As described previously, it was believed that octanoyl-ACP was the substrate for LipA, and that the substrate for LipB was lipoyl-ACP. Reactions contained, therefore, octanoyl-ACP, LipA, SAM, and sodium dithionite as the requisite reductant. Upon incubation for designated times, LipB and apo E<sub>2</sub> of the PDC were added to catalyze transfer of the suspected lipoyl group from lipoyl-ACP to the target lysyl residue of the LCP for analysis of product formation by both enzymatic and mass spectrometric methods. Analysis by enzymatic methods monitored the ability of lipoyl-PDC to catalyze reduction of 3-acetylpyridine adenine dinucleotide (APAD) in the presence of thiamine, pyruvate, CoA, and cysteine. APAD is an analog of nicotinamide adenine dinucleotide that possesses a higher molar absorptivity in its reduced form, and the rate of its reduction is directly proportional to the concentration of lipoyl-PDC present in the assay. Therefore, the amount of lipoyl-PDC produced in a LipA assay could be determined by extrapolating from appropriate standard curves displaying the rates of APAD reduction as a function of known concentrations of lipoyl-PDC. Using this method of quantification, the authors showed that LipA catalyzed the formation of  $0.032 \pm 0.010$  mol of product per mol of LipA polypeptide. Formation of lipoyl-PDC was also verified by matrix-assisted laser desorption ionization (MALDI) mass spectrometry; a species of molecular mass 8974.10, corresponding to the PDC lipoyl-E<sub>2</sub> protein, was observed only in the complete reaction. In reactions that lacked the requisite one-electron reductant, only the octanoylated E<sub>2</sub> protein was observed. Despite the assumption that lipoyl-ACP was the product of the LipA reaction, the authors were unable to detect formation of this species using a variety of methods, which included PAGE in the presence of urea, hydrophobic chromatography, immobilized metal affinity chromatography (IMAC), size-exclusion chromatography, and MALDI mass spectrometry. Given what is now known about the nature of the true substrate for LipA, this is not unexpected. Unbeknownst to the authors at the time, LipA turnover did not commence until the addition of LipB and apo-PDC to the assay mixture, which initiated transfer of the octanoyl group from octanoyl-ACP to the LCP. LipA, then, presumably acted on the octanoyl-LCP substrate.

Although the authors also reported that octanoyl-PDC did not support LipA turnover, a finding that had been reached in an earlier study as well,<sup>42</sup> a subsequent detailed study from the same laboratories provided evidence to the contrary.<sup>109</sup> First, it was shown that *lipB* null strains of *E. coli* could grow on octanoic acid in lieu of lipic acid if they still possessed both *LplA* and *LipA*. This finding suggested that *LplA* could catalyze attachment of octanoic acid to an LCP, which could then in turn be acted upon by *LipA* to form the lipoyl cofactor. In support of this model, the authors further reported that upon accumulation of *d*<sub>15</sub>-labeled octanoyl-E<sub>2</sub> domains in a *lipA lipB fadE* ( $\lambda$ ) null strain during culturing in the presence of [*d*<sub>15</sub>]octanoic acid, *LipA*,

produced by the expression of the *lipA* gene upon transduction of the cells with phage particles containing a *lipA* cosmid, could catalyze formation of the lipoyl cofactor. The null mutation in the *lipB* gene in the strain used prevented the accumulation of unlabeled octanoyl-E<sub>2</sub> domains, whereas the *fadE* null mutation prevented the degradation of [*d*<sub>15</sub>]octanoic acid by  $\beta$ -oxidation. Complementary *in vitro* studies showed that cleavage of SAM was triggered almost exclusively when octanoyl-E<sub>2</sub> was used as substrate as compared to octanoyl-ACP.<sup>109</sup> As discussed below, a separate study, in which an octanolyated form of the H-protein of the GCS was used as substrate, reported the time-dependent formation of both 5'-dA and lipoyl-H-protein with similar rate constants.

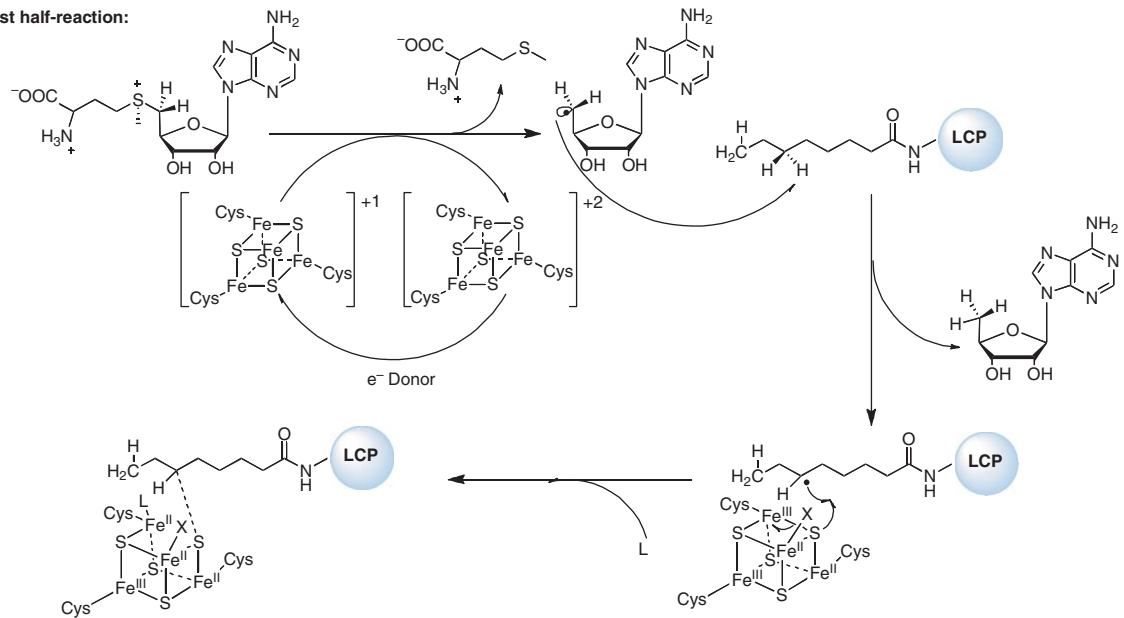
As detailed above, all characterized enzymes within the radical SAM superfamily contain a CX<sub>3</sub>CX<sub>2</sub>C motif – or a motif that is only slightly altered – in which the cysteines therein coordinate a [4Fe-4S]<sup>2+/+</sup> cluster that is requisite for turnover. LipAs, however, contain two conserved cysteine-containing motifs, the second of which is CX<sub>4</sub>CX<sub>5</sub>C. To interrogate the function of the cysteines lying in this second motif, the Booker and Krebs laboratories used a variety of spectroscopic and analytical methods in combination with site-specific replacements of the cysteines therein.<sup>144</sup> Wild-type and variant LipA proteins containing an N-terminal hexahistidine tag were produced by coexpression of the appropriate *lipA* gene in the presence of a helper plasmid, pDB1282, which encodes an operon of genes from *A. vinelandii* that are known to be involved in the biosynthesis and incorporation of Fe/S clusters into their host proteins.<sup>110</sup> Growth and expression in the presence of minimal medium precluded incorporation of unwanted metals in a protein containing eight cysteinyl residues, and also allowed for incorporation of <sup>57</sup>Fe into the protein for analysis by Mössbauer spectroscopy through its addition to the growth media at induction. The wild-type and variant LipA proteins were isolated by IMAC under anaerobic (<1 ppm O<sub>2</sub>) conditions, and were characterized with respect to their ability to catalyze turnover, their iron and acid-labile sulfide content, and their UV-vis, EPR, and Mössbauer spectra.

As-isolated wild-type LipA contained  $6.9 \pm 0.5$  mol of iron and  $6.4 \pm 0.9$  mol of sulfide per mol of protein, and possessed a UV-vis spectrum that resembled those previously reported. Mössbauer spectroscopy showed that the sample contained almost exclusively [4Fe-4S]<sup>2+</sup> clusters with at most 5% of the total iron in the form of [2Fe-2S] clusters, indicating that there were  $\sim 1.7 \pm 0.2$  [4Fe-4S] clusters per as-isolated wild-type LipA polypeptide. When the as-isolated protein was reconstituted in the presence of additional <sup>57</sup>Fe and sulfide and then characterized, it was found to contain  $13.8 \pm 0.6$  mol of iron and  $13.1 \pm 0.2$  mol of sulfide per mol of polypeptide. However, Mössbauer analysis of the reconstituted protein showed it to contain significant amounts of adventitiously bound iron in addition to [4Fe-4S] clusters ( $67 \pm 6\%$ ), indicating that only  $9.2 \pm 0.4$  mol of iron and  $8.8 \pm 0.1$  mol of sulfide per mol of polypeptide were in the form of Fe-S clusters, suggestive of two in the [4Fe-4S] configuration per polypeptide. Characterization of two Cys → Ala triple variants corroborated the conclusion reached after analysis of the wild-type enzyme. One of these variants, C68A-C73A-C79A, contained substitutions in the motif found only in LipAs, whereas the other, C94A-C98A-C101A, contained substitutions in the motif found in all radical SAM enzymes. As predicted, both as-isolated and reconstituted forms of these variant proteins contained [4Fe-4S] clusters; however, the cluster content was less than or equal to one per polypeptide, supporting the contention that both cysteinyl-containing motifs serve to coordinate [4Fe-4S] clusters that play intimate roles in catalysis. Consistent with this hypothesis, single Cys → Ala or Cys → Asp substitutions in each of the cysteines in the two motifs afforded inactive proteins, whereas substitutions of two nonconserved cysteines afforded proteins with near wild-type activity.<sup>144</sup> As will be described below, LipA catalyzes no more than one turnover and then becomes inactive. This behavior stems from the observation that the protein acts as both a catalyst and a reagent; the sulfur atom that is inserted into octanoyl-LCPs to generate the lipoyl cofactor derives from LipA itself.<sup>145</sup> Current opinion, deriving from studies on LipA<sup>145</sup> as well as two additional enzymes that catalyze sulfur insertion into inactivated C-H bonds, MiaB protein,<sup>146</sup> and BioB,<sup>147–149</sup> is that the second iron-sulfur cluster is the direct source of the inserted sulfur atom.<sup>150</sup>

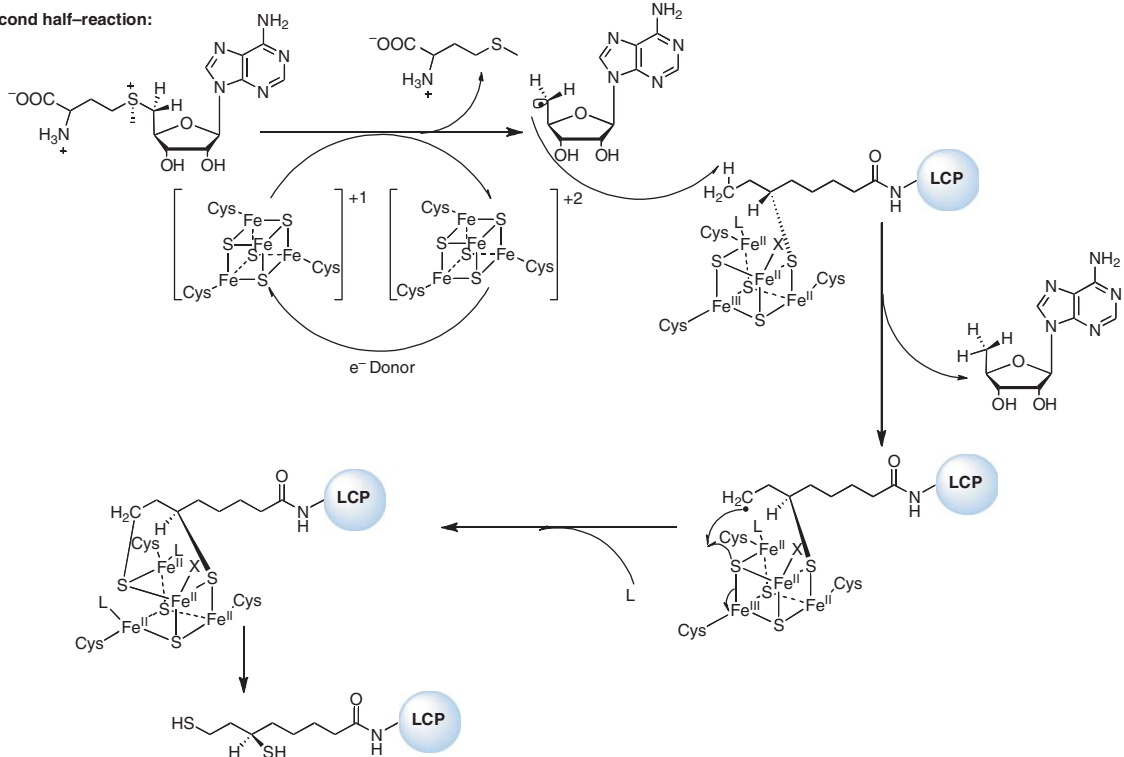
#### 7.07.4.5.3 Mechanistic studies of LipA

A working model for the mechanism of catalysis by LipA is shown in **Figure 16**. Spectroscopic and X-ray diffraction studies of several enzymes within the radical SAM superfamily have shown that SAM binds in contact to the noncysteinyl coordinated iron of the [4Fe-4S] cluster forming a bidentate chelate by its

## First half-reaction:



## Second half-reaction:



**Figure 16** Proposed mechanism of the LipA reaction. Catalysis occurs on an octanoyl substrate attached in an amide linkage to a lipoyl carrier protein (LCP). In the first half-reaction, reductive cleavage of SAM to generate a 5'-deoxyadenosyl 5'-radical allows for hydrogen atom abstraction at C6 of the fatty acyl chain. The proposed source of sulfur is a second [4Fe-4S] cluster, housed in the CX<sub>4</sub>CX<sub>5</sub>C motif conserved in lipoyl synthases. In the second half-reaction, the second 5'-deoxyadenosyl 5'-radical abstracts a hydrogen atom from C8 of the fatty acyl chain, allowing for sulfur insertion at this position. The addition of two protons affords lipoic acid in its reduced form.

amino- and carboxyl substituents.<sup>151–153</sup> An electron is passed from flavodoxin (or dithionite) to the [4Fe–4S] cluster and then into the sulfonium of SAM, inducing a reductive cleavage of the molecule to L-methionine and a 5'-deoxyadenosyl 5'-radical. The 5'-dA· abstracts a hydrogen atom from C6 of an octanoyl chain appended to an LCP, affording a substrate radical. The substrate radical then attacks one of the bridging  $\mu$ -sulfido ligands of the [4Fe–4S] cluster ligated by cysteines residing in the CX<sub>4</sub>CX<sub>5</sub>C motif, resulting in an intermediate species that is bound covalently to the Fe–S cluster. A second round of similar chemistry takes place, wherein a second equivalent of SAM is cleaved to form a 5'-dA· that abstracts a hydrogen atom from C8 of the fatty acyl chain. The resulting substrate radical attacks another of the bridging  $\mu$ -sulfido ligands of the sacrificial cluster, resulting in a species, which upon protonation, collapses to form lipoic acid in its reduced form.<sup>150</sup> An advantage of using an Fe–S cluster as the direct source of sulfur derives from the ability of iron atoms in Fe–S clusters to be amenable to one-electron exchanges, cycling between +2 and +3 oxidation states, which is necessitated by a mechanism of catalysis involving unpaired electrons on the substrate. A typical [4Fe–4S]<sup>2+</sup> cluster contains two Fe<sup>2+</sup> and two Fe<sup>3+</sup> atoms. Each attack of the substrate radical onto the bridging  $\mu$ -sulfido ligands of the cluster would be concomitant with an inner-sphere electron transfer to one of the Fe<sup>3+</sup> atoms, affording four Fe<sup>2+</sup> atoms upon one complete turnover.

The mechanism of catalysis shown in **Figure 16** makes several predictions that have been tested experimentally, the first of which is that two equivalents of 5'-dA are produced for each equivalent of lipoyl product.<sup>110</sup> In reactions using octanoyl–H-protein as the substrate and the flavodoxin/flavodoxin reductase reducing system as the source of the requisite electron, it was shown that 50  $\mu\text{mol l}^{-1}$  LipA catalyzed production of 18  $\mu\text{mol l}^{-1}$  lipoyl–H-protein and 50  $\mu\text{mol l}^{-1}$  5'-dA in a pseudo-first-order manner, exhibiting rate constants of 0.175 and 0.144  $\text{min}^{-1}$ , respectively. At each time point, the stoichiometry of 5'-dA to lipoyl–H-protein varied from  $\sim 2.2$  to 2.7, slightly greater than the expected ratio of 2.0. When [*d*<sub>15</sub>]octanoyl–H-protein was used as a substrate, with subsequent analysis by GC–MS of the 5'-dA produced, as expected most of the 5'-dA contained one deuterium atom. However, small, but significant, amounts of unlabeled 5'-dA were also produced, suggesting that LipA catalyzes slow abortive cleavage of SAM, and therefore providing rationale for the observed stoichiometry. Most surprising, however, was that in the presence of [*d*<sub>15</sub>]octanoyl–H-protein, the amplitude of 5'-dA formed was  $\sim 50\%$  of that observed when unlabeled substrate was used; and, production of the lipoyl product was undetectable, although species consistent with the molecular mass of monothiolated substrates were observed. This observation suggested that abstraction of a deuterium from one of the positions of the fatty acyl substrate was governed by a sizeable kinetic isotope effect, consistent with the observations from the Parry laboratory more than 20 years earlier.

Studies from the laboratory of Peter Roach extended the observations made by the Booker laboratory and provided evidence that sulfur insertion at C6 occurs before the insertion at C8.<sup>154</sup> These studies were carried out on a thermostable LipA from *Sulfolobus solfataricus*<sup>155</sup> using small octanoyllysine-containing peptides as substrates and monitoring of the lipoyl product by HPLC with UV detection. The use of a [8,8,8-*d*<sub>3</sub>]octanoyllysyl-containing substrate resulted in almost no formation of lipoic acid, whereas the use of a [6,6-*d*<sub>2</sub>]octanoyllysyl-containing substrate afforded the lipoylated product, indicating that the large kinetic isotope effect observed previously by the Booker laboratory derived from deuterium abstraction from the C8 position. Moreover, when unlabeled substrate was incubated with LipA, the production of lipoic acid as well as a species containing only one sulfur atom was observed. Characterization of this monothiolated species by NMR showed it to be the 6-mercaptooctanoylated product. No evidence for the 8-mercaptooctanoylated product was found, leading to the suggestion that sulfur insertion occurs at C6 before C8.<sup>154</sup>

Direct evidence supporting the contention that the sulfur atom inserted in the LipA reaction derives from LipA itself emanates from labeling studies conducted by Cicchillo and Booker.<sup>145</sup> Sodium sulfide was synthesized from natural abundance elemental sulfur (95% <sup>32</sup>S) or from <sup>34</sup>S elemental sulfur, and used as the only source of sulfur in minimal medium for the production of correspondingly labeled *E. coli* LipA. When LipA that was isolated from <sup>34</sup>S-containing media was incubated with octanoyl–H-protein under turnover conditions, the lipoyl group formed contained <sup>34</sup>S at the same natural abundance as the Na<sub>2</sub> <sup>34</sup>S used to supplement the growth medium. Since the only source of <sup>34</sup>S in the assay mixture was the LipA protein itself, this observation indicated that the protein must serve as both catalyst and reagent. When <sup>34</sup>S-labeled LipA was mixed in equimolar amounts with LipA isolated from media containing sulfur at natural abundance, two predominant forms of the lipoyl product were synthesized; one contained two <sup>34</sup>S atoms, whereas the other

contained two  $^{32}\text{S}$  atoms. The absence of forms containing one  $^{34}\text{S}$  atom and one  $^{32}\text{S}$  atom indicates that each LipA polypeptide contributes both sulfur atoms during catalysis, and that monothiolated species are not released during turnover and rebound by other LipA proteins.

#### 7.07.4.6 Lipoamidase

Lpa catalyzes the cleavage of the amide linkage that tethers lipoic acid to the target lysyl residue of an LCP. This activity was first observed in 1958, and the enzyme was originally given the name lipoyl-X hydrolase.<sup>102</sup> In early studies to characterize the reactions associated with generating protein-bound lipoamide in cell-free extracts of *S. faecalis*, it was noticed that several amides of lipoic acid, designed to act as lipoic acid antagonists because of their lack of a free carboxylate group, did not inhibit reactions of the PDC. In fact, they activated the PDC, but only in the presence of ATP, suggesting that their amide bonds were cleaved during incubation. This observation was surprising, given that certain fractions of cell-free extract exhibited much lower PDH activity than others. Incubation of this fraction with partially purified *E. coli* PDH resulted in greatly reduced PDH activity as compared to the incubation in its absence. Around the same time, the presence of an enzyme from yeast with similar properties was reported, wherein the decrease in the activity of the KDC was monitored.<sup>156</sup> These lipoyl-X hydrolases were later renamed lipoamidase, based on a study in which the enzyme was purified 100-fold from *S. faecalis* and found to operate on numerous substrates, including protein-bound lipoic acid, methyl lipoate, and some *N*-lipoyl amino acids.<sup>52</sup> Lpa from *S. faecalis* was also tested with biocytin, and showed no cleavage of the biotin appendage, differentiating itself from biotinidase, a similar enzyme capable of cleaving biotin-containing amide linkages.

More recently, Jiang and Cronan cloned the gene that encodes Lpa from *S. faecalis* (also known as *E. faecalis*) strain 10C1, and designated it *lpa*.<sup>157</sup> The toxicity of Lpa to *E. coli* made selection of viable clones a difficult task, as well as overproduction of the protein for subsequent isolation. In fact, expression of the *lpa* gene required the use of  $\lambda$ CE6 phage to introduce T7 RNA polymerase because of the extreme toxicity of the protein. The primary structure of the encoded protein suggested that it was a member of the Ser-Ser-Lys triad amidohydrolase family. Isolation of the hexahistidine-tagged Lpa allowed *in vitro* functional studies to be conducted. Substitution of the two serine residues in the Ser-Ser-Lys triad with alanine resulted in inactive proteins, whereas substitution of the lysine residue in the triad with alanine resulted in dramatically diminished activity. The purified enzyme was found to inactivate the PDC and KDC *in vitro*, and even the small molecule  $\epsilon$ -lipoyl- $\alpha$ -acetyl-lysine methyl ester served as an acceptable substrate for the enzyme.

## Acknowledgment

The research reported that was conducted by the authors of this review was supported by a grant from the National Institutes of Health (GM-63847).

### Glossary

**Acyl carrier protein** A small acidic protein that plays a key role in type II fatty acid biosynthesis, serving as the template upon which the growing fatty acyl chain is synthesized.

**Acetoin dehydrogenase complex** A complex of proteins responsible for the conversion of acetoin (3-hydroxy-2-butanone) to acetaldehyde and acetyl-CoA.

**Branched-chain  $\alpha$ -keto acid dehydrogenase complex** A lipoamide-containing multienzyme complex that functions to degrade the branched-chain amino acids isoleucine, leucine, and valine.

**Dihydrolipoic acid or dihydrolipoamide** The two-electron reduced form of lipoic acid or lipoamide containing two free sulfhydryl groups.

**1,2-Dithiolane ring** The five-membered ring generated upon intramolecular disulfide-bond formation in lipoic acid.



**E<sub>1</sub>** Thiamine diphosphate requiring subunit of pyruvate dehydrogenase,  $\alpha$ -ketoglutarate dehydrogenase, branched-chain  $\alpha$ -keto acid dehydrogenase, and acetoin dehydrogenase complexes.

**E<sub>2</sub>** Lipoyl-bearing subunit of pyruvate dehydrogenase,  $\alpha$ -ketoglutarate dehydrogenase, branched-chain  $\alpha$ -keto acid dehydrogenase, and acetoin dehydrogenase complexes.

**E<sub>3</sub>** Flavin adenine dinucleotide requiring subunit of pyruvate dehydrogenase,  $\alpha$ -ketoglutarate dehydrogenase, branched-chain  $\alpha$ -keto acid dehydrogenase, and acetoin dehydrogenase complexes.

**Glycine cleavage system** System of proteins responsible for the reversible degradation of glycine into ammonia, carbon dioxide, and methylene tetrahydrofolate.

**GLUT4** An insulin-regulated protein that functions to transport glucose.

**H-protein** Lipoyl-bearing protein component of the glycine cleavage system.

**NF- $\kappa$ B** Nuclear factor kappa-light-chain-enhancer of activated B cells. A transcription factor that responds to a variety of stress signals such as reactive oxygen species.

**Pyruvate dehydrogenase complex** System of proteins responsible for the conversion of pyruvate into acetyl-CoA and carbon dioxide.

**$\alpha$ -Ketoglutarate dehydrogenase complex** System of proteins found in the citric acid cycle that is responsible for the conversion of  $\alpha$ -ketoglutarate to succinyl-CoA and carbon dioxide.

**Lipoate-activating enzyme** A protein that catalyzes the ATP-dependent adenylation of lipoic acid, yielding lipoyl-AMP.

**Lipoyl carrier protein** Central proteins of multienzyme complexes that use the lipoyl cofactor. They contain the conserved lysyl residue to which the lipoyl cofactor is bonded covalently in an amide linkage.

**L-protein** The flavin adenine dinucleotide-requiring protein of the glycine cleavage system, which functions to reoxidize dihydrolipoamide after each turnover. In various bacteria it is identical with the E<sub>3</sub> protein.

**Lipoyltransferase** A protein that catalyzes the transfer of the lipoyl group from lipoyl-AMP – or other activated forms of the molecule – to LCPs.

**Lipoyl synthase** The protein that catalyzes the final step in the biosynthesis of the lipoyl cofactor, which is the insertion of two sulfur atoms into C–H bonds at C6 and C8 of an octanoyl-LCP.

**Lipoamide dehydrogenase** A protein that catalyzes the reoxidation of dihydrolipoamide after each turnover. Also referred to as E<sub>3</sub>- or L-protein.

**Lipoamide** Cofactor form of oxidized (cyclic disulfide) lipoic acid, in which it is attached covalently to the  $\epsilon$ -amino group of a target lysyl residue on a lipoyl carrier protein.

**LipB** Octanoyl-[acyl carrier protein]-protein transferase. This protein catalyzes the first committed step in the biosynthesis of the lipoyl cofactor, which is the transfer of an eight-carbon fatty acyl chain from octanoyl-acyl carrier protein to the  $\epsilon$  nitrogen of a target lysyl residue on various lipoyl carrier proteins.

**$\alpha$ -Lipoic acid** Canonical form of lipoic acid having the chemical formula C<sub>8</sub>H<sub>14</sub>O<sub>2</sub>S<sub>2</sub>.

**$\beta$ -Lipoic acid** Two-electrons oxidized thiosulfinate form of lipoic acid.

**Lipoyl carrier protein** A protein to which lipoic acid is covalently attached in an amide linkage to the  $\epsilon$  nitrogen atom of a specific lysyl residue. LCPs employ the lipoyl group as a cofactor in enzymatic catalysis.

**Lipoyl cofactor** The lipoyl group attached covalently to LCPs.

**Lipoamidase** Catalyzes the hydrolysis of the lipoyl cofactor, releasing free lipoic acid.

**Lipoate-protein ligase A** A protein that catalyzes the ATP-dependent activation and transfer of lipoic acid to lipoyl carrier proteins.

**P-protein** Subunit of the glycine cleavage system, which uses pyridoxal 5'-phosphate to catalyze the oxidative decarboxylation of glycine.

**Radical SAM** A superfamily of proteins in which S-adenosyl-L-methionine is used as a precursor to a 5'-deoxyadenosyl 5'-radical.

**T-protein** Tetrahydrofolate-containing subunit of the glycine cleavage system. It functions to cleave the aminomethyl group from the lipoyl cofactor, yielding ammonia and methylene tetrahydrofolate.



## References

1. A. A. Herbert; J. R. Guest, *Arch. Microbiol.* **1975**, *106*, 259–266.
2. L. J. Reed, *J. Biol. Chem.* **2001**, *276*, 38329–38336.
3. L. J. Reed; I. C. Gunsalus; G. H. F. Schnakenberg; Q. F. Soper; H. E. Boaz; S. F. Kern; T. V. Parke, *J. Am. Chem. Soc.* **1953**, *75*, 1267–1270.
4. L. Reed, *Acc. Chem. Res.* **1974**, *7*, 40–46.
5. R. M. C. Dawson; D. C. Elliott; W. H. Elliott; K. M. Jones, *Data for Biochemical Research*; Oxford University Press: New York, 1989.
6. J. A. Barltrop; P. M. Hayes; M. Calvin, *J. Am. Chem. Soc.* **1954**, *76*, 4348–4367.
7. K. Mislow; W. C. Meluch, *J. Am. Chem. Soc.* **1956**, *78*, 5920.
8. G. P. Biewenga; J. de Jong; A. Bast, *Arch. Biochem. Biophys.* **1994**, *312*, 114–120.
9. L. J. Reed; B. G. DeBusk; I. C. Gunsalus; G. H. F. Schnakenberg, *J. Am. Chem. Soc.* **1951**, *73*, 5920.
10. G. P. Biewenga; G. R. M. M. Haenen; A. Bast, An Overview of Lipoate Chemistry. In *Lipoic Acid in Health and Disease*; J. Fuchs, L. Packer, G. Zimmer, Eds.; Marcel Dekker, Inc.: New York, 1997; pp 1–32.
11. A. K. Kierner; B. Diesel, Activation of Cytoprotective Signaling Pathways by Alpha-Lipoic Acid. In *Lipoic Acid: Energy Production, Antioxidant Activity and Health Effects*; M. S. Patel, L. Packer, Eds.; CRC Press: New York, 2008.
12. B. Ke, *Biochim. Biophys. Acta* **1957**, *25*, 650–651.
13. D. R. Sanadi; M. Langley; R. L. Searls, *J. Biol. Chem.* **1959**, *234*, 178–182.
14. I. Scott; W. Duncan; V. Ekstrand, *J. Biol. Chem.* **1963**, *238*, 3928–3939.
15. G. Krause; J. Lundström; J. L. Barea; C. P. de la Cuesta; A. Holmgren, *J. Biol. Chem.* **1991**, *266*, 9494–9500.
16. T.-Y. Lin; P. S. Kim, *Biochemistry* **1989**, *28*, 5282–5287.
17. H. Scholich; M. E. Murphy; H. Sies, *Biochim. Biophys. Acta* **1989**, *1001*, 256–261.
18. V. E. Kagan; A. Shvedova; E. Serbinova; S. Khan; C. Swanson; R. Powell; L. Packer, *Biochem. Pharmacol.* **1992**, *44*, 1637–1649.
19. H. R. Rosenberg; R. Culik, *Arch. Biochem. Biophys.* **1959**, *80*, 86–93.
20. L. Packer; H. J. Tritschler; K. Wessel, *Free Radic. Biol. Med.* **1997**, *22*, 359–378.
21. L. Packer; E. H. Witt; H. J. Tritschler, *Free Radic. Biol. Med.* **1995**, *19*, 227–250.
22. G. R. M. M. Haenen; A. Bast, *Biochem. Pharmacol.* **1991**, *42*, 2244–2246.
23. V. Borcea; J. Nourooz-Zadeh; S. P. Wolff; M. Klevesath; M. Hofmann; H. Ulrich; P. Wahl; R. Ziegler; H. Trischler; B. Halliwell; P. P. Nawroth, *Free Radic. Biol. Med.* **1999**, *26* (11–12), 1495–1500.
24. O. Tirosh; C. K. Sen; S. Roy; M. S. Kobayashi; L. Packer, *Free Radic. Biol. Med.* **1999**, *26* (11–12), 1418–1426.
25. K. Hager; A. Marahrens; M. Kenkies; P. Riederer; G. Munch, Alpha-lipoic acid as a new treatment option for Alzheimer type dementia. *Arch. Geron. Geri.* **2001**, *32*, 275–282.
26. B. Halliwell; J. M. C. Gutteridge, *Free Radicals in Biology and Medicine*; Oxford University Press: Oxford, 1999.
27. G. P. Biewenga; G. R. M. M. Haenen; A. Bast, *Gen. Pharmacol.* **1997**, *29*, 233–238.
28. D. Konrad; R. Somwar; G. Sweeney; K. Yaworsky; M. Hayashi; T. Ramlal; A. Klip, *Diabetes* **2001**, *50*, 1464–1471.
29. D. Han; R. T. Hamilton; P. Y. Lam; L. Packer, Modulation of Cellular Redox and Metabolic Status by Lipoic Acid. In *Lipoic Acid: Energy Production, Antioxidant Activity and Health Effects*; M. S. Patel, L. Packer, Eds.; CRC Press: New York, 2008.
30. S. Ghosh; M. J. May; E. B. Kopp, *Annu. Rev. Immunol.* **1998**, *16*, 225–260.
31. M. A. Hofmann; A. Bierhaus; R. Ziegler; P. Wahl; P. Nawroth; H. J. Tritschler, Lipoate Effects on Atherogenesis. In *Lipoic Acid in Health and Disease*; J. Fuchs, L. Packer, G. Zimmer, Eds.; Marcel Dekker, Inc.: New York, 1997.
32. R. Schreck; K. Albermann; P. A. Baeuerle, *Free Radic. Res. Commun.* **1992**, *17*, 221–237.
33. J. R. Matthews; N. Wakasugi; J. L. Virelizier; J. Yodoi; R. T. Hay, *Nucleic Acids Res.* **1992**, *20*, 3821–3830.
34. J. Fuchs; H. Schöfer, Redox Modulation of Signal Transduction and Gene Expression in HIV Infection: The Role of the Antioxidant Lipoate. In *Lipoic Acid in Health and Disease*; J. Fuchs, L. Packer, G. Zimmer, Eds.; Marcel Dekker, Inc.: New York, 1997.
35. E. J. Duh; W. J. Maury; T. M. Folks; A. S. Fauci; A. B. Rabson, *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5974–5978.
36. V. Pande; M. J. Ramos, *Curr. Med. Chem.* **2003**, *10*, 1603–1615.
37. Y. Suzuki; B. B. Aggarwal; L. Packer, *Biochem. Biophys. Res. Commun.* **1992**, *189*, 1709–1715.
38. A. Baur; T. Harrer; M. Peukert; G. Jahn; J. R. Kalden; B. Fleckenstein, *Klin. Wochenschr.* **1991**, *69*, 722–724.
39. J. E. Cronan; X. Zhao; Y. Jiang, *Adv. Microb. Physiol.* **2005**, *50*, 103–146.
40. N. M. Nesbitt; R. M. Cicchillo; K.-H. Lee; T. L. Grove; S. J. Booker, Lipoic Acid Biosynthesis. In *Lipoic Acid: Energy Production, Antioxidant Activity and Health Effects*; M. S. Patel, L. Packer, Eds.; CRC Press: New York, 2008.
41. R. N. Perham, *Annu. Rev. Biochem.* **2000**, *69*, 961–1004.
42. K. E. Reed; J. E. Cronan, Jr., *J. Bacteriol.* **1993**, *175*, 1325–1336.
43. D. J. O'kane; I. C. Gunsalus, *J. Bacteriol.* **1948**, *56*, 499–506.
44. H. R. Skeggs; L. D. Wright; E. L. Cresson; G. D. E. MacRae; C. H. Hoffman; D. E. Wolf; K. Folkers, *J. Bacteriol.* **1956**, *72*, 519–524.
45. I. C. Gunsalus; M. I. Dolin; L. Struglia, *J. Biol. Chem.* **1952**, *194*, 849–857.
46. L. J. Reed; B. G. D. Busk; I. C. Gunsalus; G. H. F. Schnakenberg, *Science* **1951**, *114*, 93.
47. J. A. J. Brockman; E. L. R. Stokstad; E. L. Patterson; J. V. Pierce; M. Macchi; F. P. Day, *J. Am. Chem. Soc.* **1952**, *74*, 1868.
48. M. W. Bullock; J. A. Brockman, Jr.; E. L. Patterson; J. V. Pierce; E. L. R. Stokstad, *J. Am. Chem. Soc.* **1952**, *74*, 3455.
49. E. Walton; A. F. Wagner; L. H. Peterson; F. W. Holly; K. Folkers, *J. Am. Chem. Soc.* **1954**, *76*, 4748.
50. I. C. Gunsalus, *Fed. Proc.* **1954**, *13*, 715–722.
51. M. Koike; K. Suzuki, *Methods Enzymol.* **1970**, *18*, 292–298.
52. Y. Suzuki; L. J. Reed, *J. Biol. Chem.* **1963**, *238*, 4021–4025.
53. M. Koike; L. J. Reed, *J. Biol. Chem.* **1960**, *235*, 1931–1938.
54. L. J. Reed; M. Koike; M. E. Levitch; F. R. Leach, *J. Biol. Chem.* **1958**, *232*, 143–158.

55. H. Nawa; W. T. Brady; M. Koike; L. J. Reed, *J. Am. Chem. Soc.* **1960**, *82*, 896–903.
56. J. R. Robinson; S. M. Klein; R. D. Sagers, *J. Biol. Chem.* **1973**, *248*, 5319–5323.
57. M. C. Ambrose-Griffen; M. J. Danson; W. G. Griffen; G. Hale; R. N. Perham, *Biochem. J.* **1980**, *187*, 393–401.
58. K. J. Angelides; G. G. Hammes, *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 4877–4880.
59. K. Okamura-Ikeda; K. Fujiwara; Y. Motokawa, *J. Biol. Chem.* **1987**, *262*, 6746–6749.
60. R. M. Oliver; L. J. Reed, J. R. Harris Ed., Multienzyme Complexes. In *Electron Microscopy of Proteins*; Academic Press: London, 1982; Vol. 2.
61. R. N. Perham, *Biochemistry* **1991**, *30*, 8501–8512.
62. D. Voet; J. G. Voet; C. W. Pratt, *Fundamentals of Biochemistry*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, 2006.
63. N. Kruger; F. B. Oppermann; H. Lorenzl; A. Steinbuchel, *J. Bacteriol.* **1994**, *176*, 3614–3630.
64. H. Priefert; S. Hein; N. Kruger; K. Zeh; B. Schmidt; A. Steinbuchel, *J. Bacteriol.* **1991**, *173*, 4056–4071.
65. U. Neveling; R. Klasen; S. Bringer-Meyer; S. Sahn, *J. Bacteriol.* **1998**, *180*, 1540–1548.
66. J. R. Guest; H. M. Lewis, *J. Mol. Biol.* **1985**, *185*, 743–754.
67. A. Berg; A. de Kok, *Biol. Chem.* **1997**, *378*, 617–634.
68. L. J. Reed; M. L. Hackert, *J. Biol. Chem.* **1990**, *265*, 8971–8974.
69. J. D. F. Green; R. N. Perham; S. J. Ullrich; E. Appella, *J. Biol. Chem.* **1992**, *267*, 23484–23488.
70. S. L. Turner; G. C. Russell; M. P. Williamson; J. R. Guest, *Protein Eng.* **1993**, *6*, 101–108.
71. R. A. Harris; M. M. Bowker-Kinley; P. Wu; J. Jeng; K. M. Popov, *J. Biol. Chem.* **1997**, *272*, 19746–19751.
72. L. G. Korotchkina; M. S. Patel, Pyruvate Dehydrogenase Complex Regulation and Lipoic Acid. In *Lipoic Acid: Energy Production, Antioxidant Activity and Health Effects*; L. Packer, M. S. Patel, Eds.; CRC Press: Boca Raton, 2008.
73. A. Berg; J. Vervoort; A. de Kok, *J. Mol. Biol.* **1996**, *261*, 432–442.
74. F. Dardel; A. L. Davis; E. D. Laue; R. N. Perham, *J. Mol. Biol.* **1993**, *229*, 1037–1048.
75. P. M. Ricaud; M. J. Howard; E. L. Roberts; R. W. Broadhurst; R. N. Perham, *J. Mol. Biol.* **1996**, *264*, 179–190.
76. J. D. F. Green; E. D. Laue; R. N. Perham; S. T. Ali; J. R. Guest, *J. Mol. Biol.* **1995**, *248*, 328–343.
77. N. G. Wallis; R. N. Perham, *J. Mol. Biol.* **1994**, *236*, 209–216.
78. L. D. Graham; L. C. Packman; R. N. Perham, *Biochemistry* **1989**, *28*, 1574–1581.
79. R. D. Sagers; I. C. Gunsalus, *J. Bacteriol.* **1961**, *81*, 541–549.
80. D. J. Oliver; M. Neuburger; J. Bourguignon; R. Douce, *Plant Physiol.* **1990**, *94*, 833–839.
81. K. Fujiwara; K. Okamura-Ikeda; Y. Motokawa, *J. Biol. Chem.* **1984**, *259*, 10664–10668.
82. K. Hiraga; G. Kikuchi, *J. Biol. Chem.* **1980**, *255*, 11664–11670.
83. S. M. Klein; R. D. Sagers, *J. Biol. Chem.* **1967**, *242*, 297–300.
84. K. Fujiwara; K. Okamura; Y. Motokawa, *Arch. Biochem. Biophys.* **1979**, *197*, 454–462.
85. H. Kochi; G. Kikuchi, *J. Biochem.* **1974**, *75*, 1113–1127.
86. H. Kochi; G. Kikuchi, *Arch. Biochem. Biophys.* **1976**, *173*, 71–81.
87. Y. Motokawa; G. Kikuchi, *Arch. Biochem. Biophys.* **1969**, *135*, 402–409.
88. Y. Motokawa; G. Kikuchi, *J. Biochem.* **1969**, *65*, 71–75.
89. M. L. Baginsky; F. M. Huennekens, *Arch. Biochem. Biophys.* **1967**, *120*, 703–711.
90. S. M. Klein; R. D. Sagers, *J. Biol. Chem.* **1967**, *242*, 297–300.
91. J. Bourguignon; V. Merand; S. Rawsthorne; E. Forest; R. Douce, *Biochem. J.* **1996**, *313*, 229–234.
92. S. Pares; C. Cohen-Addad; L. Sieker; M. Neuburger; R. Douce, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4850–4853.
93. S. Pares; C. Cohen-Addad; L. Sieker; M. Neuburger; R. Douce, *Acta Crystallogr. D Biol. Crystallogr.* **1995**, *51*, 1041–1051.
94. C. Cohen-Addad; S. Pares; L. Sieker; M. Neuburger; R. Douce, *Struct. Biol.* **1995**, *2*, 63–68.
95. M. Faure; J. Bourguignon; M. Neuburger; D. Macherel; L. Sieker; R. Ober; R. Kahn; C. Cohen-Addad; R. Douce, *Eur. J. Biochem.* **2000**, *267*, 2890–2898.
96. C. H. Williams, Jr., Lipoamide Dehydrogenase, Glutathione Reductase, Thioredoxin Reductase, and Mercuric Reductase – A Family of Flavoenzyme Transhydrogenases. In *Chemistry and Biochemistry of Flavoenzymes*; F. Muller, Ed.; CRC Press: Boca Raton, 1992; Vol. III.
97. M. P. Patel; J. S. Blanchard, *Biochemistry* **1999**, *38*, 11827–11833.
98. A. de Kok; W. J. H. van Berkel, Lipoamide Dehydrogenase. In *Alpha-Keto Acid Dehydrogenase Complexes*; M. S. Patel, T. E. Roche, R. A. Harris, Eds.; Birkhäuser Verlag: Basel, 1996.
99. A. Mattevi; G. Obmolova; J. R. Sokatch; C. Betzel; W. G. J. Hol, *Proteins* **1992**, *13*, 336–351.
100. J. Benen; W. J. H. van Berkel; N. Dieteren; D. Arscott; C. Williams, Jr.; C. Veeger; A. de Kok, *Eur. J. Biochem.* **1992**, *207*, 487–497.
101. A. Argryrou; J. S. Blanchard; B. A. Palfey, *Biochemistry* **2002**, *41*, 14580–14590.
102. L. J. Reed; F. R. Leach; M. Koike, *J. Biol. Chem.* **1958**, *232*, 123–142.
103. D. E. Green; T. W. Morris; J. Green; J. E. Cronan, Jr.; J. R. Guest, *Biochem. J.* **1995**, *309*, 853–862.
104. K. Fujiwara; S. Takeuchi; K. Okamura-Ikeda; Y. Motokawa, *J. Biol. Chem.* **2001**, *276*, 28819–28823.
105. K. Fujiwara; K. Okamura-Ikeda; Y. Motokawa, *J. Biol. Chem.* **1994**, *269*, 16605–16609.
106. S. Brody; C. Oh; U. H. E. Schweizer, *FEBS Lett.* **1997**, *408*, 217–220.
107. S. W. Jordan; J. E. J. Cronan, *J. Biol. Chem.* **1997**, *272*, 17903–17906.
108. P. W. Majerus; A. W. Alberts; P. R. Vagelos, *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *51*, 1231–1238.
109. S. Zhao; J. R. Miller; Y. Jiang; M. A. Marletta; J. E. Cronan, Jr., *Chem. Biol.* **2003**, *10*, 1293–1302.
110. R. M. Cicchillo; D. F. Iwig; A. D. Jones; N. M. Nesbitt; C. Baleanu-Gogonea; M. G. Souder; L. Tu; S. J. Booker, *Biochemistry* **2004**, *43*, 6378–6386.
111. S. K. Mitra; D. P. Burma, *J. Biol. Chem.* **1965**, *240*, 4072–4080.
112. T. W. Morris; K. E. Reed; J. E. Cronan, Jr., *J. Biol. Chem.* **1994**, *269*, 16091–16100.
113. K. E. Reed; T. W. Morris; J. E. Cronan, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3720–3724.
114. K. Fujiwara; S. Toma; K. Okamura-Ikeda; Y. Motokawa; A. Nakagawa; H. Taniguchi, *J. Biol. Chem.* **2005**, *280*, 33645–33651.
115. D. J. Kim; K. H. Kim; H. H. Lee; S. J. Lee; J. Y. Ha; H. J. Yoon; S. W. Suh, *J. Biol. Chem.* **2005**, *280*, 38081–38089.

116. E. McManus; B. F. Luisi; R. N. Perham, *J. Mol. Biol.* **2006**, 356, 625–637.
117. J. N. Tsunoda; K. T. Yasunobu, *Arch. Biochem. Biophys.* **1967**, 118, 385–401.
118. K. Fujiwara; K. Okamura-Ikeda; Y. Motokawa, *J. Biol. Chem.* **1997**, 272, 31974–31978.
119. K. Fujiwara; H. Hosaka; M. Matsuda; K. Okamura-Ikeda; Y. Motokawa; M. Suzuki; A. Nakagawa; H. Taniguchi, *J. Mol. Biol.* **2007**, 371, 222–234.
120. K. Fujiwara; M. Suzuki; Y. Okumachi; K. Okamura-Ikeda; T. Fujiwara; E. Takahashi; Y. Motokawa, *Eur. J. Biochem.* **1999**, 260, 761–767.
121. S. T. Ali; A. J. G. Moir; P. R. Ashton; P. C. Engel; J. R. Guest, *Mol. Microbiol.* **1990**, 4, 943–950.
122. F. Dardel; L. C. Packman; R. N. Perham, *FEBS Lett.* **1990**, 264, 206–210.
123. T. W. Morris; K. E. Reed; J. E. J. Cronan, *J. Bacteriol.* **1995**, 177, 1–10.
124. S. W. Jordan; J. E. Cronan, Jr., *J. Bacteriol.* **2003**, 185, 1582–1589.
125. X. Zhao; J. R. Miller; J. E. Cronan, *Biochemistry* **2005**, 44, 16737–16746.
126. R. Vaisvila; L. J. Rasmussen; A. Løbner-Olesen; U. von Freilesleben; M. G. Marinus, *Biochim. Biophys. Acta* **2000**, 1494, 43–53.
127. N. M. Nesbitt; C. Baleanu-Gogonea; R. M. Cicchillo; K. Goodson; D. F. Iwig; J. A. Broadwater; J. A. Haas; B. G. Fox; S. J. Booker, *Protein Expr. Purif.* **2005**, 39, 269–282.
128. Q. Ma; A. N. Eddine; A. Geerlof; X. Li; J. E. Cronan; S. H. E. Kaufmann; M. Wilmanns, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 8662–8667.
129. A. A. Herbert; J. R. Guest, *J. Gen. Microbiol.* **1968**, 53, 363–381.
130. R. J. Parry, *Tetrahedron* **1983**, 39, 1215–1238.
131. L. J. Reed; T. Okaichi; I. Nakanishi, *Abstr. Int. Symp. Chem. Nat. Prod. (Kyoto)*, **1964**, 218–220.
132. R. H. White, *Biochemistry* **1980**, 19, 15–19.
133. R. H. White, *Biochemistry* **1980**, 19, 9–15.
134. R. H. White, *J. Am. Chem. Soc.* **1980**, 102, 6605–6607.
135. M. A. Hayden; I. Y. Huang; G. Iliopoulos; M. Orozco; G. W. Ashley, *Biochemistry* **1993**, 32, 3778–3782.
136. M. A. Hayden; I. Huang; D. E. Bussiere; G. W. Ashley, *J. Biol. Chem.* **1992**, 267, 9512–9515.
137. J. Cheek; J. B. Broderick, *J. Biol. Inorg. Chem.* **2001**, 6, 209–226.
138. M. Fontecave; E. Mulliez; S. Ollagnier-de Choudens, *Curr. Opin. Chem. Biol.* **2001**, 5, 506–511.
139. P. A. Frey; O. T. Magnusson, *Chem. Rev.* **2003**, 103, 2129–2148.
140. H. J. Sofia; G. Chen; B. G. Hetzler; J. F. Reyes-Spindola; N. E. Miller, *Nucleic Acids Res.* **2001**, 29, 1097–1106.
141. R. W. Busby; J. P. M. Schelvis; D. S. Yu; G. T. Babcock; M. A. Marletta, *J. Am. Chem. Soc.* **1999**, 121, 4706–4707.
142. S. Ollagnier-de Choudens; M. Fontecave, *FEBS Lett.* **1999**, 453, 25–28.
143. J. R. Miller; R. W. Busby; S. W. Jordan; J. Cheek; T. F. Henshaw; G. W. Ashley; J. B. Broderick; J. E. Cronan, Jr.; M. A. Marletta, *Biochemistry* **2000**, 39, 15166–15178.
144. R. M. Cicchillo; K.-H. Lee; C. Baleanu-Gogonea; N. M. Nesbitt; C. Krebs; S. J. Booker, *Biochemistry* **2004**, 43, 11770–11781.
145. R. M. Cicchillo; S. J. Booker, *J. Am. Chem. Soc.* **2005**, 127, 2860–2861.
146. F. Pierrel; T. Douki; M. Fontecave; M. Atta, *J. Biol. Chem.* **2004**, 279, 47555–47653.
147. G. N. L. Jameson; M. M. Cospser; H. L. Hernández; M. K. Johnson; B. H. Huynh, *Biochemistry* **2004**, 43, 2022–2031.
148. B. Tse Sum Bui; T. A. Mattioli; D. Florentin; G. Bolbach; A. Marquet, *Biochemistry* **2006**, 45, 3824–3834.
149. N. B. Ugulava; C. J. Sacanell; J. T. Jarrett, *Biochemistry* **2001**, 40, 8352–8358.
150. S. J. Booker; R. M. Cicchillo; T. L. Grove, *Curr. Opin. Chem. Biol.* **2007**, 11, 543–552.
151. D. Chen; C. Walsby; B. M. Hoffman; P. A. Frey, *J. Am. Chem. Soc.* **2003**, 125, 11788–11789.
152. G. Layer; D. W. Heinz; D. Jahn; W. D. Schubert, *Curr. Opin. Chem. Biol.* **2004**, 8, 468–476.
153. C. J. Walsby; D. Ortillo; J. Yang; M. R. Nnyepi; W. E. Broderick; B. M. Hoffmann; J. B. Broderick, *Inorg. Chem.* **2005**, 44, 727–741.
154. P. Douglas; M. Kriek; P. Bryant; P. L. Roach, *Angew. Chem.* **2006**, 118, 5321–5323.
155. P. Bryant; M. Kriek; R. J. Wood; P. L. Roach, *Anal. Biochem.* **2006**, 351, 44–49.
156. G. R. Seaman, *J. Biol. Chem.* **1959**, 234, 161–164.
157. Y. Jiang; J. E. Cronan, *J. Biol. Chem.* **2005**, 280, 2244–2256.

## Biographical Sketches



Elizabeth S. Billgren was raised in Bemus Point, NY where she graduated from Maple Grove Jr./Sr. High School in 2001 with the intention of pursuing a degree in theater. She began studies at Jamestown Community College (Jamestown, NY) where she became interested in science and graduated in 2003 with an AA in liberal arts and humanities and an AS in mathematics and science. She went on to earn a B.S. from The State University of New York at Fredonia in biology in 2005. She was first introduced to laboratory research by Dr. Wayne Yunghans and received a grant from the Biology Research Endowment Fund to complete a research project in cell biology during the summer of 2004. She then attended The Pennsylvania State University (University Park, PA) and earned an M.S. in biochemistry, microbiology, and molecular biology in the laboratory of Dr. Squire Booker in 2008. Elizabeth is currently employed by United Chemical Technologies, Inc. (Lewistown, PA) as an analytical chemist.



Robert M. Cicchillo obtained both his B.S. degree in biology (1998) and M.S. degree in chemistry (2000) from Youngstown State University. Soon after acceptance into the Biochemistry and Molecular Biology (BMB) department at The Pennsylvania State University, he joined the laboratory of Dr. Squire J. Booker where he studied the enzymatic mechanisms by which vital metabolic compounds, including lipoic acid and quinolinic acid, are biosynthesized. After earning his Ph.D. in 2006, he moved on to take a position as postdoctoral researcher in the laboratory of Dr. Wilfred A. van der Donk (University of Illinois at Urbana-Champaign) where his efforts focused on elucidating the mechanisms by which the important organophosphorous natural products fosfomycin and phosphinothricin tripeptide are biosynthesized. In the fall of 2008, he accepted a position as a senior biochemist at Dow AgroSciences in Indianapolis, where he currently resides.



Natasha M. Nesbitt was born in Flint, Michigan and received a B.S. in biochemistry from The Michigan State University in 1998. While at The Michigan State University, she obtained undergraduate research experience as a participant in the NSF Summer Research Program and as a McNair/SROP scholar in the laboratory of Professor Jack Preiss. After a year of working at Parke-Davis, she began her graduate studies at The Pennsylvania State University. In 2005 she earned her Ph.D. in biochemistry and molecular biology with Professor Squire J. Booker as an NIH Minority Predoctoral Fellow. She is currently an NIH National Research Service Award Postdoctoral Fellow at Stony Brook University in the laboratory of Professor Nicole S. Sampson. Her research focuses on elucidating the cholesterol metabolic pathway of *Mycobacterium tuberculosis*.



Squire J. Booker was raised in Beaumont, Texas, where he got excited about science during frequent trips as an elementary school student to NASA in Houston with his uncle, Albert J. Price. He received a B.A. degree with a concentration in chemistry from Austin College (Sherman, Texas) in 1987. In the summer of 1986, he participated in the first MIT Minority Summer Science Research Program, where he worked in the laboratory of the late Dr. William H. Orme-Johnson in the Department of Chemistry. He earned his Ph.D. degree from the Massachusetts Institute of Technology with Professor JoAnne Stubbe, and was supported by NSF–NATO and NIH Fellowships for postdoctoral studies in the laboratories of Dr. Daniel Mansuy (Université René Descartes, Paris, France) and Professor Perry Frey (Institute for Enzyme Research, University of Wisconsin–Madison), respectively. In 1999 he moved to The Pennsylvania State University, where he is currently an associate professor of chemistry and an associate professor of biochemistry and molecular biology. His research is concerned with understanding the mechanism and regulation of lipoic acid biosynthesis, as well as the use of iron–sulfur clusters as cofactors or cosubstrates in enzymatic catalysis.



## 7.08 Genomics and Enzymology of NAD Biosynthesis

Leonardo Sorci, Oleg Kurnasov, Dmitry A. Rodionov, and Andrei L. Osterman, Burnham Institute for Medical Research, La Jolla, CA, USA

© 2010 Elsevier Ltd. All rights reserved.

---

7.08.1	Introduction	213
7.08.2	Biochemical Transformations and Enzymes	214
7.08.2.1	Synthesis of Quinolinate	218
7.08.2.1.1	Asp–DHAP pathway	219
7.08.2.1.2	Kynurenine pathway	221
7.08.2.2	Pyridine Mononucleotides	226
7.08.2.2.1	Quinolinate phosphoribosyltransferase	227
7.08.2.2.2	Nicotinate phosphoribosyltransferase	227
7.08.2.2.3	The nicotinamide phosphoribosyltransferase	228
7.08.2.2.4	Nicotinamidase	229
7.08.2.2.5	NMN deamidase	229
7.08.2.2.6	Ribosyl nicotinamide kinase	230
7.08.2.2.7	Ribosyl nicotinamide phosphorylase	232
7.08.2.3	Formation of Pyridine Dinucleotides	233
7.08.2.3.1	NaMN adenylyltransferase of the NadD family in bacteria	234
7.08.2.3.2	NMN adenylyltransferase of the NadR family in bacteria	234
7.08.2.3.3	NaMN- and NMN-specific adenylyltransferases of the NadM family in bacteria and archaea	235
7.08.2.3.4	NaMN- and NMN-specific adenylyltransferases in eukaryotes	236
7.08.2.4	Amidation of Pyridine Nucleotides	237
7.08.2.4.1	NAD synthetase	238
7.08.2.4.2	NMN synthetase	239
7.08.2.5	NADP Formation	240
7.08.2.5.1	NAD kinase	240
7.08.3	Genomic Reconstruction of NAD Biosynthetic Pathways	241
7.08.3.1	Bacteria	241
7.08.3.1.1	Biosynthetic pathways	241
7.08.3.1.2	Transcriptional regulation	244
7.08.3.2	Archaea	246
7.08.3.3	Eukaryota	248
References		251

---

### 7.08.1 Introduction

Nicotinamide adenine dinucleotides, NAD and NADP, are indispensable cofactors involved in several redox reactions in all forms of cellular life. In addition, NAD is utilized as a co-substrate in a number of nonredox reactions playing an important role in signaling and regulatory pathways, especially in eukaryotic cells.<sup>1</sup> Although the key steps of NAD biosynthesis in model species were elucidated in classic studies of the 1950–70s (see reviews<sup>2–4</sup>), many genes and mechanistic details associated with respective biochemical transformations remained unknown until recently. Rapid progress in genome sequencing and comparative analysis led to a significant enrichment and refinement of our understanding of NAD biosynthesis in a broad range of species.



A recent burst of interest in NAD metabolism was triggered by groundbreaking discoveries of NAD and related metabolites operating as new links (e.g., mediated by NAD-dependent protein deacetylases of the Sir2/CobB family<sup>5</sup>) between the metabolic (redox) status and major regulatory programs of eukaryotic cells affecting proliferation, differentiation, longevity, and so on. These issues and, in this context, nutritional, therapeutic, and other physiological aspects of NAD metabolism were discussed in a number of insightful reviews,<sup>1,6–9</sup> and are not dealt with in this chapter. Likewise, we only briefly describe those classical enzymes and functional modules (pathways) that were previously reviewed.<sup>4,8,10</sup> Here, we focus on the most recent findings related to genes, enzymes, and pathways directly involved in NAD biosynthesis in various species. Most importantly, we attempt to capture both the modularity and diversity of NAD biosynthetic machinery emerging from the comparative analysis of scores of completely sequenced genomes. Several examples illustrating the application of the subsystem-based approach<sup>11</sup> (Chapter 7.05) for genomic reconstruction of NAD biosynthetic pathways in diverse species from all three domains of life are provided.

In this review, we consider the biosynthesis of NAD as a biochemical task with a substantial but finite number of solutions that can be implemented in the cell. Various feasible solutions (or metabolic scenarios<sup>12</sup>) can be deduced from combining individual biochemical transformations described in a subset of model organisms. The NAD biosynthesis subsystem is a collection of all known protein families (mostly enzymes) that are associated with respective reactions capturing the current knowledge of this field, as briefly outlined in Section 7.08.2 (Table 1). Specific metabolic scenarios and combinations thereof (subsystem variants<sup>13</sup>) in the organisms with completely sequenced and annotated genomes are inferred based on the presence and absence of respective genes, as illustrated by examples in Section 7.08.3. It is important to emphasize that despite the established predictive power of this approach, it provides only a preliminary assessment of a metabolic potential of a given organism. The realization and distribution of respective metabolic fluxes depend on a multitude of factors and regulatory mechanisms that may not be assessed without additional experimental data. However, recent developments in comparative genomics, led to substantial progress in our ability to infer transcriptional regulatory mechanisms in prokaryotes including identification of putative metabolic regulons (co-regulated genes) and transcription factors.<sup>14</sup> Application of this approach allowed us to substantially expand our understanding of transcriptional regulation of NAD in a broad range of divergent bacterial species, as briefly illustrated in Section 7.08.3.1.2.

## 7.08.2 Biochemical Transformations and Enzymes

The main building blocks, intermediates, and modular organization of the NAD *assembly line* are illustrated in Figure 1. Although each of the biochemical transformations was established in one or the other species, there is no single organism where all of them occurred at the same time. The biogenesis of the pyridine moiety of NAD may be accomplished by one of the two *de novo* pathways, from L-aspartate (Asp) or tryptophan (Trp), both generating pyridine-2,3-dicarboxylate (quinolinic acid (Qa)), as described in Section 7.08.2.1. Organisms that lack either pathway (e.g., some microbial pathogens) are fully dependent on the salvage of the pyridine ring supplied by various products of NAD catabolism. Among them, the most common are nicotinamide (Nm) and its deamidated analogue, nicotinic acid (Na), also known as niacin or vitamin B<sub>3</sub>. Both Nm and Na forms of B<sub>3</sub> would suppress pellagra, a human syndrome caused by the insufficient supply of dietary L-Trp, leading to a decreased level of NAD(P). These early observations led to a discovery that the pyridine ring of NAD in mammals may be supplied by either *de novo* synthesis from Trp or by niacin salvage pathways (see Bogan and Brenner<sup>6</sup> for a recent overview). Therefore, in contrast to most other vitamins, niacin is not an indispensable component of the human diet, as long as a sufficient level of proteinaceous Trp is provided. Another interesting distinction of niacin from such vitamins as B<sub>1</sub> (riboflavin) and B<sub>5</sub> (panthotenate), precursors of other universal dinucleotide cofactors, flavin adenine dinucleotide (FAD) and coenzyme A (see Chapters 7.02 and 7.11), is that it cannot be synthesized *de novo* and that its only natural source is the catabolism of the NAD cofactor. Consequently, unlike in cases of B<sub>1</sub> or B<sub>5</sub>, the salvage of exogenous pyridines requires distinct biochemical transformations that are not a part of the *de novo* route. Similar transformations are involved in pyridine recycling, which is particularly important due to the nonredox consumption of NAD by various enzymes hydrolyzing the *N*-glycoside bond between the Nm and ADP-ribose moieties of NAD.

**Table 1** Enzymes involved in NAD biosynthesis and examples of their occurrence in various species

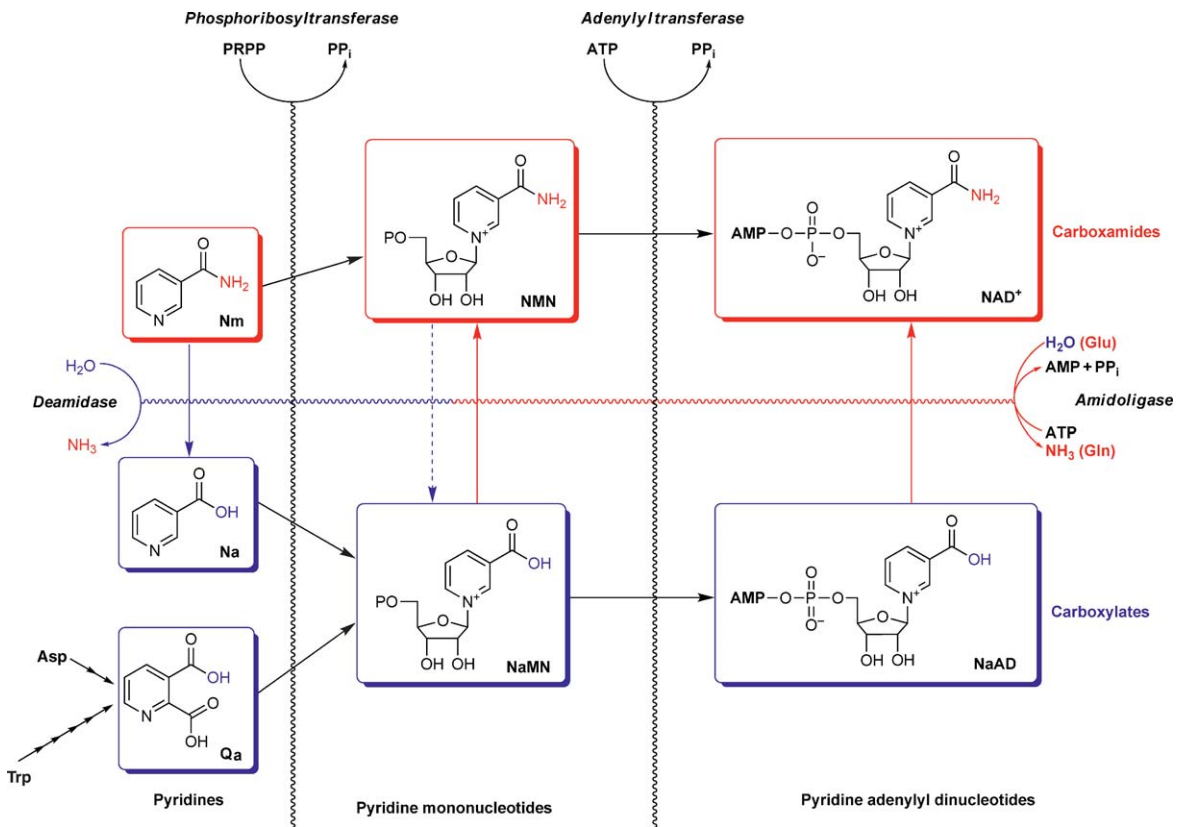
Symbol <sup>a</sup>	Enzyme family <sup>b</sup>	Gene name <sup>c</sup>	EC	FT	SA	HI	MJ	AP	AT	HS
			[B] <sup>d</sup>			[A] <sup>d</sup>		[E] <sup>d</sup>		
<i>Quinolinate synthesis from L-aspartate (Asp–DHAP pathway)</i>										
ASPOX	L-Aspartate oxidase (EC 1.4.3.16)	<i>nadB</i>	+	+	–	–	–	–	+	–
ASPDH	L-Aspartate dehydrogenase (EC 1.4.1.21)		–	–	–	–	+	–	–	–
QASYN	Quinolinate synthetase (EC 4.1.99.–)	<i>nadA</i>	+	+	–	–	+	–	+	–
<i>Quinolinate synthesis from L-tryptophan (Kyn pathway)</i>										
TRDOX	Tryptophan 2,3-dioxygenase (EC 1.13.11.11)	<i>kynA</i>	–	–	–	–	–	–	–	+
	Indoleamine 2,3-dioxygenase (EC 1.13.11.42)		–	–	–	–	–	–	–	+
KYNFA	Kynurenine formamidase (EC 3.5.1.9)		–	–	–	–	–	–	–	+
	Kynurenine formamidase, bacterial (EC 3.5.1.9)	<i>kynB</i>	–	–	–	–	–	–	–	–
KYNOX	Kynurenine 3-monooxygenase (EC 1.14.13.9)	<i>kmo</i>	–	–	–	–	–	–	–	+
KYNSE	Kynureninase (EC 3.7.1.3)	<i>kynU</i>	–	–	–	–	–	–	–	+
HADOX	3-Hydroxyanthranilate 3,4-dioxygenase (EC 1.13.11.6)	<i>haoO</i>	–	–	–	–	–	–	–	+
<i>Formation of pyridine mononucleotides</i>										
QAPRT	Quinolinate phosphoribosyltransferase (EC 2.4.2.19)	<i>nadC</i>	+	+	–	–	+	–	+	+
NAPRT	Nicotinate phosphoribosyltransferase (EC 2.4.2.11)	<i>pncB</i>	+	–	+	–	–	+	+	+
NMPRT	Nicotinamide phosphoribosyltransferase (EC 2.4.2.12)	<i>nadV</i>	–	+	–	–	–	–	–	+
NMASE	Nicotinamide deamidase (EC 3.5.1.19)	<i>pncA</i>	+	–	+	–	–	+	+	–
RNMKIN	Ribosylnicotinamide kinase (EC 2.7.1.22)	<i>nadR</i>	+	–	–	+	–	–	–	–
	Ribosylnicotinamide kinase, eukaryotic (EC 2.7.1.22)		–	–	–	–	–	–	–	+
<i>Formation of pyridine dinucleotides</i>										
NaMNAT	NaMN adenyllyltransferase, NadD family (EC 2.7.7.18)	<i>nadD</i>	+	–	+	–	–	–	–	–
	NaMN adenyllyltransferase, NadM family (EC 2.7.7.18)	<i>nadM</i>	–	–	–	–	+	+	–	–
	NaMN adenyllyltransferase, eukaryotic (EC 2.7.7.18)		–	–	–	–	–	–	+	+
NMNAT	NMN adenyllyltransferase, NadR family (EC 2.7.7.1)	<i>nadR</i>	+	–	–	+	–	–	–	–
	NMN adenyllyltransferase, NadM family (EC 2.7.7.1)	<i>nadM</i>	–	+	–	–	+	+	–	–
	NMN adenyllyltransferase, eukaryotic (EC 2.7.7.1)		–	–	–	–	–	–	+	+
<i>Amidation of pyridine nucleotides</i>										
NMNSYN	NMN synthetase (EC 6.3.1.–)	<i>nadE'</i>	–	+	–	–	–	–	–	–
NADSYN	NAD synthetase (EC 6.3.1.5)	<i>nadE</i>	+	–	+	–	+	+	+	+
	Glutamine amidotransferase chain of NADSYN		–	–	–	–	–	–	+	+
<i>Formation of NADP</i>										
NADKIN	NAD kinase (EC 2.7.1.23)	<i>nadF</i>	+	+	+	+	+	+	+	+

<sup>a</sup> Abbreviations (symbols) used for enzymes and respective enzymatic reactions are the same as in **Figure 2** and in the individual reaction diagrams (**Schemes 1–7**). Note that the same symbols may be used for distinct enzyme families as long as they catalyze the same reaction in NAD synthesis.

<sup>b</sup> Complete enzyme names and enzyme classification numbers with an indication of a protein family for those enzymatic roles that are known to be encoded by more than one protein family (nonorthologous displacements). Note that some protein families may have more than one relevant enzymatic role, either due to a broad specificity (e.g., NadM family having both NMNAT and NaMNAT activities) or due to the presence of more than one functional domain (e.g., NadR family containing NMNAT and RNMKIN domains).

<sup>c</sup> Common bacterial gene names originating mostly from *E. coli*/*Salmonella* (except for *nadV* and *nadM*).

<sup>d</sup> Examples of presence or absence (shown by '+' or '-') of genes encoding respective enzyme families are illustrated for several representative species of bacteria [B]: EC, *Escherichia coli*; FT, *Francisella tularensis*; SA, *Staphylococcus aureus*; and HI, *Haemophilus influenzae*; archaea [A]: *Methanococcus jannaschii* and *Aeropyrum permix*; and eukaryota [E]: *Arabidopsis thaliana* and *Homo sapiens*. Variations in NAD biosynthetic pathways implemented in these species are reflected in simplified diagrams in **Figures 3, 5, and 6**.



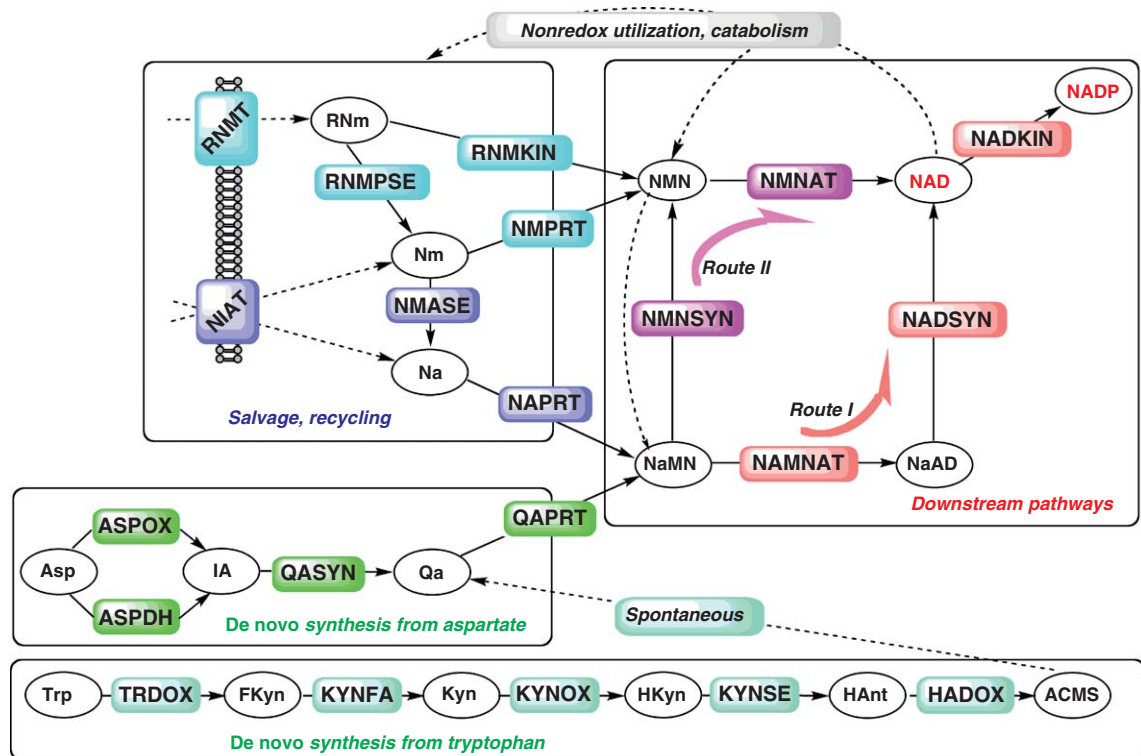
**Figure 1** Major biochemical transformations involved in NAD assembly in various species. (I) *Pyridine precursors*. Quinolinic acid (Qa) is synthesized *de novo* via one of two distinct pathways, from L-aspartate (Asp) or L-tryptophan (Trp). Nicotinic acid (Na) and nicotinamide (Nm) are the products of NAD catabolism that can be salvaged and recycled by many species. (II) *Pyridine mononucleotides*. Pyridine precursors are converted into nicotinic acid and nicotinamide mononucleotides (NaMN and NMN, respectively) via the reaction with phosphoribosyl pyrophosphate (PRPP) catalyzed by one of the three specific phosphoribosyltransferase enzymes. (III) *Pyridine dinucleotides*. Nicotinamide adenine dinucleotide (NAD) and its deamidated precursor nicotinic acid adenine dinucleotide (NaAD) are generated by respective adenylyltransferase reactions. (IV) *Amidation* of a pyridine carboxylate moiety (shown by red arrows) by an ATP-dependent amidoligase (using ammonia (NH<sub>3</sub>) or glutamine (Gln) as N-donor) typically involves NaAD or, more rarely, NaMN, but not Na or Qa. (V) *Deamidation* of pyridine carboxamide group (shown by blue arrows), a common first step in Nam salvage, was also reported to occur for the NMN intermediate, but the gene encoding the respective enzyme remains unknown. Following the arrows in this diagram, one may reconstruct *all known scenarios* of converting pyridine precursors to NAD, for example, a classical Preiss–Handler pathway: Na → NaMN → NaAD → NAD, a recently identified nondeamidating Nm salvage: Nm → NMN → NAD. Moreover, *all formally possible scenarios* (that can be deduced from this diagram) appear to be actually implemented in one or another subset of species with sequenced genomes (see Section 7.08.3 for examples).

The next stage of NAD assembly from pyridine precursors is a phosphoribosyltransferase reaction generating one of the two pyridine mononucleotide intermediates (Figure 1). Both deamidated precursors of *de novo* and salvage routes (Qa and Na, respectively) yield nicotinic acid mononucleotide (NaMN). The salvage of Nm may proceed via direct phosphoribosylation yielding nicotinamide mononucleotide (NMN) or via deamidation to Na followed by its phosphoribosylation to NaMN. Each of the three phosphoribosylation reactions are catalyzed by committed enzymes, quinolinate phosphoribosyltransferase (QAPRT), nicotinate phosphoribosyltransferase (NAPRT), and nicotinamide phosphoribosyltransferase (NMPRT), which belong to a large class of phosphoribosyl pyrophosphate (PRPP) transferases originating from a common evolutionary ancestor and sharing some structural and mechanistic features<sup>15</sup> (Section 7.08.2.2). Both types of possible interconversions between pyridine mononucleotides, (1) NMN deamidation to NaMN operating as a part of salvage machinery in bacteria<sup>16,17</sup> and (2) ATP-dependent amidation of NaMN to NMN involved in the novel Route II of NAD synthesis (see below), were described. Biochemical transformations associated with salvage and recycling of

pyridine nucleosides (not shown in **Figure 1**) include: (1) phosphorylation of ribosyl nicotinamide (RNm) or ribosyl nicotinate (RNa), yielding NMN or NaMN, respectively, and (2) NMN phosphorylase reaction yielding Nm and ribose-1-phosphate (Section 7.08.2.2).

The completion of the NAD scaffold is attained by the adenylation of pyridine mononucleotides (**Figure 1**). Representatives of four distinct adenylyltransferase families identified in various species differ in their relative substrate preferences toward NaMN vs. NMN. Despite a substantial difference at the level of primary structure, they all belong to one ‘HIGH’ superfamily of nucleotidyltransferases<sup>18</sup> sharing the overall fold and many aspects of their catalytic mechanisms (Section 7.08.2.3). The NaMN adenylyltransferase (NAMNAT) reaction is a committed and usually indispensable step of the most common *de novo* and salvage routes of NAD synthesis that proceed via the NaMN intermediate. This reaction generates another committed intermediate, nicotinic acid adenine dinucleotide (NaAD or deamido-NAD), that undergoes ATP-dependent amidation at the last step of NAD synthesis via the conventional Route I: (NaMN → NaAD → NAD, see **Figure 2**). The NMN adenylyltransferase (NMNAT) activity is usually associated with nondeamidating routes of salvage or recycling of Nm or RNm via the NMN intermediate. Recent studies in *Francisella tularensis* revealed that this activity may also be involved in the alternative Route II of *de novo* NAD synthesis (NaMN → NMN → NAD, see **Figure 2**).

The amidation of the pyridine carboxylate moiety is the last critical stage of the NAD assembly line (Section 7.08.2.4). Traditionally, this reaction was assumed to occur only at the level of a dinucleotide intermediate NaAD catalyzed by the enzyme NAD synthetase (NADSYN). However, it was recently shown that at least some members of the same enzyme family possess an unusual NMN synthetase activity (NMNSYN) efficiently catalyzing amidation of a mononucleotide intermediate NaMN.<sup>19</sup> As mentioned earlier, NMNSYN together



**Figure 2** NAD biosynthesis subsystem diagram. Major functional roles are shown by 4–6 letter abbreviations (explained in **Table 1**) over the colored background reflecting the key aspects or modules (pathways) that comprise NAD biosynthesis in various species. Catalyzed reactions are shown by solid straight arrows, and corresponding intermediate metabolites are shown as abbreviations within ovals: Asp, L-aspartate; IA, iminoaspartate; Qa, quinolinic acid; Nm, nicotinamide; Na, nicotinic acid; NaMN, nicotinic acid mononucleotide; NMN, nicotinamide mononucleotide; RNm, N-ribosylnicotinamide; NaAD, nicotinate adenine dinucleotide; NAD, nicotinamide adenine dinucleotide; NADP, NAD-phosphate; Trp, tryptophan; FKyn, N-formylkynurenine; Kyn, kynurenine; HKyn, 3-hydroxykynurenine; HAant, 3-hydroxyanthranilate; and ACMS,  $\alpha$ -amino- $\beta$ -carboxymuconic semialdehyde. Unspecified reactions (including spontaneous transformation and transport) are shown by dashed arrows.

with NMNAT enzyme comprise a relatively rare Route II of NAD synthesis in some bacteria that lack the most common NAMNAT activity. The existence of NMNSYN activity, no matter how unexpected, is not surprising as the chemical properties of pyridine carboxylate moiety would be quite similar in both mono- and dinucleotide substrates. As this is not true for a nonribosylated pyridine ring, the existence of the enzymatic amidation of Na is not anticipated, and no such activity was ever detected. Although all NADSYN/NMNSYN enzymes belong to the same family conserved in all three domains of life, they can be divided into two large groups based on the presence or absence of an additional glutaminase domain. Only those enzymes that contain this additional domain (e.g., those present in all Eukarya) can utilize glutamine (Gln) as well as ammonia ( $\text{NH}_3$ ) as a source of amide nitrogen. Enzymes of the second group (e.g., those present in all archaea) can utilize only ammonia.

Despite an apparent complexity and variations within the NAD subsystem, one may notice a remarkable symmetry of the key biochemical transformations from pyridines to NAD (Figure 1). Transitions between the three vertical planes correspond to the three stages in the biogenesis of the dinucleotide scaffold, whereas transitions between the two horizontal planes capture interconversions between the carboxylate and carboxamide forms of the respective compounds. Figure 1 infers metabolic scenarios that may be implemented in different species based on the presence and absence of genes encoding respective enzymes (e.g.,  $\text{Qa} \rightarrow \text{NaMN} \rightarrow \text{NaAD} \rightarrow \text{NAD}$  as in Route I of the *de novo* synthesis or  $\text{Nm} \rightarrow \text{NMN} \rightarrow \text{NAD}$  as in the nondeamidating salvage or recycling of Nm, etc.). Although many individual reactions in this diagram were originally described in distinct organisms, nearly all of their combinations within theoretically possible scenarios appear to actually occur in one or another subset of species. This observation from the comparative analysis of more than 700 completely sequenced genomes (Section 7.08.3) provided us with a remarkable illustration of G. Hegel's maxim that indeed, "What is rational is actual, and what is actual is rational."

A list of characterized enzyme families involved in NAD biogenesis in various species is provided in Table 1, and their involvement in various pathways that comprise this subsystem is illustrated in Figure 2. Conversion of NAD into its phosphorylated form NADP is brought about by NAD kinase (NADKIN), the most conserved enzyme in the entire subsystem, which is present even in intracellular bacterial pathogens such as *Rickettsia* spp. that lack the rest of the NAD biosynthetic machinery. Although both cofactors, NAD and NADP, are involved in mechanistically very similar hydride transfer reactions, they play quite distinct roles in the cellular metabolism. It is widely known that most of the redox enzymes utilizing NAD cofactor are involved in catabolic and energy-generating pathways, whereas NADP-utilizing enzymes play the key role in the biosynthesis of major building blocks in the cell, such as amino acids and nucleotides. This distinction provides a foundation for the global decoupling of the catabolic and anabolic aspects of cellular networks. The delicate balance between all components of the NAD cofactor pool,  $\text{NAD}^+$ , NADH,  $\text{NADP}^+$ , and NADPH, is a product of multiple reactions collectively defining the overall metabolic status of the cell.<sup>20,21</sup> The discussion of these reactions as well as those involved in nonredox consumption and utilization of NAD are outside the scope of this chapter, which is focused largely on genomic reconstruction of NAD biogenesis in various species. Before proceeding to this central topic, we briefly describe individual reactions and enzymes directly involved in NAD *de novo* biosynthesis, as well as in various salvage or recycling pathways.

### 7.08.2.1 Synthesis of Quinolinatate

The two completely unrelated pathways of Qa synthesis, from Asp and from Trp (Schemes 1 and 2), appear to be of independent evolutionary origin. Only a few other examples of convergent evolution at the level of the entire pathway are known, such as the biosynthesis of isoprenoids (deoxyxylulose phosphate vs. mevalonate pathway) and lysine (diaminopimelate vs. amino adipate pathway). The phylogenetic distribution of such alternating pathways tends to reflect the division between the three domains of life. Indeed, the Asp-dihydroxyacetone phosphate (Asp-DHAP) pathway of Qa synthesis occurs in bacteria and archaea but not in Eukaryotes that typically synthesize Qa from Trp. Among eukaryotes, plants often carry bacterial versions of alternating pathways. For example, *Arabidopsis thaliana* harbors bacterial versions of Qa and lysine biosyntheses (diaminopimelate pathway) instead of their typical eukaryotic counterparts. The synthesis of Qa from Trp, also referred to as the kynurenine (Kyn) pathway, was extensively studied, in part because of its association with a broader set of Kyn-related compounds stemming from Trp degradation and strongly



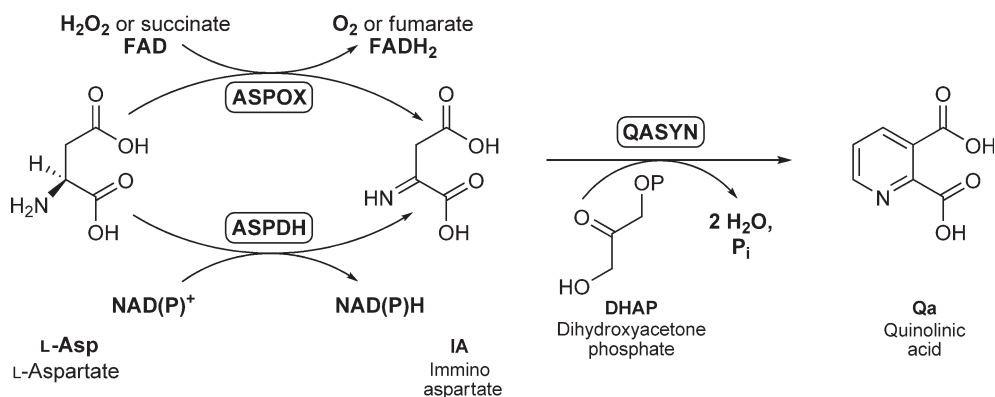
implicated in a number of neurological disorders.<sup>22</sup> Despite the early observations pointing to the existence of the Kyn pathway of Qa synthesis in *Xanthomonas* spp.,<sup>23</sup> until recently this pathway was considered strictly eukaryotic. A comparative genome analysis revealed the presence of homologues of respective genes in many bacterial species, and the proposed functional activity of these genes and of the entire pathway were experimentally confirmed.<sup>24</sup> In some of these species (including members of the Xanthomonadales group), aerobic Trp degradation is the only route of Qa synthesis.

### 7.08.2.1.1 Asp–DHAP pathway

The conversion of Asp to Qa is performed in two steps (Scheme 1), oxidation of L-Asp to an unstable intermediate, iminoaspartate (IA), followed by the condensation of IA with DHAP, a central glycolytic intermediate, to form Qa. The latter is further converted into NaMN by the QAPRT enzyme, and all three enzymatic steps are often considered single *de novo* pathway as the three respective genes (*nadB*, *nadA*, and *nadC*) have a strong tendency to co-occur (*all or none*) and form operons and regulons in prokaryotic genomes. However, since the third enzyme (QAPRT) is shared by Asp–DHAP and Kyn pathways beyond the point of their convergence at the Qa precursor, we describe it in Section 7.08.2.2 together with two other phosphoribosyltransferases (NAPRT and NMPRT).

Two nonhomologous and mechanistically distinct enzymes can catalyze the first reaction of the Asp–DHAP pathway (Figure 2), a classic flavin-dependent L-Asp oxidase (ASPOX)<sup>25</sup> and a NAD-dependent aspartate dehydrogenase (ASPDH) recently characterized in *Thermotoga maritima*.<sup>26</sup> Both enzymes belong to large superfamilies that contain representatives with different substrate specificities. The second reaction is catalyzed by the universally conserved enzyme quinolinate synthetase (QSYN). All members of the respective superfamily with a unique fold (NadA-like by SCOP classification<sup>27</sup>) appear to have the same functional activity. Therefore, this protein family may be considered a genuine signature of the Asp–DHAP pathway. The presence of a *nadA* gene homologue in a genome (always accompanied by *nadC*) unambiguously indicates the presence of the Asp–DHAP pathway even if no gene candidates for the first step can be found (see below).

**7.08.2.1.1(i) L-Aspartate oxidase** A so-called B-protein of the quinolinate synthetase enzymatic complex (composed of A and B proteins) was originally purified from *Escherichia coli* and characterized as ASPOX enzyme.<sup>25</sup> This approximately 60-kDa protein dimerized in solution and contained a single molecule of FAD cofactor involved in the oxidation of Asp to IA (Scheme 1). The respective gene (*nadB*) was identified by genetic techniques and mapped in *Salmonella typhimurium*,<sup>28,29</sup> and was later cloned from *E. coli*.<sup>30</sup> Homologous genes were identified and characterized in other bacteria,<sup>31,32</sup> archaea,<sup>33,41</sup> and plants.<sup>34</sup> Among the important enzymatic characteristics deduced mostly from the studies of the *E. coli* ASPOX are its stringent stereospecificity toward Asp (in contrast to previously described representatives of a structurally unrelated family of D-amino acid oxidases) and the observed inhibition by NAD, which appears to efficiently compete with FAD cofactor for the same binding site. The latter observation provided a rationale for the previously proposed role of NadB as a point of NAD biosynthesis feedback regulation.<sup>25,35</sup> However, this regulatory mechanism may not



Scheme 1



be universal, as FAD binding was reported to be much more tightly bound in ASPOX than the hyperthermophilic archaeon *Pyrococcus horikoshii* OT-3.<sup>33</sup> An ability of ASPOX to efficiently utilize both oxygen or fumarate as electron acceptors *in vitro* was confirmed for both bacterial and archaeal enzymes.<sup>33,36,37</sup> This unusual versatility with respect to electron acceptors explains how this pathway may operate in both aerobic and anaerobic conditions. Moreover, several lines of evidence, including a structural and mechanistic analogy with ASPOX to fumarate reductase, indicate that fumarate (rather than oxygen) is the most likely co-substrate of ASPOX, even in aerobic conditions.

ASPOX shares substantial sequence similarity with FAD-containing subunits of fumarate reductase and succinate dehydrogenase complexes. However, unlike these multisubunit enzymes, ASPOX does not seem to require a Fe-S cluster partner, directly coupling the reduction of fumarate to oxidation of Asp.<sup>38</sup> The 3D structural analysis of the *E. coli* ASPOX revealed the presence of three domains, an N-terminal dinucleotide fold domain with an insertion of an ( $\alpha + \beta$ ) capping domain and a C-terminal helical domain.<sup>39</sup> Binding of the FAD cofactor causes substantial conformational changes leading to the formation of the active site at the interface between the capping and FAD-binding domains. Several residues contributing to substrate and cofactor binding are conserved between ASPOX and fumarate reductase supporting the proposed catalytic mechanism that includes a hydride transfer from the Asp C $\alpha$  or C $\beta$  atom to FAD in conjunction with proton abstraction mediated by the Arg-290.<sup>40</sup> This is radically different from a classical D-amino acid oxidase mechanism, where the electron transfer involves an  $\alpha$ -amino group of a substrate. A recently published 3D structure of ASPOX from another hyperthermophilic archaeon, *Sulfolobus tokodaii*, supports most of these conclusions while providing a structural rationale for the observed thermostability of the archaeal enzyme.<sup>41</sup>

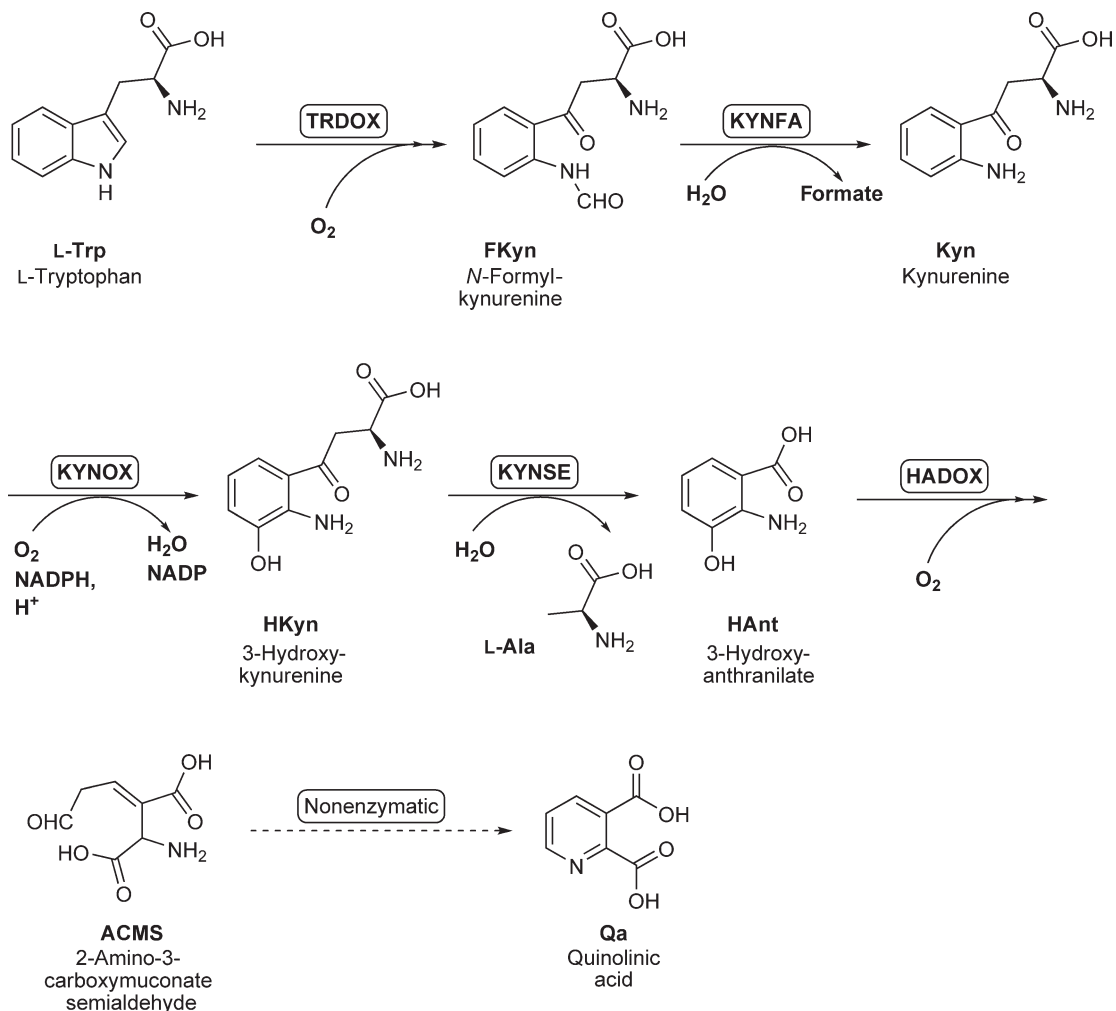
**7.08.2.1(ii) L-Aspartate dehydrogenase** ASPDH was originally characterized as an approximately 26-kDa product of the *T. maritima* gene (TM1643) from a three-gene operon that also encoded QSYN and QAPRT enzymes.<sup>26</sup> ASPDH efficiently converts Asp into IA using NAD or NADP as a cofactor (**Scheme 1**), but it has no oxidase activity. The observed enzymatic activity and specificity, along with the genomic context and the apparent absence of a classical ASPOX-encoding gene in *T. maritima*, led to a conclusion that ASPDH is a nonhomologous replacement of the ASPOX enzyme in Qa synthesis. A similar pattern of gene occurrence and chromosomal clustering was observed in a number of archaeal genomes (including all methanogenes and *Archaeoglobus fulgidus*) supporting the proposed functional role for the archaeal homologues of TM1643. This conclusion was confirmed by the experimental characterization of ASPDH from *A. fulgidus*.<sup>42</sup> The mechanistic analysis of this enzyme revealed the pro-R stereospecificity of hydride transfer from the C-4 position of the reduced NAD. Despite a mechanistic similarity, ASPDH shows no appreciable structural similarity with other pro-R stereospecific amino acid dehydrogenases, thus representing a new enzyme family. The 3D structure of *T. maritima* ASPDH revealed the presence of a dimer with each subunit composed of two domains, the N-terminal Rossmann fold domain with bound NAD-cofactor and a C-terminal ( $\alpha + \beta$ ) domain.<sup>26</sup> A recently published 3D structure of ASPDH from *A. fulgidus* suggested that the substrate binding (mimicked by the bound citrate molecule) occurs in the C-terminal domain. It is accompanied by a substantial movement of the two surface loops, bringing the substrate into close proximity with the NAD cofactor.<sup>43</sup>

Quinolinate synthase (QSYN) was initially described as the A-protein of the proposed NadAB enzymatic complex, although further studies failed to confirm the complex formation between ASPOX and QSYN. Cloning of the *nadA* gene from *E. coli* and characterization of its approximately 38-kDa product confirmed that it encoded a functional QSYN enzyme converting IA and DHAP to Qa (**Scheme 2**).<sup>30</sup> In the early enzymatic studies, QSYN was found to be functionally unstable, likely due to the presence of the Fe-S cluster, a subject of oxidative damage *in vitro* and, possibly, *in vivo*. Later, the functional activity of the recombinant NadA was successfully reconstituted *in vitro* in the absence of Fe-S clusters that did not seem to be directly involved in the QSYN mechanism.<sup>44</sup> At the same time, the importance of Fe-S biogenesis and regeneration for *in vivo* Qa biosynthesis is supported by many independent studies.<sup>45-47</sup> In addition to the studies in the *E. coli* model, the early genetic data in *Bacillus subtilis* indicated the functional importance of the *nifS* gene divergently transcribed and coregulated with the *nadBCA* operon.<sup>32,48</sup> The most remarkable evidence was recently reported for the QSYN enzyme from *A. thaliana*, whose additional SufE-like N-terminal domain was shown to be responsible for the regeneration of the highly oxygen-sensitive Fe-S cluster located at the C-terminal NadA enzymatic domain.<sup>49</sup> The actual role of Fe-S clusters in QSYN catalysis remains a subject of active research.

Similarly, the details of the QSYN catalytic mechanism, a combination of the two condensation reactions that include the elimination of two water molecules and a phosphate moiety of DHAP, are yet to be elucidated.<sup>10,25</sup> The 3D structure was so far reported only for a single representative of the QSYN family from *P. borikosbii* that was crystallized in a complex with malate presumably mimicking the IA substrate.<sup>50</sup> The protein has a unique fold with a pseudo-threefold symmetry composed of the three analogous  $\alpha/\beta/\alpha$  sandwiches, possibly reflecting two ancient gene duplication events. Despite limited mechanistic insight provided by this structure, the authors tend to support one of the previously proposed mechanisms, which includes phosphate elimination during the first condensation at the C-3 position of DHAP.<sup>50</sup> However, this interpretation is based largely on structural modeling and assumes substantial structural rearrangements that are yet to be demonstrated.

### 7.08.2.1.2 Kynurenine pathway

Formation of Qa via aerobic degradation of Trp (Kyn pathway) includes five enzymatic steps: (1) oxidation of Trp to *N*-formyl kynurenine (FKyn) by Trp 2,3-dioxygenase (TRDOX), (2) deformylation of FKyn by kynurenine formamidase (KYNFA), (3) oxidation of Kyn to 3-hydroxykynurenine (HKyn) by kynurenine 3-monooxygenase (KYNOX), (4) conversion of HKyn into 3-hydroxyanthranilate (HAnt) by kynureninase (KYNSE), and (5) oxidation of HAnt by 3-hydroxyanthranilate 3,4-dioxygenase (HADOX) to  $\alpha$ -amino- $\beta$ -carboxymuconic semialdehyde (ACMS) followed by its spontaneous cyclization to Qa (Scheme 2). This pathway and all respective



**Scheme 2**

enzymes were originally described in eukaryotes, mostly in human and yeast. In humans, various products of Trp aerobic degradation (kynurenines) are involved in a number of signaling pathways. Redistribution of fluxes between different branches of the pathway have major effects on the nervous and immune systems and may lead to a variety of severe pathological conditions. For example, the accumulation of Qa, a neurotoxin interacting with NMDA receptors in the brain, is associated with dementia and other neurological syndromes.<sup>22</sup> Physiological and pharmacological aspects of Kyn-related pathways are a subject of active biomedical research.<sup>51,52</sup> We limit the scope of our discussion to biochemical transformations immediately associated with Qa synthesis.

Although the Kyn pathway of Qa synthesis was considered a eukaryotic signature, it is not present in all branches of Eukaryota. Comparative genome analysis confirms the presence of the entire pathway in vertebrates and nematodes but not in insects. The latter group typically has only the first three enzymes of the pathway converting Trp into Hkyn, which is utilized for eye pigmentation and is further converted into xanthurenic acid by kynurenine transaminase.<sup>53</sup> Unlike mammals that can generate NAD from both Trp or niacin, insects are fully dependent on the salvage of pyridines from the diet via the NAPRT route (Preiss–Handler pathway), which appears to be conserved in all analyzed eukaryotes. Fungi reveal a patchy distribution of the Kyn pathway. For example, this pathway is present in *Saccharomyces cerevisiae*, *Candida albicans*, and *Yarrowia lipolytica* but not in *Schizosaccharomyces pombe*, *Candida glabrata*, and *Kluyveromyces lactis*. The niacin salvage pathway is present in all these species, but only the latter three are niacin auxotrophs from the set analyzed.<sup>54</sup> It is important to emphasize that in contrast to the Asp–DHAP pathway, the Kyn pathway is strictly aerobic. For example, under anaerobic conditions, *S. cerevisiae* is a Na auxotroph.<sup>55</sup> Among plants, the monocots (such as rice) seem to have both, Kyn and Asp–DHAP pathways of Qa synthesis, whereas dicots (such as *A. thaliana*) have only the latter pathway.<sup>56</sup>

Comparative genome analysis revealed an interesting distribution of Trp aerobic degradation genes in several groups of bacteria (but not in archaea). A number of species within at least two groups, Xanthomonadales (*Xanthomonas* but not *Xylella* spp.) and Bacteroidetes (*Flavobacteria* and *Cytophaga* spp.), appear to use the Kyn pathway instead of the Asp–DHAP pathway for Qa synthesis. All these species contain orthologues of the four out of five eukaryotic genes of the Kyn pathway except for KYNFA, a likely subject of nonorthologous gene replacement that remains to be elucidated. All enzymes of the bacterial Kyn pathway (including a version of KYNFA that is present in some bacteria<sup>57</sup> but not in the two groups mentioned above) were experimentally characterized, and Qa production was confirmed by genetic complementation in *E. coli*.<sup>24</sup> A complete Kyn pathway was reported as a single possible route to Qa in *Gemmata* str.<sup>24</sup> This observation is particularly intriguing especially the observed analogy between membrane-bound nucleoids characteristic of planctomycetes (such as *Gemmata*) and the eukaryotic nuclear membrane, which prompted a discussion of a possible (albeit unconfirmed) role of this bacterial group in the origin of eukaryotes.<sup>58</sup> In some bacterial species (e.g., in *Shewanella woodyi*, *Acidobacteria bacterium*, and *Salinispora* spp.), a complete set of Kyn pathway genes is present in addition to the Asp–DHAP pathway, although whether these genes are involved in the synthesis of Qa remains to be tested. For example, it was shown that a large gene cluster present in *Burkholderia cepacia* encodes a catabolic pathway that ultimately converts Trp into pyruvate and acetate.<sup>59</sup> At least two of the three species listed above (*S. woodyi* and *Salinispora* spp.) seem to have the entire set of genes from this pathway, and most of these genes are located within the same chromosomal cluster. Finally, many diverse bacteria contain incomplete sets of Kyn pathway enzymes that are insufficient for Qa synthesis, likely leading to other products of Trp degradation. For example, *Pseudomonas aeruginosa*, *Burkholderia mallei*, *Bacillus anthracis*, *Sphingopyxis alaskensis*, and several other species contain a conserved gene cluster encoding TRDOX, bacterial KYNFA, and KYNSE enzymes converting Trp into anthranilate.<sup>57</sup> In *Pseudomonas fluorescens*, these genes together with kynurenine transaminase were implicated in the synthesis of xanthurenic acid, a precursor of the siderophore quinolobactin.<sup>60</sup> Some of the Kyn pathway enzymes were implicated in the degradation of *o*-nitrobenzoate by another strain of *P. fluorescens* KU-7 and by *Arthrobacter protophormiae*.<sup>61,62</sup>

These observations indicate that the Kyn pathway may not be an entirely ‘eukaryotic invention’. It is more likely that the respective enzymes emerged in aerobic bacteria as tools for Trp catabolism, and they were not initially aimed at Qa synthesis, which was provided by a more-ancient (originally anaerobic) Asp–DHAP pathway. A recruitment of these genes for Qa synthesis (followed by a loss of the Asp–DHAP pathway) could have occurred in the environmental conditions strongly favoring the aerobic pathway. Finally, the entire set could have been acquired by a common ancestor of eukaryotes. In the course of evolution and speciation, this pathway could be lost entirely (as in some fungi and monocotyledonous plants) or partially (as in insects), while

the remaining genes could be adapted to other biological functions (such as pigmentation). On the other hand, mammals appear to have expanded the Kyn pathway, retaining its function in NAD synthesis but also engaging it in various signaling and regulatory systems. A historic perspective and basic properties of human enzymes involved in the synthesis of Qa and NAD were reviewed.<sup>1,8</sup> Therefore, in this section we mostly reflect upon the recent progress in structural and mechanistic analysis of eukaryotic enzymes and their bacterial counterparts.

**7.08.2.1.2(i) Tryptophan 2,3-dioxygenase** Representatives of the two distantly related families of approximately 40-kDa heme-containing enzymes that catalyze the first reaction of the Kyn pathway were characterized in mammals. The enzymes of the first family (here termed the TDO family) originally isolated from liver have a stringent specificity to L-Trp.<sup>63</sup> The second characterized enzyme<sup>64</sup> was termed indoleamine 2,3-dioxygenase (we refer to it as the IDO family) as it can oxidize a broader range of Trp analogues including D-Trp, tryptamine, serotonin, and so on. Representatives of both the TDO and IDO families are present in mammals, but they have different distribution patterns in other taxonomic groups. For example, insects, nematodes, and most bacteria that harbor one or another version of the Trp degradation pathway have only TDO-encoding genes, whereas yeast and other fungi use IDO in their version of the Kyn pathway.

Despite extensive physiological studies of TRDOX enzymes, their mechanistic details remained largely unexplored until recently, when representative 3D structures from both families, including IDO from human and *Shewanella oneidensis* and TDO from *Ralstonia metallidurans* and *Xanthomonas campestris*, were reported.<sup>65–67</sup> Despite a substantial difference in their amino acid sequences, members of both families share a common core domain composed of 10  $\alpha$ -helices, and the heme-binding pocket is mostly conserved between the two structures. The mechanism of the oxidative ring opening at the C2–C3 position of the indole moiety in TRDOX reaction is significantly different from the well-studied reaction catalyzed by the nonheme dioxygenases. Although the initial stage of TRDOX reaction catalysis, formation of Trp hydroperoxide by a single electron transfer, appears to be the same for TDO and IDO, further mechanistic steps may be quite different, reflecting the differences in their structures and substrate selectivity. In the case of human IDO, structural modeling and site-directed mutagenesis data argue in favor of the intermediate dioxyethane formation that does not require any interactions between the substrate and protein amino acid residues, which lead to an observed promiscuity toward various indole-containing substrates.<sup>66</sup> A detailed structure–function study of the bacterial TDO provided evidence in favor of the alternative Criegee mechanism of hydroperoxide rearrangement of the indole ring opening without dioxyethane formation.<sup>67</sup> This mechanism assumes strong interactions of amino and carboxy groups of the substrate with the enzyme leading to its high selectivity for Trp. This interpretation is supported by comparative modeling of two possible conformations (syn- and anti-) of the hydroperoxide intermediate as well as by the results of the directed mutagenesis of the three residues implicated in interactions with the substrate. The structural and biochemical analysis of another bacterial representative of the TDO family revealed a substantial structural rearrangement (induced fit) of the enzyme upon binding of the Trp substrate that leads to a complete exclusion of solvent from the active site area.<sup>65</sup> Despite these important insights, all currently proposed mechanistic models remain tentative and further analysis is required to fill the lacunae in our understanding of the TRDOX reactions.

**7.08.2.1.2(ii) Kynurenine formamidase** Although the KYNFA enzymatic activity (also referred to as aryl formamidase) associated with the second step of the Kyn pathway was described in early studies, a respective gene remained unknown for a long time. A lack of this knowledge and a relatively efficient spontaneous hydrolysis of FKyn prompted some researchers to speculate that this might be another nonenzymatic step of the Kyn pathway. However, recent studies confirmed the existence of this enzyme and its role in the Kyn pathway in various species.<sup>24,68–70</sup> Moreover, these studies, in combination with comparative genome analysis, revealed a remarkable level of interspecies variations associated with this functional role. Three structurally unrelated KYNFA families have already been identified in mammals (KYNFA-1<sup>69</sup>), bacteria (KYNFA-2<sup>24</sup>), and yeast (KYNFA-3<sup>70</sup>). At least one more family remains ‘missing’ as no representatives of the three known families could be identified in a number of bacterial species from *Xanthomonas* and *Flavobacteria* groups that contain an essential and otherwise complete Kyn pathway of Qa synthesis. Similarly, despite the functional

importance of KYNFA-1 demonstrated by the knockout of the respective gene (*AFMID* gene) in mouse model, some level of residual FKyn hydrolysis was observed in homozygous mice kidney, suggesting an existence of another, yet-unidentified KYNFA enzyme.<sup>68</sup>

The observed diversity of alternative forms of KYNFA reflects the general diversity of amidohydrolases and esterases that have evolved as several structurally and mechanistically distinct clans. Members of the KYNFA-1 family involved in the Kyn pathway of eukaryotes (except fungi) belong to the class of serine hydrolases. They share a sequence similarity with a number of characterized esterases, including conservation and verified functional importance of the classical catalytic triad (S162-H279-247). Although no structural or mechanistic studies of KYNFA-1 have been reported yet, there are many reasons to expect that this enzyme shares a characteristic  $\alpha$ -/ $\beta$ -fold and a catalytic mechanism with other members of the esterase/lipase superfamily. No obvious orthologues of the KYNFA-1 gene are present in *S. cerevisiae*, and previously this functional role was incorrectly assigned to a product of the *BNA3* gene. However, the gene *BNA3* was recently shown to encode a branching enzyme of the Kyn pathway, a pyridoxal phosphate-dependent kynurenine transaminase, whereas the KYNFA reaction was shown to be catalyzed by the product of a previously uncharacterized gene, YDR428C.<sup>70</sup> Members of this family with a rather narrow phylogenetic distribution (in yeast and some other Kyn pathway-containing fungi (e.g., *C. albicans* and *Gibberella zeae*) belong to the same esterase/lipase superfamily, but show almost no sequence similarity with the KYNFA-1 family. The 3D structure of the KYNFA-2 enzyme from *S. cerevisiae* published prior to its functional assignment<sup>71</sup> revealed the conservation of the overall  $\alpha$ -/ $\beta$ -fold and the catalytic triad (S110-H243-D211). Although both KYNFA-1 and KYNFA-2 are homologous, it is difficult to say whether the existence of these two families represent a case of an extremely divergent evolution or of an independent recruitment from the same widespread superfamily. The third family, KYNFA-3, characteristic of many bacterial species, was originally identified based on its chromosomal clustering with other genes of the Kyn pathway, and the representative recombinant enzymes from *Bacillus cereus*, *P. aeruginosa*, and *R. metallidurans* were experimentally characterized.<sup>57</sup> Members of this structurally distinct family belong to a so-called Cyclase superfamily, which includes a number of metal-dependent hydrolases with a characteristic set of conserved His, Asp, and Glu residues potentially involved in the chelation of the active site cation (most likely  $Zn^{2+}$ ) and catalysis. The recombinant enzyme from *R. metallidurans* characterized in more detail is inhibited by EDTA (but not PMSF), confirming its tentative classification as a metal-dependent hydrolase, and shows a rather stringent substrate preference for FKyn over several other aryl formamides.<sup>57</sup>

**7.08.2.1.2(iii) Kynurenine 3-monoxygenase (KYNMO)** The gene encoding an approximately 55-kDa human KYNMO enzyme (also referred to as kynurenine 3-hydroxylase) was cloned and functionally expressed.<sup>72</sup> A detailed characterization of the purified recombinant enzyme confirmed that it is a monooxygenase containing a tightly (but noncovalently) bound FAD cofactor and utilizing NADPH (and less efficiently NADH) for the cofactor regeneration coupled to oxidation of Kyn to HKyn.<sup>73</sup> The enzyme shows a stringent substrate preference for Kyn over its analogues, including a D-stereoisomer. All presently characterized KYNMO enzymes, including those from insects (e.g., *Anopheles stephensi*<sup>74</sup>) and bacteria (*Cytophaga butchinsonii*<sup>24</sup> and *P. fluorescens* str. 17400<sup>75</sup>) belong to the single well-conserved protein family. KYNMO reaction occurs at the branching point of the Kyn pathway, where an alternative reaction is a conversion of Kyn into kynurenic acid by kynurenine aminotransferase. Inhibition of KYNMO that changes a balance between these two competing enzymatic activities and decreases HKyn and Qa levels appears to be of therapeutic potential for a number of neurological disorders.<sup>51</sup> The presence of KYNMO in an organism is not always associated with Qa synthesis. As already mentioned, in insects HKyn is converted into xanthurenic acid,<sup>53</sup> whereas in some species of *Pseudomonas* it is an intermediate in the synthesis of a siderophore.<sup>60</sup> So far, no structural or mechanistic analyses were reported for the KYNMO family. However, a distant sequence similarity and conservation of the cofactors binding motifs suggest that KYNMO shares many structural and mechanistic aspects with other external flavoprotein aromatic hydroxylases. Extensively studied representatives of this family (e.g., hydroxybenzoate hydroxylase) catalyze hydroxylation of electron-rich aromatic substrates by molecular oxygen via the formation of the hydroperoxyflavin intermediate.



**7.08.2.1.2(iv) Kynureninase** This enzyme belongs to one of the largest and well-studied superfamilies of PLP-dependent aspartate aminotransferases. Despite the traditional KYNSE designation, the physiologically relevant reaction in the context of Qa synthesis is the hydrolysis of HKyn rather than Kyn. Nevertheless, all characterized eukaryotic and bacterial representatives of a broadly conserved KYNSE family display both these activities *in vitro*. In many bacteria that do not contain the KYNMO enzyme and harbor the aerobic degradation of Trp to anthranilate (as described in Kurnasov *et al.*<sup>57</sup>), the Hkyn-hydrolyzing activity of KYNSE is obsolete. However, in those bacterial and eukaryotic species that have a complete set of genes of Qa synthesis, a premature hydrolysis of Kyn, before its 3-hydroxylation, could potentially create a problem. Despite the high sequence similarity of bacterial KYNSE enzymes that are not involved in Qa synthesis (also termed *inducible*) to their eukaryotic counterparts (referred to as *constitutive*), the characterized representatives of these two groups strongly differ in their relative preference for hydroxylated versus nonhydroxylated substrates. For example, the recombinant KYNSE from *P. fluorescens* has an approximately 100-fold higher catalytic efficiency with Kyn compared to HKyn.<sup>76</sup> Conversely and in keeping with its main physiological function, the  $K_m$  of human recombinant KYNSE is approximately 20-fold lower for HKyn than for Kyn. The detailed mechanistic and structure–function analysis of approximately 45-kDa recombinant KYNSE from *P. fluorescens* confirmed that its catalytic mechanism is generally similar to the other representatives of the subgroup IVa of aminotransferases such as serine-pyruvate aminotransferase.<sup>76–78</sup> As in all PLP-dependent aminotransferases, the catalytic cycle begins by the Schiff base formation with the  $\alpha$ -amino group of the alanine moiety of Kyn (or HKyn) and proceeds via a series of intermediates driven by the C $\alpha$ -proton abstraction and hydration of the ketone carbonyl group with the formation of gem-diolate followed by the elimination of anthranilate, another (C4'-) proton abstraction, formation of quinonoid intermediate, C $\alpha$ -reprotonation, and, ultimately, the release of alanine and regeneration of a catalytically competent cofactor. The most remarkable feature of the proposed mechanism is that all these steps are mediated by a single residue (most likely the PLP-binding Lys-227 residue) operating as a general base catalyst. Structural analysis allowed us to elucidate the role of several other conserved residues in interaction with the cofactor and substrate. The 3D structure of the human KYNSE confirmed an anticipated high level of structural conservation and revealed a substantial conformational rearrangement in the positioning of the two structural domains, probably reflecting a transition between open and close conformations characteristic of the aspartate aminotransferase family.<sup>79</sup> Substrate modeling in the human KYNSE active site pointed to possible interactions that contribute to the enzyme preference for HKyn over the off-pathway hydrolysis of Kyn.

**7.08.2.1.2(v) 3-Hydroxyanthranilate 3,4-dioxygenase** HADOX catalyzes the last enzymatic step of the Qa synthesis, conversion of HAnt into ACMS, is a nonheme ferrous dioxygenase structurally that is mechanistically distinct from both TDO and IDO dioxygenase families involved in the first step of the same pathway. HADOX enzymes compose a functionally homogeneous superfamily with a cupin barrel fold, whose representatives are very highly conserved over the large phylogenetic distance. For example, the approximately 22-kDa HADOX enzyme from *R. metallidurans*,<sup>24</sup> the first characterized bacterial representative of the HADOX family, shares 37% sequence identity with the human enzyme. Recombinant human HADOX<sup>80</sup> and the enzyme purified from the beef liver<sup>81</sup> have basic characteristics similar to their counterparts from bacteria and yeast.<sup>82</sup> Mammalian HADOX enzymes contain approximately 100-aa C-terminal extension, which does not seem to be important for the enzymatic activity. According to one of the hypotheses, this segment may represent a second cupin domain contributing to the enzyme stabilization via formation of a bicupin monomer instead of a monocupin dimer characteristic of bacterial HADOX.<sup>82</sup> Published 3D structures of bacterial and yeast enzymes<sup>82,83</sup> revealed their anticipated structural similarity and contributed to mechanistic interpretation of the HADOX catalysis based on analogy with other cupin nonheme dioxygenases such as catechol extradiol dioxygenase.<sup>84</sup> Each subunit in a microbial HADOX dimer contains two iron binding sites – the buried catalytic site with conserved His and Glu residues as ligands and another forming a Fe–S cluster with four Cys residues on the surface. The catalytic cycle is initiated by the formation of a five-coordinated Fe(II) complex with HAnt stabilized by additional H-bonds with several conserved residues in the enzyme active site. Proximal binding of the dioxygen molecule leads to the electron transfer from the metal center and the formation of a superoxide intermediate. The negatively charged dioxygen attacks the C3 atom of HAnt, leading to the formation of an unsaturated seven-membered lactone intermediate via a series of rearrangements. The latter is hydrolyzed by



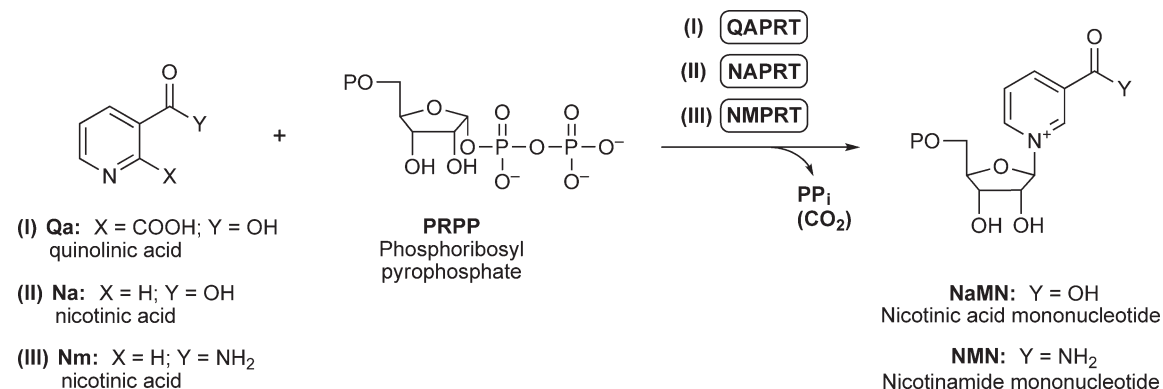
the iron(II)-bound hydroxide ion. Studies with recombinant bacterial HADOX led to the elucidation of the mechanism-based action of its 4-chloro-3-hydroxyanthranilate inhibitor.<sup>85</sup> This substrate analogue with therapeutic potential<sup>22</sup> inhibits HADOX via formation of a complex with a dioxygen-bound form of the enzyme that leads to the generation of the superoxide and oxidation of the iron center. In the inactivated enzyme, four cysteine residues of the metal center form two proximal disulfide bonds.

Finally, the recent study confirmed that the last stage of Qa formation from ACMS is indeed a spontaneous (nonenzymatic) pericyclic reaction that proceeds via ACMS *cis-trans* isomerization about C2–C3 and C4–C5 double bonds followed by the electrocyclization of the enol tautomer.<sup>86</sup>

### 7.08.2.2 Pyridine Mononucleotides

In contrast to the diversity of biochemical transformations associated with the synthesis of Qa, the conversion of pyridine precursors into pyridine mononucleotides always follows the same route, a phosphoribosyltransferase reaction with the PRPP intermediate supplied by the pentose phosphate pathway (Scheme 3). Moreover, the three known families of phosphoribosyltransferase enzymes specific for one of the three known pyridine precursors, QAPRT, NAPRT, and NMPRT, are conserved between taxons and appear to share their evolutionary ancestry and certain aspects of their catalytic mechanisms. As already mentioned, only Qa can be synthesized *de novo* by one of the two pathways described above, whereas Nm and Na are generated exclusively as products of NAD cofactor degradation. Those organisms that lack Qa *de novo* synthesis (e.g., many bacterial pathogens, as well as some species of archaea, fungi, and insects) are niacin auxotrophs that are completely dependent on the supply of exogenous precursors. Utilization of Na proceeds via NAPRT, the first enzyme in a classical Preiss–Handler pathway,<sup>87</sup> whereas two distinct routes were described for Nm. The first route that merges with the Preiss–Handler pathway upon the conversion of Nm into Na by the nicotinamidase (NMASE) is present in many bacteria, archaea, and eukaryotes (but apparently not in vertebrates). In the second, which was recently uncovered and is a relatively rare nondeamidating pathway, Nm is directly phosphoribosylated to NMN by NMPRT.

The niacin uptake mechanism has not been fully elucidated. A high-affinity Na permease of the DAL5 family was identified in *S. cerevisiae*,<sup>88</sup> but it appears to be present in only yeast and some other fungi. An energy-dependent niacin uptake was reported in *E. coli*,<sup>89</sup> but the respective transporter gene has not been identified. Therefore, it is impossible to rule out that niacin uptake may occur via a passive diffusion instead of (or at least in addition to) an active transport.<sup>4</sup> Indeed, the inactivation of a recently identified transporter gene, *niaP*, in *B. subtilis* only partially suppressed the niacin uptake.<sup>90</sup> Members of the NiaP family are conserved in many divergent bacteria, and a long-range NiaP homologue was found in the human genome. Although the respective protein product was described as a synaptic vesicle 2-related protein (SVOP) expressed in the brain and the endocrine cells,<sup>91</sup> its actual molecular function remained unclear, and it is tempting to speculate that it may play a role of a niacin transporter. The existence of the active transport of Na in human liver cells was recently reported.<sup>92</sup>



**Scheme 3**

In addition to salvage of exogenous niacin, NAPRT and NMPRT enzymes are involved in pyridine recycling pathways that play an important role in NAD homeostasis inside the cells. An endogenous Nm is generated by a variety of enzymes utilizing NAD as a co-substrate in nonredox reactions, such as ADP ribosyltransferase, poly-ADP ribosyltransferase, cADP ribose synthase, and NAD-dependent protein deacetylase of the Sir2/CobB family. Rapidly growing interest in these enzymes is due to the emerging understanding of their involvement with a variety of regulatory and signaling pathways, mostly in the eukaryotic cells (for a recent review, see Belenky *et al.*<sup>1</sup>). An additional flux of endogenous niacin may be generated by a phosphorylase reaction in the course of salvage or recycling of pyridine nucleosides. The salvage of RNm via RNm phosphorylase (RNMPSE) was recently described in yeast.<sup>93</sup> There are many reasons, including the evidence from comparative genomics, to anticipate the existence of a similar pathway in other species, including many bacteria. Pyridine nucleosides may also be directly phosphorylated to respective mononucleotides as described for some bacteria, yeast, and mammals.<sup>94–96</sup> In some species, Na is generated as a byproduct of NaMN conversion into  $\alpha$ -ribazole-phosphate, a precursor of the cobalamin lower ligand, catalyzed by the CobT enzyme.<sup>97</sup>

Although Nm and Na are the most common exogenous pyridine precursors, some bacterial pathogens may have developed a capability of salvaging Qa. For example, *Streptococcus pneumoniae* and *Streptococcus pyogenes* (but not other Streptococci) contain a QAPRT encoding gene (*nadC*) in the absence of any other genes of Qa *de novo* synthesis. This conjecture, which is yet to be tested, may provide a new link between these opportunistic pathogens and certain human pathological conditions associated with accumulation of Qa.

#### 7.08.2.2.1 Quinolate phosphoribosyltransferase

QAPRT catalyzes the last reaction leading to the *de novo* synthesis of NaMN, the transfer of the phosphoribosyl moiety of PRPP to Qa yielding NaMN, PP<sub>i</sub>, and CO<sub>2</sub>. QAPRT is ubiquitous and has been isolated and characterized from bacteria,<sup>98–100</sup> plants,<sup>101–103</sup> and mammals.<sup>104–107</sup> Interest in the human QAPRT was triggered by the observed neurotoxicity of Qa associated with a number of neurodegenerative diseases.<sup>108,109</sup> A gene, *nadC*, encoding an approximately 36-kDa QAPRT was for the first time identified in *Salmonella*.<sup>110</sup> Its orthologues are present in the genomes of hundreds of diverse species having one of the two pathways of Qa synthesis without a single detected case of nonorthologous displacement (see NAD(P) biosynthesis subsystem in The SEED database<sup>11</sup>). In many prokaryotic genomes, the gene *nadC* occurs in the same operon and/or is coregulated with *nadAB* genes of the Asp–DHAP pathway.

QAPRT was the first structurally characterized representative of type II PRTase enzymes, whose seven-stranded  $\alpha/\beta$  barrel architecture is distinct from the typical Rossmann fold of type I PRTase.<sup>111</sup> Structural analysis of the enzyme from several sources in the apo form and complexed with substrate (Qa) or product (NaMN)<sup>111–116</sup> and kinetic analysis of QAPRT from *E. coli*<sup>98</sup> and *Salmonella*<sup>117</sup> have provided insights into its mechanism of catalysis. QAPRT is a functional homodimer, and its active site is formed at the interface of the N- and C-terminal domains supplied by different subunits. The phosphoribosyltransferase reaction proceeds via the sequential ordered kinetic mechanism, with Qa productive binding in the active site preceding interaction with PRPP. The catalytic mechanism of QAPRT was proposed by analogy with orotate phosphoribosyltransferase and orotidine monophosphate decarboxylase enzymes that catalyze two consecutive reactions with an overall outcome similar to that of the NadC enzyme.<sup>10</sup> This mechanism includes activation of the PP<sub>i</sub> leaving group of PRPP via coordinated Mg<sup>2+</sup> ions and stabilization of the reactive oxocarbenium intermediate via interactions with the two conserved Arg residues. Interaction of the oxocarbenium ion with Qa apparently leads to the formation of an unstable quinolate mononucleotide that undergoes rapid decarboxylation at C-2, likely assisted by a proton transfer from the proximal Lys residue of the enzyme. A conservation of the QAPRT enzyme overall structural organization and the functional importance of implicated conserved Arg and Lys residues are supported by site-directed mutagenesis and structural data recently reported for the human enzyme.<sup>114</sup> Kinetic analysis of the human QAPRT revealed a fine regulation of its activity by the PRPP substrate inhibition.<sup>114</sup>

#### 7.08.2.2.2 Nicotinate phosphoribosyltransferase

NAPRT activity was originally described in the context of the mammalian Preiss–Handler pathway.<sup>87</sup> The respective gene (*pncB*) encoding an approximately 45-kDa protein was cloned and sequenced from *E. coli* and

*Salmonella*.<sup>118,119</sup> In many bacteria and archaea, the *pncB* gene occurs in the same operon and/or regulon with the *pncA* gene encoding the NMASE enzyme. In many bacterial pathogens, Nm salvage via the PncA–PncB route is the single and essential pathway supplying NaMN precursor for NAD synthesis.<sup>120</sup> All the genomes (including plants and yeast) that contain NMASE (PncA) also contain NAPRT (PncB). On the other hand, many bacterial (as well as all mammalian) genomes containing NAPRT do not seem to contain PncA orthologues. The latter gene pattern is consistent with the classical version of the Preiss–Handler pathway that utilizes Na but not Nm. On the other hand, the existence of the alternative NMASE in some species lacking PncA orthologues cannot be ruled out. The NAPRT enzyme from *Salmonella* was a subject of the most detailed kinetic and mechanistic studies.<sup>121–124</sup> The chemical reaction catalyzed by NAPRT (**Scheme 3**) is very similar to the reaction of QAPRT except for the decarboxylation step. However, the most remarkable and unique feature of NAPRT is its ability to couple the energy of ATP hydrolysis to increase the efficiency of catalysis by more than 10-fold and to shift the equilibrium of this reversible reaction toward the NaMN product. A mechanistic interpretation of this interesting phenomena includes autophosphorylation of the conserved His-219 residue that increases enzyme affinity toward the PRPP substrate and the dephosphorylation-coupled increase in the efficiency of the product release.<sup>122</sup> The 3D structural analysis of the yeast NAPRT provided an additional support for this model.<sup>125</sup> Despite a substantial amino acid sequence divergence, NAPRT shares the overall architecture characteristic of the type II phosphoribosyltransferases with QAPRT (and to a larger extent with NMPRT, see below), supporting the common evolutionary ancestry proposed for all three phosphoribosyltransferases involved in NAD synthesis.<sup>15</sup> An additional C-terminal domain of NAPRT supplies an ATP binding site involved in autophosphorylation of the conserved His-219 (His-232 in the yeast enzyme). It is assumed that the autophosphorylation contributes to the charge stabilization of the oxocarbenium intermediate. The phospho-histidine group appears to functionally compensate for the lack of two negatively charged carboxylates in the mechanism of QAPRT (C-2 of the Qa substrate and Glu-214 residue occupying the same position in QAPRT as His-219 in NAPRT). However, ATP activation may not be characteristic of all representatives of the NAPRT family. For example, the equivalent His residue is absent in a recently reported 3D structure of NAPRT from *Thermoplasma acidophilum*.<sup>126</sup> Several bacterial NAPRT structures were recently deposited in PDB (e.g., 2IM5, 1YBE, 2F7F), and their analysis would contribute to better understanding of this enzyme family.

A detailed molecular characterization of the human NAPRT was reported only recently, and it was shown that the activity of this enzyme (in contrast to NMPRT) is not a subject of feedback inhibition by the physiological levels of NAD.<sup>127</sup> This finding may provide a rationale for the observation that the increased levels of Na, but not Nm, in circulation leads to the increase in the NAD cellular pool, emphasizing the different roles of these two forms of niacin in nutrition and signaling in mammals (for a review, see Bogan and Brenner<sup>6</sup>).

### 7.08.2.2.3 The nicotinamide phosphoribosyltransferase

NMPRT enzyme that converts Nm directly into NMN was originally characterized as a product of the *nadV* gene in *Haemophilus ducreyi*.<sup>128</sup> This organism, as most other members of Pasteurellaceae, lacks the machinery for amidation of the pyridine carboxylate (as well as most other enzymes of NAD synthesis, see Section 7.08.3.1). The presence of the NMPRT enzyme in some (but not all) members of this group allows them to salvage Nm by an alternative nondeamidating route, directly coupling NMN synthesis with its further conversion into NAD by the NMN-specific adenylyltransferase (Section 7.08.2.3). NMPRT orthologues were identified and functionally characterized in some diverse bacteria and bacteriophages<sup>129,130</sup> and in vertebrates, including human,<sup>131</sup> but not in other eukaryotes or archaea. This sporadic phylogenetic distribution and a remarkably high degree of sequence conservation between the vertebrate and bacterial NMPRT enzymes point to an interesting evolutionary history of this enzyme family that may include a rather unusual horizontal gene transfer (HGT) from vertebrates to bacteria. The observed presence of the *nadV* gene in a number of analyzed T4-related bacteriophages may reflect upon its role in regeneration of NAD that is heavily consumed at certain stages of the phage development.<sup>132</sup> It is tempting to speculate that these and related bacteriophages could have played a role in dissemination of the *nadV* gene among bacteria (S. Gerdes, personal communication). In most (but not all) bacterial species containing the *nadV* gene it is located in an operon with a gene encoding a fusion enzyme composed of the NMNAT (NadM family) and ADP ribose pyrophosphatase

domains (Section 7.08.2.3). This operon was shown to play a role in the recycling of endogenous Nm rather than in the salvage of exogenous Nm in *Synechocystis* sp.<sup>129</sup>

The human NMPRT is an approximately 52-kDa protein that was originally characterized as a cytokine-like protein PBEF (for pre-B-cell colony-enhancing factor) that promotes B-cell precursors' maturation in the presence of interleukin-7 (IL-7) and stem cell factor (SCF). This protein was later shown to act as an adipokine secreted by visceral fat and was renamed as visfatin.<sup>133</sup> After the elucidation of the NadV function in *H. ducreyi*, the NMPRT activity of the human recombinant PBEF/visfatin protein was confirmed. This enzyme was shown to be a rate-limiting step in the NAD biosynthesis from Nm and an activator of the NAD-dependent protein deacetylase of the Sir2 family in mammals.<sup>134</sup> A connection between the enzymatic and signaling functions of NMPRT/PBEF/visfatin in NAD biosynthesis and insulin secretion emerged from the apparent role of the excreted form of NMPRT in the extracellular synthesis of NMN having a regulatory effect on insulin secretion by  $\beta$ -cells.<sup>135</sup> A possible role of the human NMPRT as a longevity factor extending the lifespan of human vascular smooth muscle cells by the activation of Sir1 and by restraining the accumulation of p53 was reported.<sup>136</sup> This enzyme is also considered a potential target for cancer therapy as it appears to be upregulated in tumors that have a high rate of NAD<sup>+</sup> turnover due to the increased ADP-ribosylation activity.<sup>137</sup>

Unlike some members of the NAPRT family, the characterized mammalian NMPRT enzymes displayed no evidence of the ATP activation behavior. The 3D structures of human and murine NMPRTs were solved in apoform and in complex with the reaction product NMN.<sup>138–140</sup> These data showed that despite a substantial overall similarity with NAPRT, this enzyme is a homodimer. A structural comparison of NMPRT with two other enzyme families suggested that the signature Asp-219 residue of NMPRT is a key determinant of its substrate preference for the amidated form of the pyridine substrate. The 3D structure of NMPRT in complex with its potent small-molecule inhibitor APO866 (also known as FK-866) revealed that it binds in a tunnel at the interface of the NMPRT dimer directly competing with the Nm substrate.<sup>138,141</sup>

Several other enzymes known to be directly involved in the biogenesis of pyridine mononucleotides via salvage and recycling of pyridines and pyridine nucleosides are less ubiquitous than the three families of phosphoribosyltransferases described above. Here, we briefly describe some of the best-studied salvage enzymes that are captured in **Figure 2** and **Table 1**.

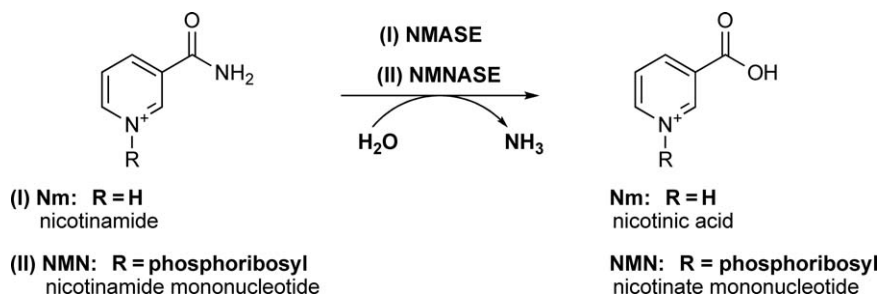
#### 7.08.2.2.4 Nicotinamidase

NMASE, an approximately 20-kDa product of the *E. coli* gene *pncA* (formerly *ydjB*), is a thiol-dependent amidase converting Nm into Na (**Scheme 4**) as a first step of a classic two-step Nm salvage pathway (PncA–PncB) that is present in many prokaryotes, plants, and nonvertebrates (e.g., fungi, nematodes, and insects). Although the *pncA* mutations affecting Nm salvage were described in early studies,<sup>142</sup> the identification of the actual gene and characterization of its protein product from *E. coli* and *Mycobacterium tuberculosis* were first reported in 1996.<sup>143,144</sup> The study of NMASE was to a large extent triggered by the role of this enzyme in the mechanism-based activation of the antituberculosis drug pyrazinamide (PZA), a close analogue of Nm that is known to be hydrolyzed by NMASE (historically referred to as pyrazinamidase) into pyrazinoic acid, an inhibitor of mycolic acid synthesis in *M. tuberculosis*. Mutations of the *pncA* gene are some of the most common causes of the observed PZA resistance in tuberculosis patients.<sup>145</sup>

A mechanistic interpretation of the NMASE catalytic mechanism as analogous to *N*-carbamoylsarcosine amidase, a well-studied enzyme that belongs to the same family, is supported by the 3D structure of NMASE from *P. borikoshii*.<sup>146</sup> In addition to the overall conservation of the characteristic  $\alpha$ -/ $\beta$ -fold, this structure confirmed the presence of the classical catalytic Cys-Lys-Asp triad in the active site. Unexpectedly, this structure also revealed the presence of Zn<sup>2+</sup> in the active site, which appears to interact with the Asp residue of the catalytic triad, possibly contributing to the proton transfer and activation of the active site thiolate for the nucleophilic attack on the amide bond of Nm.

#### 7.08.2.2.5 NMN deamidase

The existence of this enzymatic activity (EC 3.5.1.42) converting NMN into NaMN (**Scheme 4**), was documented in previous studies (as reviewed in Magni *et al.*<sup>2</sup>). It is also supported by physiological studies in *Salmonella* suggesting that NMNase-driven deamidation of NMN, rather than its NMNAT-driven adenylation



Scheme 4

to NAD, plays a key role in NmR salvage mediated by the NmR kinase (NMRKIN, see below).<sup>17,147</sup> This reaction might be of particular importance for those bacterial species that have the upstream components of the NmR salvage machinery (PnuC transporter and NMRKIN) but lack the NMNAT activity as noted for *P. aeruginosa*.<sup>95</sup> Finally, NMN is believed to be generally toxic for bacteria, at least partially due to its inhibitory effect on NAD-dependent bacterial DNA ligase. Therefore, an additional role of NMNSE (and/or NMN phosphatase) may be in rapid detoxification of NMN that may accumulate as a byproduct of DNA ligase and, possibly, some other enzymes with NAD pyrophosphatase activity. However, despite many attempts, a gene encoding NMNSE remains *missing*. For example, recent studies showed that a *pncC* locus assigned NMNSE activity by the earlier genetic studies,<sup>16</sup> in fact, encodes a periplasmic alkaline phosphatase AphA that has an NMN phosphatase activity.<sup>17</sup>

#### 7.08.2.2.6 Ribosyl nicotinamide kinase

A bacterial form of RNMKIN was predicted by bioinformatics analysis and experimentally verified to be encoded within the C-terminal domain of a multifunctional NadR protein that is present in most Enterobacteriaceae and Pasteurellaceae.<sup>95</sup> RNMKIN converts RNm into NMN (Scheme 5) in the first step of the pyridine mononucleoside salvage pathway following its uptake via PnuC transporter<sup>17,148</sup> (see Figure 2). The original prediction was based on the genomic reconstruction of the so-called V-factor pathway, which is the only route of NAD synthesis in *H. influenzae*, and the observation of the kinase-like domain fused with the NMNAT domain, which is responsible for the second and the last step of this pathway (Figure 3(d)). The kinetic analysis of the RNMKIN from *Haemophilus influenzae* and *Salmonella* revealed a stringent substrate preference for RNm over its deamidated analogue and an appreciable feedback inhibition by NAD.<sup>94,147</sup> The 3D structure of the bifunctional (NMNAT/RNMKIN) NadR protein from *H. influenzae* revealed that its RNMKIN domain has a three-layered  $\alpha$ -/ $\beta$ -/ $\alpha$ -fold similar to typical P-loop kinases, such as yeast thymidylate kinase (TMK). The similarity included conservation of certain active site elements, pointing to a likely similarity of the catalytic mechanism.<sup>149</sup> Despite the fusion of the two domains involved in the two consecutive reactions (RNMKIN and NMNAT), their spatial arrangement<sup>149</sup> in *H. influenzae* NadR protein and mutagenesis data<sup>95</sup> gave no indication of mechanistic coupling between them.

The existence of the human RNMKIN activity was suggested by early studies further supported by the mechanism-based activation of RNm synthetic analogues pursued as potential anticancer drugs such as tiazofurin and benzamide riboside.<sup>150,151</sup> Nevertheless, the first eukaryotic (yeast and human) genes encoding RNMKIN were identified only recently, and their role in NAD metabolism was demonstrated biochemically and genetically.<sup>94,152</sup> A detailed structure–function analysis of eukaryotic RNMKIN enzymes that have no homology with their bacterial counterparts revealed a substantial difference, not only in the overall structure, but also in substrate specificity.<sup>96</sup> In contrast to bacterial RNMKIN, the eukaryotic enzyme has a high activity with tiazofurin and shows a dual specificity toward both pyridine nucleosides, RNm, and its deamidated analogue RNa (Scheme 5). The demonstration of this novel ribosyl nicotinate kinase (RNAKIN) activity and the confirmed existence of the novel metabolite RNa *in vivo* prompted the authors to propose the previously unknown pathway of pyridine nucleoside salvage or recycling.<sup>96</sup> These and some other data bring into the spotlight an emerging physiologic and nutritional importance of the pyridine nucleoside precursors (that were found to be rather abundant in milk).<sup>6,93</sup> The existence of RNAKIN enzymatic activity and its potential



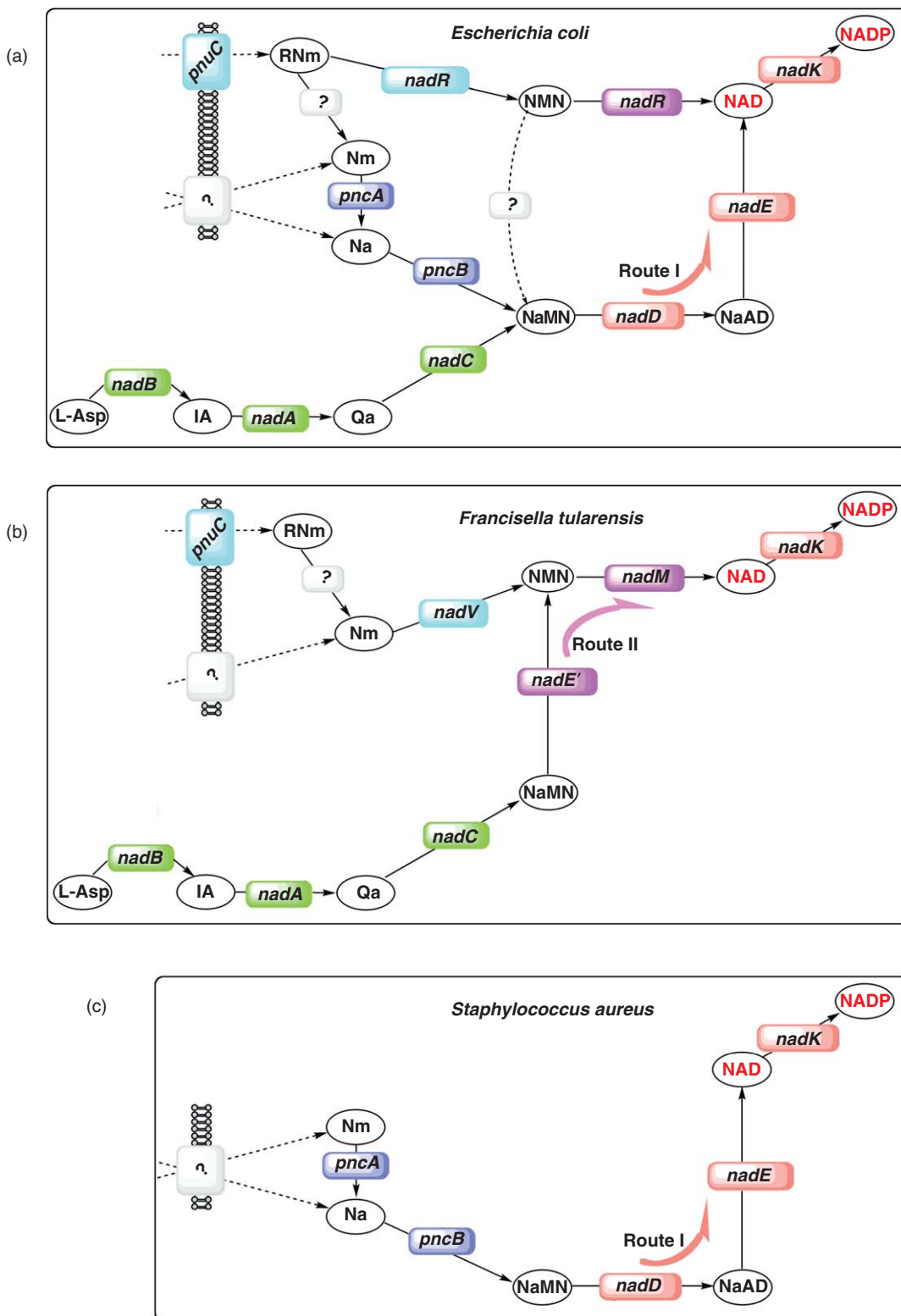
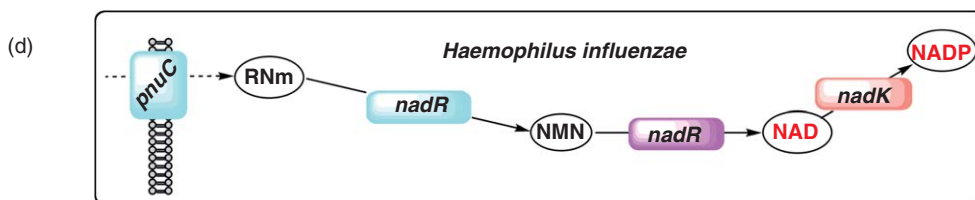
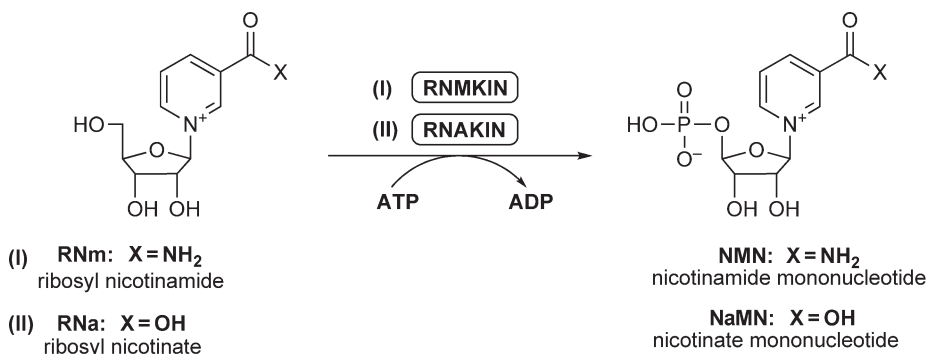


Figure 3 (Continued)





**Figure 3** NAD biosynthetic pathways in bacteria. Examples of genomic reconstruction of NAD biosynthesis in selected species ((a)–(d)) are illustrated by the simplified variants of the general subsystem diagram (shown in **Figure 2**). Only those functional roles (as well as respective reactions and metabolites) that correspond to the genes identified in a given species are shown. Gene names that are shown instead of functional role abbreviations are as given in **Table 1**. (a) In *E. coli*, the *de novo* biosynthesis from Asp is supplemented by two salvage pathways: from Nm/Na (via PncA–PncB) and from RNm (via PnuC–NadR). An alternative route of NMN salvage via NMN deamidase was demonstrated, but the corresponding gene has not been identified (shown by ‘?’). The uptake mechanism for Nm and Na in most bacteria remains unclear (as reflected by ‘?’). (b) *Francisella tularensis* is a rare example of Route II converting NaMN into NAD via NMN intermediate. Another distinctive feature is the presence of a nondeamidating Nm salvage pathway (via NadV) instead of a more common PncA–PncB pathway. This organism also contains a PnuC transporter but lacks the RNMKIN enzyme pointing to the most likely pathway via RNMPSE. (c) *Staphylococcus aureus* and many other Gram-positive pathogens lack *de novo* biosynthesis, fully relying on Nm salvage. (d) *H. influenzae* and many other Pasteurellaceae display one of the most radical truncations of NAD biosynthetic machinery. An obligatory salvage of the exogenous V-factors proceeds via the PnuC–NadR pathway of RNm salvage.

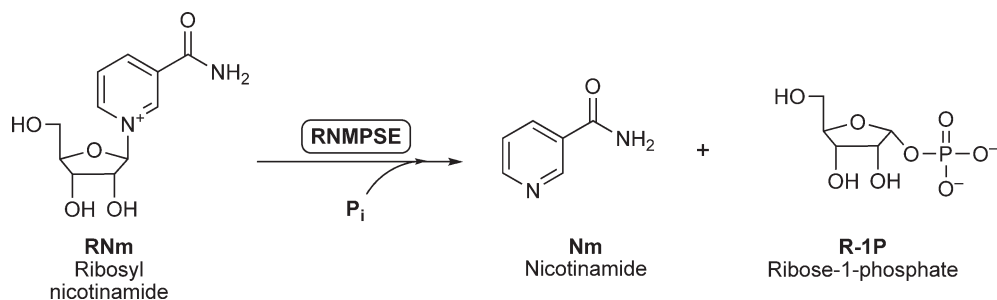


**Scheme 5**

importance for the recycling of pyridine nucleosides in plants were recently reported by a study in mungbean seedlings.<sup>153</sup> The 3D structure of the human RNMKIN, a member of the deoxynucleoside kinase and nucleoside monophosphate kinase superfamily in complex with NMN and in a ternary complex with ADP and tiazofurin together with site-directed mutagenesis, provided additional insights to its active site organization and catalytic mechanism.<sup>154</sup>

#### 7.08.2.2.7 Ribosyl nicotinamide phosphorylase

A possible involvement of the pyridine nucleoside phosphorylase in pyridine cycles was discussed in earlier reviews.<sup>2–4</sup> However, the first detailed report of the role of RNMPSE activity (**Scheme 6**), including the identification of genes contributing to the RNMKIN-independent salvage of RNm (**Figure 2**) in *S. cerevisiae*, appeared only recently.<sup>93</sup> One of these genes encodes a yeast homologue of mammalian purine phosphorylase (Pnp1) that apparently may utilize RNm (and possible RNa) as one of the phosphorolysis substrates yielding a free pyridine and ribose-1-phosphate. The same report implicated another enzyme (Urh1) as a putative nucleosidase that also contributes to the generation of Nm from RNm by a different mechanism similar to the hydrolysis of pyrimidine nucleosides by uridine- and cytidine-specific nucleoside hydrolases. Despite a likely physiological relevance of these additional pathways for recycling and salvage of pyridines, the respective enzymes have not yet been fully characterized, and we did not include them in **Table 1**. Identification and accurate annotation of the respective enzymes across diverse genomes is a challenging task



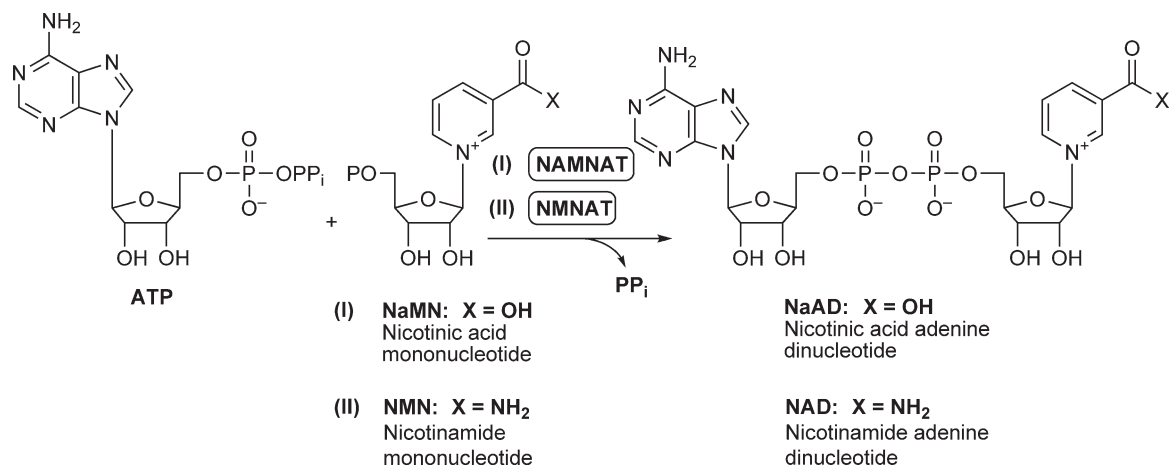
Scheme 6

as they may substantially vary between species and many of them may not be fully committed to NAD metabolism, displaying varying degrees of preference toward a broad spectrum of nucleoside substrates. Identification of gene candidates for relevant enzymes may be assisted by genome and functional context analysis, which are particularly efficient for prokaryotic genomes (see examples for NAD metabolism<sup>120,129,155</sup>). For example, *S. enterica* subsp *enterica* and *S. bongori* contain a conserved chromosomal cluster encoding an RNm transporter PnuC and a paralogue of uridine phosphorylase (Udp), which may be considered a candidate for the role of bacterial RNMPSE.

In brief, NaMN, a nearly universal intermediate of NAD synthesis, is generated mainly by phosphoribosylation of Qa or Na. Of these two pyridine precursors, Qa is generated by one of the two routes of *de novo* synthesis, whereas Na (vitamin B<sub>3</sub> or niacin) is a product of NAD degradation. Na may be supplied directly from the medium or generated by the deamidation of Nm within the cells. A direct phosphoribosylation of Nm to NMN contributes to an alternative route of Nm recycling and salvage in vertebrates and some bacteria. Salvage of pyridine nucleosides may proceed by their direct phosphorylation to respective mononucleotides or by their further degradation to pyridine precursors.

### 7.08.2.3 Formation of Pyridine Dinucleotides

The last universal stage of building the NAD scaffold is the adenylation of pyridine nucleotides (NaMN or NMN) by a committed adenylyltransferase enzyme that consumes one molecule of ATP to generate a pyridine dinucleotide (NaAD or NAD) releasing inorganic pyrophosphate (PP<sub>i</sub>) (Figure 1 and Scheme 7). Enzymes catalyzing both these reactions belong to a large nucleotidyltransferase superfamily (also known as the HIGH superfamily for the conserved active site sequence motif<sup>18</sup>) and share many aspects of structural organization and catalytic mechanism. Four distinct protein families with NaMN adenylyltransferase (XVI, NAMNAT)



Scheme 7

or/and NMN adenylyltransferase (XVII, NMNAT) activity and different phylogenetic distribution are presently known. Some of these families have a very strong preference for NaMN (the NadD family conserved in most bacteria) or NMN (the NadR family present in a limited set of bacteria), whereas other families have a comparable NAMNAT and NMNAT catalytic efficiency (the eukaryotic PNAT family and the archaeal group within the NadM family). These variations in specificity (that may occur even within a family, see below) are not surprising as both pyridine mononucleotide substrates (NaMN and NMN) are structurally similar and their adenylation propensities are not directly affected by the presence of a carboxylate or carboxamide group in the pyridine moiety. However, since NAMNAT and NMNAT play distinct roles in NAD metabolism, a genomic reconstruction approach often helps infer the most likely activity required by an organism. For example, NAMNAT activity is required for the conventional Route I of NAD synthesis (NaMN  $\rightarrow$  NaAD  $\rightarrow$  NAD, **Figure 2**) that is present in the overwhelming majority of species in all three kingdoms of life. However, a few species that do not have the NADSYN enzyme (see Section 7.08.2.4) use the NMNAT activity to directly convert NMN into NAD by salvaging Nm (via NMPRT as in *H. ducreyi*) or RNm (via RNMKIN as in *H. influenzae*) as the main or the only route to NAD. Additional examples of variations in NAD biosynthetic pathways in different groups of species will be described in Section 7.08.3. Comparative analysis of more than 700 complete genomes integrated in The SEED database suggests that the four evolutionarily related NAMNAT/NMNAT families mentioned above likely cover the entire phylogenetic space. The only species that lack any representatives of these enzyme families are bacterial endosymbionts with extremely truncated genomes, such as *Blochmannia*, *Rickettsia*, and *Chlamydia* groups that apparently developed an ability to uptake and salvage intact NAD(P) from the eukaryotic host cells.<sup>120,156,157</sup> The basic enzymatic, structural, and mechanistic features of all four NAMNAT/NMNAT families as well as some aspects of their physiology were recently reviewed.<sup>158</sup> Therefore, in this section we only briefly introduce them and highlight some new data.

#### 7.08.2.3.1 NaMN adenylyltransferase of the NadD family in bacteria

Although the chromosomal locus (*nadD*) encoding the NAMNAT was mapped in the early genetic studies of *Salmonella*,<sup>159</sup> the actual *nadD* gene (formerly *ybeN*) was cloned from *E. coli* and experimentally verified first in 2000.<sup>160</sup> This approximately 25-kDa enzyme revealed a very strong preference for NaMN vs. NMN substrate consistent with its biological function in NAD synthesis via NaMN intermediate (>1000-fold based on comparison of  $k_{\text{cat}}/K_{\text{m}}$  values for each substrate). With some quantitative variations, a similar bias in substrate preference was confirmed for other characterized representatives of the NadD family that is ubiquitous across the bacterial kingdom.<sup>129,156,161–164</sup> The *nadD* gene was proven essential in several model bacteria, and it is considered as an attractive target for the development of novel anti-infective compounds.<sup>120,156</sup>

The first 3D structure of the *E. coli* NadD enzyme revealed a version of the dinucleotide-binding (Rossmann) fold composed of the central seven-stranded  $\beta$ -sheet surrounded by  $\alpha$ -helices.<sup>164</sup> The mechanism of NadD catalysis is likely similar to other nucleotidyltransferases of this class, including a nucleophilic attack of the 5'-phosphoryl group of NaMN on the  $\alpha$ -phosphate of ATP facilitated by  $\text{Mg}^{2+}$  that is coordinated by a conserved (HXGH) segment. A high-resolution 3D structure of NadD complexed with its NaAD product provided a basis for the interpretation of its strict preference for NaMN substrate over NMN. A comparison with the 3D structure of the human PNAT (see below) confirmed an anticipated overall similarity but also revealed sufficient differences in the active site area to allow for the development of selective NadD inhibitors. This conjecture was confirmed by our recent studies that allowed us to develop potent NadD inhibitors that have almost no effect on the enzymatic activity of the human PNAT enzymes (L. Sorci *et al.*, unpublished).

#### 7.08.2.3.2 NMN adenylyltransferase of the NadR family in bacteria

The NMNAT activity was originally inferred by bioinformatics analysis<sup>165</sup> and experimentally identified for the central domain (aa 64–234) of the multifunctional NadR protein from *E. coli*. The three-domain version of the NadR protein that exists only in Enterobacteriaceae contains an N-terminal HTH domain, which is involved in DNA binding, mediating its role as a transcriptional regulator of NAD biosynthesis (see Section 7.08.3.1.2). The C-terminal domain harboring the RNMKIN activity is conserved in all members of the NadR family (see above). The NMNAT domain of the NadR family displays a very strong preference for NMN over NaMN, which is consistent with its role in salvage or recycling of amidated pyridine precursor.<sup>95,147</sup> The 3D

structure of NadR from *H. influenzae* confirmed an overall similarity of the NMNAT domain with other members of the superfamily, especially with the NadM family characteristic of archaea and a limited group of bacteria (see below). The analysis of interactions with the bound NAD molecule revealed the presence of the key contacts with its Nm moiety, providing a structural basis for the observed strict preference for NMN over NaMN.

The biological role of the NadR NMNAT activity domain in RNm salvage is of indisputable physiological importance in *H. influenzae* and other species that contain a two-domain version of the NadR protein without a regulatory HTH domain. However, evidence shows that this enzymatic activity may not be directly involved in RNm salvage in Enterobacteriaceae where the NMNAT domain could have adopted a function of an effector domain modulating the DNA-binding activity of the HTH domain in response to changes in the NAD pool. This conclusion is supported on the one hand by a relatively low NMNAT activity of *Salmonella* and *E. coli* NadR compared to its *H. influenzae* counterpart, and by an observation that this activity is neither required nor sufficient for supporting the growth of the organism when RNm salvage operates as the only route of NAD synthesis.<sup>147</sup> On the other, the RNMKIN activity of NadR is essential for RNm salvage as it contributes to capturing the NMN intermediate within the cell for its further utilization by alternative pathways, for example, via the NMNASE. Two-domain NMNAT/RNMKIN representatives of the NadR family (without an HTH domain) occur sporadically in a limited set of diverse bacteria, for example, in Mycobacteria, in *Leuconostoc*, and in some nonpathogenic Streptococci. In all cases (except for Pasteurellaceae), NadR is present in addition to (not instead of) the housekeeping NadD enzyme. Therefore, in all these cases NadR enzymes are likely involved in salvage (or recycling) of RNm that may be obtained via PnuC-mediated uptake (as in *Leuconostoc*) or generated via other, yet unknown, mechanisms.

### 7.08.2.3.3 NaMN- and NMN-specific adenylyltransferases of the NadM family in bacteria and archaea

Members of this family that were originally cloned from the archaeon *Methanococcus jannaschii* (hence the name NadM) and cyanobacterium *Synechocystis* sp., were historically the first identified genes of adenylyltransferases involved in NAD metabolism.<sup>166,167</sup> In a similar manner, the first 3D structure within the NAMNAT/NMNAT superfamily was reported for another member of the NadM family from *Methanobacterium thermoautotrophicum*.<sup>168</sup> The structural analysis of the enzyme complex with NAD followed by the site-directed mutagenesis provided the first insights into its catalytic mechanism and the functional role of the conserved HXGH sequence motif. Although a detailed kinetic comparison of the two pyridine nucleotide substrates was not reported, the authors provided a structural rationale for both types of activities, NAMNAT and NMNAT. Our recent steady-state kinetic analysis of the recombinant NadM enzyme from *M. jannaschii* confirmed its almost equivalent catalytic efficiency toward both substrates NaMN and NMN (L. Sorci, unpublished). This finding is consistent with the genomic reconstruction of NAD metabolism in all sequenced archaeal genomes that contain NadM as the only candidate gene for NAMNAT activity for Route I of NAD synthesis from the NaMN precursor (Figure 5). The actual physiological role of the NMNAT activity in archaea is less clear as the analyzed archaeal genomes do not seem to contain any enzymes or pathways of NAD synthesis proceeding via NMN intermediate (for additional discussion, see Section 7.08.3.2).

On the other hand, the bacterial version of the NadM enzyme from *Synechocystis* sp. displayed a strong preference for NMN over NaMN.<sup>167</sup> This type of substrate specificity is consistent with the proposed role of the enzyme in Nm recycling together with the NMPRT enzyme, whose gene *nadV* is located in the same operon with the *nadM* gene.<sup>129</sup> This conclusion was supported by the finding that the *nadD* gene, which is also present in the *Synechocystis* sp. genome, was essential for survival of the organism, whereas the *nadM* gene could be deleted without any detectable phenotype. A similar gene pattern is observed in several groups of divergent bacteria (e.g., *Xanthomonas* and *Burkholderia* spp.) that also contain both the *nadD* gene and the *nadM*-*nadV* operon, which is probably involved in Nm recycling/salvage pathways (as discussed in Section 7.08.2.2). Another distinctive feature of nearly all NadM family members in bacteria is the presence of an additional C-terminal domain that belongs to a Nudix hydrolase family and has an ADP-ribose pyrophosphatase activity (ADPRP).<sup>167</sup> ADP ribose is a side product of NAD glycohydrolase activity of various NAD-consuming enzymes (such as protein deacetylase of the Sir2/CobB family), and ADPRP activity is thought to play a role in its recycling. In *F. tularensis*, which does not have a NadD enzyme, NadM is the only housekeeping

adenylyltransferase that is involved in both salvage of Nm via NMPRT and *de novo* synthesis via Route II (see **Figure 3(b)** and Section 7.08.3.1). Despite dramatically different physiological roles, the 3D structures of NMNAT domains of both bacterial enzymes were highly similar to each other and to the previously published 3D structure of the archaeal NadM.<sup>169</sup> Nevertheless, some differences that were observed in the active site area involved in the pyridine binding provided a rationale for the difference in substrate preferences between bacterial and archaeal enzymes. However, despite the observed correlation between substrate preferences (strongly prevailing NMNAT versus dual NMNAT/NAMNAT specificity) and the presence or absence of the ADPRP domain, there is no evidence that these distinctive features of bacterial and archaeal enzymes are interrelated. A recent structural analysis of the two bacterial NMNAT/ADPRP enzymes, from *Synechocystis* and *F. tularensis*, did not reveal any functional interactions between NMNAT and the ADPRP domain.<sup>169</sup> Moreover, a comparative genomic reconstruction of NAD metabolism across all sequenced bacterial genomes suggests that some members of the bacterial NadM family may have a physiological role of NAMNAT (as in archaea). For example, of the two *nadM* copies present in the *Deinococcus* branch, one occurs in the conserved operon with *pncB* (NAPRT enzyme generating NaMN) and *nadE* (NADSYN utilizing NaAD). This genomic context along with the absence of the *nadD* gene in *Deinococcus* spp. genomes allow us to infer a role of the housekeeping NAMNAT in Route I of NAD synthesis for this *nadM* paralogue.

Overall, one may observe that the NadM family reveals a high degree of functional versatility affording variations in both substrate specificity and physiological role. These variations are prominent within the bacterial kingdom where the members of the NadM family are sparsely distributed. This is in contrast with archaea where NadM appears to be a completely universal housekeeping enzyme in the main route of NAD synthesis (an archaeal equivalent of NadD).

#### 7.08.2.3.4 NaMN- and NMN-specific adenylyltransferases in eukaryotes

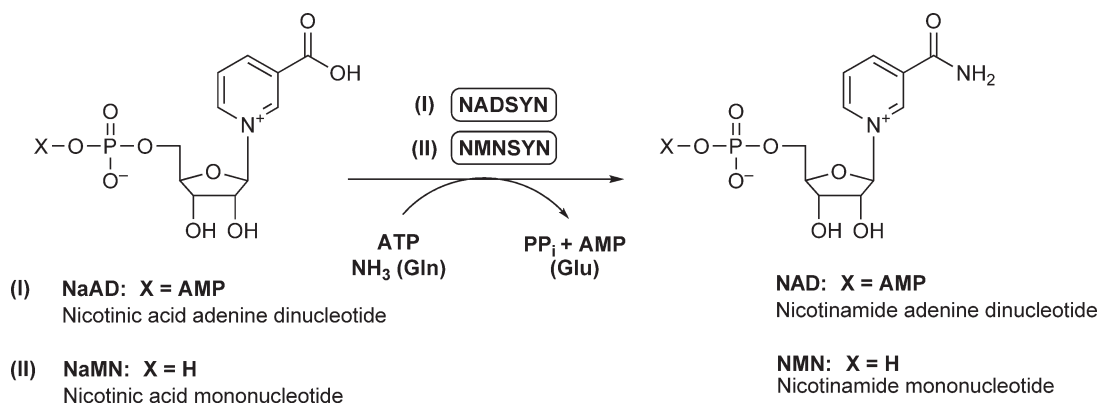
The dual substrate specificity (NAMNAT/NMNAT) with nearly equal catalytic efficiency for both pyridine mononucleotides appears to be characteristic of this family. Here, we refer to it as the PNAT family (for Pyridine monoNucleotide AdenylylTTransferase) to distinguish it from the three other families described above. The dual specificity and some other properties of the purified eukaryotic PNATs were established in biochemical studies (as reviewed in Magni *et al.*<sup>2</sup>) more than a decade before the identification of the first representative gene of the PNAT family in *S. cerevisiae*.<sup>170</sup> This organism contains two PNAT isoforms, and the second yeast PNAT was later cloned and overexpressed and the purified protein was characterized in detail.<sup>171</sup> A single PNAT gene present in *A. thaliana* was recently characterized.<sup>172</sup> The PNAT family is ubiquitous among all groups of eukaryotes, and respective genes were identified in every eukaryotic genome sequenced so far. At the same time, not a single case of PNAT occurrence was detected in approximately 700 analyzed prokaryotic genomes (see the 'NAD and NADP biosynthesis' subsystem in The SEED database.<sup>11</sup> In human, three PNAT isoforms have been identified and characterized, displaying distinct tissue distribution. The dual NAMNAT/NMNAT specificity that was confirmed for all these isoforms is physiologically significant as it allows these enzymes to participate in both *de novo* and salvage or recycling routes of NAD biosynthesis via either NMN or NaMN intermediates. In two recent comparative studies, the three human isozymes were further characterized with respect to their catalytic properties, kinetic mechanism, substrate preferences, and cellular compartmentalization.<sup>151,173</sup> A strict compartment-specific expression of the isoforms was confirmed, with PNAT-1 found in the nucleus, PNAT-2 in the Golgi complex, and PNAT-3 in mitochondria. Despite a substantial similarity of the three PNAT isoforms, an enzymatic assay taking advantage of some differences in their substrate and metal specificities was established to assess a relative contribution of each isozyme to NAD biosynthesis in various samples.<sup>151</sup> An additional interest in human PNAT enzymes is associated with implications in cancer and neurodegenerative diseases. Human PNAT-1 plays a key role in the metabolic activation of the anticancer drug tiazofurin and benzamide riboside by converting them into their active forms (TAD and BAD, respectively). In certain tumor cell lines, the observed tiazofurin resistance is due to the low NMNAT activity.<sup>174,175</sup> Conflicting reports exist about the neuroprotective effect of PNAT against axonopathy.<sup>1,7</sup> Two recent studies in the fly model indicated that PNAT confers neuroprotection independently of its enzymatic activity.<sup>176,177</sup> The 3D structures of PNAT-1 and PNAT-2 were reported in apo form and in various complexes with NMN, NAD, NaAD, ATP, and TAD analogues.<sup>178–181</sup> The analysis of these rich structural data improved mechanistic understanding of this family, including the structural basis of the dual



substrate specificity (as reviewed in Magni *et al.*<sup>158</sup>). The overall architecture of the PNAT monomer reveals a substantial similarity with other NaMNAT and NMNAT families. The closest similarity is observed with the bacterial NadD family, whereas two other families, NadM and NadR, are much more distant from both NadD and PNAT while being relatively close to each other.

#### 7.08.2.4 Amidation of Pyridine Nucleotides

The amidation of pyridine carboxylate is an essential stage in the biogenesis of the functional NAD cofactor. For a long time this reaction was thought to occur only as the last step, after the assembly of the complete pyridine dinucleotide scaffold. The respective reaction catalyzed by NAD synthetase (NADSYN) involves a transient adenylation of the pyridine carboxylate in the presence of ATP and magnesium, followed by the displacement of the AMP-moiety by ammonia (NH<sub>3</sub>).<sup>182</sup> NADSYN, originally characterized as a product of gene *nadE* in *Salmonella* and *E. coli*,<sup>183,184</sup> is the second-most-conserved enzyme in NAD synthesis (after NADKIN, Section 7.08.2.5). Members of the NadE family are found in all analyzed archaeal and eukaryotic genomes and in the overwhelming majority of bacteria without a single case of ‘missing gene’ or nonorthologous gene displacement. Among the few exceptions are bacterial species that lack the amidation machinery (as well as many other aspects of NAD biosynthesis) and that are dependent on the salvage of amidated precursors (e.g., in Pasteurellaceae, as discussed above and in Section 7.08.3.1) or even of the entire NAD cofactor (as in some intracellular endosymbionts of the *Rickettsia* and *Chlamydia* groups). Despite the overall conservation, two aspects of the NadE family appear to be a subject of variation between species. The most important one is the presence or absence of an additional glutamine amidotransferase (GAT) functional domain that allows NADSYN to utilize L-Gln as an *in situ* source of NH<sub>3</sub>. This domain, a member of the nitrilase family (Nit),<sup>185,186</sup> is present in all eukaryotic but absent in all archaeal members of the NadE family. Bacterial genomes are split into two groups, those that have a single-domain NADSYN, such as the enzymes from *E. coli* or *B. subtilis*<sup>184,187</sup> utilizing only NH<sub>3</sub> but not Gln, and those that have a two-domain GAT/NADSYN with an additional N-terminal GAT domain, such as in *M. tuberculosis* or *Synechocystis* sp. enzymes<sup>129,188</sup> utilizing both NH<sub>3</sub> and Gln for the amidation of NaAD. Another recently recognized variation within the NadE family is the existence of a divergent NadE subfamily (termed NadE’) with the NMN synthetase (NMNSYN) activity (Scheme 8). A single characterized representative of this subfamily from *F. tularensis* efficiently catalyzes amidation of the pyridine mononucleotide (NaMN) rather than the dinucleotide (NaAD) (L. Sorci *et al.*, submitted for publication). Although the NMNSYN activity appears to be confined within a small NadE’ subfamily, the importance of this finding is that it revealed a precedent of an alternative Route II of NAD synthesis where the amidation step occurs earlier to (and not after) the adenylation (Figure 3(b)).



**Scheme 8**



#### 7.08.2.4.1 NAD synthetase

A prototypical single-domain NADSYN (a *short form* lacking a GAT domain) from *E. coli* is an approximately 30-kDa product of the essential gene *nadE* (formerly *efg*) forming a homodimer in solution.<sup>184</sup> This enzyme and all other characterized representatives of this group that were cloned from various bacteria (e.g., *B. subtilis*,<sup>187</sup> *B. stearothermophilus*,<sup>189</sup> *B. anthracis*,<sup>190</sup> *H. pylori*,<sup>191</sup> and *Corynebacterium glutamicum* (K. Shatalin and A. Osterman, unpublished)) and archaea (*M. jannaschii* (L. Sorci and A. Osterman, unpublished)) efficiently catalyze ATP-dependent conversion of NaAD into NAD using only NH<sub>3</sub> but not Gln *in vitro*. Moreover, a genetic study in *Salmonella* provided convincing evidence that NH<sub>3</sub> and not Gln is the actual *in vivo* N-source for NADSYN reaction in this organism.<sup>192</sup> This finding may be tentatively generalized to many other bacterial species having a short form of *nadE* as a single NADSYN gene in their genomes. Although an alternative interpretation, the existence of a yet-unknown GAT subunit (rather than a domain) that would form a complex with a short NadE subunit, was considered, the detailed genome analysis does not support this hypothesis beyond a few cases. Although a two-subunit (rather than a two-domain) arrangement is characteristic of some ATP-dependent amidotransferase families,<sup>182</sup> in those cases both subunits (glutaminase and amidoligase) are nearly always located next to each other on the prokaryotic chromosome. A comparative analysis across hundreds of prokaryotic genomes failed to reveal any novel candidate GAT genes in the vicinity of short *nadE* genes. Among a few exceptions are *Thermus* and *Sulfolobus* spp., which contain a close homologue of a typical nitrilase-like GAT-domain encoded in the operon with a short *nadE* gene. Both genes from *Thermus thermophilus* (TTC1539 and TTC1538) were cloned and overexpressed, and their products were shown to form a tight enzymatic complex utilizing both NH<sub>3</sub> and Gln in a NADSYN reaction, whereas the amidoligase subunit alone displayed only NH<sub>3</sub>-dependent activity (K. Shatalin and A. Osterman, unpublished). A few other species (e.g., *Chlorobium tepidum*) have a close homologue of the nitrilase-like GAT domain encoded in their genomes remotely from a short *nadE* gene, and in some of them a formation of a two-subunit NADSYN with glutaminase activity may be anticipated. However, most of the other representatives of the nitrilase superfamily that are present in many prokaryotic genomes do not seem to be related to NADSYN by sequence similarity nor by chromosomal proximity. These observations are consistent with a notion that a short form of NADSYN that is present in all archaeal and many bacterial species represents a second example (after Gln synthase) of an amidoligase that uses NH<sub>3</sub> as N-source *in vivo*.

The first and the most detailed structural and mechanistic studies within this group of enzymes were reported for NADSYN from *B. subtilis*.<sup>193,194</sup> The availability of high-resolution 3D structures of several other representatives of this group from Gram-positive<sup>190</sup> and Gram-negative<sup>191,195</sup> bacteria further contributed to structure–function understanding of NADSYN. Both subunits of the NADSYN homodimer contribute to the formation of the two NaAD/NAD binding sites located at the dimer interface, whereas each subunit contains its own ATP-binding site. The structural core of a NADSYN monomeric unit reveals a variant of a Rossmann fold characteristic of the ATP-dependent amidotransferase superfamily. The key aspects of catalytic mechanism apparently shared by all NADSYN enzymes include the important role of coordinated Mg<sup>2+</sup> ions in (1) activation of the ATP  $\alpha$ -phosphate for the nucleophilic attack by the carboxylate of NaAD, (2) stabilization of the PP<sub>i</sub>-leaving group, and (3) interaction with the AMP-leaving group facilitating the collapse of the tetrahedral intermediate formed after the addition of NH<sub>3</sub>.

The first and the best characterized representative of an eukaryotic Gln-utilizing NADSYN group (*long form* containing GAT domain) is a product of the essential *Qns1* gene of *S. cerevisiae*.<sup>196,197</sup> Using site-directed mutagenesis, it was shown that the N-terminal nitrilase-like domain of the yeast NADSYN confers a glutaminase activity providing the NH<sub>3</sub> molecule that is channeled to the active site of the C-terminal amidoligase domain.<sup>196</sup> Although other structurally unrelated classes of GAT domains or subunits functioning with various ATP-dependent amidotransferases were previously described, NADSYN is the first and so far the only example of an amidotransferase with a nitrilase-like GAT domain. The nitrilase superfamily includes a large number of amidases with different specificities and biological functions that are broadly represented in all taxons. Despite their structural and functional diversity, members of this superfamily (including the GAT domain of NADSYN) appear to share the overall  $\alpha$ - $\beta$ - $\beta$ - $\alpha$  sandwich fold and a conserved Glu-Lys-Cys catalytic triad motif.<sup>185</sup> Further studies on the yeast NADSYN revealed a kinetic cooperativity between the two domains, which explained a very low level of futile Gln consumption by this enzyme in the absence of NaAD substrate. The observed approximately 50-fold activation of the GAT domain is manifested largely at the level

of  $k_{\text{cat}}$ , and it appears to be triggered by a series of interactions between several amino acids that compose a predicted  $\text{NH}_3$  channel.<sup>198</sup> Two-domain orthologues of Qns1 are present in all sequenced eukaryotic genomes. The gene encoding human NADSYN (NADSYN1) was cloned and overexpressed, and its ability to utilize Gln as N-source *in vitro* was confirmed.<sup>199</sup> The same study mistakenly reported the existence of a second ‘bacterial-like’ human gene termed NADSYN2, which, in fact, was a result of contamination by bacterial DNA, most likely from a species of *Pseudomonas*.<sup>200</sup>

Although NADSYNs containing nitrilase-like GAT domains are widespread in bacteria, they are relatively less studied compared to classical single-domain enzymes. NADSYN from *M. tuberculosis* was characterized in most detail,<sup>188,201</sup> including site-directed mutagenesis and kinetic analysis that confirmed the role of the N-terminal domain and its catalytic triad in Gln utilization. Most probably, a good number of the bacterial *nadE* homologues containing a nitrilase-like domain belong to the group of Gln-utilizing NADSYN. However, some exceptions are possible, as in the case of NADSYN from *Rhodobacter capsulatus* that was reported to be a strictly  $\text{NH}_3$ -utilizing enzyme despite the presence of a nitrilase-like N-terminal domain.<sup>184</sup> Two groups of *long* and *short* bacterial NADSYNs display a mosaic phylogenetic distribution (with some species containing both types of NADSYNs, e.g., *T. maritima* or *Burkholderia pseudomallei*), pointing to a complex evolutionary history of the NadE family. For example, in the group of Enterobacteriaceae, *Yersinia* and several related species (*Serratia*, *Photobacterium*, and *Proteus*) contain a long form of *nadE*, whereas *E. coli*, *Salmonella*, and *Shigella* spp. have a short *nadE* gene. A comparative sequence analysis suggests that the long NADSYN may be ancestral for the Enterobacteriaceae group, as its closest homologues are present in many other Gammaproteobacteria (e.g., in Alteromonadales). At the same time, the closest homologues of the short NADSYN are all from Gram-positive bacteria (from the group of Bacillales), pointing to a likely HGT followed by the loss of a functionally redundant ancestral long *nadE* gene that could have occurred prior to the speciation within the *E. coli*/*Salmonella* branch. Although the physiological significance of such a gene replacement in the group of Enterobacteriaceae and, more generally, the presence of one or another form of NADSYN in various bacteria is unclear, it may be associated with certain aspects of nitrogen metabolism in these species.

#### 7.08.2.4.2 NMN synthetase

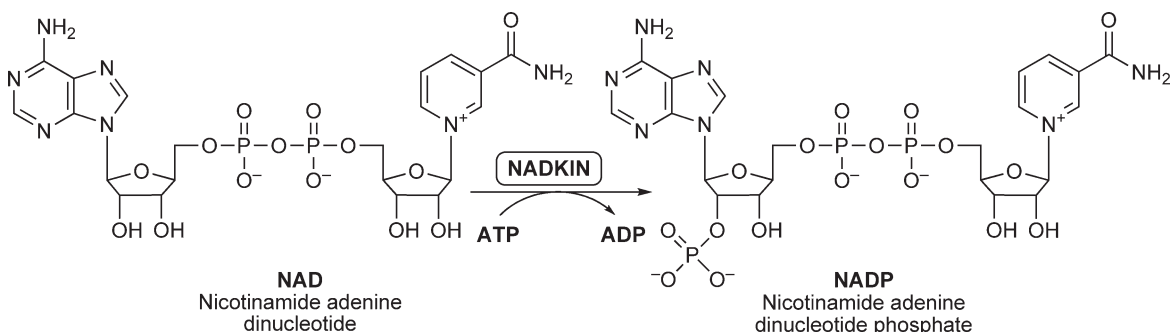
The first representative of the NMNSYN (NadE') subfamily within the NadE family was identified in *F. tularensis*.<sup>19,169</sup> An unusual substrate preference (for NaMN instead of NaAD) was inferred and experimentally verified for a single-domain member of the NadE family identified in the genome of this bacterial pathogen. This inference was based on the genomic reconstruction of NAD biosynthesis in *F. tularensis* that included the *de novo* synthesis of NaMN in the absence of the NaMNAT enzyme (of the NadD family), which would be able to provide the NaAD substrate for the regular NADSYN reaction. As already mentioned, the novel NMNSYN together with the NMNAT enzyme (of the NadM family) compose an alternative Route II of NAD synthesis from NaMN (Section 7.08.2.3). Despite the obvious chemical analogy of both reactions that are shown in **Scheme 8**, the existence of NMNSYN enzymatic activity was not reported in any of the previously studied species. This is not surprising, as all these species appear to harbor the conventional Route I of NAD synthesis that requires only NADSYN activity. In agreement with that, the NMNSYN activity of several tested NadE enzymes from other Gram-positive and Gram-negative bacteria was at least 1000-fold lower than their NADSYN activity. At the same time, the recombinant purified NadE' enzyme from *F. tularensis* displayed an approximately 60-fold preference for the NaMN over the NaAD substrate.<sup>19</sup> The 3D structure of this enzyme revealed a substantial similarity with other bacterial NadE enzymes at the level of the overall fold, arrangement of subunits in the homodimer, and conservation of active site residues involved in interactions with ATP,  $\text{NH}_3$ , and the nicotinosyl moiety of the substrate. At the same time, some structural differences were identified that provided a rationale for the unusual substrate preference of the NadE' subfamily.<sup>19</sup> A global comparative sequence analysis within the NadE family across more than 500 completely sequenced bacterial genomes suggested that the NadE' subfamily includes the enzymes from two species of Pasteurellaceae, *Mannheimia succinoproducens* and *Actinobacillus succinogenes*. Genomic reconstruction of NAD biogenesis in these species is consistent with their proposed NMNSYN activity (Section 7.08.3.1).

### 7.08.2.5 NADP Formation

The only known route leading to the formation of NADP cofactor is via the phosphorylation of NAD by NAD kinase (NADKIN), as shown in **Scheme 9**. Members of the NADKIN family are present in nearly all species with completely sequenced genomes, except some bacterial endosymbionts (such as *Chlamydia* spp.) that also lack the rest of the NAD biosynthetic machinery and salvage both NAD and NADP cofactors from their eukaryotic host. Despite such a high level of conservation, the absence of NADKIN orthologues in some deep-branched bacterial genomes (e.g., *Thermus* and *Deinococcus* groups) suggests the existence of another yet-unknown nonorthologous form of NADKIN.

#### 7.08.2.5.1 NAD kinase

The first gene representing the NADKIN family was identified and cloned from *M. tuberculosis* and termed *ppnK* for the ability of the respective approximately 30-kDa recombinant protein to utilize polyphosphate (poly(P)) as an alternative phosphoryl donor in conversion of NAD into NADP.<sup>202</sup> The same dual (ATP/poly(P)) specificity was later confirmed for a number of other NADKINs from Gram-positive bacteria and archaea.<sup>203–206</sup> However, the enzymes from *E. coli* (gene *nadK*, formerly *yjfB*<sup>207</sup>), *S. enterica*,<sup>20</sup> and *Sphingomonas* sp. *AI*,<sup>208</sup> as well as their eukaryotic orthologues,<sup>209–211</sup> can utilize only ATP and, less efficiently, other nucleoside triphosphates. Structure–function studies of the two bacterial members of the NADKIN family from *M. tuberculosis* and *Listeria monocytogenes*<sup>205,212,213</sup> and of an archaeal enzyme from *A. fulgidus* confirmed that this family shares the overall fold as well as some aspects of the active site organization and catalytic mechanism with other members of the phosphofructokinase (Pfk) superfamily.<sup>214</sup> The NADKIN N-terminal domain contains the Gly-Gly-Asp-Gly sequence motif highly conserved in the entire Pfk superfamily. However, the proposed mechanistic role of its universally conserved Asp residue is in the activation of the 2'-hydroxyl group of the adenylyl moiety of NAD substrate rather than in ATP activation, as in other Pfk enzymes, providing a rationale for the 2'-OH specificity of NAD phosphorylation by NADKIN.<sup>205</sup> A substrate-assisted catalysis via participation of the NAD diphosphate moiety in coordination of catalytic Mg<sup>2+</sup> ions was proposed. However, some important aspects of the NADKIN catalytic mechanism, such as the alternating ATP/poly(P) substrate specificity, remain to be elucidated. Allosteric regulation of NADKIN plays an important physiological role in *Salmonella* by modulating a balance between the pools of NAD and NADP cofactors that are involved in different aspects of the cellular metabolism.<sup>20</sup> The NADKIN activity appears to be suppressed by NADPH or NADH during aerobic or anaerobic growth, respectively. The enzyme activation in response to the drop in either reduced cofactor's pool allows cells to sustain the stress and maintain the required level of reductive biosynthesis.<sup>20</sup> Another type of allosteric regulation by Qa was described for NADKIN from *B. subtilis*.<sup>203</sup> A physiological role of NADKIN in the regulation of NAD(P) metabolism in human cells was studied using cell lines with decreased and increased expression of the single gene encoding this enzyme.<sup>215</sup> Properties of eukaryotic NADKIN enzymes, including multiple isoforms that are encoded in yeast and plant genomes, and their involvement with the regulation of the cellular redox balance, signaling, and oxidative stress resistance were recently reviewed.<sup>21</sup> NADKIN was proven essential for survival in a number of pathogenic bacteria, and this enzyme is considered a potential target for the development of new antibiotics. Despite a substantial similarity between bacterial and human enzymes, some differences



**Scheme 9**

revealed by the detailed structure–function analysis open opportunities for the design of selective inhibitors and drugs.<sup>216</sup>

### 7.08.3 Genomic Reconstruction of NAD Biosynthetic Pathways

Sequencing and comparative analysis of multiple genomes opened new opportunities for the accurate reconstruction of metabolic pathways and networks in a broad range of previously uncharacterized species. Metabolic reconstruction was introduced and established as one of the most powerful techniques of comparative genomics that allows us to infer biochemical transformations in the organism based solely on the presence of genes tentatively assigned respective functional roles (primarily enzymes).<sup>217,218</sup> It is important to emphasize that this approach, in addition to furnishing the first approximation of an organism's metabolic potential, yields an important set of constraints that help improve the accuracy of underlying genomic annotations. Therefore, both aspects of genome analysis – functional annotation of genes and reconstruction of respective pathways – taken together, constitute an iterative process of ever-increasing accuracy and coverage.<sup>219</sup> Among the main factors contributing to the success of this enterprise is an imposed consistency, which is quite straightforward for metabolic, especially biosynthetic, pathways that are presumed to be contiguous with balanced and physiologically relevant inputs and outputs. Therefore, finding a gap (a *missing gene*) in an otherwise complete pathway would help detect an omission or a mistake in the underlying genomic annotations that can be corrected by additional homology searches. However, a gap in a pathway that is observed in a group of genomes would suggest that these genomes may contain an alternative (nonorthologous) form of a gene with the same functional role. Several cases of nonorthologous gene replacements were briefly discussed in the previous section. Genome context analysis techniques, most importantly clustering of functionally coupled genes on the prokaryotic chromosome, domain fusions, phylogenetic profiling, and analysis of conserved regulatory sites, help tentatively identify gene candidates to fill in missing genes that can be further experimentally tested.<sup>220</sup> These techniques are particularly efficient for gene and pathway discovery in prokaryotes. Examples illustrating their application for the identification of previously uncharacterized bacterial genes of NAD metabolism include KYNFA,<sup>57</sup> NAMNAT,<sup>120</sup> RNMKIN,<sup>95</sup> niacin transporter NiaP,<sup>90</sup> and transcriptional regulator NrtR<sup>155</sup> (see Section 7.08.3.1.2). At the same time, the observed pathway gaps or other inconsistencies, such as genes whose assigned functions do not seem to fit in the context of reconstructed pathways, in some cases help infer novel pathway variants, as illustrated here by the identification of Route II of NAD synthesis in *F. tularensis*. A combination of comparative genomic techniques mentioned above with reconstruction of pathways and metabolic scenarios is captured in a subsystems-based approach implemented within The SEED genomic platform.<sup>11,12</sup> General principles of this approach and its implications for the analysis of vitamin and cofactor metabolism are illustrated in Chapter 7.05 of this volume. In this section, we provide several examples of genomic reconstruction of NAD biosynthetic pathways in diverse species. These examples (illustrated in **Table 1** and **Figures 3, 5, and 6**) were extracted from the global analysis of the NAD(P) biosynthesis subsystem across the entire collection of more than 700 completely sequenced genomes integrated in The SEED platform and available online (at <http://seed-viewer.theseed.org/>).

#### 7.08.3.1 Bacteria

##### 7.08.3.1.1 Biosynthetic pathways

NaMN is the most common mononucleotide intermediate (a hub) in NAD biogenesis. For example, in *E. coli* all three pyridine precursors are converted into NaMN (**Table 1** and **Figure 3(a)**). Qa produced by the *de novo* Asp–DHAP pathway (genes *nadB* and *nadA*) is converted into NaMN by QAPRT (gene *nadC*). Salvage of both forms of niacin proceeds via NAPRT (gene *pncB*) either directly upon or after deamidation by NMDSE (gene *pncA*). Overall, more than 90% of approximately 680 analyzed bacterial genomes contain at least one of the pathways leading to the formation of NaMN. Most of them (~480 genomes) have the entire set of *nadBAC* genes for NaMN *de novo* synthesis from Asp that are often clustered on the chromosome and/or are co-regulated by the same transcription factors (see Section 7.08.3.1.2). Among the examples provided in **Table 1**, *F. tularensis* (**Figure 4(c)**) has all three genes of this *de novo* pathway forming a single operon-like cluster and supporting the



growth of this organism in the absence of any pyridine precursors in the medium. More than half the genomes with the Asp–DHAP pathway also contain a deamidating niacin salvage pathway (genes *pncAB*) as do many representatives of the  $\alpha$ -,  $\beta$ -, and  $\gamma$ -Proteobacteria, Actinobacteria, and Bacillus/Clostridium group. As already emphasized, the genomic reconstruction approach provides an assessment of the metabolic potential of an organism, which may or may not be realized under given conditions. For example, *E. coli* and *B. subtilis* can utilize both *de novo* and PncAB Nm salvage pathways under the same growth conditions, whereas in *M. tuberculosis* (having the same gene pattern) the latter pathway was considered nonfunctional, so that the entire NAD pool is generated by the *de novo* NadABC route. However, a recent study demonstrated the functional activity of the Nm salvage pathway *in vivo*, under hypoxic conditions in infected macrophages.<sup>221</sup> This study also implicated the two downstream enzymes of NAD synthesis (NAMNAT and NADSYN) as attractive chemotherapeutic targets to treat acute and latent forms of tuberculosis.

In approximately 100 species, including many Cyanobacteria (e.g., *Synechococcus* spp.), Bacteroidetes (e.g., *Cblorobium* spp.) and Proteobacteria (e.g., *Caulobacter crescentus*, *Zymomonas mobilis*, *Desulfovibrio* spp., and *Shewanella* spp. representing  $\alpha$ -,  $\beta$ -,  $\delta$ -, and  $\gamma$ -groups, respectively) the Asp–DHAP pathway is the only route to NAD biogenesis. Among them, nearly all *Helicobacter* spp. (except *H. hepaticus*), contain only the two genes *nadA* and *nadC* but lack the first gene of the pathway (*nadB*), which is a likely subject of nonorthologous gene replacement. One case of NadB (ASPOX) replacement by the ASPDH enzyme in *T. maritima* (and methanogenic archaea) was discussed in Section 7.08.2.1. However, no orthologues of the established ASPDH could be identified in *Helicobacter* spp. as well as in approximately 15 other diverse bacterial species that have the *nadAC* but lack the *nadB* gene (e.g., all analyzed *Corynebacterium* spp. except for *C. diphtheriae*). Therefore, the identity of the ASPOX or ASPDH enzyme in these species is still unknown, representing one of the few remaining cases of ‘locally missing genes’<sup>220</sup> in the NAD subsystem. All other bacterial species contain either both the *nadA* and *nadB* genes (plus *nadC*) or none.

In a limited number of bacteria (~20 species), mostly in the two distant groups of Xanthomonadales (within  $\gamma$ -Proteobacteria) and Flavobacteriales (within Bacteroidetes), the Asp–DHAP pathway of Qa synthesis is replaced by the Kyn pathway. As described in Section 7.08.2.1.2, four out of five enzymes (TRDOX, KYNOX, KYNSE, and HADOX) in the bacterial version of this pathway are close homologues of the respective eukaryotic enzymes, whereas the KYNFA gene is a subject of multiple nonorthologous replacements. Although the identity of one alternative form of KYNFA (gene *kynB*) was established in a group of bacteria that have a partial Kyn pathway for Trp degradation to anthranilate (e.g., in *P. aeruginosa* or *B. cereus*<sup>57</sup>), none of the known KYNFA homologues are present in Xanthomonadales or Flavobacteriales. In a few species (e.g., *Salinispora* spp.) a complete gene set of the Kyn pathway genes co-occurs with a complete Asp–DHAP pathway. Further experiments would be required to establish to what extent and under what conditions these two pathways contribute to Qa formation. As discussed, the QAPRT enzyme is shared by both *de novo* pathways, and a respective gene, *nadC* is always found in the genomes containing one or the other pathway. Similarly, gene *nadC* always co-occurs with Qa *de novo* biosynthetic genes with one notable exception of two groups of Streptococci, *S. pneumoniae* and *S. pyogenes*. Although all other members of the Lactobacillales group also lack the Qa *de novo* biosynthetic machinery and rely entirely on niacin salvage, only these two human pathogens contain a *nadC* gene. The functional significance of this ‘out of context’ gene is unknown, but it is tempting to speculate that it may be involved in a yet-unknown pathway of Qa salvage from the human host.

Among approximately 150 bacterial species that lack *de novo* biosynthesis genes and rely on deamidating salvage of niacin (via NAPRT), the majority (~100) are from the group of Firmicutes. Such a functional variant (illustrated for *Staphylococcus aureus* in Figure 4(b)) is characteristic of many bacterial pathogens, both Gram-positive and Gram-negative (e.g., *Brucella*, *Bordetella*, and *Campylobacter* spp. from  $\alpha$ -,  $\beta$ -, and  $\delta$ -Proteobacteria, *Borrelia*, and *Treponema* spp. from Spirochaetes). Most of the genomes in this group contain both *pncA* and *pncB* genes that are often clustered on the chromosome and/or are co-regulated (see Section 7.08.3.1.2). In some cases (e.g., within Mollicutes and Spirochaetales), only the *pncB*, but not the *pncA* gene, can be reliably identified, suggesting that either of these species can utilize only the deamidated form of niacin (Na) or that some of them contain an alternative (yet-unknown) NMASE.

Although the nondeamidating conversion of Nm into NMN (via NMPRT) appears to be present in approximately 50 bacterial species (mostly in  $\beta$ - and  $\gamma$ -Proteobacteria), it is hardly ever the only route of NAD biogenesis in these organisms. The only possible exception is observed in *Mycoplasma genitalium* and

*M. pneumoniae* that contain the *nadV* gene as the only component of pyridine mononucleotide biosynthetic machinery. In some species (e.g., in *Synechocystis* spp.), the NMPRT–NMNAT route is committed primarily to the recycling of endogenous Nm. On the other hand, in *F. tularensis* (Figure 4(c)), NMPRT (gene *nadV*) together with NMNAT (of the *nadM* family) constitute the functional nondeamidating Nm salvage pathway as it supports the growth of the *nadE'*-mutant on Nm but not on Na (L. Sorci *et al.*, unpublished). A similar nondeamidating Nm salvage pathway implemented by NMPRT and NMNAT (of the *nadR* family) is present in some (but not all) species of Pasteurellaceae in addition to (but never instead of) the RNm salvage pathway (see below), as initially demonstrated for *H. ducreyi*.<sup>128</sup>

A two-step conversion of NaMN into NAD via a NaAD intermediate (Route I in Figure 2) is present in the overwhelming majority of bacteria. The signature enzyme of Route I, NAMNAT of the NadD family is present in nearly all approximately 650 bacterial species that are expected to generate NaMN via *de novo* or salvage pathways (as illustrated by Figures 3(a) and 3(b)). All these species, without a single exception, also contain NADSYN (encoded by either a short or a long form of the *nadE* gene), which is required for this route. The species that lack the NadD/NadE signature represent several relatively rare functional variants, including:

1. *Route I of NAD synthesis (NaMN → NaAD → NAD) variant* via a bifunctional NAMNAT/NMNAT enzyme of the NadM family is common for archaea (see Section 7.08.3.2), but it appears to be present in only a handful of bacteria, such as *Acinetobacter*, *Deinococcus*, and *Thermus* groups. Another unusual feature of the latter two groups is the absence of the classical NADKIN, a likely subject of a nonorthologous replacement that remains to be elucidated.
2. *Route II of NAD synthesis (NaMN → NMN → NAD)*. This route is implemented by a combination of the NMNAT of either the NadM family (as in *F. tularensis*) or the NadR family (as in *M. succinoproducens* and *A. succinogenes*) with NMNSYN of the NadE' family. The case of *F. tularensis* described in Section 7.08.2.4 is illustrated in Figure 3(b). The rest of the NAD biosynthetic machinery in both species from the Pasteurellaceae group, beyond the shared Route II, is remarkably different from that in *F. tularensis*. Instead of *de novo* biosynthesis, they harbor a Na salvage pathway via NAPRT encoded by a *pncB* gene that is present in a chromosomal cluster with *nadE'*. Neither of these two genes are present in other Pasteurellaceae that lack the pyridine carboxylate amidation machinery (see below).
3. *Salvage of RNm (RNm → NMN → NAD)*. A genomic signature of this pathway, a combination of the PnuC-like transporter and a bifunctional NMNAT/RNMKIN of the NadR family, is present in many Enterobacteriaceae and in several other diverse species (e.g., in *M. tuberculosis*). However, in *H. influenzae* (Figure 3(d)) and related members of Pasteurellaceae, it is the only route of NAD biogenesis. As shown in Table 1, *H. influenzae* as well as many other members of this group have lost nearly all components of the rich NAD biosynthetic machinery that are present in their close phylogenetic neighbors (such as *E. coli* and many other Enterobacteriaceae). This pathway is an ultimate route for utilization of the so called V-factors (NADP, NAD, NMN, or RNm) that are required to support growth of *H. influenzae*. It was established that all other V-factors are degraded to RNm by a combination of periplasmic- and membrane-associated hydrolytic enzymes.<sup>222</sup> Although PnuC was initially considered an NMN transporter,<sup>223</sup> its recent detailed analysis in both *H. influenzae* and *Salmonella* confirmed that its actual physiological function is in the uptake of RNm coupled with the phosphorylation of RNM to NMN by RNMKIN.<sup>17,148,224</sup> As already mentioned, *H. ducreyi* and several other V-factor-independent members of the Pasteurellaceae group (*H. somnus*, *Actinobacillus pleuropneumoniae*, and *Actinomyces comitans*) harbor the NMNAT enzyme (NadV) that allows them to grow in the presence of Nm (but not Na) in the medium (Section 7.08.2.2).
4. *Uptake of the intact NAD*. Several groups of phylogenetically distant intracellular endosymbionts with extremely truncated genomes contain only a single enzyme, NADKIN, from the entire subsystem. Among them are all analyzed species of the *Wolbachia*, *Rickettsia*, and *Blochmannia* groups. These species are expected to uptake and utilize the intact NAD from their host while retaining the ability to convert it into NADP. Among all analyzed bacteria, only the group of *Chlamydia* does not have NADKIN and depends on the salvage of both NAD and NADP via a unique uptake system.<sup>157</sup>

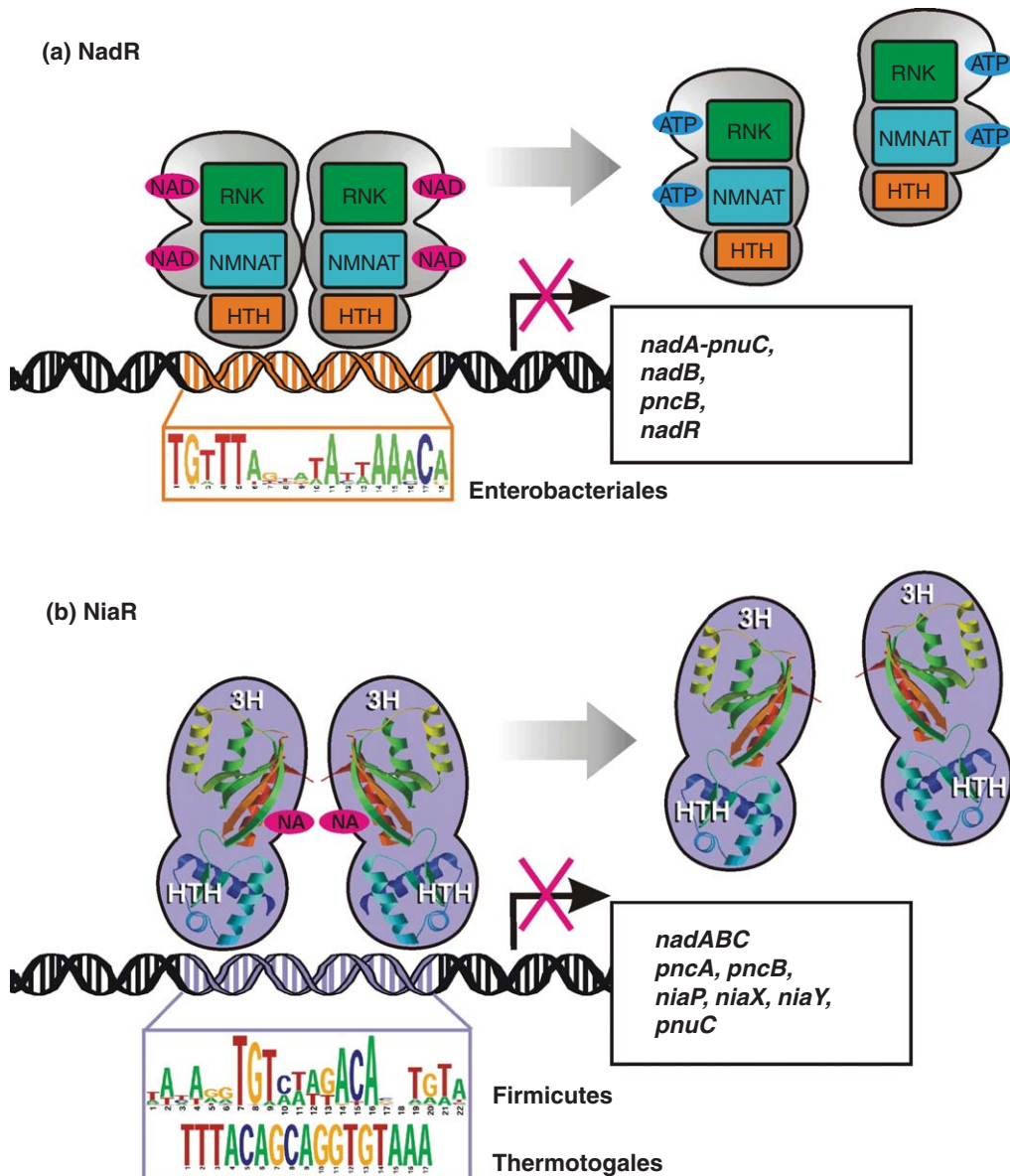
A comprehensive genomic reconstruction of the metabolic potential (gene annotations and asserted pathways) across approximately 680 diverse bacterial genomes sets the stage for the accurate cross-genome projection and prediction of regulatory mechanisms that control the realization of this potential in a variety of species and



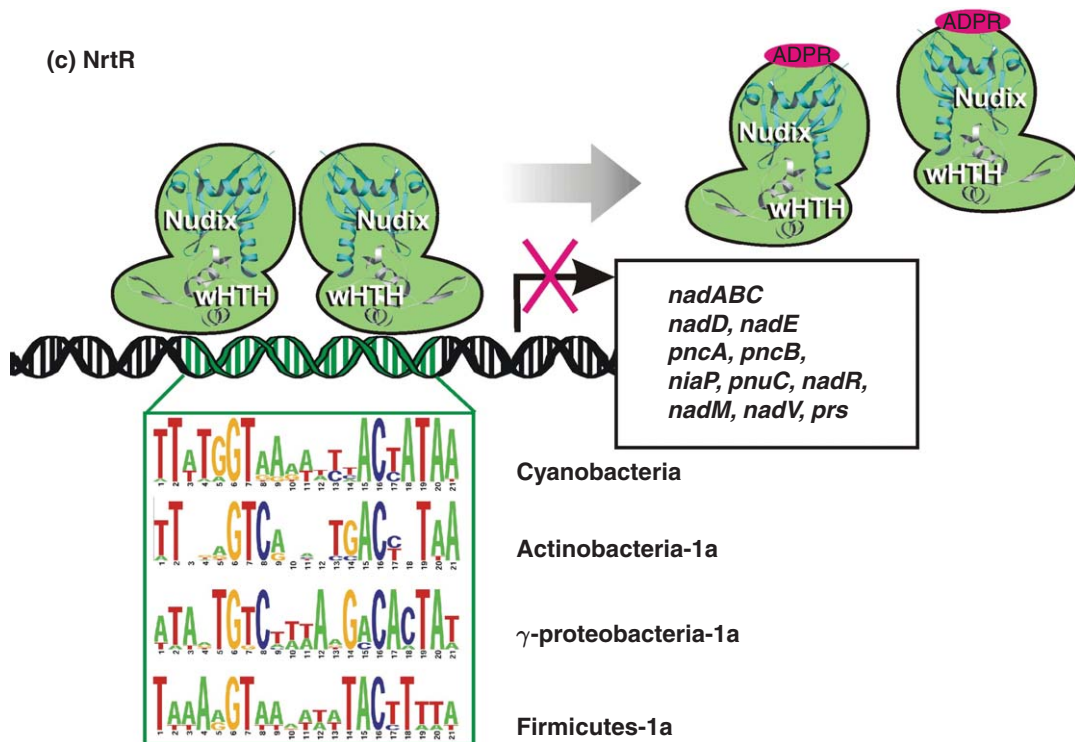
growth conditions. In the next section, we summarize the recent accomplishments in the genomic reconstruction of NAD-related regulons in bacteria.

### 7.08.3.1.2 Transcriptional regulation

Three distinct protein families, NadR,<sup>147,225,226</sup> NiaR,<sup>48,90</sup> and NrtR,<sup>155</sup> were implicated in transcriptional regulation of NAD biogenesis in various groups of bacteria. The key features of these regulators are illustrated in **Figure 4**. The results of projection and reconstruction of respective regulons by comparative genomic techniques (as reviewed in Rodionov<sup>14</sup>) are briefly described below.



**Figure 4** (Continued)



**Figure 4** Transcriptional regulation of NAD biosynthesis in bacteria. (a) NadR repressor controls initiation of transcription of its target genes (box) by specific binding to DNA motifs (a sequence logo is shown in red box) in the presence of its co-repressor, NAD. (b) NiaR functions as a transcriptional repressor of its target genes (box) in the presence of its co-repressor, Na. Firmicutes and Thermotogales have different consensus sequences of NiaR-binding motifs (logo in blue box). (c) NrtR binds its operator sites and represses the transcription of target genes (box). Various taxonomic groups of bacteria have divergent NrtR-recognition motifs (examples are shown in the green box). A product of NAD catabolism, ADP-ribose acts as an activator promoting dissociation of NrtR from DNA.

**7.08.3.1.2(i) NadR family** NAD biosynthesis in *E. coli* and *Salmonella* is controlled on the transcriptional level by the trifunctional NadR protein. It was shown that NadR operates as a transcriptional repressor for the three genomic loci containing genes involved in the *de novo* NAD biosynthesis (*nadA* and *nadB*) and in the salvage routes (*pncB*, *pnuC*, and *nadR*).<sup>147,227–230</sup> In addition to its role as transcriptional repressor mediated by the N-terminal helix-turn-helix (HTH) DNA-binding domain, NadR is also endowed with two enzymatic activities encoded in its central NMNAT domain and C-terminal RNMKIN domain (Section 7.08.3). A dimer of NadR specifically binds to an 18-bp palindromic DNA operator (see Figure 4(a) for a consensus sequence) upstream of the promoter of the regulated genes.<sup>229</sup> NAD operates as a co-repressor promoting NadR–DNA binding. In the absence of NAD, ATP operates as an antirepressor. According to a mechanistic model of NAD-dependent transcriptional regulation by NadR, the NMNAT domain operates as an effector domain that modulates the DNA-binding activity of the HTH domain, depending on the nature of the bound ligand. The HTH-containing three-domain NadR proteins are present only in the Enterobacteriaceae lineage. Orthologues of the NadR protein found in other bacterial groups lack an HTH domain, retaining only its enzymatic (but not regulatory) functions. A comparative genomic analysis of the NadR regulon revealed the conservation of its binding DNA motif and allowed the reconstruction of this regulon in all sequenced genomes of Enterobacteriaceae.<sup>225</sup> A remarkable difference in the NadR regulon content was revealed for several species. First, the *pncB* and *nadB* genes, the known regulon members in *E. coli*/*Salmonella*, lack candidate NadR-binding sites in *Yersinia*/*Erwinia* species. Second, the *nadR* gene in the latter species was found to be a subject of autoregulation as it is preceded by a high-score candidate NadR-binding site. The only component of the NadR regulon conserved in all Enterobacteriaceae genomes is the *nadA*–*pnuC* operon.

**7.08.3.1.2(ii) NiaR family** A different mechanism of regulation of NAD biosynthesis was discovered in *B. subtilis*.<sup>48</sup> A DNA-binding protein encoded by a previously uncharacterized gene *yrxA* (later termed *niaR*) was shown to operate as a Na-responsive transcriptional repressor of the *de novo* biosynthesis operon *nadABC*. Further analysis of the NiaR regulon by a combination of comparative genomic and experimental techniques allowed us to refine its DNA-binding motif and identify additional regulon members (including a novel Na/Nm transporter, NiaP) in a broad range of bacteria.<sup>90</sup> NiaR orthologues display a mosaic distribution among bacteria from the *Bacillus/Clostridium* group, and they were also detected in the *Thermotogales* lineage and in *Fusobacterium nucleatum*. In contrast to other families of transcription factors that have multiple paralogues in a genome (e.g., LacI and LysR), NiaR is always present in a single copy committed to the regulation of NAD biosynthetic genes.

The 3D structure of the NiaR orthologue from *T. maritima* (TM1602, PDB accession number 1J5Y,<sup>231</sup>) showed that its domain organization is similar to many other transcriptional regulators. The N-terminal HTH domain is remotely similar to the DNA-binding domain of the biotin repressor. The C-terminal 3H domain (containing three conserved His residues) is likely involved in the binding of the niacin co-repressor. NiaR, like many other bacterial transcription factors, tends to regulate adjacent genes, such as the *nadBAC* operon in the Bacillales order. In *Streptococcus* and *Thermotoga* species, the *niaR* gene is clustered on the chromosome with the candidate niacin transporter gene. The NiaR 22-bp palindromic DNA-binding motif (**Figure 4(b)**) is conserved upstream of NAD biosynthetic genes in all representatives of the Firmicutes group that contain NiaR orthologues. In the Thermotogales lineage, a distinct 17-bp palindromic NiaR-binding motif (**Figure 4(b)**) was tentatively identified upstream of the *nadBCA* and *niaR-niaP* operons and experimentally verified by the electrophoretic mobility shift assay.<sup>90</sup>

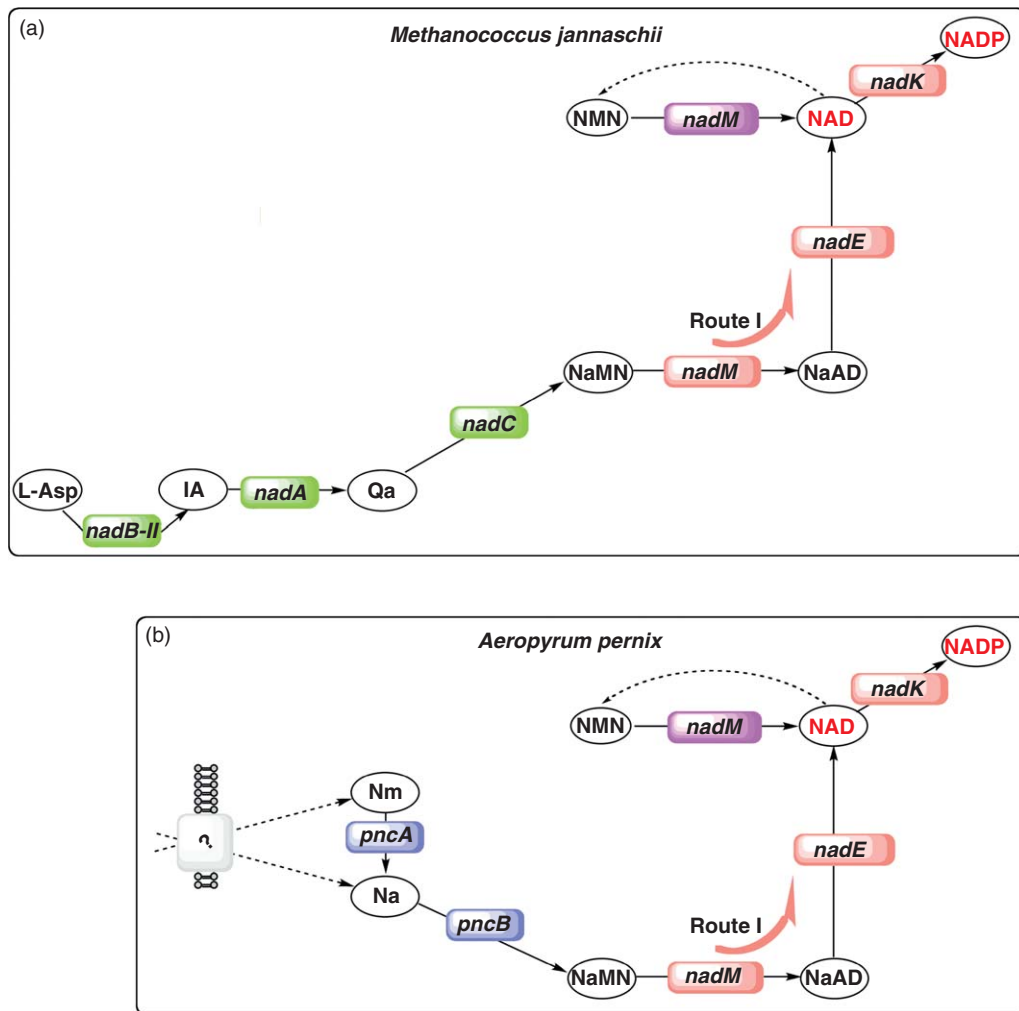
**7.08.3.1.2(iii) NrtR family** The comparative genome context analysis revealed a novel family of Nudix-related transcription factors, termed NrtR, that are responsible for the regulation of various aspects of NAD synthesis in diverse bacterial species.<sup>155</sup> In many analyzed genomes, the *nrtR* gene is located in close proximity to various NAD biosynthetic genes, such as *nadABC* in Actinobacteria, *nadDE* in Cyanobacteria, *pncBA* in some  $\gamma$ -Proteobacteria, and *nadR-pnuC* in *Streptococcus* spp. Members of the NrtR family possess an N-terminal Nudix-like domain and a C-terminal DNA-binding domain with the winged HTH-fold characteristic of many prokaryotic transcription factors.<sup>232</sup> The N-terminal Nudix-like domain of NrtR is most similar to the ADPPR enzyme (Section 7.08.2.3), although many (but not all) members of the NrtR family have lost a characteristic Nudix-hydrolase signature<sup>233,234</sup> and the respective catalytic activity. A genomic reconstruction of the NrtR regulon revealed a substantial variability in NrtR-binding DNA motifs and regulon composition between various taxonomic groups. Although 18 derived NrtR motifs differ substantially in consensus sequence, most of them share a 21-bp palindrome symmetry and a conserved core with consensus GT-N<sub>7</sub>-AC (**Figure 4(c)**). Among NrtR-regulated genes in different taxonomic groups are those involved in the *de novo* NAD biosynthesis and pyridine salvage pathways (*nadABC*, *nadD*, *nadE*, *pncAB*, *nadMV*), as well as in niacin uptake (*niaP*), RNm salvage (*pnuC-nadR*), and PRPP synthesis (*prs*).

A specific binding to the predicted DNA motifs was experimentally confirmed for the two diverse representative NrtR proteins from *Synechocystis* sp. and *S. oneidensis*.<sup>155</sup> In a series of tested intermediates associated with NAD metabolism, ADP-ribose, the product of glycohydrolytic cleavage of NAD, was found to suppress the *in vitro* binding of NrtR proteins to their DNA target sites. The cellular pool of ADP-ribose may reflect the extent of NAD consumption by various NAD glycohydrolases (such as CobB), and the increased level of this metabolite may be interpreted by a cell as a signal to replenish the cofactor pool by derepression of NAD biosynthetic genes and operons.

Although the three classes of regulons described above cover approximately 130 bacterial genomes, the regulation of NAD synthesis in many other bacterial species remains a subject of further studies.

## 7.08.3.2 Archaea

The degree of variation in archaeal NAD biosynthetic pathways is substantially lower than among bacteria, and all the identified functional variants have their exact bacterial analogues. The most common functional variant that is present in approximately 50% of the 45 analyzed archaeal genomes is illustrated by the example of *M. jannaschii* (**Figure 5(a)**). It includes both routes leading to NaMN synthesis, the *de novo* Asp-DHAP pathway



**Figure 5** Genomic reconstruction of NAD biosynthetic pathways in archaea. (a) *De novo* biosynthesis from Asp in *Methanococcus jannaschii* and many other methanogens involves an alternative (nonhomologous) enzyme – NAD-dependent ASPDH (*nadB-II*) – instead of the flavin-dependent ASPOX. In all presently available archaeal genomes, NaMNAT activity that is required for both *de novo* synthesis and Nm salvage is performed by the members of the NadM family with dual NaMNAT/ NMNAT substrate specificity. (b) The example of *Aeropyrum pernix* represents a relatively limited number of archaeal species that lack *de novo* biosynthesis and use Nm salvage as the only route of NAD synthesis.

and the deamidating salvage or recycling of Nm and/or Na via NAPRT (PncB). The set of species containing this functional variant includes 5 representatives of Crenarchaeota (the group of *Sulfolobus* and two members of Thermoproteaceae) and 17 species of Euryarchaeota (including members of Halobacteriaceae, Methanomicrobia, and Thermococcales). Among them, 10 species (including all Methanomicrobia as well as *Methanopyrus kandleri*, *Pyrobaculum aerophilum*, and *A. fulgidus*) use ASPDH instead of ASPOX (shown as gene *nadB-II* in **Figure 5**) in the first step of the Asp–DHAP pathway, whereas all others carry an orthologue of a canonical bacterial gene *nadB*. In another set of 12 genomes, predominantly euryarchaeal methanogens (representatives of Methanomicrobia, Methanobacteria, and Methanococci) and one member of Crenarchaeota (*Nitrosopumilus maritimus*), the archaeal (ASPDH-dependent) version of the *de novo* pathway is the only route of NaMN synthesis, as illustrated by the example of *M. jannaschii* (**Figure 5(a)**). *Methanosphaera stadtmanae* is the only organism in this set (and among all other analyzed archaeal genomes) that contains an orthologue of NMPRT (NadV) and thus may harbor a nondeamidating Nm salvage or recycling pathway. Another functional variant that is present in 10 archaeal genomes (5 species within Thermoprotei of



Crenarchaeota, 4 Thermoplasmata species from Euryarchaeota, and *Korarchaeum cryptofilum* from Korarchaeota) is illustrated by the example of *Aeropyrum pernix* (Figure 5(b)). These species lack *de novo* synthesis and rely fully on the salvage of Nm and/or Na as characteristic of some bacterial pathogens (compare with Figure 3(c)).

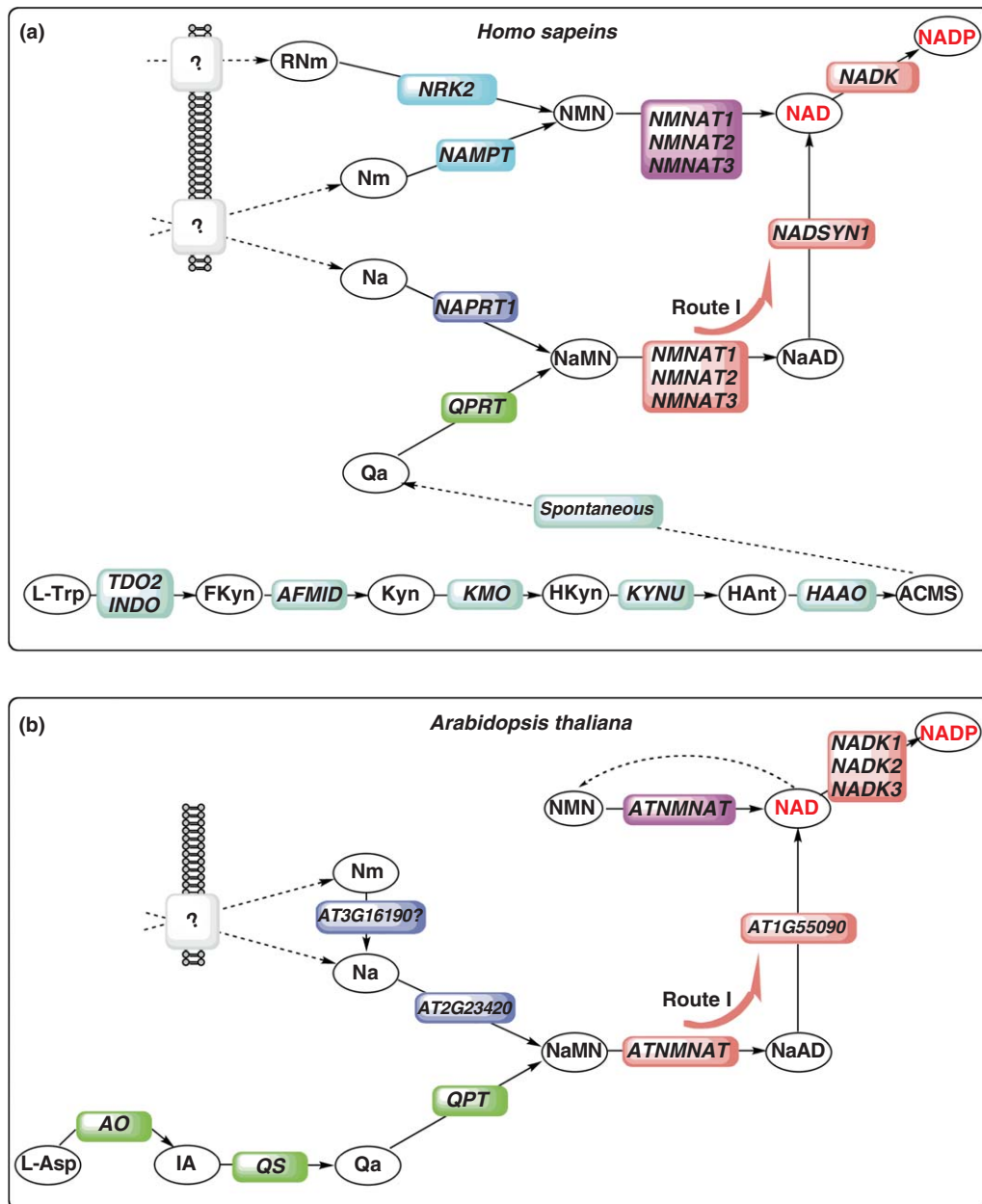
Therefore, all analyzed archaeal species synthesize NAD via the NaMN intermediate and share all downstream components of the NAD biosynthetic machinery including the NaMNAT/NMNAT enzymes of the NadM family, NADSYN and NADKIN. The only exceptional case is *Nanoarchaeum equitans*, which has an extremely truncated genome (~490 kb) and does not contain a single gene of NAD biosynthesis, reminiscent of *Chlamydia* spp. As previously discussed, the NaMNAT activity in archaea is performed by a single-domain NadM protein. A characterized representative of this family from *M. jannaschii* has a dual substrate specificity toward both NaMN and NMN substrates (L. Sorci *et al.*, unpublished). Although the physiological role of the NMNAT activity in these species (apart from the single case of *M. stadtmanae* containing the NMPRT enzyme) is largely unclear, it may be involved in a yet-unknown NMN recycling pathway. Most of the archaeal genomes contain a short form of NADSYN. However, four representatives of Euryarchaeota (two species of *Methanosaeta*, *Methanoculleus marisnigri* and *Haloquadratum walsbyi*), contain a long form of NADSYN with the N-terminal GAT domain. Among Crenarchaeota, three species of *Sulfolobus* and *Thermofilum pendens* contain chromosomal clusters where the short form of the *nadE* gene is located next to the gene encoding the GAT subunit, similar to the case of *T. thermophilus* described in Section 7.08.2.4.

Overall, the genomic reconstruction of NAD biosynthesis in archaea appears to be rather straightforward. Although only a handful of NAD biosynthetic enzymes from archaea were experimentally characterized,<sup>33,41,42,126,146,166,204,206</sup> the observed consistency leads to a substantial confidence in suggested gene and pathway assignments across the entire archaeal kingdom.

### 7.08.3.3 Eukaryota

Assessing the actual variability of NAD pathways in various branches of Eukaryota is difficult due to a small number and biased phylogenetic distribution of fully sequenced and analyzed eukaryotic genomes. However, from the analysis of 16 representative genomes (3 mammals, 1 fish, 1 frog, 4 insects, 1 nematode, 1 protozoa, 1 dicot plant, and 4 fungi), it appears that their variability is relatively higher than in archaea. This observation reflects upon the existence of an extensive NAD biosynthetic machinery, which, in the case of vertebrates, includes more than 15 distinct functional roles (Table 1 and Figure 6(a)). In that regard, NAD is quite different from many other cofactors (e.g., FAD and CoA) that are generated by vertebrates from their dietary precursors (vitamins). *Homo sapiens* and other analyzed vertebrates harbor several routes of NAD biogenesis, including the *de novo* synthesis from Trp, two distinct pyridine salvage pathways of Na (via NAPRT) and Nm (via NMPRT), as well as the pyridine nucleoside salvage (via RNMKIN and, possibly, via RNMPSE) as described in Section 7.08.2. The nematode *C. elegans* has a similar set of genes and pathways, except that it lacks the NMPRT but has the NMASE enzyme so that, in contrast with vertebrates, it may use the deamidating Nm salvage instead of the nondeamidating pathway. As discussed in Section 7.08.2.1.2, insects apparently lost two of the genes in the Kyn pathway of Qa biosynthesis, and their NAD synthesis is entirely dependent on the salvage of Nm and/or Na (via the NMASE–NAPRT pathway) or RNm (via RNMKIN). The NAD biosynthetic machinery in the protozoan parasite *Plasmodium falciparum* is truncated even further, reminding us of a similar minimization of NAD biosynthesis in bacterial pathogens. It does not contain any genes of the Kyn pathway, and it seems to contain only the Na salvage pathway (via NAPRT) merging to the regular downstream Route I of conversion of NaMN into NAD (by NaMNAT and NADSYN). *S. cerevisiae* and some (but not all) other fungi have a version of the complete Kyn pathway of Qa synthesis (see Section 7.08.2.1.2). Yeast as well as the other three analyzed genomes (*S. pombe*, *Magnaporthe grisea*, and *G. zaeae*) contain Nm/Na (NMASE–NAPRT) and RNm (RNMKIN) salvage pathways but lack the NMPRT-driven nondeamidating salvage of Nm.

The most substantial deviations are observed in *A. thaliana*, which has a bacterial-like version of Qa synthesis by the Asp–DHAP pathway instead of the Kyn pathway (Figure 6(b)). It was shown that in this organism the Qa synthesis and its conversion into NaMN occurs in plastids<sup>34</sup> consistent with an apparent bacterial origin of this pathway. The core downstream pathway and enzymes involved therein – (1) NaMNAT/NMNAT of the PNAT family; (2) a long form of NADSYN containing a GAT-domain; and (3) NADKIN – are strictly conserved in all analyzed eukaryotic genomes. The only remarkable exception is observed in *P. falciparum*,



**Figure 6** Genomic reconstruction of NAD biosynthetic pathways in eukaryotes. (a) NAD biosynthesis in *Homo sapiens* is among the best-studied eukaryotic model systems. Gene names are shown mostly as recommended by HUGO nomenclature (<http://www.genenames.org/>). In addition to the *de novo* synthesis from Trp, several salvage pathways are known, including Na salvage via the Preiss–Handler pathway and nondeamidating salvage of Nm (via NMPRT) and RNm (via RNMKIN). (b) *Arabidopsis thaliana* reveals an interesting hybrid between prokaryotic and eukaryotic features. It combines a *de novo* biosynthesis from Asp (as in bacteria and archaea) with the eukaryotic-type PNAT having a dual specificity toward NMN and NaMN.

which contains the adenylyltransferase, which is much closer to bacterial NMANT of the NadD family than to the eukaryotic PNAT.

Overall, despite the apparent variability and complexity of gene patterns associated with NAD biosynthesis, a genomic reconstruction approach revealed the existence of only several major metabolic scenarios and combinations thereof covering the entire space of approximately 770 analyzed genomes. A remarkable



consistency achieved in this analysis is illustrated by the fact that all the observed gene patterns (from 1 to 15 genes involved in NAD biosynthesis per organism) encode meaningful and essentially complete functional variants. Only isolated cases of gaps (missing genes) or apparent inconsistencies (genes out of context) could be detected as discussed in Section 7.08.2. This reflects both an essentially comprehensive knowledge of NAD biosynthesis that was accumulated in a handful of model species and our ability to accurately project this knowledge over hundreds of diverse genomes. For practical purposes, it is important to emphasize that while the initial effort on annotation and reconstruction of this subsystem required substantial manual analysis, its accurate projection over the rapidly growing body of newly sequenced genomes can now be successfully accomplished by fully automated annotation tools such as those included in the RAST server.<sup>235</sup> We believe that due to the apparent biochemical and evolutionary constraints imposed by nature on all living organisms, application of the subsystems approach would lead to similar results and conclusions with respect to the biosynthesis of other key metabolites, such as vitamins and cofactors (see Chapter 7.05).

## Acknowledgments

We are grateful to all our colleagues and collaborators involved in various aspects of the analysis of NAD biosynthesis, including G. Magni and N. Raffaelli (Polytechnic University of Marche, Ancona, Italy); H. Zhang (UT Southwestern Medical Center, Dallas, TX); T. Begley (Cornell University, Ithaca, NY); V. de Crecy-Lagard and I. Blaby (University of Florida, Gainesville, FL); and I. Rodionova, J. De Ingeniis, and X. Li (Burnham Institute for Medical Research, La Jolla, CA). Without their help and encouragement, this work would be impossible. We want to thank R. Overbeek and other colleagues at the Fellowship for Interpretation of Genomes (Burr Ridge, IL) for their help with The SEED tools, genomic annotations, and subsystem curation. We also want to thank Cindy Cook for her help with the manuscript. The genomic analysis performed in this study was supported by the grant from the National Institute of Allergy and Infectious Diseases (NIAID) 'Genomics of Coenzyme Metabolism in Bacterial Pathogens' AI066244 to Andrei Osterman.

### Abbreviations

<b>I.</b>	Metabolites
<b>ACMS</b>	$\alpha$ -amino- $\beta$ -carboxymuconic semialdehyde
<b>ADPR</b>	ADP-ribose
<b>Asp</b>	L-aspartate
<b>DHAP</b>	dihydroxyacetone phosphate
<b>FKyn</b>	formylkynurenine
<b>Gln</b>	L-glutamine
<b>HAnt</b>	3-hydroxyanthranilate
<b>HKyn</b>	3-hydroxykynurenine
<b>IA</b>	iminoaspartate
<b>Kyn</b>	kynurenine
<b>Na</b>	nicotinic acid
<b>NaAD</b>	nicotinic acid adenine dinucleotide or deamido-NAD
<b>NADP</b>	nicotinamide adenine dinucleotide phosphate or NAD phosphate
<b>NaMN</b>	nicotinic acid mononucleotide
<b>Nm</b>	nicotinamide
<b>NMN</b>	nicotinamide mononucleotide
<b>PRPP</b>	phosphoribosyl pyrophosphate
<b>Qa</b>	quinolinic acid
<b>RNa</b>	ribosyl nicotinate
<b>RNm</b>	ribosyl nicotinamide
<b>Trp</b>	L-tryptophan

<b>II.</b>	Enzymes
<b>ADPRP</b>	ADP-ribose pyrophosphatase
<b>ASPDH</b>	L-aspartate dehydrogenase
<b>ASPOX</b>	L-aspartate oxidase
<b>GAT</b>	glutamine amidotransferase
<b>HADOX</b>	3-hydroxyanthranilate 3,4-dioxygenase
<b>KYNFA</b>	kynurenine formamidase
<b>KYNOX</b>	kynurenine 3-monooxygenase
<b>KYNSE</b>	kynureninase
<b>NADKIN</b>	NAD kinase
<b>NADSYN</b>	NAD synthetase
<b>NAMNAT</b>	NaMN adenylyltransferase
<b>NAPRT</b>	nicotinic acid phosphoribosyltransferase
<b>NMASE</b>	nicotinamide deamidase
<b>NMNAT</b>	NMN adenylyltransferase
<b>NMNSYN</b>	NMN synthetase
<b>NMPRT</b>	nicotinamide phosphoribosyltransferase
<b>QAPRT</b>	quinolinic acid phosphoribosyltransferase
<b>QASYN</b>	quinolinic acid synthetase
<b>RNAKIN</b>	ribosyl nicotinate kinase
<b>RNMKIN</b>	ribosyl nicotinamide kinase
<b>RNMPSE</b>	ribosyl nicotinamide phosphorylase
<b>TRDOX</b>	tryptophan 2,3-dioxygenase
<b>III.</b>	Regulators
<b>NadR</b>	NAD regulator
<b>NiaR</b>	niacin-responsive regulator
<b>NrtR</b>	nudix-related transcriptional regulator
<b>IV.</b>	Transporters
<b>NIAT</b>	niacin transporter
<b>RNMT</b>	ribosyl-nicotinamide transporter (PnuC family)

## References

1. P. Belenky; K. L. Bogan; C. Brenner, *Trends Biochem. Sci.* **2007**, *32*, 12–19.
2. G. Magni; A. Amici; M. Emanuelli; N. Raffaelli; S. Ruggieri, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1999**, *73*, 135–182.
3. A. G. Moat; J. W. Foster, Biosynthesis and Salvage Pathways of Pyridine Nucleotides. In *Pyridine Nucleotide Coenzymes – Part B*; P. R. Dolphin, O. Avramovich, Eds.; John Wiley & Sons: New York, 1987; pp 1–20.
4. T. Penfound; J. W. Foster, Biosynthesis and Recycling of NAD. In *Escherichia Coli and Salmonella, American Society for Microbiology*; F. Neihardt, Ed.; ASM: Washington, DC, 1996; pp 721–730.
5. J. S. Smith; C. B. Brachmann; I. Celic; M. A. Kenna; S. Muhammad; V. J. Starai; J. L. Avalos; J. C. Escalante-Semerena; C. Grubmeyer; C. Wolberger; J. D. Boeke, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6658–6663.
6. K. L. Bogan; C. Brenner, *Annu. Rev. Nutr.* **2008**, *28*, 115–130.
7. J. A. Khan; F. Forouhar; X. Tao; L. Tong, *Expert Opin. Ther. Targets* **2007**, *11*, 695–705.
8. G. Magni; A. Amici; M. Emanuelli; G. Orsomando; N. Raffaelli; S. Ruggieri, *Cell. Mol. Life Sci.* **2004**, *61*, 19–34.
9. V. J. Starai; H. Takahashi; J. D. Boeke; J. C. Escalante-Semerena, *Curr. Opin. Microbiol.* **2004**, *7*, 115–119.
10. T. P. Begley; C. Kinsland; R. A. Mehl; A. Osterman; P. Dorrestein, *Vitam. Horm.* **2001**, *61*, 103–119.
11. R. Overbeek; T. Begley; R. M. Butler; J. V. Choudhuri; H. Y. Chuang; M. Cohoon; V. de Crecy-Lagard; N. Diaz; T. Disz; R. Edwards; M. Fonstein; E. D. Frank; S. Gerdes; E. M. Glass; A. Goesmann; A. Hanson; D. Iwata-Reuyl; R. Jensen; N. Jamshidi; L. Krause; M. Kubal; N. Larsen; B. Linke; A. C. McHardy; F. Meyer; H. Neuweger; G. Olsen; R. Olson; A. Osterman; V. Portnoy; G. D. Pusch; D. A. Rodionov; C. Ruckert; J. Steiner; R. Stevens; I. Thiele; O. Vassieva; Y. Ye; O. Zagnitko; V. Vonstein, *Nucleic Acids Res.* **2005**, *33*, 5691–5702.
12. M. DeJongh; K. Formsma; P. Boillot; J. Gould; M. Rycenga; A. Best, *BMC Bioinform.* **2007**, *8*, 139.
13. Y. Ye; A. Osterman; R. Overbeek; A. Godzik, *Bioinformatics* **2005**, *21*, i1–i9.
14. D. A. Rodionov, *Chem. Rev.* **2007**, *107*, 3467–3497.

15. C. Brenner, *Structure* **2005**, *13*, 1239–1240.
16. W. Cheng; J. Roth, *J. Bacteriol.* **1995**, *177*, 6711–6717.
17. J. H. Grose; U. Bergthorsson; Y. Xu; J. Sterneckert; B. Khodaverdian; J. R. Roth, *J. Bacteriol.* **2005**, *187*, 4521–4530.
18. L. Aravind; V. Anantharaman; E. V. Koonin, *Proteins*, **2002**, *48*, 1–14.
19. L. Sorci; D. Martynowski; D. A. Rodionov; Y. Eyobo; X. Zogaj; K. E. Klose; E. V. Nikolaev; G. Magni; H. Zhang; A. L. Osterman, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3083–3088.
20. J. H. Grose; L. Joss; S. F. Velick; J. R. Roth, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 7601–7606.
21. N. Pollak; C. Dolle; M. Ziegler, *Biochem. J.* **2007**, *402*, 205–218.
22. R. Schwarcz, *Curr. Opin. Pharmacol.* **2004**, *4*, 12–17.
23. A. T. Brown; C. Wagner, *J. Bacteriol.* **1970**, *101*, 456–463.
24. O. Kurnasov; V. Goral; K. Colabroy; S. Gerdes; S. Anantha; A. Osterman; T. P. Begley, *Chem. Biol.* **2003**, *10*, 1195–1204.
25. S. Nasu; F. D. Wicks; R. K. Ghoson, *J. Biol. Chem.* **1982**, *257*, 626–632.
26. Z. Yang; A. Savchenko; A. Yakunin; R. Zhang; A. Edwards; C. Arrowsmith; L. Tong, *J. Biol. Chem.* **2003**, *278*, 8804–8808.
27. A. Andreeva; D. Howorth; J. M. Chandonia; S. E. Brenner; T. J. Hubbard; C. Chothia; A. G. Murzin, *Nucleic Acids Res.* **2008**, *36*, D419–D425.
28. B. T. Cookson; B. M. Olivera; J. R. Roth, *J. Bacteriol.* **1987**, *169*, 4285–4293.
29. K. T. Hughes; B. T. Cookson; D. Ladika; B. M. Olivera; J. R. Roth, *J. Bacteriol.* **1983**, *154*, 1126–1136.
30. R. Flachmann; N. Kunz; J. Seifert; M. Gutlich; F. J. Wientjes; A. Laufer; H. G. Gassen, *Eur. J. Biochem.* **1988**, *175*, 221–228.
31. C. A. DeVries; D. J. Hassett; J. L. Flynn; D. E. Ohman, *Gene* **1995**, *156*, 63–67.
32. D. Sun; P. Setlow, *J. Bacteriol.* **1993**, *175*, 1423–1432.
33. H. Sakuraba; T. Satomura; R. Kawakami; S. Yamamoto; Y. Kawarabayasi; H. Kikuchi; T. Ohshima, *Extremophiles* **2002**, *6*, 275–281.
34. A. Katoh; K. Uenohara; M. Akita; T. Hashimoto, *Plant Physiol.* **2006**, *141*, 851–857.
35. G. Tedeschi; A. Negri; F. Cecilian; A. Mattevi; S. Ronchi, *Eur. J. Biochem.* **1999**, *260*, 896–903.
36. M. Mortarino; A. Negri; G. Tedeschi; T. Simonc; S. Duga; H. G. Gassen; S. Ronchi, *Eur. J. Biochem.* **1996**, *239*, 418–426.
37. G. Tedeschi; A. Negri; M. Mortarino; F. Cecilian; T. Simonc; L. Faotto; S. Ronchi, *Eur. J. Biochem.* **1996**, *239*, 427–433.
38. K. R. Messner; J. A. Imlay, *J. Biol. Chem.* **2002**, *277*, 42563–42571.
39. A. Mattevi; G. Tedeschi; L. Bacchella; A. Coda; A. Negri; S. Ronchi, *Structure* **1999**, *7*, 745–756.
40. R. T. Bossi; A. Negri; G. Tedeschi; A. Mattevi, *Biochemistry* **2002**, *41*, 3018–3024.
41. H. Sakuraba; K. Yoneda; I. Asai; H. Tsuge; N. Katunuma; T. Ohshima, *Biochim. Biophys. Acta* **2008**, *1784*, 563–571.
42. K. Yoneda; R. Kawakami; Y. Tagashira; H. Sakuraba; S. Goda; T. Ohshima, *Biochim. Biophys. Acta* **2006**, *1764*, 1087–1093.
43. K. Yoneda; H. Sakuraba; H. Tsuge; N. Katunuma; T. Ohshima, *FEBS J.* **2007**, *274*, 4315–4325.
44. F. Cecilian; T. Caramori; S. Ronchi; G. Tedeschi; M. Mortarino; A. Galizzi, *Protein Expr. Purif.* **2000**, *18*, 64–70.
45. R. M. Cicchillo; L. Tu; J. A. Stromberg; L. M. Hoffart; C. Krebs; S. J. Booker, *J. Am. Chem. Soc.* **2005**, *127*, 7310–7311.
46. L. Loiseau; S. Ollagnier-de Choudens; D. Lascoux; E. Forest; M. Fontecave; F. Barras, *J. Biol. Chem.* **2005**, *280*, 26760–26769.
47. S. Ollagnier-de Choudens; L. Loiseau; Y. Sanakis; F. Barras; M. Fontecave, *FEBS Lett.* **2005**, *579*, 3737–3743.
48. P. Rossolillo; I. Marinoni; E. Galli; A. Colosimo; A. M. Albertini, *J. Bacteriol.* **2005**, *187*, 7155–7160.
49. S. Ollagnier-de-Choudens; Y. Sanakis; S. E. Abdel-Ghany; C. Rousset; H. Ye; M. Fontecave; E. A. Pilon-Smits; M. Pilon, *J. Biol. Chem.* **2007**, *282*, 18254–18264.
50. H. Sakuraba; H. Tsuge; K. Yoneda; N. Katunuma; T. Ohshima, *J. Biol. Chem.* **2005**, *280*, 26645–26648.
51. R. Pellicciari; L. Amori; G. Costantino; A. Giordani; A. Macchiarulo; L. Mattoli; P. Pevarello; C. Speciale; M. Varasi, *Adv. Exp. Med. Biol.* **2003**, *527*, 621–628.
52. T. W. Stone; C. M. Forrest; G. M. Mackay; N. Stoy; L. G. Darlington, *Metab. Brain Dis.* **2007**, *22*, 337–352.
53. Q. Han; B. T. Beerntsen; J. Li, *J. Insect Physiol.* **2007**, *53*, 254–263.
54. Y. F. Li; W. G. Bao, *FEMS Yeast Res.* **2007**, *7*, 657–664.
55. C. Panozzo; M. Nawara; C. Suski; R. Kucharczyka; M. Skoneczny; A. M. Becam; J. Rytka; C. J. Herbert, *FEBS Lett.* **2002**, *517*, 97–102.
56. A. Katoh; T. Hashimoto, *Front Biosci.* **2004**, *9*, 1577–1586.
57. O. Kurnasov; L. Jablonski; B. Polanuy; P. Dorrestein; T. Begley; A. Osterman, *FEMS Microbiol. Lett.* **2003**, *227*, 219–227.
58. A. M. Poole; D. Penny, *Bioessays* **2007**, *29*, 74–84.
59. K. L. Colabroy; T. P. Begley, *J. Bacteriol.* **2005**, *187*, 7866–7869.
60. S. Matthijs; C. Baysse; N. Koedam; K. A. Tehrani; L. Verheyden; H. Budzikiewicz; M. Schafer; B. Hoorelbeke; J. M. Meyer; H. De Greve; P. Cornelis, *Mol. Microbiol.* **2004**, *52*, 371–384.
61. T. Muraki; M. Taki; Y. Hasegawa; H. Iwaki; P. C. Lau, *Appl. Environ. Microbiol.* **2003**, *69*, 1564–1572.
62. G. Pandey; D. Paul; R. K. Jain, *FEMS Microbiol. Lett.* **2003**, *229*, 231–236.
63. Y. Kotake; I. Masayama, *Z. Physiol. Chem.* **1936**, *243*, 237–244.
64. K. Higuchi; O. Hayaishi, *Arch. Biochem. Biophys.* **1967**, *120*, 397–403.
65. F. Forouhar; J. L. Anderson; C. G. Mowat; S. M. Vorobiev; A. Hussain; M. Abashidze; C. Bruckmann; S. J. Thackray; J. Seetharaman; T. Tucker; R. Xiao; L. C. Ma; L. Zhao; T. B. Acton; G. T. Montelione; S. K. Chapman; L. Tong, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 473–478.
66. H. Sugimoto; S. Oda; T. Otsuki; T. Hino; T. Yoshida; Y. Shiro, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2611–2616.
67. Y. Zhang; S. A. Kang; T. Mukherjee; S. Bale; B. R. Crane; T. P. Begley; S. E. Ealick, *Biochemistry* **2007**, *46*, 145–155.
68. V. N. Dobrovolsky; J. F. Bowyer; M. K. Pabarcus; R. H. Heflich; L. D. Williams; D. R. Doerge; B. Arvidsson; J. Bergquist; J. E. Casida, *Biochim. Biophys. Acta* **2005**, *1724*, 163–172.
69. M. K. Pabarcus; J. E. Casida, *Biochim. Biophys. Acta* **2002**, *1596*, 201–211.
70. M. Wogulis; E. R. Chew; P. D. Donohoue; D. K. Wilson, *Biochemistry* **2008**, *47*, 1608–1621.
71. J. W. Arndt; R. Schwarzenbacher; R. Page; P. Abdubek; E. Ambing; T. Biorac; J. M. Canaves; H. J. Chiu; X. Dai; A. M. Deacon; M. DiDonato; M. A. Elsliger; A. Godzik; C. Grittini; S. K. Grzechnik; J. Hale; E. Hampton; G. W. Han; J. Haugen; M. Hornsby; H. E. Klock; E. Koesema; A. Kreusch; P. Kuhn; L. Jaroszewski; S. A. Lesley; I. Levin; D. McMullan; T. M. McPhillips; M. D. Miller;

- A. Morse; K. Moy; E. Nigoghossian; J. Ouyang; W. S. Peti; K. Quijano; R. Reyes; E. Sims; G. Spraggon; R. C. Stevens; H. van den Bedem; J. Velasquez; J. Vincent; F. von Delft; X. Wang; B. West; A. White; G. Wolf; Q. Xu; O. Zagnitko; K. O. Hodgson; J. Wooley; I. A. Wilson, *Proteins* **2005**, *58*, 755–758.
72. D. Alberati-Giani; A. M. Cesura; C. Broger; W. D. Warren; S. Rover; P. Malherbe, *FEBS Lett.* **1997**, *410*, 407–412.
  73. J. Breton; N. Avanzi; S. Magagnin; N. Covini; G. Magistrelli; L. Cozzi; A. Isacchi, *Eur. J. Biochem.* **2000**, *267*, 1092–1099.
  74. M. Hirai; M. Kiuchi; J. Wang; A. Ishii; H. Matsuoka, *Insect Mol. Biol.* **2002**, *11*, 497–504.
  75. K. R. Crozier; G. R. Moran, *Protein Expr. Purif.* **2007**, *51*, 324–333.
  76. S. V. Koushik; J. A. Moore, III; B. Sundararaju; R. S. Phillips, *Biochemistry* **1998**, *37*, 1376–1382.
  77. C. Momany; V. Levdkov; L. Blagova; S. Lima; R. S. Phillips, *Biochemistry* **2004**, *43*, 1193–1203.
  78. R. S. Phillips; B. Sundararaju; S. V. Koushik, *Biochemistry* **1998**, *37*, 8783–8789.
  79. S. Lima; R. Khristoforov; C. Momany; R. S. Phillips, *Biochemistry* **2007**, *46*, 2735–2744.
  80. V. Calderone; M. Trabucco; V. Menin; A. Negro; G. Zanotti, *Biochim. Biophys. Acta* **2002**, *1596*, 283–292.
  81. D. Nandi; E. S. Lightcap; Y. K. Koo; X. Lu; J. Quancard; R. B. Silverman, *Int. J. Biochem. Cell Biol.* **2003**, *35*, 1085–1097.
  82. X. Li; M. Guo; J. Fan; W. Tang; D. Wang; H. Ge; H. Rong; M. Teng; L. Niu; Q. Liu; Q. Hao, *Protein Sci.* **2006**, *15*, 761–773.
  83. Y. Zhang; K. L. Colabroy; T. P. Begley; S. E. Ealick, *Biochemistry* **2005**, *44*, 7632–7643.
  84. E. I. Solomon; T. C. Brunold; M. I. Davis; J. N. Kemsley; S. K. Lee; N. Lehnert; F. Neese; A. J. Skulan; Y. S. Yang; J. Zhou, *Chem. Rev.* **2000**, *100*, 235–350.
  85. K. L. Colabroy; H. Zhai; T. Li; Y. Ge; Y. Zhang; A. Liu; S. E. Ealick; F. W. McLafferty; T. P. Begley, *Biochemistry* **2005**, *44*, 7623–7631.
  86. K. L. Colabroy; T. P. Begley, *J. Am. Chem. Soc.* **2005**, *127*, 840–841.
  87. J. Preiss; P. Handler, *J. Biol. Chem.* **1958**, *233*, 493–500.
  88. B. Llorente; B. Dujon, *FEBS Lett.* **2000**, *475*, 237–241.
  89. J. J. Rowe; R. D. Lemmon; G. J. Tritz, *Microbios* **1985**, *44*, 169–184.
  90. D. A. Rodionov; X. Li; I. A. Rodionova; C. Yang; L. Sorci; E. Dervyn; D. Martynowski; H. Zhang; M. S. Gelfand; A. L. Osterman, *Nucleic Acids Res.* **2008**, *36*, 2032–2046.
  91. R. Janz; K. Hofmann; T. C. Sudhof, *J. Neurosci.* **1998**, *18*, 9269–9281.
  92. H. M. Said; S. M. Nabokina; K. Balamurugan; Z. M. Mohammed; C. Urbina; M. L. Kashyap, *Am. J. Physiol. Cell Physiol.* **2007**, *293*, C1773–C1778.
  93. P. Belenky; F. G. Racette; K. L. Bogan; J. M. McClure; J. S. Smith; C. Brenner, *Cell* **2007**, *129*, 473–484.
  94. P. Bieganowski; C. Brenner, *Cell* **2004**, *117*, 495–502.
  95. O. V. Kurnasov; B. M. Polanuyer; S. Ananta; R. Sloutsky; A. Tam; S. Y. Gerdes; A. L. Osterman, *J. Bacteriol.* **2002**, *184*, 6906–6917.
  96. W. Tempel; W. M. Rabeh; K. L. Bogan; P. Belenky; M. Wojcik; H. F. Seidle; L. Nedyalkova; T. Yang; A. A. Sauve; H. W. Park; C. Brenner, *PLoS Biol.* **2007**, *5*, e263.
  97. L. A. Maggio-Hall; J. C. Escalante-Semerena, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11798–11803.
  98. R. Bhatia; K. C. Calvo, *Arch. Biochem. Biophys.* **1996**, *325*, 270–278.
  99. K. T. Hughes; J. R. Roth; B. M. Olivera, *Genetics* **1991**, *127*, 657–670.
  100. P. M. Packman; W. B. Jakoby, *J. Biol. Chem.* **1965**, *240*, 4107–4108.
  101. K. Iwai; H. Taguchi, *J. Nutr. Sci. Vitaminol. (Tokyo)* **1973**, *19*, 491–499.
  102. S. J. Sinclair; K. J. Murphy; C. D. Birch; J. D. Hamill, *Plant Mol. Biol.* **2000**, *44*, 603–617.
  103. H. Taguchi; K. Iwai, *J. Nutr. Sci. Vitaminol. (Tokyo)* **1974**, *20*, 269–281.
  104. A. C. Foster; W. C. Zinkand; R. Schwarcz, *J. Neurochem.* **1985**, *44*, 446–454.
  105. K. Iwai; H. Taguchi, *Biochem. Biophys. Res. Commun.* **1974**, *56*, 884–891.
  106. C. Kohler; L. G. Eriksson; P. R. Flood; J. A. Hardie; E. Okuno; R. Schwarcz, *J. Neurosci.* **1988**, *8*, 975–987.
  107. E. Okuno; R. J. White; R. Schwarcz, *J. Biochem.* **1988**, *103*, 1054–1059.
  108. M. P. Heyes; K. Saito; J. S. Crowley; L. E. Davis; M. A. Demitrack; M. Der; L. A. Dilling; J. Elia; M. J. Kruesi; A. Lackner; S. A. Larsen; K. Lee; H. L. Leonard; S. P. Markey; A. Martin; S. Milstein; M. M. Mouradian; M. R. Pranzatelli; B. J. Quearry; A. Salazar; M. Smith; S. E. Strauss; T. S. Swedo; W. W. Thourtelotte, *Brain* **1992**, *115* (Part 5) 1249–1273.
  109. J. F. Reinhard, Jr.; J. B. Erickson; E. M. Flanagan, *Adv. Pharmacol.* **1994**, *30*, 85–127.
  110. K. T. Hughes; A. Dessen; J. P. Gray; C. Grubmeyer, *J. Bacteriol.* **1993**, *175*, 479–486.
  111. J. C. Eads; D. Ozturk; T. B. Wexler; C. Grubmeyer; J. C. Sacchettini, *Structure* **1997**, *5*, 47–58.
  112. E. di Luccio; D. K. Wilson, *Biochemistry* **2008**, *47*, 4039–4050.
  113. M. K. Kim; Y. J. Im; J. H. Lee; S. H. Eom, *Proteins* **2006**, *63*, 252–255.
  114. H. Liu; K. Woznica; G. Catton; A. Crawford; N. Botting; J. H. Naismith, *J. Mol. Biol.* **2007**, *373*, 755–763.
  115. R. Schwarzenbacher; L. Jaroszewski; F. von Delft; P. Abdubek; E. Ambing; B. Biorac; L. S. Brinen; J. M. Canaves; J. Cambell; H. J. Chiu; X. Dai; A. M. Deacon; M. DiDonato; M. A. Elsigier; S. Eshagi; R. Floyd; A. Godzik; C. Grittini; S. K. Grzechnik; E. Hampton; C. Karlak; H. E. Klock; E. Koesema; J. S. Kovarik; A. Kreuzsch; P. Kuhn; S. A. Lesley; I. Levin; D. McMullan; T. M. McPhillips; M. D. Miller; A. Morse; K. Moy; J. Ouyang; R. Page; K. Quijano; A. Robb; G. Spraggon; R. C. Stevens; H. van den Bedem; J. Velasquez; J. Vincent; X. Wang; B. West; G. Wolf; Q. Xu; K. O. Hodgson; J. Wooley; I. A. Wilson, *Proteins* **2004**, *55*, 768–771.
  116. V. Sharma; C. Grubmeyer; J. C. Sacchettini, *Structure* **1998**, *6*, 1587–1599.
  117. H. Cao; B. L. Pietrak; C. Grubmeyer, *Biochemistry* **2002**, *41*, 3520–3528.
  118. A. Vinitzky; H. Teng; C. T. Grubmeyer, *J. Bacteriol.* **1991**, *173*, 536–540.
  119. M. G. Wubbolts; P. Terpstra; J. B. van Beilen; J. Kingma; H. A. Meesters; B. Witholt, *J. Biol. Chem.* **1990**, *265*, 17665–17672.
  120. A. L. Osterman; T. P. Begley, *Prog. Drug Res.* **2007**, *64*, 131, 133–170.
  121. J. Gross; M. Rajavel; E. Segura; C. Grubmeyer, *Biochemistry* **1996**, *35*, 3917–3924.
  122. J. W. Gross; M. Rajavel; C. Grubmeyer, *Biochemistry* **1998**, *37*, 4189–4199.
  123. C. T. Grubmeyer; J. W. Gross; M. Rajavel, *Methods Enzymol.* **1999**, *308*, 28–48.
  124. A. Vinitzky; C. Grubmeyer, *J. Biol. Chem.* **1993**, *268*, 26004–26010.

125. J. S. Chappie; J. M. Canaves; G. W. Han; C. L. Rife; Q. Xu; R. C. Stevens, *Structure* **2005**, *13*, 1385–1396.
126. D. H. Shin; N. Oganessian; J. Jancarik; H. Yokota; R. Kim; S. H. Kim, *J. Biol. Chem.* **2005**, *280*, 18326–18335.
127. N. Hara; K. Yamada; T. Shibata; H. Osago; T. Hashimoto; M. Tsuchiya, *J. Biol. Chem.* **2007**, *282*, 24574–24582.
128. P. R. Martin; R. J. Shea; M. H. Mulks, *J. Bacteriol.* **2001**, *183*, 1168–1174.
129. S. Y. Gerdes; O. V. Kurnasov; K. Shatalin; B. Polanuyer; R. Sloutsky; V. Vonstein; R. Overbeek; A. L. Osterman, *J. Bacteriol.* **2006**, *188*, 3012–3023.
130. E. S. Miller; J. F. Heidelberg; J. A. Eisen; W. C. Nelson; A. S. Durkin; A. Ciecko; T. V. Feldblyum; O. White; I. T. Paulsen; W. C. Nierman; J. Lee; B. Szczypinski; C. M. Fraser, *J. Bacteriol.* **2003**, *185*, 5220–5233.
131. A. Rongvaux; R. J. Shea; M. H. Mulks; D. Gigot; J. Urbain; O. Leo; F. Andris, *Eur. J. Immunol.* **2002**, *32*, 3225–3234.
132. R. Depping; C. Lohaus; H. E. Meyer; W. Ruger, *Biochem. Biophys. Res. Commun.* **2005**, *335*, 1217–1223.
133. A. Fukuhara; M. Matsuda; M. Nishizawa; K. Segawa; M. Tanaka; K. Kishimoto; Y. Matsuki; M. Murakami; T. Ichisaka; H. Murakami; E. Watanabe; T. Takagi; M. Akiyoshi; T. Ohtsubo; S. Kihara; S. Yamashita; M. Makishima; T. Funahashi; S. Yamanaka; R. Hiramatsu; Y. Matsuzawa; I. Shimomura, *Science* **2005**, *307*, 426–430.
134. J. R. Revollo; A. A. Grimm; S. Imai, *J. Biol. Chem.* **2004**, *279*, 50754–50763.
135. J. R. Revollo; A. A. Grimm; S. Imai, *Curr. Opin. Gastroenterol.* **2007**, *23*, 164–170.
136. E. van der Veer; C. Ho; C. O’Neil; N. Barbosa; R. Scott; S. P. Cregan; J. G. Pickering, *J. Biol. Chem.* **2007**, *282*, 10841–10845.
137. M. Hasmann; I. Schemainda, *Cancer Res.* **2003**, *63*, 7436–7442.
138. J. A. Khan; X. Tao; L. Tong, *Nat. Struct. Mol. Biol.* **2006**, *13*, 582–588.
139. R. Takahashi; S. Nakamura; T. Yoshida; Y. Kobayashi; T. Ohkubo, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2007**, *63*, 375–377.
140. T. Wang; X. Zhang; P. Bheda; J. R. Revollo; S. Imai; C. Wolberger, *Nat. Struct. Mol. Biol.* **2006**, *13*, 661–662.
141. M. K. Kim; J. H. Lee; H. Kim; S. J. Park; S. H. Kim; G. B. Kang; Y. S. Lee; J. B. Kim; K. K. Kim; S. W. Suh; S. H. Eom, *J. Mol. Biol.* **2006**, *362*, 66–77.
142. J. W. Foster; D. M. Kinney; A. G. Moat, *J. Bacteriol.* **1979**, *137*, 1165–1175.
143. R. Frothingham; W. A. Meeker-O’Connell; E. A. Talbot; J. W. George; K. N. Kreuzer, *Antimicrob. Agents Chemother.* **1996**, *40*, 1426–1431.
144. A. Scorpio; Y. Zhang, *Nat. Med.* **1996**, *2*, 662–667.
145. P. Singh; A. K. Mishra; S. K. Malonia; D. S. Chauhan; V. D. Sharma; K. Venkatesan; V. M. Katoch, *J. Commun. Dis.* **2006**, *38*, 288–298.
146. X. Du; W. Wang; R. Kim; H. Yakota; H. Nguyen; S. H. Kim, *Biochemistry* **2001**, *40*, 14166–14172.
147. J. H. Grose; U. Bergthorsson; J. R. Roth, *J. Bacteriol.* **2005**, *187*, 2774–2782.
148. E. Sauer; M. Merdanovic; A. P. Mortimer; G. Bringmann; J. Reidl, *Antimicrob. Agents Chemother.* **2004**, *48*, 4532–4541.
149. S. K. Singh; O. V. Kurnasov; B. Chen; H. Robinson; N. V. Grishin; A. L. Osterman; H. Zhang, *J. Biol. Chem.* **2002**, *277*, 33291–33299.
150. K. W. Pankiewicz, *Pharmacol. Ther.* **1997**, *76*, 89–100.
151. L. Sorci; F. Cimadamore; S. Scotti; R. Petrelli; L. Cappellacci; P. Franchetti; G. Orsomando; G. Magni, *Biochemistry* **2007**, *46*, 4912–4922.
152. B. Ma; S. J. Pan; M. L. Zupancic; B. P. Cormack, *Mol. Microbiol.* **2007**, *66*, 14–25.
153. A. Matsui; H. Ashihara, *Plant Physiol. Biochem.* **2008**, *46*, 104–108.
154. J. A. Khan; S. Xiang; L. Tong, *Structure* **2007**, *15*, 1005–1013.
155. D. A. Rodionov; J. De Ingeniis; C. Mancini; F. Cimadamore; H. Zhang; A. L. Osterman; N. Raffaelli, *Nucleic Acids Res.* **2008**, *36*, 2047–2059.
156. S. Y. Gerdes; M. D. Scholle; M. D’Souza; A. Bernal; M. V. Baev; M. Farrell; O. V. Kurnasov; M. D. Daugherty; F. Mseeh; B. M. Polanuyer; J. W. Campbell; S. Anantha; K. Y. Shatalin; S. A. Chowdhury; M. Y. Fonstein; A. L. Osterman, *J. Bacteriol.* **2002**, *184*, 4555–4572.
157. I. Haferkamp; S. Schmitz-Esser; N. Linka; C. Urbany; A. Collingro; M. Wagner; M. Horn; H. E. Neuhaus, *Nature* **2004**, *432*, 622–625.
158. G. Magni; A. Amici; M. Emanuelli; G. Orsomando; N. Raffaelli; S. Ruggieri, *Curr. Med. Chem.* **2004**, *11*, 873–885.
159. K. T. Hughes; D. Ladika; J. R. Roth; B. M. Olivera, *J. Bacteriol.* **1983**, *155*, 213–221.
160. R. A. Mehl; C. Kinsland; T. P. Begley, *J. Bacteriol.* **2000**, *182*, 4372–4374.
161. S. Han; M. D. Forman; P. Loulakis; M. H. Rosner; Z. Xie; H. Wang; D. E. Danley; W. Yuan; J. Schafer; Z. Xu, *J. Mol. Biol.* **2006**, *360*, 814–825.
162. A. M. Olland; K. W. Underwood; R. M. Czerwinski; M. C. Lo; A. Aulabaugh; J. Bard; M. L. Stahl; W. S. Somers; F. X. Sullivan; R. Chopra, *J. Biol. Chem.* **2002**, *277*, 3698–3707.
163. H. J. Yoon; H. L. Kim; B. Mikami; S. W. Suh, *J. Mol. Biol.* **2005**, *351*, 258–265.
164. H. Zhang; T. Zhou; O. Kurnasov; S. Cheek; N. V. Grishin; A. Osterman, *Structure (Camb)* **2002**, *10*, 69–79.
165. A. Mushegian, *J. Mol. Microbiol. Biotechnol.* **1999**, *1*, 127–128.
166. N. Raffaelli; M. Emanuelli; F. M. Pisani; A. Amici; T. Lorenzi; S. Ruggieri; G. Magni, *Mol. Cell. Biochem.* **1999**, *193*, 99–102.
167. N. Raffaelli; T. Lorenzi; A. Amici; M. Emanuelli; S. Ruggieri; G. Magni, *FEBS Lett.* **1999**, *444*, 222–226.
168. V. Saridakis; D. Christendat; M. S. Kimber; A. Dharamsi; A. M. Edwards; E. F. Pai, *J. Biol. Chem.* **2001**, *276*, 7225–7232.
169. N. Huang; L. Sorci; X. Zhang; C. A. Brautigam; X. Li; N. Raffaelli; G. Magni; N. V. Grishin; A. L. Osterman; H. Zhang, *Structure* **2008**, *16*, 196–209.
170. M. Emanuelli; F. Carnevali; M. Lorenzi; N. Raffaelli; A. Amici; S. Ruggieri; G. Magni, *FEBS Lett.* **1999**, *455*, 13–17.
171. M. Emanuelli; A. Amici; F. Carnevali; F. Pierella; N. Raffaelli; G. Magni, *Protein Expr. Purif.* **2003**, *27*, 357–364.
172. S. N. Hashida; H. Takahashi; M. Kawai-Yamada; H. Uchimiyu, *Plant J.* **2007**, *49*, 694–703.
173. F. Berger; C. Lau; M. Dahlmann; M. Ziegler, *J. Biol. Chem.* **2005**, *280*, 36334–36341.
174. S. Boulton; S. Kyle; B. W. Durkacz, *Br. J. Cancer* **1997**, *76*, 845–851.
175. H. N. Jayaram; A. O’Connor; M. R. Grant; H. Yang; P. A. Grieco; D. A. Cooney, *J. Exp. Ther. Oncol.* **1996**, *1*, 278–285.
176. R. G. Zhai; Y. Cao; P. R. Hiesinger; Y. Zhou; S. Q. Mehta; K. L. Schulze; P. Verstreken; H. J. Bellen, *PLoS Biol.* **2006**, *4*, e416.
177. R. G. Zhai; F. Zhang; P. R. Hiesinger; Y. Cao; C. M. Haueter; H. J. Bellen, *Nature* **2008**, *452*, 887–891.



178. S. Garavaglia; I. D'Angelo; M. Emanuelli; F. Carnevali; F. Pierella; G. Magni; M. Rizzi, *J. Biol. Chem.* **2002**, *277*, 8524–8530.
179. E. Werner; M. Ziegler; F. Lerner; M. Schweiger; U. Heinemann, *FEBS Lett.* **2002**, *516*, 239–244.
180. X. Zhang; O. V. Kurnasov; S. Karthikeyan; N. V. Grishin; A. L. Osterman; H. Zhang, *J. Biol. Chem.* **2003**, *278*, 13503–13511.
181. T. Zhou; O. Kurnasov; D. R. Tomchick; D. D. Binns; N. V. Grishin; V. E. Marquez; A. L. Osterman; H. Zhang, *J. Biol. Chem.* **2002**, *277*, 13148–13154.
182. H. Zalkin, *Adv. Enzyme Regul.* **1983**, *21*, 225–237.
183. K. T. Hughes; B. M. Olivera; J. R. Roth, *J. Bacteriol.* **1988**, *170*, 2113–2120.
184. J. C. Willison; G. Tissot, *J. Bacteriol.* **1994**, *176*, 3400–3402.
185. C. Brenner, *Curr. Opin. Struct. Biol.* **2002**, *12*, 775–782.
186. H. C. Pace; C. Brenner, *Genome Biol.* **2001**, *2*, REVIEWS0001.
187. C. Nessi; A. M. Albertini; M. L. Speranza; A. Galizzi, *J. Biol. Chem.* **1995**, *270*, 6181–6185.
188. M. Bellinzoni; E. De Rossi; M. Branzoni; A. Milano; F. A. Peverali; M. Rizzi; G. Riccardi, *Protein Expr. Purif.* **2002**, *25*, 547–557.
189. F. Yamaguchi; S. Koga; I. Yoshioka; M. Takahashi; H. Sakuraba; T. Ohshima, *Biosci. Biotechnol. Biochem.* **2002**, *66*, 2052–2059.
190. H. M. McDonald; P. S. Pruet; C. Deivanayagam; Protasevich, II; W. M. Carson; L. J. DeLucas; W. J. Brouillette; C. G. Brouillette, *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2007**, *63*, 891–905.
191. G. B. Kang; Y. S. Kim; Y. J. Im; S. H. Rho; J. H. Lee; S. H. Eom, *Proteins* **2005**, *58*, 985–988.
192. B. L. Schneider; L. J. Reitzer, *J. Bacteriol.* **1998**, *180*, 4739–4741.
193. M. Rizzi; M. Bolognesi; A. Coda, *Structure* **1998**, *6*, 1129–1140.
194. J. Symersky; Y. Vedvedjiev; K. Moore; C. Brouillette; L. DeLucas, *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2002**, *58*, 1138–1146.
195. R. Jauch; A. Humm; R. Huber; M. C. Wahl, *J. Biol. Chem.* **2005**, *280*, 15131–15140.
196. P. Bieganski; H. C. Pace; C. Brenner, *J. Biol. Chem.* **2003**, *278*, 33049–33055.
197. Y. Suda; H. Tachikawa; A. Yokota; H. Nakanishi; N. Yamashita; Y. Miura; N. Takahashi, *Yeast* **2003**, *20*, 995–1005.
198. M. Wojcik; H. F. Seidle; P. Bieganski; C. Brenner, *J. Biol. Chem.* **2006**, *281*, 33395–33402.
199. N. Hara; K. Yamada; M. Terashima; H. Osago; M. Shimoyama; M. Tsuchiya, *J. Biol. Chem.* **2003**, *278*, 10914–10921.
200. P. Bieganski; C. Brenner, *J. Biol. Chem.* **2003**, *278*, 33056–33059.
201. M. Bellinzoni; S. Buroni; M. R. Pasca; P. Guglielame; F. Arcesi; E. De Rossi; G. Riccardi, *Res. Microbiol.* **2005**, *156*, 173–177.
202. S. Kawai; S. Mori; T. Mukai; S. Suzuki; T. Yamada; W. Hashimoto; K. Murata, *Biochem. Biophys. Res. Commun.* **2000**, *276*, 57–63.
203. S. Garavaglia; A. Galizzi; M. Rizzi, *J. Bacteriol.* **2003**, *185*, 4844–4850.
204. J. Liu; Y. Lou; H. Yokota; P. D. Adams; R. Kim; S. H. Kim, *J. Mol. Biol.* **2005**, *354*, 289–303.
205. G. Poncet-Montange; L. Assairi; S. Arold; S. Pochet; G. Labesse, *J. Biol. Chem.* **2007**, *282*, 33925–33934.
206. H. Sakuraba; R. Kawakami; T. Ohshima, *Appl. Environ. Microbiol.* **2005**, *71*, 4352–4358.
207. S. Kawai; S. Mori; T. Mukai; W. Hashimoto; K. Murata, *Eur. J. Biochem.* **2001**, *268*, 4359–4365.
208. A. Ochiai; S. Mori; S. Kawai; K. Murata, *Protein Expr. Purif.* **2004**, *36*, 124–130.
209. F. Lerner; M. Niere; A. Ludwig; M. Ziegler, *Biochem. Biophys. Res. Commun.* **2001**, *288*, 69–74.
210. F. Shi; S. Kawai; S. Mori; E. Kono; K. Murata, *FEBS J.* **2005**, *272*, 3337–3349.
211. W. L. Turner; J. C. Waller; B. Vanderbeld; W. A. Snedden, *Plant Physiol.* **2004**, *135*, 1243–1255.
212. S. Mori; M. Yamasaki; Y. Maruyama; K. Momma; S. Kawai; W. Hashimoto; B. Mikami; K. Murata, *Biochem. Biophys. Res. Commun.* **2005**, *327*, 500–508.
213. N. Raffaelli; L. Finaurini; F. Mazzola; L. Pucci; L. Sorci; A. Amici; G. Magni, *Biochemistry* **2004**, *43*, 7610–7617.
214. G. Labesse; D. Douguet; L. Assairi; A. M. Gilles, *Trends Biochem. Sci.* **2002**, *27*, 273–275.
215. N. Pollak; M. Niere; M. Ziegler, *J. Biol. Chem.* **2007**, *282*, 33562–33571.
216. G. Magni; G. Orsomando; N. Raffaelli, *Mini-Rev. Med. Chem.* **2006**, *6*, 739–746.
217. H. Bono; H. Ogata; S. Goto; M. Kanehisa, *Genome Res.* **1998**, *8*, 203–210.
218. E. Selkov; N. Maltsev; G. J. Olsen; R. Overbeek; W. B. Whitman, *Gene* **1997**, *197*, GC11–GC26.
219. R. Overbeek; D. Devine; V. Vonstein, *Targets* **2003**, *2*, 138–146.
220. A. Osterman; R. Overbeek, *Curr. Opin. Chem. Biol.* **2003**, *7*, 238–251.
221. H. I. Boshoff; X. Xu; K. Tahlan; C. S. Dowd; K. Pethe; L. R. Camacho; T. H. Park; C. S. Yun; D. Schnappinger; S. Ehrh; K. J. Williams; C. E. Barry, III, *J. Biol. Chem.* **2008**, *283*, 19329–19341.
222. G. Kemmer; T. J. Reilly; J. Schmidt-Brauns; G. W. Zlotnik; B. A. Green; M. J. Fiske; M. Herbert; A. Kraiss; S. Schlor; A. Smith; J. Reidl, *J. Bacteriol.* **2001**, *183*, 3974–3981.
223. G. Liu; J. Foster; P. Manlapaz-Ramos; B. M. Olivera, *J. Bacteriol.* **1982**, *152*, 1111–1116.
224. M. Merdanovic; E. Sauer; J. Reidl, *J. Bacteriol.* **2005**, *187*, 4410–4420.
225. A. V. Gerasimova; M. S. Gelfand, *J. Bioinform. Comput. Biol.* **2005**, *3*, 1007–1019.
226. N. Zhu; B. M. Olivera; J. R. Roth, *J. Bacteriol.* **1988**, *170*, 117–125.
227. J. W. Foster; Y. K. Park; T. Penfound; T. Fenger; M. P. Spector, *J. Bacteriol.* **1990**, *172*, 4187–4196.
228. J. W. Foster; T. Penfound, *FEMS Microbiol. Lett.* **1993**, *112*, 179–183.
229. T. Penfound; J. W. Foster, *J. Bacteriol.* **1999**, *181*, 648–655.
230. N. Zhu; J. R. Roth, *J. Bacteriol.* **1991**, *173*, 1302–1310.
231. D. Weekes; M. D. Miller; S. S. Krishna; D. McMullan; T. M. McPhillips; C. Acosta; J. M. Canaves; M. A. Elsliger; R. Floyd; S. K. Grzechnik; L. Jaroszewski; H. E. Klock; E. Koesema; J. S. Kovarik; A. Kreuzsch; A. T. Morse; K. Quijano; G. Spraggon; H. van den Bedem; G. Wolf; K. O. Hodgson; J. Wooley; A. M. Deacon; A. Godzik; S. A. Lesley; I. A. Wilson, *Proteins* **2007**, *67*, 247–252.
232. L. Aravind; V. Anantharaman; S. Balaji; M. M. Babu; L. M. Iyer, *FEMS Microbiol. Rev.* **2005**, *29*, 231–262.
233. K. Okuda; H. Hayashi; Y. Nishiyama, *J. Bacteriol.* **2005**, *187*, 4984–4991.
234. K. Okuda; Y. Nishiyama; E. H. Morita; H. Hayashi, *Biochim. Biophys. Acta* **2004**, *1699*, 245–252.
235. R. K. Aziz; D. Bartels; A. A. Best; M. DeJongh; T. Disz; R. A. Edwards; K. Formsma; S. Gerdes; E. M. Glass; M. Kubal; F. Meyer; G. J. Olsen; R. Olson; A. L. Osterman; R. A. Overbeek; L. K. McNeil; D. Paarmann; T. Pacciani; B. Parrello; G. D. Pusch; C. Reich; R. Stevens; O. Vassieva; V. Vonstein; A. Wilke; O. Zagnitko, *BMC Genomics* **2008**, *9*, 75.



### Biographical Sketches



Leonardo Sorci studied biology at the Polytechnic University of Marche in Ancona, Italy, where he was awarded his M.Sc. degree and specialized in Biotechnology, in 2000. His undergraduate research, conducted under the supervision of Professor Nadia Raffaelli, concentrated on enzymatic synthesis of novel pyridine dinucleotides. He earned his Ph.D. degree in 2004 at the Institute for Biochemical Biotechnologies where he also was a post-doctoral fellow under the supervision of Professor Giulio Magni. In 2005 he joined the group of Professor Andrei Osterman at the Burnham Institute for Medical Research (La Jolla, CA, USA) where he is currently a postdoctoral associate. His research interests focus on understanding metabolic networks by combining comparative genomics tools, molecular biology, and biochemistry. Recently he has uncovered a novel metabolic route leading to NAD biosynthesis and identified small molecule inhibitors targeting NAD biosynthesis in microbial pathogens.



Oleg Kurnasov graduated in 1990 from the Moscow University, Russia with an M.S. degree in molecular biology. He continued his graduate training at the Institute of Protein Research, Pushchino, Russia working on the development and applications of the cell-free translation system. After joining the research group of Andrei Osterman at Integrated Genomics in 1999, he began his long-term study on NAD biosynthetic enzymes and pathways in various species. At present he continues his research work on bioinformatics and experimental techniques at the Burnham Institute of Medical Research.



Dmitry A. Rodionov received his M.Sc. degree from the Moscow Engineering Physics Institute in 2000 and his Ph.D. degree in molecular biology from the State Research Institute for Genetics and Selection of Industrial Microorganisms, Moscow, Russia in 2003. He was a postdoctoral fellow in the group of Professor Mikhail S. Gelfand at the Institute for Information Transmission Problems of Russian Academy of Sciences (ITP RAS, Moscow, Russia), and in the group of Professor Andrei Osterman at the Burnham Institute for Medical Research (BIMR, La Jolla, CA, USA). He participated in multiple collaborative projects as a visitor scientist at the National Center for Biotechnology Information (Bethesda, MD, USA), Unit of Microbiology and Genetics (INSA-Lyon, France), Lawrence Berkeley National Laboratory (CA, USA), and Institute for Microbiology at Humboldt University (Berlin, Germany). He is currently a research assistant professor at BIMR, CA, USA. His research interests are different aspects of comparative genomic reconstruction of metabolic pathways and regulatory networks in prokaryotes including identification of transcription factor-binding sites and RNA regulatory elements, functional annotation of genes and identification of missing genes in metabolic pathways, phylogenetic analysis of protein families, and evolution of metabolic pathways and regulatory systems.



Andrei Osterman received his Ph.D. in biochemistry from the Moscow University, Russia. The main focus of his research, including postdoctoral studies at University of Texas Southwestern in Dallas, was in the field of mechanistic and structural enzymology. In 1999 he joined Integrated Genomics, a start-up biotech firm in Chicago, to lead an effort on genomics-driven discovery of metabolic enzymes and pathways. He joined the Burnham Institute for Medical Research in La Jolla, CA as an associate professor in the Bioinformatics and Systems Biology Program in 2003. A. Osterman is one of the founders of the Fellowship for Interpretation of Genomes (FIG), and he actively participates in the Project to Annotate 1000 Genomes launched by FIG in 2003. His research group combines bioinformatics and experimental techniques to elucidate the key metabolic subsystems, such as biogenesis of vitamins and cofactors, across several diverse species with completely sequenced genomes.

## 7.09 Pyridoxal Phosphate Biosynthesis

Jeremiah W. Hanes, Steven E. Ealick, and Tadhg P. Begley, Cornell University, Ithaca, NY, USA

Ivo Tews, Heidelberg University Biochemistry Center, Heidelberg, Germany

© 2010 Elsevier Ltd. All rights reserved.

7.09.1	Introduction	259
7.09.2	Enzymology of PLP Biosynthesis via the DXP-Dependent Route	260
7.09.3	Enzymology of PLP Biosynthesis via the R5P-Dependent Route	265
7.09.4	Transport, Salvage, and Interconversion of Various Forms of Vitamin B <sub>6</sub>	267
7.09.5	Conclusions	268
References		269

### 7.09.1 Introduction

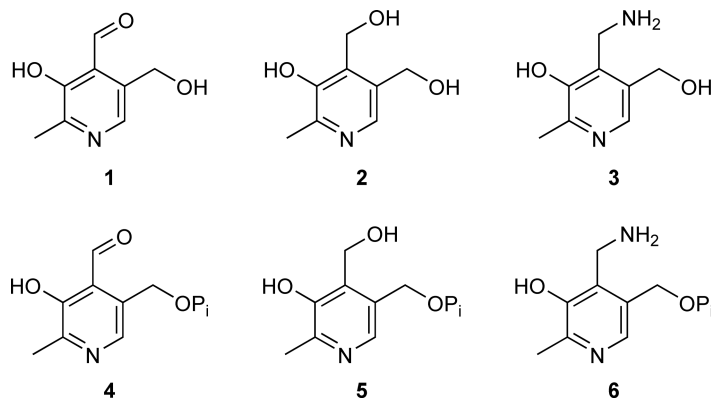
The term vitamin B<sub>6</sub> broadly refers to six compounds named pyridoxal (PL; **1**), pyridoxine (PN; **2**), pyridoxamine (PM; **3**) and their 5'-phosphorylated counterparts (PLP; **4**, PNP; **5**, and PMP; **6**, respectively (**Scheme 1**)). Pyridoxal 5'-phosphate (PLP) is typically recognized as the active form of vitamin B<sub>6</sub> and participates in a remarkable number of biochemical transformations mostly in amino acid metabolism, but also in some reactions in carbohydrate metabolism as well as in the modification of a variety of amine-containing molecules.<sup>1–3</sup>

In most cases, PLP acts as an enzyme-bound cofactor, with amino acid substrates, which aids in the catalysis of several classes of reactions including racemization, transamination, decarboxylation, and  $\alpha$ -,  $\beta$ -, and  $\gamma$ -elimination and replacement. The unifying role of PLP in these reactions is to act as an electron sink to stabilize carbanion formation at the  $\alpha$ -position of the amine-containing compound.<sup>1</sup> However, there are examples of reactions catalyzed by PLP where the role is fundamentally different. In the enzyme lysine 2,3-aminomutase (LAM), PLP catalyzes a radical-initiated isomerization reaction.<sup>4</sup> Another instance of a nontraditional role pertains to the glycogen phosphorylase family of enzymes. These enzymes use the phosphate group of PLP, instead of an external aldimine, for catalysis.<sup>5</sup> More recently, vitamin B<sub>6</sub> derivatives have been implicated as efficient antioxidants *in vivo*<sup>6,7</sup> and may act as singlet oxygen scavengers.<sup>8,9</sup>

Two *de novo* pathways for the biosynthesis of PLP are known (**Scheme 2**). The deoxyxylulose 5-phosphate (DXP)-dependent biosynthetic pathway present in *Escherichia coli* is fairly well understood from a mechanistic enzymology perspective, largely due to the successful *in vitro* reconstitution and structural characterization of the key enzymes involved.<sup>10,11</sup> Several reviews of this route are available, which provide a detailed account of some of the earlier work involved in identifying and characterizing the genes and the associated enzymatic activities.<sup>12–15</sup> Nevertheless, significant progress in recent years has been made in gaining a more in-depth understanding of the enzymes responsible for vitamin B<sub>6</sub> biosynthesis through the DXP-dependent pathway.

Overall, there are seven enzymes that employ the metabolites glyceraldehyde 3-phosphate (G3P; **7**), pyruvate **8**, and D-erythrose 4-phosphate **9** as precursors.<sup>10,16–22</sup> G3P and pyruvate are condensed by DXP synthase (DXPS) to give DXP **10**,<sup>23</sup> one of the substrates for pyridoxine 5'-phosphate (PNP) synthase (or pyridoxine auxotroph J (PdxJ)). The other substrate is 1-amino-3-hydroxyacetone phosphate **11** and is produced in a series of reactions catalyzed by erythrose 4-phosphate dehydrogenase (GapB), PdxB, PdxF, and PdxA starting with D-erythrose 4-phosphate **9**. PNP synthase is responsible for catalyzing the complex pyridine ring-forming reaction to give pyridoxine **5**,<sup>18,20,24–26</sup> which is then oxidized to PLP **4** by PdxH.<sup>27</sup> Interestingly, the occurrence of this system is largely limited to the  $\gamma$ -subdivision of proteobacteria.<sup>28</sup>

The discovery of a second pathway is owed to the observation that a gene initially termed singlet oxygen resistance 1 (SOR1) was involved in conferring singlet oxygen resistance in the fungi, *Cercospora nicotianae*<sup>7,29</sup> and *Aspergillus nidulans*.<sup>6</sup> *C. nicotianae*, a phytopathogenic fungus, produces perylenequinone toxins that are light-activated and lead to the generation of activated oxygen species. This organism is of interest due to its



**Scheme 1** Vitamin B<sub>6</sub> derivatives.

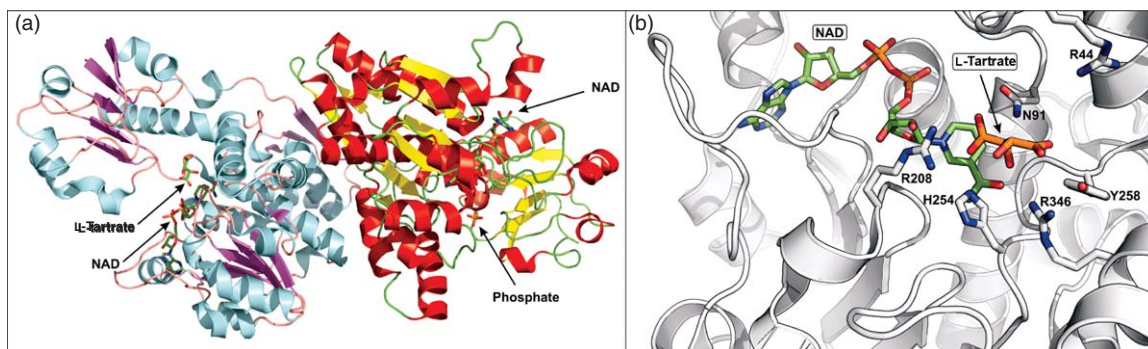
ability to avoid the toxicities associated with singlet oxygen.<sup>30</sup> It was shown that the highly conserved gene, SOR1, was essential for pyridoxine biosynthesis and occurs in organisms from widely diverse taxa in four kingdoms.<sup>31</sup> This discovery also suggested a previously unknown role for pyridoxine in singlet oxygen resistance, which was further supported by the observation that pyridoxine quenches singlet oxygen at a rate similar to that of vitamins C and E.<sup>7,8</sup> The SOR1 gene product was later renamed as Pdx1 (termed SNZ in *Saccharomyces cerevisiae*) and showed functional complementation to PdxJ using *E. coli*.<sup>32</sup> It followed that another gene, *pdx2* (SNO (SNZ proximal open reading frame) in *S. cerevisiae*), was associated with the *pdx1* gene and *C. nicotianae* lacking functional Pdx2 became auxotrophic for pyridoxine.<sup>33</sup> In *Bacillus subtilis*, the *pdx1* mutant was an auxotroph for vitamin B<sub>6</sub>; however, the *pdx2* mutant was a conditional auxotroph depending on the addition of ammonia suggesting that Pdx2 had glutaminase activity that could be substituted by the addition of ammonia to the growth medium.<sup>34</sup> The gene products copurified suggesting a physical interaction between the two subunits. The biosynthetic activities and three-dimensional structures of Pdx1 and Pdx2 were established and it was shown that the synthase subunit, Pdx1, works in conjunction with the glutaminase subunit, Pdx2, to catalyze PLP formation using D-ribose 5-phosphate **12**, glutamine **13**, and G3P **7** as substrates.<sup>34-43</sup> This pathway (termed R5P-dependent)<sup>7,28,32,34,44</sup> represents the most prevalent route and is present in archaea, fungi, plants, and the majority of bacteria.

## 7.09.2 Enzymology of PLP Biosynthesis via the DXP-Dependent Route

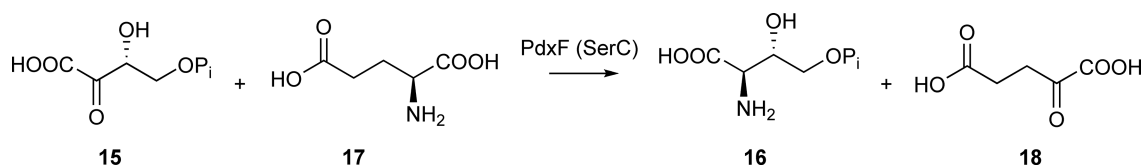
The biosynthesis of PLP in *E. coli* is composed of two main branches. In the first step toward obtaining the three-carbon substrate **11** for PNP synthase, GapB uses nicotinamide adenine dinucleotide (NAD) to catalyze the oxidation of D-erythrose 4-phosphate **9** to D-erythronate 4-phosphate **14**.<sup>45</sup> The enzyme from *E. coli* has been characterized *in vitro* and the reaction proceeds through a covalent thiohemiacetal intermediate bound at Cys149.<sup>46</sup> The enzyme is a tetramer with a subunit mass of ~37 kDa. A  $k_{\text{cat}}$  of ~20 s<sup>-1</sup>, a  $K_{\text{m}}$  of ~510 μM for D-erythrose 4-phosphate, and a  $K_{\text{m}}$  of ~800 μM for NAD were also determined. GapB shares over 40% sequence identity with glyceraldehyde 3-phosphate dehydrogenase. In the next step of the pathway, D-erythronate 4-phosphate dehydrogenase (PdxB) oxidizes **14** to (3R)-3-hydroxy-2-oxo-4-phosphonooxybutanoate **15** and also uses NAD as the electron acceptor (Scheme 2). The structure of PdxB from *Pseudomonas aeruginosa* was recently solved with NAD and tartrate (substrate analogue) bound at the active site (Figure 1). The two subunits of the homodimer were bound to different combinations of ligands and are also in dissimilar overall conformations.<sup>47</sup> The equilibrium for this reaction lies greatly on the side of **14** and the reaction is driven by the PdxF (also called SerC)-catalyzed transamination reaction, which efficiently converts **15** to 4-hydroxy-L-threonine phosphate (HTP) **16** (Scheme 3).<sup>48</sup> The enzyme from *E. coli* has been characterized *in vitro*, and has a  $k_{\text{cat}}$  of ~0.15 s<sup>-1</sup> and a  $K_{\text{m}}$  of ~110 μM for **15**. In many cases, PdxF is referred to as SerC because this enzyme also functions in the serine biosynthetic pathway. This PLP-dependent







**Figure 1** Crystal structure of *P. aeruginosa* PdxB (PDB code: 2o4c). (a) An illustration of the homodimer with ligands bound. The two subunits are in dissimilar overall conformations. One is bound to a substrate analogue, L-tartrate, and the other is bound to phosphate, while both are bound to NAD. (b) The active site showing residues within 6 Å of the substrate analogue L-tartrate.

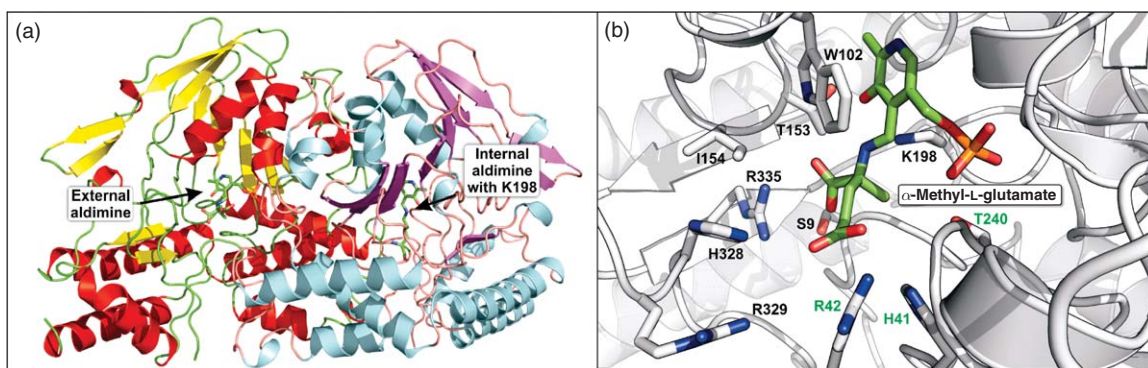


**Scheme 3** Reaction catalyzed by PdxF.

transamination uses the amino group of glutamate 17.<sup>49</sup> SerC catalyzes the conversion of 3-phosphohydroxypyruvate to O-phospho-L-serine and also uses glutamate as the nitrogen source. It is interesting that PLP biosynthesis uses a PLP-dependent enzyme.

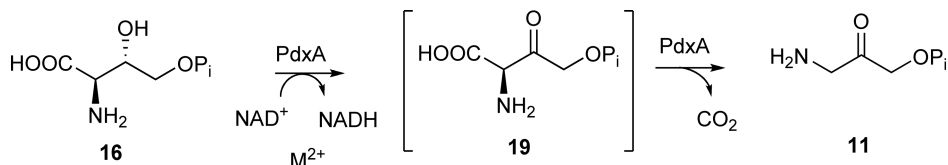
The structure of PdxF from *E. coli* has been determined to 2.3 Å resolution and is a homodimer that binds PLP through an aldimine with Lys198 (Figure 2).<sup>50</sup>

PdxA catalyzes the oxidative decarboxylation of HTP 16 to the unstable 1-amino-3-hydroxyacetone phosphate 11 (Scheme 4).<sup>16,19,51</sup> The recombinant enzyme from *E. coli* has a subunit molecular weight of 35.9 kDa and is a homodimer with a  $k_{cat}$  of  $\sim 1.4 s^{-1}$ . The  $K_m$  values have not been determined to our knowledge. This reaction likely occurs in a stepwise manner with the oxidation from the alcohol to ketone 19 occurring prior to the decarboxylation. PdxA can use either NAD or nicotinamide adenine dinucleotide phosphate



**Figure 2** Crystal structure of *E. coli* PdxF (SerC) (PDB code: 1bjo). (a) An illustration of the homodimer with bound ligands. Both subunits have PLP bound, one as an external aldimine with the analogue  $\alpha$ -methyl-L-glutamate, and the other as an internal aldimine with K198. (b) The active site containing the external aldimine identifying residues within 5 Å of  $\alpha$ -methyl-L-glutamate. Three residues (R42, H41, and T240) are from the other subunit and denoted by green text.

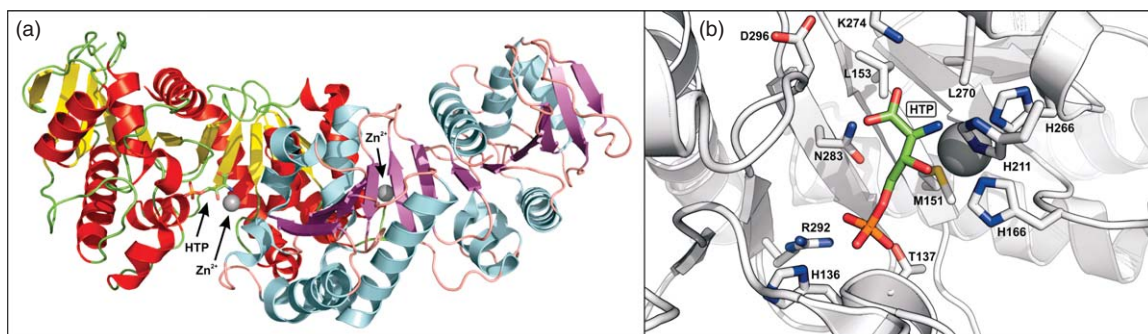




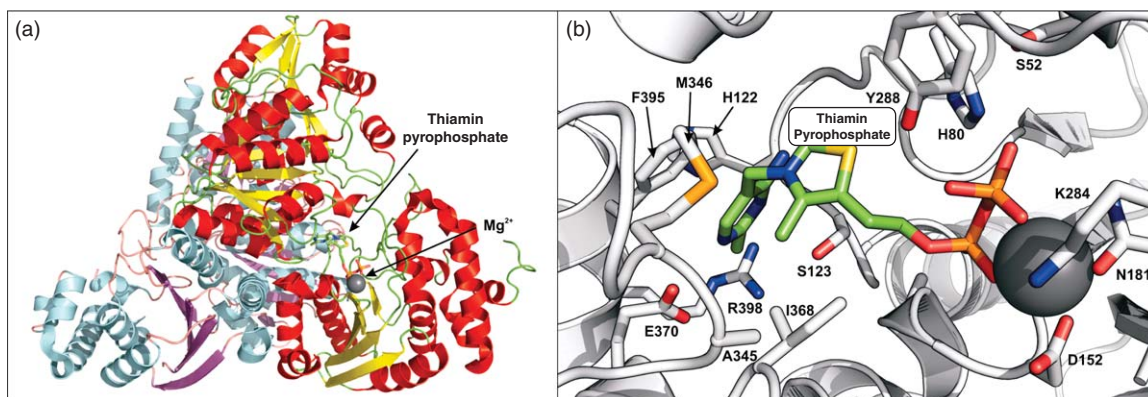
**Scheme 4** Reaction catalyzed by PdxA.

(NADP) as a redox cofactor and requires a divalent metal ion.<sup>10</sup> The addition of 1 mM ethylenediaminetetraacetic acid (EDTA) abolishes the oxidation of HTP and the full activity of the apoenzyme can be restored using 1 mM  $\text{Mn}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Mg}^{2+}$ , or  $\text{Ca}^{2+}$ . Half of the original activity is restored upon addition of 1 mM  $\text{Ni}^{2+}$  or 2 mM  $\text{Zn}^{2+}$ .<sup>10,51</sup> The structure of the enzyme from *E. coli* has been determined and is similar to isocitrate dehydrogenase and isopropylmalate dehydrogenase (Figure 3).<sup>19</sup>

The five-carbon backbone of PNP is derived from DXP, which is produced from G3P and pyruvate in a condensation reaction catalyzed by DXPS. The recombinant enzyme from *E. coli* has been characterized and shown to have a  $k_{\text{cat}}$  of  $\sim 270 \text{ s}^{-1}$  and a  $K_m$  of  $\sim 240 \mu\text{M}$  for G3P and  $\sim 96 \mu\text{M}$  for pyruvate. DXPS uses thiamin pyrophosphate as a cofactor, requires a divalent metal ion such as  $\text{Mg}^{2+}$  or  $\text{Mn}^{2+}$ , and operates as a homodimer with 67.6 kDa subunits. Extensive biochemical studies on this enzyme have been carried out in part because it is an attractive target for the development of antibiotics, antimalarials, and herbicides,<sup>52</sup> and also might be useful for the industrial production of ubiquinone-10.<sup>53</sup> Recently, the *E. coli* and *Deinococcus radiodurans* DXPS structures were published (Figure 4).<sup>52</sup> In addition to pyridoxine biosynthesis, DXPS is responsible for providing DXP for isoprenoid and thiamin biosynthesis.<sup>23</sup>



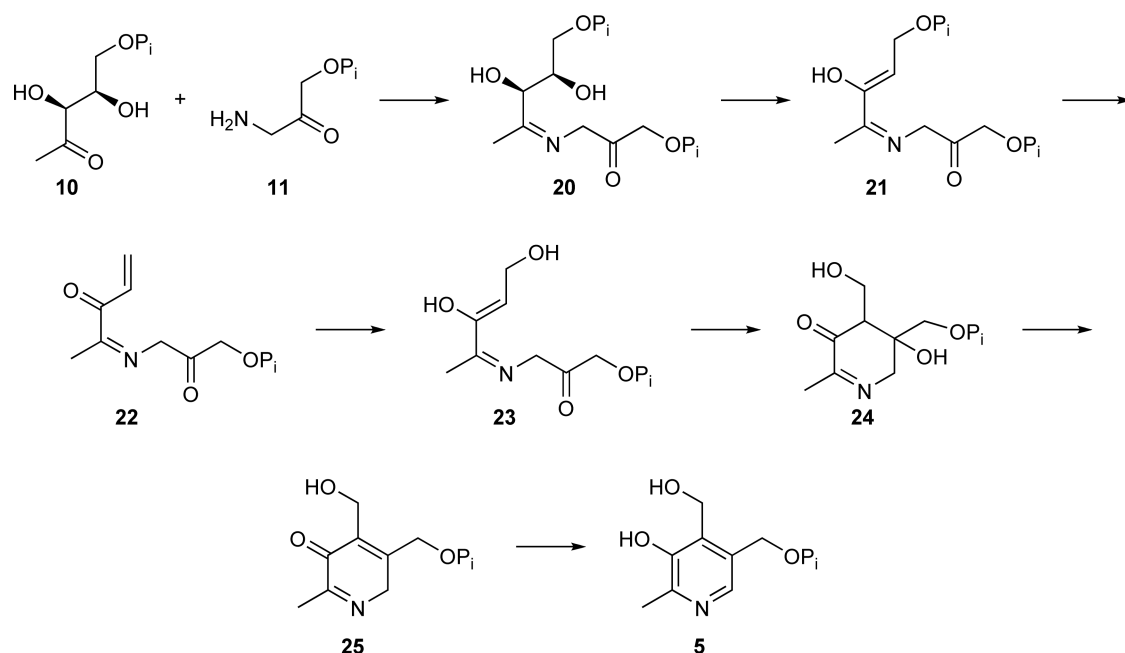
**Figure 3** Crystal structure of *E. coli* PdxA (PDB code: 1ps6). (a) An illustration of the homodimer with ligands bound. Both subunits have  $\text{Zn}^{2+}$  bound. The subunit in red and yellow is bound to HTP as depicted. (b) The active site containing HTP identifying residues within 4 Å of the ligand and  $\text{Zn}^{2+}$ .



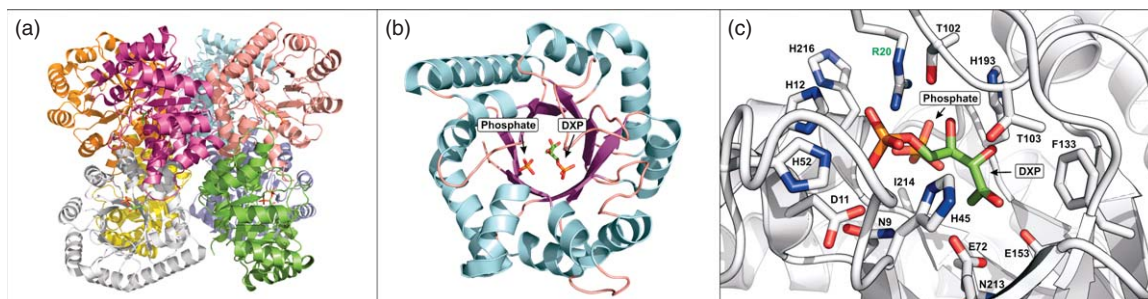
**Figure 4** Crystal structure of *E. coli* DXPS (PDB code: 2o1s). (a) An illustration of the homodimer with thiamin pyrophosphate and  $\text{Mg}^{2+}$  bound. The cofactor is visible in only one subunit. (b) The active site identifying residues within 4 Å of thiamin pyrophosphate and  $\text{Mg}^{2+}$ .

The two branches of the DXP-dependent pathway are bridged by PNP synthase (PdxJ). Due to the limitations imposed by the instability of 1-amino-3-hydroxyacetone phosphate **11**, detailed *in vitro* studies of PdxJ have been difficult. Nonetheless, a report of a  $K_m$  of  $\sim 27 \mu\text{M}$  and a  $k_{\text{cat}}$  of  $\sim 4 \text{ min}^{-1}$  for DXP was enabled by preparing 1-amino-3-hydroxyacetone phosphate *in situ* using PdxA and NAD.<sup>10</sup> A series of elegant  $^{18}\text{O}$ -labeling experiments were accomplished where  $[3,4-^{18}\text{O}_2]\text{DXP}$  was synthesized and converted into pyridoxine.<sup>17</sup> The C3 oxygen atom was retained whereas the C4 oxygen atom was lost during the reaction. Because an earlier mechanistic proposal suggested that the C4 oxygen atom would ultimately be bound to the phosphorus atom of phosphate, a  $^{31}\text{P}$  nuclear magnetic resonance (NMR) experiment was performed. The  $^{31}\text{P}$  NMR signal showed no upfield isotopic shift due to the incorporation of the C4  $^{18}\text{O}$  atom. Therefore, it was proposed that this oxygen atom is likely lost due to an elimination reaction. In addition to performing the reaction with labeled substrate, it was also run in buffer containing 30%  $\text{H}_2^{18}\text{O}$ . This showed not only that the phosphate is not hydrolyzed, but also that the C4' oxygen of pyridoxine is derived from the solvent. A mechanistic proposal consistent with these observations is shown in **Scheme 5**. In this proposal, DXP and 1-amino-3-hydroxyacetone phosphate react to form imine **20**. The elimination of water from C4 of DXP leads to **21**, which allows the elimination of phosphate to give **22**. Addition of water followed by an intramolecular aldol condensation results in ring closure to give **24**. The elimination of water and aromatization results in PNP **5**.

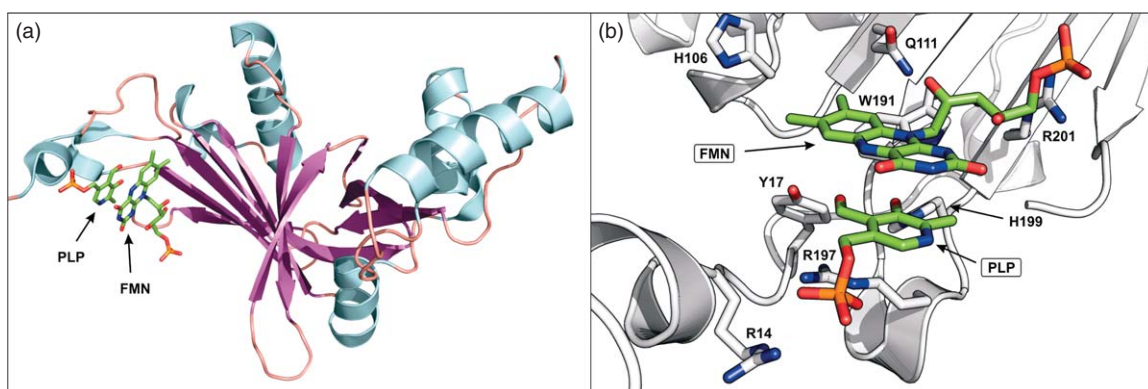
Additional mechanistic information concerning the PNP synthase reaction was obtained by a series of X-ray crystallography studies where various ligand-bound states of the enzyme have been reported.<sup>54</sup> In addition to the unliganded protein,<sup>26</sup> the structure of the enzyme has been solved with DXP,<sup>20</sup> phosphate and G3P,<sup>18</sup> and product bound at the active site.<sup>26</sup> PNP synthase is an octamer with monomers composed of a  $(\beta/\alpha)_8$  barrel motif. The ligand-binding sites are positioned between subunits with an arginine (Arg20) from another subunit interacting with a phosphate presumably bound in the 1-amino-3-hydroxyacetone phosphate site (**Figures 5(a)–5(c)**). The structure shown in **Figure 5** was particularly interesting because the monomers of the octamer were in three different states, revealing snapshots of the catalytic mechanism. An open conformation with no ligands bound, a partially open conformation with DXP bound, and a closed conformation with both DXP and phosphate bound were captured in a single experiment at a resolution of  $1.96 \text{ \AA}$ .<sup>20</sup> It was concluded that the phosphate was positioned in the binding pocket where the other substrate,



**Scheme 5** Proposed reaction mechanism catalyzed by PNP synthase (PdxJ).



**Figure 5** Crystal structure of *E. coli* PNP synthase with DXP and phosphate bound in the active site (PDB code: 1m5w). (a) An illustration of the octamer. Six of the subunits have DXP bound and two of those six also have phosphate bound. (b) Illustration of an individual monomer of the octamer with bound DXP and phosphate. (c) The active site identifying residues within 4 Å of DXP and phosphate. R20 interacts with the phosphate ion and is from another subunit (labeled with green text).



**Figure 6** Crystal structure of *E. coli* PdxH with PLP and FMN bound (PDB code: 1jnw). (a) An illustration of the monomer. (b) The active site identifying residues within 4 Å of PLP and FMN.

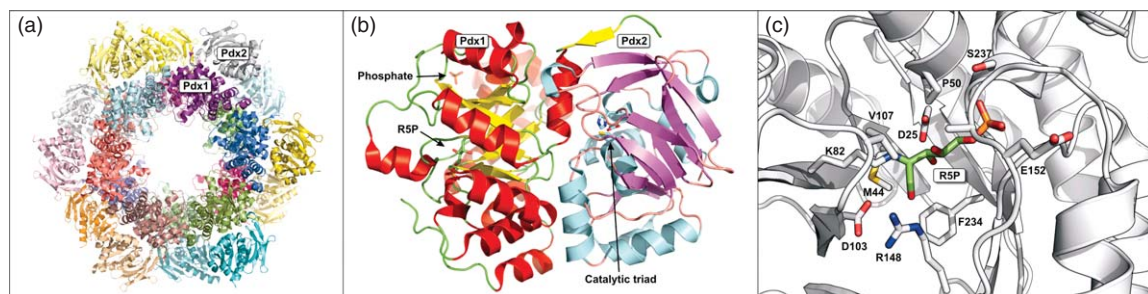
1-amino-3-hydroxyacetone phosphate, would bind during normal turnover. Structures of the enzyme are available in multiple conformational states that could explain an ordered access by substrates or water to the active site.

The final step in this pathway is the oxidation of PNP to PLP and is carried out by PdxH. The recombinant enzyme from *E. coli* has been studied *in vitro* and is a 51 kDa homodimer that utilizes flavin mononucleotide (FMN) as a cofactor. PdxH can use either PNP or pyridoxamine 5'-phosphate (PMP) as a substrate with a  $K_m$  of  $\sim 2$  and  $105 \mu\text{M}$  and  $k_{\text{cat}}$  of  $\sim 0.8$  and  $1.7 \text{ s}^{-1}$  for PNP and PMP, respectively.<sup>21</sup> The structures of the enzyme from *E. coli* as well as homologues from *Mycobacterium tuberculosis* and humans have been solved.<sup>55–58</sup> The *E. coli* enzyme with PLP and FMN bound is shown in **Figure 6**. PdxH is involved in both the biosynthetic and the salvage pathways and is further discussed in a section describing the transport, salvage, and interconversion of the various forms of vitamin B<sub>6</sub>.

### 7.09.3 Enzymology of PLP Biosynthesis via the R5P-Dependent Route

As stated above, the only other known route for the biosynthesis of PLP is dependent upon R5P, G3P, and glutamine, and is catalyzed by a single enzyme complex. The PLP synthase holoenzyme complex consists of 12 synthase (Pdx1) subunits and 12 glutaminase (Pdx2) subunits that form into a cogwheel-type structure (**Figure 7**).<sup>36,39,41,42,59</sup> Pdx2 subunits decorate the exterior of this cogwheel and catalyze the hydrolysis of glutamine using a Glu-His-Cys catalytic triad to give ammonia. Ammonia then diffuses through a hydrophobic channel to the active site of Pdx1 where the condensation reaction occurs. The Pdx1 monomers have a classic





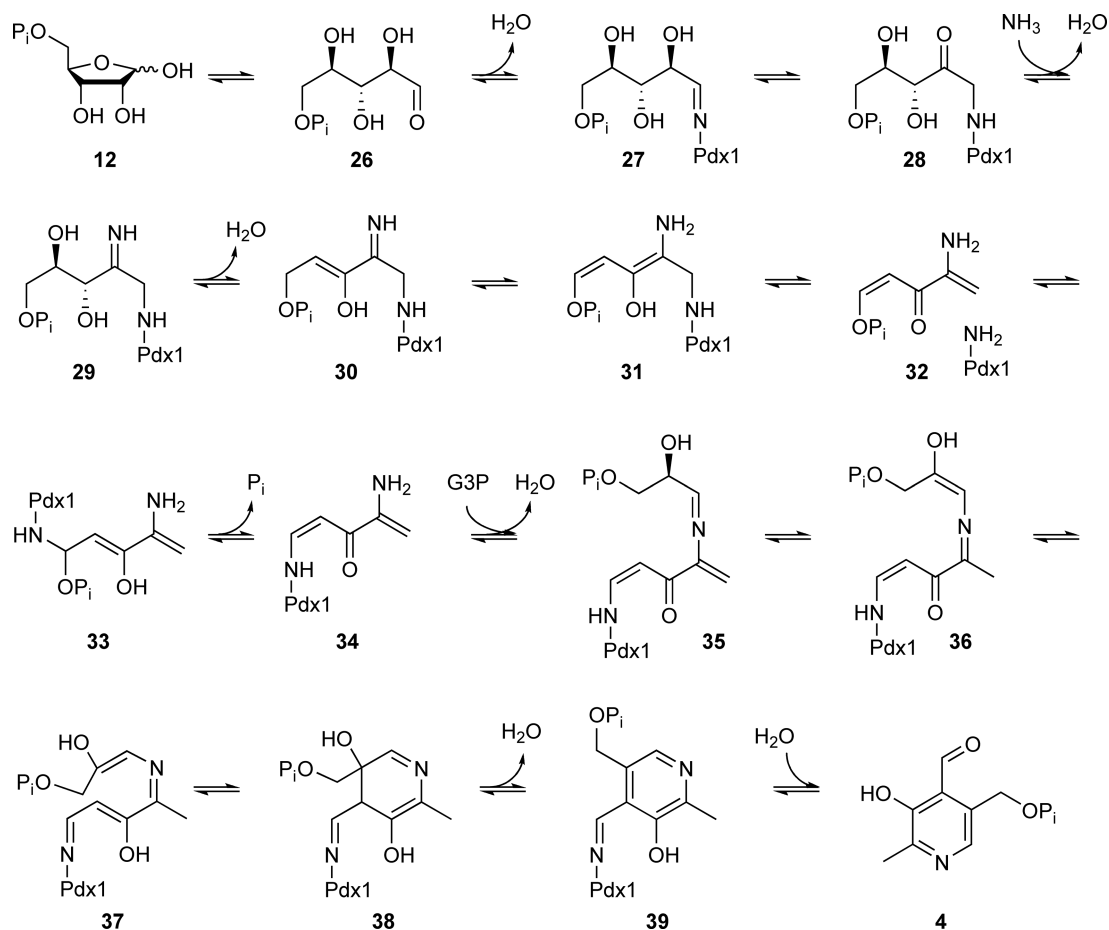
**Figure 7** Crystal structure of PLP synthase. (a) An illustration of the 24-subunit complex of Pdx1 and Pdx2 from *B. subtilis* (PDB code: 2nv2). The inner ring, where one subunit is denoted as Pdx1, forms a dodecamer. Six subunits are visible and six subunits form another symmetrical ring toward the rear. The outer ring is composed of 12 Pdx2 subunits, which are in contact with Pdx1. No Pdx2/Pdx2 interaction is observed in the structure. (b) Illustration of one Pdx1 subunit bound to one Pdx2 subunit for the enzyme from *Thermotoga maritima* (PDB code: 2iss). Shown are R5P and phosphate, which are bound in what is termed P1 and P2, respectively. The catalytic triad of Pdx2 is shown to illustrate where the hydrolysis of glutamine is likely to occur. (c) The active site of Pdx1 identifying residues within 4 Å of R5P. At the time this structure was published, it was thought that the R5P-derived ligand was bound through an imine between K82 (K81 in *B. subtilis*) and C2 of the open ring form of R5P. It is now thought that K82 is bound through C1 and a ketone is present at C2 in the R5P-bound form of the enzyme.<sup>63</sup>

( $\beta/\alpha$ )<sub>8</sub>-fold with the axis of each  $\beta$ -barrel nearly perpendicular to the sixfold axis. Pdx2 is a class I glutaminase and is characterized structurally as an  $\alpha/\beta$  three-layer sandwich. A great deal of research has gone into the identification of the thermodynamic driving force leading to Pdx1 and Pdx2 complex formation. These studies have revealed that complex formation is highly influenced by the binding of glutamine to Pdx2. The dissociation constant of the two subunits is reduced from  $\sim 6.9$  to  $0.3 \mu\text{M}$  upon addition of glutamine as the complex converts from the binary to the ternary Michaelis complex.<sup>35,60</sup> Another in-depth thermodynamic study was performed using the proteins from the malaria-causing human pathogen *Plasmodium falciparum*.<sup>60</sup> The original model of Pdx2 glutaminase activation by oxyanion hole formation was substantiated using a combination of structural analysis, site-directed mutagenesis, and biochemical and biophysical characterization.<sup>61</sup> The regulation of catalytic activity by oligomerization has recently been investigated.<sup>62</sup>

The reconstitution of the active Pdx1/Pdx2 complex, and the characterization of PLP as the reaction product revealed that the pentose and triose precursors to PLP were R5P and G3P, respectively. The enzyme complex also has the ability to use ribulose 5-phosphate as a substrate, although at a reduced maximum catalytic rate.<sup>40</sup>

The steady-state kinetics of the reaction were measured for the three substrates and a  $k_{\text{cat}}$  of  $\sim 7.6 \text{ min}^{-1}$  and a  $K_{\text{m}}$  of  $\sim 1 \mu\text{M}$  were determined for the hydrolysis of glutamine in the presence of both subunits.<sup>43</sup> The optimal ratio of Pdx1 to Pdx2 for the hydrolysis was approximately 1:1. A  $k_{\text{cat}}$  of  $\sim 0.02 \text{ min}^{-1}$  and a  $K_{\text{m}}$  of  $\sim 68 \mu\text{M}$  were estimated for R5P and a  $K_{\text{m}}$  of  $\sim 77 \mu\text{M}$  was estimated for G3P. A pull-down assay, taking the advantage of histidine-tagged Pdx2 and the glutaminase inhibitor acivicin, suggested that the interaction of Pdx1 and Pdx2 is transient. An interaction between the subunits occurred only prior to incubation with the inhibitor and was eliminated upon addition of acivicin.<sup>43</sup> In another study, a chromophoric reaction intermediate was observed that accumulated in the active site of Pdx1 upon addition of only glutamine and R5P. The intermediate was converted to PLP upon addition of G3P, was shown to have a mass of 95 Da, and was covalently attached to Lys81.<sup>64</sup> An earlier reaction intermediate was also detected that could be observed bound to the same lysine residue and had a mass of 212 Da corresponding to the mass of R5P minus one water.<sup>64</sup> The mechanism of PLP synthase has been the topic of several recent studies and is beginning to emerge. A proposal that is consistent with all available data is shown in **Scheme 6**.

Ring opening of R5P to aldehyde **26** allows for imine formation with an active site lysine residue and isomerization to give **28**. The hydrolysis of glutamine provides ammonia, which then reacts with **28** to form imine **29**. Dehydration is followed by tautomerization. Elimination of the lysine amine from C1 then gives **32**. A conjugate addition of the same lysine at C5 and the elimination of phosphate, results in the chromophoric intermediate **34** that has been characterized extensively. Addition of G3P gives **35**, which then undergoes two tautomerization reactions and an electrocyclic ring closure resulting in **38**. Aromatization



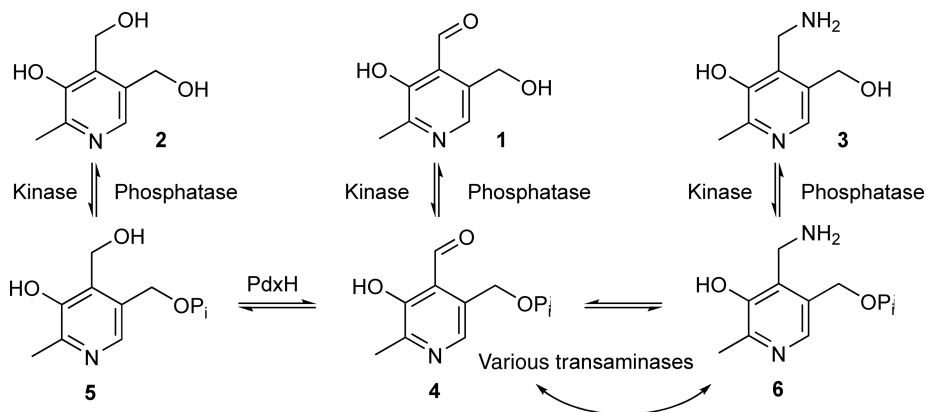
**Scheme 6** Mechanistic proposal for PLP biosynthesis via the R5P-dependent route.

and imine hydrolysis complete the formation of PLP. The intriguing C1 to C5 lysine migration (**31–33**) is supported by the NMR characterization of three intermediates along the reaction pathway.<sup>63</sup> Additionally, the chromophoric intermediate **34** was further characterized biochemically and was trapped using the common biological reducing agent tris(2-carboxyethyl)phosphine (TCEP).<sup>65,66</sup>

Even though there is now an extensive body of research supporting this mechanism, the proposed mechanism will undoubtedly evolve as more information is uncovered regarding this complex and intriguing pathway.

### 7.09.4 Transport, Salvage, and Interconversion of Various Forms of Vitamin B<sub>6</sub>

Each of the six PLP vitamers can be interconverted in a variety of organisms as shown in **Scheme 7**. It has generally been thought that, for bacteria, the salvaged forms of vitamin B<sub>6</sub> are the vitamers lacking 5'-phosphate groups. This notion is accurate for the vast majority of physiologically significant small molecules, as transport is usually carried out only after the hydrolytic action of nonspecific phosphatases secreted to the extracellular environment. These metabolites are then rephosphorylated intracellularly by specific kinases. Surprisingly, little is known regarding the capability of *E. coli* to transport the various forms of vitamin B<sub>6</sub>. However, the bacterium has been shown to harbor genes encoding two kinase isoforms (termed PL kinase 1 and PL kinase 2) proficient at performing the Mg<sup>2+</sup>/ATP-dependent phosphorylation of PN, PM, and PL.<sup>67–69</sup> Once phosphorylated, PNP and PMP are readily converted to PLP in most organisms by PdxH, an FMN-dependent oxygen-requiring enzyme, which also operates in the *de novo* biosynthetic pathway.<sup>21,27,70</sup>



**Scheme 7** Interconversion of vitamin B<sub>6</sub> vitamers.

This enzyme does not catalyze the analogous oxidation on vitamers lacking the 5'-phosphate. Various transaminases involved in amino acid metabolism catalyze the interconversion of PLP and PMP. At present there is no known pathway for the catabolism of vitamin B<sub>6</sub> in *E. coli*. However, the catabolic pathway in *Mesorhizobium loti* MAFF303099 has been well characterized.<sup>71-79</sup>

### 7.09.5 Conclusions

While PLP biosynthesis and salvage are now mature research areas, many questions still remain unanswered: Many details of the enzymology of the biosynthetic pathways still remain to be resolved and it is not yet clear if all of the biosynthetic or catabolic pathways have been identified. In addition, the PLP transport system in bacteria has not yet been identified and PLP, while covalently attached to its cognate enzyme, has not yet been exploited in activity-based proteomics.

#### Glossary

**apoenzyme** The polypeptide portion of an enzyme that requires a cofactor for full function.

**auxotrophic** The inability of an organism to synthesize a nutrient that is required for its growth.

**external aldimine** A form of pyridoxal 5'-phosphate that is involved in an imine linkage with a small molecule ligand.

**glutaminase** An enzyme that catalyzes the hydrolysis of glutamine to yield glutamate and ammonia.

**internal aldimine** An enzyme-bound form of pyridoxal 5'-phosphate that is involved in an imine linkage with a protein lysine residue.

**phytopathogenic** An organism that is pathogenic to a plant.

**proteobacteria** Microorganisms such as *Escherichia coli* and *Salmonella typhimurium* of the phylum proteobacteria.

#### Abbreviations

**DXP** deoxyxylulose 5-phosphate

**DXPS** DXP synthase



<b>FMN</b>	flavin mononucleotide
<b>G3P</b>	glyceraldehyde 3-phosphate
<b>HTP</b>	4-hydroxy-L-threonine phosphate
<b>LAM</b>	lysine 2,3-aminomutase
<b>PL</b>	pyridoxal
<b>PLP</b>	pyridoxal 5'-phosphate
<b>PM</b>	pyridoxamine
<b>PMP</b>	pyridoxamine 5'-phosphate
<b>PN</b>	pyridoxine
<b>PNP</b>	pyridoxine 5'-phosphate
<b>SNZ</b>	snooze
<b>SOR1</b>	singlet oxygen resistance 1
<b>TCEP</b>	tris(2-carboxyethyl)phosphine

## References

1. A. C. Eliot; J. F. Kirsch, *Annu. Rev. Biochem.* **2004**, *73*, 383–415.
2. H. Hayashi, *J. Biochem.* **1995**, *118*, 463–473.
3. R. A. John, *Biochim. Biophys. Acta* **1995**, *1248*, 81–96.
4. P. A. Frey, *Annu. Rev. Biochem.* **2001**, *70*, 121–148.
5. N. B. Livanova; N. A. Chebotareva; T. B. Eronina; B. I. Kurganov, *Biochemistry (Mosc.)* **2002**, *67*, 1089–1098.
6. A. H. Osmani; G. S. May; S. A. Osmani, *J. Biol. Chem.* **1999**, *274*, 23565–23569.
7. M. Ehrenshaft; P. Bilski; M. Y. Li; C. F. Chignell; M. E. Daub, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9374–9378.
8. P. Bilski; M. Y. Li; M. Ehrenshaft; M. E. Daub; C. F. Chignell, *Photochem. Photobiol.* **2000**, *71*, 129–134.
9. S. K. Jain; G. Lim, *Free Radic. Biol. Med.* **2001**, *30*, 232–237.
10. D. E. Cane; S. Du; J. K. Robinson; Y. Hsiung; I. D. Spenser, *J. Am. Chem. Soc.* **1999**, *121*, 7722–7723.
11. B. Laber; W. Maurer; S. Scharf; K. Stepusin; F. S. Schmidt, *FEBS Lett.* **1999**, *449*, 45–48.
12. R. E. Hill; I. D. Spenser, Biosynthesis of Vitamin B<sub>6</sub>. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*; F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, H. E. Umbarger, Eds.; ASM Press: Washington, DC, 1996; pp 695–703.
13. E. E. Snell; B. E. Haskell, Metabolism of Water-Soluble Vitamins. In *Comprehensive Biochemistry*; M. Florkin, E. H. Stoltz, Eds.; Elsevier: New York, 1971; pp 47–71.
14. W. B. Dempsey, Synthesis of Pyridoxal Phosphate. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*; F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, H. E. Umbarger, Eds.; ASM Press: Washington, DC, 1987; pp 539–543.
15. R. E. Hill; I. D. Spenser, The Biosynthesis of Vitamin B<sub>6</sub>. In *Pyridoxal Phosphate: Chemical, Biochemical and Medical Aspects*; D. Dolphin, R. Poulson, O. Avramovic, Eds.; John Wiley & Sons, Inc.: New York, 1986; pp 417–476.
16. J. Banks; D. E. Cane, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1633–1636.
17. D. E. Cane; S. Du; I. D. Spenser, *J. Am. Chem. Soc.* **2000**, *122*, 4213–4214.
18. M. Garrido-Franco; B. Laber; R. Huber; T. Clausen, *J. Mol. Biol.* **2002**, *321*, 601–612.
19. J. Sivaraman; Y. Li; J. Banks; D. E. Cane; A. Matte; M. Cygler, *J. Biol. Chem.* **2003**, *278*, 43682–43690.
20. J. I. Yeh; S. Du; E. Pohl; D. E. Cane, *Biochemistry* **2002**, *41*, 11649–11657.
21. G. Zhao; M. E. Winkler, *J. Bacteriol.* **1995**, *177*, 883–891.
22. T. B. Fitzpatrick; N. Amrhein; B. Kappes; P. Macheroux; I. Tews; T. Raschle, *Biochem. J.* **2007**, *407*, 1–13.
23. G. A. Sprenger; U. Schorken; T. Wiegert; S. Grolle; A. A. deGraaf; S. V. Taylor; T. P. Begley; S. BringerMeyer; H. Sahm, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12857–12862.
24. M. Tazoe; K. Ichikawa; T. Hoshino, *Biosci. Biotechnol. Biochem.* **2002**, *66*, 934–936.
25. M. G. Franco; R. Huber; F. S. Schmidt; B. Laber; T. Clausen, *Acta Crystallogr. D: Biol. Crystallogr.* **2000**, *56*, 1045–1048.
26. M. G. Franco; B. Laber; R. Huber; T. Clausen, *Structure* **2001**, *9*, 245–253.
27. M. Di Salvo; Y. Yang; G. Zhao; M. E. Winkler; V. Schirch, *Protein Expr. Purif.* **1998**, *13*, 349–356.
28. G. Mittenhuber, *J. Mol. Microbiol. Biotechnol.* **2001**, *3*, 1–20.
29. M. Ehrenshaft; K. R. Chung; A. E. Jennis; M. E. Daub, *Curr. Genet.* **1999**, *34*, 478–485.
30. M. E. Daub; M. Ehrenshaft, *Annu. Rev. Phytopathol.* **2000**, *38*, 461–490.
31. M. Ehrenshaft; A. E. Jennis; K. R. Chung; M. E. Daub, *Mol. Cell* **1998**, *1*, 603–609.
32. D. K. Wetzel; M. Ehrenshaft; S. A. Denslow; M. E. Daub, *FEBS Lett.* **2004**, *564*, 143–146.
33. M. Ehrenshaft; M. E. Daub, *J. Bacteriol.* **2001**, *183*, 3383–3390.
34. B. R. Belitsky, *J. Bacteriol.* **2004**, *186*, 1191–1196.
35. M. Neuwirth; K. Flicker; M. Strohmeier; I. Tews; P. Macheroux, *Biochemistry* **2007**, *46*, 5131–5139.

36. M. Strohmeier; T. Raschle; J. Mazurkiewicz; K. Rippe; I. Sinning; T. B. Fitzpatrick; I. Tews, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 19284–19289.
37. M. Tambasco-Studart; I. Tews; N. Amrhein; T. B. Fitzpatrick, *Plant Physiol.* **2007**, *144*, 915–925.
38. C. Wrenger; M. L. Eschbach; I. B. Muller; D. Warnecke; R. D. Walter; *J. Biol. Chem.* **2005**, *280*, 5242–5248.
39. F. Zein; Y. Zhang; Y. N. Kang; K. Burns; T. P. Begley; S. E. Ealick, *Biochemistry* **2006**, *45*, 14609–14620.
40. K. E. Burns; Y. Xiang; C. L. Kinsland; F. W. McLafferty; T. P. Begley, *J. Am. Chem. Soc.* **2005**, *127*, 3682–3683.
41. J. A. Bauer; E. M. Bennett; T. P. Begley; S. E. Ealick, *J. Biol. Chem.* **2004**, *279*, 2704–2711.
42. M. Gengenbacher; T. B. Fitzpatrick; T. Raschle; K. Flicker; I. Sinning; S. Muller; P. Macheroux; I. Tews; B. Kappes, *J. Biol. Chem.* **2006**, *281*, 3633–3641.
43. T. Raschle; N. Amrhein; T. B. Fitzpatrick, *J. Biol. Chem.* **2005**, *280*, 32291–32300.
44. A. Sakai; M. Kita; T. Katsuragi; N. Ogasawara; Y. Tani, *J. Biosci. Bioeng.* **2002**, *93*, 309–312.
45. G. Zhao; A. J. Pease; N. Bharani; M. E. Winkler, *J. Bacteriol.* **1995**, *177*, 2804–2812.
46. S. Boschi-Muller; S. Azza; D. Pollastro; C. Corbier; G. Branlant, *J. Biol. Chem.* **1997**, *272*, 15106–15112.
47. J. Y. Ha; J. H. Lee; K. H. Kim; d.J. Kim; H. H. Lee; H. K. Kim; H. J. Yoon; S. W. Suh, *J. Mol. Biol.* **2007**, *366*, 1294–1304.
48. G. Zhao; M. E. Winkler, *J. Bacteriol.* **1996**, *178*, 232–239.
49. C. Drewke; M. Klein; D. Clade; A. Arenz; R. Muller; E. Leistner, *FEBS Lett.* **1996**, *390*, 179–182.
50. G. Hester; W. Stark; M. Moser; J. Kallen; Z. Markovic-Housley; J. N. Jansonius, *J. Mol. Biol.* **1999**, *286*, 829–850.
51. D. E. Cane; Y. J. Hsiung; J. A. Cornish; J. K. Robinson; I. D. Spenser, *J. Am. Chem. Soc.* **1998**, *120*, 1936–1937.
52. S. Xiang; G. Usunow; G. Lange; M. Busch; L. Tong, *J. Biol. Chem.* **2007**, *282*, 2676–2682.
53. J. K. Lee; D. K. Oh; S. Y. Kim, *J. Biotechnol.* **2007**, *128*, 555–566.
54. M. Garrido-Franco, *Biochim. Biophys. Acta* **2003**, *1647*, 92–97.
55. B. K. Biswal; M. M. Cherney; M. Wang; C. Garen; M. N. James, *Acta Crystallogr. D: Biol. Crystallogr.* **2005**, *61*, 1492–1499.
56. B. K. Biswal; K. Au; M. M. Cherney; C. Garen; M. N. James, *Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* **2006**, *62*, 735–742.
57. M. L. di Salvo; M. K. Safo; F. N. Musayev; F. Bossa; V. Schirch, *Biochim. Biophys. Acta* **2003**, *1647*, 76–82.
58. F. N. Musayev; M. L. di Salvo; T. P. Ko; V. Schirch; M. K. Safo, *Protein Sci.* **2003**, *12*, 1455–1463.
59. J. Zhu; J. W. Burgner; E. Harms; B. R. Belitsky; J. L. Smith, *J. Biol. Chem.* **2005**, *280*, 27914–27923.
60. K. Flicker; M. Neuwirth; M. Strohmeier; B. Kappes; I. Tews; P. Macheroux, *J. Mol. Biol.* **2007**, *374*, 732–748.
61. S. Wallner; M. Neuwirth; K. Flicker; I. Tews; P. Macheroux, *Biochemistry* **2009**, *48*, 1928–1935.
62. T. Raschle; D. Speziga; W. Kress; C. Moccand; P. Gehrig; N. Amrhein; E. Weber-Ban; T. B. Fitzpatrick, *J. Biol. Chem.* **2009**, *284*, 7706–7718.
63. J. W. Hanes; I. Keresztes; T. P. Begley, *Nat. Chem. Biol.* **2008**, *4*, 425–430.
64. T. Raschle; D. Arigoni; R. Brunisholz; H. Rechsteiner; N. Amrhein; T. B. Fitzpatrick, *J. Biol. Chem.* **2007**, *282*, 6098–6105.
65. J. W. Hanes; K. E. Burns; D. G. Hilmey; A. Chatterjee; P. C. Dorrestein; T. P. Begley, *J. Am. Chem. Soc.* **2008**, *130*, 3043–3052.
66. J. W. Hanes; I. Keresztes; T. P. Begley, *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 2102–2105.
67. J. H. Park; K. Burns; C. Kinsland; T. P. Begley, *J. Bacteriol.* **2004**, *186*, 1571–1573.
68. Y. Yang; H. C. Tsui; T. K. Man; M. E. Winkler, *J. Bacteriol.* **1998**, *180*, 1814–1821.
69. M. L. di Salvo; S. Hunt; V. Schirch, *Protein Expr. Purif.* **2004**, *36*, 300–306.
70. C. Notheis; C. Drewke; E. Leistner, *Biochim. Biophys. Acta* **1995**, *1247*, 265–271.
71. T. Mukherjee; K. M. McCulloch; S. E. Ealick; T. P. Begley, *Biochemistry* **2007**, *46*, 13606–13615.
72. T. Mukherjee; C. Kinsland; T. P. Begley, *Bioorg. Chem.* **2007**, *35*, 458–464.
73. T. Mukherjee; D. G. Hilmey; T. P. Begley, *Biochemistry* **2008**, *47*, 6233–6241.
74. F. Ge; N. Yokochi; Y. Yoshikane; K. Ohnishi; T. Yagi, *J. Biochem.* **2008**, *143*, 603–609.
75. Y. Yoshikane; N. Yokochi; M. Yamasaki; K. Mizutani; K. Ohnishi; B. Mikami; H. Hayashi; T. Yagi, *J. Biol. Chem.* **2008**, *283*, 1120–1127.
76. B. Yuan; N. Yokochi; Y. Yoshikane; K. Ohnishi; T. Yagi, *J. Biosci. Bioeng.* **2006**, *102*, 504–510.
77. N. Yokochi; S. Nishimura; Y. Yoshikane; K. Ohnishi; T. Yagi, *Arch. Biochem. Biophys.* **2006**, *452*, 1–8.
78. J. Funami; Y. Yoshikane; H. Kobayashi; N. Yokochi; B. Yuan; K. Iwasaki; K. Ohnishi; T. Yagi, *Biochim. Biophys. Acta* **2005**, *1753*, 234–239.
79. B. Yuan; Y. Yoshikane; N. Yokochi; K. Ohnishi; T. Yagi, *FEMS Microbiol. Lett.* **2004**, *234*, 225–230.

### Biographical Sketches



Jeremiah W. Hanes earned his B.Sc. degree from the University of Nebraska–Lincoln in 2000 and then his Ph.D. degree from the University of Texas, Austin, in 2004 under Kenneth Johnson. He then carried out postdoctoral studies at Cornell University with Tadhg Begley. Much of his research has been in the field of mechanistic enzymology, specifically motor proteins and enzymes involved in cofactor biosynthesis. Jeremiah is currently a senior scientist in the enzymology department at Pacific Biosciences, a start-up company developing third-generation DNA sequencing technology.



Steven E. Ealick is the William T. Miller Professor of Chemistry and Chemical Biology at Cornell University. He received his B.S. degree in chemistry from Oklahoma State University in 1972 and his Ph.D. degree in physical chemistry from the University of Oklahoma in 1976. He is also the director of the Northeast Collaborative Access Team at the Advanced Photon Source, Argonne National Laboratory. His research focuses on protein structure and function using X-ray crystallography, including enzyme mechanism, structure-based drug design, and protein evolution. He has published widely in the areas of thiamin biosynthesis, purine biosynthesis, polyamine biosynthesis, and cytokine structure and function. He is also interested in the development of instrumentation for synchrotron beam lines used for protein crystallography.



Tadhg P. Begley obtained his B.Sc. degree from The National University of Ireland in Cork in 1977 and his Ph.D. degree from the California Institute of Technology under P. Dervan in 1982. He carried out postdoctoral studies at the University of Geneva with W. Oppolzer and at MIT with C. Walsh. After 23 years in the Cornell Chemistry Department, he recently moved to Texas A&M University where he is the Derek H. R. Barton Professor of Chemistry. Begley's research is focused on the mechanistic enzymology of complex organic transformations, particularly those found on the vitamin biosynthetic pathways.



Ivo Tews obtained his Ph.D. degree at EMBL Hamburg in 1996 and is currently a research assistant at the Heidelberg University Biochemistry Center. He has worked on the biochemistry and crystallographic 3D structure determination of proteins involved in vitamin B<sub>6</sub> biosynthesis from bacteria, yeast, and *Plasmodium falciparum*, the causative agent of malaria. The Tews laboratory has also carried out research on cellular signaling by adenylyl cyclases in the tuberculosis pathogen *Mycobacterium tuberculosis* and regulated protein transport across biomembranes.

## 7.10 Pyridoxal 5'-Phosphate-Dependent Enzymes: Catalysis, Conformation, and Genomics

**Samanta Raboni, Francesca Spyrakis, Barbara Campanini, Alessio Amadasi, Stefano Bettati, Alessio Peracchi, and Andrea Mozzarelli**, University of Parma, Parma, Italy; Present address: University of Padua, Padua, Italy

**Roberto Contestabile**, University of Rome 'La Sapienza,' Rome, Italy

© 2010 Elsevier Ltd. All rights reserved.

<b>7.10.1</b>	<b>Introduction</b>	274
7.10.1.1	From Vitamin B <sub>6</sub> to Pyridoxal 5'-Phosphate	274
7.10.1.2	Overview of PLP Catalysis	275
7.10.1.2.1	Dunathan's hypothesis	278
7.10.1.3	Families of PLP-Enzymes Based on Evolutionary Relationships and Fold Types of PLP-Enzymes Based on Three-Dimensional Structures	279
7.10.1.4	List of Previous Reviews on PLP-Enzymology	280
<b>7.10.2</b>	<b>Reactivity of Amino Acids at <math>\alpha</math>-Position</b>	280
7.10.2.1	Amino Acid Racemization	280
7.10.2.1.1	Alanine racemase	280
7.10.2.2	Decarboxylation	282
7.10.2.2.1	Ornithine decarboxylase	284
7.10.2.2.2	DOPA decarboxylase	285
7.10.2.2.3	Dialkylglycine decarboxylase and diaminopimelate decarboxylase	286
7.10.2.3	$\alpha$ -Elimination and $\alpha$ -Replacement Reactions	286
7.10.2.3.1	Serine hydroxymethyltransferase	287
7.10.2.3.2	Threonine aldolase	289
7.10.2.3.3	$\alpha$ -Oxoamine synthases	290
7.10.2.3.4	1-Aminocyclopropane-1-carboxylate deaminase	290
7.10.2.4	Transamination Reaction	291
7.10.2.4.1	Aminotransferases	291
<b>7.10.3</b>	<b>Reactivity of Amino Acids at <math>\beta</math>-Position</b>	295
7.10.3.1	$\beta$ -Elimination Reactions	295
7.10.3.1.1	Tyrosine phenol-lyase	295
7.10.3.1.2	Tryptophan indole-lyase	298
7.10.3.1.3	Cystalysin	298
7.10.3.1.4	Cysteine desulfurases	299
7.10.3.1.5	Cystathionine $\beta$ -lyase	300
7.10.3.1.6	MaY	301
7.10.3.2	$\beta$ -Replacement Reactions	301
7.10.3.2.1	Tryptophan synthase	302
7.10.3.2.2	O-acetylserine sulfhydrylase	304
7.10.3.2.3	Cystathionine $\beta$ -synthase	305
<b>7.10.4</b>	<b>Reactivity of Amino Acids at <math>\gamma</math>-Position</b>	306
7.10.4.1	$\gamma$ -Elimination Reactions	306
7.10.4.1.1	Cystathionine $\gamma$ -lyase	306
7.10.4.1.2	Methionine $\gamma$ -lyase	308
7.10.4.2	$\gamma$ -Replacement Reactions	308
7.10.4.2.1	Threonine synthase	308
7.10.4.2.2	Cystathionine $\gamma$ -synthase	309



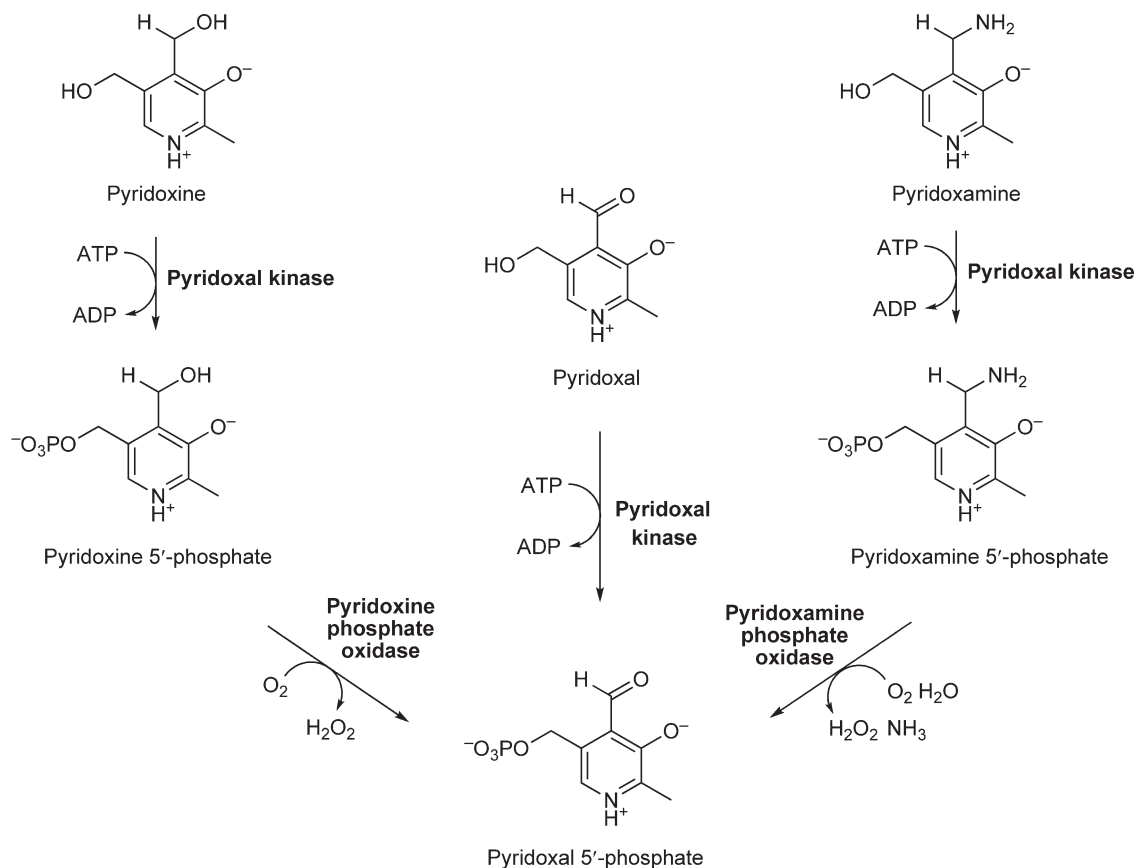
<b>7.10.5</b>	<b>Radical Reactions in PLP-Enzymes</b>	312
7.10.5.1	Aminomutases	312
7.10.5.1.1	Lysine 2,3-aminomutase	312
7.10.5.1.2	CDP-6-deoxy-D-glycero-L-threo-4-hexulose-3-dehydrase	315
<b>7.10.6</b>	<b>Reactivity with Inhibitors</b>	315
7.10.6.1	Mechanism-Based Inhibitors	317
7.10.6.1.1	Activated electrophiles	317
7.10.6.1.2	Activated nucleophiles	318
7.10.6.1.3	Aromatization mechanism-based inhibitors	318
7.10.6.2	Alternative Strategies	325
7.10.6.2.1	Noncovalent inhibitors	325
7.10.6.2.2	Coenzyme–substrate analogues	326
7.10.6.2.3	Allosteric inhibitors	327
<b>7.10.7</b>	<b>Role of Conformational Changes in Substrate and Reaction Specificity and Allosteric Regulation</b>	327
7.10.7.1	Aspartate Aminotransferase and the Open/Closed Transition in PLP-Dependent Enzymes	327
7.10.7.2	Tryptophan Synthase	328
7.10.7.3	O-Acetylserine Sulfhydrylase	329
<b>7.10.8</b>	<b>Genomics, Evolution, Structure, and Reactivity/Specificity</b>	329
7.10.8.1	How Old is PLP?	329
7.10.8.2	How Did PLP-Dependent Enzymes Emerge and Evolve?	330
7.10.8.3	A Case Study: Fold-Type I Enzymes	332
7.10.8.4	Exploring the Modern PLP-Dependent Enzymes: Can Sequence be Used to Infer Function?	333
<b>References</b>		337

## 7.10.1 Introduction

Pyridoxal 5'-phosphate (PLP)-dependent enzymes represent about 4% of the enzymes classified by the Enzyme Commission. The versatility of PLP in carrying out a large variety of reactions exploiting the electron sink effect of the pyridine ring, the conformational changes accompanying the chemical steps and stabilizing distinct catalytic intermediates, and the spectral properties of the different coenzyme–substrate derivatives signaling the reaction progress are some of the key features of PLP and PLP-dependent catalytic action. Progress in PLP-dependent enzymes functional genomics and development of inhibitors with therapeutic activity highlight the continuous novelty of an 'old' class of enzymes.

### 7.10.1.1 From Vitamin B<sub>6</sub> to Pyridoxal 5'-Phosphate

Pyridoxine, pyridoxamine, and pyridoxal are the three B<sub>6</sub> vitamers found in plants and bacteria (**Figure 1**). The vitamers are characterized by an alcohol, an amine, and an aldehyde moiety at position 4' of the pyridine ring, respectively. They are phosphorylated at position C5' by a kinase and oxidized at position C4' by the action of a specific oxidase to form PLP, the biological active form of vitamin B<sub>6</sub> (**Figure 1**). Vitamin B<sub>6</sub> is synthesized in microorganisms and plants through two independent pathways,<sup>1</sup> one starting from D-erythrose 4-phosphate and deoxyxylulose 5-phosphate and the second starting from ribose 5-phosphate, glutamine, and glyceraldehyde-3-phosphate. Vitamin B<sub>6</sub> and its derivatives are absorbed in humans from food and, upon dephosphorylation by a specific phosphatase and phosphorylation by a kinase, are made available as PLP to the enzymes, whose activity depends on this coenzyme (**Figure 2**).<sup>2</sup> In some enzymes, vitamin B<sub>6</sub> is bound as pyridoxamine 5'-phosphate (PMP) (**Figure 1**), which is transformed to PLP during the catalytic cycle.

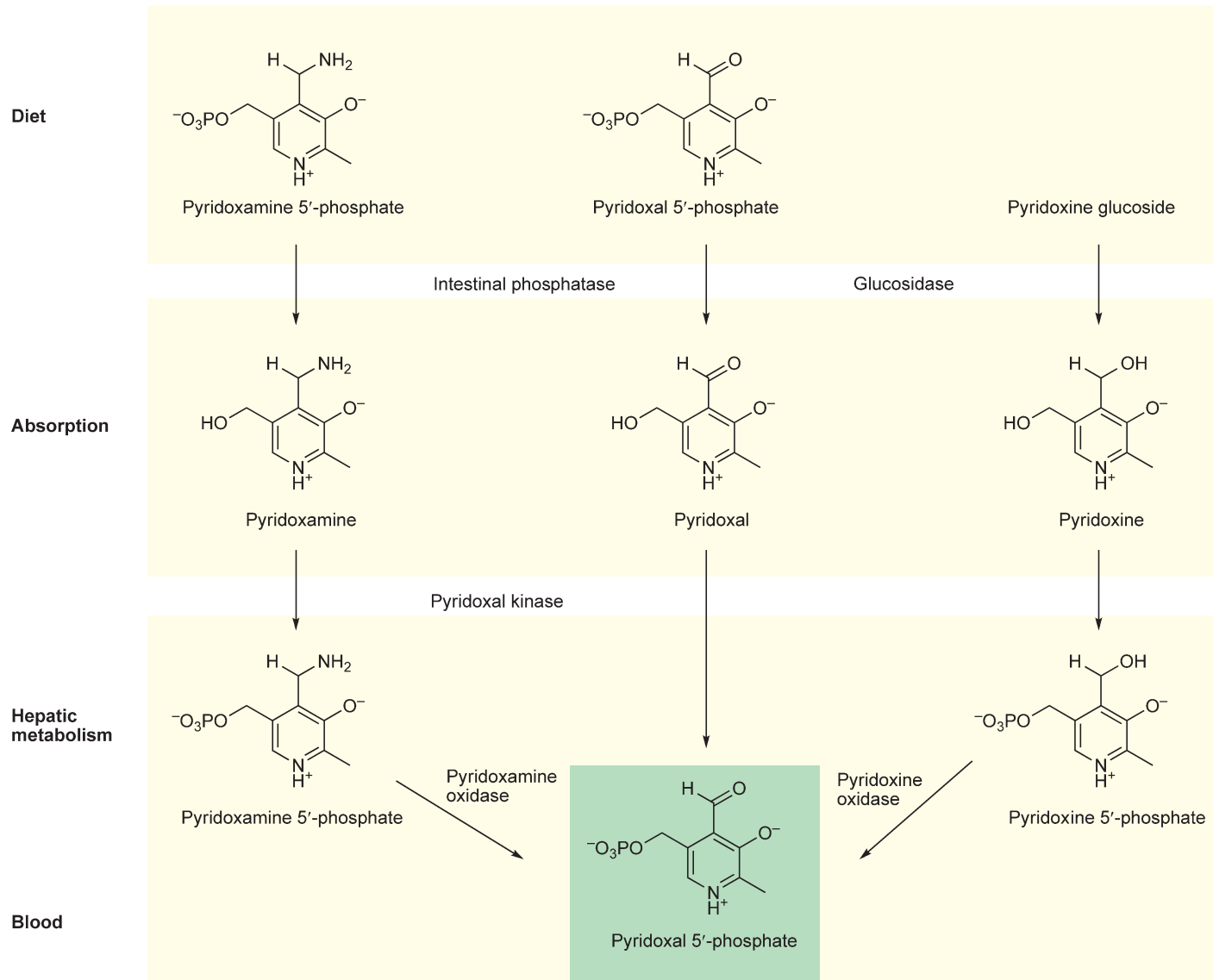


**Figure 1** Interconversion of B<sub>6</sub> vitamers with their 5'-phospho forms.

The recommended uptake of vitamin B<sub>6</sub> is 1.5 mg day<sup>-1</sup>. A reduced level of vitamin B<sub>6</sub>, due to deficiency of PLP-generating enzymes, causes neurological disorders such as seizures, which can be treated by supplementation of vitamin B<sub>6</sub> or PLP. Moreover, a variety of diseases are associated with the deficiency of specific PLP-dependent enzymes.<sup>2</sup> A typical case is cystathionine  $\beta$ -synthase, the enzyme that converts homocysteine into cystathionine (see Section 7.10.3.2.3). Its absence causes an increase in the plasma concentration of homocysteine, a recognized risk factor for atherosclerosis and cardiovascular diseases.<sup>3,4</sup> Furthermore, some PLP-dependent enzymes are targets of commercial drugs that inhibit enzyme action,<sup>5</sup> such as GABA-aminotransferase inactivated by vigabatrin to treat epilepsy<sup>6</sup> or DOPA decarboxylase (DDC) inactivated by carbidopa to treat Parkinson's disease.<sup>7</sup> The mechanisms of action of these compounds are discussed in Section 7.10.6. Other PLP-dependent enzymes are under investigation for the treatment of several diseases.<sup>5</sup>

### 7.10.1.2 Overview of PLP Catalysis

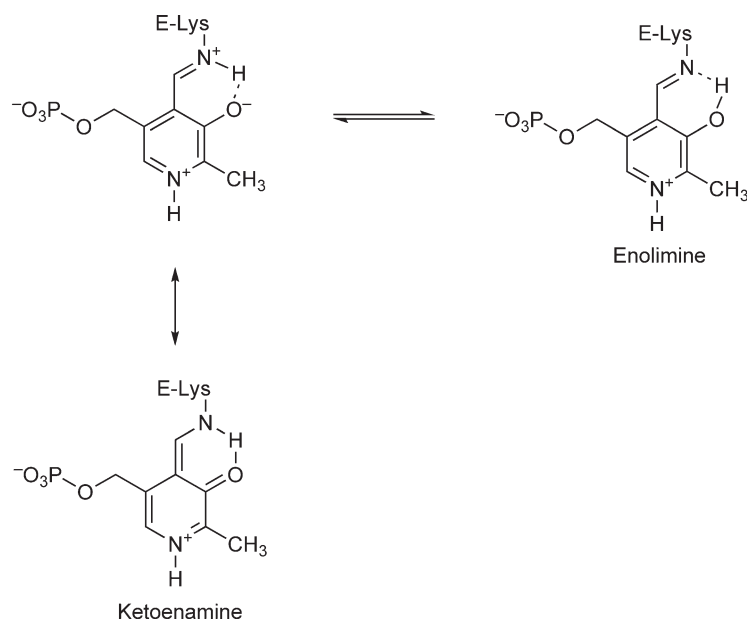
PLP is the coenzyme of more than 160 enzymes catalyzing a broad variety of reactions, predominantly involving an amino acid, an amine, or an oxoacid. Some of the selective examples are (1) the transfer of the alpha amine group from an amino acid to a ketoacid, catalyzed by transaminases, (2) the elimination or replacement of the substituent in the beta position of an amino acid or other compounds, catalyzed by lyases and synthases, respectively, and (3) the removal of the alpha carboxylic group from amino acids or other compounds, catalyzed by decarboxylases. More complex reactions are also carried out, involving, for example,



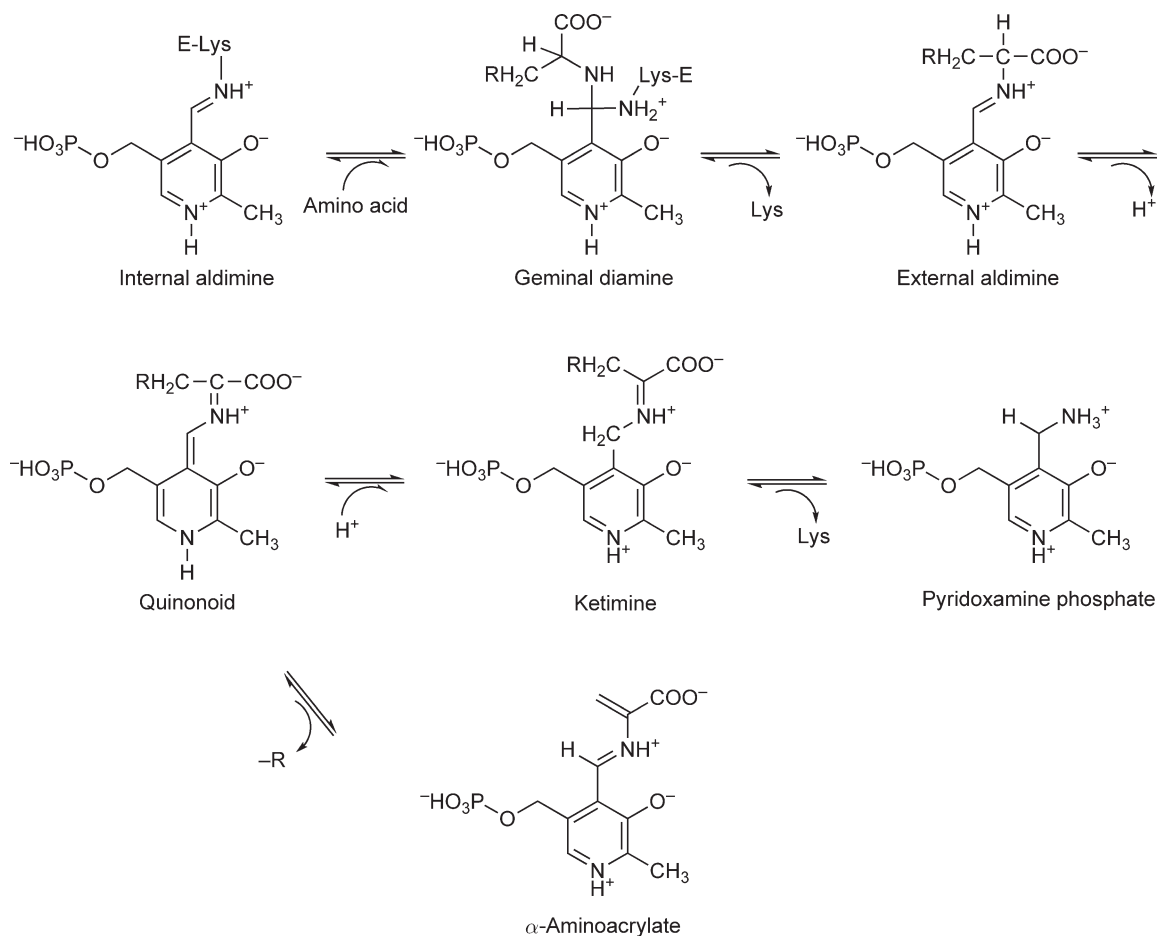
**Figure 2** Conversion of dietary B<sub>6</sub> vitamins to intracellular PLP cofactor.

substrate transamination and decarboxylation, such as in dialkyl glycine decarboxylase. In all these cases, catalysis takes place via a series of chemical transformations that exploit the electron sink properties of the pyridine ring. The catalytic mechanism of these enzymes is discussed in Sections 7.10.2–7.10.5. A distinct case is glycogen phosphorylase that degrades glycogen to produce glucose-1-phosphate, where PLP intervenes via an acid–base-assisted catalysis brought about by its 5' phosphate moiety. Glycogen phosphorylase catalysis will not be discussed.

In spite of the striking differences in the catalytic mechanisms, PLP-dependent enzymes share the common feature that the coenzyme is bound via a Schiff base to the  $\epsilon$ -amine group of an active-site lysine (**Figure 3**), leading to the internal aldimine species. The groups attached to the imine double bonds are consistently in a cisoid position, in agreement with calculations that indicate this geometric configuration is particularly stable. The internal aldimine can be distributed between two tautomeric forms, the ketoenamine absorbing at 400–420 nm, and the enolimine absorbing at 320–340 nm (**Figure 3**). The latter species is favored in apolar environments,<sup>8</sup> thus, the absorption band of a native PLP-dependent enzyme might exhibit a single or two overlapping bands. Furthermore, the spectrum of the internal aldimine depends on the protonation state of the imine N and of the pyridine N1 that, in turn, depends on the corresponding  $pK_a$  value of these groups. Frequently, these  $pK_a$  values have been found to be  $>9$ , thus both N atoms are protonated under physiological conditions. In most of the PLP-dependent enzymes, a negatively charged aspartate stabilizes the protonated N1 of the pyrimidine ring. In tryptophan synthase (TRPS), this residue is replaced by a neutral serine, whereas in alanine racemase (AlaR) this residue is a positively charged arginine, probably instrumental to a very high discrimination between  $\alpha$  and C reprotonation from the opposite side and protonation at C4', leading to transamination (see below). The first catalytic step, common to most PLP-dependent enzymes, requires a nucleophilic attack of the substrate-containing amine to the imine double bond of the internal aldimine with formation of a gem-diamine, accompanied by the change in the C4' hybridization from  $sp^2$  to  $sp^3$ . In order to achieve this reaction, the incoming substrate protonated amine is deprotonated by a protein basic residue. The gem-diamine evolves in the substrate's Schiff base, called the external aldimine. In several PLP-dependent enzymes, the transaldimination is accompanied by a conformational change from an open to a closed active-site state, that can control either substrate or reaction specificity. Moreover, PLP ring may tilt up to  $15^\circ$ , as observed in aspartate aminotransferase (AAT).<sup>9,10</sup> The relevance of these dynamic events are discussed in Section 7.10.7. On the substrate, external aldimine distinct chemical transformations take place depending on the specific



**Figure 3** The ketoenamine and enolimine tautomers of the internal aldimine.



**Figure 4** Structures of catalytic intermediates in PLP-catalyzed reactions.

enzyme. These involve the formation of a sequence of catalytic intermediates, whose structure depends on the specific bonds of the substrate that are affected by catalysis (Figure 4). Reactions at the  $\alpha$ -carbon of an amino acid are discussed in Section 7.10.2, and those of the  $\beta$ -carbon in Section 7.10.3, and reactions at the  $\gamma$ -position in Section 7.10.4. Catalytic intermediates that are common to several PLP-dependent enzymes are quinonoid species,  $\alpha$ -aminoacrylate, ketimine, and PMP. Each intermediate is characterized by distinct spectroscopic properties that allow monitoring the catalytic time courses and the determination of substrate, substrate analogues, and inhibitor binding affinity.<sup>8</sup>

Most of PLP-catalyzed reactions involve both nucleophilic and electrophilic attacks and proton transfers. In some cases, PLP reactions proceed via the formation of radical species. These are discussed in Section 7.10.5. Several compounds react with PLP enzymes leading to activity inhibition. These reactions are explained in Section 7.10.6.

The cooperation between the coenzyme and the residues within the active sites dictates the catalytic reaction and the substrate specificity, leading to the well-known versatility of PLP chemistry, unparalleled by any other coenzyme. Recently, the catalytic versatility of the PLP-dependent enzymes has been found to be expanded and modulated by the copresence of other coenzymes, as heme, *S*-adenosyl methionine (SAM), and cobalamine.

#### 7.10.1.2.1 Dunathan's hypothesis

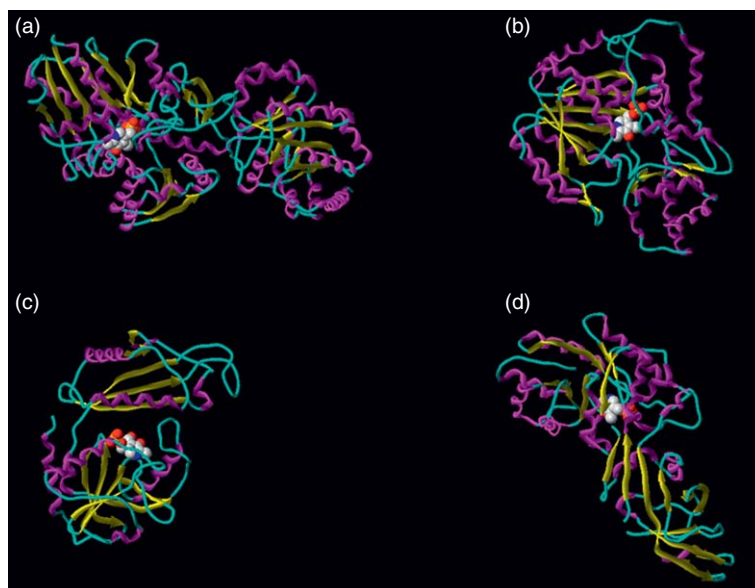
A key feature of PLP catalysis is the ability of the pyridine ring to stabilize negative charges storing electrons at the protonated N1 nitrogen, via the formation of a quinonoid species (Figure 4). The negative charge develops



when, upon binding of an amino-containing substrate to the internal aldimine and formation of the external aldimine, one of the  $\alpha$ -carbon bonds is heterolytically broken, generating a carbanion species. This unstable and reactive species is stabilized by a double-bond conjugation with the pyridine ring and formation of an extended molecular orbital. In order to achieve the electron delocalization, the  $p$  orbital of the  $\alpha$ -carbon containing the electron doublet must be perpendicular to the plane of the aromatic ring, that is, parallel to the ring  $\pi$  orbitals. This hypothesis, which explains the stereospecificity of the  $\alpha$ -bond breakage, was originally proposed by Dunathan<sup>11</sup> and later confirmed by structural data. Thus, while the coenzyme allows the catalysis, the protein matrix plays the essential role of selecting the substrate and anchoring it within the active site with the proper orientation for the  $\alpha$  scissile bond.

### 7.10.1.3 Families of PLP-Enzymes Based on Evolutionary Relationships and Fold Types of PLP-Enzymes Based on Three-Dimensional Structures

Several studies have explored the evolutionary relationship within the PLP-dependent enzymes family comparing either the nucleotide or amino acid sequences<sup>12–14</sup> or the three-dimensional structures<sup>15–17</sup> (see Section 7.10.8). On the basis of the structure, PLP-dependent enzymes are divided into fivefold types (Figure 5 and Table 1). Fold-type I is the most representative. It consists of two domains, a large domain and a small domain, which sandwich the coenzyme (Figure 5(a)). The prototype of fold-type I is AAT. Enzymes catalyzing reactions at  $\alpha$ - and  $\gamma$ -carbon belong to this fold type. Fold-type II contains only enzymes catalyzing  $\beta$ -elimination and  $\beta$ -replacement reactions. The prototype of the class is TRPS. Fold-type II consists of two domains of almost equal dimensions (Figure 5(b)). Fold-type III contains a few enzymes and represents AlaR. Fold-type III consists of two domains, an  $\alpha/\beta$  barrel and  $\beta$ -strand domain (Figure 5(c)). Fold-type IV represented by D-amino acid aminotransferase (DAAT) consisting of a small and a large domain (Figure 5(d)). Fold-type V represented by glycogen phosphorylase. In several cases of dimeric enzymes, the two active sites face each other, with residues of one subunit contributing to the formation of the opposite active site. All known PLP-enzymes are oligomeric, ranging from the prevalent dimeric form to dodecamers, and the monomeric form is catalytically inactive.



**Figure 5** Three-dimensional structures of the PLP-enzyme representative of the four fold types. Fold-type V, represented by the glycogen phosphorylase, is omitted.

**Table 1** Fold types and reaction types of selected PLP-dependent enzymes

Fold type	Representative enzymes	<i>C<math>\alpha</math></i> bond cleaved			Reaction
		–H	–COO <sup>–</sup>	–R	
I	<i>Aspartate aminotransferase</i>	+			Transamination
	<i>Phosphoserine aminotransferase</i>	+			Transamination
	<i>Ornithine aminotransferase</i>	+			Transamination
	<i>Selenocysteinyl-tRNA synthase</i>	+			$\beta$ -Replacement
	<i>Tryptophanase</i>	+			$\beta$ -Elimination
	<i>Cystathionine <math>\gamma</math>-lyase</i>	+			$\gamma$ -Elimination
	<i>Alanine racemase (fungal)</i>	+			Racemization
	<i>Glutamate decarboxylase</i>		+		Decarboxylation
II	<i>Serine hydroxymethyltransferase</i>			+	Side-chain cleavage
	<i>Tryptophan synthase</i>	+			$\beta$ -Replacement
	<i>Serine dehydratase</i>	+			$\beta$ -Elimination
	<i>Threonine synthase</i>	+			$\gamma$ -Replacement
III	<i>Serine racemase</i>	+			Racemization
	<i>Alanine racemase (bacterial)</i>	+			Racemization
	<i>Diaminopimelate decarboxylase</i>		+		Decarboxylation
IV	<i>D-Threonine aldolase</i>			+	Side-chain cleavage
	<i>D-Alanine aminotransferase</i>	+			Transamination
V	<i>4-Amino-4-deoxychorismate lyase</i>	(+)			$\beta$ -Elimination
	<i>Glycogen phosphorylase</i>		n/a		
VI	<i>D-Lysine 5,6-aminomutase</i>		n/a		
VII	<i>Lysine 2,3-aminomutase</i>		n/a		

n/a = not applicable.

#### 7.10.1.4 List of Previous Reviews on PLP-Enzymology

Over the years, PLP enzymology has been the subject of several reviews, dealing predominantly with the spectroscopy and chemistry,<sup>8</sup> catalytic mechanisms,<sup>18–20</sup> evolutionary relationships, structure–function relationships,<sup>16,17</sup> and reactivity with drugs and inhibitors.<sup>5,21</sup> The more investigated PLP-enzymes have also been specifically reviewed. These include the transaminases,<sup>22</sup> TRPS,<sup>23–28</sup> and *O*-acetylserine sulphydrylase (OASS).<sup>29,30</sup>

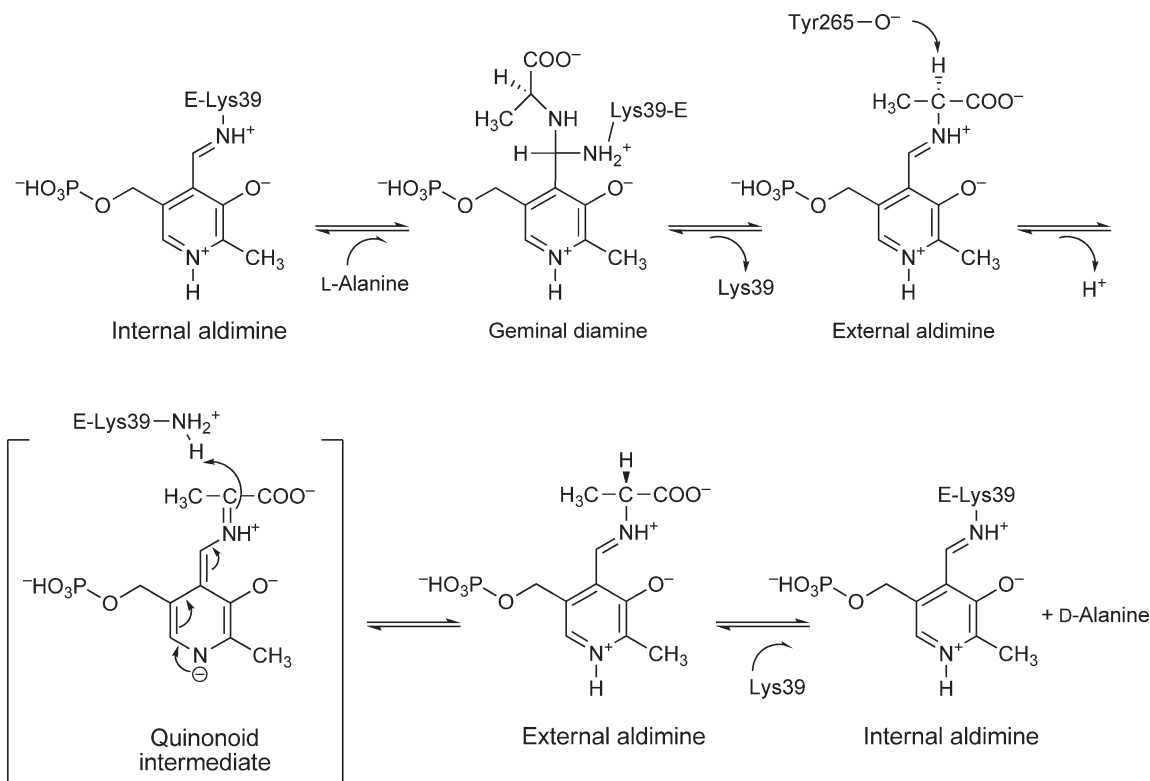
### 7.10.2 Reactivity of Amino Acids at $\alpha$ -Position

#### 7.10.2.1 Amino Acid Racemization

In nature, L-amino acids are one of the most widely used building blocks and metabolic precursors. In many instances, however, organisms benefit from the use of amino acids of opposite stereochemistry. Relevant examples are D-alanine and D-glutamate, key components of the bacterial cell wall, and D-serine, that acts as an endogenous agonist of the NMDA receptor in the human brain. These compounds are produced through the action of racemases, which evolved to catalyze inversions of stereochemistry. All the amino acid racemases known to date operate via an initial removal of the substrate  $\alpha$ -proton.<sup>31</sup> In the amino acid zwitterionic form, the  $pK_a$  of this proton has been estimated to be around 30,<sup>32</sup> thus this reaction is particularly difficult. Several racemases are PLP-dependent, although there are racemases operating in a PLP-independent manner. These include glutamate racemase, proline racemase, and aspartate racemase.<sup>31,33</sup>

##### 7.10.2.1.1 Alanine racemase

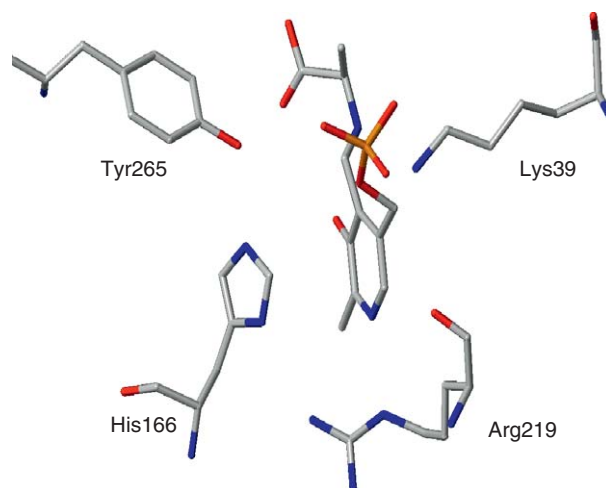
AlaR, the bacterial enzyme that catalyzes the racemization of L- and D-alanine, is the most investigated PLP-dependent amino acid racemase. It is also the only PLP-dependent racemase whose structure has been solved.



**Figure 6** Proposed reaction mechanism for AlaR. The hypothetical quinonoid intermediate is shown within brackets.

The enzyme is a homodimer belonging to the fold-type III of PLP-dependent enzymes.<sup>34–37</sup> The catalytic mechanism of AlaR has been investigated in detail because the enzyme is a target for the design of antibiotics and a better knowledge of its action might allow the development of mechanism-based inhibitors more specific than those discovered to date.<sup>5</sup> The mechanism of AlaR provides an example of PLP-catalyzed racemization reactions. As in all PLP-enzymes, the cofactor of AlaR is bound through an aldimine linkage to the  $\epsilon$ -amino group of an active-site lysine (Lys39), forming the internal aldimine (Figure 6). The amino group of the incoming substrate, L-alanine, displaces Lys39 forming an external aldimine, via a geminal diamine intermediate. The following two-base racemization mechanism steps have been proposed to involve Lys39 and Tyr265, that lie on opposite sides of the substrate.<sup>38</sup> C $\alpha$  deprotonation is carried out by Tyr265 to form a carbanion that is reprotonated by Lys39 to generate the external aldimine with inverted stereochemistry. Then, the internal aldimine between PLP and Lys39 is regenerated through transimination with release of D-alanine.<sup>37–40</sup>

Structural evidence and site-directed mutagenesis studies strongly support the two-base mechanism involving the phenolate anion of Tyr265 for the deprotonation of the substrate C $\alpha$ .<sup>39,41–43</sup> This two-base mechanism offers advantages for the control of reaction specificity, because the second catalytic residue is pre-positioned to rapidly donate a proton to the antipodal face of the intermediate, reducing the competition with alternative reaction pathways. However, a low pK<sub>a</sub> is required for Tyr265 to be an efficient base catalyst. The origin of the low pK<sub>a</sub> of the Tyr phenol group lies in the interactions that this residue undertakes in the AlaR active site. Tyr265 interacts with Arg219 via a bond with His166. This hydrogen bond network can lower the pK<sub>a</sub> of Tyr by transferring positive charge from Arg219 to His166 (Figure 7).<sup>39</sup> The positively charged Arg219 residue, in contrast with the negatively charged (transaminases) or neutral (TRPS) residue of other PLP-enzymes, interacts with the pyridine nitrogen (N1) of the cofactor, probably preventing its protonation.<sup>37</sup> This impairs the electron sink properties of the PLP ring casting some doubts on the presence of the quinonoid intermediate, typical of many PLP-catalyzed reactions.<sup>19</sup> This intermediate has not been detected



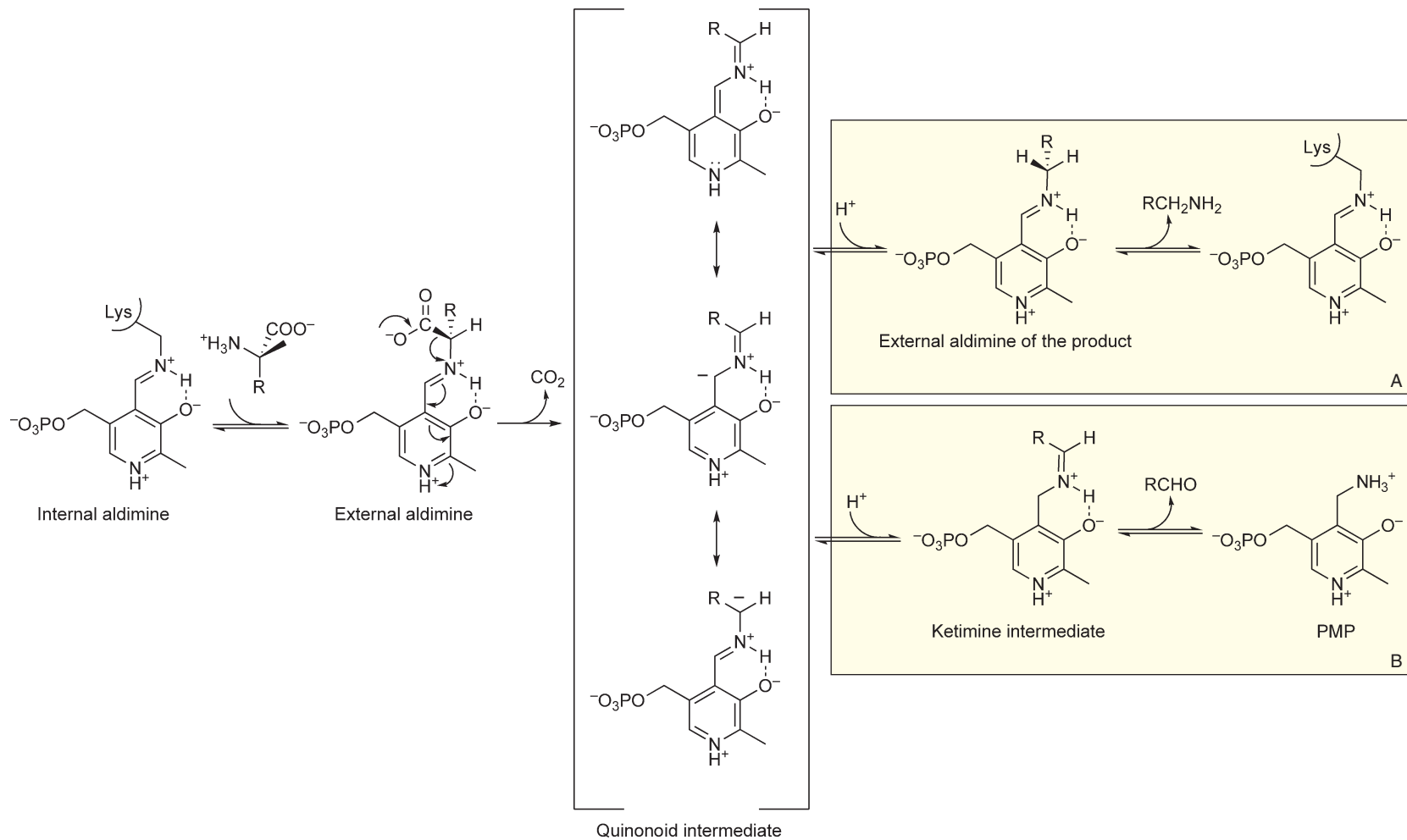
**Figure 7** AlaR bound with the external aldimine adduct, Ala-PLP (PDB code: 1L6G). For clarity some residues have been hidden.

in steady-state and stopped-flow absorbance spectra with wild-type AlaR. In the absence of the quinonoid intermediate, or in the presence of an unstable quinonoid intermediate, the carbanion formed from the  $\alpha$ -proton abstraction is expected to be very unstable. For this reason, an  $S_E2$  mechanism has also been considered, where the incoming proton develops some bonding with  $C\alpha$  in a concerted transition state.<sup>19</sup> Multiple hydrogen kinetic isotope effect studies have contradicted this hypothesis giving evidence of a stepwise mechanism with formation of a carbanionic intermediate.<sup>44</sup> Further calculations of free energy profiles of AlaR<sup>45</sup> have shown that, without protonation of N1, the quinonoid intermediate is at least  $4 \text{ kcal mol}^{-1}$  higher in energy than the aldimine intermediates, giving reason to the impossibility to observe this species by absorption spectroscopy. Toney proposed that this high energy of the quinonoid plays a role in reducing the frequency of transamination, the primary side reaction. According to this interpretation, the high-energy quinonoid intermediate undergoes reprotonation more rapidly than required for Lys39 or Tyr265 to move in proximity with the coenzyme  $C4'$  to yield transamination.<sup>46</sup> Thus, the action of the protein matrix to stabilize or destabilize the carbanionic quinonoid intermediate could be a second strategy to control reaction specificity in PLP-dependent enzymes.

### 7.10.2.2 Decarboxylation

The elimination of a  $\text{CO}_2$  molecule from the substrate catalyzed by decarboxylases requires the stabilization of a carbanionic intermediate, a task often performed by enzymatic cofactors. Indeed, decarboxylation reactions on  $\alpha$ -amino acids are catalyzed mainly by PLP-dependent enzymes, with a small fraction of reactions catalyzed by enzymes that use a pyruvoyl cofactor.<sup>47</sup> PLP-dependent decarboxylases are largely widespread among both eukaryotes and prokaryotes where they participate in the biosynthesis of biological amines (e.g., dopamine, histamine, and serotonin) and polyamines. In addition, in prokaryotes, inducible PLP-dependent decarboxylases take part in the regulation of intracellular pH.<sup>48</sup>

The general reaction mechanism for PLP-dependent  $\alpha$ -amino acid decarboxylation is shown in **Figure 8**. After the transamination reaction, a molecule of  $\text{CO}_2$  is eliminated from the external aldimine to yield the quinonoid intermediate. Contrary to the other reactions catalyzed by PLP-dependent enzymes, such as transamination, the decarboxylation is energetically irreversible.<sup>20</sup> Amino acid decarboxylases are remarkable enzymes in achieving the largest enhancement of reaction rate among enzyme-catalyzed reactions.<sup>47</sup> Structural data for some PLP-dependent decarboxylases suggest that this huge rate enhancement ( $\sim 10^{20}$ ) is partly due to the destabilization of the reaction ground state by the hydrophobic and/or negatively charged groups present in the carboxylate binding pocket. (see Jackson,<sup>49</sup> O' Leary,<sup>50</sup> and Jackson<sup>51</sup> and references therein)



**Figure 8** General reaction mechanism for PLP-dependent  $\alpha$ -amino acid decarboxylases.

Protonation at  $C_{\alpha}$  of the quinonoid intermediate (mechanism A) leads to the formation of the external aldimine of the product and eventually to the release of the amine that usually represents the rate-limiting step of the reaction.<sup>52</sup> Protonation at C4' (mechanism B) constitutes the main side reaction and leads to the formation of a ketimine intermediate. Actually, the rate of the abortive transamination is very low (about 0.01% of the turnovers for ornithine decarboxylase (ODC)),<sup>53</sup> due to the geometry of the active site that disallows proximity of acidic groups to  $C_{\alpha}$ .<sup>20</sup> The hydrolysis of the ketimine leads to the release of an aldehydic product with formation of PMP.

Interestingly, the decarboxylation reaction occurs on the *si* face of the cofactor,<sup>54</sup> leading to retention of configuration about  $C_{\alpha}$  in the case of fold-type I enzymes, whereas fold-type III decarboxylases bind the substrate with the carboxylate group on the *re* face of the cofactor and lead to an inversion of configuration about  $C_{\alpha}$ .<sup>49,54</sup>

The analysis of the sequences of 54 PLP-dependent decarboxylases led Christen and coworkers to the identification of four evolutionary distinct groups.<sup>55</sup> Groups I–III were demonstrated to belong to fold-type I,<sup>15</sup> whereas proteins of sequence group IV belong to fold-type III. Group I comprised only of glycine decarboxylase at the time of the study. Group II is the larger one and comprises, among others, histidine decarboxylase (HDC), glutamate decarboxylase (GDC), and DDC. This group is formed by decarboxylases distantly related to transaminases.<sup>55</sup> Groups III and IV comprised of basic amino acid decarboxylases, for example, ODC, lysine decarboxylase (LDC), arginine decarboxylase (ADC), and diaminopimelate decarboxylase (DAPDC). Interestingly, sequences of ADC and ODC are present in both groups, with no obvious rationale underlying the group division. This suggests that even enzymes with the same substrate specificity have evolved along multiple lineages.<sup>55</sup>

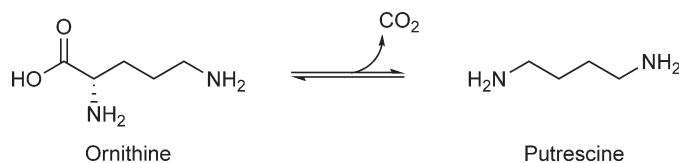
Finally, some PLP-dependent decarboxylases are well assessed drug targets, as for example, ODCs and DDCs (see below) and many more are being developed as targets of antibacterial drugs, as is the case of DAPDC – a bacterial enzyme that has no counterpart in higher eukaryotes.<sup>56</sup>

#### 7.10.2.2.1 Ornithine decarboxylase

ODC is the most thoroughly studied PLP-dependent decarboxylase and its reaction mechanism can be considered the prototype for basic amino acid decarboxylases.<sup>52,53</sup> ODC catalyzes the rate-limiting step in the biosynthesis of polyamines,<sup>57</sup> that is, the decarboxylation of the urea cycle intermediate L-ornithine to give putrescine (Figure 9). Polyamines are thought to play an important role in the support of cell growth and therefore in the control of cellular proliferation. In fact, *Trypanosoma brucei* ODC is the target of DFMO, a drug used in the treatment of African sleeping sickness.<sup>58</sup> Furthermore, ODC is the potential target of drugs against proliferative diseases.<sup>59,60</sup>

In eukaryotes, turnover of ODC is controlled by binding to an antizyme at the dimer interface with formation of an ODC–antizyme heterodimer that targets the decarboxylase to degradation by the 26S proteasome.<sup>61–63</sup>

To date, structures of ODC from *T. brucei* (TbODC),<sup>53,64,65</sup> *Lactobacillus 30a* (LbODC),<sup>66,67</sup> *Homo sapiens*,<sup>68</sup> *Mus musculus*,<sup>54</sup> and of the dual function enzyme from *Vibrio vulnificus*<sup>69</sup> have been solved. Prokaryotic ODCs fall under fold-type I, whereas eukaryotic ODCs fall under fold-type III.<sup>15</sup> The active form of the enzyme is a dimer with the two active sites placed at the subunit interface. The dimer can further oligomerize into a hexamer of dimers in the case of LbODC. The formation of the dodecamer is prevented by a single point mutation, G121Y.<sup>66</sup> Structural studies and information derived from site-directed mutants have helped in the identification of key catalytic residues. The active-site lysine, K69 (TbODC nomenclature), binds the cofactor



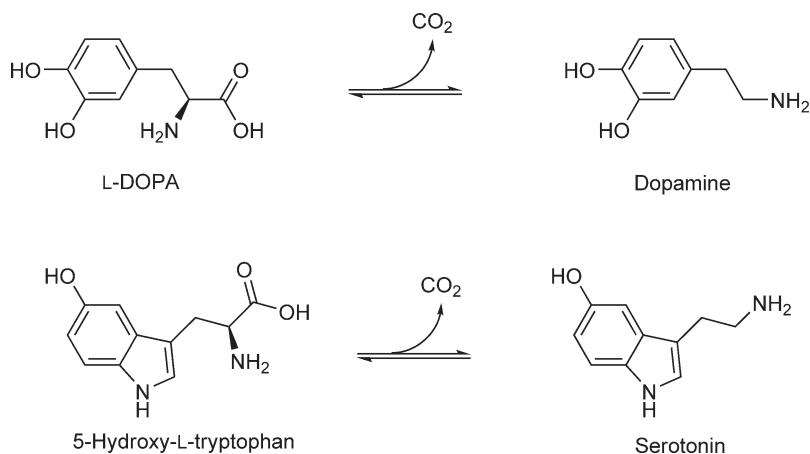
**Figure 9** Decarboxylation of L-ornithine leads to the polyamine putrescine, the precursor of spermidine and sperimine in the polyamine biosynthetic pathway.



and is also thought to play a role in the catalytic mechanism, as demonstrated by functional studies.<sup>70,71</sup> Investigations of K69R and K69A mutants demonstrate that the active-site lysine controls substrate selectivity via the proper positioning of the carboxylate moiety.<sup>70</sup> K69 lies in the carboxylate binding pocket in the unligated structure, but is displaced upon substrate binding.<sup>49,51</sup> Structural data with the substrate analogue D-ornithine strongly suggests that substrate binds to PLP with the leaving group on the *re* face of the cofactor.<sup>49</sup> Consequently, in the ligated form, the C<sub>α</sub> proton lies on the opposite face of ornithine with respect to the side chain of K69, and its abstraction by the active-site lysine is disfavored. E274 stabilizes the positive charge on the protonated pyridinium nitrogen and hence improves the electron sink properties of the cofactor.<sup>64,72</sup> In fact, the E274A mutant shows a large carbon isotope effect,<sup>71</sup> consistent with a slower decarboxylation step. D361 and D332 are considered, from the X-ray diffraction data in the presence of the inhibitor DFMO, to be important for substrate binding and discrimination.<sup>64</sup> In ODC, the typical large rate enhancement observed in decarboxylases is partly due to the destabilization of the ground state by hydrophobic and acidic residues lining the carboxylate binding pocket, that is, F397, Y389, and D361. The role played by hydrophobic residues in the carboxylate binding pocket is demonstrated by studies on the F397A mutant. In the mutant, the decarboxylation step is slowed down by a factor of about 10<sup>3</sup> and turns out to be rate limiting.<sup>49</sup> The strict control of the protonation step leads to a very high specificity for the elimination reaction and to a very low occurrence of the transamination side reaction.<sup>53</sup> C360 probably plays an important role in the control of reaction specificity. The mutation of C360 to alanine or serine increases to 90% the amount of the transamination product  $\gamma$ -aminobutyraldehyde.<sup>53</sup> Furthermore, structural data (Jackson and references therein) showed that the side chain of C360 rotates in and out of the active site as a function of the ligation state of the enzyme, with the residue being oriented inside the active site when the product is bound. This observation and data from mutational studies<sup>53</sup> suggest that C360 could serve as a general acid catalyst whose role in the protonation of C<sub>α</sub> is strictly controlled to prevent the reaction to take place before the decarboxylation step.

#### 7.10.2.2.2 DOPA decarboxylase

DDC catalyzes the conversion of L-3,4-dihydroxyphenylalanine (L-DOPA) into dopamine (Figure 10), a neurotransmitter found in the nervous system and peripheral tissues of both vertebrates and invertebrates<sup>73</sup> and also in plants where it is implicated in the biosynthesis of benzyloquinoline alkaloids.<sup>74</sup> DDC also catalyzes the decarboxylation of tryptophan, phenylalanine, and tyrosine and of 5-hydroxy-L-tryptophan to give 5-hydroxytryptamine (serotonin), and, therefore, is also referred to as aromatic amino acid decarboxylase. Inhibitors of DDC, for example, carbiDOPA and benserazide, are currently used in the treatment of Parkinson's disease to increase the amount of L-DOPA in the brain.



**Figure 10** Synthesis of the neurotransmitters dopamine and serotonin through decarboxylation of the substrates L-DOPA and 5-hydroxy-L-tryptophan, respectively.

DDC belongs to fold-type I of PLP-dependent enzymes<sup>15,75</sup> and to group II of PLP-dependent decarboxylases.<sup>55</sup> In the human DDC, as observed for all fold-type I enzymes, the aspartate involved in salt bridge formation with the pyridine nitrogen (D271) precedes the amino acid sequence in the active-site lysine (K303).

The binding of unprotonated L-DOPA to DDC is favored over the binding of the protonated form, due to the hydrophobic nature of the enzyme active site.<sup>76</sup> The hydrophobicity of the active site is suggested by the high population of the enolimine tautomer of the internal aldimine<sup>76</sup> that is favored over the ketoenamine tautomer in apolar environments.<sup>77</sup> As a consequence, the bimolecular association rate constant at pH 7 is unusually low (about  $10^5 \text{ mol l}^{-1} \text{ s}^{-1}$ ), due to the presence of a large fraction of protonated amino acid at neutral pH<sup>76</sup> whose  $\alpha$ -amino group  $pK_a$  decreases from 8.7 to 6.6 upon binding.<sup>78</sup> Binding of L-DOPA to the protonated internal aldimine shifts the enolimine to ketoenamine tautomeric equilibrium toward the latter, and leads to the formation of an unprotonated external aldimine that absorbs maximally at 380 nm.<sup>76,78</sup> Similar to many PLP-dependent enzymes, DDC undergoes a conformational change associated with catalysis that involves the open to closed transition of a mobile loop, not detectable in the three-dimensional structure.<sup>75,79,80</sup> Experiments of limited proteolysis suggest that the formation of the external aldimine is associated with loop closure and to the stabilization of a more rigid conformation.<sup>79</sup> The loop hosts two conserved residues among PLP-dependent decarboxylases: Y332 and R334.

The unprotonated external aldimine has weaker electron-withdrawing properties. Thus, the catalytic process must rely on the high exergonic properties of the decarboxylation reaction.<sup>78</sup> The quinonoid intermediate does not accumulate significantly, so that a prompt and facile protonation at  $C_\alpha$  is envisaged. The protonation step could involve the conserved Y332 whose acidity is increased by the action of another conserved residue, H192.<sup>75</sup> Alternatively, the fast protonation of the quinonoid intermediate could involve the O3' proton shuttled by the side chain of the active-site lysine.<sup>78</sup>

In addition to half-transamination toward D-amino acid, that constitutes the main side reaction catalyzed by PLP-dependent decarboxylases, DDC was demonstrated to catalyze the oxidative deamination of aromatic amines<sup>81–83</sup> with the attack of molecular oxygen to the quinonoid.<sup>81</sup>

#### **7.10.2.2.3 Dialkylglycine decarboxylase and diaminopimelate decarboxylase**

Dialkylglycine decarboxylase (DGD) is unusual as it catalyzes both a decarboxylation and a transamination during its normal catalytic cycle. DGD uses stereoelectronic effects to control its unusual reaction specificity.<sup>84,85</sup> The three-dimensional structure of DGD showed that the enzyme possesses two binding sites for monovalent cations (MVCs).<sup>86,87</sup> In particular, one site, located near the active site, binds potassium ions and controls the catalytic activity. The other is located at the carboxyl terminus of an alpha helix and probably has a structural role.

DAPDC catalyzes the last step of lysine biosynthesis in bacteria. Lysine is used by bacteria not only for protein synthesis but also in cellular processes fundamental for bacterial viability,<sup>88</sup> thus, drugs targeted against DAPDC could be efficient drug antituberculosis. DAPDC is the only decarboxylase, which acts on a carbon atom in D-configuration.<sup>89</sup> The three-dimensional structure of DAPDC has been solved confirming its classification as a fold-type III PLP-dependent enzyme.<sup>89</sup> Moreover, it was proposed that substrate specificity derives from the recognition of the L-chiral center of the substrate diaminopimelate and a system of ionic 'molecular rulers' that dictate substrate length.<sup>88</sup>

#### **7.10.2.3 $\alpha$ -Elimination and $\alpha$ -Replacement Reactions**

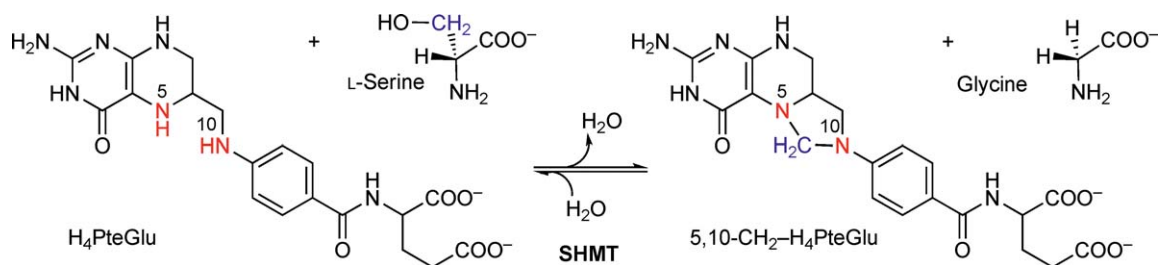
The cleavage or formation of a  $C\alpha-C\beta$  bond completes the possibilities of labilization of substituents at the  $\alpha$ -carbon of the substrates. On the one hand, serine hydroxymethyltransferase (SHMT) and threonine aldolase (TA) catalyze the aldol cleavage of the side chain of hydroxyamino acids, yielding glycine and the corresponding aldehyde. On the other hand, oxoamine synthases catalyze ester condensation reactions, that join acyl groups from CoA derivatives to external aldimines of either glycine or serine. An unusual cyclopropane ring-opening reaction, consisting of the cleavage of the  $C\alpha-C\beta$  bond of 1-aminocyclopropane-1-carboxylate (ACC), is catalyzed by ACC deaminase.

### 7.10.2.3.1 Serine hydroxymethyltransferase

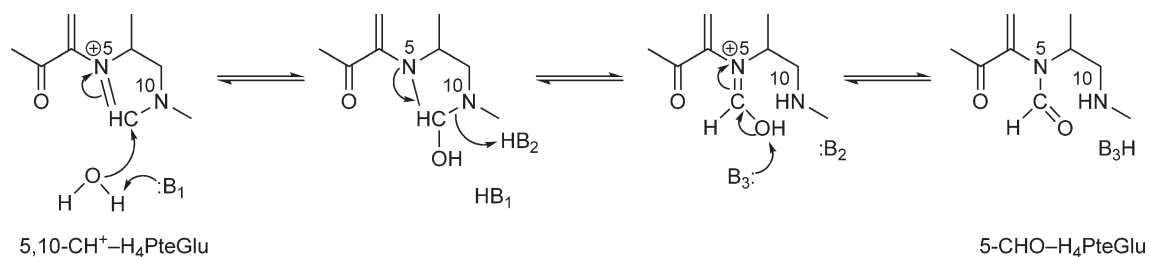
SHMT catalyzes the reversible transfer of  $C_{\beta}$  of serine to tetrahydropteroylglutamate ( $H_4PteGlu$ ) to form glycine and 5,10-methylene- $H_4PteGlu$  (Figure 11). This reaction represents the main source of one-carbon units required for the biosynthesis of purines, thymidylate, methionine, and several neurotransmitters.<sup>90</sup> SHMT has been intensively investigated because of the interest aroused by the complex mechanism of the hydroxymethyltransferase reaction and its broad substrate and reaction specificity. The enzyme has also been shown to catalyze, in the presence of glycine, the conversion of 5,10-methenyl- $H_4PteGlu$  into 5-formyl- $H_4PteGlu$  (Figure 12),<sup>91–93</sup> and evidence has accumulated that this activity is entirely responsible for the presence of 5-formyl- $H_4PteGlu$  in the cell.<sup>94,95</sup> Moreover, *in vitro* and in the absence of  $H_4PteGlu$ , SHMT catalyzes decarboxylation of aminomalonnate and retroaldol cleavage of several L- $\beta$ -hydroxyamino acids at rates sometimes approaching and even exceeding those of the hydroxymethyltransferase reaction.<sup>96,97</sup>

As L-serine was shown to be slowly converted by SHMT into glycine and formaldehyde in the absence of  $H_4PteGlu$ ,<sup>98,99</sup> for a long time the enzyme has been thought to catalyze both the cleavage of hydroxyamino acids and the hydroxymethyltransferase reaction through a retro-aldol cleavage mechanism (Figure 13(a)). According to this hypothesis, serine cleavage would be followed by condensation of formaldehyde and  $H_4PteGlu$  (Figure 13(b)), which is also known to take place rapidly and spontaneously.<sup>100</sup> The retro-aldol cleavage mechanism typically entails the presence of a catalytic base, which in SHMT must be located on the *re* face of the cofactor, to abstract the proton from the hydroxyl group of substrates and a periplanar conformation of the atoms involved in the transition state. Since the antiperiplanar conformation is energetically favored, it is predicted that starting from (3*R*)-[3-<sup>3</sup>H]serine and  $H_4PteGlu$  the product will be (11*S*)-[11-<sup>3</sup>H]methylene- $H_4PteGlu$ , if the released formaldehyde does not rotate freely, or a racemic mixture of 11*S* and 11*R* if it does.<sup>96</sup> However, two independent studies reported that the principal product of such a reaction is labeled at the 11*R* position.<sup>101,102</sup>

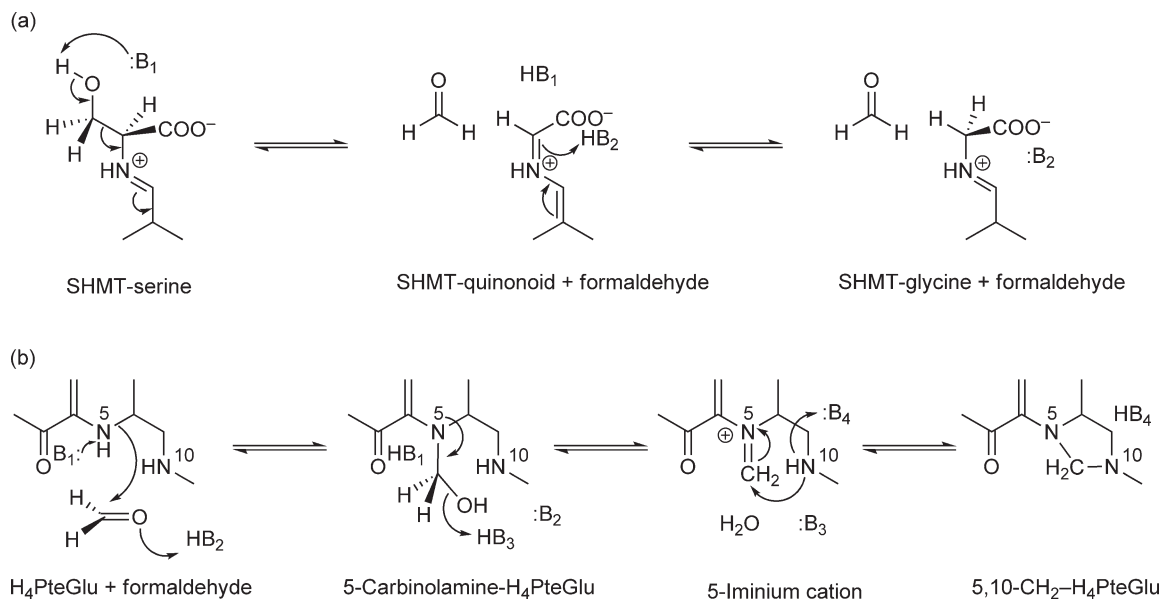
Several reaction mechanisms, alternative to the retro-aldol cleavage, have been proposed on the basis of kinetic, stereochemical, and reaction specificity studies and exhaustively reviewed.<sup>96,103</sup> However, only the relatively recent resolution of crystal structures of the enzyme from several sources<sup>104–108</sup> and characterization of site-specific mutants have resolved some of the crucial questions concerning the catalytic mechanism of SHMT.



**Figure 11** The serine hydroxymethyltransferase reaction.



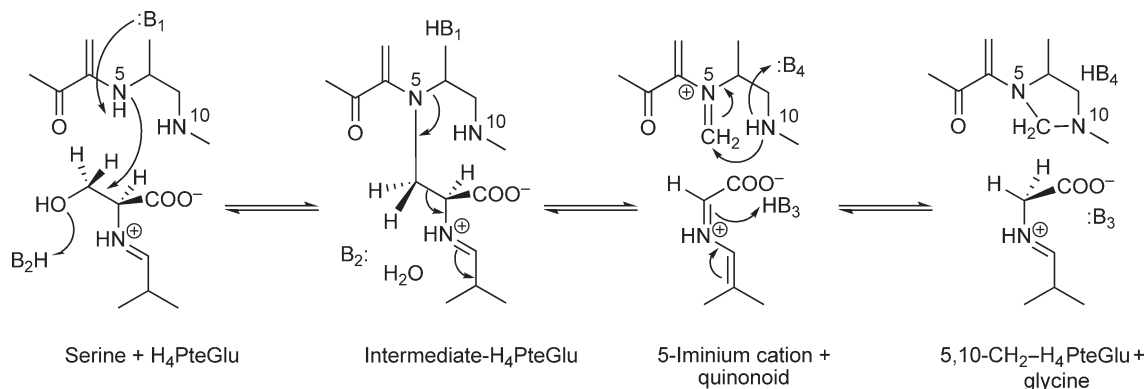
**Figure 12** Proposed mechanism for the hydrolysis of 5,10-methenyl- $H_4PteGlu$  to 5-formyl- $H_4PteGlu$  catalyzed by SHMT.



**Figure 13** Retroaldol cleavage of serine (a), followed by condensation between formaldehyde and H<sub>4</sub>PteGlu (b).

All SHMTs catalyze the folate-independent cleavage of both *erythro* and *threo* isomers of  $\beta$ -hydroxyamino acids, favoring the *erythro* forms. Modeling of both isomers of *L*- $\beta$ -phenylserine into the active site indicated that the *erythro* form can easily bind only in the synperiplanar conformation; conversely, the *threo* form can need to assume the antiperiplanar conformation in order to be accommodated without steric clashes of the phenyl group.<sup>109</sup> An analogous situation is observed for the isomers of *L*-threonine. It is therefore plausible that hydroxyamino acids are cleaved by SHMT in both synperiplanar and antiperiplanar conformations.

This apparent lack of stereochemical constraints and the fact that investigators were not able to identify the strong basic group required by the retro-aldol mechanism<sup>110–116</sup> suggests that removal of the proton from the hydroxyl group is not the first step in both folate-dependent and -independent cleavage of substrates.<sup>109</sup> The resolution of *Bacillus stearothermophilus* SHMT, in complex with *L*-serine in the synperiplanar conformation, led to the proposal of a single-step nucleophilic displacement mechanism, which avoids formation of formaldehyde.<sup>104</sup> However, this does not address the folate-independent cleavage of hydroxyamino acids, which is assumed to take place through the retro-aldol mechanism. According to this mechanism (Figure 14), which requires the presence of a proton donor at the active site, N5 of H<sub>4</sub>PteGlu attacks C3 of serine and as a



**Figure 14** Single-step nucleophilic displacement mechanism.

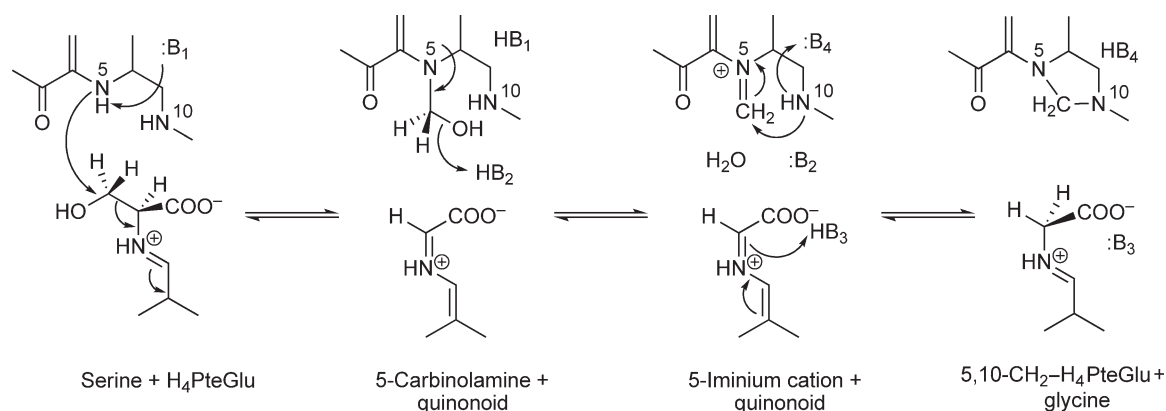
consequence the hydroxyl group is protonated and eliminated as water; the covalent adduct formed between H<sub>4</sub>PteGlu and serine collapses into products without formation of formaldehyde. However, the examination of the structures of SHMT complexes with serine, and with glycine and 5-formyl-H<sub>4</sub>PteGlu, revealed that N5 of the folate substrate is not in a good position to carry out an S<sub>N</sub>2 nucleophilic substitution of the serine hydroxyl. Moreover, in the reverse reaction, the S<sub>N</sub>2 substitution of N5 of the covalent adduct with a weak base, such as water, would be very unfavorable.<sup>109</sup>

These geometric and chemical difficulties, together with the results of recent site-directed mutagenesis studies on the role of E75 in rabbit cytosolic SHMT (E53 in *B. stearothermophilus* SHMT, E57 in *E. coli* SHMT, and E74 in sheep SHMT), led to a new proposal, which satisfies the stereochemical requirements and is in agreement with most experimental results.<sup>117</sup> E75 is located at hydrogen bond distance from the hydroxyl group of substrates and carbinolamine intermediate and is in the protonated form.<sup>107,109</sup> Mutation of E75 of sheep SHMT into Gln and Lys had already provided evidence that this residue is crucial for the folate-dependent cleavage of substrates, but has no apparent role in the folate-independent cleavage of *L*-allo-threonine.<sup>115</sup> Mutation of E75 of rabbit SHMT into either Gln or Leu lower the hydroxymethyltransferase activity by several 100-fold and abolish the methenyl-H<sub>4</sub>PteGlu hydrolase activity. However, the mutant enzymes catalyze the cleavage of hydroxyamino acids more efficiently, including serine cleavage into glycine and formaldehyde.<sup>109</sup> These data enabled a modified retro-aldol mechanism (Figure 15), in which the nucleophilic attack of H<sub>4</sub>PteGlu N5 to serine C3 determines the breakage of the C $\alpha$ -C $\beta$  bond, with the elimination of the quinonoid, which is then protonated to give glycine, and formation of the carbinolamine intermediate. Protonation of the carbinolamine hydroxyl group by E75 catalyzes the formation of the iminium cation, that develops into 5,10-methylene-H<sub>4</sub>PteGlu.<sup>109,117</sup> The mechanism of 5,10-methenyl-H<sub>4</sub>PteGlu hydrolysis evidently requires the same acid-base catalysis. At present, experimental evidence cannot help discriminating between this concerted mechanism and the equally possible one in which the breakage of the C $\alpha$ -C $\beta$  bond precedes the formation of the N5-C $\beta$  bond.

SHMT also catalyzes racemization of alanine and transamination of both its enantiomers. The particular reaction catalyzed by SHMT is mainly determined by the structure of the amino acid substrate. In the case of serine or glycine, the true substrates, SHMT does not catalyze any of the alternate reactions. The currently accepted model attributes this reaction specificity to the existence of open and closed active-site conformations. The physiological substrates generate the closed conformation, whereas alternate substrates react while the enzyme remains in the open conformation, which permits reaction paths leading to decarboxylation, transamination, and racemization.<sup>118</sup>

### 7.10.2.3.2 Threonine aldolase

TA is another fold-type I enzyme, which catalyzes the cleavage of the C $\alpha$ -C $\beta$  bond of hydroxyamino acids.<sup>119</sup> TA and SHMT have very similar catalytic properties, being able to catalyze the same set of reactions, that is, the retro-aldol cleavage of several different erythro and threo *L*- $\beta$ -hydroxyamino acids and the transamination



**Figure 15** Modified, concerted retroaldol cleavage mechanism proposed for the SHMT reaction.

and racemization of alanine.<sup>120</sup> The evolutionary analysis and the structural comparison between SHMTs and TAs suggest that the enzymes are closely related. Moreover, AlaR from *Cocliobolus carbonum* was predicted to be a member of fold-type I and a strict relative of TA and SHMT.<sup>120</sup> These observations make SHMT, TA, and fungal AlaR a good system for revealing the structural determinants of the reaction specificity in PLP-dependent enzymes. Interestingly, threonine aldolases and AlaRs are also found among fold-type III enzymes. The presence of both activities in two distinct, evolutionary unrelated families represents an interesting example of convergent evolution. The independent emergence of the same catalytic properties in families characterized by completely different folds may not have been determined by chance, but by the similar structural features required to catalyze PLP-dependent aldolase and racemase reactions.<sup>121</sup>

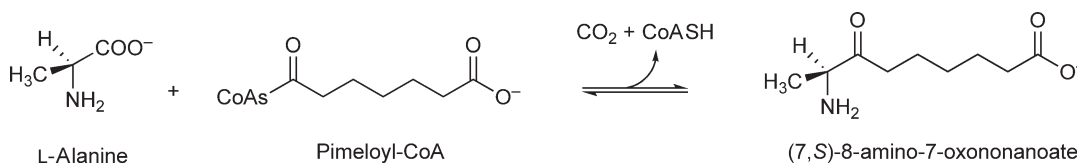
#### 7.10.2.3.3 $\alpha$ -Oxoamine synthases

The  $\alpha$ -oxoamine synthases family is a small group of fold-type I enzymes that catalyze Claisen condensations between amino acids and acyl-CoA thioesters (Figure 16). Members of this family are (1) 8-amino-7-oxononanoate (AON) synthase (AONS), which catalyzes the first committed step in the biosynthesis of biotin,<sup>122</sup> (2) 5-aminolevulinate synthase (ALAS), responsible for the condensation between glycine and succinyl-CoA,<sup>123</sup> which yields aminolevulinate, the universal precursor of tetrapyrrolic compounds, (3) serine palmitoyltransferase (SPT), which catalyzes the first reaction in sphingolipids synthesis,<sup>124</sup> and (4) 2-amino-3-ketobutyrate CoA ligase (KBL), involved in the threonine degradation pathway.<sup>125</sup> With the exception of the reaction catalyzed by KLB, all condensation reactions involve a decarboxylase step.

The understanding of the chemical details of the  $\alpha$ -oxoamine synthases mechanism derived from a series of complementary investigations was included radiolabeling studies on ALAS,<sup>126</sup> kinetic studies on AONS<sup>127,128</sup> and ALAS,<sup>129–131</sup> and detailed crystallographic studies on KBL<sup>125</sup> and AONS.<sup>128,132</sup> The reaction catalyzed by AONS starts with a transaldimination that forms the external aldimine with L-alanine. The following step is the binding of pimeloyl-CoA to the enzyme, which determines a conformational change of the dimeric protein and appears to favor the formation of a quinonoid intermediate through  $\alpha$ -proton abstraction (Figure 17 – step 1). Then, pimeloyl-CoA condenses with the quinonoid and HSCoA is liberated (Figure 17 – step 2). Pimeloyl addition takes place from the face opposite that of deprotonation. Therefore, after condensation, the position of the carboxylate is equivalent to that previously held by the  $\alpha$ -proton. The  $\beta$ -ketoacid intermediate formed in the Claisen condensation<sup>133</sup> decarboxylates giving a second quinonoid (Figure 17 – step 3), which is then protonated yielding AON external aldimine (Figure 17 – step 4). Modeling of the  $\beta$ -ketoacid into the active-site conformation that AONS assumes, when the product aldimine is formed, showed that the C8-carboxylate bond is not aligned with the  $\pi$  orbitals of the cofactor, as the Dunathan's stereoelectronic hypothesis requires.<sup>133</sup> On the other hand, decarboxylation may be promoted by the C7 ketone, which easily accepts an electron and is stabilized by an adjacent histidine residue. However, the formation of a quinonoid intermediate upon decarboxylation suggests that the cofactor also plays an important role.

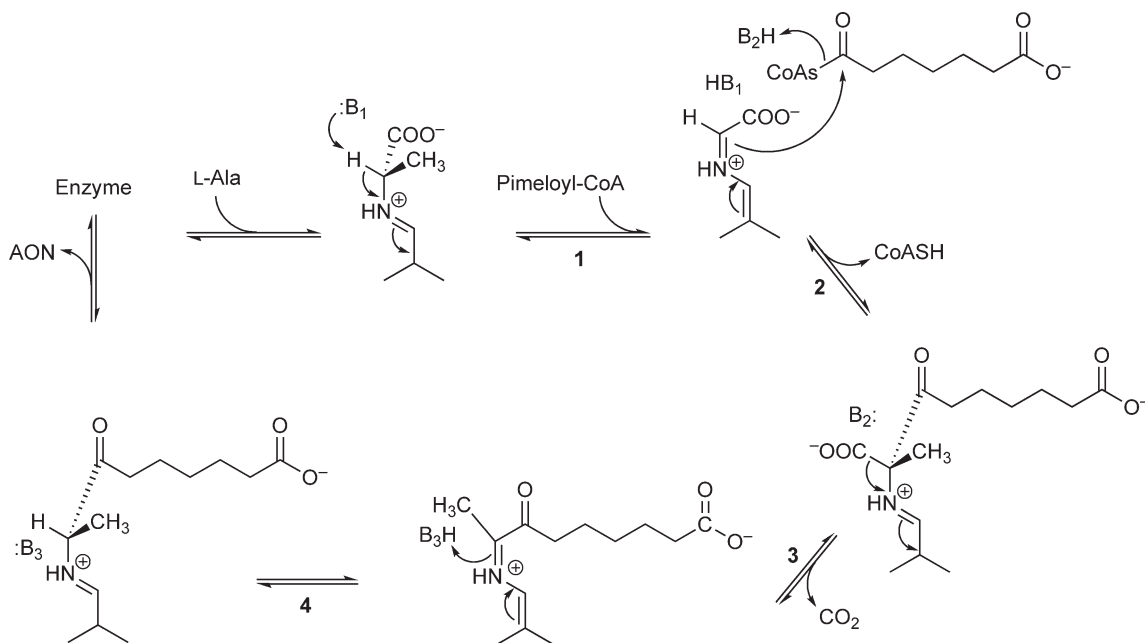
#### 7.10.2.3.4 1-Aminocyclopropane-1-carboxylate deaminase

In plants, ACC is the precursor of ethylene, a hormone that regulates various aspects of plant development and growth.<sup>134</sup> Several bacteria that promote plant growth are able to convert ACC into  $\alpha$ -ketobutyrate and ammonia, thanks to the action of ACC deaminase.<sup>135</sup> This is a unique PLP-dependent enzyme since it catalyzes the cleavage of a C $\alpha$ -C $\beta$  bond without the formation of a carbanionic intermediate. Indeed, ACC has no removable  $\alpha$ -hydrogens and its carboxyl group is retained in the product. In the light of this consideration, Walsh *et al.*<sup>136</sup> proposed two main reaction mechanisms, the first of which is initiated by a direct  $\beta$ -hydrogen abstraction, while the second proceeds through a nucleophilic attack carried out by an enzyme residue on the  $\beta$ -carbon of ACC. Both



**Figure 16** The reaction catalyzed by 8-amino-7-oxononanoate synthase, a member of the  $\alpha$ -oxoamine synthases family.





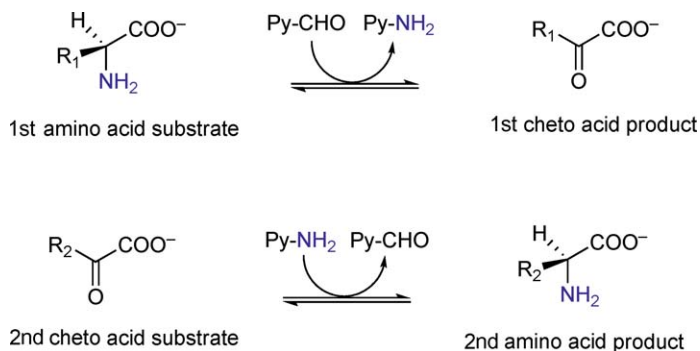
**Figure 17** Mechanism of the AONS reaction.

deprotonation and nucleophilic attack result in the opening of the cyclopropane ring. The crystal structure of the enzyme, which belongs to the fold-type II group, has been solved from *Pseudomonas*<sup>137</sup> and the yeast *Hansenula saturnus*.<sup>138</sup> Mechanistic studies based on site-directed mutagenesis and crystallographic characterization of reaction intermediates generated with a modified substrate (1-amino-2-methylenecyclopropane-1-carboxylic acid) supported both mechanisms.<sup>137,139</sup> However, because of the more sound basic chemistry, in terms of electronic configuration and thermodynamic favorability, the nucleophilic attack mechanism appears more likely.<sup>140</sup>

## 7.10.2.4 Transamination Reaction

### 7.10.2.4.1 Aminotransferases

Aminotransferases catalyze the reversible exchange of a primary amino group between amino and oxo compounds, via a double-displacement reaction in which the cofactor shuttles between its aldehyde and amino forms (Figure 18). Over a long period, AAT has been the most investigated PLP-dependent enzyme. It has been

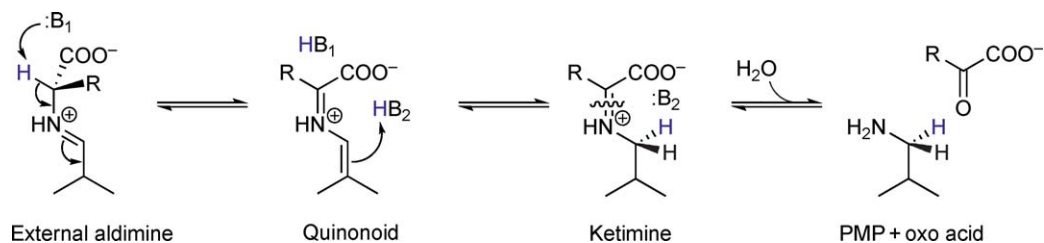


**Figure 18** The two half reactions of the transamination process. The cofactor (Py) shuttles between the pyridoxal phosphate form (Py-CHO), bound to the enzyme as internal aldimine, and the pyridoxamine phosphate (PMP) form (Py-NH<sub>2</sub>).

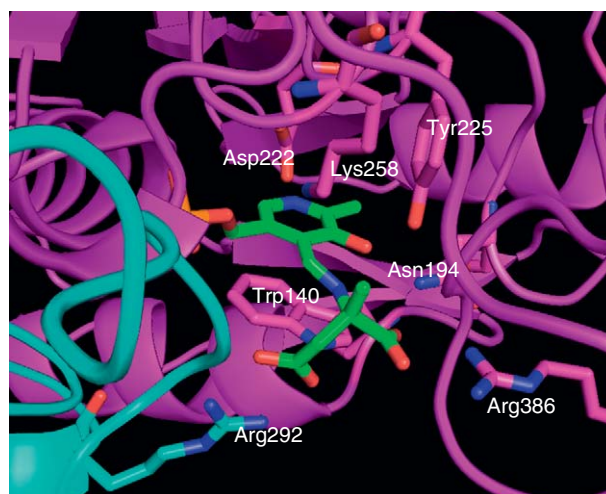
considered a model in the study of aminotransferases, among all PLP-dependent enzymes, since its activity was discovered in pigeon breast muscle in 1937.<sup>22</sup> The crystal structures of chicken mitochondrial and cytosolic AAT were the first three-dimensional structures of PLP-dependent enzymes to be solved,<sup>10,141</sup> explaining many of the experimental mechanistic data accumulated in the previous years.<sup>142</sup> With the advent of site-directed mutagenesis and the overexpression technology, the attention of researchers focused on the enzyme from *Escherichia coli*, leading to many important findings, some of which were of general relevance for PLP-based catalysis.

The crucial feature of the transamination mechanism is a tautomerization reaction, or 1,3 prototropic shift, in which a proton is transferred from the substrate carbon, which is directly bound to the amino group to the C4' of the cofactor (Figure 19). In this process, the external aldimine of the first amino acid substrate is isomerized into a ketimine, which is then hydrolyzed, liberating the related oxo acid product and the enzyme with the cofactor in the PMP form. The second half transamination reaction takes place via a reverse mechanism, starting from the enzyme in the PMP form and the second oxo acid substrate.

In AAT, the lysine residue (K258) that forms the internal aldimine is eliminated in the transaldimination reaction and optimally positioned on the *si* face of the cofactor to catalyze the 1,3 prototropic shift (Figure 20), acting as B<sub>1</sub> and B<sub>2</sub> in Figure 19. Several targeted mutagenesis investigations confirmed the role of proton/donor acceptor of K258<sup>144–147</sup> and extended the catalytic function of this residue to the transaldimination process.<sup>148</sup> In aminotransferases, reaction specificity must be guaranteed by avoiding that the quinonoid intermediate is protonated from the *re* side of the cofactor, which is accessible from the solvent. No acid residues are present in AAT on this side of PLP. However, water molecules act as proton donors, leading to the very slow racemization of substrates.<sup>149,150</sup> This represents the principal side reaction catalyzed by AAT, which is minimized by closure of the enzyme triggered by substrate binding,<sup>151</sup> as the internal aldimine is converted into the external aldimine, which excludes water from the *re* side of the cofactor. The closure of the enzyme also



**Figure 19** Mechanism of a half transamination reaction involving a generic amino acid with side chain R.



**Figure 20** Cartoon representation of the active site of dimeric pig cytosolic AAT (subunits are colored in magenta and cyan) in complex with L-2-methylaspartate, in which the external aldimine of PLP and the amino acid residues cited in the text are represented by a stick model.<sup>143</sup>

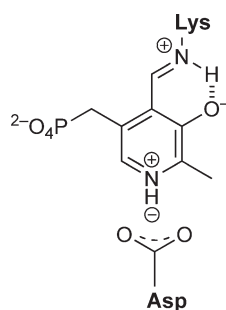
corresponds to a tilt of the PLP pyridine ring, which in unliganded AAT is angled at about  $10^\circ$ , with respect to the plane of the indole ring of W140,<sup>9</sup> and becomes nearly parallel to it when the quinonoid intermediate is formed.<sup>10</sup> Additionally, W140 is involved in the binding of the distal carboxylate group of 4-carbon dicarboxylic substrates. Mutagenesis experiments involving W140<sup>152</sup> supported the idea that the improved stacking interactions between PLP and the indole ring favors formation of the quinonoid intermediate.<sup>10</sup>

In AAT, D222 is located within the hydrogen bond and the strong ionic interaction distance from the pyridine nitrogen of PLP ( $N_1$ ). This interaction maintains the  $N_1$  of PLP in the protonated form (Figure 21) and therefore enhances the electron-withdrawing capacity of the cofactor. The role of D200 has been demonstrated by site-directed mutagenesis experiments and in particular by the characterization of mutant D222A, reconstituted with the coenzyme analogue *N*(1)-methylated PLP, in which the pyridine nitrogen bears a fixed positive charge.<sup>153,154</sup> The relevance of the stabilizing interaction is also evident from the observation that the aspartate residue is the only structurally and functionally conserved residue of fold-type I enzymes, the active-site lysine being not structurally conserved.<sup>155</sup>

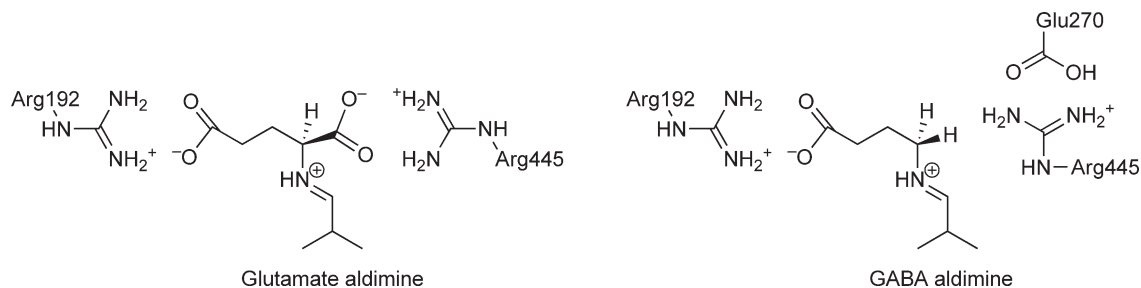
The electronic distribution within the coenzyme needs to be regulated during the catalytic process. In AAT, Tyr225 and Asn194 contribute to this regulation by making hydrogen bonds to  $O3'$  of the cofactor. Similar interactions are found in many other fold-type I enzymes. Site-directed mutagenesis investigations on AAT suggested that the role of these interactions is to lower the  $pK_a$  of the imine nitrogen of the internal aldimine bond, so as to stabilize its unprotonated form in the neutral pH range.<sup>156,157</sup> N194 and Y225 were also proposed to be part of a hydrogen-bonding network, that also involves R386 and the  $\alpha$ -carboxylate group of the substrates, which has a role to facilitate substrate binding and transaldimination.<sup>158</sup> When ligands bind, R386 moves by 2.8 Å to form hydrogen bonds and a salt bridge with the  $\alpha$ -carboxylate and a hydrogen bond with N194  $O\delta 1$ .<sup>10</sup> In this process, which takes place when the enzyme is converted into the closed conformation, the  $pK_a$  of the internal aldimine rises from 6.2 to 8.2 and five water molecules are expelled from the active site. At neutral pH, this change of  $pK_a$  determines the withdrawal of a proton from the amino group of the incoming substrate, which is then able to attack the  $C4'$  of PLP.

AAT is exceptional among aminotransferases in that all of its substrates are dicarboxylic compounds of similar size, that can be easily bound by two arginine residues properly positioned in the active site: R386 and R292\* (asterisk indicates that the residue is contributed by the neighboring subunit), which make hydrogen bonds and salt bridges to the proximal and distal carboxylate groups, respectively.<sup>159–162</sup> Many other aminotransferases, however, use substrate–product pairs with different structures and therefore show a dual specificity, being able to accommodate both pairs, and at the same time refuse nonsubstrate amino and oxo compounds. For these enzymes substrate binding is a much more complex problem, which has been solved using different strategies.

Histidinol phosphate aminotransferase, that catalyzes the transamination between histidinol phosphate and oxoglutarate, accommodates the side chains of the two different substrates in spatially distinct binding sites. While both the phosphate and  $\alpha$ -carboxylate moieties of the substrates bind to the same arginine residue, the oppositely charged imidazole and  $\gamma$ -carboxylate have separate binding sites.<sup>163</sup> Most aminotransferases, however, evolved so as to bind substrates at the same site. Since many of them use glutamate and oxoglutarate as one substrate/product pair, the common binding site must be able to accept alternatively a charged  $\gamma$ -carboxylate and a side chain with either neutral or positively charged character. Two different solutions to this problem are



**Figure 21** Stabilizing interaction between the structurally and functionally conserved aspartate residue of fold-type I enzymes and the protonated pyridine nitrogen of PLP.



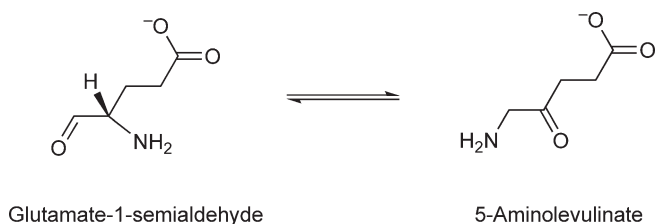
**Figure 22** Schematic example of the arginine switch in GABA aminotransferase.

found in aminotransferases: arginine switches and hydrogen bond networks. The first strategy concerns the ability to switch the arginine residue that interacts with the carboxylate of one substrate out of the active site when the other substrate is bound. Such a tactic is employed by  $\gamma$ -aminobutyric acid (GABA) aminotransferase<sup>164,165</sup> and ornithine aminotransferase,<sup>166</sup> which use both  $\omega$ -amino acids and glutamate as substrates. In GABA aminotransferase, similar to AAT, two arginine residues interact with the carboxylate groups of glutamate. However, when GABA is bound, the arginine residue that interacts with the proximal carboxylate of glutamate, swings out to form a salt bridge with a glutamate residue of the active site (Figure 22). Another example of the arginine switch is found in tyrosine aminotransferase, that binds aromatic amino acids such as tyrosine, phenylalanine, and tryptophan, as well as dicarboxylic aspartate and glutamate.<sup>167</sup>

Interestingly, the arginine switch mechanism was first recognized when it was artificially induced in AAT. When AAT was mutated in six distinct positions, a substantial increase of activity with aromatic substrates was observed.<sup>168</sup> The crystal structure of the engineered enzyme showed that the aromatic side chains could be accommodated at the active site as a result of R292 movement.<sup>169</sup> A similar observation was made on an AAT mutant form whose substrate specificity was broadened using direct evolution techniques, in order to include branched chain and aromatic amino acids.<sup>170–172</sup>

Aminotransferases belonging to subfamily I $\gamma$ ,<sup>173</sup> which also bind glutamate as substrate, do not have a conserved arginine residue, equivalent to R292\*, that binds the distal carboxylate. In *Pyrococcus horikoshii* AAT, a member of subfamily I $\gamma$ , the distal carboxylate of glutamate is recognized directly by a tyrosine residue, which is part of a hydrogen bond network.<sup>174</sup> Modeling of substrates into tyrosine aminotransferase from the same organism showed that this other member of subfamily I $\gamma$  also binds glutamate via an extended hydrogen-bonding network;<sup>175</sup> this network simply rearranges when tyrosine or other uncharged substrates are bound. The lack of a flexible, positively charged arginine residue do not only allow the alternate binding of negatively charged and uncharged substrate, but also permits to distinguish between very similar compounds, such as glutamate and aspartate. The specificity ratio ( $k_{\text{cat}}/K_{\text{m}}^{\text{Glu}}/k_{\text{cat}}/K_{\text{m}}^{\text{Asp}}$ ) for tyrosine aminotransferase from *Pyrococcus horikoshii* is 3400, compared to 0.27 for the *E. coli* enzyme.<sup>175</sup> A similar substrate binding mode is found in branched chain amino acid aminotransferase, a fold-type III enzyme.<sup>176</sup> In this enzyme, an arginine residue is present in the side chain binding pocket, but it makes a monodentate interaction with the distal carboxyl group of glutamate and is extensively hydrogen bonded to the neighboring residues, which are mainly aromatic. This situation allows the arginine residue to remain in position when the branched chain substrates bind, without the need of a large rearrangement of the hydrogen bond network.

DAAT, another fold-type III aminotransferase, and fold-type I AAT represent a very interesting example of convergent evolution. Several features of the active sites of AAT and DAAT are similar, even though, for some aspects, they are mirror images and this evidently determines the opposite stereospecificity of the two enzymes.<sup>177</sup> In DAAT, K145 forms the internal aldimine and, analogously to K258 of AAT, is positioned so as to act in the 1,3 prototropic shift. A glutamate residue<sup>181</sup> stabilizes the protonated state of the pyridine nitrogen. Y31 binds to O3' of PLP, as Y225 does in AAT. However, K145 lies on the *re* face of PLP in DAAT, which, opposite to AAT, exposes the *si* face of the cofactor to the solvent and directs the *re* face to the protein. The position of R98\* in DAAT corresponds to that of R386 in AAT, being nearest to the O3' of PLP on the side opposite to the phosphate group. As a consequence, in the external aldimine, the C $\alpha$ -H bond to be cleaved is perpendicular to the *re* face of the cofactor.



**Figure 23** The reaction catalyzed by GSAM.

Glutamate-1-semialdehyde-1,2-mutase (GSAM) catalyzes the isomerization of L-glutamate-1-semialdehyde (GSA) into 5-aminolevulinic acid (ALA), the universal precursor of tetrapyrroles such as chlorophyll, heme, and coenzyme B<sub>12</sub> (Figure 23). This enzyme is only present in plants, algae, and most bacteria, which obtain GSA from the reduction of glutamyl-tRNA<sup>Glu</sup>, catalyzed by glutamyl-tRNA<sup>Glu</sup> reductase.<sup>178</sup> In animals, fungi, and some bacteria, ALA is synthesized directly through the condensation of glycine and succinyl-CoA, a reaction catalyzed by a member of the  $\alpha$ -oxoamine synthases family, ALA synthase.<sup>179</sup>

The three-dimensional structure of GSAM, which has been solved from three different sources,<sup>180–182</sup> and the mechanism of the reaction it catalyzes strictly resemble those of the aminotransferases.<sup>183</sup> However, GSAM catalyzes an intramolecular exchange of the carbonyl and amino groups, which are both present on the substrate GSA. The catalytic cycle begins with GSA and the cofactor as PMP and is completed without the involvement of a second amino acid substrate (Figure 24). A key event in the reaction mechanism is the isomerization between different forms of the 4,5-diaminolevulinic acid (DAVA) aldimine, which can dissociate from the enzyme, leaving the latter in the free PLP form and consequently slowing down the reaction. It is not clear whether this detrimental dissociation, which is clearly detectable during the catalyzed reaction *in vitro*,<sup>184</sup> is an essential feature of the catalytic mechanism and if it actually takes place *in vivo*. However, a conformationally mobile loop located at the active-site entrance of the enzyme appears to act as a lid which limits DAVA dissociation.<sup>185</sup>

Interestingly, Moser *et al.*,<sup>186</sup> upon their resolution of glutamyl-tRNA<sup>Glu</sup> reductase crystal structure, proposed a model whereby the large void of the V-shaped enzyme may be occupied by GSAM. According to this hypothesis, GSA, which is an unstable compound at neutral pH, would be produced at the active site of the reductase and directly channeled to GSAM, allowing for the efficient synthesis of ALA. A recent crystallographic study on GSAM showed that the mobile loop may actually undergo conformational changes linked to the catalytic turnover, controlling accessibility to the active site. The same study proposed a cross-talk mechanism between the two subunits of GSAM that would be involved in substrate channeling and catalysis.<sup>187</sup>

## 7.10.3 Reactivity of Amino Acids at $\beta$ -Position

### 7.10.3.1 $\beta$ -Elimination Reactions

$\beta$ -elimination from the C $\alpha$  of the PLP external aldimine releases the  $\beta$ -substituent of the amino acid and produces the aminoacrylate intermediate, which undergoes hydrolysis to form pyruvate and ammonia. Many enzymes, such as tyrosine phenol-lyase (TPL), tryptophan indole-lyase (Trpase), cystalysin, and cysteine desulfurases catalyze  $\beta$ -elimination as their main reaction, but a number of PLP-enzymes belonging to the  $\beta$ - and  $\gamma$ -family perform it as a side reaction.

#### 7.10.3.1.1 Tyrosine phenol-lyase

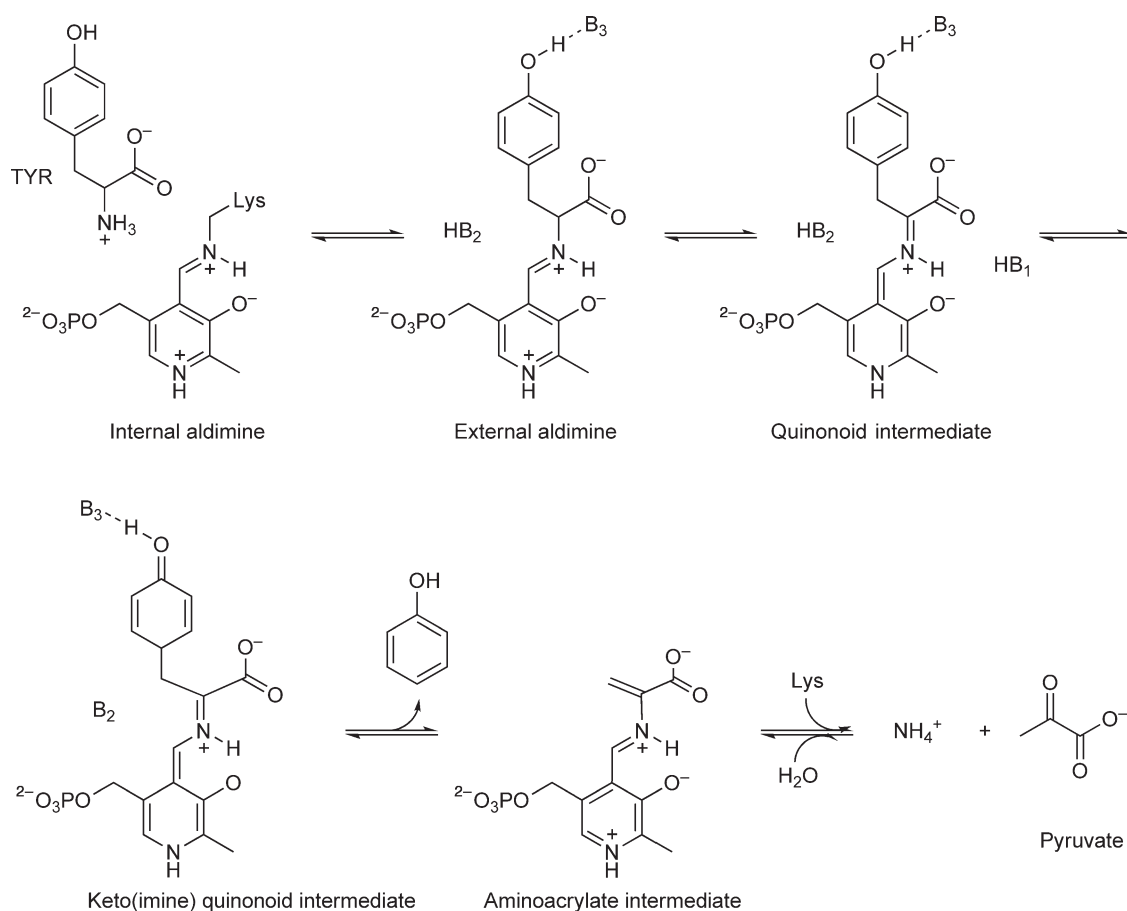
TPL catalyzes the reversible hydrolytic cleavage of L-tyrosine to phenol and ammonium pyruvate.<sup>188</sup> This enzyme resembles Trpase in sequence and structure but shows strict specificity for its physiological substrates, L-tyrosine. TPL is found primarily in enterobacteria,<sup>189</sup> although it occurs in some arthropods.<sup>190,191</sup> It consists of four chemically identical subunits and binds four molecules of PLP per tetramer.<sup>192,193</sup> Subunits are arranged in two catalytic dimers. The coenzyme-binding residue is Lys257 (*Citrobacter freundii* numbering) located at the interface between the large and small domains of the catalytic dimer. One MVC site is located at the interface between subunits adjacent to the active site, but the cation is not in direct contact with either substrate or





PLP.<sup>191</sup> Glu69 forms part of the MVC-binding site and is responsible for the transmission of activation from the MVC site to the active site. The change in the structure of Glu69 is transmitted to the cofactor site through Tyr71, which is hydrogen bonded to the PLP.<sup>194,195</sup> Binding of MVC affects a conformational equilibrium between low and high activity forms, with  $K^+$ ,  $NH_4^+$ , or  $Rb^+$  favoring the high activity form.  $Na^+$  exhibits an inhibition effect.<sup>193,196,197</sup> Most of the residues in the cofactor binding site and the substrate binding pocket are strictly conserved in all known sequences of TPL and resemble those found in Trpase.

The proposed catalytic mechanisms of TPL and Trpase are very similar. The reactions involve the elimination of formally unactivated carbon leaving groups and imply that the enzymes have to activate the aromatic ring.<sup>198</sup> The PLP is bound to the resting enzyme as a Schiff base (Figure 25). L-Tyrosine may bind with protonated or unprotonated amine<sup>199</sup> to form an external aldimine, with release of Lys257. MVCs do not significantly affect the equilibrium for external aldimine formation. The external aldimine is subsequently deprotonated at the  $\alpha$ -carbon most probably by the  $\epsilon$ -amino group of Lys257 ( $B_1$  in the figure) to form the quinonoid intermediate.<sup>200</sup> Asn185 participates in the stabilization of the keto quinonoid intermediate.<sup>201</sup> Formation of the cyclohexadienone tautomer of the phenol requires both general acid and general base catalysis. Tyr71 acts as a general acid ( $HB_2$ ) and transfers a proton to the C-1 position of the aromatic ring, possibly mediated through a water molecule.<sup>195</sup> Arg381 is in a position to function as the second catalytic base ( $B_3$ ), which interacts with the phenolic OH of the tyrosine and, by abstraction of the proton, facilitates the tautomerization of the phenol to the cyclohexadienone intermediate and subsequent elimination of phenol. Alternatively, this group may assist protonation at C-1 by hydrogen bonding to the phenol.<sup>194,202</sup> The rates of quinonoid intermediate formation and phenol elimination are accelerated by MVC binding. Phenol elimination



**Figure 25** Mechanism of the  $\beta$ -elimination catalyzed by tyrosine phenol-lyase.

gives the  $\alpha$ -aminoacrylate intermediate, which releases iminopyruvate that, ultimately, undergoes hydrolysis to yield ammonium pyruvate. The reaction specificity of TPL and Trpase is controlled by supplementary basic groups in the active site acting during the  $\beta$ -elimination step, rather than at the stage of substrate binding.<sup>202</sup> Active-site residues that control the substrate specificity are Thr124, which is critical to properly orient the phenol ring for tautomerization, and Phe448, which is involved in a conformational change from open to closed forms.<sup>203,204</sup>

TPL also catalyzes the reverse reaction and may perform other  $\alpha,\beta$ -elimination and  $\beta$ -replacement reactions. In addition to L-tyrosine, L- and D-serine, S-methyl-L-cysteine,<sup>188</sup> S-ethyl-L-cysteine, S-(*o*-nitrophenyl)-L-cysteine, O-benzoyl-L-serine,<sup>205</sup> and  $\beta$ -chloro-L-alanine also act as substrates for  $\beta$ -elimination. TPL also catalyzes racemization of alanine.<sup>188,206</sup>

### 7.10.3.1.2 Tryptophan indole-lyase

Trpase is a bacterial PLP-dependent lyase that catalyzes *in vivo* the degradation of L-tryptophan to yield indole and ammonium pyruvate. The enzyme can also synthesize L-tryptophan in conditions of excess pyruvate, ammonia, and a moderate supply of indole. Despite the extreme specificity *in vivo* for its physiological substrates, the enzyme also catalyzes  $\alpha,\beta$ -elimination and  $\beta$ -replacement reactions *in vitro* on several  $\beta$ -substituted L-amino acid residues,<sup>207,208</sup> including S-(*o*-nitrophenyl)-L-cysteine,<sup>205,209</sup> S-alkyl-L-cysteines,<sup>188,210</sup>  $\beta$ -chloro-L-alanine,<sup>195,210</sup> L-serine,<sup>188</sup> and O-acyl-L-serine.<sup>205</sup> Trpase is very similar in sequence and structure to TPL. The active form of Trpase is a tetramer and each subunit binds one molecule of PLP.<sup>197</sup> The active site is located near the subunit interface in the catalytic dimer. A catalytic cleft is formed by the large and small domains of one subunit and the large domain of the neighboring subunit. PLP is located at the bottom of the catalytic cleft and is tightly bound to the protein through several specific interactions, mainly with residues from the large domain. Most of the residues in the cofactor binding site and the substrate binding pocket are strictly conserved between Trpase and TPL. The PLP binding residue is Lys266 (*P. vulgaris* numbering). A large conformational change with closure of the Trpase subunit accompanies binding of PLP to the apo form.<sup>208</sup> Both PLP and MVC contribute to the stability of the catalytic dimer. The MVC binding site is located on the interface of the subunits. The large distance between the PLP moiety and MVC binding site excludes the possibility of direct involvement in catalysis. The structure of the binding site is very similar to TPL. In Trpase, NH<sub>4</sub><sup>+</sup>, K<sup>+</sup>, or Rb<sup>+</sup> are required for activity, since they increase the intersubunit cohesion, the affinity of the apo enzymes for PLP and of holo enzymes for substrates, affect the switch between different conformations and tautomeric species of the holo enzyme,<sup>208,211–213</sup> while Na<sup>+</sup> inhibits activity.<sup>197,214</sup>

The proposed mechanisms of Trpase is very similar to TPL. The catalytic mechanism includes as principal stages: (1) formation of the external aldimine upon reaction with tryptophan; (2) abstraction of the  $\alpha$ -proton of the substrate by the  $\epsilon$ -amino group of the PLP-binding Lys,<sup>200</sup> leading to formation of the intermediate quinonoid structure; (3) removal of the aromatic side group as a result of the concerted process of protonation of the indole fragment and cleavage of the C–C bond, resulting in the formation of the  $\alpha$ -aminoacrylate intermediate.<sup>21</sup> TRPase mechanism is an example of the rare SE<sub>2</sub> mechanism, instead of the typical Friedel–Crafts electrophilic aromatic substitution type.<sup>198</sup> The general acid catalyst in the elimination of leaving groups from quinonoid intermediates is probably Tyr72,<sup>215</sup> which also plays a role in cofactor binding in the absence of substrate. Asp133 and His458 (*P. vulgaris* numbering) are important in reaction specificity and substrate binding. His463 (*E. coli* numbering) may also act as a general base, either directly or through a hydrogen-bonded network with other residues, and/or active-site water.<sup>215,216</sup>

### 7.10.3.1.3 Cystalsin

Cystalsin belongs to the group of PLP-dependent C <sub>$\beta$</sub> -S <sub>$\gamma$</sub>  lyases and catalyzes the  $\alpha,\beta$ -elimination of L-cysteine to pyruvate, ammonia, and H<sub>2</sub>S. Cystalsin is isolated from the oral pathogen *Treponema denticola* and is the putative virulence factor responsible for periodontitis. The enzyme exhibits hemolytic activity and can modify hemoglobin.<sup>217</sup> The tertiary fold and the dimeric arrangement are similar to those of aminotransferases. A coenzyme-induced conformational change produces a holo form with less exposed hydrophobic surface than the apo form.<sup>218</sup> The PLP cofactor in the active site forms a Schiff base with the  $\epsilon$ -amino group of Lys238.<sup>219</sup>

A catalytic mechanism for the  $\alpha,\beta$ -elimination has been proposed.<sup>220,221</sup> Upon binding of the substrate cysteine, the Michaelis complex is rapidly converted into the external aldimine, followed by the  $\alpha$ -proton

abstraction to yield an  $\alpha$ -carbanion stabilized as the quinonoid intermediate. Subsequent elimination of  $\text{H}_2\text{S}$  generates the aminoacrylate. Protonation of the aminoacrylate and reverse transaldimination then form iminopropionate and regenerate the enzyme-bound PLP. Hydrolysis of iminopropionate to pyruvate and ammonia presumably occurs outside the active site.<sup>219</sup> Lys238 is proposed to act as the general base, which abstracts the  $\text{C}\alpha$  proton from the substrate and as the general acid that protonates the  $\beta$ -leaving group.<sup>222</sup> Tyr64 positions the Lys  $\epsilon$ -amino group in a proper orientation for catalysis.<sup>223</sup>

Cystalytin has a relatively broad substrate specificity. In addition to L-cysteine, cystalytin accommodates  $\beta$ -chloroalanine, cysteine methyl ester, *S*-ethylcysteine, and *S*-methylcysteine as substrates for  $\beta$ -elimination reactions.<sup>224</sup> In addition cystalytin exhibits a high catalytic versatility. It is a cyst(e)ine C–S lyase, catalyzes AlaR reaction,<sup>221</sup> the  $\beta$ -desulfination of L-cysteine sulfinic acid and the  $\beta$ -decarboxylation of L-aspartate and oxalacetate.<sup>225</sup>  $\alpha,\beta$ -Elimination and racemization probably share the step leading to the quinonoid intermediate, and the same catalytic residues.<sup>223</sup>

#### 7.10.3.1.4 Cysteine desulfurases

Cysteine desulfurases decompose L-cysteine through a  $\beta$ -elimination reaction to L-alanine and sulfane sulfur via the formation of an enzyme-bound persulfide intermediate. The persulfide sulfur is subsequently incorporated into the biosynthetic pathways of sulfur-containing biofactors. Thus, these enzymes play an important role in the biosynthesis of Fe–S clusters, thiamine, thionucleosides in tRNA, biotin, lipoic acid, molybdopterin, and NAD. They are also proposed to be involved in cellular iron homeostasis and in the biosynthesis of selenoproteins.<sup>226</sup> Cys desulfhydrase are widely distributed among prokaryotes<sup>227</sup> and the mitochondria of eukaryotes.<sup>228</sup> Cys desulfhydrase-related sequences are classified into two different groups based on sequence similarity.<sup>229</sup> NifS and IscS are members of group I, which is the more homogeneous group. Group II proteins have far more divergent sequences than group I,<sup>230</sup> differ from group I in four sequence regions<sup>229</sup> and display a greater biochemical versatility. In *E. coli*, two group II proteins exist, which were originally named CsdB<sup>231</sup> or SufS<sup>232</sup> and cysteine sulfinate desulfinate (CSD).<sup>229</sup> NifS and IscS seem to contribute to minor pathways of FeS cluster biosynthesis, most probably by mobilizing sulfur from L-cysteine as substrate, while Suf proteins are thought to play important roles in sulfur metabolism during stress. CSD is implicated in selenocysteine metabolism and selenophosphate synthesis, forming L-alanine and selenium and shows broad substrate specificity for a number of cysteine analogues.<sup>226,229,233</sup>

A PLP cofactor attached to the lysine in an His-Lys-X-X-X-Pro-X-Gly-X-Gly motif is a crucial feature of these proteins. Additionally crucial is a conserved cysteinyl residue, which serves as the persulfide site. These proteins belong to fold-type I of PLP-dependent enzymes and are homodimers. Each monomer is subdivided into a large domain with one molecule of PLP in aldimine linkage with a Lys residue and a small domain, where the critical cysteinyl residue is located in the middle of a loop. An extended lobe in CDS contains the conserved Cys and constitutes one side of the entrance to the active site. This lobe in CSD may be responsible for the ability of the enzyme to discriminate between selenocysteine and cysteine.<sup>226,233,234</sup> NifS binds and transforms the cysteine substrate in a manner usual for PLP-containing enzymes up to the stage of the central quinonoid intermediate. Cysteine desulfuration is initiated by the formation of a Schiff base between cysteine and PLP, followed by the abstraction of sulfur from the substrate and formation of an enzyme-bound cysteine persulfide and alanine via a ketimine intermediate. The cysteine residue acts as a nucleophile and attacks the sulfhydryl group of the ketimine intermediate.<sup>226,235,236</sup> Cysteine desulfuration in the SufS protein proceeds by a mechanism analogous to that of NifS/IscS group I, but the reaction is slower.<sup>237</sup>

Cys desulfurases were also reported to catalyze the decomposition of selenocysteine to L-alanine and elemental selenium with varying efficiency. The catalytic mechanism of both cysteine desulfuration and selenocysteine deselenation is similar. The PLP-binding Lys is the base that accepts the  $\text{C}\alpha$  proton of the substrate and reprotonates the intermediate to form a ketimine species. The selenohydril group of L-selenocysteine is probably deprotonated and present in an anionic form. The deprotonation of the selenohydril group may be facilitated by a His residue, located in the active site. The Cys residue is not essential for deselenation process.<sup>237</sup> Selenium is, then, released spontaneously from the ketimine intermediate.<sup>226,231,238,239</sup>

Selenocysteine lyase (SCL) from higher eukaryotes is the mammalian counterpart to bacterial enzymes and is thought to play a vital role in the metabolism of selenium in the cell.<sup>231,240</sup> SCL has been shown to be similar in amino acid composition to the NifS family, thus suggesting a similar mechanism used to generate elemental

sulfur or selenium intermediates between bacteria and multicellular eukaryotes. SCL is highly specific for L-selenocysteine and its activity toward L-cysteine is negligibly small.<sup>238</sup>

In the cyanobacterium *Synechocystis*, C-DES, a L-cysteine/cystine C-S-lyase with limited sequence homology to the NifS family, exploits a  $\beta$ -elimination reaction on cystine to catalyze the formation of hydrogen sulfide, free cysteine persulfide, pyruvate, and ammonia. The C-DES-catalyzed formation of cysteine persulfide represents an alternative pathway of S<sub>0</sub> production to NifS proteins. The acceptor of the activated sulfur during FeS cluster synthesis and the mode of sulfur transfer into the FeS cluster await identification.<sup>241,242</sup>

C-DES is a homodimeric protein and each monomer of C-DES comprises a small domain and a large domain containing the cofactor binding site. The active site of C-DES is located near the center of the monomer-monomer interface. Contrary to other PLP enzymes, the active-site crevice is composed equally of residues from both subunits in the dimer. PLP is bound covalently to Lys223. The substrate-binding pocket of C-DES is a highly polar canyon that accommodates amino acid substrates of various lengths. Two conserved arginines are involved in substrate binding through hydrogen bonding to cystine carboxylates and are a hallmark for the preferential binding of bipolar substrates.<sup>241,242</sup> Their repositioning, in association with Trp251, reorients cystine into a bent conformation necessary for C-S bond cleavage and forms a hydrophobic pocket that stabilizes the cysteine persulfide product.<sup>243</sup> C-DES lacks the catalytic cysteinyl residue of NifS-like proteins. The enzyme displays a strong substrate preference for L-cystine over L-cysteine.<sup>241,244</sup> In the proposed catalytic mechanism, the internal aldimine reacts with cystine to form the external aldimine via a gem-diamine.<sup>242</sup> The most probable candidate to deprotonate the substrate amino group for nucleophilic attack of the internal aldimine is His114. In the presence of cystine, external aldimine does not accumulate and rapidly decays to a transient quinonoid species and stable tautomers of the  $\alpha$ -aminoacrylate. In the presence of cysteine, the transient formation of a dipolar species precedes the selective and stable accumulation of the external aldimine, without formation of the  $\alpha$ -aminoacrylate Schiff base under reducing conditions. Specificity and reactivity are modulated by the protein matrix. Tight substrate selection prevents the desulfuration of cysteine and the accumulation of the toxic molecule hydrogen sulfide.<sup>245</sup> The  $\epsilon$ -amino group of Lys223 is thought to catalyze the C $\alpha$  proton abstraction, the protonation of the leaving group, and the reverse transaldimination. One Cys residue, Cys112, distinct from the catalytic cysteine of NifS, was proposed to accept a sulfur from the product cysteine persulfide.<sup>242</sup> Cys112 lies in a pocket connected to the active site through a narrow tunnel that allows the diffusion of molecules not bigger than a sulfide. Transfer of sulfur to a secondary site might avoid steric interference with an incoming cystine substrate molecule and favor the flow of sulfur to acceptor proteins that may bind at the pocket. Moreover, the labile cysteine persulfide would be protected until the acceptor protein docks at the exit tunnel.<sup>246</sup>

#### **7.10.3.1.5 Cystathionine $\beta$ -lyase**

Cystathionine  $\beta$ -lyase (CBL) catalyzes the penultimate step in microbial and plant methionine biosynthesis, that is, the cleavage of L-cystathionine to L-homocysteine, pyruvate, and ammonia. CBL belongs to the fold-type I of PLP-dependent enzyme and shares extensive homology with cystathionine  $\gamma$ -synthase (CGS).<sup>247</sup> Both bacterial and plant CBLs are tetramers composed of four identical subunits. Each active site contains one cofactor molecule bound within a cleft formed between the dimer interfaces. Residues from both monomers participate in substrate binding and catalysis. A Schiff base is formed between the cofactor and the holoenzyme at the  $\epsilon$ -amino group of a Lys.<sup>247-250</sup> The substrate-binding pocket of CBL from *E. coli* is smaller than that from *A. thaliana*, and seems to be optimized for cystathionine, whereas in *A. thaliana* it can accommodate longer substrate with equal affinity. Substrate binding to the *E. coli* enzyme seems to occur with higher affinity, due to stronger interaction at the  $\alpha$ -carboxyl site. A relatively broad substrate specificity toward sulfur-containing amino acids was observed for CBL from spinach, *E. coli*, and *S. typhimurium* enzymes. For all these enzymes, the highest rates of  $\beta$ -cleavage activity were obtained with cystathionine and djenkolate, whereas cystine and lanthionine proved to be poor substrates. In plant CBL, the nature of the sulfur-containing bridge (thioether, thioacetal, or disulfide) and, therefore, the spatial separation of the two carboxy groups of the substrate influences the rate of the  $\alpha,\beta$ -elimination process and the efficiency of the cleavage.<sup>247-250</sup>

The substrate cystathionine undergoes transaldimination with the enzyme-bound cofactor PLP. Many interactions properly position the substrate in a productive binding mode. Two conserved Arg residues anchor the carboxylate groups of the cystathionine substrate and act as general  $\alpha$ -carboxylate docking site, whereas the

$\alpha$ -amino group interacts with the ionized PLP hydroxyl group. Ser339 (*E. coli* numbering) interacts with the sulfur atom and breaks the pseudosymmetry of the cystathionine.<sup>247</sup> Tyr181 (*A. thaliana* numbering), a conserved residue in all cystathionine-utilizing enzymes, abstracts a proton from the charged  $\alpha$ -amino group of the bound substrate. After formation of the external aldimine, the PLP-binding Lys abstracts the proton from  $C\alpha$ , which produces the quinonoid intermediate and transfers the proton directly from  $C\alpha$  to  $S\gamma$  of the substrate, thereby inducing  $C\beta$ -S bond cleavage. Subsequent elimination of homocysteine generates the PLP derivative of aminoacrylate. Aminoacrylate is converted into iminopropionate by proton addition on  $C\beta$ . This proton may originate from the first transaldimination step and be transferred to aminoacrylate via the Tyr residue. Hydrolysis of iminopropionate to pyruvate and ammonia presumably occurs outside the active site of CBL.<sup>247,250</sup>

The enzyme has been postulated by Clausen *et al.*,<sup>247</sup> to exist in an open conformation and undergo a conformational change to a closed enzyme form upon substrate binding. However, large domain movements may not be required during the course of the reaction, since inhibitor-bound and -unbound CBL structures do not significantly differ in the  $C\alpha$  positions.<sup>251</sup>

#### 7.10.3.1.6 MalY

MalY belongs to the group of PLP-dependent  $C\beta$ - $S\gamma$  lyases, which catalyzes the elimination of a  $C\beta$ -substituent to generate the aminoacrylate-PLP Schiff base, with formation of ammonium ion and pyruvate. In *E. coli*, it is a maltose regulon modulator. MalY utilizes amino acids containing a  $C\beta$ -S linkage as substrates, producing a sulfur-containing molecule. It seems to prefer cystine over cystathionine as the main substrate.<sup>252</sup> MalY shows homology to transsulfuration enzymes of the  $\gamma$ -family of PLP and exhibits the same  $C\beta$ - $S\gamma$  lysis activity as CBL, including the same pH optimum, substrate preference and kinetic constants. MalY is a dimeric protein and each subunit of the dimer is composed of a large PLP-binding domain and a small domain similar to aminotransferases. The open-closed transition common to many PLP enzymes seems not relevant to MalY. Unique to the active site of the MalY is the binding mode of the PLP-phosphate group, which interacts primarily with five water molecules. Only one protein residue anchors the cofactor. This, in addition to the absence of the network of interactions stabilizing the pyridine positive charge, is indicative of weak binding of the cofactor and of a larger, more flexible, and highly accessible substrate-binding pocket. At present, the physiological substrate of MalY is unknown. This unique phosphate-binding mode may indicate interaction with a sterically demanding substrate, such as an amino acid-sugar conjugate with a sulfur atom in the  $\beta$ -position.

The cofactor is covalently attached by an aldimine linkage to Lys233, which is part of the conserved PLP-binding site. Reaction with the substrate leads to the formation of the external aldimine, followed by abstraction of the substrate  $C\alpha$  proton by the released deprotonated amino group of Lys233. The resulting carbanionic intermediate is stabilized by the electron-withdrawing capacity of the pyridine ring, which is mediated by Asp201 and Tyr121. Finally, Ser36 guides the amino group of Lys233 to the substrate  $S\gamma$ . After protonation of  $S\gamma$  and release of product homocysteine (if cystathionine is the substrate) or cysteine persulfide (cystine as substrate), reverse transaldimination leads to iminopropionate, which is hydrolyzed to pyruvate and ammonia outside the active site.

MalY acts as a repressor of the maltose regulon. The PLP domain contains unique regions responsible for protein interaction and recognition of major importance for its repressor function.<sup>253</sup>

#### 7.10.3.2 $\beta$ -Replacement Reactions

In  $\beta$ -replacement reactions, the  $\beta$ -substituent of an amino acid substrate is replaced by a new  $\beta$ -substituent. For the three enzymes (TRPS, OASS, and cystathionine  $\beta$ -synthase (CBS)) whose mechanisms are discussed in this section, the catalytic reaction is composed of two distinct half-reactions. The  $\beta$ -elimination is followed by a  $\beta$ -addition where nucleophilic agents, indole, sulfide, and homocysteine, respectively, react with the  $\alpha$ -aminoacrylate Schiff base to form the final product, L-tryptophan, L-cysteine, and L-cystathionine, respectively.



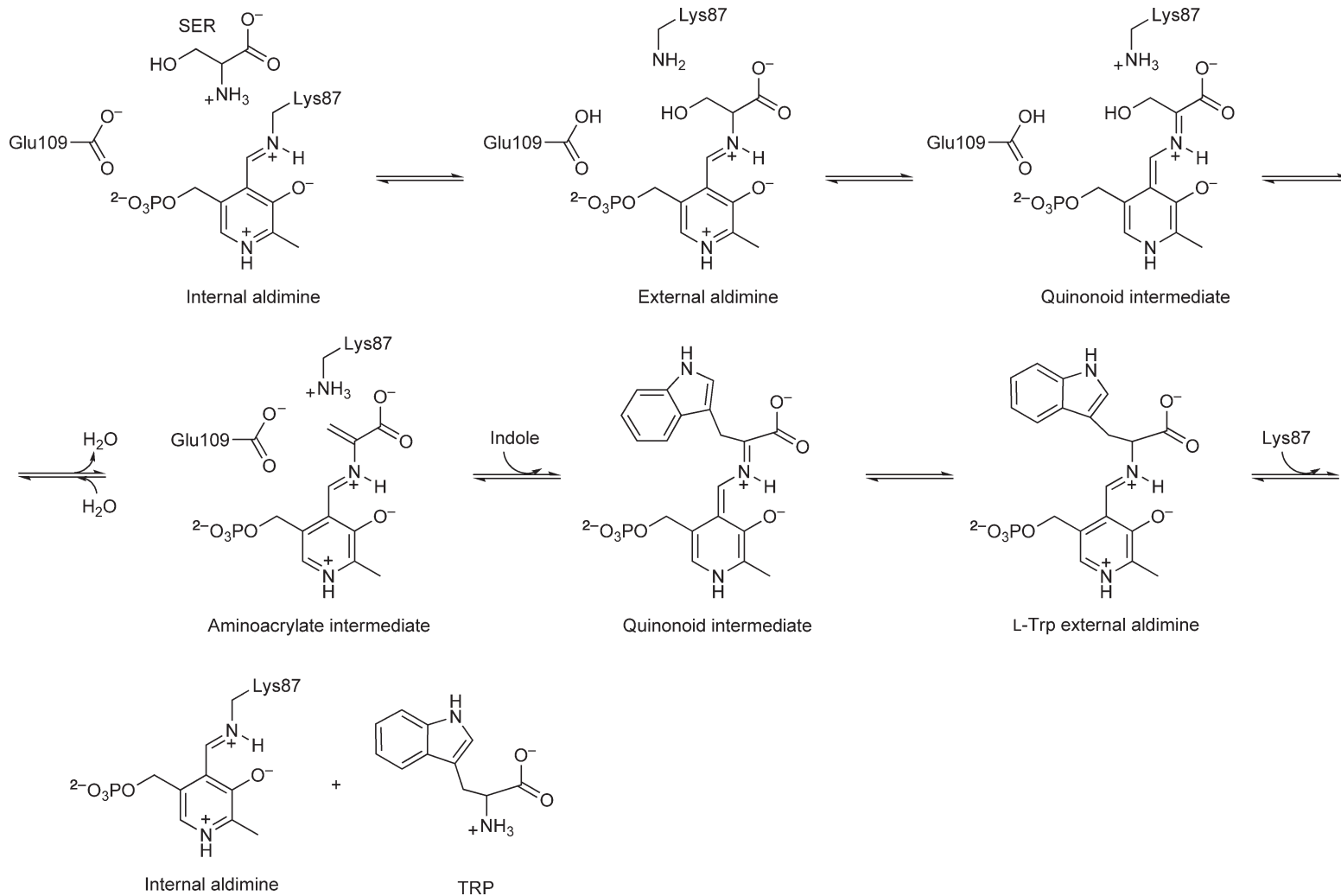
### 7.10.3.2.1 Tryptophan synthase

TRPS are bifunctional tetrameric enzymes that catalyzes the last two steps in the biosynthesis of L-tryptophan in bacteria, plants, and fungi, that is, the conversion of indole 3-glycerol phosphate and serine to tryptophan and water.<sup>23–25,26–28</sup> In bacteria, the two reactions are catalyzed by separate  $\alpha$  and  $\beta$ -subunits, which combine to form a stable  $(\alpha\beta)_2$ -heterotetramer. The  $\alpha_2\beta_2$  complex is nearly linear; the two  $\beta$ -subunits are tightly associated in the center of the molecule and the two  $\alpha$ -subunits are situated on opposite sides of the  $\beta$ -dimer.<sup>254</sup> Subunit association increases the substrate affinity and the activity of the complex. The formation of the complex is accompanied by subunit conformational changes and the subunit interface mediates the activation of the  $\alpha$ - and  $\beta$ -activity and the communication between the active sites. Mutual activation is brought about by stabilization of the respective active conformational states of each component. In yeast and molds, the two reactions are catalyzed by separate  $\alpha$ - and  $\beta$ -sites in a single bifunctional polypeptide chain.<sup>255</sup> TRPS catalyzes the cleavage of indole-3-glycerol phosphate to indole and glyceraldehyde-3-phosphate at the  $\alpha$ -site. Indole is then transferred through an intermolecular tunnel to the  $\beta$ -active site, where it is condensed with serine in a PLP-assisted  $\beta$ -replacement reaction.<sup>23,24</sup> The tunnel, that extends from the active site of the  $\alpha$ -subunit to the binding site for PLP, allows indole to diffuse between the two active sites and prevents its escape to the solvent.<sup>256</sup> The tunnel is opened by the movement of two aromatic residues that act as a molecular gate.<sup>257–259</sup> PLP is located at the interface between the two domains of the  $\beta$ -subunit and forms a Schiff base bond with the  $\epsilon$ -amino group of  $\beta$ Lys87 (*S. typhimurium* numbering). L-Serine binds to the internal aldimine forming a gem-diamine complex that rapidly converts into the external aldimine (Figure 26). In the  $\alpha_2\beta_2$ , this intermediate rapidly decays by  $\alpha$ -proton abstraction yielding an unstable quinonoid that eliminates a hydroxide ion to give  $\alpha$ -aminoacrylate intermediates. The release of the hydroxyl moiety on the  $\beta$ -carbon completes the first catalytic stage, the  $\beta$ -elimination reaction. The equilibrium distribution between the external aldimine and aminoacrylate is shifted in favor of the former species at low temperature, high pH, and in the presence of  $\text{Na}^+$  and  $\text{K}^+$ , whereas  $\text{Li}^+$ ,  $\text{Cs}^+$ , and  $\text{NH}_4^+$  favor the latter species.<sup>258,260–264</sup> In the second stage of the catalytic reaction, indole reacts with the  $\alpha$ -aminoacrylate double bond forming a quinonoid species. By accepting a proton, the quinonoid species interconverts to the external aldimine of L-tryptophan.<sup>23</sup> The release of L-tryptophan completes the catalytic cycle, regenerating the internal aldimine. An ordered sequential mechanism is generally favored for the  $\beta$ -subunit with L-serine binding first. This mechanism is supported by the requirement that L-serine binds before indole, since L-serine forms a specific binary complex with the active site of the  $\beta$ -protomer, whereas indole does not.  $\beta$ Lys87 has been proposed as the active-site base involved in abstraction of the  $\alpha$ -proton of the external aldimine and in the protonation of the tryptophan quinonoid intermediate prior to release of L-tryptophan from its external aldimine.  $\beta$ Glu109 is thought to accept a proton from the  $\alpha$ -amine of L-serine and donate it to the leaving hydroxide. The carboxylate of the  $\beta$ Glu109 is also thought to activate the indole toward nucleophilic attack on the aminoacrylate.<sup>265–269</sup> The overall stereochemistry of the  $\beta$ -replacement reaction was found to proceed with retention of configuration at the  $\beta$ -carbon of the amino acid substrate.<sup>270</sup> The stereochemistry of the elimination/addition reactions catalyzed by TRPS in the formation of tryptophan from serine and indole was proposed to be synchronous with both the  $\alpha$ -proton and the  $\beta$ -substituent from the same faces to generate the  $\alpha$ -aminoacrylate intermediate.<sup>271,272</sup>

MVCs are important effectors of TRPS. The MVC binding site is located in the  $\beta$ -subunit separate from the catalytic site.<sup>254,258</sup> Sodium ions strongly favor the accumulation of the external aldimine species by stabilizing a partially open conformation of the enzyme, whereas cesium ions stabilize the  $\alpha$ -aminoacrylate and the closed state.<sup>261,264</sup> MVCs bound to the  $\beta$ -subunit also proved to alter the conformation of the side chains of the residues acting as molecular gate.<sup>258</sup>

Substrate channeling in bienzyme complex is regulated by allosteric interactions transmitted between the  $\alpha$ - and  $\beta$ -sites that synchronize the  $\alpha$ - and  $\beta$ -reactions.<sup>273,274</sup> These allosteric signals are triggered by covalent bonding changes in the reacting substrate at the  $\beta$ -site and ligand-binding interactions at the  $\alpha$ -site. Binding of  $\alpha$ -site ligands modulates the affinity of  $\beta$ -subunit for serine and the distribution of the catalytic intermediates formed at the  $\beta$ -active site,<sup>23,24</sup> while the formation of the highly reactive  $\alpha$ -aminoacrylate species leads to a dramatic enhancement in the rate of IGP cleavage at the  $\alpha$ -site and indole passage through the tunnel. Indole, thus, is not produced and does not reach the  $\beta$ -site before a highly reactive aminoacrylate intermediate has been formed,<sup>269,275</sup> preventing the accumulation of indole at the  $\alpha$ -site or in the tunnel. Signal transmission also switches the  $\alpha$ - and  $\beta$ -sites between an open, low-activity state to a closed high-activity state.<sup>257,276–280</sup>





**Figure 26** Mechanism of the tryptophan synthase  $\beta$ -reaction. The enzyme-substrate intermediates formed in the reaction and the residues involved in catalysis are shown.

The transition between these states is predominantly associated with the rearrangement of structural elements of the  $\alpha$ - and  $\beta$ -subunit. Distinct pathways for intersubunit communication exist, which specifically involve different structural elements of  $\alpha$ -subunit. The  $\alpha$ -ligand-triggered intersubunit communication seems to rely on a single hydrogen bond formed between  $\beta$ Ser178 and  $\alpha$ Gly181. Formation of this hydrogen bond reduces loop mobility and stabilizes the closed conformation of the  $\alpha$ -subunit.<sup>278,281–284</sup> The movement of the loop is propagated to the  $\beta$ -subunit, to the  $\alpha$ -site catalytic residue to be oriented for efficient catalysis and to a second loop in the  $\alpha$ -subunit involved in the intersubunit activation that leads to a reciprocal modulation of the activity of the  $\alpha$ - and  $\beta$ -subunits within the tetramer.

Several L-serine-related amino acids are substrates for the  $\beta$ -replacement reaction, while indole analogues or derivatives of indole or a variety of alkanethiols can act as nucleophiles. TRPS also catalyzes  $\beta$ -elimination reactions with L-serine and with 6-chloro-L-alanine, in which the aminoacrylate is hydrolyzed to pyruvic acid and ammonia. Although the  $\beta$ -subunit of TRPS catalyzes both types of reaction, subunit association reduces the  $\beta$ -elimination reaction rate of the  $\alpha_2\beta_2$  complex.<sup>285,286</sup>

### 7.10.3.2.2 *O*-acetylserine sulphydrylase

OASS catalyzes the replacement of the  $\beta$ -acetoxy group of *O*-acetyl-L-serine (OAS) with sulfide to generate L-cysteine, the final step in the pathway for the *de novo* biosynthesis of L-cysteine in enteric bacteria and plants.<sup>29</sup> In bacteria two isoforms, OASS-A, and OASS-B are known, with the A-isozyme expressed under aerobic and the B-isozyme expressed under anaerobic conditions. The overall fold is very similar for the two isoforms, both being homodimers with one PLP per subunit.<sup>287,288</sup> The cofactor in the active-site cleft is in Schiff base linkage with the  $\epsilon$ -amino group of a Lys. The isozymes also share a common chemical mechanism.<sup>30</sup> The B-isozyme has a higher turnover number, a lower  $K_m$  for OAS and is less substrate selective for both the amino acid and nucleophilic substrates. The reaction is an elimination–addition reaction to give an overall  $\beta$ -substitution reaction.<sup>289</sup> The kinetic mechanism is Ping–Pong Bi–Bi. The first substrate, OAS, binds to the enzyme as the monoanion with the  $\alpha$ -amine unprotonated to carry out a nucleophilic attack on C-4' of the internal Schiff base. The geminal-diamine intermediates and successively the external Schiff base are rapidly and transiently formed. Lys41 (*S. typhimurium* numbering) acts as a general base to abstract the  $\alpha$ -proton in the  $\beta$ -elimination reaction with release of the first product, acetate, at the end of the first half reaction and formation of the  $\alpha$ -aminoacrylate intermediates. The active site partially opens to release the product and allow entry of bisulfide. Bisulfide then adds as the second substrate to the enzyme and attacks C $\beta$  of  $\alpha$ -aminoacrylate to give the cysteine external Schiff base in the second half-reaction. Protonated Lys41 acts as a general acid to donate a proton to C $\alpha$  as the product external Schiff base is formed. The free enzyme is finally regenerated through a gem-diamine intermediate, releasing L-cysteine as the final product.<sup>29,30</sup> The enzyme cycles through different conformations as it catalyzes its reaction. The internal aldimine exists in an open conformation, the external aldimine in a closed conformation and the  $\alpha$ -aminoacrylate in a conformation that differs from those of the internal and external aldimines.<sup>30,287,290–293</sup> The conformational change to close the site is triggered by binding of the substrate  $\alpha$ -carboxylate to the substrate-binding loop (residues 68–71), specifically to the side chain of Ser69 in OASS-A, which moves by more than 7 Å as the active site closes. The first half-reaction limits the overall reaction, while the second half-reaction is very fast and likely to be diffusion limited. The overall stereochemistry of the  $\beta$ -replacement reaction of OASS has been found to proceed with retention of configuration at the  $\beta$ -carbon of the substrate.<sup>294</sup> The reaction is an antielimination, with both the  $\alpha$ -proton and the  $\beta$ -substituent leaving in opposite directions. An E<sub>2</sub> or concerted elimination mechanism for the elimination step of OASS has been proposed based on the absence of quinonoid intermediate during catalysis and on the presence of a good leaving group, which does not require general acid–base chemistry during elimination. OASS is, thus, an exception, since a requirement for acid–base chemistry in the elimination reaction to protonate the leaving group and a quinonoid intermediate have been observed in most PLP-dependent enzymes catalyzing  $\beta$ -elimination reactions.<sup>295–297</sup>

In the absence of the nucleophilic substrate, the enzyme also exhibits a  $\beta$ -elimination reaction, in which transimination of the  $\alpha$ -aminoacrylate external Schiff base occurs giving free enzyme and free  $\alpha$ -aminoacrylate, which decomposes in solution to pyruvate and ammonia.

OASS-A represents a new class of PLP-dependent enzymes regulated by small anions.<sup>298</sup> An allosteric anion-binding site is located at the dimer interface. The conformation generated by chloride binding to the

allosteric site differs from either the open and the closed conformation. Upon binding of chloride the enzyme undergoes a structural reorganization, which prevents stabilization of the external Schiff base and subsequent chemistry and finally results in an inhibited conformation. Sulfide is a competitive substrate inhibitor of OASS<sup>299,300</sup> and the allosteric binding site may be the binding site for this physiological inhibitor. No allosteric binding site at the dimer interface is evident in OASS-B, although functional data indicate an inhibitory effect of chloride ions on OASS-B activity.<sup>301</sup>

OASS-A is known to form a holoenzyme complex with serine acetyltransferase (SAT), both in bacteria<sup>302,303</sup> and plants.<sup>304,305</sup> The biological function of the complex and the mechanism of reciprocal regulation of the constituent enzymes are still poorly understood, but SAT can inhibit OASS catalytic activity with a double mechanism: the stabilization of a closed conformation less accessible to the natural substrate and the competition with OAS for binding to the enzyme active site. SAT binding to OASS involves the anchoring of the SAT C-terminal decapeptide in the OASS active site.<sup>306,307</sup> OASS-B is reported not to bind *E. coli* SAT. The structure near the active site of OASS-B differs from that of OASS-A, which might be responsible for the loss in the ability of OASS-B to participate in complex formation with SAT. At present, the physiological significance of the inability of OASS-B to interact with SAT is unknown.

In *Mycobacterium tuberculosis*, three isozymes endowed with sulfhydrylase activity have recently been described: CysK1, CysK2, and CysM.<sup>308,309</sup> Whereas CysK1 is specific for *O*-acetylserine, CysM was recently found to be specific for *O*-phosphoserine<sup>310,311</sup> and to interact with a small protein, denoted CysO.<sup>308,312</sup> CysO carries sulfur in the form of a thiocarboxylated C-terminus and delivers the final product L-cysteine after the proteolytic cleavage of a covalent adduct by a metalloprotease. This finding provides evidence that in *M. tuberculosis* there is a pathway for cysteine biosynthesis that is independent of OAS and sulfide, and might be exploited for the design of specific inhibitors.

#### 7.10.3.2.3 Cystathionine $\beta$ -synthase

CBS catalyzes a  $\beta$ -replacement reaction condensing homocysteine and serine to form cystathionine. This reaction is the committed step in the synthesis of cysteine from methionine by transsulfuration. CBS is the most common locus for mutations associated with homocystinuria, an inherited metabolic disorder. In humans, accumulation of homocysteine is correlated with a number of different disease pathologies, such as cardiovascular diseases, neural tube defects and Alzheimer's disease.<sup>313–315</sup> It is closely related to other members of fold-type II, OASS and the  $\beta$ -subunit of TRPS. Human CBS is a unique heme containing PLP-dependent enzyme, binding one heme and one PLP per subunit. It is a homotetramer, with individual monomers displaying a modular organization: an N-terminal heme domain followed by a catalytic core that houses the PLP cofactor and by a C-terminal regulatory domain.<sup>316</sup> The heme in mammalian CBS displays unusual electronic properties, since it is a six-coordinate low-spin ferric heme, with histidine and cysteine as the axial ligands.<sup>317,318</sup> Perturbation of the heme environment, despite high PLP saturation, is detrimental to the catalytic activity.<sup>319</sup> Information regarding the oxidation state of the heme is transmitted to PLP at a distant site. The role of the heme in mammals is an open question. The absence of heme from the homologous yeast enzyme points to a regulatory role for this cofactor. A redox-regulatory role has been proposed.<sup>320–323</sup>

Human CBS shows complex regulation. It is allosterically activated by SAM,<sup>324</sup> which binds stoichiometrically to each subunit. The binding site is presumed to be in the C-terminal domain, since a deletion of this region prevents SAM binding and leads to a loss of SAM responsiveness.<sup>316</sup> It partially alleviates the intrasteric inhibition on the N-terminal catalytic core of the protein exerted by the C-terminal domain. Thus, the full-length enzyme exhibits a 'basal' level of activity, which can be stimulated by SAM to give the 'activated' state of the enzyme. Deletion of the inhibitory C-terminal domain results in an increase relative to the basal level and gives a 'superactivated' state of the enzyme.<sup>314</sup> The allosteric activation by SAM in addition to the presence of the heme cofactor distinguishes CBS orthologues from higher versus lower eukaryotes. CBS from yeast and *T. cruzi* is not responsive to SAM.<sup>325–327</sup> The C-terminal domain is also important for maintaining the oligomeric state of the protein. The yeast enzyme is a homo-octamer and deletion of the C-terminal regulatory domain leads to the formation of highly active dimeric enzymes.<sup>326</sup> Mutations in this domain correlate with hereditary diseases.<sup>323,328</sup>

The PLP resides as an internal aldimine forming a Schiff base with a Lys residue in the active site. The mechanism is predicted to involve the formation of PLP-bound species, in analogy with other enzymes

catalyzing  $\beta$ -replacement reactions. Addition of serine results in the formation of the external aldimine, which undergoes proton abstraction at the  $\alpha$ -carbon by the PLP-binding Lys followed by elimination of the hydroxyl group of L-serine to generate an aminoacrylate intermediate.<sup>329</sup> In the human enzyme, the aminoacrylate intermediate does not appear to accumulate to a significant extent.<sup>316,317,330</sup> No quinonoid intermediate has been detected in the yeast enzyme, suggesting that the conversion of aminoacrylate into cystathionine is the rate-limiting step in the overall reaction.<sup>330</sup> Nucleophilic attack by the thiolate of homocysteine on the aminoacrylate and reprotonation at C $\alpha$  generate the external aldimine of cystathionine. A final transaldimination reaction releases the final product, cystathionine.<sup>323</sup> The reaction is reversible but L-cystathionine formation is strongly favored *in vivo*. Similar to other related  $\beta$ -replacement enzymes, the proposed mechanism is ping-pong, in which serine binding and release of water are followed by homocysteine binding and release of cystathionine.<sup>326</sup>

The aminoacrylate is stable in yeast-truncated CBS and it does not decompose to NH<sub>3</sub> and pyruvate. Thr81, Ser82, Thr85, Gln157, and Tyr158 residues contribute to suppression of the competing  $\beta$ -elimination reaction by maintaining the correct arrangement of the active-site residues and stabilizing the closed conformation.<sup>331</sup> Although CBS successfully restricts the chemistry of the PLP cofactor to the  $\beta$ -replacement reaction, the substrate specificity of the enzyme is not tightly controlled.<sup>315,332</sup> L-cysteine, 3-chloro-L-alanine, S-hydroxyethyl-L-cysteine, L-Ser-O-sulfate, and L-allothreonine are accepted by CBS in place of L-serine, whereas  $\beta$ -mercaptoethanol and H<sub>2</sub>S can replace the nucleophilic L-homocysteine.<sup>333–335</sup>

## 7.10.4 Reactivity of Amino Acids at $\gamma$ -Position

### 7.10.4.1 $\gamma$ -Elimination Reactions

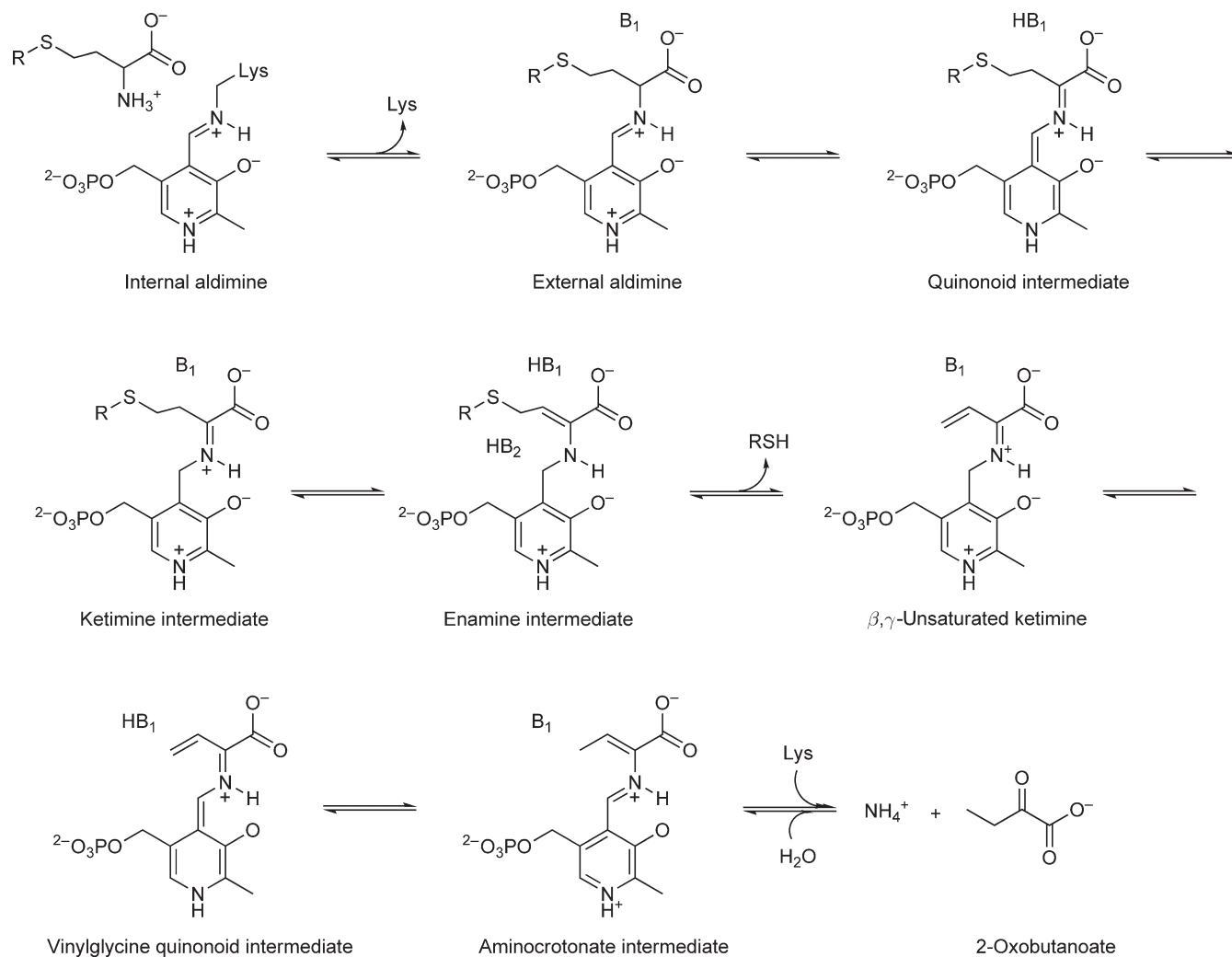
The mechanism of  $\gamma$ -elimination reaction involves abstraction of the C $\alpha$  and C $\beta$  protons and removal of the  $\gamma$ -leaving group of the substrate side chain, with formation of  $\alpha$ -ketobutyrate and ammonia. Cystathionine  $\gamma$ -lyase (CGL) and methionine  $\gamma$ -lyase (MGL) are representative enzymes of this class.

#### 7.10.4.1.1 Cystathionine $\gamma$ -lyase

CGL catalyzes the  $\gamma$ -cleavage of cystathionine to yield cysteine,  $\alpha$ -ketobutyrate, and ammonia. CGL is present in actinobacteria, fungi, and mammals.

In humans, the enzyme is linked to cystathioninuria, cystinosis, and, although less frequently than CBS, homocystinuria. These metabolic disorders potentially result in mental or physical impairments. In addition, malignant lymphoid cells show markedly reduced levels of CGL.<sup>336,337</sup> CGL shows extensive sequence homology to CBL and CGS.<sup>14</sup> CGL are homotetramers and carry one PLP cofactor per monomer covalently bound through a Schiff base to an active-site lysine.<sup>338</sup> CGLs from yeast or humans are virtually identical at their active sites with CBL and CGS from *E. coli*.<sup>339</sup> A molecular mechanism was deduced on the basis of the structural similarity to the *E. coli* CBL active site.<sup>247</sup> In the first step, substrate binds to the active-site PLP cofactor by formation of a Schiff base (Figure 27).  $\alpha$ -Proton abstraction by the active-site lysine residue (B<sub>1</sub> residue in the figure), followed by reprotonation at C-4' leads to the reversible formation of a ketimine intermediate. This species subsequently undergoes  $\beta$ -proton abstraction by the same lysine. An unidentified active-site residue (HB<sub>2</sub> residue in the figure) acts as the proton donor in the breakage of the C $\gamma$ -S bond, with release of cysteine and formation of a vinylglycine ketamine intermediate. The  $\gamma$ -methylene group of the resulting vinylglycine ketamine intermediate is reprotonated, yielding the PLP-derivative of aminocrotonate. Reaction with the active-site lysine regenerates the internal aldimine and releases the aminocrotonate intermediate. The release step is rate-limiting for the  $\gamma$ -elimination reaction of CGL from yeast.<sup>340</sup> Nonenzymatic hydrolysis of the aminocrotonate intermediate produces ammonia and 2-oxobutanoate.

Human CGL displays an interesting substrate specificity with clear preference of C-S over S-S bond breakage: L-cysteine and L-cystine are converted orders of magnitudes more slowly than the natural substrate L-cystathionine. The yeast enzyme attacks the C- $\beta$ -S bond of L-cystine or L-cysteine,<sup>338</sup> whereas the streptomyces enzyme is quite active toward L-cystine. A CGL enzyme from *Lactococcus lactis* was reported to consist of at least six identical subunits and have a broad substrate specificity and relatively low specific activity toward L-cystathionine compared to bacterial CGL.<sup>341</sup> In humans, L-cystathionine is split almost exclusively in



**Figure 27** General mechanism of the  $\gamma$ -elimination reaction. In CGL, the substrate is cystathionine and  $R = CH_2CH(NH_3^+)COO^-$ ; in MGL, the substrate is methionine and  $R = CH_3$ .

a CGL-specific manner, whereas the yeast enzyme harbors pronounced CBL activity.<sup>338</sup> CGL from yeast is reported to synthesize L-cystathionine through a  $\gamma$ -replacement reaction.<sup>342,343</sup> The reaction specificity seems context-dependent, such that the reaction catalyzed *in vivo* depends on the substrate supplied. *In vivo*, CGL catalyzes cysteine desulfhydration under physiologic conditions in smooth muscle. The  $\beta$ -cleavage of cystine by CGL yields pyruvate, ammonia, and thiocysteine that is converted into hydrogen sulfide, which functions as a smooth muscle relaxant.<sup>344,345</sup>

#### 7.10.4.1.2 Methionine $\gamma$ -lyase

MGL catalyzes the  $\alpha,\gamma$ -elimination reaction of methionine to  $\alpha$ -ketobutyrate, methanethiol, and ammonia. MGL has been isolated from a number of bacteria, including *Pseudomonas putida*, *Aeromonas* sp., *Clostridium sporogenes*, *P. taetrolens*, and *Brevibacterium linens*, from the primitive protozoa *Entamoeba histolytica* and *Trichomonas vaginalis*, but is not believed to be present in yeast, plants, or mammals.<sup>346–349</sup> Two MGL isoforms have been isolated from *T. vaginalis*<sup>350</sup> and *Entamoeba histolytica*,<sup>351</sup> which differ in substrate specificity, overall charge, and catalytic properties. They show a high degree of sequence identity to MGL from *Pseudomonas putida*. MGL has demonstrated antitumor efficacy in a number of methionine-dependent cancer cell lines.<sup>352</sup>

MGL belongs to the evolutionary  $\gamma$ -family of PLP-dependent enzymes.<sup>14</sup> The enzyme forms a homotrimer. Two monomers associate to build the active dimer. Each monomer contains one PLP molecule as a cofactor, which is covalently linked to the  $\epsilon$ -amino group of a Lys residue. The arrangement of the PLP-binding residues is almost identical in all known MGL structures.<sup>353–355</sup> The spatial fold of subunits, with three functionally distinct domains, a N-terminal domain, a large PLP binding domain, and a C-terminal domain, and their quarternary arrangement, is similar to those of CGL, CBL, and CGS from *E. coli*.

The overlay of the active sites of *T. vaginalis* MGL and yeast CGL provides structural explanations for the differential substrate preference. The conserved amino acids Phe44, Met84, Cys110, and Val331 (*E. histolytica* numbering) were present in all known MGLs and generally missing in other PLP-dependent enzymes. The phenyl group of Phe44 forms a hydrophobic contact area together with the side chains of Leu55 and Ile59 to bind the methyl group of methionine. Cys110 (Cys116 in MGL from *P. putida*) is involved in mediating substrate binding and in determining substrate specificity, but is not essential for activity of the enzyme.<sup>356</sup> The equivalent cysteine residue is also present in the two *Trichomonas* MGL. The N-terminal region, which plays a role in substrate recognition, is very flexible in MGL from *P. putida*. A Tyr114 residue is conserved in all sequences of the  $\gamma$ -family enzymes and facilitates the elimination of the substrate, but it is unclear if it serves as a general acid catalyst.<sup>357</sup> Both MGLs from *Trichomonas* and *P. putida* catalyze methionine, homocysteine, cysteine, and OAS but are inactive toward cystathionine. The *Pseudomonas* and the *Aeromonas* enzymes also use cystathionine as a substrate.<sup>347,358</sup> Only L-amino acids are active substrates of all other MGLs except for an enzyme from *Aspergillus*, which was reported to also catalyze the  $\gamma$ -elimination of D-methionine.<sup>359</sup> MGL also catalyzes  $\gamma$ -replacement of L-methionine and various thiols as well as  $\alpha,\beta$ -elimination and  $\beta$ -replacement of S-substituted L-cysteines.<sup>360,361</sup> Differences in the active sites organization between the two trichomonad enzymes are thought to have important implications for discrimination between  $\gamma$ -elimination and  $\beta$ -elimination reactions.<sup>350</sup>

#### 7.10.4.2 $\gamma$ -Replacement Reactions

Only two enzymes (threonine synthase (TS) and CGS) are known to catalyze the  $\gamma$ -replacement reaction, which is composed of two distinct half-reactions. The mechanism involves the elimination of the  $\gamma$ -leaving group, followed by a Michael addition, where water or cysteine reacts with the  $\beta,\gamma$ -unsaturated ketimine to form the final product, L-threonine and L-cystathionine, respectively. In the case of TS, the addition is on the  $\beta$ -carbon.

##### 7.10.4.2.1 Threonine synthase

TS is a fold-type II PLP-dependent enzymes, which catalyzes the ultimate step in threonine biosynthesis, the  $\gamma$ -replacement reaction of O-phosphohomoserine (OPHS) yielding threonine and inorganic phosphate. In plants, OPHS is the branching point for threonine and methionine biosynthesis and allosteric activation by



SAM, a product of methionine decomposition, coordinates the flux between both synthetic pathways. In bacteria and fungi, the activity of TS seems not regulated by any effector.<sup>362–365</sup>

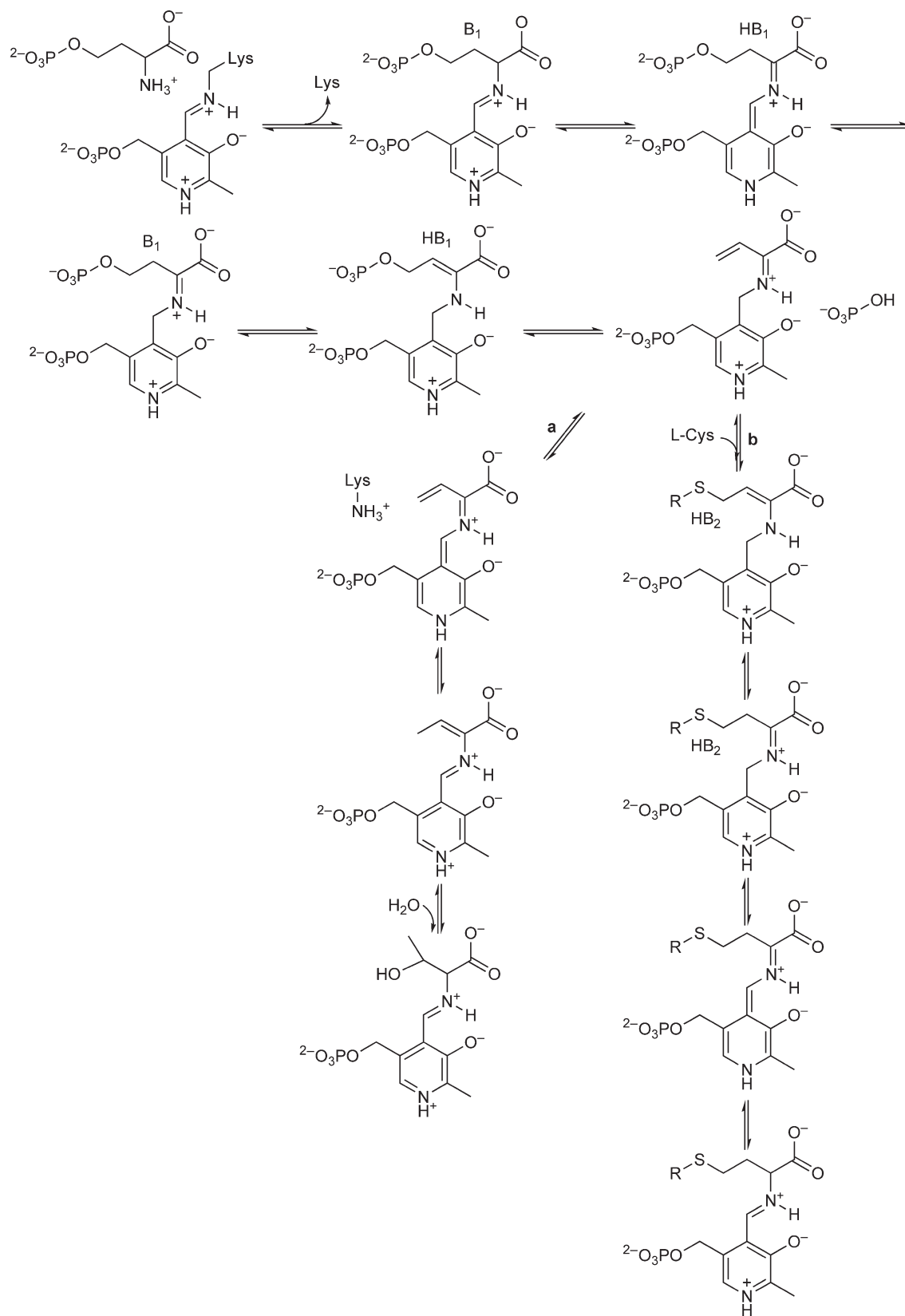
TS from higher plants is a homodimer, whereas TS from yeast and fungi functions in monomeric state.<sup>363,365</sup> Each dimer has four domains in each monomer. Domain 1 and the swap domain are specific to TS, whereas all proteins from the fold-type II share a common fold for the other two domains. The active site is located in a deep crevice at the interface of domains 2 and 3 of each monomer.<sup>366</sup> Plant TS has an additional N-terminal domain-binding SAM, whereas TSs from bacteria or fungi lack the SAM binding domain and may have an additional C-terminal extension.<sup>367</sup> Binding of the cofactor is mainly achieved by the PLP-binding lysine and the highly specific pocket that anchors the PLP phosphate group. Similar to other members of the  $\beta$ -family, a glycine-rich region interacts with the phosphate group of the cofactor. The binding pocket of yeast TS lined with basic residues, mainly Lys and Arg, and amide nitrogens generate a positive surface potential, which guide the phosphate group of the OPHS substrate to the active site, whereas Asp163 orients the amino group of the incoming substrate for transaldimination.<sup>363</sup> TS from bacteria and fungi shows an active site and a PLP binding similar to those of other fold-type II PLP-dependent enzymes. On the contrary, the TS structure from *A. thaliana*, reveals a PLP orientation never previously observed for a type II PLP-dependent enzyme, where the pyridine ring of the PLP occupies the site of the substrate phosphate and partly blocks the substrate site. PLP is bound in an inactive conformation in the absence of its allosteric activator. Binding of SAM triggers the rotation of PLP to its active conformation and the reorganization of the active site. Once the PLP is in its active position, the active-site conformation is the same as for all enzymes in the family, and the reaction probably proceeds in the same way as for nonallosteric TS enzymes. The SAM-binding site in TS has a highly conserved plant-specific unique fold.<sup>368</sup>

In the proposed mechanism, the nucleophilic attack of the amino group of the substrate on C4' of the cofactor yields the external aldimine and the release of the neutral side chain of PLP-binding Lys (Figure 28 – pathway a). This is the only residue in the immediate environment of the bound substrate that can act as an acid–base catalyst. Thus, it abstracts the  $\alpha$ -proton on the C $\alpha$  atom of the substrate to yield the carbanion. The negative charge flows into the pyridine ring of the cofactor to stabilize the carbanion intermediate although the charge delocalization is less effective than that in the typical quinonoid form. After the  $\alpha$ -proton elimination, Lys adds a proton to the C4' atom of the carbanion intermediate, yielding a ketimine intermediate. Lys abstracts the proton on the C $\beta$  atom to produce the enamine and transfers it to the phosphate to yield a  $\beta,\gamma$ -unsaturated ketimine and inorganic phosphate. The proton is abstracted from C4' of the cofactor and transferred to C $\gamma$  of the unsaturated ketimine.<sup>363,367</sup> The released inorganic phosphate was proposed as alternative proton carrier to the C $\gamma$  atom instead of the Lys side-chain amino group, distant to act as a proton donor. According to this hypothesis, Lys eliminates a proton on the C4' of the cofactor, the phosphate accepts the proton from Lys and passes it to the C $\gamma$  atom in a concerted way, producing an  $\alpha,\beta$ -unsaturated aldimine. Lys abstracts a proton from a water molecule located at the *si*-face side of the cofactor forming the OH<sup>-</sup> anion. The OH<sup>-</sup> attacks the C $\beta$  atom from the *si*-face side and Lys adds its proton to the C $\alpha$  atom of the substrate, resulting in the formation of the external aldimine. The enzyme maintains the closed form or a pseudo-closed form until the product threonine is expelled from the active site to guarantee the stereochemical control on water addition. The presence of the side chain of Lys and of the phosphate ion sterically directs the site of addition. Finally, the cofactor recovers the Schiff base with Lys and releases the product, threonine.<sup>363,367</sup>

#### 7.10.4.2.2 Cystathionine $\gamma$ -synthase

CGS catalyzes the  $\gamma$ -replacement reaction of an activated form of L-homoserine with L-cysteine, leading to cystathionine. *O*-Succinyl-L-homoserine (L-OSHS), *O*-acetyl-L-homoserine (OAHS), and *O*-phospho-L-homoserine (OPHS) are substrates for CGS from bacteria, fungi, and plants, respectively.<sup>332</sup> The plant enzyme is also able to convert the microbial substrates, albeit at much higher  $K_m$  values.<sup>369</sup> This reaction is the first step in the transsulfuration pathway that converts L-Cys into L-homocysteine, the immediate precursor of L-methionine. The *O*-activated L-homoserine substrate is situated at a metabolic branch point between L-Met and L-Thr biosynthesis, and which substrate is used by CGS depends on the species. In analogy with TS, CGS is tightly regulated by SAM concentration in plants.<sup>370</sup>

CGS is a member of the  $\gamma$ -subfamily of fold-type I. Comparison of the structure of *E. coli* CGS with that of yeast CGL and *E. coli* CBL demonstrated that there is sequence and structural conservation in the active sites of



**Figure 28** General mechanism of the  $\gamma$ -replacement reaction. In CGS, when the substrate is OASH, the phosphate group is substituted by CH<sub>3</sub>CO, when it is OSHS, the group is COO<sup>-</sup>CH<sub>2</sub>CH<sub>2</sub>CO. R = CH<sub>2</sub>CH(NH<sub>3</sub><sup>+</sup>)COO<sup>-</sup>.

these enzymes. The different reaction specificity and chemical reaction types supported by the enzymes are thought to be a consequence of their different substrate-binding properties. The active CGS has been identified as a homotetramer with one PLP cofactor bound per monomer through a Schiff base linkage to an active-site lysine. The four monomers assemble into two catalytically active dimers. Each CGS subunit can be divided into three domains: an N-terminal domain composed of an extended loop structure, a large PLP-binding domain and a five-stranded  $\beta$ -sheet C-terminal domain never observed in any other related PLP enzyme structure. The loop protrudes much further into the active site than in CBL, narrowing the active-site cleft and leading to a reorientation of a structural element, thereby opening a second entrance to the active site. This region contains neither catalytically important residues nor amino acids implicated directly in substrate binding but presumably influences the overall shape and the electrostatic curvature of the active site and plays a steric role for substrate binding or product release. The active site of CGS is situated at the subunit interface and comprises residues from two subunits.<sup>339,371,372</sup> The consensus PLP-binding site and several catalytically critical residues are conserved. Besides the covalent bond, the cofactor is anchored predominantly to the active site through its phosphate group. In addition, Asp173 (*E. coli* nomenclature) forms a strong hydrogen bond/salt bridge to the PLP pyridine nitrogen, stabilizing its positive charge and increasing the electrophilic character of the cofactor. Arg361 (*E. coli* nomenclature) is largely conserved in the  $\alpha$ - and  $\gamma$ -family and binds the  $\alpha$ -carboxylate group of the incoming substrate. Other residues postulated to bind L-OSHS are Arg48, Tyr101, Arg106, Ser326, while Asp42, Arg49, and Glu325 bind L-Cys (*E. coli* nomenclature). It is probable that a conformational change in the protein is required to allow access to the aldimine.<sup>332</sup> After productive binding of the substrate, its amino group must be deprotonated for the nucleophilic attack on C4' of the internal aldimine. At optimal pH, Tyr101 should exist as phenolate due to the two neighboring positive charges (Arg48, NH of the internal aldimine) and is proposed to deprotonate the incoming substrate and initiate transaldimination. After transaldimination, the PLP-binding Lys is responsible for the proton transfer from C $_{\alpha}$  to C4' of the external aldimine. Tyr46 (*E. coli* nomenclature) assists the positioning of the protonated Lys amino group near C4' after the initial  $\alpha$ -proton abstraction. The side chain of Lys can swing to either C $_{\alpha}$ , C4', or C $_{\beta}$  of the substrate and it is therefore also proposed to deprotonate C $_{\beta}$  to initiate  $\gamma$ -cleavage. Tyr101 also facilitates the elimination of the  $\gamma$ -substituent with the formation of a  $\beta,\gamma$ -unsaturated ketimine. The protonated Schiff base makes this intermediate electron deficient and activates the  $\gamma$ -carbon toward Michael nucleophilic addition. Nucleophilic attack by the thiol group of cysteine on the  $\beta,\gamma$ -unsaturated ketimine intermediate and reversal of the previous steps results in a net  $\gamma$ -replacement reaction and release of cystathionine (Figure 28 – pathway b).<sup>371</sup> The mechanism is ordered with L-homoserine ester associating before L-Cys at L-Cys concentrations greater than  $K_M^{L-Cys}$  and becomes ping-pong with  $\gamma$ -substituent elimination preceding L-Cys association at L-Cys concentrations lower than  $K_M^{L-Cys}$ .<sup>373,374</sup>

The enzyme also catalyzes a competing  $\gamma$ -elimination reaction, in which L-homoserine ester is hydrolyzed to succinate, acetate, or phosphate (depending on the activated substrate), NH<sub>3</sub>, and  $\alpha$ -ketobutyrate. The  $\beta,\gamma$ -unsaturated ketimine has been identified as the common partitioning intermediate between the  $\gamma$ -elimination and  $\gamma$ -replacement reactions.<sup>375</sup> Tautomerization of the  $\beta,\gamma$ -unsaturated ketimine involving removal of the C-4' proton, formation of the highly conjugated electron-rich vinylglycine quinonoid, and subsequent protonation at the  $\gamma$ -carbon yields the aminocrotonate derivative of PLP. Subsequent transaldimination of the aminocrotonate derivative involving the  $\epsilon$ -amino group of a lysine side chain releases the  $\alpha$ -iminobutyric acid from the active site, where it undergoes hydrolysis to yield  $\alpha$ -oxobutyrate and ammonia. The total absence of the aminocrotonate derivative of PLP distinguishes the  $\gamma$ -elimination reactions catalyzed by the plant and bacterial CGS. Despite the similarities in the amino acid sequence, the environment of PLP in the *E. coli* and *A. thaliana* enzymes are different. The plant enzyme is more specific for the  $\gamma$ -replacement than for the  $\gamma$ -elimination of OAHs. Thus, a low rate of the  $\gamma$ -elimination reaction in plants might prevent consumption of OPHS through a futile process.<sup>376</sup>

CGS catalyzes *in vitro* a variety of other nonphysiological reactions such as  $\beta$ -replacement,  $\alpha$ -, or  $\beta$ -proton-exchange reactions and CBL-like  $\beta$ -elimination reactions, with very low efficiency.<sup>377,378</sup> *In vitro*, CGS is able to accept sulfide as the nucleophilic substrate in place of L-Cys, thus maintaining the original sulfhydrylase activity.<sup>379</sup> A microbial CGS exhibits a broad substrate specificity because several thiol compounds such as sodium sulfide, homocysteine, and various alkyl and aryl mercaptans can replace cysteine.<sup>380</sup>

## 7.10.5 Radical Reactions in PLP-Enzymes

The strong electron-withdrawing capacity of PLP accounts for the role played by this cofactor in a plethora of enzymatic reactions catalyzed by vitamin B<sub>6</sub>-dependent enzymes through the stabilization of carbanionic intermediates. Nevertheless, it is well assessed that in some cases PLP can play unusual roles, as for example, as a general acid/base catalyst in the phosphorolysis of glycogen catalyzed by glycogen phosphorylase. A less appreciated function of vitamin B<sub>6</sub>-derivatives is the stabilization of radical intermediates, as is the case with some aminomutases and CDP-6-deoxy-*D-glycero-L-threo*-4-hexulose-3-dehydrase (often referred to as E<sub>1</sub> in the literature).<sup>381</sup> The stabilization of a species with an unpaired electron exploits the electron sink properties of PLP, albeit only in the case of E<sub>1</sub> the unpaired electron is truly delocalized over the pyridine ring.<sup>382</sup> The catalytic properties of PLP-dependent enzymes that catalyze radical reactions are mostly determined by the copresence of other coenzymes, including vitamin B<sub>6</sub>-derivatives, NADH, FAD, Fe-S clusters, adenosylcobalamin (AdoCbl), and *S*-adenosyl-L-methionine.

### 7.10.5.1 Aminomutases

To date many PLP-dependent enzymes catalyzing 1,2-migrations of amino groups through a radical chemistry have been characterized (see Ruzicka and Frey<sup>383</sup> and references therein). PLP-dependent aminomutases (Table 2) catalyze very similar reactions, where a 5'-deoxyadenosyl radical, generated by the homolytic cleavage of either SAM or AdoCbl abstracts a proton from the amino acid substrate. The reaction proceeds through a cyclic radical intermediate (see below) bound to PLP which is subsequently hydrolyzed to regenerate the 5'-deoxyadenosyl radical, the product, and the internal aldimine of the coenzyme.

The more thoroughly characterized PLP-dependent aminomutase is lysine 2,3-aminomutase (LAM), therefore its reaction mechanism and structure will be discussed in Section 7.10.5.1.1.

#### 7.10.5.1.1 Lysine 2,3-aminomutase

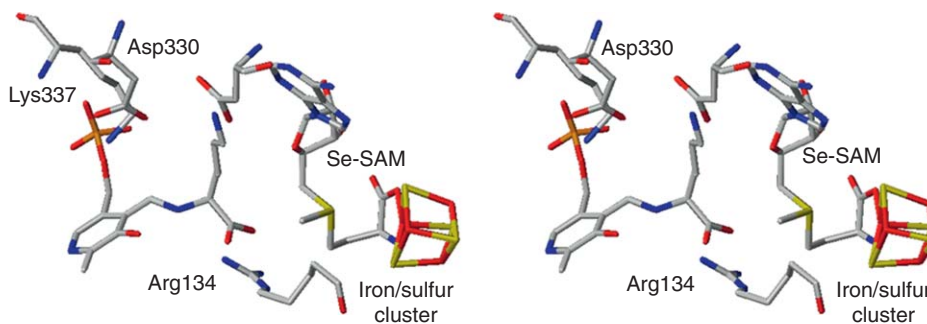
LAM, which has been isolated from many bacterial sources, catalyzes the isomerization of lysine and  $\beta$ -lysine.<sup>384,390</sup> The isomerization of lysine takes place with inversion of the configuration,<sup>382</sup> for example, the 3-*pro-R* hydrogen of lysine is transferred to the 2-*pro-R* position of  $\beta$ -lysine.  $\beta$ -Lysine is thought to play different roles, for example, as a source of acetyl-CoA and ammonia in *Clostridia*<sup>391</sup> or as a secondary metabolite involved in the modification of antibiotic molecules in *Streptomyces* (see Agnihotri and Liu<sup>382</sup> and references therein).

LAM belongs to the family of radical-SAM enzymes<sup>392</sup> that share the common feature of using SAM as a source of 5'-adenosyl radicals. LAM is the only member of the *radical-SAM* family to exploit the properties of PLP to stabilize the radical intermediate.

LAM from *Clostridium subterminale* in complex with the substrate crystallizes as a homotetramer<sup>393</sup> with domain-swapped dimers stabilized by structural zinc ions. Each monomer contains one active site, which is a channel defined by a ( $\beta/\alpha$ )<sub>6</sub>-crescent and binds a molecule of PLP, a molecule of SAM, and a [4Fe-4S] cluster. In the active site the external aldimine of PLP with lysine, SAM, and the Fe-S cluster lie spatially close (Figure 29),<sup>393</sup> supporting the idea that the formation of the 5'-deoxyadenosyl radical and the abstraction of the

**Table 2** PLP-dependent aminomutases acting through radical chemistry

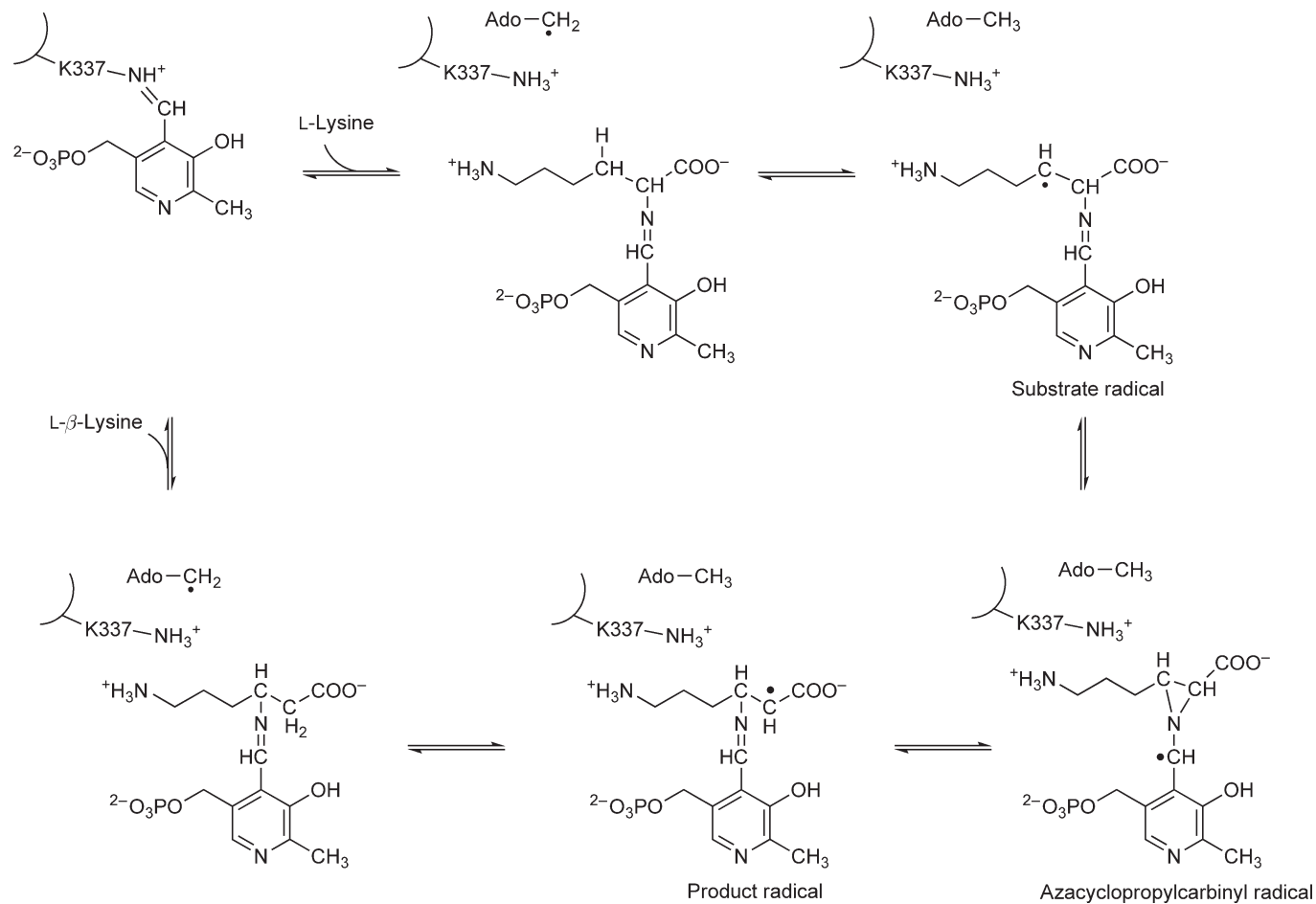
Aminomutase	SAM/AdoCbl dependent	Reaction product	Cellular function	Reference(s)
Lysine 2,3-aminomutase	SAM-dependent	$\beta$ -Lysine	Antibiotic synthesis. Source of carbon and nitrogen	384
Arginine 2,3-aminomutase	SAM-dependent	$\beta$ -Arginine	Antibiotic synthesis	385
Glutamate 2,3-aminomutase	SAM-dependent	$\beta$ -Glutamate	Osmoregulator	383
Lysine 5,6-aminomutase	AdoCbl-dependent	2,5-Diaminohexanoic acid	Lysine fermentation	386–388
Ornithine 4,5-aminomutase	AdoCbl-dependent	2,4-Diaminopentanoic acid	Stickland fermentation	389



**Figure 29** Ball-and-stick, wall-eyed stereodiagram figure of the active site of LAM. This wall-eyed stereodiagram emphasizes the stereochemical relationships of the three cofactors, with PLP bound as lysyl external aldimine.

C3 hydrogen from the cofactor-bound lysine take place without the need for the conformational changes observed in AdoCbl-dependent enzymes.<sup>386,393</sup>

The binding mode of both PLP and [4Fe-4S] cluster is peculiar and gives a structural support to the reaction mechanism. PLP is bound to the N-terminus of the active-site channel and is held in place by a number of hydrogen bonds, mainly involving the phosphate group of the cofactor. Many water molecules are observed in the active site, one of which is involved in an unprecedented hydrogen bond with the pyridinium nitrogen. This interaction stabilizes the deprotonated form of the pyridinium nitrogen whose electron sink properties weaken. As shown later, the radical intermediate stabilization is attained through the delocalization of the unpaired electron over the imine bond and does not involve the pyridinium ring. A conserved sequence of the type CXXXCXXC, observed in all the members of the radical-SAM enzyme family,<sup>394</sup> is involved in the coordination of three out of four iron atoms of the [4Fe-4S] cluster. This coordination mode leaves one iron atom with one accessible coordination site, which is occupied by the carboxylate and the  $\alpha$ -amino group of SAM.<sup>394,395</sup> The reaction mechanism of LAM involves the exploitation of the chemical properties of widespread cofactors, such as PLP, SAM, and a Fe-S cluster in an unprecedented way. The reaction catalyzed by LAM is divided into two main steps: 5'-deoxyadenosyl radical formation and lysine/ $\beta$ -lysine isomerization. The first reaction takes place on the [4Fe-4S] cluster and involves SAM as a source of 5'-deoxyadenosyl radicals. For a long time SAM has been considered as the methyl-donor in biological reactions, whereas many of the chemical properties of this compound have been far less appreciated. Now, it is proved that all the functional groups bound to the sulfur atom of SAM have distinct biological destinations owing to the electrophilic character of carbon atoms bound to sulfonium centers.<sup>394</sup> In LAM, SAM is a donor of 5'-deoxyadenosyl radicals, a function which is more frequently performed by AdoCbl, a fact that led Barker to the definition of SAM as 'a poor's man adenosylcobalamin.'<sup>396</sup> The major issue when SAM is considered as a donor of 5'-deoxyadenosyl radicals concerns the mechanism of scission of SAM into 5'-deoxyadenosine radical and methionine, a process requiring an electron. By no means can the reaction involve the direct homolysis of a C-S bond, that is estimated to have a bond dissociation energy of 60 kcal mol<sup>-1</sup>,<sup>396</sup> twice as much as the bond dissociation energy for the homolysis of the Co-C bond in adenosylcobalamin.<sup>390</sup> The source of reducing power is the [4Fe-4S]<sup>+1</sup> cluster. In the inactive enzyme, the Fe-S cluster is in the [4Fe-4S]<sup>+2</sup> oxidation state, which is reduced to the [4Fe-4S]<sup>+1</sup> form upon SAM binding under reducing conditions.<sup>384</sup> In the resting enzyme PLP is bound to the side chain of a lysine residue through a Schiff base linkage.<sup>397</sup> The binding of the substrate lysine leads to the formation of the external aldimine through a classic transimination reaction (Figure 30).<sup>398</sup> The 5'-deoxyadenosyl radical, which is thought to be bound no further than 3.8 Å from the lysyl side chain,<sup>393</sup> abstracts the 3-*pro-R* hydrogen leading to the substrate radical. The unpaired electron is delocalized into the imine bond of the Schiff base leading to the cyclic intermediate. The reversible opening of the azacyclopropylcarbinyl radical leads, in the forward direction, to the formation of the product radical. The presence of the product radical on the reaction pathway has been proved by EPR spectroscopy<sup>399</sup> owing to its enhanced stability with respect to the substrate radical. The high steady-state concentration of this species (about 10% of the concentration of the enzyme)<sup>396</sup> is attributed to the stabilization of the unpaired electron by delocalization into the carboxyl group.<sup>396</sup> The identification of the substrate radical has been accomplished



**Figure 30** Mechanism for L-lysine/L-β-lysine isomerization catalyzed by LAM. The 5'-deoxyadenosyl radical ( $\text{Ado}-\text{CH}_2\cdot$ ) is generated from the homolytic cleavage of SAM at the [4Fe-4S] cluster site. Reproduced with permission from P. A. Frey; M. D. Ballinger; G. H. Reed, *Biochem. Soc. Trans.* **1998**, 26 (3), 304-310.



using the substrate analogue 4-thia-L-lysine.<sup>400,401</sup> The abstraction of an hydrogen from 5'-deoxyadenosine leads to the regeneration of the 5'-deoxyadeanosyl radical and to the formation of the external aldimine of  $\beta$ -lysine, that is subsequently hydrolyzed to  $\beta$ -lysine and PLP.

#### 7.10.5.1.2 CDP-6-deoxy-D-glycero-L-threo-4-hexulose-3-dehydrase

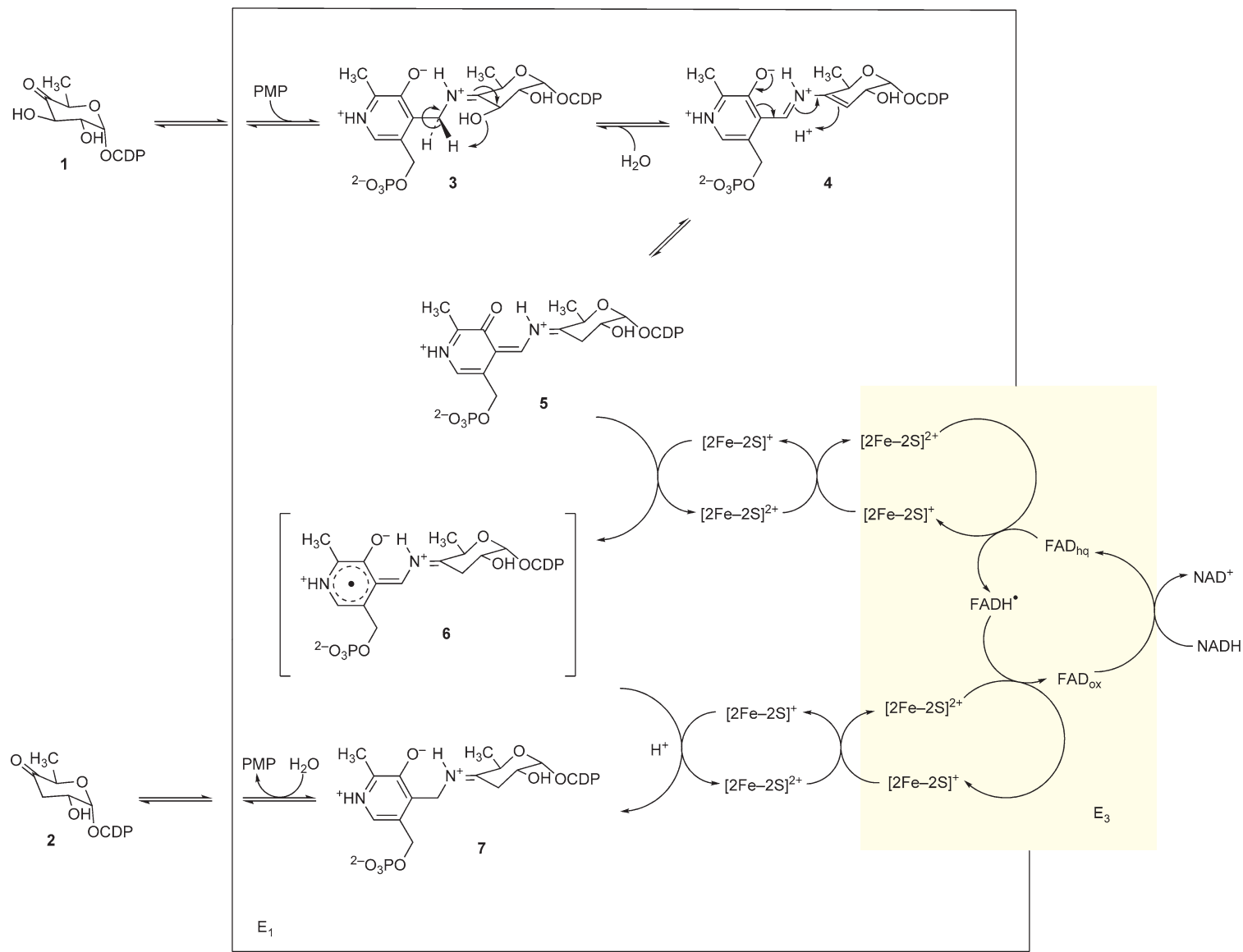
The reaction catalyzed by CDP-6-deoxy-D-glycero-L-threo-4-hexulose-3-dehydrase is a further example of how the amazing versatility of vitamin B<sub>6</sub>-derived cofactors has been exploited by nature. It is the only known example to date of a dehydration reaction catalyzed by a PMP-dependent protein. E<sub>1</sub> depends on a [2Fe-2S] cluster for its activity. E<sub>1</sub> is part of the 3,6-dideoxyhexoses synthetic pathway in bacteria.<sup>402</sup> In *Yersinia pseudotuberculosis* E<sub>1</sub> catalyzes the C3 deoxygenation of CDP-6-deoxy-D-glycero-L-threo-4-hexulose in the biosynthetic pathway of L-ascarylose. The reducing power needed to reconstitute the active form of E<sub>1</sub> (e.g., regeneration of PMP) is released by an enzyme of the flavodoxin-NADP<sup>+</sup> reductase family, CDP-6-deoxy-D-glycero-L-threo-4-hexulose-3-dehydrase reductase (E<sub>3</sub>).<sup>403,404</sup>

A few signatures in the amino acid sequence of E<sub>1</sub> are suggestive of a protein, which is evolutionarily distinct from all the known PMP- and Fe-S-dependent enzymes, which may thus constitute the representative of a new enzyme subclass.<sup>405</sup> The sequence of E<sub>1</sub> shows some similarities with those of known PLP-dependent enzymes, albeit with one important difference: the active-site Lys involved in internal aldimine formation with the cofactor is absent. Its position is occupied by a His residue (His220), conserved among the dehydrases.<sup>405,406</sup> E<sub>1</sub> is one of few examples of PMP-dependent enzymes that bind a [2Fe-2S] cluster.<sup>405,407</sup> The unprecedented cluster ligation pattern is of the type C-X<sub>1</sub>-C-X<sub>7</sub>-C-X<sub>16</sub>-H,<sup>405,408</sup> and is believed to strongly influence the properties of the cluster, as for example, its high redox potential (-209 mV) and the strong coupling with the Fe-S cluster of E<sub>3</sub>.<sup>409</sup> Recently, GABA transaminase has been shown to contain a iron-sulfur cluster with unknown function.<sup>407</sup>

The removal of the C3 hydroxyl group from CDP-6-deoxy-D-glycero-L-threo-4-hexulose (**1**) (**Figure 31**) to give CDP-3,6-deoxy-D-glycero-L-threo-4-hexulose (**2**) is a challenging reaction, due to the lack of a labile hydrogen that would facilitate the 1,2 elimination of a water molecule. The first step in the reaction catalyzed by E<sub>1</sub> is therefore the binding of **1** through a Schiff base linkage to PMP. The C4' hydrogen of adduct **3** is acidic and can be abstracted by a base leading to a  $\Delta^{3,4}$ -glucose-en intermediate (**4**). It was previously thought that the base responsible for the abstraction of C4' hydrogen was His220, but recent results suggest a possible role of this residue as a general acid that makes 3-OH a better leaving group.<sup>410</sup> The regeneration of PMP from intermediate **4** and the release of the product is achieved through a two-step reduction with the formation of the radical intermediate **6**. The Fe-S cluster of E<sub>1</sub> functions as a one-electron carrier, which is reduced by FAD on E<sub>3</sub>. Interestingly, the switch from two-electron chemistry to one-electron chemistry is allowed by the delocalization properties of PMP, which stabilizes an anionic intermediate via a quinonoid species (**5**) that can subsequently undergo a one-electron reduction with formation of a radical delocalized over the pyridine ring. The structure of the radical intermediate **6** has been assessed by EPR spectroscopy and pulsed ENDOR.<sup>411</sup> Subsequent hydrolysis of the reduction product leads to the regeneration of PMP and the release of the product **2**.

### 7.10.6 Reactivity with Inhibitors

PLP-dependent enzymes represent a very interesting pharmaceutical target for the development of mechanism-based and noncovalent inhibitors. The high interest demonstrated by the scientific community toward these enzymes is justified by their involvement in numerous cellular processes, including pathogens and plants metabolic pathways and in several human diseases. Different strategies have been exploited in order to produce effective and selective drugs, comprising irreversible mechanism-based inhibitors and, more recently, noncovalent or allosteric inhibitors. Some representative examples for each category are given below.



**Figure 31** Reaction mechanism for dehydration of CDP-6-deoxy-D-glycero-L-threo-4-hexulose catalyzed by E<sub>1</sub>. The reduction of E<sub>1</sub> by its reductase E<sub>3</sub> is also shown. Reproduced with permission from G. Agnihotri; H. W. Liu, *Bioorg. Chem.* **2001**, 29 (4), 234–257; G. Agnihotri; Y. N. Liu; B. M. Paschal; H. W. Liu, *Biochemistry* **2004**, 43 (44), 14265–14274.

### 7.10.6.1 Mechanism-Based Inhibitors

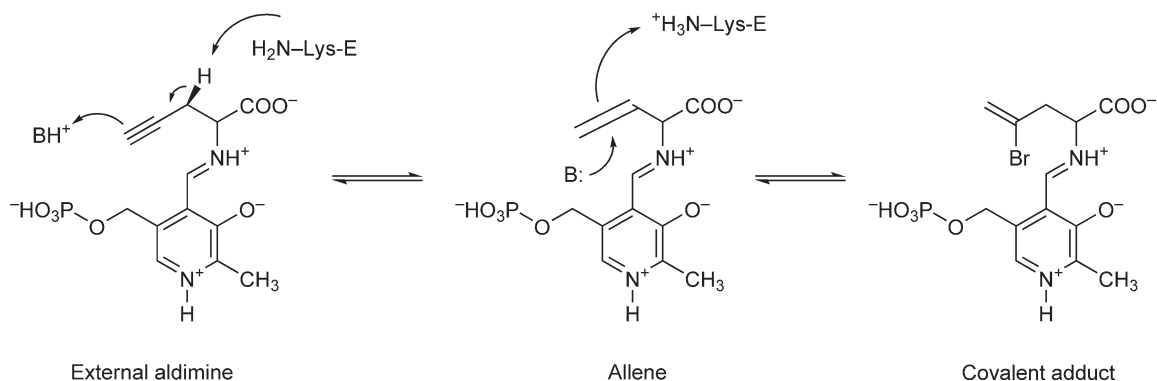
Among the group of mechanism-based inhibitors, three different categories can be identified: (1) activated electrophiles acting through a Michael addition mechanism, (2) activated nucleophiles acting through an enamine mechanism, and (3) compounds reacting with PLP and forming stable complexes through an aromatization inactivation mechanism.<sup>412</sup>

#### 7.10.6.1.1 Activated electrophiles

The presence of highly electrophilic moieties or groups able to be electrophilically activated through the interaction with PLP is typical of several irreversible PLP-dependent enzyme inhibitors and leads to the formation of covalent bonds with active-site nucleophilic residues.<sup>412</sup> A very large category of activated electrophile inhibitors is constituted by vinyl compounds, usually presenting a  $\beta,\gamma$ -unsaturated imine Michael acceptor, such as vigabatrin and the tetrazole bioisostere of vigabatrin.<sup>413,414</sup> These compounds are irreversible  $\gamma$ -aminobutyric acid aminotransferase (GABA-AT) inhibitors, susceptible of the nucleophilic attack of Lys329 on the penultimate carbon of the vinyl bond. In addition, two irreversible TS inhibitors: DL-Z-2-amino-5-phosphono-3-pentenoic acid and DL-E-2-amino-5-phosphono-3-pentenoic acid belong to this category and, differently from DL-2-amino-3-[(phosphonomethyl)amino]propanoic acid, exploit the activation of an electrophilic group, rather than the formation of a stable aromatic adduct<sup>415</sup> (see Section 7.10.6.1.3).

An alternative process, always involving the formation of an activated electrophile, is adopted by propargylglycine, a well-known inhibitor of PLP-dependent enzymes involved in the trans-sulfuration pathways. The mechanism used by propargylglycine to inhibit CGS, CGL,<sup>416</sup> and MGL involves the attack of PLP and the formation of an allene intermediate, subsequently transformed into a stable ketimine derivative, by the Michael addition of an enzyme nucleophile residue (Figure 32).

An acetylenic group is also present in  $\gamma$ -ethynyl-GABA,<sup>417</sup> one of the first GABA-AT mechanism-based inhibitor, able to attack also glutamate-1-semialdehyde-1,2-mutase (GSAM).<sup>418</sup> The formation of an allene intermediate is also involved in the irreversible reaction of AlaR with 3-halovinylglycines, mechanism-based inhibitors, which use the haloethylene moiety to generate a reactive allene during catalysis. Differently from the other nonvinyl  $\beta$ -substituted Ala-R inhibitors, the reaction proceeds through the formation of an electrophilic allene, rather than a nucleophilic aminoacrylate.<sup>419</sup> Finally, several  $\beta$ -halosubstituted derivatives adopt an activated electrophile strategy. This is the case of  $\beta,\beta,\beta$ -trifluoroalanine with CGL<sup>416</sup> and CBL,<sup>420</sup> where the formation of a Schiff base between the inhibitor and the enzyme-cofactor complex is followed by the elimination of hydrogen fluoride and the generation of an activated Michael acceptor species such as an  $\alpha,\beta$ -unsaturated imine, subsequently attacked by an enzyme nucleophilic residue.<sup>247,420</sup> Analogously, the interaction between ODC and  $\alpha$ -difluoromethylornithine (DFMO) generate a conjugated imine that irreversibly interacts with Cys 360.<sup>421</sup> DFMO is conceivably able to bind also to diaminopimelate decarboxylase (DAPDC).<sup>89</sup> A similar mechanism has been suggested for the interaction of  $\alpha$ -fluoromethylhistidine with HDC,<sup>422,423</sup> and for the



**Figure 32** Mechanism of inactivation by propargylglycine.<sup>5</sup>

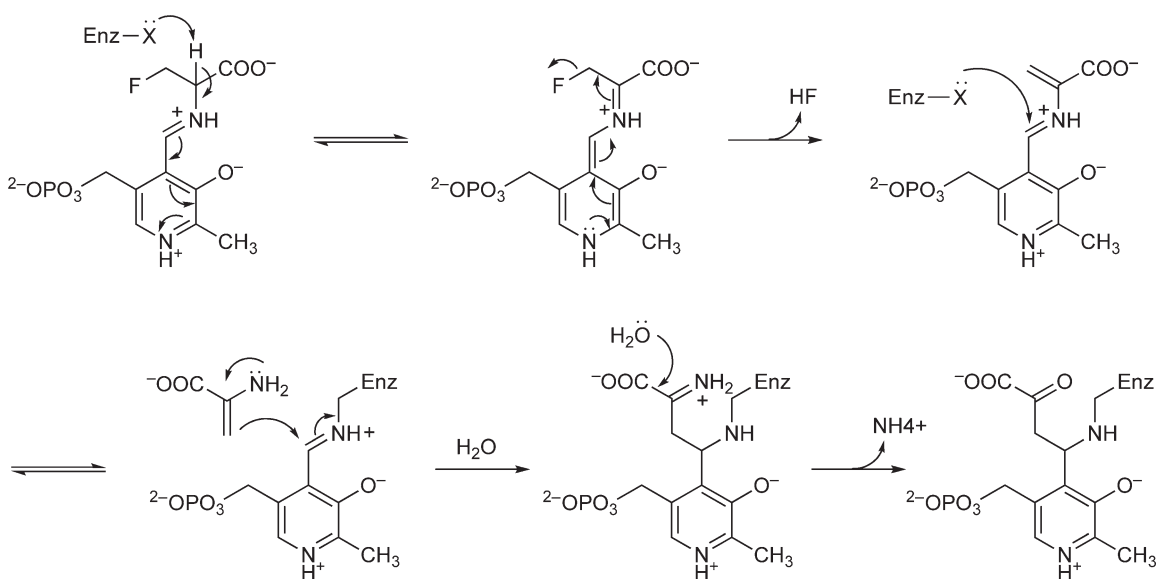
irreversible complexes formed between AON synthase (KAPAS) and trifluoroalanine,<sup>424</sup> and di- and tri-halogenated alanine and Ala-R (see the Ala-R case section for further details and schemes).<sup>425,426</sup>

### 7.10.6.1.2 Activated nucleophiles

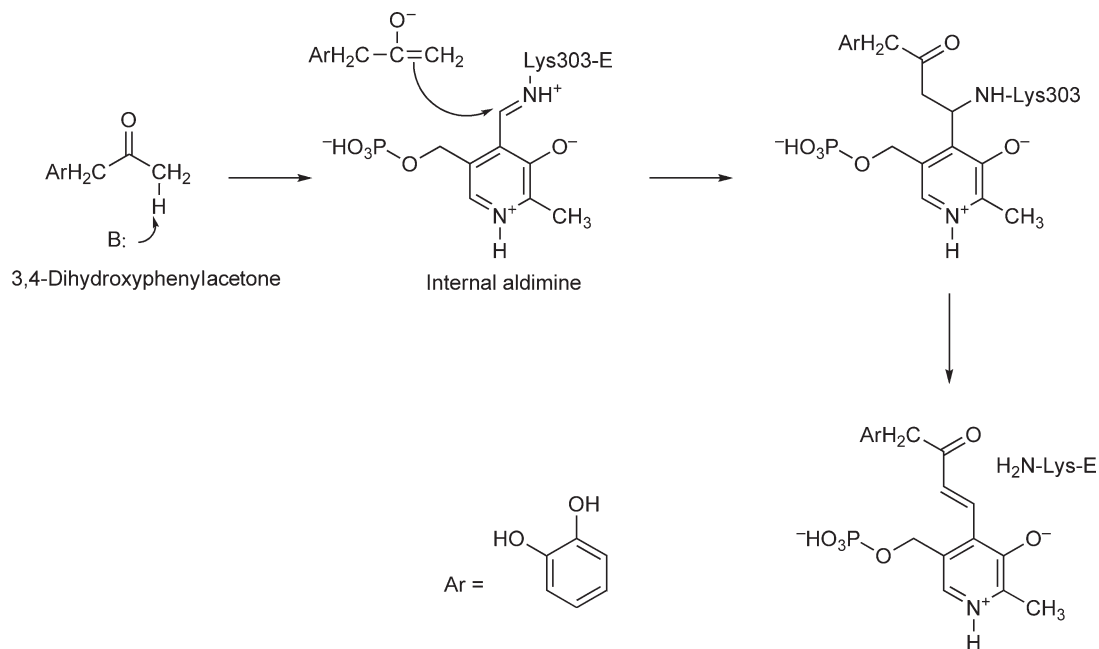
Inhibitors able to generate activated nucleophile species irreversibly inactivate the target enzyme because a stable bond is formed between PLP and the enzyme. This strategy, also known as enamine inactivation mechanism, usually evolves through the formation of a nucleophilic enamine intermediate, able to covalently attack the internal aldimine.<sup>412</sup> Activated nucleophile ligands are often characterized by the presence of good  $\beta$ -leaving groups, such as (1)  $\beta$ -fluoroalanine, a monohalogenated structural analogue of D-Ala and well-known AlaR suicide inhibitor (see **Figure 33** for the inactivation mechanism),<sup>412,426–428</sup> (2) the  $\beta$ -halo-D-amino acids, able to potentially inhibit DAAT,<sup>429</sup> (3) the 4-amino-5-fluoro-pentanoic acid<sup>430,431</sup> and some of its rigid analogues<sup>432–434</sup> inactivating  $\gamma$ -aminobutyric acid aminotransferase (GABA-AT), and (4) 5-fluoromethylornithine, the first specific irreversible inhibitor of L-ornithine: 2-oxoacid aminotransferase (OAT).<sup>435</sup> Being GABA-AT and OAT enzymes characterized by similar catalytic mechanisms,<sup>436,437</sup> the inactivation reactions are analogous. In fact, GABA-AT inhibitors are also potent OAT inactivators.<sup>438,439</sup> A highly reactive intermediate able to alkylate the enzyme target is also generated during the attack to DDC by  $\alpha$ -chloromethyl and  $\alpha$ -fluoromethyl derivatives of DOPA,<sup>440–443</sup>  $\alpha$ -vinylDOPA and  $\alpha$ -acetylenic DOPA.<sup>444</sup> An interesting mixed mechanism, implying the activation of a nucleophile and the formation of a covalent bond only with the cofactor, operates in the reaction of the DDC inhibitor  $\alpha$ -methylDOPA, a well-known antihypertensive agent (**Figure 34**).<sup>82</sup> 3,4-Dihydroxyphenylacetone is produced through a decarboxylation-dependent transamination of  $\alpha$ -methylDOPA, catalyzed by DDC. Subsequently, the enzyme is inactivated by the covalent bond formed between the generated nucleophile and the cofactor, with the C3 of the ketone attacking the aldehydic carbon of PLP, generating an aldol-type adduct and trapping the PLP-lysine aldimine of the active site.<sup>82</sup>

### 7.10.6.1.3 Aromatization mechanism-based inhibitors

This class usually include inhibitors able to form stable aromatic products through enzyme-catalyzed aromatization, or stable ketimine intermediates.<sup>412</sup> The generated adducts involve the formation of a covalent interaction between the substrate and the cofactor, but not with the protein, differently from the two previous

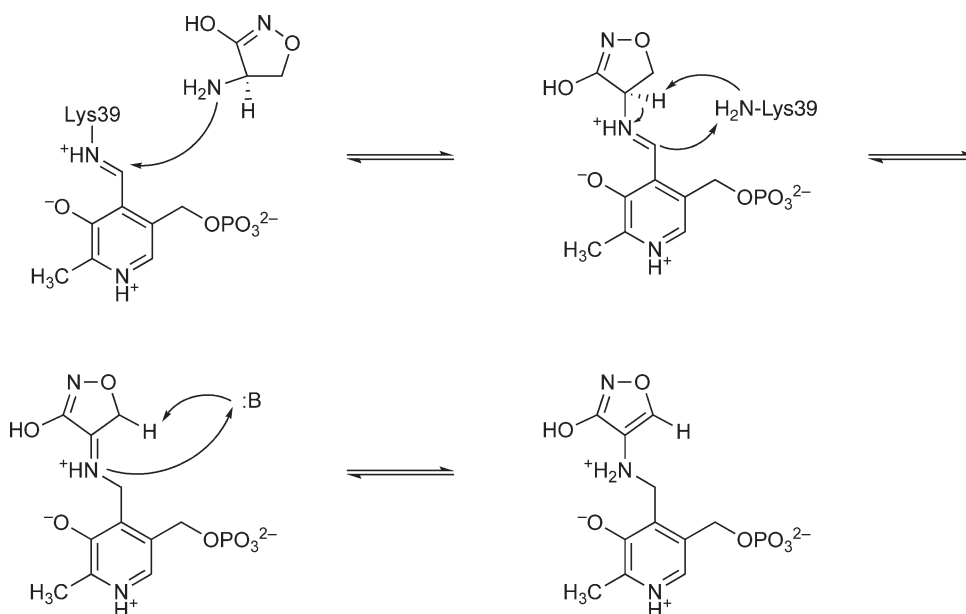


**Figure 33** Mechanism of inactivation of Ala-R by monohalogenated alanine analogues. Reproduced with permission from A. C. Eliot; J. F. Kirsch, *Annu. Rev. Biochem.* **2004**, *73*, 383–415; W. S. Faraci; C. T. Walsh, *Biochemistry* **1989**, *28* (2), 431–437.



**Figure 34** Mechanism of DDC inactivation by  $\alpha$ -methylDOPA.<sup>82</sup>

classes. It derives that the enzyme can usually be reactivated by dialysis in the presence of PLP.<sup>412</sup> An example of aromatization mechanism-based inhibitor is represented by D-cycloserine, a cyclic structural analogue of D-Ala, able to inhibit AlaR<sup>445,446</sup> and DAA,<sup>447</sup> through an initial transamination step and a subsequent tautomerization, generating a stable aromatic product, as reported in **Figure 35**. Other inhibitors adopting this type of strategy are represented by gabaculine, a vigabatrin-like inhibitor able to inactivate GABA-AT,<sup>448</sup> DAA, and OAT<sup>438</sup> and other fluorinated conformationally restricted vigabatrin analogues (see GABA-AT, case



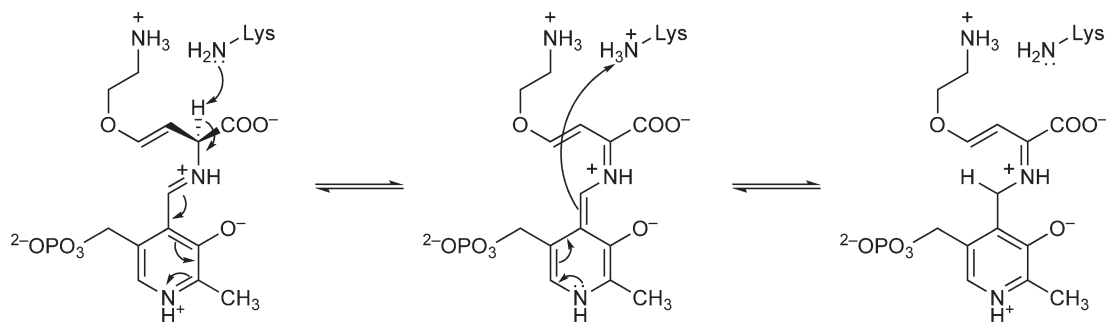
**Figure 35** Aromatization mechanism of AlaR inactivation by D-cycloserine.<sup>445</sup>

section). Aminoclenomycin is able to irreversibly inactivate DAPAS via a reaction forming an external aldimine, a quinonoid, and a ketimine intermediate and a stable aromatic ring derivative that does not dissociate from the active site.<sup>449,450</sup> Analogously, the DL-2-amino-3-[(phosphonomethyl)amino]propanoic acid forms with PLP of TS a stable spirocyclic derivative.<sup>415</sup>

The ketimine intermediate-based approach is exploited by those inhibitors, which are not able to generate an aromatic intermediate, but form a stable not hydrolyzable ketimine, such as aminoethoxyvinylglycine (AVG) when attacking ACC synthase,<sup>451</sup> as reported in **Figure 36**; rhizobitoxine, a CBL inhibitor,<sup>452</sup> or DL-*E*-2-amino-5-phosphono-3-pentenoic acid, a CGS inhibitor.<sup>416</sup> Despite the presence of a vinyl group, AVG does not react following the electrophile-activated mechanism previously described, but forms stable PLP-ketimine that slowly dissociates. AVG is able to act as an inhibitor of different PLP-dependent enzymes involved in the transsulfuration pathways, such as CGS,<sup>416</sup> CBL,<sup>453</sup> CGL, and cystalysin<sup>220</sup> via a slow-binding two-step mechanism.<sup>453</sup> A subclass of this category is represented by reversible inhibitors, constituted by  $\alpha$ -methyl, amino-oxy, or hydrazine analogues, able to bind the PLP cofactor, and to form a stable aldimine which does not react further.<sup>412</sup> This is, for example, the case of the aminoxy compounds *O*-benzylhydroxylamine, *O*-*t*-butylhydroxylamine, carboxymethoxylamine, and *O*-allylhydroxylamine, which act by forming a stable Schiff base with the PLP cofactor.<sup>454</sup> This species represents a starting point for the development of new branched-chain amino acid aminotransferase (BCAT) inhibitors to be used in tuberculosis treatment.<sup>455</sup> Analogously, a transaldimination of the native enzyme occurs when alanine phosphonate reversibly interacts with Ala-R,<sup>456</sup> forming an Ala-P-PLP Schiff-base bond. In contrast, the transaldimination, occurring during the reaction of GSAM with the novel efficient inhibitor 2,3-diaminopropyl sulfate, seems to lead to the formation of a stable and inactive quinonoid structure.<sup>457</sup> Conceivably, several other inhibitors adopt this strategy, such as phenelzine, a compound able to inhibit alanine aminotransferase (AGXT), and to induce, consequently, an increase in alanine brain levels,<sup>458</sup> the SHMT inhibitors sulfonyl fluoride triazine derivatives,<sup>459</sup> the hydrazide, sulfonylhydrazide or oxime-based inhibitors of CBL,<sup>460</sup> or the three DDC inhibitors L- $\alpha$ -methyl- $\alpha$ -hydrazino-3,4-2,3,4-trihydroxyphenylpropionic acid (carbidopa), *N*-(DL-seryl) *N'*-2,3,4-trihydroxybenzyl-hydrazine (benserazide), and 2,3,4-trihydroxybenzylhydrazine (Ro 4-5127), which bind to the PLP cofactor by forming a hydrazone linkage through their hydrazine moiety. However, though benserazide demonstrated to behave as a poor inhibitor, carbidopa and Ro 4-5127 were found to be powerful irreversible inhibitors, thus a different, and not yet deciphered, inactivation mechanism might be exploited.<sup>461-464</sup>

Two examples of enzymes inactivated by inhibitors following alternative mechanisms with respect to the above reported are given below.

The 'alanine racemase case'. Ala-R is a fold-type III PLP-dependent enzyme implicated in the conversion of L-Ala into D-Ala, the amino acid required for the synthesis of prokaryote cell walls. The absence of Ala-R from eukaryotes, and the absolute requirement of D-alanine, makes it an eligible target for the development of selective antibiotics.<sup>34,35</sup> At present, most of the AlaR inhibitors are mono-, di-, or trihalogenated structural analogues of D-Ala and act using different mechanism-based strategies, able to generate both activated

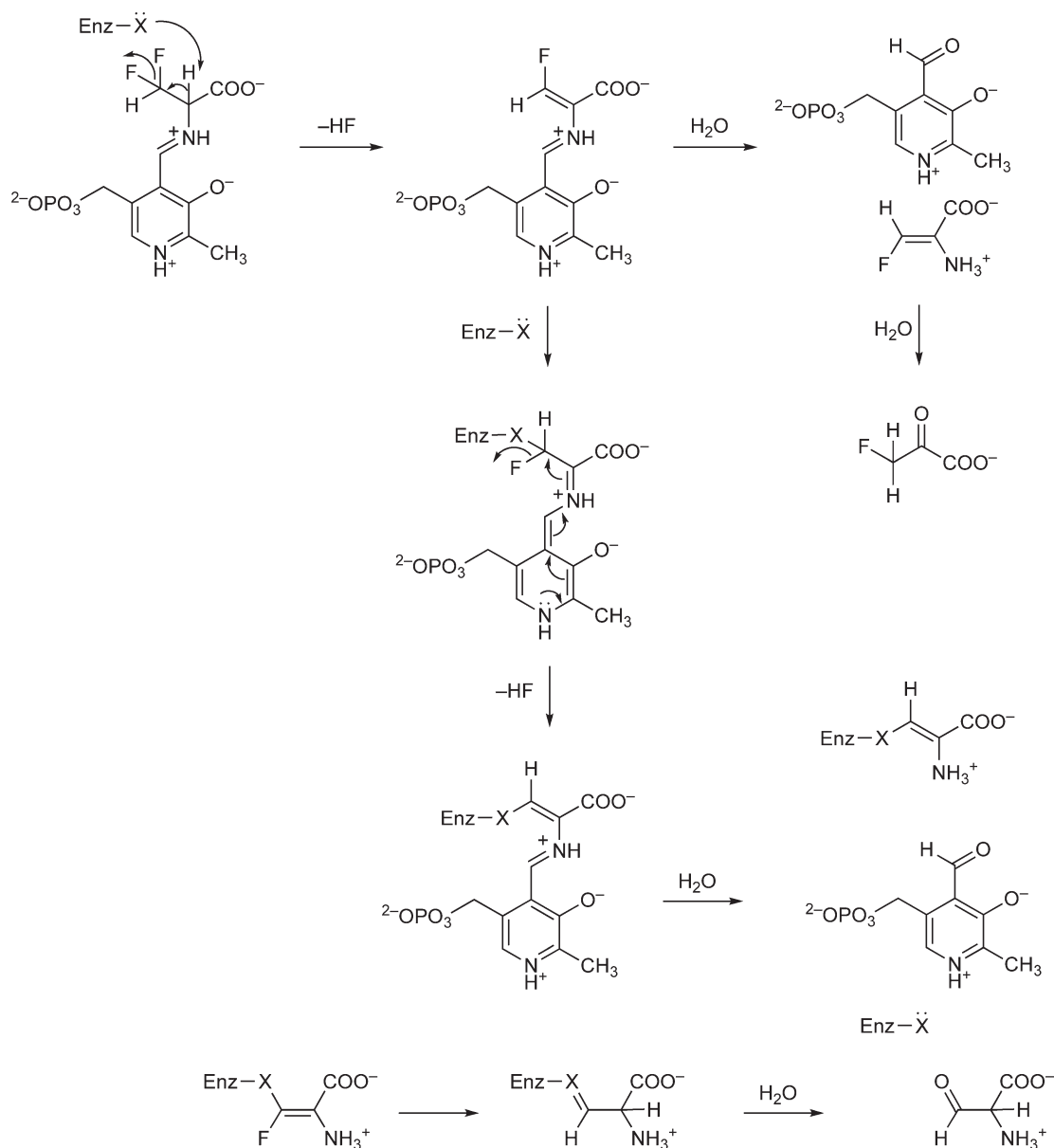


**Figure 36** Ketamine-based inactivation of 1-aminocyclopropane-1-carboxylate synthase by AVG.<sup>451</sup>

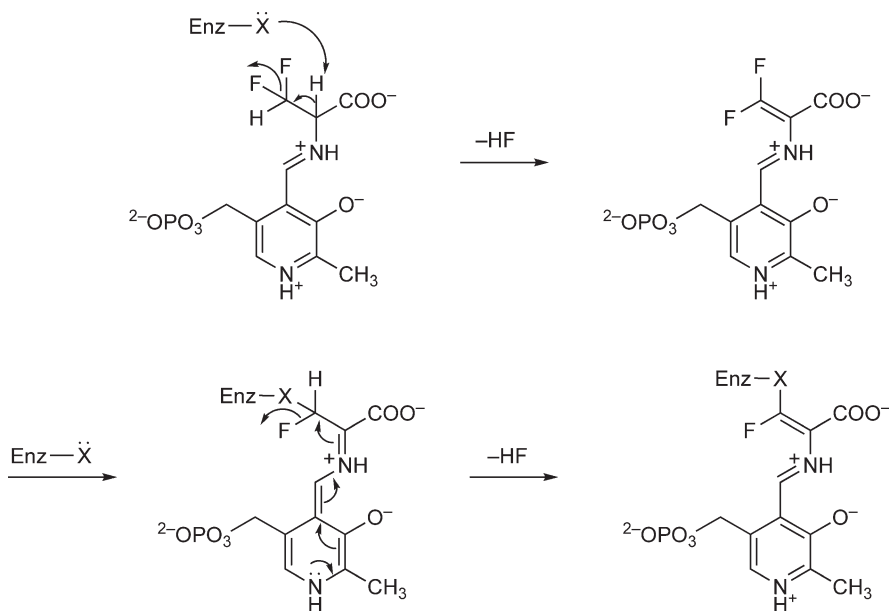


electrophiles or nucleophiles. As reported in **Figure 33**, the presence of a monohalogenated alanine induces the formation of a potent nucleophile group (aminoacrylate), released by transaldimination, and able to subsequently attack and alkylate the PLP–enzyme complex.<sup>412,426</sup> A different mechanism, implying the nucleophilic attack of the enzyme at the olefinic terminus was previously proposed.<sup>425</sup>

Although monohalogenated inhibitors adopt a nucleophile activated strategy, di- and trihalogenated derivatives operate through the formation of activated electrophile species. In particular, as shown in **Figure 37**, the halogenated vinyl adduct, resulting from the attack of a difluoroalanine to the cofactor, could reasonably be the electrophile target of enzyme nucleophilic residues. The loss of the remaining fluorine yields a vinyl-X adduct, that might be hydrolyzed, depending on the nature of the enzyme nucleophilic group.<sup>425</sup>



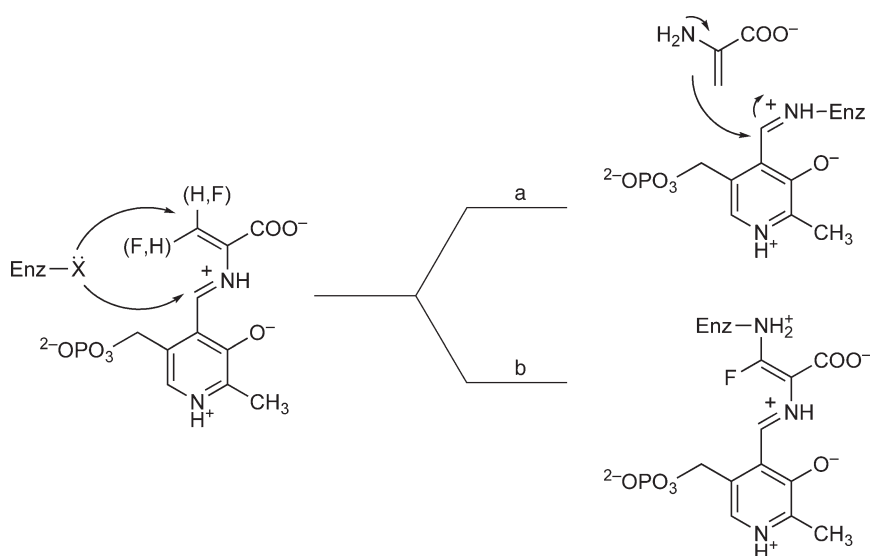
**Figure 37** Mechanism of inactivation of Ala-R by dihalogenated alanine analogues.<sup>425</sup>



**Figure 38** Mechanism of inactivation of Ala-R by trihalogenated alanine analogues. Reproduced with permission from W. S. Faraci; C. T. Walsh, *Biochemistry* **1989**, 28 (2), 431–437.

The inactivation of Ala-R by trihalogenated alanine analogues is shown in **Figure 38**. According to the mechanism proposed by Faraci and Walsh,<sup>426</sup> the addition of the halogenated analogue leads to an initial transaldimination, followed by the elimination of the first fluoride ion and formation of the  $\beta$ -difluoro- $\alpha,\beta$ -unsaturated imine complex. The subsequent nucleophilic attack of the enzyme on the olefinic terminus induces the loss of the second fluoride ion and the formation of a stable inactive complex.<sup>426</sup>

The two distinct mechanisms, based on activated nucleophile and activated electrophile, are summarized in **Figure 39**. According to the presence of mono or ditrihaloalanine analogues, two different reaction pathways could thus be followed. If two hydrogens are present on the  $\beta$ -carbon, the olefinic terminus is not enough

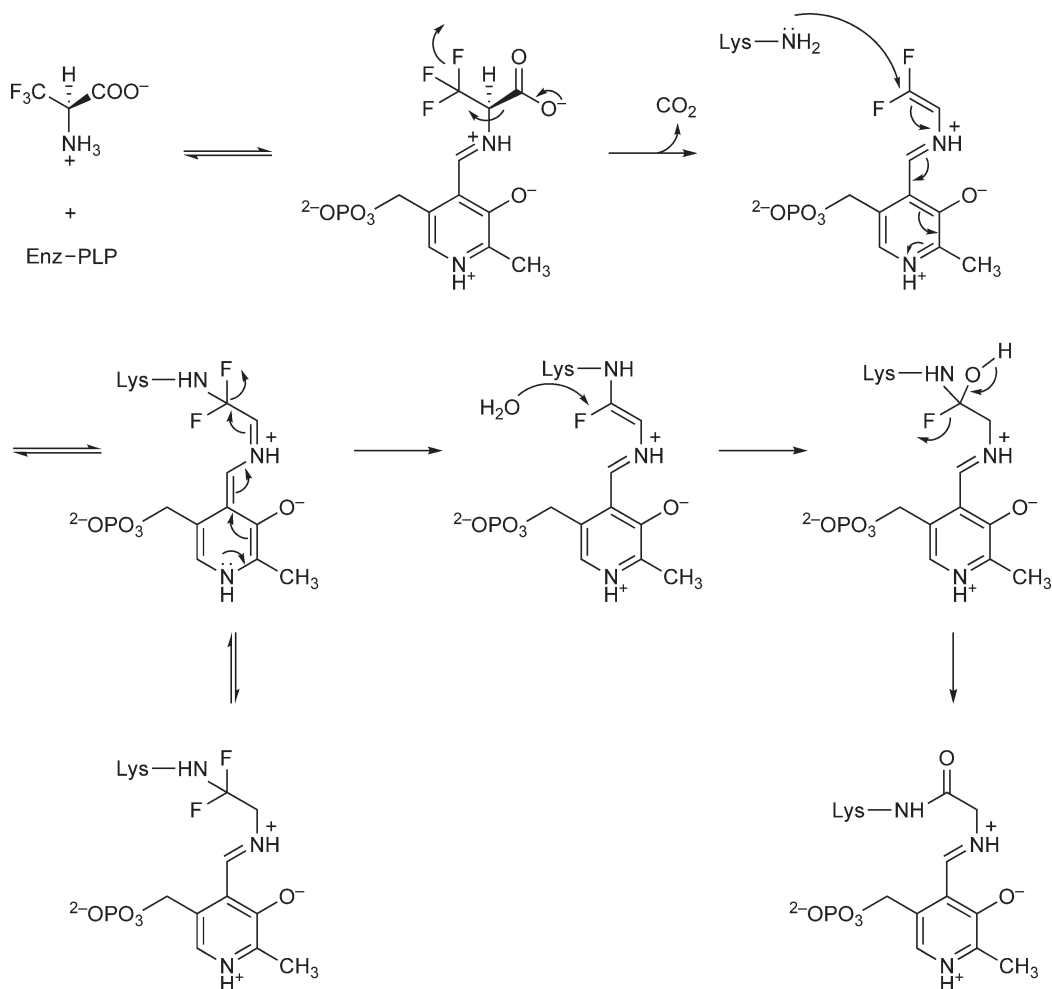


**Figure 39** Different possible inactivation mechanisms of Ala-R by halogenated alanine.<sup>426</sup>

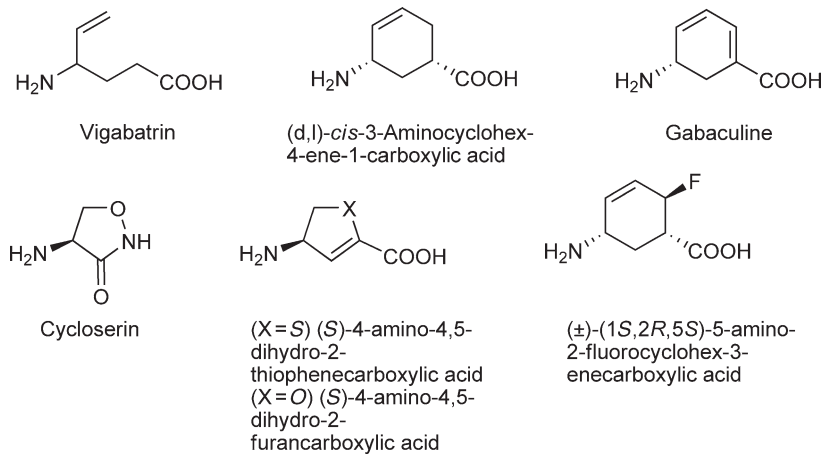
electrophilic to be attacked by the lysine residue (pathway a) and the nucleophilic attack takes place on the aldimine carbon, leading to a standard transaldimination process, and to the formation of an activated nucleophile aminoacrylate. On the contrary, when two fluorine substituents are present on the  $\beta$ -carbon, the lysine residue could perform a Michael addition on the highly electrophilic difluoroeneamino terminus (pathway b). Therefore, the selection of the mechanism reaction is essentially guided by the different electrophilicity of fluoride or hydrogen substituent at the  $\beta$ -carbon of the eneamino-PLP complex.<sup>426</sup>

More recently, an alternative inactivation mechanism of KAPAS (AON synthase) by trifluoroalanine has been proposed (Figure 40).<sup>424</sup> The authors suggested that the decarboxylation and the loss of fluoride occur immediately after the aldimine formation. The derived 2,3-unsaturated imino species is then attacked by the enzyme residue Lys236 and a water molecule, generating a gem hydroxyfluoro species which loses the last fluoride ion and forms a stable amide.

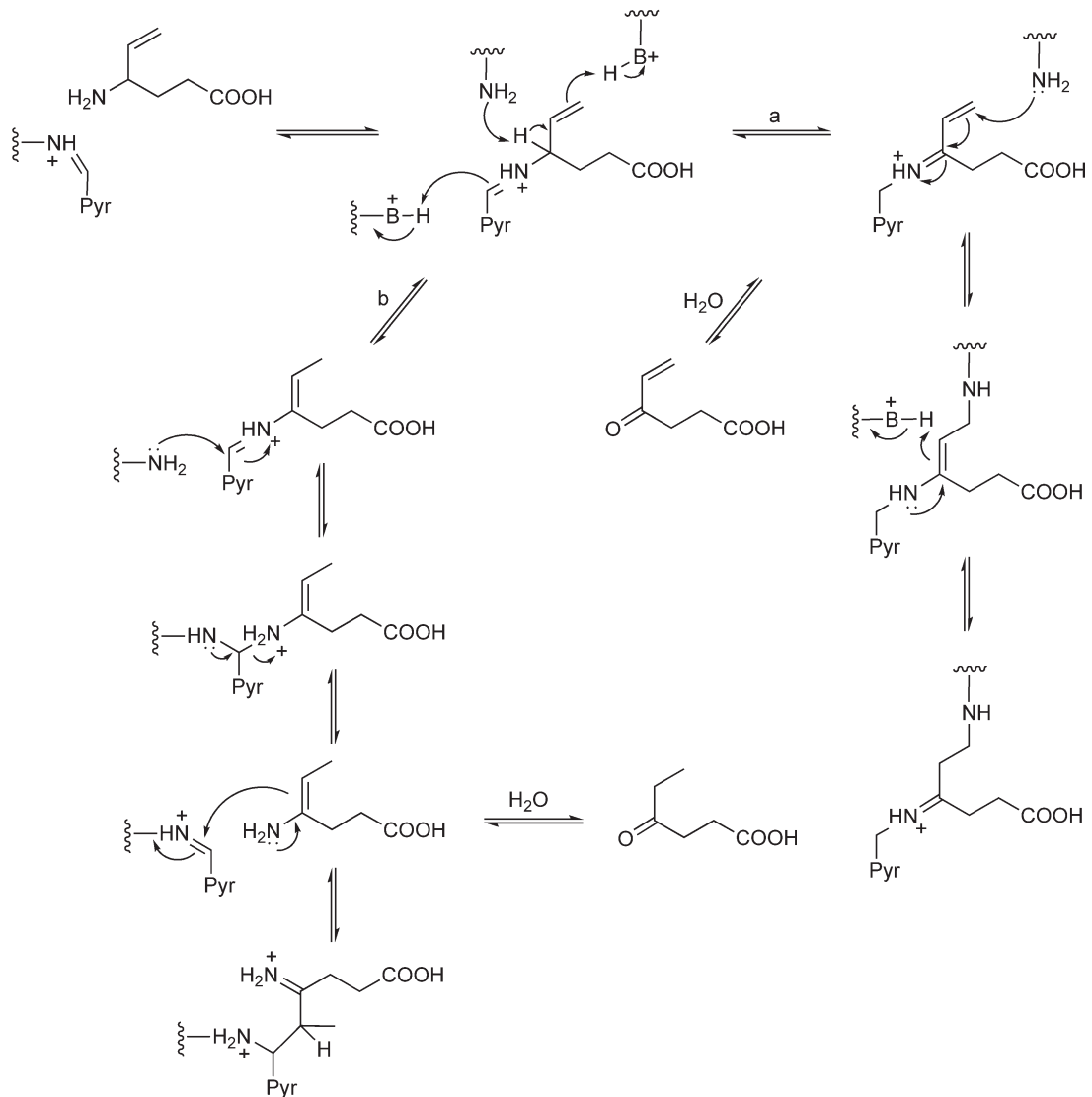
*The GABA-AT case.* Inhibition reactions of GABA-AT, caused by different compounds, follow the above reported mechanisms. Since low GABA levels demonstrated to be related to several pathologies, such as epilepsy,<sup>465,466</sup> Parkinson's disease,<sup>467</sup> Huntington's chorea,<sup>468</sup> Alzheimer's disease,<sup>469</sup> and tardive dyskinesia,<sup>470</sup> GABA-AT represents one of the most studied and investigated PLP-dependent enzymes. The well-known anticonvulsant agent vigabatrin (4-amino-5-hexenoic acid or  $\gamma$ -vinyl GABA),<sup>413,414</sup> acts by following two principal inactivation pathways (Figures 41 and 42): a Michael addition mediated by Lys329 (reaction a) and the formation of an enamine able to nucleophilically attack the internal aldimine (reaction b).<sup>471</sup>



**Figure 40** Mechanism of inactivation of 8-amino-7-oxononoate synthase by trifluoroalanine.<sup>424</sup>



**Figure 41** Chemical structures of GABA-AT inhibitors.



**Figure 42** Michael addition mechanism (reaction a) and enamine mechanism (reaction b) of inactivation of GABA-AT by vigabatrin.<sup>471</sup>

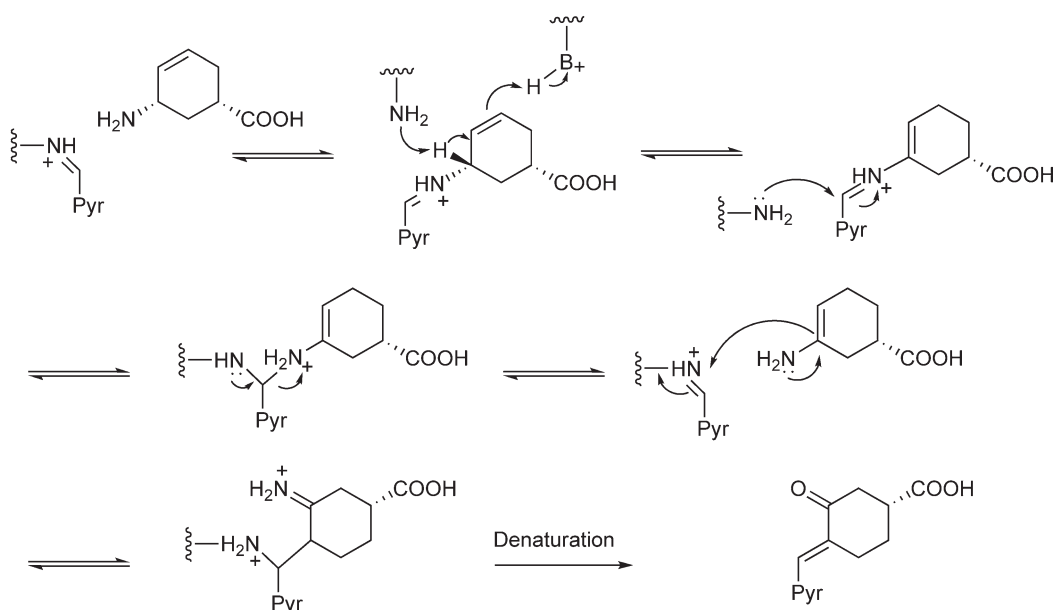
This mechanism is also followed by fluorine-containing conformationally restricted analogues of vigabatrin, based on a cyclopentanecarboxylic moiety, such as (1*R*,4*S*)-4-amino-3-trifluoromethylcyclopent-2-enecarboxylic acid.<sup>472</sup> The strategy of conformation restriction has been generally adopted in order to produce more potent and less degradable drugs, by stabilizing the biologically active conformer and concomitantly eliminating the metabolized species.<sup>434</sup> Thus, for example, the conformationally restricted vigabatrin analogue ((*D,L*)-*cis*-3-aminocyclohex-4-ene-1-carboxylic acid), inhibits GABA-AT by using only an enamine-based mechanism (**Figures 41 and 43**), exploiting the strategy of a stereospecific control.<sup>471</sup> The third aromatization-based mechanism is followed by other GABA-AT inhibitors, such as gabaculine<sup>448</sup> (**Figures 41 and 44**), cycloserin,<sup>473</sup> and fluorinated conformationally restricted vigabatrin analogues, as (*S*)-4-amino-4,5-dihydro-2-thiophenecarboxylic acid,<sup>474</sup> (*S*)-4-amino-4,5-dihydro-2-furancarboxylic acid,<sup>475</sup> and ( $\pm$ )-(1*S*,2*R*,5*S*)-5-amino-2-fluorocyclohex-3-enecarboxylic acid<sup>476</sup> (**Figure 41**). Recently, several five- or six-membered analogues have been developed in order to irreversibly inactivate GABA-AT, with potency even higher than vigabatrin.<sup>433,477</sup>

## 7.10.6.2 Alternative Strategies

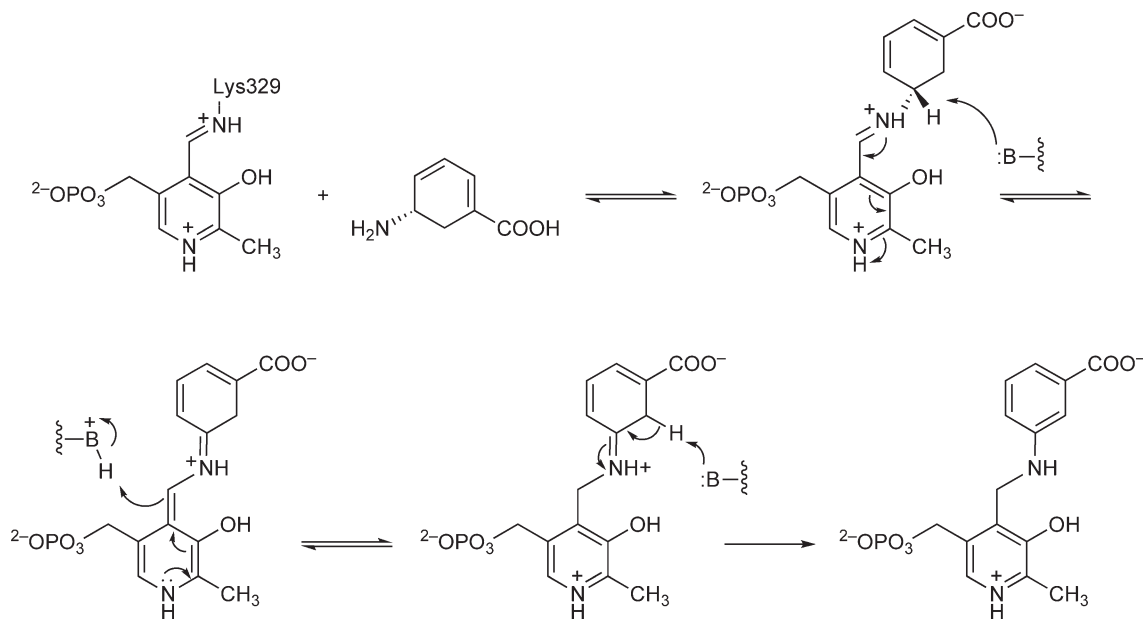
### 7.10.6.2.1 Noncovalent inhibitors

The development of molecules able to inhibit the target enzyme through noncovalent interaction represents one of the most investigated strategy, since it would avoid the occurrence of the adverse effects associated with an irreversible mechanism of inactivation. Nevertheless, at the moment, the class of compounds able to inhibit PLP-dependent enzyme in a reversible and noncovalent way is one of the less populated, in particular with respect to the category of mechanism-based inhibitor.

A noncovalent bond is formed by: (1) the CGS inhibitor 5-carboxymethylthio-3-(3'-chlorophenyl)-1,2,4-oxadiazole,<sup>416</sup> which binds to three active-site groups, thus disrupting the stacking of PLP with the aromatic side chain of a tyrosine residue, (2) maleate, which inhibits AAT by binding in the aspartate site,<sup>478</sup> (3) the competitive kynurenine aminotransferase inhibitor 4-(2-aminophenyl)-4-oxobutyric acid,<sup>479</sup> and (4) the antibiotic pantocinB able to inhibit *N*-acetylornithine aminotransferase (AcOAT) in a competitive way.<sup>480</sup> An interesting alternative to irreversible GABA-AT inhibition is represented by some taurine analogues such as ( $\pm$ )piperidine-3-sulfonic acid (PSA), 2-aminoethylphosphonic acid (AEP), ( $\pm$ )2-acetylamincyclohexane sulfonic acid (ATAHS), and 2-aminobenzenesulfonate (ANSA), which exploit a mixed-type inhibition, or a pure competitive strategy. In the



**Figure 43** Enamine mechanism of inactivation of GABA-AT by (*D,L*)-*cis*-3-aminocyclohex-4-ene-1-carboxylic acid.<sup>471</sup>

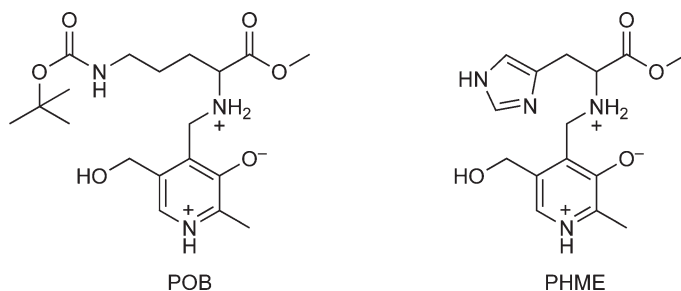


**Figure 44** Aromatization mechanism of inactivation of GABA-AT by Gabaculine.<sup>434</sup>

first case, they presumably induce conformational changes in the enzyme structure, whereas in the second case they bind directly into the active site, but they do not possess reactive groups such as the vinylic vigabatrin moiety, which can lead to an irreversible inhibition.<sup>481</sup> The strategy of noncovalent inhibition has also been exploited for the development of compounds able to bind to the  $\alpha$ -subunit active site of TRPS. Thus, a series of indole-acetyl amino acids and a series of arylthioalkylphosphonate compounds,<sup>482</sup> analogues of the substrate transition state, were synthesized and demonstrated to behave as good inhibitors with micromolar  $IC_{50}$ s.<sup>483</sup> The crystallographic structures of several TRPS-inhibitor complexes were determined.<sup>484</sup>

#### 7.10.6.2.2 Coenzyme-substrate analogues

The design of transition-state inhibitors was exploited for the development of new human decarboxylase coenzyme-based inhibitors. These coenzyme-substrate analogues are synthesized as precursors, subsequently activated through intracellular phosphorylation by pyridoxal kinase and hydrolysis of the carboxymethyl ester by endogenous esterase. The derived phosphopyridoxyl-amino acids analogues cannot be further transformed and are thus able to potently and specifically inhibit the newly synthesized apo-enzymes. In particular, the ODC inhibitor POB (BOC-protected pyridoxyl-ornithine conjugate; **Figure 45**) demonstrated to inhibit tumor cells proliferation better than the well known  $\alpha$ -difluoromethylornithine,<sup>485</sup> whereas PHME (pyridoxyl-histidine methyl ester conjugate; **Figure 45**) specifically inhibits HDC, without affecting the activity of human ODC and cell viability.<sup>486</sup>



**Figure 45** Coenzyme-substrate analogues.



### 7.10.6.2.3 Allosteric inhibitors

Among the category of nonmechanism-based inhibitors, the allosteric effectors represent a potentially interesting alternative for several PLP-dependent enzymes, even if very few allosteric effectors are currently available. G418 sulfate (Geneticin) behaves as a weak noncompetitive inhibitor of ODC. It binds at the junction between the  $\beta$ -sheet and the  $\alpha/\beta$ -barrel domain and inducing a significant disorder in a loop region at the dimer interface, where fundamental catalytic residues are located.<sup>31</sup> Analogously, five compounds recently identified by high-throughput screening seem to inhibit the bacterial cystathionine  $\beta$ -lyase both in the free and substrate-bound form. Therefore, they bind to a site different from the catalytic site.<sup>460</sup> The design of allosteric effectors has also been suggested as a possible strategy to enhance the activity of CBS and inhibit OASS-A.<sup>484</sup> Another possibility to inhibit OASS could be represented by the design of molecules able to interfere with the formation of the regulatory complex of the enzyme with SAT.<sup>484</sup> These compounds are generated from various available crystallographic structures and from the observation that the C-terminal decapeptide of SAT acts similar to competitive inhibitor of OASS-A.<sup>487</sup> Finally, the inhibition of the enzyme dimerization has been proposed, even if not yet applied, as a possible different way to inactivate AlaR.<sup>488</sup>

## 7.10.7 Role of Conformational Changes in Substrate and Reaction Specificity and Allosteric Regulation

The protein matrix confers substrate and reaction specificity and greatly enhances the catalytic efficiency of enzymes, including PLP-dependent enzymes. Dynamic processes contribute to these effects. For many PLP enzymes, substrate binding, and/or the formation of specific reaction intermediates have been shown to induce conformational transitions, or to selectively stabilize alternative conformations, that play a major role in controlling the efficiency and specificity of the catalyzed reaction. In some cases, allosteric effectors, intersubunit communications or protein-protein interactions influence catalysis by long-range dynamic effects. We report a few exemplary cases: AAT, as a paradigm case of how ligand binding stabilizes a closed conformation entailing large domain movements, and TRPS and OASS. For TRPS, distinct communication pathways are involved in different allosteric effects, as intersubunit activation and ligand-induced regulation, exploiting conformational transitions between open and closed states to regulate enzyme activity. In the case of OASS, similar conformational dynamics appear to underlie a higher level of functional regulation, that is brought about by interprotein interactions with serine acetyl-transferase, the preceding enzyme in L-cystein biosynthesis. Furthermore, many enzymes belonging to the fold-type II contain regulatory domains that bind diverse allosteric effectors (see Eliot and Kirsch<sup>19</sup> and references therein), such as S-adenosylmethionine (AdoMet) for cystathionine  $\beta$ -synthase<sup>315</sup> and plant TS<sup>366,368</sup> or amino acids with feedback regulatory effects, for threonine deaminase.<sup>489</sup> In TS from *Arabidopsis thaliana*, binding of AdoMet to two sites at the dimer interface triggers a large reorganization of active-site loops in one monomer of the structural dimer and allows the displacement of PLP to its active conformation.<sup>368</sup> Binding of AMP or CMP to biodegradative threonine deaminase from *Salmonella typhimurium* triggers the tetramerization of the enzyme, that is otherwise in a dimeric form. Conformational changes induced by binding of the ligand nucleotides at the dimer interface appear to be essential for tetramerization, and differences in the tertiary and quaternary structure in the absence and the presence of CMP probably account for the observed increase of substrate affinity and enzyme activation.<sup>490</sup>

### 7.10.7.1 Aspartate Aminotransferase and the Open/Closed Transition in PLP-Dependent Enzymes

Aspartate aminotransferase (AAT) is the first PLP-dependent enzyme for which the three-dimensional structure has been determined<sup>10,141,142</sup> and is the prototype of fold-type I PLP-enzymes. Each subunit of the AAT homodimer has a large and a small domain. The coenzyme is bound to the large (N-terminal) domain and located in a pocket at the subunit interface, so that residues from each monomer contribute to the formation of both active sites. The proximal and distal carboxylate group of the dicarboxylic substrates bind to Arg386 and Arg292, respectively, the latter contributed by the opposite subunit.<sup>142</sup> Early crystallographic structures

determined for the unligated form, and for the complex with substrate analogues,<sup>16,142</sup> indicated that ligand binding induces a bulk movement of the small domain, completely reverted upon formation of the pyridoxamine form of the coenzyme (see Section 7.10.2.4). Structural changes occur within the small domain itself: an helix in mitochondrial AAT, besides moving with the core of the small domain, presents an additional shift and rotation that closes the active-site entrance.<sup>151</sup> The equilibrium between the crystallographically detected 'open' and 'closed' conformations of AAT was probed in solution with several techniques, including peptide hydrogen–deuterium exchange,<sup>491</sup> fluorescence spectroscopy,<sup>492</sup> and limited proteolysis.<sup>493</sup> This was the first example of what appears to be a very common feature in PLP-dependent enzymes, the regulation of activity by substrates, inhibitors, and allosteric effectors through modulation of the equilibrium between two conformational states: an open conformation, endowed with low, if any, catalytic activity, and a closed, catalytically active conformation. AATs from different sources differ for the scope of the underlying structural dynamics: in subgroup Ia AATs (from yeast, chicken, pig, *E. coli*, etc.)<sup>494</sup> the small domain as a whole exhibits a large-amplitude rigid-body rotation (e.g., 12–14° for chicken mitochondrial AAT,<sup>492</sup> 5–6° in *E. coli* AAT,<sup>495,496</sup> whereas in subgroup Ib AATs (e.g., *Thermus thermophilus* HB8 AAT) only the N-terminal region of the small domain approaches the active site to interact with the substrate<sup>497,498</sup>).

A list of PLP-dependent enzymes for which open and closed conformations have been crystallographically detected includes human kynureninase,<sup>499</sup> SHMT,<sup>104,106</sup> OASS,<sup>293,298</sup> TRPS,<sup>500</sup> bacterial TS,<sup>367</sup> and TPL.<sup>204</sup> In many other cases the existence of open and closed conformations has been inferred on the basis of functional and spectroscopic data. However, there is evidence that this is not a universal attribute of PLP-dependent enzymes, many of which do not appear to require a closed structure (see Jansonius<sup>16</sup> and references therein). For instance, despite structural similarity to TRPS and OASS, open/closed transitions were excluded for yeast TS, based on the presence of additional structural elements fixing the PLP-binding domain to the C domain.<sup>363</sup> Different from the plant enzyme, yeast TS is not activated by AdoMet and is active in the monomeric form.

An open enzyme conformation facilitates substrate access to the active site and product release, whereas a closed conformation could maximize enzyme–substrate interactions by improving contact with the distorted substrate conformation in the transition state, provided that the free energy gain of improved interactions exceeds the cost of the enzyme conformational change.<sup>47</sup> Moreover, the stabilization of the closed conformation is expected to enhance substrate and reaction specificity by restricting the bound substrate to a specific orientation, optimizing the positioning of catalytic residues to promote the desired reaction and limit side-reactions, and protecting catalytic intermediates from the solvent.<sup>19,118,501</sup> For instance, the closure of the active-site crevice of AAT upon ligand binding prevents water molecules to access the quinonoid intermediate, thus preventing protonation of the C $\alpha$  from the wrong (*re*) face and limiting the racemization side reaction.<sup>150</sup> It has been argued that the conformational changes induced by substrate binding do not necessarily account for an increased substrate specificity.<sup>19,20,502,503</sup> Possible exceptions occur when the conformational change results in the substrate being surrounded on all sides by the enzyme, and when a substrate binding or product release step is rate limiting for a good substrate, whereas a chemical step is rate limiting for an alternative poor substrate.<sup>502</sup> The latter is at least partially the case for the release of the product oxaloacetate upon reaction of AAT with aspartate and  $\alpha$ -ketoglutarate.<sup>504</sup>

### 7.10.7.2 Tryptophan Synthase

A key feature of TRPS is the presence of  $\alpha$  and  $\beta$  active sites, separated by about 20 Å, catalyzing two successive reactions (see Section 7.10.3.2.1). To ensure that the overall catalytic process is productive and efficient, cross talks between sites are mandatory. Two distinct pathways of intersubunit communication have been elucidated and found to be endowed with specialized roles.<sup>25</sup> The first one is predominantly associated with the reciprocal subunit activation, that is the catalytic activities of the complex is 30–100-fold higher than those in the isolated subunits. This communication involves the so-called COMM domain of the  $\beta$ -subunit that interacts with  $\alpha$  loop 2, containing  $\alpha$ -active-site residues.<sup>278,505</sup> The second pathway is predominantly involved in the allosteric regulation and is based on the interaction between  $\alpha$ loop6 and the COMM domain. Ligands of one subunit control the open–closed conformation equilibrium of the other subunit. The occupancy of the  $\alpha$ -active site by  $\alpha$ -subunit allosteric ligands triggers the stabilization of the closed state of the  $\alpha$ -subunit through the formation of a single hydrogen bond between the NH of  $\alpha$ Gly181 and the carbonyl oxygen of

$\beta$ Ser178.<sup>278,281–283</sup> In the open state, the  $\alpha$ loop 6 is undetectable by X-ray methods due to its mobility. Molecular dynamics simulations were carried out on the wild type and mutants, in the absence and presence of  $\alpha$ -subunit ligands to simulate the open–closed conformations, thus suggesting an explanation of the reduced activity of the  $\alpha$ -active site for some of the mutants.<sup>284</sup>

### 7.10.7.3 O-Acetylserine Sulfhydrylase

The structure of the functional homo-dimer of OASS has been determined in different conformations, in the absence and presence of substrate analogues and allosteric effectors.<sup>287,293,298</sup> Binding of L-cysteine or of other substrate analogues is followed by the formation of the external aldimine of PLP that is, in turn, associated with a conformational change from an open (inactive) to a closed (active) structure. This structural transition is probably triggered by the binding of the  $\alpha$ -carboxylate of the amino acid substrate to an  $\alpha$ -carboxyl subsite in the enzyme active site.<sup>506</sup> Furthermore, binding of chloride to an allosteric anion-binding site, located at the dimer interface, stabilizes a conformation that differs both from the open conformation of the internal aldimine and the closed one of external aldimine.<sup>298</sup> This state inhibits the formation of the external Schiff base and the subsequent chemistry. It was suggested that the chloride ion may behave as an analogue of sulfide, the physiological inhibitor of OASS.

OASS plays a key role in sulfur metabolism, controlling the cysteine biosynthetic pathway both directly and indirectly. A direct control is exerted through the formation of bi-enzyme complexes with ATP sulfurylase and SAT. The latter complex, grouping the last two enzymes in cysteine biosynthesis, is usually referred to as cysteine synthase. The characterization of the effects of the cysteine synthase complex on the conformational and functional properties of OASS have been investigated exploiting the fluorescence properties of PLP.<sup>307</sup> Binding of SAT triggers the transition from an open to a closed conformation, indicating that SAT can inhibit the catalytic activity of OASS in two ways, by competing with the substrate OAS for binding to the enzyme active site, and by stabilizing a closed conformation that is less accessible to the natural substrate.

## 7.10.8 Genomics, Evolution, Structure, and Reactivity/Specificity

As outlined in the previous chapters, the functional versatility of PLP-dependent enzymes is remarkable and unrivaled among enzymes that depend on organic cofactors. Such versatility is illustrated by the fact that, as of December 2007, the Enzyme Commission<sup>36</sup> had catalogued 161 PLP-dependent activities, corresponding to over 4% of all classified activities. To this total, one should add a substantial number of activities that have been described but not yet registered.<sup>13</sup>

The functional diversity of PLP-dependent enzymes elicits a series of questions. How early did these enzymes appear in the course of evolution? How did their impressive functional diversification come about? What are the general relationships between structure and function within this group of enzymes? And, can sequence or structural information be used to infer function?

### 7.10.8.1 How Old is PLP?

PLP-dependent enzymes are ubiquitous in biology. A substantial number of genes coding for PLP-dependent enzymes are found in all free-living organisms whose genomes have been sequenced to date and, actually, the fraction of such genes is greater in prokaryotes with small genomes.<sup>13</sup> The universal diffusion of PLP-dependent enzymes indicates that the metabolic importance of PLP was well established in the last common ancestor of Eubacteria, Archaea, and Eukarya. This inference is buttressed by the recent recognition that the main metabolic pathway for the *de novo* biosynthesis of PLP, based on just two enzymes termed Pdx1 and Pdx2, is remarkably conserved in plants, fungi, Archaea, and most bacteria.<sup>1,507</sup> The conservation of this biosynthetic route among modern-day organisms (the only notable exceptions are animals (which are unable to synthesize PLP *de novo*) and proteobacteria of the gamma subgroup such as *E. coli* (which are postulated to have originally lost the pathway for PLP biosynthesis, only to develop, in

recent times, a new and more complex biosynthetic route)) suggests, at the very least, that the pathway originated before the separation between the three kingdoms of life.

PLP itself, however, is presumably much older than this. There are general arguments suggesting that organic cofactors and coenzymes represent 'biochemical fossils' from very primitive stages in the history of life.<sup>508</sup> In particular, the fact that several cofactors show nucleotide-like features is often used to support the occurrence of an 'RNA world,' that is, a very early phase of biotic evolution in which RNA molecules were capable of self-replication and of a rudimentary form of metabolism.<sup>509,510</sup> Within this hypothetical world, cofactors and coenzymes would have helped expand the chemical repertoire of catalytic RNAs.<sup>511</sup>

The general arguments about the antiquity of cofactors apply to PLP. The nonenzymatic synthesis of pyridoxal under prebiotic conditions is considered possible,<sup>512,513</sup> whereas the presence of a 5' phosphate group could hint to an ancestral attachment of the cofactor to RNA molecules.<sup>511</sup> Furthermore, there are specific grounds to assume that PLP arrived on the evolutionary scene before the emergence of proteins. In fact, in current metabolism, PLP-dependent enzymes play a central role in the synthesis and interconversion of amino acids, and thus they are closely related to protein biosynthesis. In an early phase of biotic evolution, free PLP could have played many of the roles now fulfilled by PLP-dependent enzymes, since the cofactor by itself can catalyze (albeit at a low rate) reactions such as amino acid transaminations, racemizations, decarboxylations, and eliminations.<sup>514,515</sup> This suggests that the appearance of PLP may have preceded (and somehow eased) the transition from primitive 'RNA-based' life forms to more modern organisms dependent on proteins.

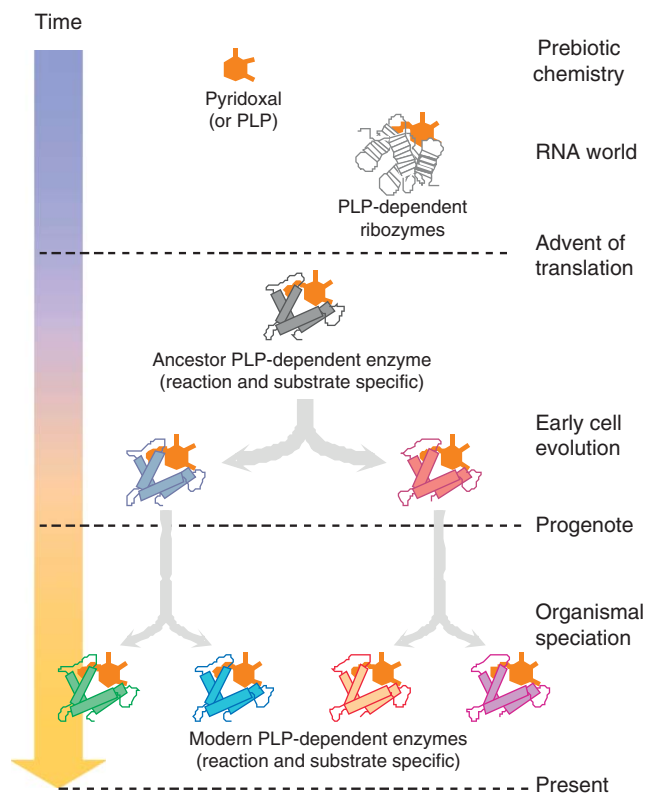
### 7.10.8.2 How Did PLP-Dependent Enzymes Emerge and Evolve?

PLP-dependent enzymes, as we know them today, could have started to appear and to differentiate after the 'invention' of translation. A few years ago, Christen and Mehta<sup>516</sup> provided a neat overview of the emergence and differentiation of these biocatalysts over the course of evolution, as can be inferred from the structural and functional properties of the extant enzymes.<sup>12</sup> More recent studies have only integrated, but not revolutionized, the picture provided by those authors.

Christen and Mehta suggested that the progenitor PLP-dependent enzyme(s) could be endowed with very minimal features: a binding site for PLP (which was presumably linked through a Schiff base to a Lys residue (Covalent linkage of the cofactor to a Lys residue is observed in the vast majority of PLP-dependent enzymes, suggesting a very ancient origin of this feature and a strong selective pressure for its conservation. One simple explanation may be that covalent binding prevents the dissociation and rapid loss of the cofactor, especially from active sites that are highly exposed to the solvent.<sup>517</sup> In many reactions, the colocalization of PLP with a flexible and ionizable Lys side chain can also represent a significant mechanistic advantage. However, this is not a strict prerequisite for catalysis, since there are a few PLP-dependent enzymes (acting on amino substrates) that lack an active-site lysine and that do not bear a covalently bound cofactor.<sup>406,518</sup> Conversely, glycogen phosphorylases bind PLP through an internal Schiff base, even though they do not use the pyridoxal carbonyl group or the lysine side chain for catalysis<sup>519</sup>)) and an adjacent, rudimentary binding site for an amino acid.<sup>12,516</sup> The two adjacent sites might have been enough to allow and favor formation of the external aldimine, which is the obligatory starting point for all reactions catalyzed by PLP-dependent enzymes on amino acid substrates. Such a postulated ancestral enzyme would have shown a low substrate specificity (binding many amino acids) and it would also have been catalytically promiscuous, that is, it would have carried out most, if not all, of the reaction types already accessible to free PLP.

Subsequent optimization of the noncovalent interactions between the protein and the PLP-substrate adduct may be assumed to have led to (1) a specialization for reaction type, (2) a specialization for substrate specificity, and (3) an improved catalytic efficiency. As discussed below, it is probable that differentiation of the ancestor enzyme(s) for reaction specificity preceded to some extent the acquisition of substrate specificity, even though the two types of specialization have kept occurring in parallel for the whole evolutionary history of PLP-dependent enzymes (**Figure 46**).

Achieving reaction specificity in PLP-catalyzed reactions means, in the first place, favoring the selective cleavage of one of the three bonds that connect the amino acid C $\alpha$  to the  $\alpha$ -proton, to the  $\beta$ -carbon, or to the carboxylate group. According to Dunathan's hypothesis<sup>11</sup> the bond to be cleaved has to be oriented perpendicular to the plane of PLP, to maximize overlap with the  $\pi$  system of the cofactor and stabilize the



**Figure 46** A schematic overview of the molecular evolution of PLP-dependent enzymes. In prebiotic times and in very early biotic systems, free pyridoxal (or PLP) presumably have catalyzed a number of reactions that are performed today by PLP-dependent enzymes. The occurrence of PLP-dependent ribozymes within the postulated ‘RNA world’ has also been suggested.<sup>511</sup> Subsequent to the invention of translation, PLP-dependent enzymes started to appear and to evolve. For simplicity, the development of only one structural family is shown, but since the main fold-type groups of PLP-dependent enzymes presumably had independent origins, the course of events outlined here may have occurred in parallel more than once, each time starting from an ancestor protein with a different fold. The ancestor PLP enzyme (presumably an allrounder catalyst) originated the lineages of reaction and substrate-specific enzymes that occur in recent species. Similarly, many PLP-dependent enzymes (including several aminotransferases and amino acid synthases) were already completely specialized for reaction and substrate specificity before the three biological kingdoms branched off from each other 1500 to 1000 million years ago.<sup>12,516</sup> Such an early specialization is consistent with the conclusion that in the progenote (the last universal ancestor of contemporary cells) the major metabolic pathways were already established.<sup>520</sup> In several instances, however, specialization of the enzymes for substrate and even for reaction specificity seems to have overlapped with organismal speciation.

negative charge that accumulates during cleavage. Extant PLP-dependent enzymes achieve their reaction specificity by modulating, through specific interactions, the geometry around  $C\alpha$ , so as to favor the preferential cleavage of one particular bond.<sup>17</sup> Thus, the functional diversification of the progenitor PLP-dependent enzyme(s) must have entailed the development of active sites where only specific geometries were allowed, limiting catalytic promiscuity and augmenting specificity. Further and later modifications of the active site may have led to specialization for reaction types involving the side-chain carbons ( $\beta$ - and  $\gamma$ -eliminations and replacements).

Indeed, an analysis of present-day PLP-dependent enzymes is fully consistent with the idea that substantial differences in catalytic specificity can be achieved through relatively small alterations of the same structural scaffold. In fact, despite their functional variety, all structurally characterized PLP-dependent enzymes belong to just six or seven distinct structural groups,<sup>15,17</sup> (520 missing)<sup>393,521</sup> which are likely to represent independent evolutionary lineages.<sup>12,516</sup> In contrast to earlier suggestions,<sup>14</sup> it is now appreciated that each of the major structural classes contains representatives of multiple reaction types (**Table 1**) even though some limitations



are apparent. For example, in all known fold-type II enzymes, formation of the external aldimine is invariably followed by removal of the  $\alpha$ -proton, so that these enzymes do not catalyze decarboxylation or side-chain cleavage reactions.

### 7.10.8.3 A Case Study: Fold-Type I Enzymes

Examining the evolutionary history of fold-type I enzymes seems especially helpful to envisage how functional specialization developed.<sup>12,516</sup> This group of PLP-dependent enzymes is by far the largest and the most functionally diverse (Figure 47). It is also the most diffused in modern-day organisms: an examination of about 20 complete genomes of free-living species, from methanogens to man, showed that only two PLP-dependent enzymes were common to all these organisms, and both were fold-type I enzymes (AAT and SHMT).<sup>13</sup> Collectively, these observations suggest that the fold-type I group is probably the most ancient lineage of PLP-dependent enzymes.

Christen and Mehta<sup>516</sup> reconstructed a pedigree of these enzymes, uncovering the pattern of evolutionary events that underlies the present functional specialization.<sup>12</sup> The pedigree of fold-type I proteins (Figure 47) suggests that the ancestor enzymes may have been specific for reactions with covalency changes limited to C $\alpha$ , in particular transaminations and decarboxylations. One of the earliest evolutionary events seems to have been the separation between a lineage of catalysts devoted to decarboxylations and another lineage mostly specialized for transamination reactions. This second lineage further diverged into reaction-specific subgroups, such as three main classes of aminotransferases (ATI, ATII, and ATIV), a group of CoA-dependent acyltransferases (such as ALAS), some evolutionary branches mostly specialized in  $\beta$ - or  $\gamma$ -eliminations and replacements.<sup>12,516</sup>

Within each of these sublineages, the type of catalyzed reaction was mostly conserved, whereas the substrates varied. For example, all the enzymes in the ATII subfamily are aminotransferases, with the only apparent exception of dialkylglycine decarboxylase, which, however, catalyzes a decarboxylation-dependent transamination (the enzyme cannot directly transaminate its substrate, which lacks an  $\alpha$ -proton, but proceeds to decarboxylate it and then catalyzes a transamination with the decarboxylation product).<sup>46</sup> Hence, the evolutionary tree shows that, in general, specialization for reaction type came first, whereas the last and shortest phase of the evolutionary history involved specialization for substrate specificity.

To explain this apparent order of events, Christen and Mehta<sup>516</sup> put forward two distinct rationales, one metabolic and one mechanistic.<sup>12</sup> Metabolically, they argued that, in a primitive organism, a reaction-specific but substrate-aspecific enzyme (i.e., a catalyst that accelerates one particular type of reaction for multiple substrates) could provide more evolutionary advantages than a substrate-specific but reaction-aspecific enzyme (Note that a substrate-aspecific enzyme is not necessarily an enzyme with low affinity for its substrates. A nonstrict substrate specificity seems even an inherent property of PLP-dependent enzymes such as the aminotransferases, which often must bind, for example, an aromatic or basic amino acid/oxo acid pair in the first half-reaction, and then accept glutamate and 2-oxoglutarate in the second half of the reaction.<sup>19</sup> The facility to interact with different ligands is illustrated by the rather broad substrate specificity of many extant aminotransferases.<sup>522–525</sup>). Mechanistically, they suggested that a change in substrate specificity might entail only minor alterations in an active site (consistent with the relatively short evolutionary times needed to achieve this specialization), whereas a change in reaction type could require more extensive structural adaptations.

It is clear, however, that in some lineages, specialization for (or change of) reaction specificity has occurred till recent times, overlapping with specialization for substrate. This is well exemplified by the case of L-threonine-0-3-phosphate decarboxylase (PThr decarboxylase; Figure 47), a bacterial enzyme involved in cobalamin biosynthesis which belongs to the ATI evolutionary branch and is most closely related to histidinol phosphate aminotransferases.<sup>526</sup> ACC synthase (Figure 47) is another enzyme belonging to the ATI lineage (and most similar to aspartate and aromatic aminotransferases) which nevertheless catalyzes an  $\alpha,\gamma$ -elimination reaction.<sup>527</sup>

At a fundamental level, the possibility of evolving, within the same phylogenetic sublineage, enzymes that carry out different reaction types (and, conversely, the evolution of enzymes with the same reaction specificity within completely different lineages) seem intrinsic to the versatility of PLP itself and to the common features of the reactions they catalyze, whose mechanisms include, almost invariably, the formation of carbanionic intermediates stabilized by the cofactor.<sup>18</sup>



#### 7.10.8.4 Exploring the Modern PLP-Dependent Enzymes: Can Sequence be Used to Infer Function?

The current flood of genomic sequences has offered new opportunities to detect and analyze the distribution of enzyme families within and between genomes. Among other things, all this sequence information helps outline a picture of the appearance and evolution of enzyme functions, as described above. The identification and classification of PLP-dependent enzymes in complete genomes seems particularly favored: structurally, these enzymes belong to a limited number of fold groups, which facilitates their recognition. Functionally, they have been studied for decades and the types of reactions they can perform are well understood. Despite this, there are significant limitations to functionally assign these enzymes based on sequence.

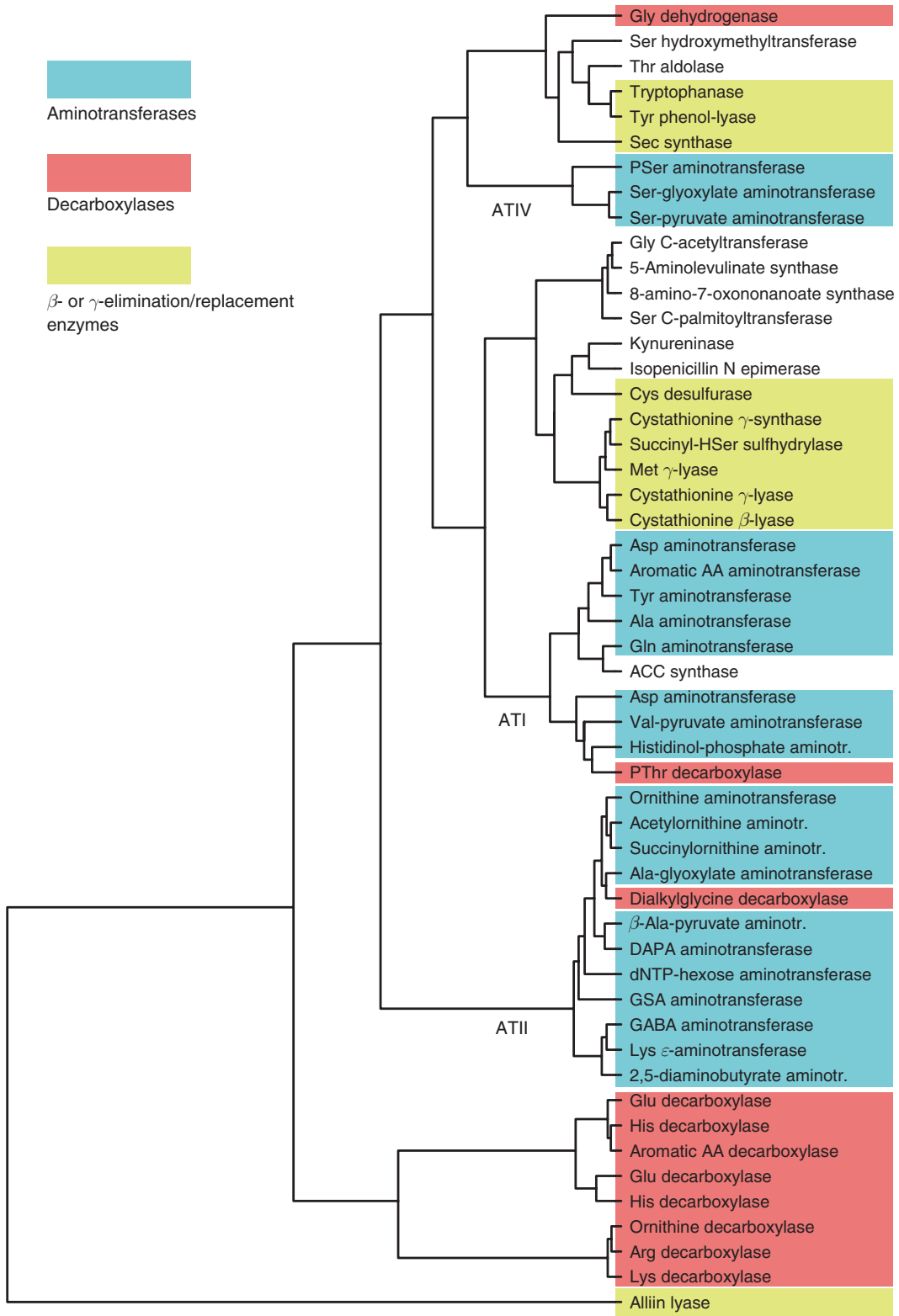
In fact, while homology criteria are well-suited for detecting structural similarities and drawing phylogenetic relationships, they are much less reliable for establishing protein function, even when sequence identity is >50%.<sup>528</sup> The problem seems particularly severe for PLP-dependent enzymes, for reasons that have in part been anticipated above. First, as just noted, the constant mechanistic features of PLP catalysis can ease both convergent evolution (i.e., the appearance of enzymes with identical or very similar activities along independent lineages) and divergent evolution (i.e., the evolution of enzymes with very different activities from a common ancestor). Second, the apparently facile changes in substrate specificity (Figure 47) complicate the sequence-based assignment of a given PLP-dependent enzyme to a given substrate. Finally, it is now appreciated that PLP-dependent enzymes are involved in a surprising variety of cellular processes, so that even for an enzyme whose catalytic activity can be assigned with good confidence based on sequence, the actual biological function may remain uncertain.

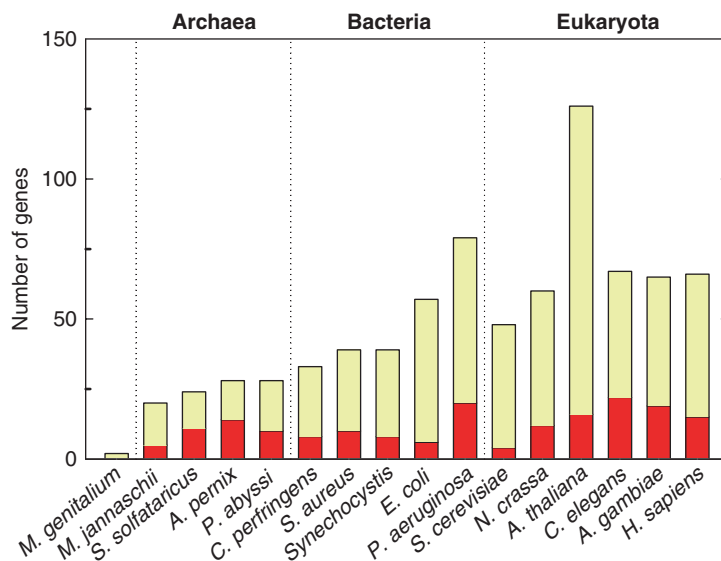
Enzymes that possess the same catalytic activity but are evolutionarily unrelated and play very diverse biological functions can be found, for example, in a group of PLP-dependent enzymes with desulhydrase activity, that is, enzymes that release H<sub>2</sub>S from thiol amino acids. Bacteria employ desulhydrases not only in the metabolism of sulfurated amino acids and in the adaptation to new nutrient sources,<sup>529,530</sup> but also, sometimes, as virulence factors.<sup>531,532</sup> An *E. coli* enzyme with desulhydrase activity is even known to act as a modulator of gene expression, although this function seems to be unrelated to catalysis.<sup>533</sup> Sulfide production by PLP-dependent enzymes is also important in vertebrates, where H<sub>2</sub>S has been shown to act as a neuromodulator.<sup>534</sup>

PLP-dependent desulhydrases necessarily show very similar mechanisms, but often come from independent evolutionary lineages. For example, although most bacterial L-cysteine desulhydrases are fold-type I enzymes belonging to the same evolutionary branch as cystathionine β- and γ-lyases, L-cysteine desulhydrase from *Fusobacterium nucleatum* is a member of the fold-type II group and its closest sequence homologue is a cysteine synthase.<sup>532</sup> D-cysteine desulhydrase from *E. coli* is also a fold-type II enzyme; not strictly related to other desulhydrases but resembling instead an ACC deaminase.<sup>529</sup>

The difficulties in establishing function on the basis of sequence similarity explain why a substantial fraction of the genes that have been identified as encoding PLP-dependent enzymes remain functionally unclassified or only tentatively classified, even in the cases of model organisms. For example, in 2003, Percudani and Peracchi<sup>13</sup> performed a search of genes coding for PLP-dependent enzymes in a series of complete genomes, concluding that enzymatic function could not be reliably assigned in up to 50% of the cases (Figure 48). In many cases, these genes could code for 'novel' PLP-dependent enzymes, that is, enzymes possessing activities not yet described or characterized.

Indeed, over the past few years, a number of such 'novel' enzymes have been identified and characterized at the molecular level. These include several aminotransferases, involved in the metabolism of NTP-linked amino sugars;<sup>517,518,535–538</sup> a yeast enzyme with L-threo-3-hydroxyaspartate dehydratase activity;<sup>539</sup> L-cysteate sulfo-lyase, a bacterial enzyme that degrades cysteic acid;<sup>540</sup> phenylacetaldehyde synthase (a specialized tyrosine decarboxylase from plants, that is capable of stoichiometrically coupling the decarboxylation reaction with the further conversion of tyramine into phenylacetaldehyde);<sup>541</sup> a metazoan aspartate racemase;<sup>542</sup> a mammalian enzyme with phosphothreonine phospho-lyase activity;<sup>543</sup> and finally L,L-diaminopimelate aminotransferase, an enzyme involved in L-lysine biosynthesis in plants and *Chlamydia*.<sup>544</sup> Overall, these studies confirm that the functional versatility of PLP-dependent enzymes is far from being fully explored, and stress that the 'functional genomics' of these enzymes remain a challenge in the future.





**Figure 48** The total number of PLP-dependent genes in a set of representative genomes (Reproduced with permission from R. Percudani; A. Peracchi, *EMBO Rep.* 2003, 4 (9), 850–854); the fraction of sequences that could not be unambiguously classified in terms of function is shown in red. The organisms shown are: *Mycoplasma genitalium* (not a free-living organism), *Methanococcus jannaschii*, *Sulfolobus solfataricus*, *Aeropyrum permix*, *Pyrococcus abyssi*, *Clostridium perfringens*, *Staphylococcus aureus*, *Synechocystis* sp. PCC 6803, *Escherichia coli*, *Pseudomonas aeruginosa*, *Saccharomyces cerevisiae*, *Neurospora crassa*, *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Anopheles gambiae*, and *Homo sapiens*. The relatively high number of PLP-dependent genes in *A. thaliana* can be explained in part by the occurrence, for many enzymes, of at least three isoforms (cytoplasmic, mitochondrial, and plastidic) encoded by different genes.

In 1995, Grishin *et al.*<sup>15</sup> classified the PLP-dependent enzymes whose structures were then available into five structurally unrelated groups (or ‘fold types’), distinguished by roman numerals.<sup>17</sup> It is now well recognized that each of the four major structural groups (fold-types I–IV) contains representatives of multiple reaction types; the examples shown here are limited (when possible) to reactions involving amino acid substrates.

To the original classification, we have added two minor groups of enzymes, for which structures have only recently been solved: ‘fold-type VI’ indicates a group whose prototype is lysine 5,6-aminomutase,<sup>521</sup> whereas ‘fold-type VII’ refers to a group whose prototype structure is represented by LAM.<sup>393</sup> The reactions catalyzed by both aminomutases do not seem to involve the cleavage of C $\alpha$  bonds, but require the formation of PLP-bound radicals, and exploit a second cofactor (cobalamin for D-lysine aminomutase; S-adenosyl-methionine for LAM). Although fold-types I–V are diffused in all kingdoms of life, enzymes with fold-types VI and VII have been described only in a small subset of bacteria.

**Figure 47** Evolutionary pedigree of 52 functionally validated fold-type I of enzymes (Reproduced with permission from P. K. Mehta; P. Christen, *Adv. Enzymol. Relat. Areas Mol. Biol.* 2000, 74, 129–184; P. Christen; P. K. Mehta, *Chem. Rec.* 2001, 1 (6), 436–447). Different background colors are used to highlight the major reaction types catalyzed by extant enzymes. The distinction of the fold-type I aminotransferases into three subgroups (ATI, ATII, and ATIV) is commonly adopted in the literature (a fourth group of aminotransferases, termed ATIII, is composed of fold-type IV enzymes). Note that some enzymes with the same activity may appear more than once in the pedigree, when they belong to clearly different lineages. For example, the subgroup ATI includes at least two distinct lineages of aspartate aminotransferase, one composed mainly by eukaryotic sequences, the other including exclusively prokaryotic sequences (eubacterial and archaeabacterial). The sequences of the first lineage appear to be more closely related to those of aromatic-amino-acid aminotransferase, tyrosine aminotransferase, and alanine aminotransferase than to the aspartate aminotransferases of the second group.<sup>12,516</sup>

## Acknowledgment

This work was supported by a grant from the Italian Ministry of University and Research (COFIN2007 to A. M.).

## Glossary

**apo-enzyme** an enzyme without a cofactor.

**catalyst** a chemical substance that increases the rate of a chemical reaction and is not consumed by the reaction itself unlike the substrates that participate in the chemical reaction.

**coenzyme** a subclass of cofactors, derived from hydrophilic vitamins, like pyridoxal 5'-phosphate derived from vitamin B<sub>6</sub>.

**cofactor** organic compounds or metal ions that are bound to an enzyme and are required for catalysis.

**enzymes** biomolecules that catalyze a specific chemical reaction, increasing the rate by at least 10<sup>6</sup> fold.

**holo-enzyme** an enzyme with bound cofactor.

**isoenzyme** multiple forms of enzymes that differ in the amino acid sequence.

**substrate** a molecule that binds to an enzyme active site and is transformed in products by its catalytic action.

## Abbreviations

<b>AAT</b>	aspartate aminotransferase
<b>ACC</b>	1-aminocyclopropane-1-carboxylate
<b>AdoCbl</b>	adenosylcobalamin
<b>AEP</b>	2-aminoethylphosphonic acid
<b>AGXT</b>	alanine aminotransferase
<b>ALA</b>	5-aminolevulinic acid
<b>AlaR</b>	alanine racemase
<b>ALAS</b>	5-aminolevulinate synthase
<b>ANSA</b>	2-aminobenzenesulfonate
<b>AON</b>	8-amino-7-oxononanoate
<b>AONS</b>	8-amino-7-oxononanoate synthase
<b>ATAHS</b>	(±)2-acetylaminocyclohexane sulfonic acid
<b>AVG</b>	aminoethoxyvinylglycine
<b>BCAT</b>	branched-chain amino acid aminotransferase
<b>CBL</b>	cystathionine β-lyase
<b>CBS</b>	cystathionine β-synthase
<b>CGL</b>	cystathionine γ-lyase
<b>CGS</b>	cystathionine γ-synthase
<b>DAAT</b>	D-amino acid aminotransferase
<b>DAPDC</b>	diaminopimelate decarboxylase
<b>DDC</b>	DOPA decarboxylase
<b>DFMO</b>	α-difluoromethylornithine
<b>E<sub>1</sub></b>	CDP-6-deoxy-D-glycero-L-threo-4-hexulose-3-dehydrase
<b>E<sub>3</sub></b>	CDP-6-deoxy-D-glycero-L-threo-4-hexulose-3-dehydrase reductase
<b>GABA</b>	γ-aminobutyric acid
<b>GABA-AT</b>	γ-aminobutyric acid aminotransferase
<b>GSA</b>	glutamate-1-semialdehyde
<b>GSAM</b>	glutamate-1-semialdehyde-1,2-mutase
<b>H<sub>4</sub>PteGlu</b>	tetrahydropteroylglutamate or tetrahydrofolate
<b>HDC</b>	histidine decarboxylase

<b>KAPAS</b>	8-amino-7-oxononanoate synthase
<b>KBL</b>	2-amino-3-ketobutyrate CoA ligase
<b>LAM</b>	lysine 2,3-aminomutase
<b>LbODC</b>	<i>Lactobacillus 30a</i> ornithine decarboxylase
<b>L-DOPA</b>	L-3,4-dihydroxyphenylalanine
<b>L-OSHS</b>	O-succinyl-L-homoserine
<b>MGL</b>	methionine $\gamma$ -lyase
<b>MVC</b>	monovalent cation
<b>OAHS</b>	O-acetyl-L-homoserine
<b>OAS</b>	O-acetyl-L-serine
<b>OASS</b>	O-acetylserine sulfhydrylase
<b>OAT</b>	L-ornithine:2-oxoacid aminotransferase
<b>ODC</b>	ornithine decarboxylase
<b>OPHS</b>	O-phosphohomoserine
<b>OPHS</b>	O-phospho-L-homoserine
<b>PHME</b>	pyridoxyl-histidine methyl ester conjugate
<b>PMP</b>	pyridoxamine 5'-phosphate
<b>PSA</b>	( $\pm$ )piperidine-3-sulfonic acid
<b>SAM</b>	S-adenosyl-L-methionine
<b>SAT</b>	serine acetyltransferase
<b>SCL</b>	selenocysteine lyase
<b>SHMT</b>	serine hydroxymethyltransferase
<b>SPT</b>	serine palmitoyltransferase
<b>TA</b>	threonine aldolase
<b>TbODC</b>	<i>Trypanosoma brucei</i> ornithine decarboxylase
<b>TPL</b>	tyrosine phenol-lyase
<b>Trpase</b>	tryptophan indole-lyase
<b>TRPS</b>	tryptophan synthase
<b>TS</b>	threonine synthase

## References

1. T. B. Fitzpatrick; N. Amrhein; B. Kappes; P. Macheroux; I. Tews; T. Raschle, *Biochem. J.* **2007**, *407* (1), 1–13.
2. P. Clayton, *J. Inherit. Metab. Dis.* **2006**, *29* (2–3), 317–326.
3. S. Cook; O. M. Hess, *Handb. Exp. Pharmacol.* **2005**, (170), 325–338.
4. K. S. McCully, *Am. J. Clin. Nutr.* **2007**, *86* (5), 1563S–1568S.
5. A. Amadasi; M. Bertoldi; R. Contestabile; S. Bettati; B. Cellini; M. L. di Salvo; C. Borri-Voltattorni; F. Bossa; A. Mozzarelli, *Curr. Med. Chem.* **2007**, *14* (12), 1291–1324.
6. J. W. Wheless; R. E. Ramsay; S. D. Collins, *Neurotherapeutics* **2007**, *4* (1), 163–172.
7. R. Pahwa, *J. Am. Med. Dir. Assoc.* **2006**, *7* (7 Suppl. 2), 4–10.
8. R. G. Kallen; T. Korpela; A. E. Martell; Y. Matsushima; C. M. Metzler; D. E. Metzler; V. Yuru; I. M. Rolston; F. A. Saviu; Y. M. Torchinski; H. Ueno. In *Transaminases*; Wiley and Sons: New York, 1985; pp 38–108.
9. C. A. McPhalen; M. G. Vincent; J. N. Jansonius, *J. Mol. Biol.* **1992**, *225* (2), 495–517.
10. G. C. Ford; G. Eichele; J. N. Jansonius, *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77* (5), 2559–2563.
11. H. C. Dunathan, *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55* (4), 712–716.
12. P. K. Mehta; P. Christen, *Adv. Enzymol. Relat. Areas Mol. Biol.* **2000**, *74*, 129–184.
13. R. Percudani; A. Peracchi, *EMBO Rep.* **2003**, *4* (9), 850–854.
14. F. W. Alexander; E. Sandmeier; P. K. Mehta; P. Christen, *Eur. J. Biochem.* **1994**, *219* (3), 953–960.
15. N. V. Grishin; M. A. Phillips; E. J. Goldsmith, *Protein Sci.* **1995**, *4* (7), 1291–1304.
16. J. N. Jansonius, *Curr. Opin. Struct. Biol.* **1998**, *8* (6), 759–769.
17. G. Schneider; H. Kack; Y. Lindqvist, *Structure* **2000**, *8* (1), R1–R6.
18. R. A. John, *Biochim. Biophys. Acta* **1995**, *1248* (2), 81–96.
19. A. C. Eliot; J. F. Kirsch, *Annu. Rev. Biochem.* **2004**, *73*, 383–415.
20. H. Hayashi, *J. Biochem.* **1995**, *118* (3), 463–473.

21. R. S. Phillips; B. Sundararaju; N. G. Faleev, *J. Am. Chem. Soc.* **2000**, *122*, 1008–1014.
22. P. Christen; D. E. Metzler, *Transaminases*; Wiley: New York, 1985; p xxiv, 643p.
23. E. W. Miles, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1991**, *64*, 93–172.
24. E. W. Miles, *Subcell. Biochem.* **1995**, *24*, 207–254.
25. A. Mozzarelli; S. Bettati, *Chem. Rec.* **2006**, *6*, 275–287.
26. T. R. Barends; M. F. Dunn; I. Schlichting, *Curr. Opin. Chem. Biol.* **2008**, *12*, 1–8.
27. M. F. Dunn; D. Niks; H. Ngo; T. R. Barends; I. Schlichting, *Trends Biochem. Sci.* **2008**, *33*, 254–264.
28. S. Raboni; S. Bettati; A. Mozzarelli, *Cell. Mol. Life Sci.* **2009**, submitted.
29. W. M. Rabeh; P. F. Cook, *J. Biol. Chem.* **2004**, *279* (26), 26803–26806.
30. C. H. Tai; P. F. Cook, *Adv. Enzymol. Relat. Areas Mol. Biol.* **2000**, *74*, 185–234.
31. M. E. Tanner, *Acc. Chem. Res.* **2002**, *35* (4), 237–246.
32. A. Rios; T. L. Amyes; J. P. Richard, *J. Am. Chem. Soc.* **2000**, *122*, 9373–9385.
33. T. Yoshimura; N. Esaki, *J. Biosci. Bioeng.* **2003**, *96*, 103–109.
34. P. LeMagueres; H. Im; A. Dvorak; U. Strych; M. Benedik; K. L. Krause, *Biochemistry* **2003**, *42* (50), 14752–14761.
35. P. LeMagueres; H. Im; J. Ebalunode; U. Strych; M. J. Benedik; J. M. Briggs; H. Kohn; K. L. Krause, *Biochemistry* **2005**, *44* (5), 1471–1481.
36. M. Noda; Y. Matoba; T. Kumagai; M. Sugiyama, *J. Biol. Chem.* **2004**, *279* (44), 46153–46161.
37. J. P. Shaw; G. A. Petsko; D. Ringe, *Biochemistry* **1997**, *36* (6), 1329–1342.
38. G. F. Stamper; A. A. Morollo; D. Ringe, *Biochemistry* **1998**, *37* (29), 10438–10445.
39. S. Sun; M. D. Toney, *Biochemistry* **1999**, *38* (13), 4058–4065.
40. A. Watanabe; T. Yoshimura; B. Mikami; H. Hayashi; H. Kagamiyama; N. Esaki, *J. Biol. Chem.* **2002**, *277* (21), 19166–19172.
41. A. Watanabe; Y. Kurokawa; T. Yoshimura; N. Esaki, *J. Biochem. (Tokyo)* **1999**, *125* (6), 987–990.
42. A. Watanabe; Y. Kurokawa; T. Yoshimura; T. Kurihara; K. Soda; N. Esaki, *J. Biol. Chem.* **1999**, *274* (7), 4189–4194.
43. A. Watanabe; T. Yoshimura; B. Mikami; N. Esaki, *J. Biochem. (Tokyo)* **1999**, *126* (4), 781–786.
44. M. A. Spies; M. D. Toney, *Biochemistry* **2003**, *42* (17), 5099–5107.
45. M. A. Spies; J. J. Woodward; M. R. Watnik; M. D. Toney, *J. Am. Chem. Soc.* **2004**, *126* (24), 7464–7475.
46. M. D. Toney, *Arch. Biochem. Biophys.* **2005**, *433* (1), 279–287.
47. B. G. Miller; R. Wolfenden, *Annu. Rev. Biochem.* **2002**, *71*, 847–885.
48. E. A. Boeker; E. E. Snell, *The Enzymes*, 3rd ed.; Academic Press: New York, 1972; Vol. 6.
49. L. K. Jackson; H. B. Brooks; D. P. Myers; M. A. Phillips, *Biochemistry* **2003**, *42* (10), 2933–2940.
50. M. H. O'Leary; G. J. Piazza, *Biochemistry* **1981**, *20* (10), 2743–2748.
51. L. K. Jackson; E. J. Goldsmith; M. A. Phillips, *J. Biol. Chem.* **2003**, *278* (24), 22037–22043.
52. H. B. Brooks; M. A. Phillips, *Biochemistry* **1997**, *36* (49), 15147–15155.
53. L. K. Jackson; H. B. Brooks; A. L. Osterman; E. J. Goldsmith; M. A. Phillips, *Biochemistry* **2000**, *39* (37), 11247–11257.
54. A. D. Kern; M. A. Oliveira; P. Coffino; M. L. Hackert, *Structure* **1999**, *7* (5), 567–581.
55. E. Sandmeier; T. I. Hale; P. Christen, *Eur. J. Biochem.* **1994**, *221* (3), 997–1002.
56. P. P. McCann; A. E. Pegg, *Pharmacol. Ther.* **1992**, *54* (2), 195–215.
57. C. W. Tabor; H. Tabor, *Annu. Rev. Biochem.* **1984**, *53*, 749–790.
58. J. Pepin; F. Milord; C. Guern; P. J. Schechter, *Lancet* **1987**, *2* (8573), 1431–1433.
59. A. E. Pegg; L. M. Shantz; C. S. Coleman, *J. Cell. Biochem. Suppl.* **1995**, *22*, 132–138.
60. R. A. Casero, Jr.; L. J. Marton, *Nat. Rev. Drug Discov.* **2007**, *6* (5), 373–390.
61. S. Hayashi; Y. Murakami; S. Matsufuji, *Trends Biochem. Sci.* **1996**, *21* (1), 27–30.
62. P. Coffino, *Nat. Rev. Mol. Cell Biol.* **2001**, *2* (3), 188–194.
63. M. Zhang; C. M. Pickart; P. Coffino, *EMBO J.* **2003**, *22* (7), 1488–1496.
64. N. V. Grishin; A. L. Osterman; H. B. Brooks; M. A. Phillips; E. J. Goldsmith, *Biochemistry* **1999**, *38* (46), 15174–15184.
65. L. K. Jackson; J. Baldwin; R. Akella; E. J. Goldsmith; M. A. Phillips, *Biochemistry* **2004**, *43* (41), 12990–12999.
66. J. Vitali; D. Carroll; R. G. Chaudhry; M. L. Hackert, *Acta Crystallogr. D Biol. Crystallogr.* **1999**, *55* (Pt. 12), 1978–1985.
67. C. Moman; S. Ernst; R. Ghosh; N. L. Chang; M. L. Hackert, *J. Mol. Biol.* **1995**, *252* (5), 643–655.
68. J. J. Almrud; M. A. Oliveira; A. D. Kern; N. V. Grishin; M. A. Phillips; M. L. Hackert, *J. Mol. Biol.* **2000**, *295* (1), 7–16.
69. J. Lee; A. J. Michael; D. Martynowski; E. J. Goldsmith; M. A. Phillips, *J. Biol. Chem.* **2007**, *282* (37), 27115–27125.
70. A. L. Osterman; H. B. Brooks; L. Jackson; J. J. Abbott; M. A. Phillips, *Biochemistry* **1999**, *38* (36), 11814–11826.
71. T. Swanson; H. B. Brooks; A. L. Osterman; M. H. O'Leary; M. A. Phillips, *Biochemistry* **1998**, *37* (42), 14943–14947.
72. A. L. Osterman; L. N. Kinch; N. V. Grishin; M. A. Phillips, *J. Biol. Chem.* **1995**, *270* (20), 11797–11802.
73. L. Nagy; L. Hiripi, *Neurochem. Int.* **2002**, *41* (1), 9–16.
74. P. J. Facchini; K. L. Huber-Allanach; L. W. Tari, *Phytochemistry* **2000**, *54* (2), 121–138.
75. P. Burkhardt; P. Dominici; C. Borri-Voltattorni; J. N. Jansonius; V. N. Malashkevich, *Nat. Struct. Biol.* **2001**, *8* (11), 963–967.
76. H. Hayashi; H. Mizuguchi; H. Kagamiyama, *Biochemistry* **1993**, *32* (3), 812–818.
77. S. Shaltiel; M. Cortijo, *Biochem. Biophys. Res. Commun.* **1970**, *41* (3), 594–600.
78. H. Hayashi; F. Tsukiyama; S. Ishii; H. Mizuguchi; H. Kagamiyama, *Biochemistry* **1999**, *38* (47), 15615–15622.
79. S. Ishii; H. Hayashi; A. Okamoto; H. Kagamiyama, *Protein Sci.* **1998**, *7* (8), 1802–1810.
80. M. Bertoldi; P. Frigeri; M. Paci; C. B. Voltattorni, *J. Biol. Chem.* **1999**, *274* (9), 5514–5521.
81. M. Bertoldi; B. Cellini; B. Maras; C. B. Voltattorni, *FEBS Lett.* **2005**, *579* (23), 5175–5180.
82. M. Bertoldi; P. Dominici; P. S. Moore; B. Maras; C. B. Voltattorni, *Biochemistry* **1998**, *37* (18), 6552–6561.
83. M. Bertoldi; P. S. Moore; B. Maras; P. Dominici; C. B. Voltattorni, *J. Biol. Chem.* **1996**, *271* (39), 23954–23959.
84. W. Liu; M. D. Toney, *Biochemistry* **2004**, *43* (17), 4998–5010.
85. X. Zhou; S. Kay; M. D. Toney, *Biochemistry* **1998**, *37* (16), 5761–5769.
86. M. D. Toney; E. Hohenester; S. W. Cowan; J. N. Jansonius, *Science* **1993**, *261* (5122), 756–759.
87. E. Hohenester; J. W. Keller; J. N. Jansonius, *Biochemistry* **1994**, *33* (46), 13561–13570.



88. S. S. Ray; J. B. Bonanno; K. R. Rajashankar; M. G. Pinho; G. He; H. De Lencastre; A. Tomasz; S. K. Burley, *Structure* **2002**, 10 (11), 1499–1508.
89. K. Gokulan; B. Rupp; M. S. Pavelka, Jr.; W. R. Jacobs, Jr.; J. C. Sacchettini, *J. Biol. Chem.* **2003**, 278 (20), 18588–18596.
90. R. L. Blakley; S. J. Benkovic, *Folates and Pterins*; Wiley: New York, 1984; p v.
91. P. Stover; V. Schirch, *Biochemistry* **1992**, 31 (7), 2155–2164.
92. P. Stover; V. Schirch, *Biochemistry* **1992**, 31 (7), 2148–2155.
93. P. Stover; V. Schirch, *J. Biol. Chem.* **1990**, 265 (24), 14227–14233.
94. W. B. Holmes; D. R. Appling, *J. Biol. Chem.* **2002**, 277 (23), 20205–20213.
95. H. L. Kruschwitz; D. McDonald; E. A. Cossins; V. Schirch, *J. Biol. Chem.* **1994**, 269 (46), 28757–28763.
96. R. B. Matthews; J. T. Drummond, *Chem. Rev.* **1990**, 90, 1275–1290.
97. M. Sinnott, *Comprehensive Biological Catalysis: A Mechanistic Reference*; Academic Press: San Diego, 1998.
98. M. S. Chen; L. V. Schirch, *J. Biol. Chem.* **1973**, 248 (23), 7979–7984.
99. M. S. Chen; L. V. Schirch, *J. Biol. Chem.* **1973**, 248 (10), 3631–3635.
100. R. G. Kallen; W. P. Jencks, *J. Biol. Chem.* **1966**, 241 (24), 5851–5863.
101. C. M. Tatum, Jr.; P. A. Benkovic; S. J. Benkovic; R. Potts; E. Schleicher; H. G. Floss, *Biochemistry* **1977**, 16 (6), 1093–1102.
102. M. A. Vanoni; S. Lee; H. G. Floss, *J. Am. Chem. Soc.* **1990**, 112, 3987–3992.
103. V. Schirch, *Mechanism of Folate-Requiring Enzymes in One-Carbon Metabolism*; Academic Press: San Diego, 1998; Vol. 2, pp 211–252.
104. V. Trivedi; A. Gupta; V. R. Jala; P. Saravanan; G. S. Rao; N. A. Rao; H. S. Savithri; H. S. Subramanya, *J. Biol. Chem.* **2002**, 277 (19), 17161–17169.
105. D. M. Szebenyi; X. Liu; I. A. Kriksunov; P. J. Stover; D. J. Thiel, *Biochemistry* **2000**, 39 (44), 13313–13323.
106. J. N. Scarsdale; S. Radaev; G. Kazanina; V. Schirch; H. T. Wright, *J. Mol. Biol.* **2000**, 296 (1), 155–168.
107. J. N. Scarsdale; G. Kazanina; S. Radaev; V. Schirch; H. T. Wright, *Biochemistry* **1999**, 38 (26), 8347–8358.
108. S. B. Renwick; K. Snell; U. Baumann, *Structure* **1998**, 6 (9), 1105–1116.
109. D. M. Szebenyi; F. N. Musayev; M. L. di Salvo; M. K. Safo; V. Schirch, *Biochemistry* **2004**, 43 (22), 6865–6876.
110. J. M. Quashnock; J. F. Chlebowski; M. Martinez-Carrion; L. Schirch, *J. Biol. Chem.* **1983**, 258 (1), 503–507.
111. L. Schirch; K. D. Schnackerz, *Biochem. Biophys. Res. Commun.* **1978**, 85 (1), 99–106.
112. S. Hopkins; V. Schirch, *J. Biol. Chem.* **1986**, 261 (7), 3363–3369.
113. J. R. Jagath; B. Sharma; N. A. Rao; H. S. Savithri, *J. Biol. Chem.* **1997**, 272 (39), 24355–24362.
114. D. Schirch; S. Delle Fratte; S. Iurescia; S. Angelaccio; R. Contestabile; F. Bossa; V. Schirch, *J. Biol. Chem.* **1993**, 268 (31), 23132–23138.
115. J. V. Rao; V. Prakash; N. A. Rao; H. S. Savithri, *Eur. J. Biochem.* **2000**, 267 (19), 5967–5976.
116. R. Contestabile; S. Angelaccio; F. Bossa; H. T. Wright; N. Scarsdale; G. Kazanina; V. Schirch, *Biochemistry* **2000**, 39 (25), 7492–7500.
117. V. Schirch; D. M. Szebenyi, *Curr. Opin. Chem. Biol.* **2005**, 9 (5), 482–487.
118. V. Schirch; K. Shostak; M. Zamora; M. Guatam-Basak, *J. Biol. Chem.* **1991**, 266 (2), 759–764.
119. H. Ogawa; T. Gomi; M. Fujioka, *Int. J. Biochem. Cell Biol.* **2000**, 32 (3), 289–301.
120. R. Contestabile; A. Paiardini; S. Pascarella; M. L. di Salvo; S. D'Aguanno; F. Bossa, *Eur. J. Biochem.* **2001**, 268 (24), 6508–6525.
121. A. Paiardini; R. Contestabile; S. D'Aguanno; S. Pascarella; F. Bossa, *Biochim. Biophys. Acta* **2003**, 1647 (1–2), 214–219.
122. F. C. Neidhardt, *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*; American Society for Microbiology: Washington, DC, 1987.
123. G. C. Ferreira; P. J. Neame; H. A. Dailey, *Protein Sci.* **1993**, 2 (11), 1959–1965.
124. H. Ikushiro; H. Hayashi; H. Kagamiyama, *Biochemistry* **2004**, 43 (4), 1082–1092.
125. A. Schmidt; J. Sivaraman; Y. Li; R. Larocque; J. A. Barbosa; C. Smith; A. Matte; J. D. Schrag; M. Cygler, *Biochemistry* **2001**, 40 (17), 5151–5160.
126. Z. Zaman; P. M. Jordan; M. Akhtar, *Biochem. J.* **1973**, 135 (2), 257–263.
127. O. Ploux; A. Marquet, *Eur. J. Biochem.* **1996**, 236 (1), 301–308.
128. S. P. Webster; D. Alexeev; D. J. Campopiano; R. M. Watt; M. Alexeeva; L. Sawyer; R. L. Baxter, *Biochemistry* **2000**, 39 (3), 516–528.
129. G. A. Hunter; G. C. Ferreira, *J. Biol. Chem.* **1999**, 274 (18), 12222–12228.
130. G. A. Hunter; G. C. Ferreira, *Biochemistry* **1999**, 38 (12), 3711–3718.
131. J. Zhang; G. C. Ferreira, *J. Biol. Chem.* **2002**, 277 (47), 44660–44669.
132. D. Alexeev; M. Alexeeva; R. L. Baxter; D. J. Campopiano; S. P. Webster; L. Sawyer, *J. Mol. Biol.* **1998**, 284 (2), 401–419.
133. O. Kerbarh; D. J. Campopiano; R. L. Baxter, *Chem. Commun. (Camb.)* 60–62.
134. D. O. Adams; S. F. Yang, *Proc. Natl. Acad. Sci. U.S.A.* **1979**, 76 (1), 170–174.
135. B. R. Glick; D. M. Penrose; J. Li, *J. Theor. Biol.* **1998**, 190 (1), 63–68.
136. C. Walsh; R. A. Pascal, Jr.; M. Johnston; R. Raines; D. Dikshit; A. Krantz; M. Honma, *Biochemistry* **1981**, 20 (26), 7509–7519.
137. S. Karthikeyan; Z. Zhao; C. L. Kao; Q. Zhou; Z. Tao; H. Zhang; H. W. Liu, *Angew. Chem. Int. Ed. Engl.* **2004**, 43 (26), 3425–3429.
138. M. Yao; T. Ose; H. Sugimoto; A. Horiuchi; A. Nakagawa; S. Wakatsuki; D. Yokoi; T. Murakami; M. Honma; I. Tanaka, *J. Biol. Chem.* **2000**, 275 (44), 34557–34565.
139. T. Ose; A. Fujino; M. Yao; N. Watanabe; M. Honma; I. Tanaka, *J. Biol. Chem.* **2003**, 278 (42), 41069–41076.
140. N. Hontzeas; C. E. Hontzeas; B. R. Glick, *Biotechnol. Adv.* **2006**, 24 (4), 420–426.
141. V. V. Borisov; S. N. Borisova; N. I. Sosfenov; B. K. Vainshtein, *Nature* **1980**, 284 (5752), 189–190.
142. J. F. Kirsch; G. Eichele; G. C. Ford; M. G. Vincent; J. N. Jansonius; H. Gehring; P. Christen, *J. Mol. Biol.* **1984**, 174 (3), 497–525.
143. S. Rhee; M. M. Silva; C. C. Hyde; P. H. Rogers; C. M. Metzler; D. E. Metzler; A. Arnone, *J. Biol. Chem.* **1997**, 272, 17293–17302.
144. B. A. Malcolm; J. F. Kirsch, *Biochem. Biophys. Res. Commun.* **1985**, 132 (3), 915–921.
145. S. Kuramitsu; Y. Inoue; S. Tanase; Y. Morino; H. Kagamiyama, *Biochem. Biophys. Res. Commun.* **1987**, 146 (2), 416–421.
146. M. Ziak; R. Jaussi; H. Gehring; P. Christen, *Eur. J. Biochem.* **1990**, 187 (2), 329–333.
147. M. D. Toney; J. F. Kirsch, *J. Biol. Chem.* **1991**, 266 (35), 23900–23903.
148. M. D. Toney; J. F. Kirsch, *Biochemistry* **1993**, 32 (6), 1471–1479.

149. S. Kochhar; P. Christen, *Eur. J. Biochem.* **1988**, 175 (2), 433–438.
150. S. Kochhar; P. Christen, *Eur. J. Biochem.* **1992**, 203 (3), 563–569.
151. C. A. McPhalen; M. G. Vincent; D. Picot; J. N. Jansonius; A. M. Lesk; C. Chothia, *J. Mol. Biol.* **1992**, 227 (1), 197–213.
152. H. Hayashi; Y. Inoue; S. Kuramitsu; Y. Morino; H. Kagamiyama, *Biochem. Biophys. Res. Commun.* **1990**, 167 (2), 407–412.
153. T. Yano; Y. Hinoue; V. J. Chen; D. E. Metzler; I. Miyahara; K. Hirotsu; H. Kagamiyama, *J. Mol. Biol.* **1993**, 234 (4), 1218–1229.
154. T. Yano; S. Kuramitsu; S. Tanase; Y. Morino; H. Kagamiyama, *Biochemistry* **1992**, 31 (25), 5878–5887.
155. H. Kack; J. Sandmark; K. Gibson; G. Schneider; Y. Lindqvist, *J. Mol. Biol.* **1999**, 291 (4), 857–876.
156. K. Inoue; S. Kuramitsu; A. Okamoto; K. Hirotsu; T. Higuchi; Y. Morino; H. Kagamiyama, *J. Biochem. (Tokyo)* **1991**, 109 (4), 570–576.
157. J. M. Goldberg; R. V. Swanson; H. S. Goodman; J. F. Kirsch, *Biochemistry* **1991**, 30 (1), 305–312.
158. T. Yano; T. Mizuno; H. Kagamiyama, *Biochemistry* **1993**, 32 (7), 1810–1815.
159. A. T. Danishefsky; J. J. Onnuffer; G. A. Petsko; D. Ringe, *Biochemistry* **1991**, 30 (7), 1980–1985.
160. Y. Inoue; S. Kuramitsu; K. Inoue; H. Kagamiyama; K. Hiromi; S. Tanase; Y. Morino, *J. Biol. Chem.* **1989**, 264 (16), 9673–9681.
161. H. Hayashi; S. Kuramitsu; Y. Inoue; Y. Morino; H. Kagamiyama, *Biochem. Biophys. Res. Commun.* **1989**, 159 (1), 337–342.
162. C. N. Cronin; J. F. Kirsch, *Biochemistry* **1988**, 27 (12), 4572–4579.
163. K. Haruyama; T. Nakai; I. Miyahara; K. Hirotsu; H. Mizuguchi; H. Hayashi; H. Kagamiyama, *Biochemistry* **2001**, 40 (15), 4633–4644.
164. M. D. Toney; S. Pascarella; D. De Biase, *Protein Sci.* **1995**, 4 (11), 2366–2374.
165. P. Storici; G. Capitani; D. De Biase; M. Moser; R. A. John; J. N. Jansonius; T. Schirmer, *Biochemistry* **1999**, 38 (27), 8628–8634.
166. P. Storici; G. Capitani; R. Muller; T. Schirmer; J. N. Jansonius, *J. Mol. Biol.* **1999**, 285 (1), 297–309.
167. A. Okamoto; Y. Nakai; H. Hayashi; K. Hirotsu; H. Kagamiyama, *J. Mol. Biol.* **1998**, 280 (3), 443–461.
168. J. J. Onuffer; J. F. Kirsch, *Protein Sci.* **1995**, 4 (9), 1750–1757.
169. V. N. Malashkevich; J. J. Onuffer; J. F. Kirsch; J. N. Jansonius, *Nat. Struct. Biol.* **1995**, 2 (7), 548–553.
170. S. C. Rothman; J. F. Kirsch, *J. Mol. Biol.* **2003**, 327 (3), 593–608.
171. S. Oue; A. Okamoto; T. Yano; H. Kagamiyama, *J. Biol. Chem.* **1999**, 274 (4), 2344–2349.
172. S. Oue; A. Okamoto; T. Yano; H. Kagamiyama, *J. Biochem. (Tokyo)* **2000**, 127 (2), 337–343.
173. R. A. Jensen; W. Gu, *J. Bacteriol.* **1996**, 178 (8), 2161–2171.
174. H. Ura; K. Harata; I. Matsui; S. Kuramitsu, *J. Biochem. (Tokyo)* **2001**, 129 (1), 173–178.
175. I. Matsui; E. Matsui; Y. Sakai; H. Kikuchi; Y. Kawarabayasi; H. Ura; S. Kawaguchi; S. Kuramitsu; K. Harata, *J. Biol. Chem.* **2000**, 275 (7), 4871–4879.
176. M. Goto; I. Miyahara; H. Hayashi; H. Kagamiyama; K. Hirotsu, *Biochemistry* **2003**, 42 (13), 3725–3733.
177. S. Sugio; G. A. Petsko; J. M. Manning; K. Soda; D. Ringe, *Biochemistry* **1995**, 34 (30), 9661–9669.
178. C. G. Kannangara; S. P. Gough; P. Bruyant; J. K. Hooper; A. Kahn; D. von Wettstein, *Trends Biochem. Sci.* **1988**, 13 (4), 139–143.
179. G. C. Ferreira; J. Gong, *J. Bioenerg. Biomembr.* **1995**, 27 (2), 151–159.
180. M. Hennig; B. Grimm; R. Contestabile; R. A. John; J. N. Jansonius, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, 94 (10), 4866–4871.
181. X. Lv; J. Fan; H. Ge; Y. Gao; X. Zhang; M. Teng; L. Niu, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2006**, 62 (Pt. 5), 483–485.
182. J. O. Schulze; W. D. Schubert; J. Moser; D. Jahn; D. W. Heinz, *J. Mol. Biol.* **2006**, 358 (5), 1212–1220.
183. P. Christen; D. E. Metzler, *Transaminases*; Wiley: New York; Chichester, 1985; p xxiv, 643p.
184. R. J. Tyacke; J. L. Harwood; R. A. John, *Biochem. J.* **1993**, 293 (Pt. 3), 697–701.
185. R. Contestabile; S. Angelaccio; R. Maytum; F. Bossa; R. A. John, *J. Biol. Chem.* **2000**, 275 (6), 3879–3886.
186. J. Moser; W. D. Schubert; V. Beier; I. Bringemeier; D. Jahn; D. W. Heinz, *EMBO J.* **2001**, 20 (23), 6583–6590.
187. J. Stetefeld; M. Jenny; P. Burkhard, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103 (37), 13688–13693.
188. H. Kumagai; H. Yamada; H. Matsui; H. Ohkishi; K. Ogata, *J. Biol. Chem.* **1970**, 245 (7), 1767–1772.
189. H. Enei; H. Matsui; K. Yamashita; S. Okumura; H. Yamada, *Agric. Biol. Chem.* **1972**, 36, 1861–1868.
190. S. S. Duffey; M. S. Blum, *Insect. Biochem.* **1977**, 7, 57–65.
191. A. A. Antson; T. V. Demidkina; P. Gollnick; Z. Dauter; R. L. von Tersch; J. Long; S. N. Berezhnoy; R. S. Phillips; E. H. Harutyunyan; K. S. Wilson, *Biochemistry* **1993**, 32 (16), 4195–4206.
192. V. K. Kazakov; I. Tarusina; I. V. Myagkikh; T. Demidkina, *Biokhimiya (Moscow)* **1987**, 52, 1319–1323.
193. B. Sundararaju; H. Chen; S. Shilcutt; R. S. Phillips, *Biochemistry* **2000**, 39 (29), 8546–8555.
194. B. Sundararaju; A. A. Antson; R. S. Phillips; T. V. Demidkina; M. V. Barbolina; P. Gollnick; G. G. Dodson; K. S. Wilson, *Biochemistry* **1997**, 36 (21), 6502–6510.
195. H. Y. Chen; T. V. Demidkina; R. S. Phillips, *Biochemistry* **1995**, 34 (38), 12276–12283.
196. T. V. Demidkina; I. V. Myagkikh, *Biochimie* **1989**, 71 (4), 565–571.
197. M. N. Isupov; A. A. Antson; E. J. Dodson; G. G. Dodson; I. S. Dementieva; L. N. Zakomirdina; K. S. Wilson; Z. Dauter; A. A. Lebedev; E. H. Harutyunyan, *J. Mol. Biol.* **1998**, 276 (3), 603–623.
198. R. S. Phillips; T. V. Demidkina; N. G. Faleev, *Biochim. Biophys. Acta* **2003**, 1647 (1–2), 167–172.
199. D. M. Kiick; R. S. Phillips, *Biochemistry* **1988**, 27 (19), 7333–7338.
200. V. V. Kulikova; L. N. Zakomirdina; I. S. Dementieva; R. S. Phillips; P. D. Gollnick; T. V. Demidkina; N. G. Faleev, *Biochim. Biophys. Acta* **2006**, 1764 (4), 750–757.
201. M. V. Barbolina; R. S. Phillips; P. D. Gollnick; N. G. Faleev; T. V. Demidkina, *Protein Eng.* **2000**, 13 (3), 207–215.
202. N. G. Faleev; S. B. Ruvinov; T. V. Demidkina; I. V. Myagkikh; M. Gololobov; V. I. Bakhmutov; V. M. Belikov, *Eur. J. Biochem.* **1988**, 177 (2), 395–401.
203. T. V. Demidkina; M. V. Barbolina; N. G. Faleev; B. Sundararaju; P. D. Gollnick; R. S. Phillips, *Biochem. J.* **2002**, 363 (Pt. 3), 745–752.
204. D. Milic; D. Matkovic-Calogovic; T. V. Demidkina; V. V. Kulikova; N. I. Sinitzina; A. A. Antson, *Biochemistry* **2006**, 45 (24), 7544–7552.
205. R. S. Phillips, *Arch. Biochem. Biophys.* **1987**, 256 (1), 302–310.
206. H. Chen; R. S. Phillips, *Biochemistry* **1993**, 32 (43), 11591–11599.
207. W. A. Newton; Y. Morino; E. E. Snell, *J. Biol. Chem.* **1965**, 240, 1211–1218.

208. E. E. Snell, *Adv. Enzymol.* **1975**, *42*, 287–333.
209. C. H. Suelter; J. Wang; E. E. Snell, *FEBS Lett.* **1976**, *66*, 230–232.
210. T. Watanabe; E. E. Snell, *J. Biochem. (Tokyo)* **1977**, *82*, 733–745.
211. A. Hogberg-Raubaud; O. Raubaud; M. E. Goldberg, *J. Biol. Chem.* **1975**, *250*, 3352–3358.
212. C. H. Suelter; E. E. Snell, *J. Biol. Chem.* **1977**, *252*, 1852–1857.
213. T. Toraya; T. Nihira; S. Fukui, *Eur. J. Biochem.* **1976**, *69*, 411–419.
214. Y. Morino; E. E. Snell, *J. Biol. Chem.* **1967**, *242*, 5591–5601.
215. R. S. Phillips; N. Johnson; A. V. Kamath, *Biochemistry* **2002**, *41* (12), 4012–4019.
216. T. V. Demidkina; L. N. Zakomirdina; V. V. Kulikova; I. S. Dementieva; N. G. Faleev; L. Ronda; A. Mozzarelli; P. D. Gollnick; R. S. Phillips, *Biochemistry* **2003**, *42* (38), 11161–11169.
217. L. Chu; J. L. Ebersole; G. P. Kurzban; S. C. Holt, *Infect. Immun.* **1997**, *65* (8), 3231–3238.
218. B. Cellini; R. Montoli; A. Bossi; M. Bertoldi; D. V. Laurents; C. B. Voltattorni, *Arch. Biochem. Biophys.* **2006**, *455* (1), 31–39.
219. M. Bertoldi; B. Cellini; T. Clausen; C. B. Voltattorni, *Biochemistry* **2002**, *41* (29), 9153–9164.
220. H. I. Krupka; R. Huber; S. C. Holt; T. Clausen, *EMBO J.* **2000**, *19* (13), 3168–3178.
221. M. Bertoldi; B. Cellini; A. Paiardini; M. Di Salvo; C. Borri Voltattorni, *Biochem. J.* **2003**, *371* (Pt. 2), 473–483.
222. M. Bertoldi; B. Cellini; S. D'Aguzzo; Borri; C. Voltattorni, *J. Biol. Chem.* **2003**, *278* (39), 37336–37343.
223. B. Cellini; M. Bertoldi; R. Montoli; C. Borri Voltattorni, *Biochemistry* **2005**, *44* (42), 13970–13980.
224. G. P. Kurzban; L. Chu; J. L. Ebersole; S. C. Holt, *Oral Microbiol. Immunol.* **1999**, *14* (3), 153–164.
225. B. Cellini; M. Bertoldi; C. Borri Voltattorni, *FEBS Lett.* **2003**, *554*, 306–310.
226. H. Mihara; N. Esaki, *Appl. Microbiol. Biotechnol.* **2002**, *60* (1–2), 12–23.
227. L. Zheng; V. L. Cash; D. H. Flint; D. R. Dean, *J. Biol. Chem.* **1998**, *273* (21), 13264–13272.
228. R. Lill; G. Kispal, *Trends Biochem. Sci.* **2000**, *25* (8), 352–356.
229. H. Mihara; T. Kurihara; T. Yoshimura; K. Soda; N. Esaki, *J. Biol. Chem.* **1997**, *272* (36), 22417–22424.
230. J. Tachezy; L. B. Sanchez; M. Muller, *Mol. Biol. Evol.* **2001**, *18* (10), 1919–1928.
231. H. Mihara; M. Maeda; T. Fujii; T. Kurihara; Y. Hata; N. Esaki, *J. Biol. Chem.* **1999**, *274* (21), 14768–14772.
232. S. I. Patzer; K. Hantke, *J. Bacteriol.* **1999**, *181* (10), 3307–3309.
233. D. Kessler, *FEMS Microbiol. Rev.* **2006**, *30* (6), 825–840.
234. H. Mihara; T. Fujii; S. Kato; T. Kurihara; Y. Hata; N. Esaki, *J. Biochem.* **2002**, *131* (5), 679–685.
235. L. Zheng; R. H. White; V. L. Cash; D. R. Dean, *Biochemistry* **1994**, *33* (15), 4714–4720.
236. L. Zheng; R. H. White; V. L. Cash; R. F. Jack; D. R. Dean, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90* (7), 2754–2758.
237. L. Loiseau; S. Ollagnier-de-Choudens; L. Nachin; M. Fontecave; F. Barras, *J. Biol. Chem.* **2003**, *278* (40), 38352–38359.
238. H. Mihara; T. Kurihara; T. Watanabe; T. Yoshimura; N. Esaki, *J. Biol. Chem.* **2000**, *275* (9), 6195–6200.
239. H. Mihara; T. Kurihara; T. Yoshimura; N. Esaki, *J. Biochem.* **2000**, *127* (4), 559–567.
240. G. M. Lacourciere; T. C. Stadtman, *J. Biol. Chem.* **1998**, *273* (47), 30921–30926.
241. T. Lang; D. Kessler, *J. Biol. Chem.* **1999**, *274* (1), 189–195.
242. T. Clausen; J. T. Kaiser; C. Steegborn; R. Huber; D. Kessler, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (8), 3856–3861.
243. J. T. Kaiser; S. Bruno; T. Clausen; R. Huber; F. Schiaretta; A. Mozzarelli; D. Kessler, *J. Biol. Chem.* **2003**, *278* (1), 357–365.
244. I. Leibrecht; D. Kessler, *J. Biol. Chem.* **1997**, *272* (16), 10442–10447.
245. E. G. Mueller, *Nat. Chem. Biol.* **2006**, *2* (4), 185–194.
246. B. Campanini; F. Schiaretta; S. Abbruzzetti; D. Kessler; A. Mozzarelli, *J. Biol. Chem.* **2006**, *281* (50), 38769–38780.
247. T. Clausen; R. Huber; B. Laber; H. D. Pohlenz; A. Messerschmidt, *J. Mol. Biol.* **1996**, *262* (2), 202–224.
248. S. Ravanel; D. Job; R. Douce, *Biochem. J.* **1996**, *320* (Pt. 2), 383–392.
249. A. Martel; C. Bouthier de la Tour; F. Le Goffic, *Biochem. Biophys. Res. Commun.* **1987**, *147* (2), 565–571.
250. U. Breiting; T. Clausen; S. Ehler; R. Huber; B. Laber; F. Schmidt; E. Pohl; A. Messerschmidt, *Plant Physiol.* **2001**, *126* (2), 631–642.
251. L. J. Ejim; J. E. Blanchard; K. P. Koteva; R. Sumerfield; N. H. Elowe; J. D. Chechetto; E. D. Brown; M. S. Junop; G. D. Wright, *J. Med. Chem.* **2007**, *50* (4), 755–764.
252. E. Zdych; R. Peist; J. Reidl; W. Boos, *J. Bacteriol.* **1995**, *177* (17), 5035–5039.
253. T. Clausen; A. Schlegel; R. Peist; E. Schneider; C. Steegborn; Y. S. Chang; A. Haase; G. P. Bourenkov; H. D. Bartunik; W. Boos, *EMBO J.* **2000**, *19* (5), 831–842.
254. C. C. Hyde; S. A. Ahmed; E. A. Padlan; E. W. Miles; D. R. Davies, *J. Biol. Chem.* **1988**, *263* (33), 17857–17871.
255. W. H. Matchett; J. A. DeMoss, *J. Biol. Chem.* **1975**, *250* (8), 2941–2946.
256. X. Huang; H. M. Holden; F. M. Raushel, *Annu. Rev. Biochem.* **2001**, *70*, 149–180.
257. P. S. Brzovic; Y. Sawa; C. C. Hyde; E. W. Miles; M. F. Dunn, *J. Biol. Chem.* **1992**, *267* (18), 13028–13038.
258. S. Rhee; K. D. Parris; S. A. Ahmed; E. W. Miles; D. R. Davies, *Biochemistry* **1996**, *35* (13), 4211–4221.
259. S. B. Ruvinov; X. J. Yang; K. D. Parris; U. Banik; S. A. Ahmed; E. W. Miles; D. L. Sackett, *J. Biol. Chem.* **1995**, *270* (11), 6357–6369.
260. A. Peracchi; S. Bettati; A. Mozzarelli; G. L. Rossi; E. W. Miles; M. F. Dunn, *Biochemistry* **1996**, *35* (6), 1872–1880.
261. A. Peracchi; A. Mozzarelli; G. L. Rossi, *Biochemistry* **1995**, *34* (29), 9459–9465.
262. E. Woehl; M. F. Dunn, *Biochemistry* **1999**, *38* (22), 7131–7141.
263. E. Woehl; M. F. Dunn, *Biochemistry* **1999**, *38* (22), 7118–7130.
264. E. U. Woehl; M. F. Dunn, *Biochemistry* **1995**, *34* (29), 9466–9476.
265. E. W. Miles; H. Kawasaki; S. A. Ahmed; H. Morita; S. Nagata, *J. Biol. Chem.* **1989**, *264* (11), 6280–6287.
266. Z. Lu; S. Nagata; P. McPhie; E. W. Miles, *J. Biol. Chem.* **1993**, *268* (12), 8727–8734.
267. S. Raboni; A. Mozzarelli; P. F. Cook, *Biochemistry* **2007**, *46* (45), 13223–13234.
268. P. S. Brzovic; A. M. Kayastha; E. W. Miles; M. F. Dunn, *Biochemistry* **1992**, *31* (4), 1180–1190.
269. K. S. Anderson; E. W. Miles; K. A. Johnson, *J. Biol. Chem.* **1991**, *266* (13), 8020–8033.
270. G. E. Skye; R. Potts; H. G. Floss, *J. Am. Chem. Soc.* **1974**, *96* (5), 1593–1595.
271. E. Schleicher; K. Mascaró; R. Potts; D. R. Mann; H. B. Floss, *J. Am. Chem. Soc.* **1976**, *98* (4), 1043–1044.



272. C.-H. Tai; P. F. Cook, *Acc. Chem. Res.* **2001**, *34*, 49–59.
273. M. F. Dunn; V. Aguilar; P. Brzovic; W. F. Drewe, Jr.; K. F. Houben; C. A. Leja; M. Roy, *Biochemistry* **1990**, *29* (37), 8598–8607.
274. F. M. Raushel; J. B. Thoden; H. M. Holden, *Acc. Chem. Res.* **2003**, *36* (7), 539–548.
275. C. A. Leja; E. U. Woehl; M. F. Dunn, *Biochemistry* **1995**, *34* (19), 6552–6561.
276. P. S. Brzovic; C. C. Hyde; E. W. Miles; M. F. Dunn, *Biochemistry* **1993**, *32* (39), 10404–10413.
277. P. S. Brzovic; K. Ngo; M. F. Dunn, *Biochemistry* **1992**, *31* (15), 3831–3839.
278. T. R. Schneider; E. Gerhardt; M. Lee; P. H. Liang; K. S. Anderson; I. Schlichting, *Biochemistry* **1998**, *37* (16), 5394–5406.
279. A. Osborne; Q. Teng; E. W. Miles; R. S. Phillips, *J. Biol. Chem.* **2003**, *278* (45), 44083–44090.
280. E. W. Miles; S. Rhee; D. R. Davies, *J. Biol. Chem.* **1999**, *274* (18), 12193–12196.
281. A. Marabotti; D. De Biase; A. Tramonti; S. Bettati; A. Mozzarelli, *J. Biol. Chem.* **2001**, *276* (21), 17747–17753.
282. S. Raboni; S. Bettati; A. Mozzarelli, *J. Biol. Chem.* **2005**, *280* (14), 13450–13456.
283. M. Weyand; I. Schlichting; P. Herde; A. Marabotti; A. Mozzarelli, *J. Biol. Chem.* **2002**, *277* (12), 10653–10660.
284. F. Spyralis; S. Raboni; P. Cozzini; S. Bettati; A. Mozzarelli, *Biochim. Biophys. Acta* **2006**, *1764* (6), 1102–1109.
285. S. A. Ahmed; S. B. Ruvinov; A. M. Kayastha; E. W. Miles, *J. Biol. Chem.* **1991**, *266* (32), 21548–21557.
286. E. W. Miles, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1979**, *49*, 127–186.
287. P. Burkhard; G. S. Rao; E. Hohenester; K. D. Schnackerz; P. F. Cook; J. N. Jansonius, *J. Mol. Biol.* **1998**, *283* (1), 121–133.
288. M. T. Claus; G. E. Zocher; T. H. Maier; G. E. Schulz, *Biochemistry* **2005**, *44* (24), 8620–8626.
289. C. H. Tai; S. R. Nalabolu; J. W. Simmons, III; T. M. Jacobson; P. F. Cook, *Biochemistry* **1995**, *34* (38), 12311–12322.
290. G. D. McClure, Jr.; P. F. Cook, *Biochemistry* **1994**, *33* (7), 1674–1683.
291. S. Benci; S. Vaccari; A. Mozzarelli; P. F. Cook, *Biochemistry* **1997**, *36* (49), 15419–15427.
292. S. Benci; S. Vaccari; A. Mozzarelli; P. F. Cook, *Biochim. Biophys. Acta* **1999**, *1429* (2), 317–330.
293. P. Burkhard; C. H. Tai; C. M. Ristroph; P. F. Cook; J. N. Jansonius, *J. Mol. Biol.* **1999**, *291* (4), 941–953.
294. H. G. Floss; E. Schleicher; R. Potts, *J. Biol. Chem.* **1976**, *251* (18), 5478–5482.
295. P. F. Cook, *Biochim. Biophys. Acta* **2003**, *1647* (1–2), 66–69.
296. W. F. Drewe, Jr.; M. F. Dunn, *Biochemistry* **1985**, *24* (15), 3977–3987.
297. R. S. Phillips, *Biochemistry* **1991**, *30* (24), 5927–5934.
298. P. Burkhard; C. H. Tai; J. N. Jansonius; P. F. Cook, *J. Mol. Biol.* **2000**, *303* (2), 279–286.
299. P. F. Cook; R. T. Wedding, *J. Biol. Chem.* **1976**, *251* (7), 2023–2029.
300. C. H. Tai; S. R. Nalabolu; T. M. Jacobson; D. E. Minter; P. F. Cook, *Biochemistry* **1993**, *32* (25), 6433–6442.
301. A. Chattopadhyay; M. Meier; S. Ivaninskii; P. Burkhard; F. Speroni; B. Campanini; S. Bettati; A. Mozzarelli; W. M. Rabeh; L. Li; P. F. Cook, *Biochemistry* **2007**, *46* (28), 8315–8330.
302. K. Mino; T. Yamanoue; T. Sakiyama; N. Eisaki; A. Matsuyama; K. Nakanishi, *Biosci. Biotechnol. Biochem.* **1999**, *63* (1), 168–179.
303. K. Mino; T. Yamanoue; T. Sakiyama; N. Eisaki; A. Matsuyama; K. Nakanishi, *Biosci. Biotechnol. Biochem.* **2000**, *64* (8), 1628–1640.
304. M. Droux; M. L. Ruffet; R. Douce; D. Job, *Eur. J. Biochem.* **1998**, *255* (1), 235–245.
305. M. Wirtz; O. Berkowitz; M. Droux; R. Hell, *Eur. J. Biochem.* **2001**, *268* (3), 686–693.
306. K. Mino; K. Hiraoka; K. Imamura; T. Sakiyama; N. Eisaki; A. Matsuyama; K. Nakanishi, *Biosci. Biotechnol. Biochem.* **2000**, *64* (9), 1874–1880.
307. B. Campanini; F. Speroni; E. Salsi; P. F. Cook; S. L. Roderick; B. Huang; S. Bettati; A. Mozzarelli, *Protein Sci.* **2005**, *14* (8), 2115–2124.
308. K. E. Burns; S. Baumgart; P. C. Dorrestein; H. Zhai; F. W. McLafferty; T. P. Begley, *J. Am. Chem. Soc.* **2005**, *127*, 11602–11603.
309. R. Schnell; W. Oehlmann; M. Singh; G. Schneider, *J. Biol. Chem.* **2007**, *282*, 23473–23481.
310. D. Agren; R. Schnell; W. Oehlmann; M. Singh; G. Schneider, *J. Biol. Chem.* **2008**, *283* (46), 31567–31574.
311. S. E. O'Leary; C. T. Jurgenson; S. E. Ealick; T. P. Begley, *Biochemistry* **2008**, *47* (44), 11606–11615.
312. C. T. Jurgenson; K. E. Burns; T. P. Begley; S. E. Ealick, *Biochemistry* **2008**, *47* (39), 10354–10364.
313. J. P. Kraus; M. Janosik; V. Kozich; R. Mandell; V. Shih; M. P. Sperandio; G. Sebastio; R. de Franchis; G. Andria; L. A. Kluijtmans; H. Blom; G. H. Boers; R. B. Gordon; P. Kamoun; M. Y. Tsai; W. D. Kruger; H. G. Koch; T. Ohura; M. Gaustadnes, *Hum. Mutat.* **1999**, *13* (5), 362–375.
314. R. Banerjee; R. Evande; O. Kabil; S. Ojha; S. Taoka, *Biochim. Biophys. Acta* **2003**, *1647* (1–2), 30–35.
315. E. W. Miles; J. P. Kraus, *J. Biol. Chem.* **2004**, *279* (29), 29871–29874.
316. S. Taoka; L. Widjaja; R. Banerjee, *Biochemistry* **1999**, *38* (40), 13155–13161.
317. S. Taoka; B. W. Lepore; O. Kabil; S. Ojha; D. Ringe; R. Banerjee, *Biochemistry* **2002**, *41* (33), 10454–10461.
318. M. Meier; M. Janosik; V. Kery; J. P. Kraus; P. Burkhard, *EMBO J.* **2001**, *20* (15), 3910–3916.
319. S. Ojha; J. Wu; R. LoBrutto; R. Banerjee, *Biochemistry* **2002**, *41* (14), 4649–4654.
320. S. Taoka; S. Ohja; X. Shan; W. D. Kruger; R. Banerjee, *J. Biol. Chem.* **1998**, *273* (39), 25179–25184.
321. S. Taoka; M. West; R. Banerjee, *Biochemistry* **1999**, *38* (22), 7406.
322. S. Taoka; E. L. Green; T. M. Loehr; R. Banerjee, *J. Inorg. Biochem.* **2001**, *87* (4), 253–259.
323. R. Banerjee; C. G. Zou, *Arch. Biochem. Biophys.* **2005**, *433* (1), 144–156.
324. J. D. Finkelstein; W. E. Kyle; J. L. Martin; A. M. Pick, *Biochem. Biophys. Res. Commun.* **1975**, *66* (1), 81–87.
325. K. H. Jhee; P. McPhie; E. W. Miles, *J. Biol. Chem.* **2000**, *275* (16), 11541–11544.
326. K. H. Jhee; P. McPhie; E. W. Miles, *Biochemistry* **2000**, *39* (34), 10548–10556.
327. T. Nozaki; Y. Shigeta; Y. Saito-Nakano; M. Imada; W. D. Kruger, *J. Biol. Chem.* **2001**, *276* (9), 6516–6523.
328. K. H. Jhee; W. D. Kruger, *Antioxid. Redox Signal.* **2005**, *7* (5–6), 813–822.
329. R. Evande; S. Ojha; R. Banerjee, *Arch. Biochem. Biophys.* **2004**, *427* (2), 188–196.
330. K. H. Jhee; D. Niks; P. McPhie; M. F. Dunn; E. W. Miles, *Biochemistry* **2001**, *40* (36), 10873–10880.
331. S. M. Aitken; J. F. Kirsch, *Biochemistry* **2004**, *43* (7), 1963–1971.
332. S. M. Aitken; J. F. Kirsch, *Arch. Biochem. Biophys.* **2005**, *433* (1), 166–175.
333. K. H. Jhee; D. Niks; P. McPhie; M. F. Dunn; E. W. Miles, *Biochemistry* **2002**, *41* (6), 1828–1835.
334. A. E. Braunstein; E. V. Goryachenkova, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1984**, *56*, 1–89.

335. I. Willhardt; B. Wiederanders, *Anal. Biochem.* **1975**, 63 (1), 263–266.
336. J. A. Sturman; N. G. Beratis; L. Guarini; G. E. Gaull, *J. Biol. Chem.* **1980**, 255 (10), 4763–4765.
337. J. R. Uren; R. Ragin; M. Chaykovsky, *Biochem. Pharmacol.* **1978**, 27 (24), 2807–2814.
338. S. Yamagata; R. J. D'Andrea; S. Fujisaki; M. Isaji; K. Nakamura, *J. Bacteriol.* **1993**, 175 (15), 4800–4808.
339. A. Messerschmidt; M. Worbs; C. Steegborn; M. C. Wahl; R. Huber; B. Laber; T. Clausen, *Biol. Chem.* **2003**, 384 (3), 373–386.
340. S. Yamagata; T. Yasugahira; Y. Okuda; T. Iwama, *J. Biochem.* **2003**, 134 (4), 607–613.
341. P. G. Bruinenberg; G. De Roo; G. Limsowtin, *Appl. Environ. Microbiol.* **1997**, 63, (2), 561–566.
342. B. Ono; N. Ishii; K. Naito; S. Miyoshi; S. Shinoda; S. Yamamoto; S. Ohmori, *Yeast* **1993**, 9 (4), 389–397.
343. S. Yamagata; T. Akamatsu; T. Iwama, *Appl. Environ. Microbiol.* **2004**, 70 (6), 3766–3768.
344. D. Cavallini; B. Mondovi; C. De Marco; A. Sciosciasantoro, *Arch. Biochem. Biophys.* **1962**, 96, 456–457.
345. D. Cavallini; B. Mondovi; C. De Marco; A. Scioscia-Santoro, *Enzymologia* **1962**, 24, 253–266.
346. K. Soda, *Meth. Enzymol.* **1987**, 143, 453–459.
347. N. Esaki; K. Soda, *Meth. Enzymol.* **1987**, 143, 459–465.
348. B. C. Lockwood; G. H. Coombs, *Biochem. J.* **1991**, 279 (Pt. 3), 675–682.
349. O. W. Griffith, *Meth. Enzymol.* **1987**, 143, 366–376.
350. A. E. McKie; T. Edlind; J. Walker; J. C. Mottram; G. H. Coombs, *J. Biol. Chem.* **1998**, 273 (10), 5549–5556.
351. M. Tokoro; T. Asai; S. Kobayashi; T. Takeuchi; T. Nozaki, *J. Biol. Chem.* **2003**, 278 (43), 42717–42727.
352. K. Miki; W. Al-Refai; M. Xu; P. Jiang; Y. Tan; M. Bouvet; M. Zhao; A. Gupta; T. Chishima; H. Shimada; M. Makuuchi; A. R. Moossa; R. M. Hoffman, *Cancer Res.* **2000**, 60 (10), 2696–2702.
353. D. V. Mamaeva; E. A. Morozova; A. D. Nikulin; S. V. Revtovich; S. V. Nikonov; M. B. Garber; T. V. Demidkina, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2005**, 61 (Pt 6), 546–549.
354. H. Motoshima; K. Inagaki; T. Kumasaka; M. Furuichi; H. Inoue; T. Tamura; N. Esaki; K. Soda; N. Tanaka; M. Yamamoto; H. Tanaka, *J. Biochem.* **2000**, 128 (3), 349–354.
355. D. Kudou; S. Misaki; M. Yamashita; T. Tamura; T. Takakura; T. Yoshioka; S. Yagi; R. M. Hoffman; A. Takimoto; N. Esaki; K. Inagaki, *J. Biochem.* **2007**, 141 (4), 535–544.
356. T. Nakayama; N. Esaki; H. Tanaka; K. Soda, *Biochemistry* **1988**, 27 (5), 1587–1591.
357. H. Inoue; K. Inagaki; N. Adachi; T. Tamura; N. Esaki; K. Soda; H. Tanaka, *Biosci. Biotechnol. Biochem.* **2000**, 64 (11), 2336–2343.
358. H. Hori; K. Takabayashi; L. Orvis; D. A. Carson; T. Nobori, *Cancer Res.* **1996**, 56 (9), 2116–2122.
359. J. Ruiz-Herrera; R. L. Starkey, *J. Bacteriol.* **1969**, 99 (3), 764–770.
360. H. Tanaka; N. Esaki; K. Soda, *Biochemistry* **1977**, 16 (1), 100–106.
361. H. Tanaka; N. Esaki; K. Soda, *Enzyme Microb. Technol.* **1985**, 7, 530–537.
362. G. Curien; D. Job; R. Douce; R. Dumas, *Biochemistry* **1998**, 37 (38), 13212–13221.
363. M. Garrido-Franco; S. Ehler; A. Messerschmidt; S. Marinkovic; R. Huber; B. Laber; G. P. Bourenkov; T. Clausen, *J. Biol. Chem.* **2002**, 277 (14), 12396–12405.
364. J. Giovannelli; K. Veluthambi; G. A. Thompson; S. H. Mudd; A. H. Datko, *Plant Physiol.* **1984**, 76 (2), 285–292.
365. B. Laber; W. Maurer; C. Hanke; S. Grafe; S. Ehler; A. Messerschmidt; T. Clausen, *Eur. J. Biochem.* **1999**, 263 (1), 212–221.
366. K. Thomazeau; G. Curien; R. Dumas; V. Biou, *Protein Sci.* **2001**, 10 (3), 638–648.
367. R. Omi; M. Goto; I. Miyahara; H. Mizuguchi; H. Hayashi; H. Kagamiyama; K. Hirotsu, *J. Biol. Chem.* **2003**, 278 (46), 46035–46045.
368. C. Mas-Droux; V. Biou; R. Dumas, *J. Biol. Chem.* **2006**, 281 (8), 5188–5196.
369. S. Ravanel; M. Droux; R. Douce, *Arch. Biochem. Biophys.* **1995**, 316 (1), 572–584.
370. Y. Chiba; R. Sakurai; M. Yoshino; K. Ominato; M. Ishikawa; H. Onouchi; S. Naito, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100 (18), 10225–10230.
371. T. Clausen; R. Huber; L. Prade; M. C. Wahl; A. Messerschmidt, *EMBO J.* **1998**, 17 (23), 6827–6838.
372. C. Steegborn; A. Messerschmidt; B. Laber; W. Streber; R. Huber; T. Clausen, *J. Mol. Biol.* **1999**, 290 (5), 983–996.
373. E. L. Holbrook; R. C. Greene; J. H. Krueger, *Biochemistry* **1990**, 29 (2), 435–442.
374. S. M. Aitken; D. H. Kim; J. F. Kirsch, *Biochemistry* **2003**, 42 (38), 11297–11306.
375. P. Brzovic; E. L. Holbrook; R. C. Greene; M. F. Dunn, *Biochemistry* **1990**, 29 (2), 442–451.
376. S. Ravanel; B. Gakiere; D. Job; R. Douce, *Biochem. J.* **1998**, 331 (Pt. 2), 639–648.
377. S. Guggenheim; M. Flavin, *J. Biol. Chem.* **1969**, 244 (22), 6217–6227.
378. S. Guggenheim; M. Flavin, *J. Biol. Chem.* **1969**, 244 (13), 3722–3727.
379. Y. Hacham; U. Gophna; R. Amir, *Mol. Biol. Evol.* **2003**, 20 (9), 1513–1520.
380. H. Kanzaki; M. Kobayashi; T. Nagasawa; H. Yamada, *Eur. J. Biochem.* **1987**, 163 (1), 105–112.
381. P. Gonzalez-Porque; J. L. Strominger, *J. Biol. Chem.* **1972**, 247 (21), 6748–6756.
382. G. Agnihotri; H. W. Liu, *Bioorg. Chem.* **2001**, 29 (4), 234–257.
383. F. J. Ruzicka; P. A. Frey, *Biochim. Biophys. Acta* **2007**, 1774 (2), 286–296.
384. P. A. Frey; O. T. Magnusson, *Chem. Rev.* **2003**, 103 (6), 2129–2148.
385. P. C. Prabhakaran; N. T. Woo; P. S. Yorgey; S. J. Gould, *J. Am. Chem. Soc.* **1988**, 110 (17), 5785–5791.
386. F. Berkovitch; E. Behshad; K. H. Tang; E. A. Enns; P. A. Frey; C. L. Drennan, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 15870–15875.
387. C. H. Chang; P. A. Frey, *J. Biol. Chem.* **2000**, 275, 106–114.
388. K. H. Tang; A. Harms; P. A. Frey, *Biochemistry* **2002**, 41 (27), 8767–8776.
389. H. P. Chen; S. H. Wu; Y. L. Lin; C. M. Chen; S. S. Tsay, *J. Biol. Chem.* **2001**, 276, 44744–44750.
390. P. A. Frey; M. D. Ballinger; G. H. Reed, *Biochem. Soc. Trans.* **1998**, 26 (3), 304–310.
391. C. S. Thressa, Lysine Metabolism by Clostridia. In *Advances in Enzymology and Related Areas of Molecular Biology*; M. Alton, Ed.; 2006; pp 413–448.
392. H. J. Sofia; G. Chen; B. G. Hetzler; J. F. Reyes-Spindola; N. E. Miller, *Nucleic Acids Res.* **2001**, 29 (5), 1097–1106.
393. B. W. Lepore; F. J. Ruzicka; P. A. Frey; D. Ringe, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102 (39), 13819–13824.
394. M. Fontecave; M. Atta; E. Mulliez, *Trends Biochem. Sci.* **2004**, 29 (5), 243–249.

395. D. Chen; C. Walsby; B. M. Hoffman; P. A. Frey, *J. Am. Chem. Soc.* **2003**, *125* (39), 11788–11789.
396. P. A. Frey, *FASEB J.* **1993**, *7*, 662–670.
397. D. Chen; P. A. Frey, *Biochemistry* **2001**, *40* (2), 596–602.
398. M. Moss; P. A. Frey, *J. Biol. Chem.* **1987**, *262* (31), 14859–14862.
399. M. D. Ballinger; G. H. Reed; P. A. Frey, *Biochemistry* **1992**, *31* (4), 949–953.
400. W. Wu; K. W. Lieder; G. H. Reed; P. A. Frey, *Biochemistry* **1995**, *34* (33), 10532–10537.
401. J. Miller; V. Bandarian; G. H. Reed; P. A. Frey, *Arch. Biochem. Biophys.* **2001**, *387* (2), 281–288.
402. H. W. Liu; J. S. Thorson, *Annu. Rev. Microbiol.* **1994**, *48*, 223–256.
403. V. P. Miller; J. S. Thorson; O. Ploux; S. F. Lo; H. W. Liu, *Biochemistry* **1993**, *32* (44), 11934–11942.
404. O. Ploux; Y. Lei; K. Vatanen; H. W. Liu, *Biochemistry* **1995**, *34* (13), 4159–4168.
405. G. Agnihotri; Y. N. Liu; B. M. Paschal; H. W. Liu, *Biochemistry* **2004**, *43* (44), 14265–14274.
406. Y. Lei; O. Ploux; H. W. Liu, *Biochemistry* **1995**, *34* (14), 4643–4654.
407. P. Storici; D. De Biase; F. Bossa; S. Bruno; A. Mozzarelli; C. Peneff; R. B. Silverman; T. Schirmer, *J. Biol. Chem.* **2004**, *279* (1), 363–373.
408. P. Smith; P. H. Szu; C. Bui; H. W. Liu; S. C. Tsai, *Biochemistry* **2008**, *47* (24), 6329–6341.
409. K. D. Burns; P. A. Pieper; H. W. Liu; M. T. Stankovich, *Biochemistry* **1996**, *35* (24), 7879–7889.
410. Q. Wu; Y. N. Liu; H. Chen; E. J. Molitor; H. W. Liu, *Biochemistry* **2007**, *46* (12), 3759–3767.
411. C. W. T. Chang; D. A. Johnson; V. Bandarian; H. Zhou; R. LoBrutto; G. H. Reed; H. W. Liu, *J. Am. Chem. Soc.* **2000**, *122* (17), 4239–4240.
412. A. Eliot; J. Kirsch, *Annu. Rev. Biochem.* **2004**, *73*, 383–415.
413. M. J. Jung; B. Lippert; B. W. Metcalf; P. Bohlen; P. J. Schechter, *J. Neurochem.* **1977**, *29* (5), 797–802.
414. B. Lippert; B. W. Metcalf; M. J. Jung; P. Casara, *Eur. J. Biochem.* **1977**, *74* (3), 441–445.
415. B. Laber; K. P. Gerbling; C. Harde; K. H. Neff; E. Nordhoff; H. D. Pohlenz, *Biochemistry* **1994**, *33* (11), 3413–3423.
416. C. Steegborn; B. Laber; A. Messerschmidt; R. Huber; T. Clausen, *J. Mol. Biol.* **2001**, *311* (4), 789–801.
417. J. Burke; R. B. Silverman, *J. Am. Chem. Soc.* **1991**, *113*, 9329–9340.
418. R. J. Tyacke; R. Contestabile; B. Grimm; J. L. Harwood; R. A. John, *Biochem. J.* **1995**, *309* (Pt. 1), 307–313.
419. N. A. Thornberry; H. G. Bull; D. Taub; K. E. Wilson; G. Gimenez-Gallego; A. Rosegay; D. D. Soderman; A. A. Patchett, *J. Biol. Chem.* **1991**, *266* (32), 21657–21665.
420. R. B. Silverman; R. H. Abeles, *Biochemistry* **1976**, *15* (21), 4718–4723.
421. R. Poulin; L. Lu; B. Ackermann; P. Bey; A. E. Pegg, *J. Biol. Chem.* **1992**, *267* (1), 150–158.
422. T. Watanabe; A. Yamatodani; K. Maeyama; H. Wada, *Trends Pharmacol. Sci.* **1990**, *11* (9), 363–367.
423. M. T. Olmo; F. Sanchez-Jimenez; M. A. Medina; H. Hayashi, *J. Biochem.* **2002**, *132* (3), 433–439.
424. D. Alexeev; R. L. Baxter; D. J. Campopiano; O. Kerbarh; L. Sawyer; N. Tomczyk; R. Watt; S. P. Webster, *Org. Biomol. Chem.* **2006**, *4* (7), 1209–1212.
425. E. A. Wang; C. Walsh, *Biochemistry* **1981**, *20* (26), 7539–7546.
426. W. S. Faraci; C. T. Walsh, *Biochemistry* **1989**, *28* (2), 431–437.
427. E. Wang; C. Walsh, *Biochemistry* **1978**, *17* (7), 1313–1321.
428. B. Badet; D. Roise; C. T. Walsh, *Biochemistry* **1984**, *23* (22), 5188–5194.
429. T. S. Soper; J. M. Manning, *J. Biol. Chem.* **1981**, *256* (9), 4263–4268.
430. R. B. Silverman; M. A. Levy, *Biochemistry* **1981**, *20* (5), 1197–1203.
431. R. B. Silverman; B. J. Invergo, *Biochemistry* **1986**, *25* (22), 6817–6820.
432. J. Qiu; R. B. Silverman, *J. Med. Chem.* **2000**, *43* (4), 706–720.
433. Y. Pan; J. Qiu; R. B. Silverman, *J. Med. Chem.* **2003**, *46*, 5292–5293.
434. Z. Wang; R. B. Silverman, *Bioorg. Med. Chem.* **2006**, *14* (7), 2242–2252.
435. G. Daune; F. Gerhart; N. Seiler, *Biochem. J.* **1988**, *253* (2), 481–488.
436. B. W. Metcalf, *Biochem. Pharmacol.* **1979**, *28* (11), 1705–1712.
437. J. A. Williams; G. Bridge; L. J. Fowler; R. A. John, *Biochem. J.* **1982**, *201* (1), 221–225.
438. M. J. Jung; N. Seiler, *J. Biol. Chem.* **1978**, *253* (20), 7431–7439.
439. E. D. Jones; J. M. Basford; R. A. John, *Biochem. J.* **1983**, *209* (1), 243–249.
440. J. Kollonitsch; L. M. Perkins; A. A. Patchett; G. A. Doldouras; S. Marburg; D. E. Duggan; A. L. Maycock; S. D. Aster, *Nature* **1978**, *274* (5674), 906–908.
441. B. W. Metcalf; P. Bey; C. Danzin; M. J. Jung; P. Casara; J. Ververt, *J. Am. Chem. Soc.* **1978**, *100*, 2551–2553.
442. M. G. Palfreyman; C. Danzin; P. Bey; M. J. Jung; G. Ribereau-Gayon; M. Aubry; J. P. Vevert; A. Sjoerdsma, *J. Neurochem.* **1978**, *31* (4), 927–932.
443. M. J. Jung; M. G. Palfreyman; J. Wagner; P. Bey; G. Ribereau-Gayon; M. Zraika; J. Koch-Weser, *Life Sci.* **1979**, *24* (11), 1037–1042.
444. G. Ribereau-Gayon; C. Danzin; M. G. Palfreyman; M. Aubry; J. Wagner; B. W. Metcalf; M. J. Jung, *Biochem. Pharmacol.* **1979**, *28* (8), 1331–1335.
445. T. D. Fenn; G. F. Stamper; A. A. Morollo; D. Ringe, *Biochemistry* **2003**, *42* (19), 5775–5783.
446. T. D. Fenn; T. Holyoak; G. F. Stamper; D. Ringe, *Biochemistry* **2005**, *44* (14), 5317–5327.
447. D. Peisach; D. Chipman; P. Van Ophem; J. Manning; D. Ringe, *J. Am. Chem. Soc.* **1998**, *120*, 2268–2274.
448. M. Fu; R. B. Silverman, *Bioorg. Med. Chem.* **1999**, *7* (8), 1581–1590.
449. J. Sandmark; S. Mann; A. Marquet; G. Schneider, *J. Biol. Chem.* **2002**, *277* (45), 43352–43358.
450. S. Mann; O. Ploux, *FEBS J.* **2006**, *273* (20), 4778–4789.
451. G. Capitani; D. L. McCarthy; H. Gut; M. G. Grutter; J. F. Kirsch, *J. Biol. Chem.* **2002**, *277* (51), 49735–49742.
452. J. Giovannelli; L. D. Owens; S. H. Mudd, *Plant Physiol.* **1973**, *51* (3), 492–503.
453. T. Clausen; R. Huber; A. Messerschmidt; H. D. Pohlenz; B. Laber, *Biochemistry* **1997**, *36* (41), 12633–12643.
454. T. Beeler; J. E. Churchich, *J. Biol. Chem.* **1976**, *251* (17), 5267–5271.
455. E. S. Venos; M. H. Knodel; C. L. Radford; B. J. Berger, *BMC Microbiol.* **2004**, *4* (1), 39.



456. V. Copie; W. S. Faraci; C. T. Walsh; R. G. Griffin, *Biochemistry* **1988**, *27* (14), 4966–4970.
457. R. Contestabile; T. Jenn; M. Akhtar; D. Gani; R. A. John, *Biochemistry* **2000**, *39* (11), 3091–3096.
458. V. A. Tanay; M. B. Parent; J. T. Wong; T. Paslawski; I. L. Martin; G. B. Baker, *Cell. Mol. Neurobiol.* **2001**, *21* (4), 325–339.
459. K. Snell; D. Riches, *Cancer Lett.* **1989**, *44* (3), 217–220.
460. L. Ejim; J. Blanchard; K. Koteva; R. Sumerfield; N. Elowe; J. Chechetto; E. Brown; M. Junop; G. Wright, *J. Med. Chem.* **2007**, *50*, 755–764.
461. C. Borri-Voltattorni; A. Minelli; P. Borri, *Experientia* **1977**, *33* (2), 158–160.
462. C. Borri-Voltattorni; A. Minelli; P. Borri, *FEBS Lett.* **1977**, *75* (1), 277–280.
463. C. Borri Voltattorni; A. Minelli; P. Borri, *Life Sci.* **1981**, *28* (1), 103–108.
464. D. E. Schwartz; R. Brandt, *Arzneimittelforschung* **1978**, *28* (2), 302–307.
465. R. A. Bakay; A. B. Harris, *Brain Res.* **1981**, *206* (2), 387–404.
466. K. G. Lloyd; C. Munari; P. Worms; L. Bossi; J. Bancaud; J. Talairach; P. L. Morselli, *Adv. Biochem. Psychopharmacol.* **1981**, *26*, 199–206.
467. N. Nishino; H. Fujiwara; S. A. Noguchi-Kuno; C. Tanaka, *Jpn. J. Pharmacol.* **1988**, *48* (3), 331–339.
468. J. Butterworth; C. M. Yates; J. Simpson, *J. Neurochem.* **1983**, *41* (2), 440–447.
469. T. Aoyagi; T. Wada; M. Nagai; F. Kojima; S. Harada; T. Takeuchi; H. Takahashi; K. Hirokawa; T. Tsumita, *Chem. Pharm. Bull. (Tokyo)* **1990**, *38* (6), 1748–1749.
470. L. M. Gunne; J. E. Haggstrom; B. Sjoquist, *Nature* **1984**, *309* (5966), 347–349.
471. S. Choi; P. Storic; T. Schirmer; R. B. Silverman, *J. Am. Chem. Soc.* **2002**, *124* (8), 1620–1624.
472. H. Lu; R. B. Silverman, *J. Med. Chem.* **2006**, *49*, 7404–7412.
473. G. Olson; M. Fu; S. Lau; K. Rinehart; R. B. Silverman, *J. Am. Chem. Soc.* **1998**, *120*, 2256–2267.
474. J. Adams; T. Chen; B. Metcalf, *J. Org. Chem.* **1985**, *50*, 2730–2736.
475. J. Burkhardt; G. Holbert; B. W. Metcalf, *Tetrahedron Lett.* **1984**, *25*, 5267–5270.
476. Z. Wang; H. Yuan; D. Nikolic; R. B. Van Breemen; R. B. Silverman, *Biochemistry* **2006**, *45* (48), 14513–14522.
477. H. Yuan; R. B. Silverman, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1651–1654.
478. P. Christen; D. Metzler, *Transaminases*; Wiley: New York, 1985.
479. F. Rossi; S. Garavaglia; G. B. Giovanzana; B. Arca; J. Li; M. Rizzi, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (15), 5711–5716.
480. S. Brady; S. Wright; J. Lee; A. Sutton; C. Zumoff; R. Wodzinski; S. Beer; J. Clardy, *J. Am. Chem. Soc.* **1999**, *121*, 11912–11913.
481. L. Ricci; M. Frosini; N. Gaggelli; G. Valensin; F. Machetti; G. Sgaragli; M. Valoti, *Biochem. Pharmacol.* **2006**, *71* (10), 1510–1519.
482. M. Weyand; I. Schlichting; A. Marabotti; A. Mozzarelli, *J. Biol. Chem.* **2002**, *277* (12), 10647–10652.
483. J. Finn; C. Langevine; I. Birk; J. Birk; K. Nickerson; S. Rodaway, *Bioorg. Med. Chem. Lett.* **1999**, *9* (16), 2297–2302.
484. A. Amadasi; M. Bertoldi; R. Contestabile; S. Bettati; B. Cellini; M. di Salvo; C. Borri-Voltattorni; F. Bossa; A. Mozzarelli, *Curr. Med. Chem.* **2007**, *14*, 1291–1324.
485. F. Wu; D. Grossenbacher; H. Gehring, *Mol. Cancer Ther.* **2007**, *6*, 1831–1839.
486. F. Wu; J. Yu; H. Gehring, *FASEB J.* **2007**, *22*, 1–9.
487. B. Huang; M. W. Vetting; S. L. Roderick, *J. Bacteriol.* **2005**, *187* (9), 3201–3205.
488. K. Yokoigawa; Y. Okubo; K. Soda, *FEMS Microbiol. Lett.* **2003**, *221* (2), 263–267.
489. D. T. Gallagher; G. L. Gilliland; G. Xiao; J. Zondlo; K. E. Fisher; D. Chinchilla; E. Eisenstein, *Structure* **1998**, *6* (4), 465–475.
490. D. K. Simanshu; H. S. Savithri; M. R. Murthy, *J. Biol. Chem.* **2006**, *281* (51), 39630–39641.
491. K. Pfister; E. Sandmeier; W. Berchtold; P. Christen, *J. Biol. Chem.* **1985**, *260* (21), 11414–11421.
492. D. Picot; E. Sandmeier; C. Thaller; M. G. Vincent; P. Christen; J. N. Jansonius, *Eur. J. Biochem.* **1991**, *196* (2), 329–341.
493. M. I. Arnone; L. Birolo; M. Giamberini; M. V. Cubellis; G. Nitti; G. Sannia; G. Marino, *Eur. J. Biochem.* **1992**, *204* (3), 1183–1189.
494. A. Okamoto; R. Kato; R. Masui; A. Yamagishi; T. Oshima; S. Kuramitsu, *J. Biochem.* **1996**, *119* (1), 135–144.
495. A. Okamoto; T. Higuchi; K. Hirotsu; S. Kuramitsu; H. Kagamiyama, *J. Biochem.* **1994**, *116* (1), 95–107.
496. J. Jager; M. Moser; U. Sauder; J. N. Jansonius, *J. Mol. Biol.* **1994**, *239* (2), 285–305.
497. T. Nakai; K. Okada; S. Akutsu; I. Miyahara; S. Kawaguchi; R. Kato; S. Kuramitsu; K. Hirotsu, *Biochemistry* **1999**, *38* (8), 2413–2424.
498. N. Watanabe; M. M. Cherney; M. J. van Belkum; S. L. Marcus; M. D. Flegel; M. D. Clay; M. K. Deyholos; J. C. Vederas; M. N. James, *J. Mol. Biol.* **2007**, *371* (3), 685–702.
499. S. Lima; R. Khristoforov; C. Momany; R. S. Phillips, *Biochemistry* **2007**, *46* (10), 2735–2744.
500. S. Rhee; K. D. Parris; C. C. Hyde; S. A. Ahmed; E. W. Miles; D. R. Davies, *Biochemistry* **1997**, *36* (25), 7664–7680.
501. R. Wolfenden, *Mol. Cell. Biochem.* **1974**, *3* (3), 207–211.
502. D. Herschlag, *Bioorg. Chem.* **1988**, *16*, 62–96.
503. A. Pasternak; A. White; C. J. Jeffery; N. Medina; M. Cahoon; D. Ringe; L. Hedstrom, *Protein Sci.* **2001**, *10* (7), 1331–1342.
504. J. M. Goldberg; J. F. Kirsch, *Biochemistry* **1996**, *35* (16), 5280–5291.
505. R. Rowlett; L. H. Yang; S. A. Ahmed; P. McPhie; K. H. Jhee; E. W. Miles, *Biochemistry* **1998**, *37* (9), 2961–2968.
506. P. Burkhardt; C. H. Tai; C. M. Ristroph; P. F. Cook; J. N. Jansonius, *J. Mol. Biol.* **1999**, *291* (4), 941–953.
507. G. Mittenhuber, *J. Mol. Microbiol. Biotechnol.* **2001**, *3* (1), 1–20.
508. A. Eschenmoser, *Angew. Chem. Intl. Ed. Engl.* **1988**, *27* (5), 5–39.
509. *The RNA World*, 2nd ed.; CSHL Press: Cold Spring Harbor, 1999.
510. S. A. Benner; A. D. Ellington; A. Tauer, *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86* (18), 7054–7058.
511. V. R. Jadhav; M. Yarus, *Biochimie* **2002**, *84* (9), 877–888.
512. S. M. Austin; T. G. Waddell, *Orig. Life. Evol. Biosph.* **1999**, *29* (3), 287–296.
513. N. Aylward; N. Bofinger, *Biophys. Chem.* **2006**, *123* (2–3), 113–121.
514. D. K. Dalling; D. M. Grant; W. J. Horton; R. D. Sagers, *J. Biol. Chem.* **1976**, *251* (23), 7661–7668.
515. R. F. Zabinski; M. D. Toney, *J. Am. Chem. Soc.* **2001**, *123* (2), 193–198.
516. P. Christen; P. K. Mehta, *Chem. Rec.* **2001**, *1* (6), 436–447.

517. I. C. Schoenhofen; V. V. Lunin; J. P. Julien; Y. Li; E. Ajamian; A. Matte; M. Cygler; J. R. Brisson; A. Aubry; S. M. Logan; S. Bhatia; W. W. Wakarchuk; N. M. Young, *J. Biol. Chem.* **2006**, *281* (13), 8907–8916.
518. P. D. Cook; J. B. Thoden; H. M. Holden, *Protein Sci.* **2006**, *15* (9), 2093–2106.
519. D. Palm; H. W. Klein; R. Schinzel; M. Buehner; E. J. Helmreich, *Biochemistry* **1990**, *29* (5), 1099–1107.
520. C. Ouzounis; N. Kyrpides, *FEBS Lett.* **1996**, *390* (2), 119–123.
521. F. Berkovitch; E. Behshad; K. H. Tang; E. A. Enns; P. A. Frey; C. L. Drennan, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (45), 15870–15875.
522. Q. Han; J. Fang; J. Li, *Biochem. J.* **2001**, *360* (Pt. 3), 617–623.
523. C. Nowicki; G. R. Hunter; M. Montemartini-Kalisz; W. Blankenfeldt; H. Hecht; H. M. Kalisz, *Biochim. Biophys. Acta* **2001**, *1546* (2), 268–281.
524. D. E. Ward; W. M. de Vos; J. van der Oost, *Archaea* **2002**, *1* (2), 133–141.
525. J. Kim; D. Kyung; H. Yun; B. K. Cho; J. H. Seo; M. Cha; B. G. Kim, *Appl. Environ. Microbiol.* **2007**, *73* (6), 1772–1782.
526. C. G. Cheong; C. B. Bauer; K. R. Brushaber; J. C. Escalante-Semerena; I. Rayment, *Biochemistry* **2002**, *41* (15), 4798–4808.
527. G. Capitani; E. Hohenester; L. Feng; P. Storici; J. F. Kirsch; J. N. Jansonius, *J. Mol. Biol.* **1999**, *294* (3), 745–756.
528. B. Rost, *J. Mol. Biol.* **2002**, *318* (2), 595–608.
529. J. Soutourina; S. Blanquet; P. Plateau, *J. Biol. Chem.* **2001**, *276* (44), 40864–40872.
530. S. Auger; M. P. Gomez; A. Danchin; I. Martin-Verstraete, *Biochimie* **2005**, *87* (2), 231–238.
531. H. I. Krupka; R. Huber; S. C. Holt; T. Clausen, *EMBO J.* **2000**, *19* (13), 3168–3178.
532. H. Fukamachi; Y. Nakano; M. Yoshimura; T. Koga, *FEMS Microbiol. Lett.* **2002**, *215* (1), 75–80.
533. T. Clausen; A. Schlegel; R. Peist; E. Schneider; C. Steegborn; Y. S. Chang; A. Haase; G. P. Bourenkov; H. D. Bartunik; W. Boos, *EMBO J.* **2000**, *19* (5), 831–842.
534. H. Kimura, *Mol. Neurobiol.* **2002**, *26* (1), 13–19.
535. I. C. Schoenhofen; D. J. McNally; E. Vinogradov; D. Whitfield; N. M. Young; S. Dick; W. W. Wakarchuk; J. R. Brisson; S. M. Logan, *J. Biol. Chem.* **2006**, *281* (2), 723–732.
536. S. D. Breazeale; A. A. Ribeiro; C. R. Raetz, *J. Biol. Chem.* **2003**, *278* (27), 24731–24739.
537. S. Vijayakumar; A. Merckx-Jacques; D. B. Ratnayake; I. Gryski; R. K. Obhi; S. Houle; C. M. Dozois; C. Creuzenet, *J. Biol. Chem.* **2006**, *281* (38), 27733–27743.
538. R. K. Obhi; C. Creuzenet, *J. Biol. Chem.* **2005**, *280* (21), 20902–20908.
539. M. Wada; S. Nakamori; H. Takagi, *FEMS Microbiol. Lett.* **2003**, *225* (2), 189–193.
540. K. Denger; T. H. Smits; A. M. Cook, *Biochem. J.* **2006**, *394* (Pt. 3), 657–664.
541. Y. Kaminaga; J. Schnepf; G. Peel; C. M. Kish; G. Ben-Nissan; D. Weiss; I. Orlova; O. Lavie; D. Rhodes; K. Wood; D. M. Porterfield; A. J. Cooper; J. V. Schloss; E. Pichersky; A. Vainstein; N. Dudareva, *J. Biol. Chem.* **2006**, *281* (33), 23357–23366.
542. K. Abe; S. Takahashi; Y. Muroki; Y. Kera; R. H. Yamada, *J. Biochem.* **2006**, *139* (2), 235–244.
543. S. Donini; R. Percudani; A. Credali; B. Montanini; A. Sartori; A. Peracchi, *Biochem. Biophys. Res. Commun.* **2006**, *350* (4), 922–928.
544. A. J. McCoy; N. E. Adams; A. O. Hudson; C. Gilvarg; T. Leustek; A. T. Maurelli, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (47), 17909–17914.

### Biographical Sketches



Samanta Raboni received her degree in pharmaceutical chemistry *summa cum laude* at the University of Parma, Italy, in 1998. She was awarded Ph.D. in biochemistry in 2005 by the University of Torino, Italy, for her thesis entitled ‘Stability and allosteric regulation of tryptophan synthase.’ The elucidation of functional and regulatory properties of pyridoxal 5-phosphate-dependent enzymes is her main area of research.



Francesca Spyraakis graduated in medicinal chemistry in 2003 and received her Ph.D. in biochemistry and molecular biology in 2007 at the University of Parma, Parma, Italy. Her research activities are devoted to develop and test alternative scoring methods for the prediction of binding free energy of biomolecular interactions, in collaboration with the Virginia Commonwealth University, Richmond, USA, to apply molecular dynamic methods for the detection of ligand binding pathways, in collaboration with the University of Barcelona, Spain, and for the modeling of protein flexibility in drug design, in collaboration with the University of Modena and Reggio Emilia, Italy.



Barbara Campanini graduated *cum laude* in pharmaceutical chemistry and technologies, University of Parma in 1998. In 2002, she received her Ph.D. in molecular biology and pathology, University of Parma defending a dissertation on 'Structural determinants of the stability of the pyridoxal 5'-phosphate-dependent enzyme *O*-acetylserine sulfhydrylase.' Since 2006, she works as a research scientist at the University of Parma. Her research interests include the functional characterization of PLP-dependent enzymes involved in cysteine metabolism and the preparation of variants of the green fluorescent protein for structural and spectroscopic studies.



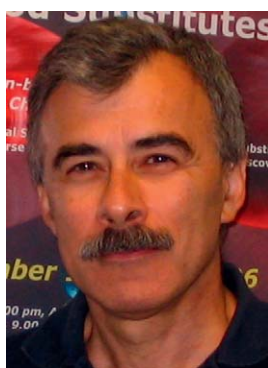
Alessio Amadasi graduated with honors in pharmaceutical chemistry and technology in 2004. He gained a Ph.D. in biochemical sciences working on *in silico* prediction of water–protein free energy of interaction, at the Department of Biochemistry and Molecular Biology, University of Parma, Italy. In 2007, he utilized his time at the laboratory of Professor Glen E. Kellogg and Donald J. Abraham at the Institute for Structural Biology and Drug Discovery of the Virginia Commonwealth University, USA. Amadasi expertises in the field of molecular modeling and structure-based drug design applied to the discovery of new PLP-dependent enzyme inhibitors.



Stefano Bettati received his MS degree from the University of Parma in 1992, defending a thesis on ‘The allosteric regulation of the enzyme tryptophan synthase,’ and his Ph.D. degree in 1998 from the University of Modena, by working on hemoglobin. He was a postdoctoral fellow at the laboratory of Chemical Physics, National Institutes of Health, Bethesda, MD, USA, from 1999 to 2000. At present he is a full professor at Applied Physics, Faculty of Medicine, University of Parma, Italy. His current research interests are mainly in structure, function, and dynamics relationships of PLP-dependent enzymes, heme proteins, and green fluorescent proteins.



Alessio Peracchi obtained a degree in veterinary medicine from the University of Parma in 1989, and later a Ph.D. in molecular biotechnology at the Catholic University of the Sacred Heart, Piacenza, Italy in 1994, and conducted postdoctoral research at Stanford University. In 1998, he became a permanent researcher at the University of Parma, and since 2000 he is an associate professor of biochemistry at the same institution. In 2001, he was nominated EMBO Young Investigator by the European Molecular Biology Organization. His research centers on the function and evolution of catalytic biomacromolecules, contributing to the field with studies on protein enzymes, ribozymes, and deoxyribozymes (catalytic DNAs).



Andrea Mozzarelli received his degree in chemistry in 1974. He is a full professor of biochemistry, University of Parma, Parma, Italy. His main research interest is in the elucidation of structure–dynamics–function relationships of hemoglobins, pyridoxal 5-phosphate-dependent enzymes, and green fluorescent proteins. Protein immobilization through crystallization or encapsulation in silica gels has been exploited as a novel tool to select, isolate, and characterize distinct tertiary and quaternary conformations, using spectroscopic and microspectroscopic methods. Modeling and advanced computational methods are applied to determine free energy of ligand–protein interactions and to design protein ligands. Presently, he is a member of the Editorial Board of *Biochimica et Biophysica Acta – Proteins and Proteomics*, *Current Medicinal Chemistry*, and *Burger's Medicinal Chemistry*, 7th Ed. He is the organizer of the international courses 'From Structural Genomics to Drug Discovery,' held at the University of Parma, Parma, Italy, since 2000, every other year. He was the Chairman of the Conference on 'New challenges in protein science,' held in Parma in 2008 in honor of Dr. W. A. Eaton. He is the president of the XII International Symposium on Blood Substitutes, Parma, Italy, 25–28 August 2009.



Roberto Contestabile was educated at the 'Sapienza' University of Rome, where he graduated in biological sciences. He received his Ph.D. in biochemistry from the School of Molecular and Medical Biosciences of the University of Wales, Cardiff, UK. His research interests focus on the mechanism, structure, and folding of enzymes using pyridoxal 5'-phosphate as a cofactor.



## 7.11 Coenzyme A Biosynthesis and Enzymology

Erick Strauss, Stellenbosch University, Stellenbosch, South Africa

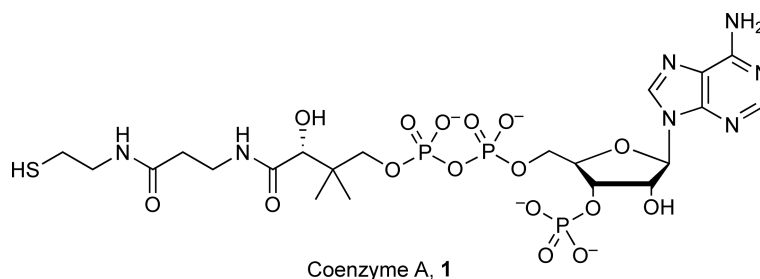
© 2010 Elsevier Ltd. All rights reserved.

7.11.1	Introduction	352
7.11.2	Coenzyme A and the Elucidation of Its Biosynthetic Pathway: A Historical Perspective	353
7.11.2.1	Discovery and Early Studies of Pantothenic Acid	353
7.11.2.2	Discovery and Early Studies of CoA and CoA Enzymology	353
7.11.2.3	Early Studies of CoA Biosynthesis	354
7.11.3	CoA Biosynthesis	355
7.11.3.1	Pantothenic Acid Biosynthesis	357
7.11.3.2	Pantothenate (and Pantetheine) Transport	357
7.11.3.3	Pantothenate Kinase ( <i>coaA</i> and <i>coaX</i> Gene Products)	358
7.11.3.3.1	Type I pantothenate kinases	358
7.11.3.3.2	Type II pantothenate kinases	359
7.11.3.3.3	Type III pantothenate kinases	360
7.11.3.3.4	Archaeal pantothenate kinases	361
7.11.3.3.5	Distribution of PanK types	362
7.11.3.3.6	Pantothenate kinase as a salvage enzyme	363
7.11.3.4	PPCS ( <i>coaB</i> Gene Product)	364
7.11.3.5	Phosphopantothenoylcysteine Decarboxylase ( <i>coaC</i> Gene Product)	366
7.11.3.6	Phosphopantetheine Adenylyltransferase ( <i>coaD</i> Gene Product)	369
7.11.3.7	Dephospho-Coenzyme A Kinase ( <i>coaE</i> Gene Product)	371
7.11.3.8	Genomic Variations in the CoA Biosynthetic Pathway	372
7.11.3.9	Essentiality of Genes Encoding the CoA Biosynthetic Enzymes	374
7.11.4	CoA Utilization, Degradation, and Recycling	374
7.11.4.1	Enzymes Involved in the Consumption of CoA	374
7.11.4.1.1	Phosphopantetheinyl transferase/ACP synthase	374
7.11.4.1.2	Triphosphoribosyl-dephospho-CoA formation and utilization	377
7.11.4.2	Enzymes Involved in the Degradation and Recycling of CoA	378
7.11.4.2.1	CoA diphosphatases	378
7.11.4.2.2	Pantetheinase	379
7.11.4.2.3	Pantothenase	380
7.11.4.2.4	CoA disulfide reductase	380
7.11.5	CoA-Dependent Enzymes	382
7.11.5.1	Enzymes Involved in the Biosynthesis of CoA Thioesters (Acyl-CoAs)	382
7.11.5.1.1	Ligases: Coupling activated carboxylic acids to CoA directly	382
7.11.5.1.2	Oxidoreductases: CoA thioester formation through aldehyde oxidation	384
7.11.5.1.3	Oxidoreductases: CoA thioester formation through oxidative decarboxylation of $\alpha$ -keto acids	384
7.11.5.1.4	Carbon monoxide dehydrogenase/acetyl-CoA synthase	385
7.11.5.2	Acylating Enzymes: CoA Thioesters Acting as Acyl Donors toward Noncarbon Acceptors	385
7.11.5.2.1	Nitrogen- and oxygen-based acceptors ( <i>N</i> - and <i>O</i> -acyltransferases)	385
7.11.5.2.2	Sulfur-based acceptors (transacylases)	388
7.11.5.2.3	Acyl exchange reactions (CoA-transferases)	388
7.11.5.2.4	Water as acceptor (thioesterases)	390
7.11.5.3	Condensing Enzymes: CoA Thioesters Acting as Enolate Nucleophiles toward Carbonyl-Based Electrophiles	391

7.11.5.3.1	Acyl-CoAs (or acyl-ACPs) as electrophiles: True Claisen condensations (thiolases)	391
7.11.5.3.2	Ketones, aldehydes, and carboxylates as electrophiles	392
7.11.5.4	Lyases: Using CoA to Break Carbon–Carbon bonds	395
7.11.5.5	Enzymes Modifying the Acyl Moiety of CoA Thioesters	396
<b>7.11.6</b>	<b>Chemical Biology Tools Based on CoA Enzymology</b>	<b>396</b>
7.11.6.1	Preparation of CoA Analogues	397
7.11.6.2	Applications of CoA Analogues in Chemical Biology	399
7.11.6.2.1	CoA analogues as mechanistic and reactivity probes, and as enzyme inhibitors	399
7.11.6.2.2	CoA analogues as reporter labels	399
<b>7.11.7</b>	<b>Drug Development Efforts Targeting CoA Enzymology</b>	<b>402</b>
<b>7.11.8</b>	<b>Conclusion</b>	<b>403</b>
<b>References</b>		<b>403</b>

### 7.11.1 Introduction

Since the first discovery of coenzyme A (CoA, **1**) by Fritz Lipmann in 1945 as part of his seminal studies on biological acetylation,<sup>1</sup> it has become clear that this molecule may arguably be one of the most widely used of all the enzyme cofactors. A number of attempts have been made to quantify the extent to which it is used; for example, a favorite quote in many introductions to studies on CoA states that 4% of all known enzymes use it as a cofactor. The precise fraction is probably much larger, although it would be difficult to defend any number in a serious argument. A survey of the BRENDA database of enzymes shows that at least 9% of all the enzymes in this database (as distinguished by EC number and origin) use CoA in one form or another – whether this is a more accurate reflection of the cofactor's ubiquity is hard to say. Nonetheless, CoA's role as an essential cofactor in the intermediate metabolism of all living organisms certainly cannot be denied.



One of the reasons for its widespread use is based on the fact that CoA – and the CoA-derived phosphopantetheine tether found on various carrier proteins – acts as the major carrier of metabolites containing carboxylic acid groups: short and long carbon chain acids, Krebs cycle metabolites, and amino acids are the most common compounds carried. However, at any given moment most CoA molecules are involved in the transfer of acetyl groups among a variety of both small and macromolecules, and consequently it is from this role that the cofactor derives the name that Lipmann originally bestowed on it: Coenzyme ‘A,’ to indicate its involvement in ‘activation of acetate.’<sup>2</sup>

CoA fulfills its main biological functions by using its terminal thiol to form thioester linkages with carboxylic acid groups.<sup>3</sup> The scope of reactions in which CoA is involved can therefore almost completely be rationalized in terms of the unique chemistry and reactivity of thiols and thioesters in comparison to other common biologically relevant functional groups. The increased acidity of thiols ( $pK_a \sim 9\text{--}10$ ) means that the thiolate anion, a very good nucleophile, is easily formed in a biological context and also facilitates the acylation reactions in which CoA itself is involved. Thioesters, in turn, are good electrophiles toward both heteroatom nucleophiles (such as alcohols and amines) and carbon nucleophiles (such as enolates), with both kinetic and thermodynamic driving forces favoring acyl transfer. However, in spite of this reactivity profile, thioesters are not any more prone to hydrolysis than its oxoester counterparts, thus making them the ideal carrier of acyl

groups in an aqueous environment.<sup>4,5</sup> Moreover, the CoA thioesters are also good nucleophiles, often forming the very enolates that will react with electrophilic acyl-CoAs by removal of the thioester's  $\alpha$ -proton, which is markedly more acidic ( $pK_a \sim 21$ ) than those of other carboxylic acid derivatives. This duality in its reactivity is another reason for the broad-based utility of CoA as a biological cofactor.

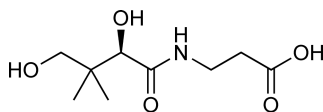
Although the chemical structure of CoA was already determined in 1953,<sup>6</sup> and its first total synthesis reported a few years later,<sup>7</sup> its biosynthetic pathway was fully elucidated and reconstituted only during the past decade.<sup>8–10</sup> This major advancement has allowed a flood of new technologies based on CoA and CoA-based enzymology to be developed. In a similar manner, the essentiality of CoA has led many to propose the CoA-biosynthesizing and CoA-utilizing enzymes as ideal targets for drug development.<sup>11</sup> This chapter will therefore not only report on the diversity of chemical reactions catalyzed by CoA-based enzymes, but also highlight the recent advances that have been made in our understanding of CoA biosynthesis, in the exploitation of CoA analogues as tools and probes in a variety of non-CoA-related studies, and finally also in the development of CoA-directed antimicrobials and drugs.

### 7.11.2 Coenzyme A and the Elucidation of Its Biosynthetic Pathway: A Historical Perspective

More than 60 years have passed since the first discovery of CoA and the sheer number of studies on CoA-related enzymology that has since been published seems daunting at first glance. Considering the recent trend of often disregarding the findings published in literature older than a decade or two, it is therefore perhaps appropriate to provide a brief perspective of the early advances that were made in this field.

#### 7.11.2.1 Discovery and Early Studies of Pantothenic Acid

The history of CoA begins with the discovery of pantothenic acid **2** as growth stimulant in *Saccharomyces cerevisiae* by Williams *et al.*<sup>12</sup> in 1933. It derives its name from the Greek word *pantoben*, meaning 'from all sides,' 'from every quarter,' or 'from everywhere' since it soon became apparent that this vitamin was widely present in living organisms, all of which seemed to respond to it as a growth factor. However, only in 1940 it was established that various growth factors identified through nutritional studies – among others, chick antidermatitis factor and rat growth factor – were indeed all pantothenic acid.<sup>13,14</sup> At this point it was formally classified as a B-complex vitamin, and later specifically as vitamin B<sub>5</sub>. Readers interested in these early biological studies on water-soluble vitamins are referred to the relevant sections of the annual reviews that were published in the volumes of the *Annual Review of Biochemistry* during this period.



Pantothenic acid, **2**

The purification of pantothenic acid from sheep liver subsequently allowed its structure to be determined, with these studies showing it to be made up from  $\beta$ -alanine and pantoic acid (or 2,4-dihydroxy-3,3-dimethylbutanoic acid) units bound together with an amide linkage.<sup>15</sup> Several successful attempts to the synthesis of **2** were published soon after. Although it was known that only the dextrorotary isomer was biologically active, the absolute configuration of the chiral hydroxyl group was established only much later.<sup>16</sup>

#### 7.11.2.2 Discovery and Early Studies of CoA and CoA Enzymology

The discovery of acetyl phosphate in the late 1930s prompted Fritz Lipmann to first consider this molecule as the main form of active acetate in intermediary metabolism.<sup>17</sup> However, the studies that were conducted by Lipmann and his coworkers to test this hypothesis soon showed that another heat-stable cofactor rather fulfilled this role;<sup>18,19</sup> this finding eventually led to the discovery and characterization of CoA as the predominant

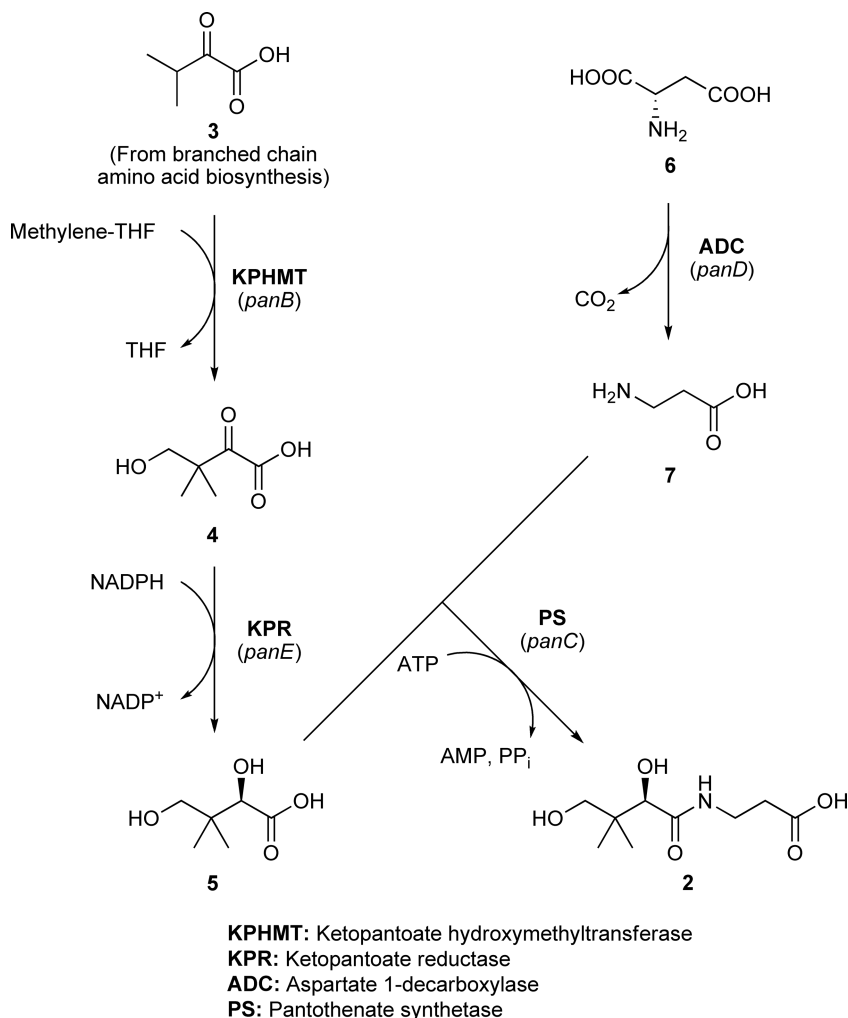
cofactor involved in biological acetylation. Initial studies showed pantothenic acid to be one of the main constituents of CoA and also indicated the presence of an adenosine moiety. However, it involved the combined efforts of several groups using enzymatic and chemical degradation studies to establish the position of the pyrophosphate linkage, the identity of the sulfur-containing moiety, and the presence and point of attachment of the 3'-phosphate. The structure of CoA was finally confirmed as depicted in structure 1.<sup>6</sup> In recognition of this finding, Lipmann received the Nobel Prize in Physiology or Medicine (together with Hans Krebs, who received it for his 'discovery of the citric acid cycle' – another key CoA-dependent process) in 1953. He recounts the details and progression of these discoveries in a paper based on his Nobel Prize address,<sup>1</sup> in a review,<sup>20</sup> and in an autobiographical reflection.<sup>21</sup>

CoA's discovery was nearly immediately followed by the purification of acetyl-CoA by Lynen and Reichert who confirmed that the acetyl group was bound to CoA through an energy-rich thioester.<sup>22,23</sup> The identity of acetyl-CoA as the main form of 'active acetate' in metabolism was subsequently confirmed through several studies of acetyl-CoA-dependent enzymes, many of which helped to lay the foundation for modern mechanistic enzymology. Feodor Lynen was intensely involved in many of these studies, and they eventually led him and Konrad Bloch being awarded the Nobel Prize in Physiology or Medicine in 1964 'for their discoveries concerning the mechanism and regulation of the cholesterol and fatty acid metabolism.'

### 7.11.2.3 Early Studies of CoA Biosynthesis

The first indications of the individual steps of the pantothenic acid biosynthetic pathway were provided by Wieland and Moller,<sup>24</sup> when as early as 1941 they showed that a yeast extract produced pantothenic acid in low yield from pantolactone and  $\beta$ -alanine. Maas and his coworkers was later able to study the activity in cell-free *Escherichia coli* extracts and gathered extensive information about the enzyme, showing that pantothenic acid is formed through the ATP-mediated coupling of pantoic acid 5 and  $\beta$ -alanine 7 (Scheme 1) with AMP and pyrophosphate forming as side products.<sup>25,26</sup> The biosynthesis of pantoic acid was established to proceed from  $\alpha$ -ketoisovaleric acid 3, which was also shown to be the precursor of the branched-chain amino acid valine.<sup>27</sup> On the contrary, several different proposals for the formation of  $\beta$ -alanine were put forward because the manner in which it was produced seemed to vary between organisms; however, evidence strongly suggested that in bacteria it was formed by the decarboxylation of aspartic acid 6. By the late 1950s, the pantothenic acid biosynthetic pathway was proposed to proceed in a form nearly identical to the one known to operate in bacteria today.<sup>28</sup>

In a similar fashion, the five steps of the CoA biosynthesis were elucidated based on various studies on the potential intermediates and precursors of the pathway.<sup>29,30</sup> Initially, two different potential pathways were proposed which only differed in the timing of the 4'-phosphorylation of the pantothenate moiety of CoA: in the first, pantetheine 12 was suggested to form through the coupling of pantothenic acid 2 and cysteine, followed by decarboxylation of the resulting pantothenoylcysteine. A kinase would then act on pantetheine to give 4'-phosphopantetheine 10. On the contrary, it was also possible for pantothenic acid to be phosphorylated first to give 4'-phosphopantothenic acid 8, after which the formation and decarboxylation of 4'-phosphopantothenoylcysteine (PPC) 9 followed. The distinction between these pathways mainly centered on the apparent substrate specificity of the kinase preparations that were used to perform the 4'-phosphorylation reaction. Most of the bacterial enzymes seemed to be specific for pantothenic acid, whereas some (although not all) animal enzymes acted on pantetheine.<sup>31–33</sup> Some preparations – such as the ones from *Morganella morganii* (previously *Proteus morganii*) and from rat liver and kidney – phosphorylated both.<sup>34</sup> Furthermore, the identification of pantetheine 12 as being identical to the long-studied *Lactobacillus bulgaricus* factor (LBF) also seemed to indicate a central role for this compound in CoA metabolism.<sup>35</sup> However, Brown's studies of the pathway using *M. morganii* and rat liver extracts indicated that both systems only coupled 4'-phosphopantothenic acid 8 (and not pantothenic acid 2) to cysteine, and only decarboxylated PPC 9 (and not pantothenoylcysteine).<sup>36</sup> These results were subsequently independently verified using rat liver extracts. Interestingly, Brown's study also showed that the bacterial coupling enzyme specifically required CTP to mediate the coupling reaction, whereas the mammalian enzyme showed similar activity with all the nucleoside triphosphates.<sup>37</sup> Subsequently, the last two enzymes of the pathway were partially purified and characterized from pig and rat liver.<sup>38,39</sup> Based on these findings, the steps of the CoA pathway were therefore formulated by Brown in 1960 as it is shown in Scheme 2, the currently accepted form of the pathway.<sup>40</sup>

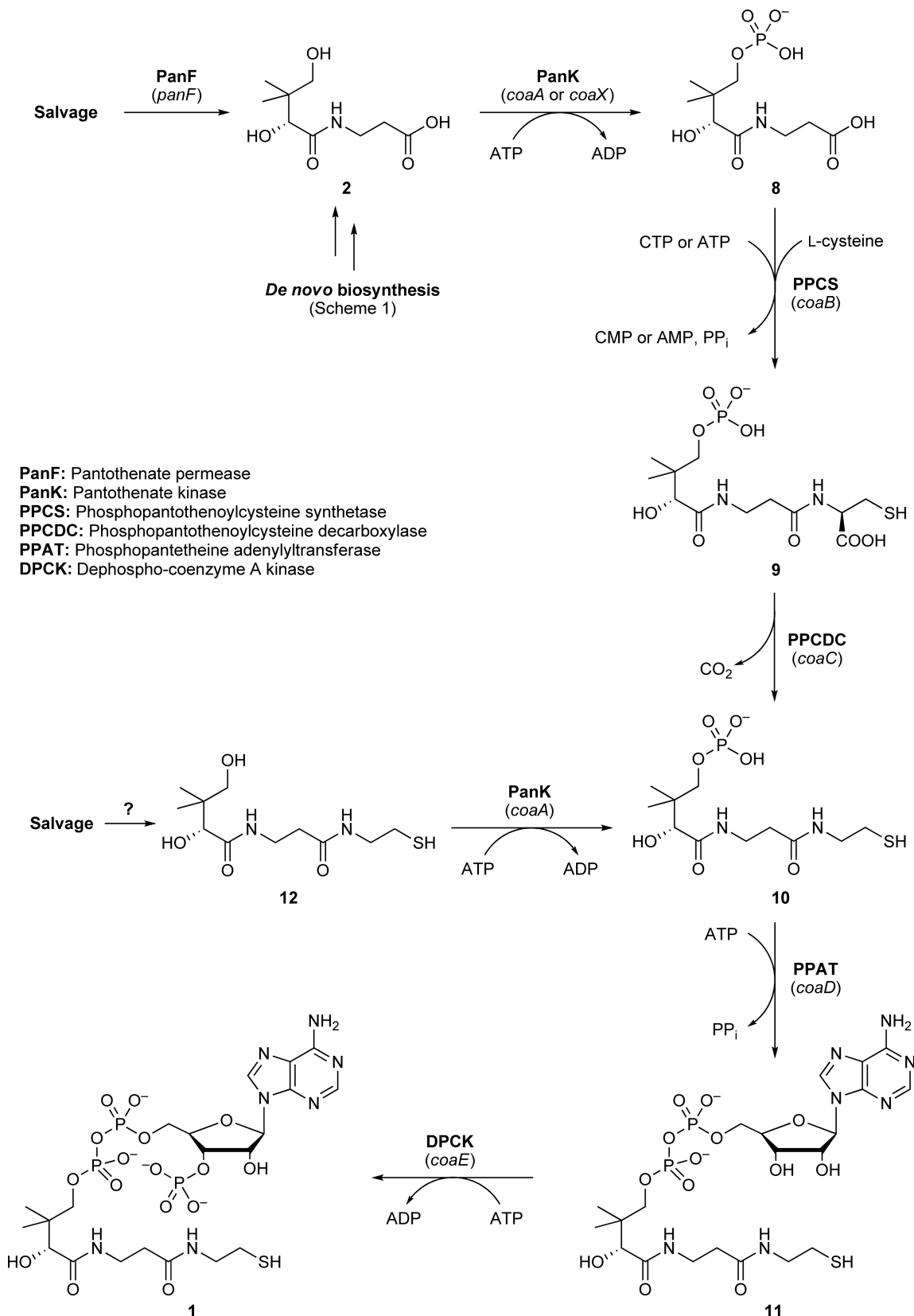


**Scheme 1** The pantothenate biosynthetic pathway in *E. coli*. Enzyme names are abbreviated in bold, whereas the respective genes encoding them are given in italics.

### 7.11.3 CoA Biosynthesis

It is surprising that, in light of the many historic studies of the pantothenic acid and CoA biosynthetic enzymes and their transformations, the cloning and full characterization of all these enzymes are relatively recent events. The first gene encoding a biosynthetic enzyme from these pathways that were identified was *coaA*, encoding pantothenate kinase (PanK) in *Salmonella typhimurium*, in 1979.<sup>41</sup> The corresponding protein in *E. coli* was subsequently the first CoA biosynthetic protein to be cloned and expressed in 1992.<sup>42</sup> In contrast, the last enzyme – phosphopantothenoylcysteine synthetase (PPCS, part of the bifunctional CoaBC protein) in *E. coli* – was only cloned in 2001.<sup>9</sup> However, since then we have witnessed a tremendous growth in our understanding of the enzymes of these pathways has been witnessed.

After pantothenate biosynthesis (**Scheme 1**) or uptake, the CoA biosynthetic pathway universally consists of five enzymatic steps as shown in **Scheme 2**. In the following sections, our current knowledge of each of the biosynthetic steps and other aspects of CoA biosynthesis will be highlighted with a special emphasis on their mechanism and structure. The diversity of genomic variations in the CoA pathway will also be discussed. Many excellent reviews on the CoA biosynthetic pathway have appeared in recent years. These differ in focus, and the readers are encouraged to refer to them as a complement to the content of this chapter whether their interest is



**Scheme 2** The universal coenzyme A biosynthetic pathway. Enzyme names are abbreviated in bold, whereas the respective genes encoding them are given in italics.



a mechanistic one,<sup>10</sup> more centered on the genetics, enzymology, and regulation of the pathway,<sup>8</sup> or specifically focused on the various drug development efforts that have targeted the pathway.<sup>11</sup>

### 7.11.3.1 Pantothenic Acid Biosynthesis

Pantothenic acid **2** is a common precursor for CoA biosynthesis in all organisms, which have to produce this vitamin either through *de novo* biosynthesis or by salvaging it from the environment.<sup>10,43</sup> All plants, fungi, and most bacteria (notable exceptions being *Streptococcus pneumoniae*, *Lactobacillus lactis*, and *Haemophilus influenzae*) are able to biosynthesize pantothenate themselves, often producing much more than is required to satisfy their own CoA biosynthetic requirements.<sup>8</sup> *Escherichia coli*, for example, produces 15 times more than it needs and secretes the excess into the environment.<sup>44</sup> This general oversupply of pantothenate sustains the needs of all other organisms that lack the capacity to produce pantothenate.

The enzymes and transformations that are generally involved in pantothenate biosynthesis are shown in **Scheme 1**. First  $\alpha$ -ketoisovalerate **3**, a common biosynthetic precursor from which the branched-chain amino acids are also produced, is transformed into ketopantoic acid **4** by the action of ketopantoate hydroxymethyltransferase (KPHMT). This enzyme uses  $N^8, N^{10}$ -methylene tetrahydrofolate as a cofactor and source of the hydroxymethyl group. Next ketopantoic acid is reduced to pantoic acid **5** by the NADPH-dependent ketopantoate reductase (KPR). The enzyme pantothenate synthetase (PS) catalyzes the coupling of pantoic acid **5** with  $\beta$ -alanine **7**. This enzyme uses ATP to activate the carboxylate of pantoic acid by the formation of a transient acyl-adenylylate, which is subsequently attacked by the amine of  $\beta$ -alanine. Since PS is the last enzyme of the pathway, and since the activity has been shown to be essential in nearly all pantothenate-biosynthesizing organisms under conditions in which the availability of the vitamin is limited, this activity has received considerable attention as a potential antimicrobial drug target. The source of  $\beta$ -alanine differs depending on the organism: in *E. coli*, the decarboxylation of aspartic acid **6** as catalyzed by the pyruvoyl-dependent aspartate 1-decarboxylase (ADC) enzyme is the only means through which this amino acid is produced. In the yeast *S. cerevisiae*,  $\beta$ -alanine is obtained from the degradation of the polyamines spermine and spermidine, whereas in some other organisms it may be obtained through the degradation of uracil, by an alternative propionate catabolism pathway, or from cysteine.<sup>45</sup>

The enzymes and transformation of the pantothenic acid biosynthetic pathway have been thoroughly reviewed elsewhere, and readers are referred to this detailed account for more information on the various steps summarized above.<sup>45</sup>

### 7.11.3.2 Pantothenate (and Pantetheine) Transport

Most organisms, regardless of their ability to produce pantothenic acid, have some transport system that actively mediates pantothenate uptake.<sup>46</sup> The *E. coli* pantothenate permease (*panF* gene product) is the best characterized of these transport systems.<sup>47,48</sup> Studies have shown that it is a unidirectional sodium-based symporter, which is highly specific for pantothenate ( $K_t$  of  $0.4 \mu\text{mol l}^{-1}$ ), and has maximum velocity of  $1.6 \text{ pmol min}^{-1}$  per  $10^8$  cells.<sup>48,49</sup> This rate of transport is not affected by the intracellular CoA concentration but does increase 10-fold when the permease is overexpressed.<sup>47</sup> Overexpression of the protein also results in an increase in the intracellular concentration of pantothenate but still leaves the CoA levels unaffected, indicating that pantothenate transport does not regulate CoA biosynthesis. The observation that the transporter is unidirectional was deduced from studies that showed that *panF* mutants are still able to excrete the vitamin, which indicates that an as-yet unidentified system is responsible for pantothenate efflux.<sup>50</sup>

In the yeast *S. cerevisiae*, pantothenate is transported by the product of the *FEN2* gene product, a proton-coupled symporter that has no similarity to the *E. coli* PanF transporter.<sup>51,52</sup> It also has high affinity for pantothenate, with a  $K_t$  of  $3.5 \mu\text{mol l}^{-1}$ . In *Schizosaccharomyces pombe*, the homologous Liz1 protein is active in pantothenate transport.<sup>53</sup> Both these transporters are required for pantothenate uptake when the concentrations of the vitamin in the extracellular medium are low. In higher organisms pantothenate transport is mainly mediated by a sodium-dependent multivitamin transporter, which also transports biotin and lipoate.<sup>54–56</sup> Interestingly, although human erythrocytes are normally impermeable to pantothenate, the vitamin can enter cells infected by the malaria parasite *Plasmodium falciparum* that induces the so-called new permeation

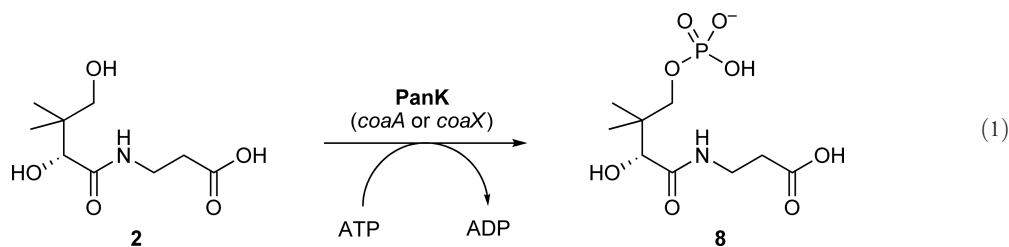
pathways (NPP) in the membrane.<sup>57</sup> These pathways are involved in the transport of numerous low-molecular-weight solutes that are required by the parasite; however, pantothenate is one of the few that can enter the erythrocyte exclusively in this manner. Once inside, the intracellular parasite imports pantothenate using a low-affinity ( $K_t \sim 23 \text{ mmol l}^{-1}$ ) proton-coupled transporter, following which it is immediately phosphorylated by a high-affinity PanK.<sup>58</sup> This transporter bears little resemblance to the mammalian transport system.

It has been demonstrated that pantethine (the disulfide of pantetheine **12**) will support the growth of several bacteria – mostly Lactobacilli – in media lacking pantothenic acid.<sup>59</sup> *Escherichia coli* has been shown to also be able to import exogenous pantetheine for its CoA biosynthetic needs.<sup>60</sup> Its phosphorylated version, 4'-phosphopantetheine **10**, a product of both CoA degradation and turnover of the prosthetic group of the acyl carrier protein (ACP),<sup>61,62</sup> can also be exported from *E. coli* cells but not reimported. However, no transporter specific for pantetheine or 4'-phosphopantetheine has been identified as yet. Considering the amphipathic character of the pantetheine molecule, and in light of recent studies that have shown that a large variety of pantothenamide analogues are freely taken up by *E. coli*,<sup>63–65</sup> it is most probable that pantetheine enters the cell through diffusion-based mechanisms.

A recent review provides a more general discussion of pantothenate transport in all organisms.<sup>8</sup>

### 7.11.3.3 Pantothenate Kinase (*coaA* and *coaX* Gene Products)

PanK (EC 2.7.1.33) catalyzes the first and committed step of CoA biosynthesis: the ATP-dependent phosphorylation of pantothenic acid **2** to form 4'-phosphopantetheinate **8** (Equation (1)).



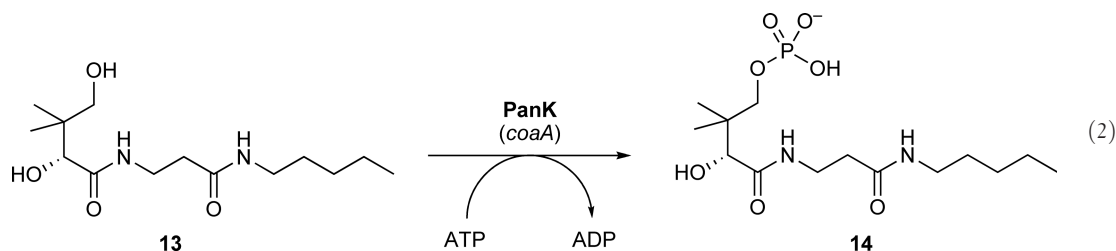
PanK is the best studied of all the CoA biosynthetic enzymes. The *E. coli* enzyme was the first to be expressed and purified, followed by PanKs from numerous other organisms that represent eubacterial, fungal, plant, and human sources.<sup>8</sup> It recently has been shown that despite their common catalytic function, PanKs can be grouped into at least three distinct types that can be distinguished based on sequence homology, structural fold, and catalytic and inhibition properties.<sup>66</sup> The distinction between these types was first thought to be strictly along phylogenetic lines, with clearly defined eukaryotic and prokaryotic types. However, it has now become evident that such a simple classification is not possible as some bacterial PanKs are more closely related to what was originally considered to be strictly eukaryotic enzymes. Moreover, some organisms may even harbor two different PanK types, as in the case of *Bacillus subtilis* and *Mycobacterium tuberculosis*. To highlight the differences and similarities between the various PanK types, our current knowledge of each type is discussed in the following subsections, followed by an analysis of their distribution across all kingdoms.

#### 7.11.3.3.1 Type I pantothenate kinases

The best representative of type I PanKs is the *E. coli* enzyme, the first PanK to be fully characterized and the protein considered to be the prototypical bacterial PanK.<sup>42</sup> It is encoded by the *coaA* gene<sup>67</sup> (as first identified in *S. typhimurium*<sup>41</sup>), and as a result the protein is also often referred to as CoaA. The *E. coli* PanK-I protein exists as a homodimer with  $K_m$  values for its two substrates – pantothenate and ATP – of 36 and 136  $\mu\text{mol l}^{-1}$ ,<sup>68</sup> although other values in a similar range have been reported in the literature.<sup>63,65</sup> The reaction proceeds by an ordered sequential mechanism, with ATP binding before pantothenate. ATP binding was also found to be highly cooperative.<sup>68</sup> An important characteristic of the PanK-I enzymes is their inhibition by free CoA (and by CoA thioesters, albeit to a lesser extent) both *in vivo* and *in vitro*.<sup>69</sup> This inhibition by the end products of the pathway is not only one of the most important ways whereby cells can regulate their intracellular CoA concentration, but also extremely efficient, as highlighted by the observation that a 76-fold overexpression of the protein only

results in a 2.6-fold increase in the CoA concentration.<sup>42</sup> Various kinetic and site-directed mutagenesis studies have shown that this inhibition by CoA is competitive with ATP binding.<sup>70</sup> Superimposition of the crystal structures of the enzyme in complex with either CoA or AMPPNP, a nonhydrolyzable ATP analogue, shows that although both molecules indeed bind in pockets that are spatially situated close to another, their mode of binding is distinctly different.<sup>71</sup> However, the pyrophosphate moiety of CoA occupies the same sites as the  $\beta$ - and  $\gamma$ -phosphates of AMPPNP, which would explain the competition between CoA and ATP, as well as the inhibition of PanK activity upon binding of CoA. The same structures also show why acyl-CoAs are less effective as inhibitors, as CoA's free thiol fits tightly in a hydrophobic pocket lined by four aromatic residues leaving little space for the acyl group to be accommodated in. Similarly, comparison of the CoA complex structure with a ternary structure in which ADP and pantothenate are bound to the enzyme, as well as with a model structure containing the pantothenate analogue *N*-pentyl pantothenamide, indicates that pantothenate and the substrate are accommodated in the same binding groove as the one in which the pantothenate moiety of CoA is found.<sup>72</sup> The pantothenate-binding site is also highly flexible, with the result that the binding of each of these ligands produces a distinct conformation in the protein. These structures all show that *E. coli*'s PanK-I belongs to the family of P-loop kinases, whose members mainly include the various nucleoside and nucleotide kinases.<sup>73,74</sup> This is one of the defining characteristics of this type of PanK.

The binding and reactivity of type I PanKs toward the pantothenamide family of pantothenic acid analogues deserves special mention. When these compounds were first discovered the primary candidate, *N*-pentyl pantothenamide, was shown to inhibit *E. coli* cell growth with a minimum inhibitory concentration of  $\sim 2 \mu\text{mol l}^{-1}$ ;<sup>75</sup> later this growth inhibition was shown to be due to its biosynthetic conversion into the corresponding CoA analogue, ethyldethia-CoA, which acts as an antimetabolite of CoA in normal cell metabolism.<sup>63,76</sup> A full account of the family of pantothenamides' mode of action and the current drug development efforts that focus on their use as potential antibacterial agents is given in a later section. However, it is important to note that all these compounds are active as antimetabolites primarily because they act as alternative substrates of PanK enzymes, and especially of type I PanKs; the phosphorylation of *N*-pentyl pantothenamide **13** is a representative example of this alternate reaction (Equation (2)). The extent of cross-reactivity that a specific PanK enzyme exhibits toward such pantothenamides is therefore one of the important characteristics whereby these enzymes are defined, especially in light of drug development efforts. For example, organisms that harbor a PanK that is unable to perform the phosphorylation reaction of the one shown in Equation (2) cannot be targeted with a strategy that aims to exploit the *in vivo* formation of CoA-based antimetabolites.



Other type I PanK enzymes that have been characterized are the *M. tuberculosis* enzyme,<sup>77,78</sup> of which the crystal structure has been determined, and the enzyme from *Corynebacterium* (previously *Brevibacterium*) *ammoniaenes*,<sup>79,80</sup> an industrially relevant organism.

### 7.11.3.3.2 Type II pantothenate kinases

Type II PanKs are mainly found in eukaryotic sources, including fungi (*Aspergillus nidulans*),<sup>81</sup> plants (*Arabidopsis thaliana* and spinach),<sup>82–84</sup> insects (*Drosophila melanogaster*),<sup>85,86</sup> and mammals (mouse, rat, and human).<sup>87–93</sup> Many of these organisms have more than one PanK-encoding gene, with the different genes exhibiting tissue-specific expression levels. For example, in humans *PANK1* is mostly expressed in the liver and the kidney, and to a lesser extent in the heart and the muscle, whereas *PANK2* is expressed in most tissues, where it is localized to the mitochondria.<sup>94–96</sup> On the contrary, *PANK3* is mainly expressed in the liver, followed by the heart and the skeletal muscle, where *PANK4* expression predominates.<sup>89</sup> Some of these genes also result in the production of two related but distinct PanK isoforms, as in the case of human and mouse *PANK1*, which has two gene

products, PanK1 $\alpha$  and PanK1 $\beta$ .<sup>90–92</sup> In combination, this diversity of genes and expression patterns often leads to a large variety of related PanK proteins in a single organism that can be functionally distinguished, as demonstrated by the fact that the genetic disorder known as PanK-associated neurodegeneration (PKAN) has only been linked to mutations in *PANK2*.<sup>89,97</sup> Similar to the type I PanK enzymes, PanK-II proteins also experience feedback inhibition by the end products of the CoA biosynthetic pathway, although in their case CoA thioesters (especially acetyl-CoA) are generally more potent inhibitors than free CoA.<sup>91,92,98,99</sup> PanK-II enzymes also accept the pantothenamides as alternative substrates.<sup>100</sup>

Interestingly, some Gram-positive bacteria that encode PanK proteins, which – based on primary sequence homology – are more closely related to the eukaryotic type II PanKs mentioned above than to the prototypical bacterial PanK-I enzymes discussed in the previous section. The most notable example of such a bacterial type II PanK is the enzyme from *Staphylococcus aureus*.<sup>101,102</sup> However, this enzyme diverges from the generalized PanK-II profile in that it has a much lower  $K_m$  for ATP (34  $\mu\text{mol l}^{-1}$ , compared with most other PanKs that have values in excess of 100  $\mu\text{mol l}^{-1}$ ) and, most importantly, does not experience any feedback inhibition by CoA or its thioesters.<sup>101,102</sup> The structural basis for this lack of inhibition can be localized to the substitution of two residues in the putative acetyl-CoA binding pocket as identified by comparison with the human PanK3 structure.<sup>103</sup> The first substitution, Ala $\rightarrow$ Tyr, is believed to introduce steric bulk into the potential binding site of the allosteric inhibitor, whereas the second, Trp $\rightarrow$ Arg, disrupts the hydrophobic pocket normally created for the binding of CoA's thiol (or the acetyl group of the thioester). This unique feature – SaPanK-II is the only known type I or type II PanK enzyme that is refractory to feedback inhibition – may be a result of *S. aureus*'s distinctive redox biology, as unlike most organisms it does not rely on glutathione and an NADPH-dependent glutathione reductase to maintain its intracellular redox balance.<sup>104</sup> Instead, it uses CoA and CoA disulfide formation to relieve oxidative stress, which would require much higher concentrations of the cofactor than is needed under normal metabolic conditions. Despite all these differences, the *S. aureus*'s enzyme still accepts pantothenamides as alternative substrates, and this class of potential antimetabolites has received special attention in the development of antistaphylococcal agents.<sup>100,101</sup>

Although PanK-II enzymes can easily be distinguished from type I PanKs based on the differences in their primary sequences, they also have distinctly different structural folds. The structures of the human PanK1 $\alpha$  and PanK3 isoforms and the bacterial PanK-II from *S. aureus*, as well as structural fold predictions,<sup>73,103,105</sup> indicate that PanK-II enzymes belong to the acetate and sugar kinase/heat-shock protein 70/actin (ASKHA) superfamily of kinases; they are not P-loop kinases like type I PanKs.

#### 7.11.3.3.3 Type III pantothenate kinases

The first proposal for the existence of a third PanK type was made based on genome-wide analyses which showed that some bacteria, like *Helicobacter pylori* and *Pseudomonas aeruginosa*, clearly had genes encoding homologues of the last four *E. coli* CoA biosynthetic enzymes, but lacked a gene encoding a recognizable PanK.<sup>46,106</sup> This was the case regardless of whether the prototypical *E. coli* PanK-I or any type II PanK was used to identify a potential homologue. A patent subsequently reported a discovery in the Gram-positive bacterium *B. subtilis* of a gene distinct from its predicted type I PanK-encoding *coaA* gene, but which was also able to rescue a temperature-sensitive *E. coli* *coaA* mutant.<sup>107</sup> Moreover, the study showed that *B. subtilis* *coaA* mutants remained viable, but that double mutants in which the second putative PanK-encoding gene was also disrupted failed to grow. The second PanK-encoding gene was therefore denoted as *coaX*, to differentiate it from the more common *coaA* gene. Homology searches based on the *coaX* sequence showed that most of the bacteria that were thought to have the so-called missing PanKs, indeed had CoaX homologues that bore no sequence similarity to other PanK-I or PanK-II proteins.

The *B. subtilis* CoaX-protein, as well as its homologue in *H. pylori*, was subsequently overexpressed, purified, and fully characterized.<sup>108</sup> This study showed that the proteins do not only have distinct primary sequences, but also kinetic parameters that differ markedly from those of the previously known type I and type II PanKs. For example, although these enzymes have  $K_m$  values for pantothenate in a range similar to those of other PanKs, their  $K_m$  values for ATP were found to be in the millimolar range, and as high as  $\sim 10 \text{ mmol l}^{-1}$  for *H. pylori* CoaX. Various other high-energy phosphate-containing molecules have subsequently been tested as potential alternative phosphate donors, but only purine nucleotides showed detectable activity and none of these are as high as ATP. CoaX-like PanKs are also not feedback inhibited by CoA or acetyl-CoA, and do not accept the

pantothenamide class of pantothenate analogues as alternative substrates. In combination, these attributes indicated that the CoaX proteins are indeed a third type of PanK, and can be denoted as PanK-III.

Structural studies have confirmed that similar to PanK-II enzymes, type III PanKs also belong to the ASKHA superfamily of kinases.<sup>66,105,109</sup> Structures of the *Thermotoga maritima* and *P. aeruginosa* PanK-III enzymes also showed that when pantothenate binds to the protein both its C4' hydroxyl and its carboxylate form direct interactions with protein residues. The pantothenate-binding pocket is therefore enclosed on both ends, making it impossible for the enzyme to accept either the pantotheine moiety of CoA as feedback inhibitor or the extended hydrophobic alkyl chain of the pantothenamide substrate analogues. Moreover, a binary complex of the *T. maritima* PanK-III with ATP showed that the enzyme makes very few contacts with either the adenine or the ribose moieties of the nucleotide, explaining the observed high  $K_m$  for ATP. Finally, the structures also indicate that although PanK-II and PanK-III enzymes both belong to the ASKHA kinase family and both have conserved folds and key catalytic residues, there are significant differences in the ATP- and pantothenate-binding sites between the two types.

Recent studies have brought a number of other characteristics of PanK-III enzymes to light, showing that they are dependent on monovalent cations (such as  $\text{NH}_4^+$  and  $\text{K}^+$ ) for activity, and that like *E. coli*'s PanK-I enzyme binding of the substrates are most probably sequential. However, in contrast to the *E. coli* enzyme pantothenate probably binds before ATP, and phosphorylation is more likely to proceed by an associative mechanism with an active site Asp residue acting as catalytic base.<sup>66,105,109</sup> Isothermal titration calorimetry (ITC) experiments of the *T. maritima* and *H. pylori* enzymes show that the protein binds pantothenate with high affinity ( $K_d \sim 2.7\text{--}6.4 \mu\text{mol l}^{-1}$ ) whereas the ATP analogue AMPPNP binds with a much higher  $K_d$  value ( $\sim 3 \text{ mmol l}^{-1}$ ) in agreement with the high observed  $K_m$  values. Interestingly, the category A biodefense pathogen *Bacillus anthracis* was also shown to possess a type III PanK even though it had previously been suggested that this bacterium has a eukaryotic-like type II PanK. However, expression of the putative PanK-II with either N- or C-terminal histidine tags failed to produce protein with demonstrable *in vitro* activity.<sup>110</sup> The corresponding gene also failed to rescue a temperature-sensitive *E. coli* *coaA* mutant when cloned *in trans*. This suggests that *B. anthracis* is solely reliant on its PanK-III enzyme for CoA biosynthesis, which would be in agreement with the observation that (like *S. aureus*) it relies on CoA as its major low-molecular-weight thiol.

The major characteristics of the three types of PanK discussed above are summarized in comparative fashion in **Table 1**.

#### 7.11.3.3.4 Archaeal pantothenate kinases

To date no PanK enzyme from any archaeobacterial species has been isolated or otherwise characterized. Homology searches of the available sequenced archaeal genomes using sequences of any one of the three PanK types as search terms have also failed to identify potential candidates of putative PanK-encoding genes. This

**Table 1** Generalized comparison of the main attributes of the three characterized PanK types

Attribute	Type I PanK	Type II PanK		Type III PanK
	( <i>E. coli</i> CoaA)	( <i>A. nidulans</i> PanK)	( <i>S. aureus</i> CoA – atypical)	( <i>H. pylori</i> CoaX)
$K_m$ (pantothenate)	$36 \pm 4 \mu\text{mol l}^{-1}$	$60 \mu\text{mol l}^{-1}$	$23 \mu\text{mol l}^{-1}$	$5.5 \mu\text{mol l}^{-1}$
$K_m$ (ATP)	$136 \pm 15 \mu\text{mol l}^{-1}$	$145 \mu\text{mol l}^{-1}$	$34 \mu\text{mol l}^{-1}$	$7.9 \text{ mmol l}^{-1}$
$k_{\text{cat}}$	$0.30 \pm 0.13 \text{ s}^{-1}$	$1.95 \text{ s}^{-1}$	$1.65 \pm 0.09 \text{ s}^{-1}$	$2.09 \pm 0.26 \text{ s}^{-1}$
Cofactor requirements	$\text{Mg}^{2+}$	$\text{Mg}^{2+}$	$\text{Mg}^{2+}$	$\text{Mg}^{2+}$ & $\text{K}^+$ or $\text{NH}_4^+$
Feedback inhibitor	CoA (less by CoA thioesters)	Acetyl-CoA	None	None
Pantothenamide analogues	Substrates	Substrates	Substrates	No effect
Structural fold <sup>73,74</sup>	P-loop kinase	ASKHA superfamily	ASKHA superfamily	ASKHA superfamily
References	68	81	101	108, 109



suggests that an as-yet uncharacterized type of PanK predominates in the archaeobacteria. Genchel has used gene cluster analysis to identify putative small-molecule kinase-encoding genes that are linked with the genes of other known CoA biosynthetic enzymes.<sup>111</sup> Such genes were suggested to encode candidate PanK enzymes; however, their activity still has to be verified by biochemical or genetic means.

#### 7.11.3.3.5 Distribution of PanK types

As mentioned above, the distribution of PanK types was long thought to strictly follow the domains of life, with all eukaryotes having type II PanKs and all bacteria having PanK-I enzymes. With the discovery of type II enzymes in *S. aureus* and the previously unknown group of type III PanKs such a narrowly defined characterization is obviously not appropriate. Although it is true that all eukaryotes only have PanK-II enzymes, all three types are found in bacteria, and often also in combinations. The SEED genomic integration platform (<http://seed-viewer.theseed.org/>; *Subsystems*: Coenzyme A Biosynthesis), which has previously been used to establish the existence of the five-step universal CoA biosynthetic pathway in the majority of these organisms, contains data from more than 400 complete or nearly complete genomes and can be used to survey the phylogenetic distribution of PanK types in bacteria.<sup>112,113</sup> A summary of the results of such a comprehensive survey in 13 major bacterial groups is depicted in **Table 2** and shows that type III enzymes in fact account for the only identifiable PanK activity in 8 of the 13 groups. PanK-III enzymes are also found in combination with type I PanKs in four of the remaining five groups, the exception being the Chlamydiae in which no PanK can as yet be identified.<sup>106,112</sup> It is also especially interesting to note that many pathogenic bacteria only have type III PanKs; however, it remains uncertain whether this observation is relevant in terms of the virulence of these organisms. Nonetheless, it does point to PanK-III proteins as promising targets for drug development, especially considering their divergence from type II (and therefore human) enzymes.<sup>109</sup> The Firmicutes is the only

**Table 2** Phylogenetic distribution of PanK types in bacteria<sup>a</sup>

Bacterial group	PanK type present				
	Only type I	Only type II	Only type III	Type I + type III combination	Type II + type III combination
Actinobacteria <sup>b</sup>	✓		✓	✓	
Aquificae			✓		
Bacteroidetes/Chlorobi			✓		
Chlamydiae <sup>c</sup>					
Chloroflexi	✓		✓		
Cyanobacteria			✓		
Deinococcus/Thermus			✓		
Firmicutes <sup>d</sup>	✓	✓	✓	✓	✓
Fusobacteria			✓		
Planctomycetes			✓		
Proteobacteria:					
α	✓		✓		
β			✓		
δ/ε			✓		
γ	✓		✓		
Unclassified			✓		
Spirochaetes			✓		
Thermotogae			✓		

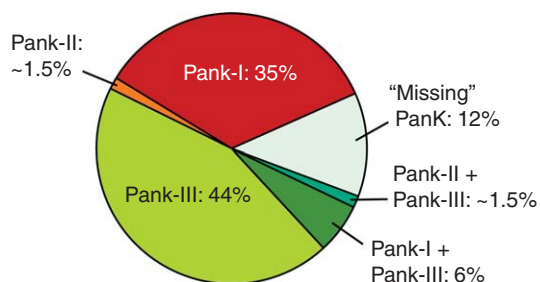
<sup>a</sup> For a more detailed distribution of different types of PanK in individual species, see Yang *et al.*<sup>66</sup>

<sup>b</sup> *Mycobacterium* spp. and *Streptomyces* spp. are predicted to have both type I and type III PanKs.

<sup>c</sup> No PanK candidate has been identified in the *Chlamydiae* to date.

<sup>d</sup> Most bacilli have both type I and type III PanKs, except *B. anthracis*, *B. cereus*, and *Oceanobacillus* spp. which seem to have combinations of type II and type III, although recent reports show that the type II protein is inactive.<sup>110</sup> Staphylococci are the only bacteria that have only type II enzymes.<sup>66</sup>



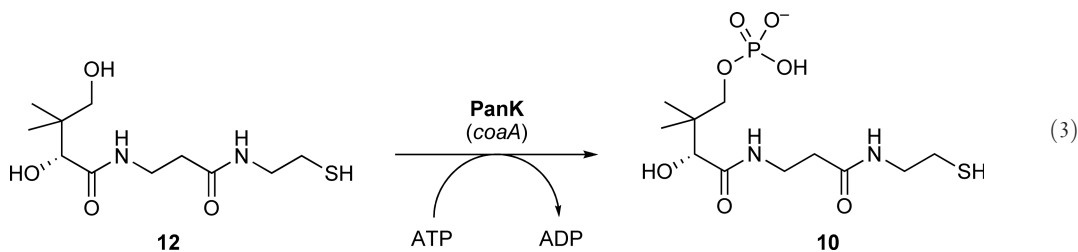


**Figure 1** The distribution of the PanK types in more than 280 eubacterial species with sequenced genomes representing 13 bacterial groups as indicated by the percentage representation of each type (either alone or in combination), as well as the proportion of organisms in which the PanK activity is considered to be missing.

bacterial group that has members containing PanK-II enzymes: in the Staphylococci it is predicted to be the only PanK, whereas in *B. anthracis*, *Bacillus cereus*, and *Oceanobacillus* spp. it was predicted to occur in combination with a type III PanK. However, a recent study of PanK activity in *B. anthracis* showed that this organism's putative type II PanK did not exhibit *in vitro* or *in vivo* activity, raising the question whether this is indeed a meaningful functional distinction.<sup>110</sup> Taken together, this analysis shows that in contrary to the previously held notion that *E. coli*'s PanK-I enzyme is the prototypical bacterial PanK, it is rather the PanK-III enzymes that can be considered to be the predominant type of PanK in all eubacteria. This observation is made more obvious by a graphical representation of the distribution of the PanK enzyme types – either alone or in combination – in the ~280 distinct bacterial species that have been sequenced (Figure 1). This distribution is based on the identification of each PanK by sequence homology searches and by cluster analysis. However, in some of these cases the organisms are predicted to harbor a PanK activity, but the responsible enzyme cannot be identified by bioinformatics methods; in such cases, the PanK is considered to be 'missing'.<sup>106</sup> A large proportion (~12%) of bacterial species with sequenced genomes fall into this group, indicating that there may be more variability in the sequences of the three PanK types than is currently known. Furthermore, as noted above, no PanK enzyme has yet been characterized from an archaeobacterial source, suggesting the existence of a possible fourth type of PanK, PanK-IV.

#### 7.11.3.3.6 Pantothenate kinase as a salvage enzyme

One of the alternative activities of PanK enzymes that have only recently been formally characterized is their ability to phosphorylate the advanced CoA biosynthetic intermediate, pantetheine **12** (Equation (3)).



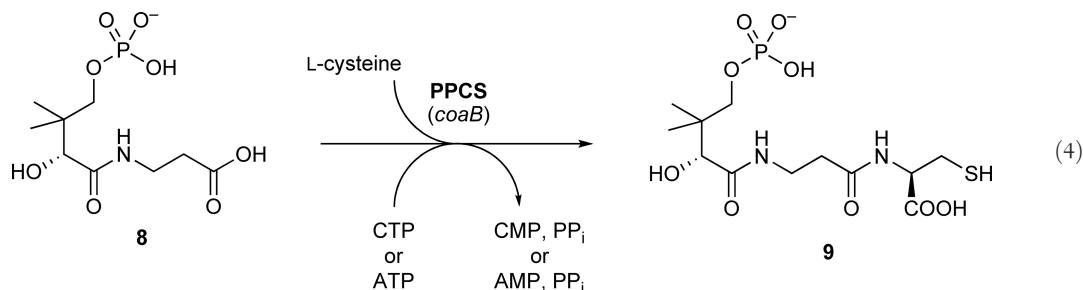
The sections highlighting the earliest studies on CoA enzymology recounted the initial proposal that pantetheine – and not pantothenic acid – is the CoA biosynthetic intermediate that undergoes 4'-phosphorylation. This conclusion was reached based on the isolation of the so-called pantetheine kinase from pigeon liver<sup>32</sup> and the finding that the *M. morgani* PanK enzyme also acts on pantetheine.<sup>34,36</sup> In combination with the studies that showed that *E. coli* strains are able to utilize pantetheine to sustain their CoA biosynthetic needs,<sup>60</sup> these observations all suggest that in many organisms a pantetheine shunt exists as a salvage pathway for CoA biosynthesis, as shown in Scheme 2. The main requirement for such a salvage pathway would be the presence of a PanK enzyme with low substrate specificity that is able to perform the phosphorylation of pantetheine in addition to that of its natural substrate, pantothenate. Such a lack of substrate specificity has since been

demonstrated for most well-studied type I and type II PanK enzymes (especially the proteins from *E. coli* and *S. aureus*) that act on a variety of pantothenamide analogues as alternative substrates.<sup>65,100,114</sup> The alternate activity of the *E. coli* PanK-I enzyme has been characterized and was shown to have a  $K_m$  for pantetheine of  $91 \mu\text{mol l}^{-1}$  with a corresponding  $k_{\text{cat}}$  of  $0.32 \text{ s}^{-1}$ . These values translate into a specificity constant ( $k_{\text{cat}}/K_m$ ) of  $3.53 \text{ s}^{-1} \text{ mmol l}^{-1}$ , nearly 6 times lower than the value of its natural substrate.<sup>65,115</sup>

The action of PanK as a salvage kinase has been proposed to be essential in pathogenic organisms such as the *Mycoplasma* spp. and *Treponema pallidum*, in which no PPCS- or phosphopantothenoylcysteine decarboxylase (PPCDC)-encoding genes can be identified as yet.<sup>112</sup> In the absence of these activities such organisms would be completely reliant on the ability of their PanK enzymes to phosphorylate pantetheine to provide **10** as substrate of the phosphopantetheine adenylyltransferase (PPAT) enzyme. However, both these bacteria are predicted to have type III PanKs, which have shown to have high substrate specificity and which do not act on the pantothenamides as substrates. In fact, *in vitro* tests of the heterologously expressed and purified *T. pallidum* PanK-III enzyme failed to demonstrate any observable activity on pantetheine as substrate, although it was active on pantothenate (L. A. Brand and E. Strauss, unpublished results). This suggests that only PanK-I (*coaA* gene products) and type II PanK enzymes are able to act as salvage kinases and that there must be another means whereby organisms such *T. pallidum* satisfies its CoA biosynthetic needs.

#### 7.11.3.4 PPCS (*coaB* Gene Product)

The second enzyme of the CoA biosynthetic pathway, PPCS (EC 6.3.2.5), catalyzes the  $\text{Mg}^{2+}$ -dependent formation of 4'-phosphopantothenoylcysteine **9** (PPC) from 4'-phosphopantothenate **8** and L-cysteine (Equation (4)).<sup>10</sup> It is often also referred to as the CoaB protein. Two forms of the enzyme can be distinguished. The first, which predominates in bacteria, uses CTP to activate the carboxylate of **8** through the formation of an acyl-cytidylate intermediate.<sup>9</sup> This bacterial form of the enzyme is normally fused to the next enzyme of the pathway, PPCDC, to form a bifunctional CoaBC protein (*coaBC* product, previously known as *dfp*). In contrast, the enzyme that mainly occurs in eukaryotes is monofunctional and uses ATP for carboxylate activation.<sup>82,116–119</sup> Interestingly, the PPCS enzymes from the Streptococci and Enterococci are similarly predicted to be monofunctional.<sup>106</sup>



The PPCS enzyme was first cloned and characterized in *E. coli* as a result of an attempt to isolate the PPCDC enzyme in this bacterium.<sup>9</sup> This study showed that there is close interaction between the PPCS and the PPCDC enzymes, and suggested that 4'-phosphopantothenic acid **8** is the physiological substrate for the isolated protein. Subsequent cloning of the *coaBC* gene confirmed the bifunctional nature of the CoaBC protein and allowed its overexpression and kinetic characterization. Using a coupled assay based on the release of pyrophosphate the following apparent kinetic parameters were determined:  $K_m(4'\text{-phosphopantothenate})$   $55 \mu\text{mol l}^{-1}$ ,  $K_m(\text{CTP})$   $106 \mu\text{mol l}^{-1}$ ,  $K_m(\text{L-cysteine})$   $109 \mu\text{mol l}^{-1}$ , and average  $k_{\text{cat}}$   $2.75 \text{ s}^{-1}$ .<sup>120</sup> These values are lower than those determined by using a  $\text{CO}_2$  assay that couples the production of product **9** to its decarboxylation by the PPCDC enzyme; this suggests that the association of the PPCS and PPCDC activities was not as close as originally thought.<sup>9</sup> This conclusion was confirmed by a study which showed that the product of the PPCS reaction is released from the protein before it returns to a different active site for transformation by the PPCDC activity, suggesting that the fusion of the two activities in bacteria does not have major functional significance.<sup>121</sup> The requirement for CTP, which was first described by Brown in his studies of

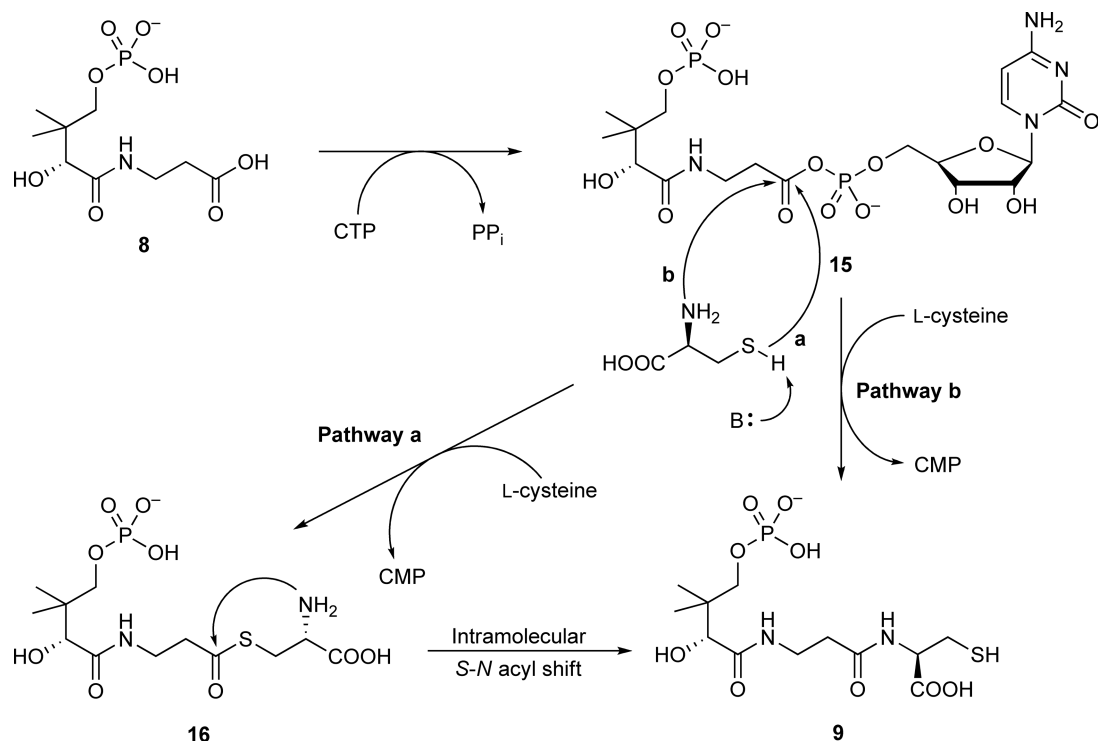
the partially purified protein from *M. morgani*,<sup>36,37</sup> is a unique characteristic of bacterial PPCS enzymes as no other synthetase enzyme that uses CTP in transient fashion has been described. Little is known about the metabolic relevance of this preference.

Studies on the CoaB domains of the bifunctional CoaBC proteins from *E. coli*<sup>122,123</sup> and *Methanocaldococcus jannaschii*,<sup>124</sup> the eukaryotic CoaB proteins from human<sup>116</sup> and *A. thaliana*,<sup>82</sup> and the singular monofunctional bacterial CoaB from *Enterococcus faecalis*,<sup>125</sup> have since been performed. These confirm the preference of all the bacterial (both eubacterial and archaeal) proteins for CTP and the eukaryotic enzymes for ATP. They also show that the nucleoside triphosphate is cleaved to the nucleoside monophosphate and pyrophosphate during the reaction, which suggested that activation of the substrate carboxylate occurred by the formation of a transient acyl-cytidylate (or acyl-adenylate). Studies of site-directed mutants of the *E. coli* protein, which are defective in the utilization and coupling of cysteine, showed that a phosphopantothenoyl-cytidylate intermediate indeed accumulates, confirming this proposal.<sup>122</sup> The same result was found for the archaeal protein<sup>124</sup> and strongly suggested by the studies of *A. thaliana* PPCS (AtCoaB).<sup>82</sup> These results are in contrast with the original studies of Abiko who found that the rat liver enzyme uses ATP and forms ADP and phosphate during the reaction, which suggested the intermediacy of an acyl-phosphate intermediate in this case.<sup>117–119</sup> However, no independent verification of this result has been made, and it is unlikely that a PPCS enzyme using such an alternate form of activation exists uniquely in rats.

Crystal structures of the native monofunctional human CoaB protein<sup>126</sup> and of site-directed mutants of the CoaB domain of *E. coli*'s bifunctional CoaBC have been determined.<sup>127</sup> The structures of the *E. coli* protein with CTP, CMP, or the acyl-cytidylate intermediate bound have also been solved. A direct structural comparison of the bacterial and eukaryotic proteins shows that both molecules are dimers, although the human protein contains an insertion that introduces a large new dimer interface contact area, which should significantly increase dimer stability. In bacterial proteins, such added stability would be largely unnecessary as their major oligomeric interactions occur through the CoaC domains of the native bifunctional protein (see Section 7.11.3.5). The nucleotide-binding site of the human enzyme is larger, as it would be expected for an ATP-utilizing enzyme, and possibly also more flexible. One face of the binding pocket provides hydrogen bonds and a shape that would complement either the adenine or the cytosine moieties of ATP and CTP, respectively. Finally, both enzymes seem to share the same mode of substrate binding and also mechanism, with the carboxylate of phosphopantothenic acid **8** ideally positioned for attack of the  $\alpha$ -phosphate of either CTP or ATP. Although no structure with cysteine bound was determined, analysis of the structure with the phosphopantothenoyl-cytidylate intermediate bound shows the limited options for binding of the cysteine in a manner that would allow direct nucleophilic attack on the activated carboxyl.

A single detailed kinetic study of the PPCS enzyme from *E. faecalis* was conducted,<sup>125</sup> which showed that the enzyme's kinetic mechanism can be described as Bi Uni Uni Bi Ping Pong, with CTP as the first substrate to bind, followed by the binding of **8**, the release of pyrophosphate, the binding of cysteine, the release of product **9**, and finally the release of CMP. It, therefore, essentially has a mechanism identical to that of the pantothenic acid biosynthetic enzyme, PS.<sup>128–131</sup> In the same study, the presence of the acyl-cytidylate intermediate was directly confirmed through <sup>31</sup>P NMR studies that showed the transfer of oxygen from [*carboxyl*-<sup>18</sup>O]-**8** to CMP. Taken together, these data allow for the catalytic mechanism shown in **Scheme 3** to be put forward for bacterial PPCS enzymes. Eukaryotic enzymes are expected to follow the same mechanism, except that the initial activation of the carboxyl group would be mediated by ATP and not CTP.

An important distinction has to be made concerning the mode of nucleophilic attack of the cysteine on the phosphopantothenoyl-cytidylate intermediate **15**, as two pathways are possible for this attack. In the first (pathway **a** in **Scheme 3**), cysteine's thiol reacts with the activated carboxyl to form a thioester intermediate **16**, which will subsequently undergo an intramolecular *S*→*N* acyl shift to give **9**. In the second, the amino group of cysteine attacks the activated acyl group of **15** directly (pathway **b**). Because no structure of any PPCS enzyme has been determined to date in which cysteine or an analogue is bound in the active site, and because analysis of the available structures gives little indication of the potential binding interactions the amino acid might have with the protein, it is still unclear which of these pathways predominantly operates.

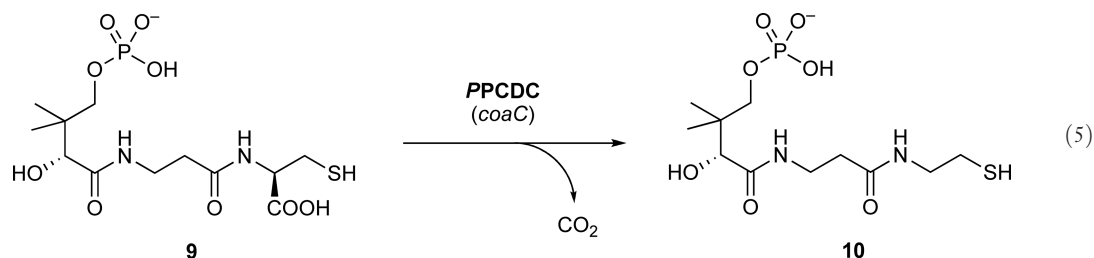


**Scheme 3** Proposed mechanism for the bacterial PPCS enzyme based on current knowledge. Two pathways are possible for the formation of the final product **9** from the acyl-cytidylate intermediate **15**, and it is still unclear which predominates.

Their distinction is important, however, especially in light of the finding that the enzyme is highly selective for the incorporation of cysteine compared with its close structural analogue serine, from which it differs only in the exchange of the thiol group for a hydroxyl.<sup>132</sup> Experiments have shown that the protein will only incorporate cysteine even in the presence of excess serine with a cysteine/serine selectivity ratio of at least  $5.0 \times 10^5$  (the assay's limit of detection). This high level of selectivity is to be expected as the incorporation of serine may lead to the formation of the potentially toxic analogue oxy-CoA,<sup>133</sup> and raises the question as to the mechanism whereby this selectivity is achieved. One proposal has it that, by using pathway **a**, PPCS enzymes may select cysteine by using the difference in  $pK_a$  values between thiols and alcohols and the higher nucleophilicity of thiolates to form the thioester **16** first, which subsequently rearranges to the thermodynamically more stable amide. Unfortunately, no experimental validation of this proposed mode of selection has been achieved as yet.

### 7.11.3.5 Phosphopantothenoylecysteine Decarboxylase (*coaC* Gene Product)

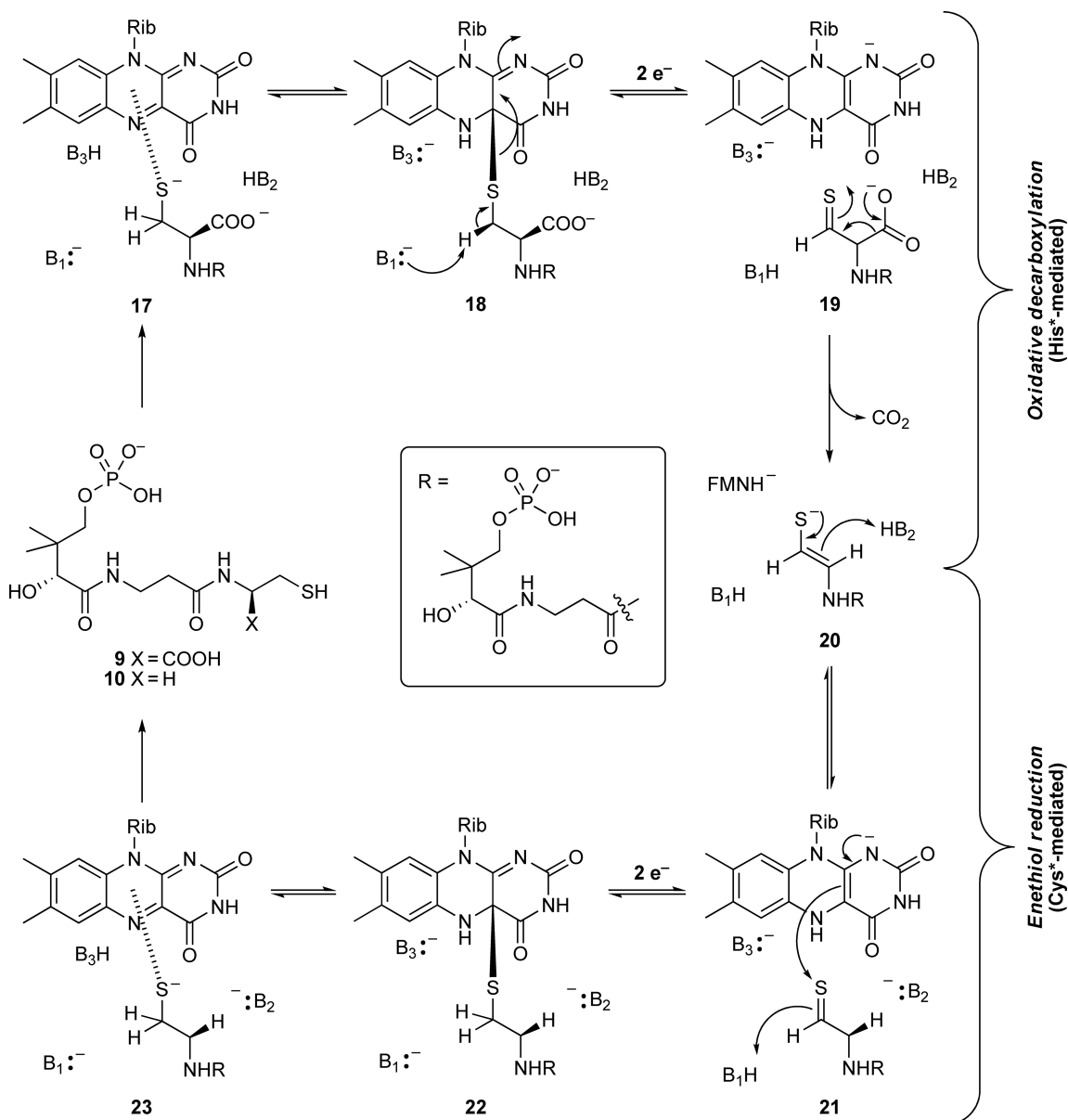
The phosphopantothenoylecysteine decarboxylase (PPCDC; EC 4.1.1.36) activity has proved to probably be the most elusive enzyme of the CoA biosynthetic pathway with regard to both its identification and its mechanism. The enzyme catalyzes the decarboxylation of the cysteine moiety of the substrate PPC **9** to 4'-phosphopantetheine **10** (Equation (5)), which would imply that during catalysis a negative charge develops on the carbon adjacent to the amide nitrogen, and it was unclear how this charge could be stabilized or how any of the known decarboxylase mechanisms could operate in its case.<sup>10</sup> Because the substrate does not have a free amine, the reaction was unlikely to be pyridoxal-5'-phosphate (PLP)-dependent; however, later studies suggested that the horse liver and *E. coli* enzymes<sup>134–137</sup> are dependent on a pyruvoyl cofactor that normally also requires the formation of an intermediate Schiff base.<sup>138</sup> The enzyme's mechanistic mystery was finally solved when the *E. coli* enzyme was cloned and characterized in 2000.<sup>139</sup>



*Escherichia coli*'s PPCDC is part of a bifunctional protein that also harbors PPCS activity and is therefore also referred to as the CoaBC protein (*coaBC* or *dfp* gene product).<sup>9,139</sup> Surprisingly, the enzyme was found not to be pyruvoyl-dependent but to have a tightly bound flavin mononucleotide as cofactor instead, which suggested the involvement of a redox process during the course of catalysis. In fact, the *E. coli* enzyme's CoaC domain exhibits a high level of sequence homology to a group of lantibiotic biosynthetic enzymes (EpiD<sup>140</sup> and MrsD<sup>141,142</sup> being the best characterized of these), which all catalyze the decarboxylation of a cysteine residue located at the C-terminus of an oligopeptide.<sup>143</sup> Although these proteins are also all flavin-dependent, they differ from PPCDC enzymes in that their substrates are oxidatively decarboxylated and form enethiol products and reduced flavin instead. It was therefore concluded that PPCDC enzymes and the lantibiotic decarboxylases such as EpiD share a similar mechanism for the initial oxidative decarboxylation of their substrates, whereby the cysteine thiol is oxidized by the flavin cofactor to form a thioaldehyde. This intermediate would undergo spontaneous decarboxylation (in a manner similar to  $\beta$ -keto acids) to give an enethiol product. However, at this point the enzymes' mechanism diverges: in EpiD the enethiol is the final product, with the resulting reduced flavin most probably reacting with oxygen to form H<sub>2</sub>O<sub>2</sub> as a by-product; in the case of PPCDC enzymes the flavin reduces the enethiol to give a cysteaminylium moiety and oxidized flavin as products. The details of these mechanistic differences have slowly emerged during recent years through various mechanistic, mutagenesis, and structural studies on both PPCDC and lantibiotic biosynthesis enzymes.

In EpiD, formation of the enethiol product has been confirmed by direct mass spectrometric and NMR analysis.<sup>144–147</sup> A crystal structure of an EpiD mutant with a pentapeptide substrate analogue bound<sup>148</sup> and the biochemical characterization of various site-directed mutants subsequently indicated the crucial involvement of a histidine residue in the oxidation of the substrate thiol.<sup>139</sup> Similar mutagenesis studies on the PPCDC activity of the *E. coli* CoaBC protein confirmed that this enzyme had a similar requirement, but also fingered a cysteine residue to be important in the subsequent reduction reaction.<sup>149</sup> Subsequently, the AtHAL3a protein of *A. thaliana* was found to have PPCDC activity, and the corresponding mutants of the monofunctional protein further supported these conclusions.<sup>150,151</sup> A crystal structure of the *A. thaliana* AtHAL3a protein in which this cysteine had been exchanged for a serine in fact found that in the absence of the cysteine the enethiol intermediate became trapped in the active site, confirming its essentiality in the second half-reaction.<sup>152</sup> The chemical mechanisms of both steps have also been investigated extensively in various mechanistic studies, which included making use of solvent exchange reactions, studying the primary kinetic isotope effects of the oxidation reaction as well as the changes in the UV absorbance spectrum of the flavin cofactor upon addition of the substrate.<sup>121</sup> Most recently, a mechanism-based inactivating agent was used to pinpoint the exact residues involved in this reduction step.<sup>153</sup> Taken together, these studies have allowed a general mechanism for PPCDC enzymes to be proposed as shown in **Scheme 4**.

The mechanism of the oxidative decarboxylation is proposed to proceed through the charge–transfer complex **17**, which forms by deprotonation of the thiol of the PPC **9** substrate. The charge–transfer complex is in equilibrium with the flavin C(4a)–thiol adduct **18**, which activates the thiol for oxidation by deprotonation of the carbon adjacent to the thiol and electron transfer to the cofactor. The thioaldehyde **19** formed in this manner can spontaneously decarboxylate to form the enethiol **20**. The essential histidine residue is proposed to be involved in these first steps, as exchanging it to an asparagine residue abolishes all activity. Based on structural evidence, this histidine most probably fulfills the role of the base-labeled B<sub>3</sub> in **Scheme 4**, which is involved in deprotonation of the thiol. However, the identity of the base B<sub>1</sub> is less clear: site-directed mutagenesis and structural studies both point to an asparagine residue for this role, as its side chain amide is in proximity to the hydrogen atoms of the carbon adjacent to the thiol. Such a conclusion seems chemically counterintuitive



**Scheme 4** Proposed mechanism for the PPCDC enzymes based on current knowledge. The reaction takes place in two parts. The first part, mediated by an active site His-residue, involves oxidation of the substrate thiol to a thioaldehyde, followed by spontaneous decarboxylation. In the second part the enethiol is reduced back to the alkanethiol to form the product; this step is dependent on an active site cysteine. The identity of the active site bases and acids labeled  $B_1$ ,  $HB_2$ , and  $B_3$  are discussed in the text.

because of the exceedingly low basicity of an amide coupled with the low acidity of the methylene protons (even when activated by formation of the flavin–thiol adduct). Nonetheless, with the limited structural information that is currently available no alternative base can be proposed, and this contradiction will therefore have to be addressed by directed mechanistic studies.

The reduction of the enethiol is the step that differentiates PPCDC enzymes from their lantibiotic biosynthesizing counterparts. Here, the key difference is proposed to center on the availability of an active site acid ( $HB_2$  in **Scheme 4**), which is able to protonate the enethiol to promote the formation of its thioaldehyde tautomer.<sup>153</sup> This proposal is based on the mechanistic evidence that point to the enethiol

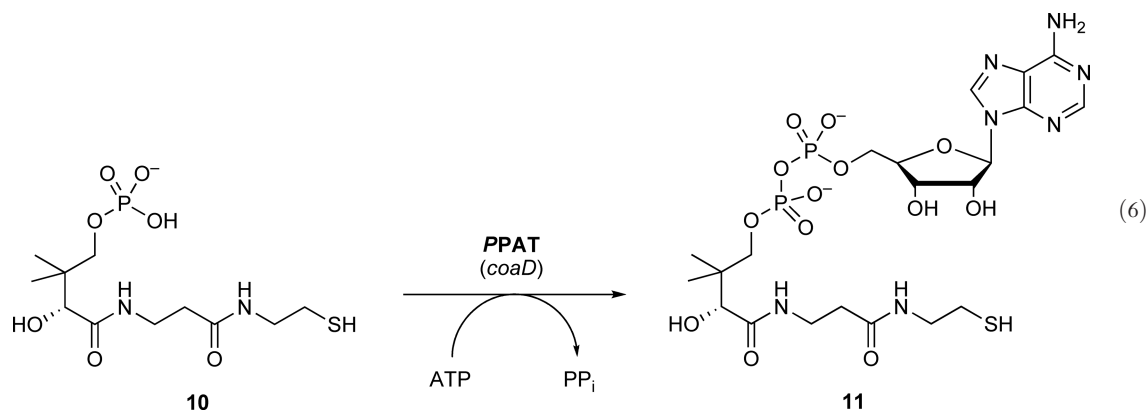


reduction occurring through the formation of the flavin-product C(4a)-thiol adduct **22**, and the charge-transfer complex **23**; this can only happen by electron transfer from the reduced flavin to the sulfur atom of the thioaldehyde **21**. However, enethiol-thione equilibria normally favor the formation of the enethiol, and without the presence of an active site acid the thioaldehyde is therefore unlikely to form.<sup>154,155</sup> All evidence suggests that the active site cysteine residue that was first identified by mutagenesis and structural studies fulfills the role of this acid. This analysis is supported by the lantibiotic biosynthesis enzymes EpiD and MrsD having the residue exchanged for the weaker acids serine and threonine, respectively, and the fact that the AtHAL3a protein in which this cysteine is mutated to serine also accumulates the enethiol intermediate. However, the proposed active site cysteine is found as part of the substrate-binding clamp of the enzyme, which is disordered unless the substrate is bound; in the mutants in which it is exchanged for serine it also seemed to be very far from the  $\alpha$ -carbon of the enethiol for optimal proton transfer.<sup>152</sup> Such ambiguity as to the identity of acid HB<sub>2</sub> was removed by a study in which a mechanism-based electrophilic trap was used to covalently label the residue in question. This experiment was performed on the human CoaC protein and found that it is indeed the corresponding cysteine residue that was labeled.<sup>153</sup> Moreover, the labeled enzyme was also found to be unable to reduce the enethiol analogue that forms upon inactivation, which lends further support to this proposal. Finally, the protonation of the enethiol proceeds in such a manner as to return the hydrogen to the same face as the one from which CO<sub>2</sub> departed (as shown in **Scheme 4**), because two independent studies have confirmed that the reaction proceeds with retention of configuration.<sup>156,157</sup>

As mentioned in the previous section on the PPCS enzyme, most prokaryotic PPCDC enzymes are part of bifunctional proteins that also have PPCS activity, the exception being the Streptococci and Enterococci, which are proposed to have separate monofunctional proteins.<sup>46,111</sup> Bacterial CoaBC proteins form homododecamers (tetramers of trimers) through interaction of their CoaC domains, a structural feature they share with the lantibiotic decarboxylases.<sup>139,148</sup> Eukaryotic proteins are all proposed to be monofunctional, trimeric proteins.<sup>152</sup> Recently, the CoaBC protein from the archaeobacterium *M. jannaschii* was also cloned and characterized, and showed that this protein is also homododecameric.<sup>124</sup> However, it lacks the typical substrate recognition sequences present in other CoaC proteins, and – although it has the active site histidine residue (H87) associated with deprotonation and oxidative decarboxylation of the substrate – no residue could be identified corresponding with the cysteine residue implicated to be the active site acid involved in the reduction of the enethiol intermediate. Instead, this protein had a glutamic acid (E168), a more acidic residue, in the corresponding position. This exchange is still in agreement with the mechanism shown in **Scheme 4** but indicates that the identities of the active site acids and bases B<sub>1</sub>, HB<sub>2</sub>, and B<sub>3</sub> are not a fixed feature of PPCDC enzymes, and that they are not necessarily always associated with His- and Cys-residues as explained above. This discovery suggests that PPCDC enzymes may exhibit some minor variations on the mechanism described here.

### 7.11.3.6 Phosphopantetheine Adenylyltransferase (*coaD* Gene Product)

The penultimate CoA biosynthetic enzyme catalyzes the reversible Mg<sup>2+</sup>-dependent adenylylation of 4'-phosphopantetheine **10** to form dephospho-CoA **11** and pyrophosphate as products (Equation (6)). It is known as phosphopantetheine adenylyltransferase (PPAT; EC 2.7.7.3) or as the CoaD protein. In most eukaryotes, the PPAT activity is fused to the last enzyme of the pathway, dephospho-CoA kinase (DPCK), to form a bifunctional PPAT/DPCK protein that is also referred to as CoA synthase.<sup>158</sup> In bacteria a single enzyme was shown to perform the reaction.<sup>159</sup> Identification and characterization of the PPAT activity of the bifunctional human protein (by comparative genomics<sup>116</sup> and by traditional methods<sup>160</sup>), of the bifunctional murine enzyme,<sup>161</sup> and of the monofunctional enzyme from *A. thaliana*,<sup>82</sup> all showed that the eukaryotic proteins share little sequence similarity with the bacterial enzymes, indicating the existence of two distinct nucleotidyltransferases that catalyze the adenylylation of 4'-phosphopantetheine. Interestingly, archaea are predicted to have an eukaryotic-type PPAT as no homologue of the bacterial enzyme can be identified in any sequenced archaeal genome.<sup>111</sup> The cloning and expression of the *Pyrococcus abyssi* enzyme has confirmed this prediction, although unlike other characterized PPAT enzymes it behaves as a monomer in solution.<sup>162,163</sup>



The bacterial CoaD protein (*coaD* gene product, previously *kdtB*) from *E. coli* was the first PPAT enzyme to be cloned and fully characterized.<sup>164</sup> The enzyme was purified with CoA tightly bound, which gave the first indication that the PPAT activity is the second enzyme in the biosynthetic pathway to be regulated by feedback inhibition. The same study, and the subsequent crystal structures of the enzyme,<sup>165–167</sup> shows that the protein exists as a homohexamer made up of a dimer of two distinct trimers in solution. Interestingly, the enzyme was shown to have half-of-sites reactivity, as both the substrates (ATP and phosphopantetheine) and the product dephospho-CoA preferentially bind to only one of these trimers. The reaction is predicted to proceed by direct nucleophilic attack of 4'-phosphate of **10** on the  $\alpha$ -phosphate of ATP in an in-line displacement mechanism that would result in the formation of a pentacoordinate transition state, which is stabilized by interactions with active site residues, especially a conserved histidine residue. CoA regulates the enzyme's activity allosterically by binding to the same site as the product **11**, but in a distinct manner. Only the binding sites of the 4'-phosphate groups of these two molecules overlap, whereas their pantetheine arms take on two different conformations and their adenylyl moieties bind to different sites because of the presence of the 3'-phosphate group of CoA. CoA binds differently to the two protomers as the binding of the first molecule influences the conformation of the other protomer. One of these binding modes resembles the product conformation closely.

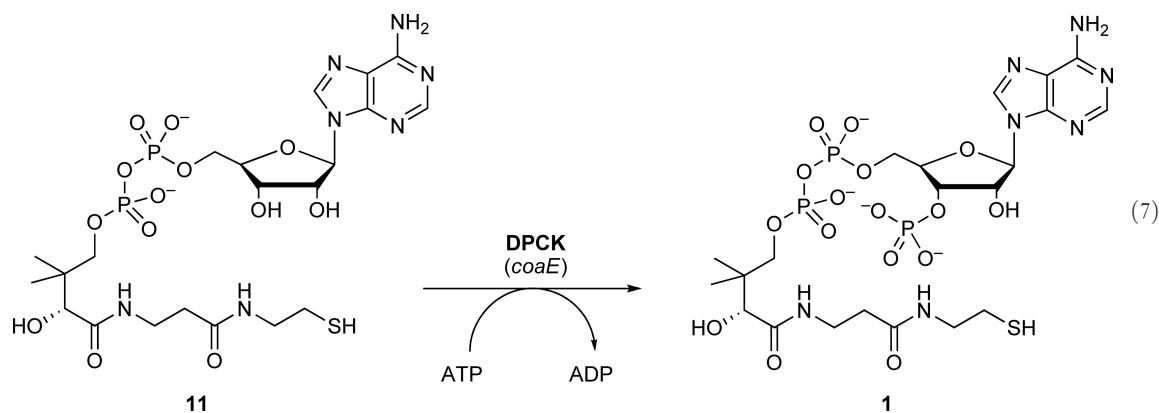
A recent detailed study of the kinetic mechanism of the *E. coli* enzyme using steady-state kinetics and ITC shows that the enzyme has a random bi–bi mechanism in which a ternary complex of enzyme, ATP, and phosphopantetheine are formed. The following apparent kinetic parameters for the forward reaction were determined:  $K_m$   $220 \pm 10 \mu\text{mol l}^{-1}$  and  $k_{\text{cat}}$   $1.59 \pm 0.01 \text{ s}^{-1}$  (for ATP), and  $K_m$   $4.7 \pm 0.5 \mu\text{mol l}^{-1}$  and  $k_{\text{cat}}$   $1.37 \pm 0.07 \text{ s}^{-1}$  (for 4'-phosphopantetheine **10**). The binding of dephospho-CoA was also shown to be competitive with both ATP and phosphopantetheine, whereas pyrophosphate was competitive with ATP (product inhibition of pyrophosphate vs. phosphopantetheine could not be established). CoA binding is competitive with ATP, phosphopantetheine, and dephospho-CoA. Interestingly, this study found that the dissociation constant of CoA from the free enzyme is  $65 \text{ nmol l}^{-1}$ , although the  $K_i$  values were much higher, in the range of  $10\text{--}50 \mu\text{mol l}^{-1}$ . This may be due to the two different modes of CoA binding observed in the binary CoA–PPAT structures.<sup>167</sup> Acetyl-CoA only exhibits weak inhibition, suggesting that free CoA (and not its thioesters) is the main feedback inhibitor of the enzyme. Taken together, the kinetic data suggest that in light of the low  $K_m$  for phosphopantetheine and the mediocre  $K_i$  for CoA, regulation of PPAT by CoA inhibition is likely to be only physiologically relevant when CoA levels are high and phosphopantetheine levels are low. Finally, the ITC experiments found that the substrates, product, and the feedback inhibitor all bind to the enzyme in a ratio of one ligand per monomer, which is in contradiction to the previous finding of half-of-sites reactivity. However, because the ITC study was performed at pH 8.0 while the previous finding was made at pH 5.0, it is possible that protonation state plays a significant role in substrate binding in this enzyme.

In addition to the structures of the *E. coli* enzyme, structures or preliminary structures of PPAT enzymes from *Thermus thermophilus* (pdb: 1OD6),<sup>168</sup> *M. tuberculosis* (pdb: 1TFU),<sup>169</sup> *E. faecalis*,<sup>170</sup> *B. subtilis* (pdb: 1O6B),<sup>171</sup> *T. maritima* (pdb: 1VLH), and the bifunctional PPAT/DPCK murine CoA synthase with acetyl-CoA bound (pdb: 2F6R) have been determined. The crystal structures of bacterial PPAT enzymes show that they display sequence homology to members of the superfamily of nucleotidyltransferase  $\alpha/\beta$  phosphodiesterases, which

include cytidylyltransferases, class I aminoacyl-tRNA synthetases, and also PS.<sup>166</sup> Members of this family all have a mononucleotide-binding fold and a conserved H/TXGH motif that contains the histidine residue that is involved in the stabilization of the transition state. Little is known about the structure of the murine enzyme and how its PPAT domain compares to that of the bacterial protein. Although a preliminary analysis suggests that it also belongs to the same superfamily,<sup>172</sup> a detailed analysis would greatly aid current drug development efforts that have targeted the essential bacterial PPAT enzyme for antimicrobial development.

### 7.11.3.7 Dephospho-Coenzyme A Kinase (*coaE* Gene Product)

Dephospho-CoA kinase (DPCK; EC 2.7.1.24) catalyzes the selective MgATP-dependent phosphorylation of the 3'-hydroxyl of the ribose moiety dephospho-CoA **11** to form CoA **1**, the final product of the biosynthetic pathway (Equation (7)). As pointed out in the previous section, most eukaryotic organisms have the DPCK activity located with PPAT on a bifunctional protein referred to as CoA synthase.<sup>38,158,173</sup> However, unlike PPAT, the DPCK domain of this protein exhibits good sequence homology with its bacterial counterparts.<sup>116,160,161</sup> In *A. thaliana*, the protein is monofunctional.<sup>82</sup>



The first bacterial DPCK to be cloned and characterized was the *E. coli* CoaE protein (*coaE* gene product, previously *yacE*), which was identified by a basic local alignment search tool (BLAST) using the N-terminal sequence of the wild-type enzyme purified from *C. ammoniagenes*.<sup>174</sup> The enzyme is a 22.6-kDa monomer in solution and exhibits apparent  $K_m$  values of  $140 \mu\text{mol l}^{-1}$  and  $740 \mu\text{mol l}^{-1}$  for ATP and dephospho-CoA, respectively. It displays poor activity with adenosine, AMP, and adenosine phosphosulfate (APS) as alternate phosphoryl acceptors, which all shows less than 8% activity compared with the natural substrate. Studies on the bifunctional PPAT/DPCK protein show that the normally reversible PPAT activity becomes irreversible when it is coupled to DPCK, most probably because the latter activity demonstrates a low  $K_m$  for dephospho-CoA of  $5.2 \pm 1.5 \mu\text{mol l}^{-1}$  (its  $K_m$  for ATP was found to be  $192 \pm 14 \mu\text{mol l}^{-1}$ ).<sup>160</sup> Similar values were determined in an independent study.<sup>116</sup>

The crystal structures of the *E. coli*,<sup>175</sup> *H. pylori*,<sup>176</sup> *T. maritima* (pdb: 2GRJ), *Aquifex aeolicus* (pdb: 2IF2), *H. influenzae*,<sup>177</sup> and *T. thermophilus*<sup>178</sup> DPCK proteins have been determined, the latter two with ATP-bound. The crystal structure of the eukaryotic bifunctional PPAT/DPCK protein from *Mus musculus* has also been determined, and the coordinates are deposited in the PDB (pdb: 2F6R). These structures show that similar to type I PanK enzymes, DPCK proteins belong to the family of P-loop kinases, which include the various nucleoside and nucleotide kinases.<sup>73,74</sup> Although no structure with the dephospho-CoA **11** substrate bound is currently available, the structure of **11** has been modeled in the putative active site of the *H. influenzae* protein. This model shows that four of the five residues that are strictly conserved in DPCK sequences are either found in the P-loop (Gly-14 and Lys-15) or involved in either adenine binding (Arg-140) or dephospho-CoA binding (Thr-8). The fifth residue is the putative catalytic base (Asp-33), which is within hydrogen-bonding distance of the 3'-hydroxyl of the ribose in the model. This strongly suggests that the Asp activates the substrate for nucleophilic attack on the  $\gamma$ -phosphate of ATP. On the contrary, *H. influenzae* and *T. thermophilus* DPCK structures show that ATP's  $\gamma$ -phosphate is about  $7 \text{ \AA}$  away from the predicted location of the 3'-OH, indicating that the protein will have to undergo considerable conformational change to bring the two substrates in

proximity. However, such induced-fit movements are not uncommon in members of this kinase family.<sup>179</sup> Recently more structures of the native *E. coli* DPCCK protein (pdb: 1VIY) and the protein in complex with ADP (pdb: 1VHL) and ATP (pdb: 1VHT) have been deposited in the PDB as part of the Structural GenomiX (SGX) bacterial genomics project.<sup>171</sup> However, no detailed comparative analysis of these structures has been published as yet.

### 7.11.3.8 Genomic Variations in the CoA Biosynthetic Pathway

Significant diversity exists among the CoA biosynthetic pathways of various organisms. For example, the discussions of the CoA biosynthesis enzymes above highlighted the existence in some cases of two or more distinct nonhomologous proteins that all exhibit the same activity. The transformation of ketopantoic acid **4** to pantoic acid **5** in the biosynthesis of pantothenic acid can also be catalyzed by one of the two proteins: a strict KPR enzyme (the *panE* gene product; EC 1.1.1.169) or the ketol-acid reductoisomerase (*ilvC* gene product; EC 1.1.1.86), which are involved in isoleucine and valine biosyntheses. As noted previously, not all organisms have the ability to synthesize pantothenic acid *de novo*, and these therefore have to salvage it from the environment, importing it into the cell by means of a pantothenate transporter. Finally, not all organisms have all the genes required for an intact CoA biosynthetic pathway. This may be because some have acquired the ability to transport advanced intermediates of the pathway across their membranes, or because we are still unable to identify these genes by current bioinformatics methods.<sup>106</sup> Taken together, these differences make for the existence of a number of variants with regard to the assembly of the CoA biosynthetic pathway.

The existence of a specific variant and the frequency with which it occurs can be explored by the extensive use of bioinformatics tools as demonstrated in a recent study.<sup>112</sup> Such an analysis shows that when both the pantothenic acid and the CoA biosynthetic pathways of a specific organism – as well as its ability to import pantothenate by means of a specific transporter – are considered, 10 different variants apart from the ‘basic’ form of the pathway can be identified. Some of these can be classified further into a number of subvariants, depending on the occurrence of a specific enzyme type when more than one option is available (e.g., PanK-III is present instead of PanK-I). The basic form was defined as a pathway that includes candidates for all the enzymes of the pantothenic acid and CoA biosynthetic pathways, as well as a pantothenate transporter for salvage of the vitamin. Of the variants that deviate from this basic form, two are abundantly represented among the organisms surveyed: variant 1 includes organisms in which no enzymes required for *de novo* pantothenate biosynthesis can be identified, whereas organisms that have variant 2 of the pathway do not have the ability to import pantothenate (as determined by the nonidentification of a putative pantothenate transporter). Together these three variants occur in ~80% of the organisms that were analyzed. Among the remaining variants are examples of both functional and nonfunctional pathways as defined by the predicted ability of an organism to produce CoA. The nonfunctional variants pose some especially interesting biochemical questions with regard to the manner in which such organisms satisfy their CoA requirements. For example, one variant (variant 3) representing the Chlamydiaeae, Rickettsiaceae, and some of the Mycoplasmaceae predict the presence of only a functional DPCCK enzyme, suggesting that these organisms are dependent on the salvage of dephospho-CoA by some unknown transporter. Another (variant 4), found in the *Mycoplasma* spp. and in *T. pallidum*, have predicted PanK-III, PPAT, and DPCCK enzymes, but no PPCS- or PPCDC-encoding genes. Such a variation predicts the operation of a pantetheine shunt, which would require the uptake of pantetheine and its subsequent phosphorylation by the PanK enzyme to produce 4'-phosphopantetheine as substrate for the rest of the pathway. A full biochemical characterization of such underrepresented variants of the CoA pathway may significantly add to our understanding of the biology of this cofactor; however, to date little work has been done in this regard. Finally, the archaea are considered to be representative of a fifth variant in which no known ADC, PS, PanK, or DPCCK enzymes can be identified by homology. However, as some studies have used bioinformatics to propose putative candidate proteins for these activities,<sup>111</sup> it may well be that the archaea could eventually be considered a mere subvariant of some of the more common variant forms.

**Figure 2** provides an overview of some of these variants (and their subvariants) by comparing the reconstructed pantothenic acid and CoA biosynthetic pathways in a variety of organisms (including some of those mentioned above) to the known pathways in *E. coli*.

Step	Enzymes involved in pantothenic acid and CoA biosynthesis	EC number	Basic form (with subvariants)				Variant 1 (no <i>de novo</i> pantothenate)		Variant 2 (no PanF)		V. 3	V. 4	V. 5	
			<i>Escherichia coli</i>	<i>Staphylococcus aureus</i>	<i>Pseudomonas aeruginosa</i>	<i>Bacillus subtilis</i>	<i>Bacillus cereus</i>	<i>Homo sapiens</i>	<i>Streptococcus pneumoniae</i>	<i>Corynebacterium glutamicum</i>	<i>Mycobacterium tuberculosis</i>	<i>Mycoplasma genitalium</i>	<i>Treponema pallidum</i>	<i>Pyrococcus furiosus</i>
<b>Pantothenic acid Biosynthesis</b>														
1	Ketopantoate hydroxymethyltransferase (KPHMT)	2.1.2.11	<i>panB</i>	+	+	+	+	-	-	+	+	-	-	+
2	Ketopantoate reductase (KPR)	1.1.1.169	<i>panE</i>	+	+	+	+	-	-	+	+	-	-	+
3	Ketol-acid reductoisomerase (KAR)	1.1.1.86	<i>ilvC</i>	+	+	+	(+)	-	-	+	-	-	-	+
4	Aspartate 1-decarboxylase (ADC)	4.1.1.11	<i>panD</i>	+	+	+	+	-	-	+	+	-	-	-
5	Pantothenate synthetase (PS)	6.3.2.1	<i>panC</i>	+	+	+	+	-	-	+	+	-	-	-
<b>Pantothenate transporter</b>														
6	Pantothenate:Na <sup>+</sup> symporter (or similar) (PanF)		<i>panF</i>	+	+	+	+	+	?	-	-	-	-	+
<b>CoA biosynthesis</b>														
	Pantothenate kinase (PanK)	2.7.1.33												?
7a	Type I		<i>coaA</i>	-	-	+	-	-	+	+	+	-	-	-
7b	Type II			<i>coaA</i>	-	-	+	+	+	-	-	-	-	-
7c	Type III				<i>coaX</i>	+	+	-	-	-	+	-	+	-
8	Phosphopantothencysteine synthetase (PPCS)	6.3.2.5	<i>coaBC</i>	+	+	+	+	+	+	+	+	-	-	+
9	Phosphopantothencysteine decarboxylase (PPCDC)	4.1.1.36												
	Phosphopantetheine adenylyltransferase (PPAT)	2.7.7.3												
10a	Form 1 (Bacterial)		<i>coaD</i>	+	+	+	+	-	+	+	+	-	+	+
10b	Form 2 (Eukaryotic)			-	-	-	-	<i>coaDE</i>	-	-	-	-	-	-
11	Dephospho-CoA kinase (DPCK)	2.7.1.24	<i>coaE</i>	+	+	+	+		+	+	+	+	+	?

**Figure 2** Reconstruction of the pantothenic acid and CoA biosynthetic pathways in some selected organisms to demonstrate the genomic variations that occur in the pathway. The pantothenic acid and CoA biosynthetic enzymes are listed on the left, with the corresponding genes in *E. coli* listed in the first column. The presence or absence of putative orthologues in other organisms is indicated with '+' and '-'. A question mark '?' indicates that the specific activity must be present for the organism to have a functional CoA pathway, but cannot be identified by bioinformatics. Bifunctional enzymes are indicated by squares bisected by a diagonal line. Organisms are grouped according to the number of predefined variations in the pathway: the 'basic form' of the pathway includes organisms capable of *de novo* pantothenic acid biosynthesis, pantothenic acid salvage through a pantothenate transporter, and CoA biosynthesis. Variant 1 includes organisms which do not have functional pantothenic acid biosynthetic pathways. Variant 2 includes organisms in which no pantothenate transporter can be identified. Variants 3 and 4 have nonfunctional pathways, although organisms having variant 4 of the pathway may be able to biosynthesize CoA by the uptake of pantotheine followed by its subsequent phosphorylation by PanK-III as a salvage kinase. Archaea are considered to represent a fifth variant because of the missing PanK and DPCK encoding genes in its CoA pathway. Various subvariants also occur within these general groupings; some of these are exemplified by the specific organisms shown in the figure. More detailed information on these variants and on the identification of the putative orthologues is provided in the original study.<sup>112</sup>



### 7.11.3.9 Essentiality of Genes Encoding the CoA Biosynthetic Enzymes

Current knowledge suggests that nearly all organisms have to produce all their CoA requirements through *de novo* biosynthesis from pantothenic acid, as none of the biosynthetic intermediates can normally be taken up by cells. Because CoA itself is also an essential cofactor, it therefore follows that at least some of the genes encoding the CoA biosynthetic enzymes should be essential in such organisms (or at least conditionally essential). The only possible exceptions are the PPCS and PPCDC enzymes, which together produce 4'-phosphopantetheine from 4'-phosphopantothenate, as they may be bypassed if the organism has a PanK with a salvage activity (i.e., the ability to act on pantetheine as substrate, see Section 7.11.3.3.6), and can obtain pantetheine from the environment. However, a close inspection of the various studies aimed at determining the set of essential genes in various organisms indicates that such predictions of essentiality are not always borne out by the experimental data. Furthermore, the definition of 'essentiality' is not a fixed one and varies between studies. Nonetheless, a comparative study of the known gene-essentiality data of specific metabolic pathways across various organisms shows that among the genes encoding the CoA biosynthetic enzymes there are nearly always at least one which can be described as essential.<sup>180</sup> A summary of the CoA biosynthetic genes that have been found to be essential by at least one study is provided in Table 3; however, the reader must take note that other studies (using different conditions and experimental techniques) may have found the same genes to be nonessential, or conditionally essential. Interestingly, one organism in which CoA biosynthesis seems to be nonessential is the bacterium with the so-called minimal genome, *Mycoplasma genitalium*. Of the five biosynthetic enzymes, this organism only has a DPCK (and even this enzyme does not seem to be essential), and it therefore must source CoA (or dephospho-CoA) from the environment.<sup>181</sup> For an up-to-date overview of the status of the gene-essentiality data, the SEED database should be consulted (<http://seed-viewer.theseed.org/>; *Subsystems: Coenzyme A Biosynthesis*).

Although these essentiality studies are extremely helpful in explaining certain phenotypes, as well as for the design of specific drug development programs, it is important to consider the specific experimental conditions under which and by which the essentiality was determined. Also, most of these studies are aimed at determining the minimum set of essential genes of the whole organism and not at determining the essentiality of their CoA biosynthetic genes *per se*. In this regard, the data should be carefully analyzed before a final conclusion about the essentiality of a specific CoA biosynthetic gene is reached. A case in point is the apparent essentiality of the *coaA* gene of *M. tuberculosis*, which encodes its type I PanK enzyme. This result, as determined by a transposon site hybridization study,<sup>183</sup> seems curious because *M. tuberculosis* also has a PanK-III enzyme; it should therefore be unaffected by the interruption of its *coaA* gene. *Bacillus subtilis* is another organism with both type I and type III PanKs, and a *coaA* mutant of this organism indeed grows normally.<sup>107</sup> In such cases, it would be prudent to confirm the findings through directed follow-up studies.

## 7.11.4 CoA Utilization, Degradation, and Recycling

Apart from the enzymes involved in the biosynthesis of CoA that were highlighted in the previous section, a number of enzymes also deplete the levels of cellular CoA by either modifying the cofactor, catalyzing its degradation, or consuming it during the phosphopantetheinylation of other enzymes or carrier proteins. Cells have to balance these processes and apply them in combination with various other strategies – such as regulation of gene expression and compartmentalization – to regulate the intracellular levels of CoA. For a detailed overview of the processes that contribute to CoA regulation readers are referred to a recent review;<sup>8</sup> this section will focus on the specific enzymes that are involved in the consumption, degradation, and recycling of the cofactor.

### 7.11.4.1 Enzymes Involved in the Consumption of CoA

#### 7.11.4.1.1 Phosphopantetheinyl transferase/ACP synthase

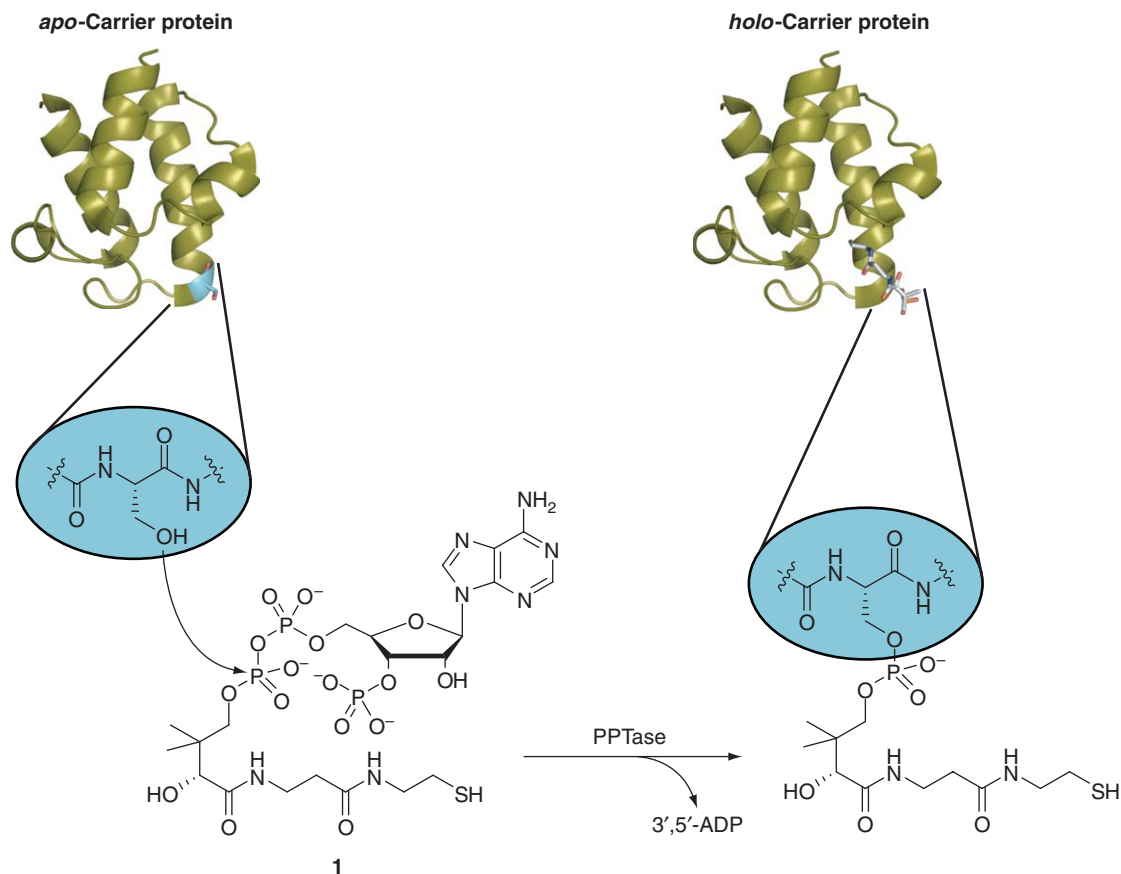
The major net consumers of CoA are the enzymes that transfer the cofactor's 4'-phosphopantetheinyl moiety to a conserved serine residue of *apo*-carrier proteins – such as the ACPs of type II fatty acid synthase (FAS) and



**Table 3** Genes encoding CoA biosynthetic enzymes that have been found to be essential by at least one study<sup>a</sup>

Organism	Enzyme			
	<i>PanK</i>	<i>PPCS/PPCDC</i>	<i>PPAT</i>	<i>DPCK</i>
<i>Escherichia coli</i>	<i>coaA</i> figl83333.1.peg.3890 <sup>b</sup>	<i>coaBC (dfp)</i> figl83333.1.peg.3575	<i>coaD (kdtB)</i> figl83333.1.peg.3570	<i>coaE (yacE)</i> figl83333.1.peg.103
<i>Bacillus subtilis</i>				<i>coaE (ytaG)</i> figl224308.1.peg.2909
<i>Mycobacterium tuberculosis</i>	<i>coaA</i> , Rv1092c figl83332.1.peg.1094	<i>dfp</i> , Rv1391 figl83332.1.peg.1393		Rv1631 figl83332.1.peg.1633
<i>Pseudomonas aeruginosa</i>	figl208964.1.peg.4279 <sup>c</sup>	figl208964.1.peg.5317	figl208964.1.peg.364	figl208964.1.peg.4529
<i>Haemophilus influenzae</i>	figl71421.1.peg.601		figl71421.1.peg.621	
<i>Streptococcus pneumoniae</i>		<i>coaC</i> (PPCDC) only figl171101.1.peg.1110		
<i>Salmonella typhimurium</i>		figl99287.1.peg.3606		
<i>Helicobacter pylori</i>				figl85962.1.peg.822

<sup>a</sup> For a full discussion of these data, refer to Gerdes *et al.*<sup>180</sup><sup>b</sup> Fig numbers refer to the specific gene's ID number in the SEED database.<sup>c</sup> Essentiality data from Liberati *et al.*<sup>182</sup>



**Figure 3** The posttranslational modification reaction catalyzed by PPTase enzymes that transfer CoA's phosphopantetheinyl moiety to a conserved serine residue of a variety of carrier proteins and other enzymes, and form 3',5'-ADP as a by-product.

polyketide synthase (PKS) systems, the acyl carrier domain of type I FAS and PKS systems, and the peptidyl carrier proteins (PCPs) of the nonribosomal peptide synthetases (NRPSs) – to form the respective active *holo*-carrier proteins and 3',5'-ADP (Figure 3).<sup>184</sup> The same posttranslational modification is also required for the activity of some enzymes, such as 10-formyltetrahydrofolate dehydrogenase,<sup>185</sup> the aldehyde oxidoreductase from *Nocardia* spp.,<sup>186</sup> and the  $\alpha$ -amino adipate reductase involved in lysine biosynthesis in fungi.<sup>187,188</sup> The proteins that catalyze this reaction are generally known as 4'-phosphopantetheinyl transferase (PPTase; EC 2.7.8.7) enzymes.

*Escherichia coli*'s *holo*-ACP synthase (AcpS) was the first PPTase to be cloned and characterized, and was subsequently shown to be a member of a new superfamily of PPTase enzymes.<sup>189,190</sup> Three groups of PPTase can be distinguished based on sequence and carrier protein specificity: the first group comprises bacterial *holo*-ACP synthases (such as AcpS) that act on carrier proteins from FAS and type II PKS systems. Enzymes of this type also show little cross reactivity with other kinds of carrier proteins. The second group is made up of the PPTases that are integrated domains of the multienzyme complex FAS of eukaryotes. It only accepts the ACP domain of eukaryotic FAS as a substrate. The third type of PPTase is exemplified by the Sfp protein from *B. subtilis*, which is involved in surfactin biosynthesis. Sfp-type PPTases modify most types of carrier proteins and are generally associated with secondary metabolism. Much effort has gone toward elucidating the structural basis for the specificity of carrier protein

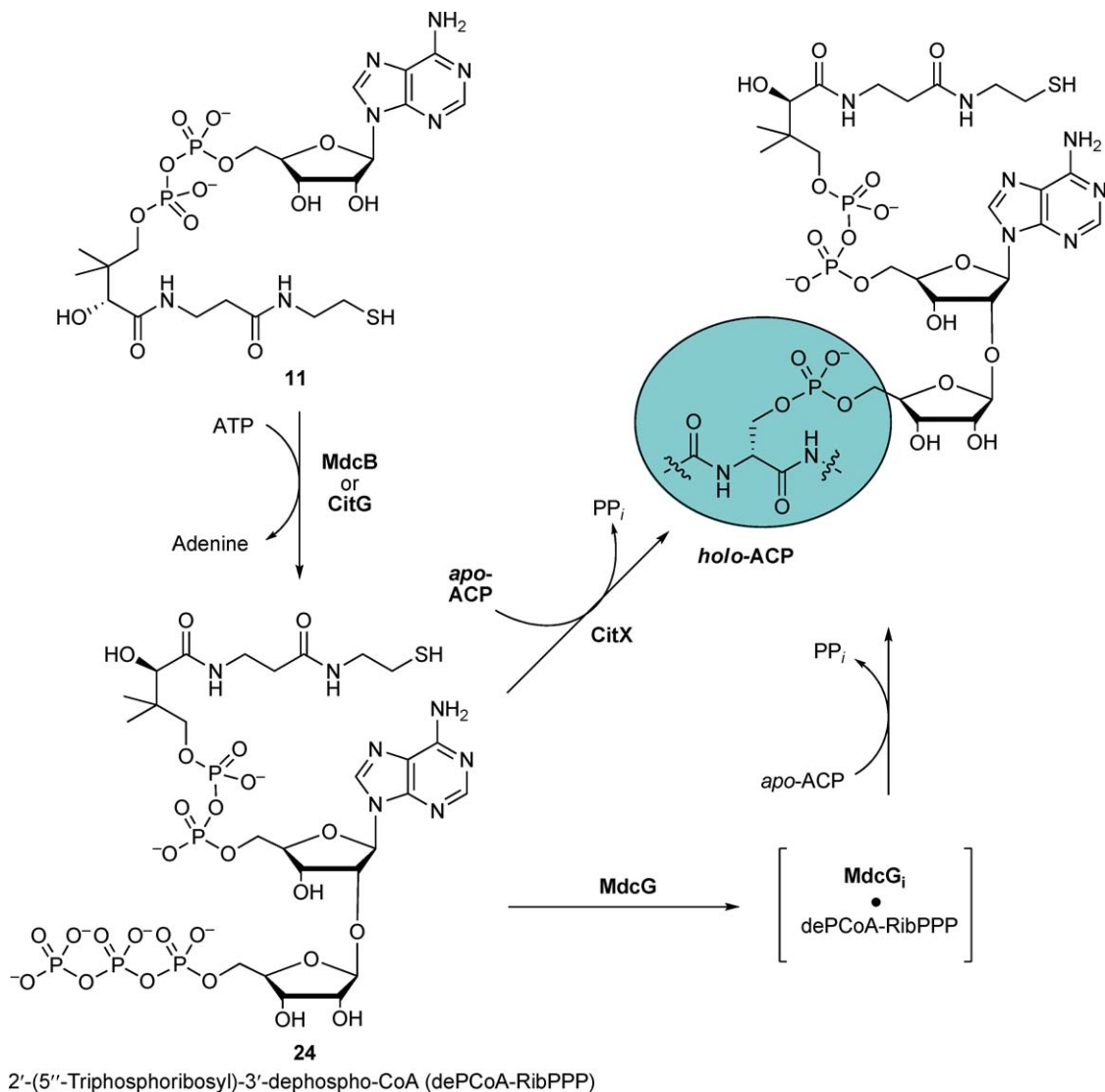
recognition by these different PPTases; such studies have been aided by the availability of crystal structures of the *B. subtilis* Sfp and AcpS enzymes,<sup>191,192</sup> and the AcpS from *S. pneumoniae*.<sup>193</sup> Site-directed mutation analysis based on these structures has shown that helix 2 of the carrier proteins are important for recognition and that the exchange of a single residue in a PCP is sufficient to allow cross-reactivity with an AcpS enzyme.<sup>194–196</sup> The study of carrier protein structure and recognition is an area of continued and growing interest because of the central role of these proteins in antibiotic biosynthesis.<sup>197,198</sup> For more information in this regard, the reader is referred to the chapters on PKS and NRPS systems in this series. The specificity of PPTases toward its other substrate, CoA, has also generated much attention. Early studies of the Sfp protein (as well as its structure) have shown that it would accept a wide range of CoA analogues as substrates because the enzyme does not have made any direct contacts with the pantetheine moiety of the cofactor.<sup>199</sup> Such observations have since been exploited in the development of various orthogonal protein-labeling strategies that will be highlighted in a later section of this chapter.

Recently, the structure of the human PPTase, which has been shown to modify both cytosolic and mitochondrial FAS, as well as the aminoadipate-semialdehyde dehydrogenase enzyme, has been solved in its native form, in complex with CoA and in complex with both CoA and the ACP domain of human FAS.<sup>200</sup> The preliminary structures of the AcpS enzymes from *M. tuberculosis*,<sup>201</sup> *S. aureus*,<sup>202</sup> and *Streptomyces coelicolor* (pdb: 2JCA; 2JBZ) have also been made available. Based on these and other PPTase structures and structure-based mutational studies, various mechanistic proposals for the enzyme's reaction have been put forward.<sup>191,192,196,200</sup> Most of these have CoA binding to the PPTase enzyme first in a Mg<sup>2+</sup>-dependent manner, although interestingly the Sfp enzyme demonstrates much higher activity in the presence of Mn<sup>2+</sup> (a 300-fold increase in the specificity constant). The carrier protein joins the complex last and positions the reactive serine in proximity to a glutamic acid residue (E151 in Sfp), which, together with an aspartate residue (D107 in Sfp), a second glutamate (E109 in Sfp), a water molecule, and CoA's pyrophosphate group, is initially involved in coordination of the Mg<sup>2+</sup> ion. The glutamate residue subsequently deprotonates the carrier protein's serine, activating it for attack on CoA's  $\beta$ -phosphate through an addition–elimination mechanism.

#### 7.11.4.1.2 Triphosphoribosyl-dephospho-CoA formation and utilization

Malonate decarboxylase (EC 4.1.1.88) and citrate lyase (EC 4.1.3.6) are both large enzyme complexes that consist of multiple subunits, the smallest of which acts as an ACP. These two complexes catalyze the decarboxylation of malonate to acetate and CO<sub>2</sub> and the Mg<sup>2+</sup>-dependent cleavage of citrate to acetate and oxaloacetate, respectively. Both have been shown to require a thiol-containing prosthetic group for activity. However, unlike the carrier proteins described in the previous section, the ACP subunits of these proteins are not phosphopantetheinylated by a reaction with CoA. Instead, they rely on a unique cofactor, 2'-(5"-triphosphoribosyl)-3'-dephospho-CoA (**24**, dePCoA-RibPPP), as source of a 2'-(5"-phosphoribosyl)-3'-dephospho-CoA prosthetic group, which is bound to a conserved serine residue of the ACP.<sup>203–207</sup> A similar prosthetic group has been identified in citramalate lyase (EC 4.1.3.22). The proposed biosynthesis and subsequent transfer reactions of the cofactor **24** to the ACPs of these complexes are shown in **Scheme 5**.

On the basis of studies of the malonate decarboxylase from *Klebsiella pneumoniae* and the citrate lyase enzyme from *E. coli*, it has been shown that 2'-(5"-triphosphoribosyl)-3'-dephospho-CoA **24** is formed by the reaction of dephospho-CoA **11** with ATP to form a unique 1 → 2 glycosidic bond between their ribose moieties. This reaction is catalyzed by either the MdcB or the CitG proteins, depending on the system. In the case of malonate decarboxylase, the formed dePCoA-RibPPP binds with a strong but noncovalent interaction to the MdcG protein, to form MdcG<sub>3</sub>. The protein complex reacts with MdcC, the ACP of the malonate decarboxylase complex, and transfers 2'-(5"-phosphoribosyl)-3'-dephospho-CoA to it presumably by reaction of the active site serine with the  $\alpha$ -phosphate of the triphosphoribosyl moiety, with concomitant release of pyrophosphate. In a similar fashion, CitX catalyzes the transfer of the same prosthetic group to the  $\gamma$ -subunit (CitD) of citrate lyase. It is currently still unclear why these enzymes rely on this



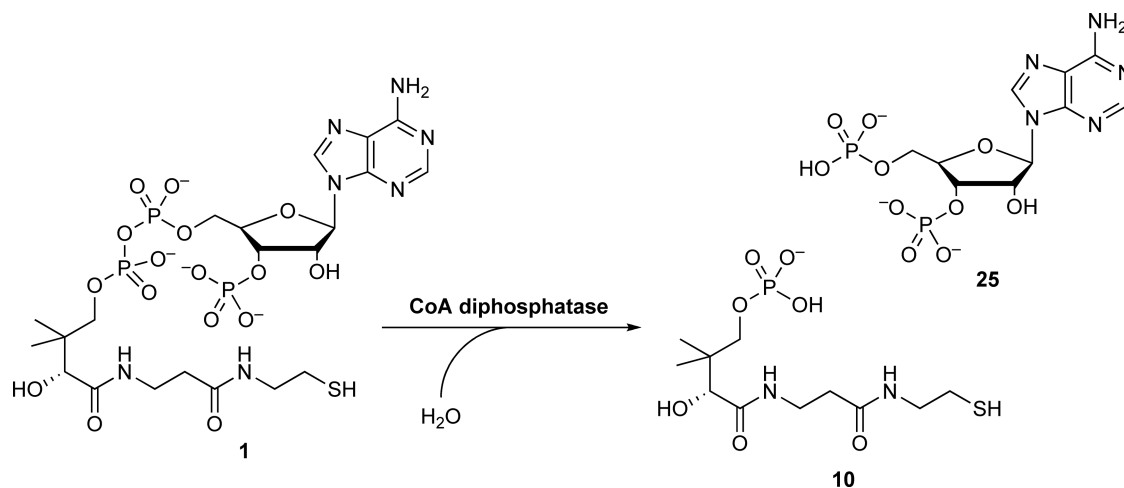
**Scheme 5** Malonate decarboxylase and citrate lyase both contain a *holo*-ACP (MdcC and CitD, respectively) that has a 2'-(5''-phosphoribosyl)-3'-dephospho-CoA prosthetic group which is attached by a phosphodiester linkage to a conserved serine residue. The prosthetic group originates from the cofactor 2'-(5''-triphosphoribosyl)-3'-dephospho-CoA **24**, which is biosynthesized from dephospho-CoA **11** and ATP by either MdcB or CitG. The posttranslational modification of the *apo*-ACP proteins is catalyzed by a complex of the enzyme MdcG with **24**, or CitX, depending on the system.

unique prosthetic group for catalysis and whether its biosynthesis from dephospho-CoA instead of CoA can be rationalized in terms of the differences in their metabolism and regulation.

#### 7.11.4.2 Enzymes Involved in the Degradation and Recycling of CoA

##### 7.11.4.2.1 CoA diphosphatases

Although little is known about CoA degradation, the enzymatic cleavage of the pyrophosphate bonds of CoA to form 4'-phosphopantetheine **10**, and 3',5'-ADP **25** is one way in which the concentration of the cofactor can be lowered (Equation (8)).



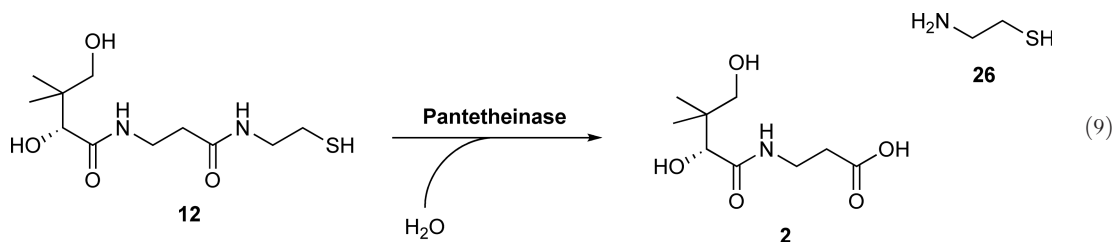
(8)

In mammals, a plasma membrane nucleotide pyrophosphatase has been shown to perform this reaction, although the enzyme's action is not considered to be specific for CoA because of its location in the cell and because it has a preference for acyl-CoAs ( $K_m$  of 8–9  $\mu\text{mol l}^{-1}$ ) and dephospho-CoA ( $K_m$  of  $\sim 20 \mu\text{mol l}^{-1}$ ) over free CoA ( $K_m$  of  $\sim 300 \mu\text{mol l}^{-1}$ ).<sup>208–211</sup> However, nudix hydrolase enzymes that act specifically as diphosphatases of CoA and its derivatives have been identified in the peroxisomes of various eukaryotic cells, including those of the yeast *S. cerevisiae*,<sup>212</sup> mouse liver and kidney,<sup>213,214</sup> and *Caenorhabditis elegans*.<sup>215</sup> Nudix hydrolases are a superfamily of  $\text{Mg}^{2+}$ -dependent enzymes that catalyze the hydrolysis of nucleoside diphosphates linked to other moieties (X). They are recognized based on a highly conserved consensus sequence of 23 residues also known as the Nudix box,  $\text{GX}_5\text{EX}_7\text{REUXEEXGU}$ , where U is a bulky hydrophobic residue and X is any residue. This Nudix box has a specific structural motif that functions both in  $\text{Mg}^{2+}$  binding and as the active site. The mechanisms of these enzymes are diverse, although most proceed by associative nucleophilic substitutions of water at one of the phosphorus atoms of the pyrophosphate moiety. The structures and mechanisms of Nudix hydrolases have recently been reviewed.<sup>216</sup>

Among these, the Pcd1p protein from yeast acts preferentially on oxidized CoA (CoA disulfide), followed by free CoA and various CoA thioesters,<sup>212</sup> whereas the mouse NUDT7 $\alpha$  and RP2p proteins exhibited highest activity toward free CoA and palmitoyl-CoA, respectively.<sup>213,214</sup> The *C. elegans* Y87G2A.14 protein also preferred free CoA as a substrate.<sup>215</sup> The  $K_m$  values for the best substrates of these enzymes ranged from 24  $\mu\text{mol l}^{-1}$  for CoA disulfide (*S. cerevisiae* Pcd1p) to  $\sim 600 \mu\text{mol l}^{-1}$  for free CoA (mouse RP2p). The structure of the CoA diphosphatase from the bacterium *Deinococcus radiodurans*, also a nudix hydrolase, has been determined and shows that the enzyme shares the basic fold of other Nudix hydrolases of which the structures have been determined.<sup>217</sup> When the substrate is modeled in the enzyme it shows interactions with the NuCoA motif that is characteristic of the enzymes that hydrolyze CoA. In this model CoA takes on an L-shaped conformation that conforms to the 90° turns of two of the strands of the NuCoA motif and promotes interactions between it and the phosphopantetheine moiety of the substrate.

#### 7.11.4.2.2 Pantetheinase

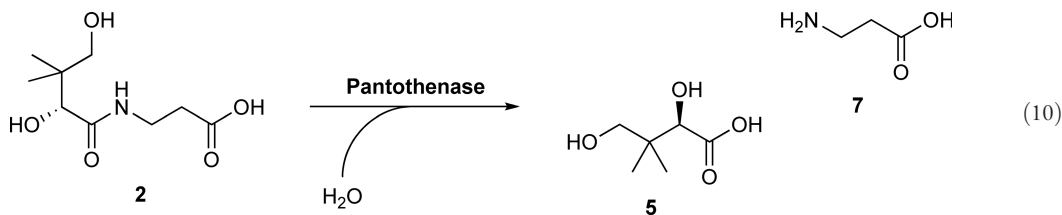
The degradation of pantetheine **12** through hydrolysis of one of its two amide bonds to form pantothenic acid **2** and cysteamine **26** is another important pathway for CoA metabolism. This reaction is catalyzed by a pantetheine hydrolase (or pantetheinase) enzyme (EC 3.5.1.92) (Equation (9)). Soluble enzymes with pantetheinase activity have been identified in and isolated from horse kidney,<sup>218</sup> pig kidney,<sup>219,220</sup> and pigeon liver.<sup>221,222</sup> In mice and humans, the activity was found to be associated with membrane-bound glycoproteins called Vanins.<sup>223–225</sup> Interestingly, expression of the *Vanin-1* gene is upregulated in response to oxidative stress, suggesting a role for the Vanin-1 protein in the response of tissues to this condition. This observation can be rationalized in terms of the antioxidant activity of the cysteamine molecules that would be produced in the process.



Characterization of the activities of the horse and pig enzymes has shown that pantetheinase is specific for the pantothenate moiety of its substrate and will not react with CoA, 4'-phosphopantetheine, or  $\beta$ -alanyl cysteamine. However, it accepts the modification of the cysteamine moiety and hydrolyzes a variety of pantetheine thioesters. The enzyme does not require any cofactors, although it is inhibited by pantetheine (pantetheine disulfide), oxidized glutathione, and thiol inhibitors (such as  $\text{Hg}^{2+}$ ), suggesting the involvement of an active site cysteine in catalysis. The pig enzyme has a molecular weight of 72 kDa and exhibits a  $K_m$  of  $20 \mu\text{mol l}^{-1}$  for pantetheine. All these studies have been performed on purified or partially purified native enzymes; no pantetheinase enzyme has been cloned to date.

#### 7.11.4.2.3 Pantothenease

A pantothenic acid hydrolase (pantothenease) activity has been isolated from *Pseudomonas fluorescens* and other *Pseudomonas* strains.<sup>226</sup> This enzyme hydrolyzes the amide bond of pantothenic acid **2** to form pantoic acid **5** (or pantoyl lactone) and  $\beta$ -alanine **7** (EC 3.5.1.22) (Equation (10)).<sup>227</sup> A detailed kinetic study of the reaction mechanism has shown that the reaction is partially reversible because of the formation of an acyl-enzyme (pantoyl-enzyme) intermediate during the course of catalysis, which may react with either water or  $\beta$ -alanine to form pantoic acid (the product hydrolysis) or pantothenic acid (the original substrate).<sup>228</sup> Such a mechanism suggests that this enzyme could act as a pantothenate synthase, as reaction of the active site serine with pantoyl lactone would result in the formation of the pantoyl-enzyme intermediate. However, no biochemical or genetic evidence is currently available to support such a hypothesis.



The enzyme was shown to be either a dimer of two 50 kDa subunits or a tetramer of 60 kDa subunits depending on its source. The determined  $K_m$  values for pantothenic acid are high, varying between 3.2 and  $\sim 15 \text{ mmol l}^{-1}$ . This may indicate that the isolated enzymes are not necessarily specific for pantothenic acid and that the observed activity could potentially be ascribed to a nonspecific serine protease.

#### 7.11.4.2.4 CoA disulfide reductase

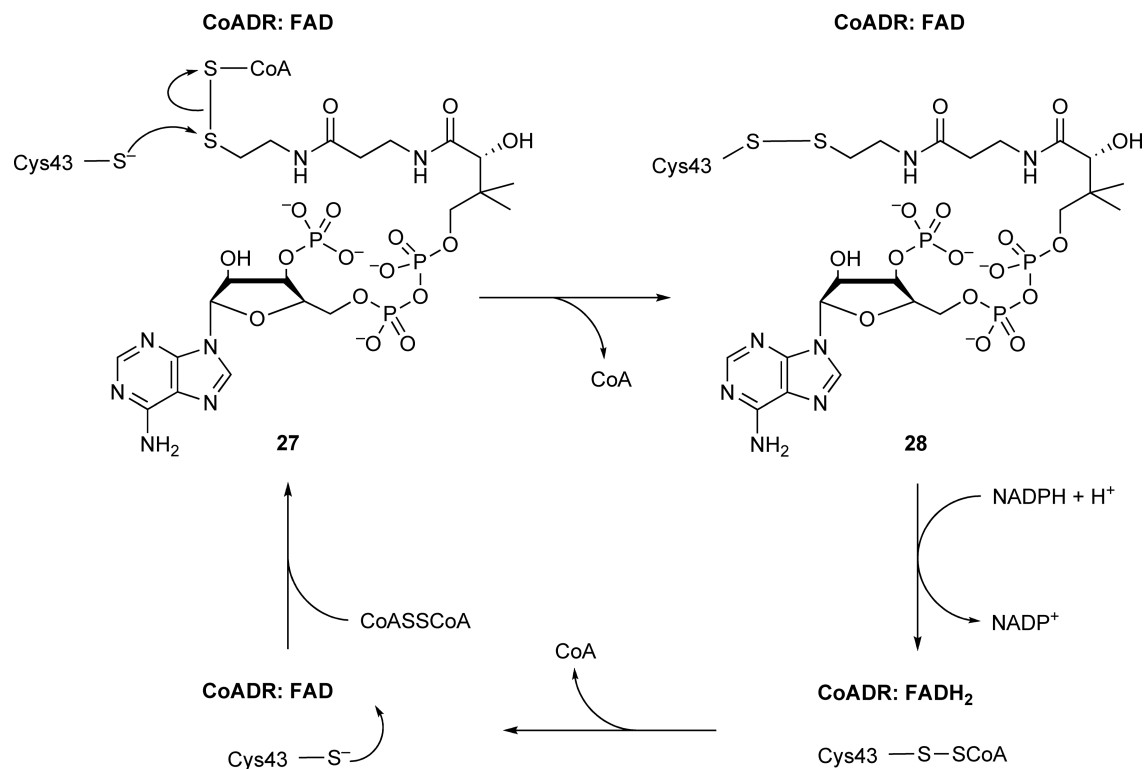
It is well known that the low-molecular-weight thiol glutathione, in combination with an NADPH-dependent glutathione reductase enzyme, plays a central role in maintaining the intracellular redox balance and thiol/disulfide ratio in most eukaryotes and Gram-negative bacteria.<sup>229</sup> However, glutathione is absent in many bacterial species, and in many of these – most notably *S. aureus*,<sup>104</sup> *Bacillus megaterium*,<sup>230</sup> *B. anthracis*,<sup>110</sup> *Borrelia burgdorferi*, and some hyperthermophilic archaea<sup>231</sup> – it has been shown that CoA is the major low-molecular-weight thiol, often occurring at millimolar concentrations in these cells.<sup>232</sup> These organisms are dependent on a CoA/CoA disulfide equilibrium to maintain their redox balance, which requires the action of a CoA disulfide reductase (CoADR; EC 1.8.1.14) enzyme to return the oxidized cofactor to its reduced form. Such CoADR enzymes have been characterized in most of the organisms listed above and were shown to be FAD-containing enzymes that preferentially reduce CoA disulfide **27** in the presence of NADPH.<sup>104,233–237</sup> However, unlike other well-known and extensively studied members of the family of flavin-dependent pyridine nucleotide-disulfide



oxidoreductases (PNDORs), CoADR enzymes do not have a redox-active disulfide motif.<sup>238</sup> Instead, these enzymes have a single active site cysteine found in a conserved active site sequence (SFXXC), which is normally associated with the NADH oxidase and NADH peroxidase enzymes from *E. faecalis* that contain cysteine–sulfenic acid redox centers (although these enzymes also belong to the PNDOR family).<sup>239</sup> When the *S. aureus* CoADR protein was purified in its native form or when expressed in *E. coli*, it was found that this cysteine (Cys43) was fully modified by forming a stable Cys43–SSCoA mixed disulfide (28). This finding was explained by experiments which showed that Cys43 of the reduced enzyme preferentially reacts with CoA disulfide and the CoA moieties of mixed CoA disulfides (such as CoASSGlutathione).<sup>104</sup> The enzyme is also essentially inactive (0.03% activity) when this cysteine residue is exchanged for serine.<sup>235</sup> Taken together, these findings confirm the catalytic role of the active site cysteine and suggest that the Cys43–SSCoA mixed disulfide acts as the redox center during catalysis. A proposed mechanism for CoA disulfide reduction is shown in Scheme 6; a more detailed account of the various steps that occur during the reaction has been described elsewhere.<sup>235</sup>

Kinetic and structural characterization of the known CoADR enzymes shows the occurrence of some variation with regard to their specificity. The *S. aureus* CoADR, a homodimer of 49 kDa subunits, has a  $K_m$  of  $11 \pm 2 \mu\text{mol l}^{-1}$  and  $1.6 \pm 0.5 \mu\text{mol l}^{-1}$  for CoA disulfide and NADPH, respectively. Although an initial study of this enzyme reported  $k_{\text{cat}}$  values in the range of  $1000 \text{ s}^{-1}$  (which would have had the enzyme operating near the limiting rate constant for a diffusion-controlled enzyme–substrate encounter),<sup>104</sup> subsequent kinetic analyses have shown the turnover number to be closer to  $27 \text{ s}^{-1}$ .<sup>235</sup> The enzyme shows 17% activity in the presence of NADH relative to NADPH. Titration of the oxidized enzyme with NADPH indicated that the enzyme exhibits differential behavior between the two active sites (one site per identical subunit) of the protein, and that only one FAD is reduced in the process.

The crystal structure of the oxidized *S. aureus* enzyme with the Cys43–SSCoA mixed disulfide bound in its active site shows the overall structure to be very similar to that of NADH peroxidase and predicts NADPH binding to be similar to that seen in other PNDOR enzymes.<sup>240</sup> CoA is bound in an extended fashion, and



**Scheme 6** General mechanism for the reduction of CoA disulfide by CoADR enzymes as based on current knowledge of the FAD-containing and NADPH-dependent CoADR from *S. aureus*.

although it makes 12 direct polar interactions with the enzyme its ribose and 3'-phosphate moieties are exposed to the solvent. This observation does not agree with the much lower activity (12–13%) of the enzyme toward dephospho-CoA as substrate. The pantetheine and pyrophosphate moieties form extensive contacts with the enzyme and is deemed very important for recognition. A model of the CoA disulfide substrate shows that the second CoA moiety can be accommodated without any steric collisions with the protein or first CoA. Mechanistically, the structure is in agreement with the proposed mechanism shown in **Scheme 6**.

A recent detailed characterization of the putative CoADR enzyme from *Pyrococcus furiosus* have shown that this protein is in fact a CoA-dependent NAD(P)H elemental sulfur oxidoreductase (NSR), which is mainly involved in the reduction of  $S^0$  to  $H_2S$  in this heterotrophic hyperthermophile.<sup>241</sup> Previous studies of the *P. furiosus* and *Pyrococcus horikoshii* enzymes have found that they catalyze the NADPH-dependent reduction of CoA disulfide with kinetic parameters similar to those of other CoADR enzymes ( $K_m$  of  $9\ \mu\text{mol l}^{-1}$  and  $30\ \mu\text{mol l}^{-1}$  for NADPH and CoA disulfide, respectively).<sup>236</sup> This CoADR activity is now believed to be a side reaction of the NSR protein, and it has been suggested that it does not have physiological significance in this organism and, by inference, other similar heterotrophic archaea. The mechanistic involvement of CoA in the  $S^0$  reduction reaction has not been elucidated as yet, but it has been proposed that CoA, together with the single active site cysteine residue, allows for two-electron reduction of  $S^0$  with the resulting CoA–Cys mixed disulfide being reduced by NADPH in a similar fashion to the first proposed step of the CoADR mechanism.

For a more detailed analysis of flavin-dependent disulfide reductases the reader is referred to the chapter on flavoenzymes in this series.

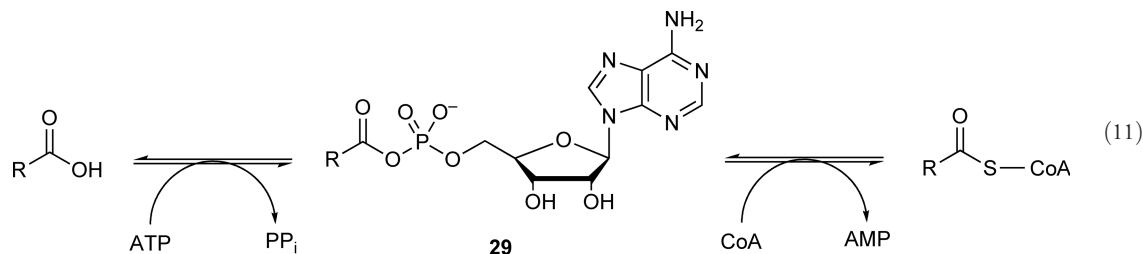
## 7.11.5 CoA-Dependent Enzymes

The chemistry and enzymology of CoA-dependent enzymes covers an extraordinarily wide range of reactions. Many of these do not even exploit the unique chemistry of the cofactor's thiol or its thioesters, but simply use CoA as a handle for binding the large array of carboxylic acids that are involved in primary and secondary metabolism. The sheer number of enzymes that fall into this category necessitates that any review of the subject has to be limited in its scope; this section will therefore provide general discussions on the mechanisms and structural characteristics of the broad classes of enzymes that are involved in the biosynthesis of CoA thioesters, in the transfer of the acyl moiety of such thioesters to other groups – either by group exchange or by condensation reactions – and on enzymes that cleave carbon–carbon bonds by making use of CoA thioester chemistry. Finally, a brief overview will be given of the types of reactions that act on CoA-bound substrates, but in which the cofactor itself has no mechanistic involvement. For a more detailed discussion of these enzymes the reader will be directed to the other relevant chapters in this series. It should be pointed out that much of what will be said with regard to CoA-dependent enzymes in this section also holds for ACP-dependent enzymes, and the reader is reminded to bear this in mind in the discussions that follow. However, where such enzymes are specifically relevant to the discussion they will be mentioned by name.

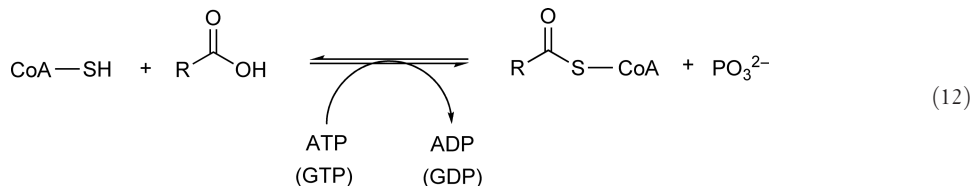
### 7.11.5.1 Enzymes Involved in the Biosynthesis of CoA Thioesters (Acyl-CoAs)

#### 7.11.5.1.1 Ligases: Coupling activated carboxylic acids to CoA directly

The acyl-CoA ligases (EC 6.2.1.-; often also referred to as acyl-CoA synthetases, or ACSs) catalyze the reversible nucleoside triphosphate-dependent formation of acyl-CoA thioesters from CoA and a free carboxylic acid. Two mechanistic types can be distinguished in this group of enzymes:<sup>242–246</sup> the first uses ATP to activate the carboxylic acid by formation of an intermediate acyl–adenylate **29**, which subsequently reacts with CoA to give the thioester and AMP (Equation (11)). These so-called AMP-forming ACS enzymes represent most of the enzymes involved in the direct coupling of a wide array of structurally diverse carboxylic acids to CoA, whether they are short-, medium-, or long-chain fatty acids,<sup>247,248</sup> xenobiotics,<sup>249</sup> amino acids (taking part in nonribosomal peptide synthesis<sup>250</sup>), or other carboxylic acid-containing metabolites.<sup>251–253</sup> This diversity is also found on a genomic level; for example, it has recently been shown that 26 distinct ACS-encoding genes can be identified in the human genome.<sup>254</sup> However, despite this variety AMP-forming ACS enzymes seem to all belong to the firefly luciferase superfamily of adenylate-forming enzymes.<sup>255–257</sup>



The second type of ACS enzyme also mainly uses ATP (although some enzymes use GTP) to activate carboxylic acids for thioester formation; however, in this case ADP and phosphate are formed as by-products (Equation (12)). Extensive mechanistic and structural studies of the ATP- and GDP-dependent succinyl-CoA synthetase enzymes (ATP-dependent: EC 6.2.1.5; GTP-dependent: EC 6.2.1.4) from mammalian and bacterial sources have shown that during the course of the reaction a phospho-enzyme intermediate is formed by transfer of the  $\gamma$ -phosphate of the NTP to an active site histidine residue.<sup>258,259</sup> Subsequently, the phosphate is transferred to the substrate's carboxylic acid to give an activated acyl-phosphate, which is finally attacked by CoA's thiol to give the acyl-CoA and phosphate. As mentioned above these reactions are normally reversible, and in this case the reverse reaction – the formation of succinate and NTP from succinyl-CoA and the corresponding NDP – is a substrate level phosphorylation (such as the one observed in the citric acid cycle) in which the high-energy thioester bond is converted into the high-energy phosphoanhydride bond of the nucleoside triphosphate. Interestingly, the equilibrium constant for the overall reaction approximates one; consequently the direction of the catalyzed reaction depends on the energy state of the cell as reflected in the prevalent ATP/ADP ratio. Structurally, ADP-forming ACS enzymes are characterized by the presence of the 'ATP-grasp' nucleotide-binding motif, which is also found in other carboxylate-amine ligases.<sup>260</sup>

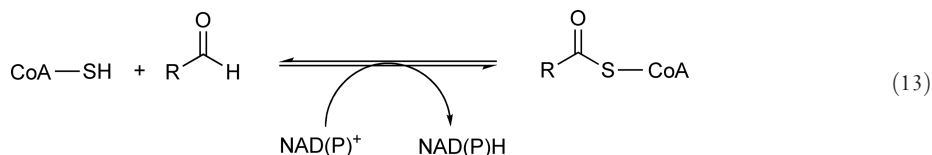


Acetyl-CoA, which is at the crossroads of a wide variety of reactions that interchange central metabolites – including fatty acids, amino acids, ethanol, and acetaldehyde, and intermediates from glycolysis and the citric acid cycle – can be formed directly from acetic acid by three distinct pathways.<sup>246</sup> The activity of these pathways (which all operate in prokaryotes) differs depending on the availability of acetate, allowing the organisms to use it as carbon and energy source regardless of its concentration in the environment. The first of these pathways is found in the reaction catalyzed by the AMP-forming acetyl-CoA synthetase (EC 6.2.1.1), which is mainly active when the available acetate concentration is low ( $<10 \text{ mmol l}^{-1}$ ). This is the only pathway for acetate activation that is also found in eukaryotes. The second pathway proceeds by means of an ADP-forming acetyl-CoA synthetase (EC 6.2.1.13), which has been described in archaeobacteria and the eukaryotic parasites *Entamoeba histolytica* and *Giardia lamblia*.<sup>245,261</sup> However, this enzyme is normally active in the reverse direction, driving the formation of ATP from ADP at the expense of acetyl-CoA hydrolysis – in the parasites mentioned above this is important as they do not have mitochondria and therefore have a fermentative energy metabolism. Finally, the acetate kinase (EC 2.7.2.1)/phosphotransacetylase (EC 2.3.1.8) pathway forms acetyl-CoA (through the formation of acetyl phosphate) when the environmental concentration of acetate is high ( $>30 \text{ mmol l}^{-1}$ ). This pathway is also reversible and is used to conserve energy and maintain the steady-state concentration of CoA in the cell.<sup>262,263</sup>

Enzymes with acyl-ACP synthetase activity (i.e., coupling carboxylic acids to the prosthetic group of ACPs) have been characterized in *E. coli* and *Vibrio harveyi*.<sup>264</sup>

### 7.11.5.1.2 Oxidoreductases: CoA thioester formation through aldehyde oxidation

Aldehyde dehydrogenase (ALDH) enzymes can be classified as two structurally unrelated families of proteins, both of which catalyze the NAD(P)H-dependent oxidation of aldehydes to carboxylic acids (or their derivatives). The mechanism of reduction is also similar in these families: an active site cysteine residue in the enzyme acts as a nucleophile and attacks the aldehyde group, followed by hydride transfer from the resulting thiohemiacetal to the nicotinamide cofactor to give an acyl–enzyme thioester intermediate. However, the families differ with regard to the ultimate fate of this thioester: in phosphorylating ALDHs the acyl group is transferred to inorganic phosphate to give an acyl-phosphate as a product; this, for example, is the case in glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Nonphosphorylating ALDH enzymes on the contrary transfer the acyl group either to water (hydrolysis) or to the thiol of CoA (transthioesterification) to give the free carboxylic acid or the CoA thioester (Equation (13)) as products.

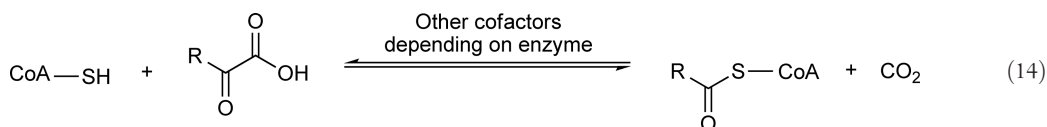


The best-studied example of a CoA-dependent nonphosphorylating ALDH is the methylmalonate-semialdehyde dehydrogenase, which has been isolated from both mammalian and bacterial sources.<sup>265,266</sup> This enzyme transforms malonate semialdehyde and methylmalonate semialdehyde into acetyl-CoA and propionyl-CoA, respectively, through an oxidation reaction as described above, followed by a decarboxylation reminiscent of other  $\beta$ -keto acids. Mechanistic studies of the *B. subtilis* enzyme have shown that it is activated by NAD<sup>+</sup> binding, that it exhibits half-of-sites reactivity (only two moles of NADH forms per tetrameric protein unit) and that the decarboxylation reaction occurs after formation of the acyl–enzyme intermediate. Acyl transfer from the enzyme to CoA completes the reaction.

A reversible CoA-dependent acetaldehyde dehydrogenase has been identified in the parasites *G. lamblia* and *E. histolytica*, where it plays an important role in the organisms' utilization of glucose as energy source through fermentative pathways, as mentioned in the previous section.<sup>267,268</sup>

### 7.11.5.1.3 Oxidoreductases: CoA thioester formation through oxidative decarboxylation of $\alpha$ -keto acids

CoA thioesters are also the products of the oxidative decarboxylation reactions of  $\alpha$ -keto acids, especially pyruvate and  $\alpha$ -ketoglutarate, from which acetyl-CoA and succinyl-CoA are formed, respectively (Equation (14)). Three distinct types of enzymes catalyze such reactions; however, the mechanistic involvement of CoA is generally rather limited for two of these, and only a brief discussion of each will be provided here. For more detailed information on these enzymes, the reader is referred to the relevant chapters on thiamin and lipoic acid enzymology and on radical enzymes in this series (see Chapters 1.08 and 7.03).

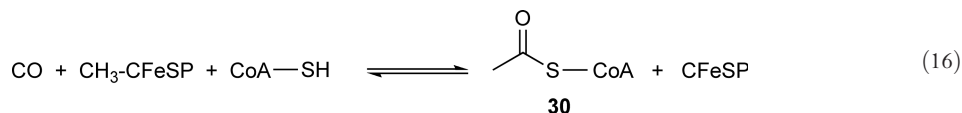
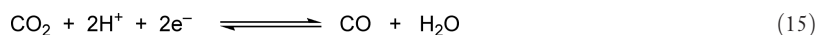


Under aerobic conditions the oxidative decarboxylation of  $\alpha$ -keto acids is catalyzed by large multienzyme dehydrogenase complexes that require NAD<sup>+</sup>, thiamin pyrophosphate (TPP), lipoic acid, and CoA for activity. The reaction proceeds by the TPP-mediated decarboxylation of the acid to give an anionic hydroxyalkyl–TPP intermediate, which subsequently reacts with enzyme-bound lipoic acid to give the acylated lipoamide. CoA acts as the final acceptor of the acyl group as it is transferred from the acylated lipoyl-bearing domain of the dehydrogenase complex. In contrast, under anaerobic conditions the same type of conversion is accomplished by ferredoxin-dependent 2-keto acid oxidoreductase enzymes, such as pyruvate:ferredoxin oxidoreductase (PFOR, EC 1.2.7.1). Although these enzymes also require TPP and CoA for activity, they have three [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+/1+</sup> clusters that act as redox centers and lead to the formation of a radical TPP intermediate (hydroxyethylidene-TPP radical in the case of PFOR). In the absence of CoA this radical intermediate slowly decays to form a

two-electron reduced enzyme, but when it is present the rate of radical decay increases by 100 000-fold. Although it is known that CoA's sulfur is the major factor in controlling this rate of decay, it is less clear how the rate of electron transfer is controlled, and by which mechanism the acyl-CoA product is finally formed.<sup>269,270</sup> Our current knowledge of PFOR enzymology was recently described in an excellent review.<sup>271</sup> Finally, acetyl-CoA is also formed from the oxidative decarboxylation of pyruvate as catalyzed by the radical *S*-adenosyl methionine (SAM) enzyme pyruvate formate-lyase (PFL, EC 2.3.1.54). This enzyme does not utilize a TPP cofactor, but also has a [Fe<sub>4</sub>S<sub>4</sub>] cluster and uses SAM to form a catalytically active protein-based glycyl radical. The reaction proceeds through a radical-mediated decarboxylation mechanism and finally forms an acetyl-enzyme thioester intermediate from which CoA accepts the acetyl group in a transthioesterification reaction.<sup>272,273</sup>

#### 7.11.5.1.4 Carbon monoxide dehydrogenase/acetyl-CoA synthase

The bifunctional carbon monoxide dehydrogenase (CODH)/acetyl-CoA synthase enzyme is a key enzyme involved in the Wood–Ljungdahl pathway of carbon fixation that operates in anaerobic bacteria. As such, it is a major player in the global carbon cycle. The CODH component of the enzyme catalyzes the reversible reduction of CO<sub>2</sub> to CO (Equation (15)), which is then channeled to the ACS active site where it reacts with CoA and a methyl group provided by the corrinoid iron–sulfur protein (CFeSP) to form acetyl-CoA **30** (Equation (16)).



X-ray crystal structures of CODH/ACS proteins show that they are large (300 kDa) homodimeric proteins with the two CODH domains ( $\beta$  subunits) at the core, and the two ACS domains ( $\alpha$  subunits) tethered to the side of each of these  $\beta$  subunits. The  $\beta$  subunits each contain one Ni and 10 Fe ions that are arranged into three FeS clusters (the so-called B-, C-, and D-clusters), whereas the  $\alpha$  subunits each contains the active site A-cluster, a [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+/1+</sup> cubane that is bridged by the sulfur of a cysteinyl residue to the proximal metal (M<sub>p</sub>) of a binuclear center. This binuclear center contains a square planar nickel ion, referred to as the distal Ni (Ni<sub>d</sub>), which is coordinated by the two thiolates and two backbone amides of a Cys-Gly-Cys motif. There has been some debate as to the identity of M<sub>p</sub>, as Ni, Cu, and Zn ions have been shown to occupy this site; however, it is now generally accepted that the A-cluster is a binuclear Ni–Ni center bridged by a cysteine thiol to a [4Fe–4S] cluster.<sup>274</sup>

Extensive structural, spectroscopic and kinetic studies of the ACS enzyme have led to the proposal of two mechanisms for acetyl-CoA formation: the first, the ‘diamagnetic mechanism,’ involves a Ni(0) state and all diamagnetic intermediates, although the second is referred to as the ‘paramagnetic mechanism’ because the proposed Ni–Fe–C intermediate is EPR active and paramagnetic. There is also some debate as to the order in which the substrates bind the active Ni center. Regardless of the mechanism, the last proposed metal-bound intermediate is an acetyl-metal species that forms by condensation of the metal-bound methyl- and carbonyl group. Mechanistically, CoA's involvement seems to be limited to the thiolysis of the acetyl-metal bond to form the final acetyl-CoA product, although some evidence exists that the sulfur of CoA coordinates with M<sub>p</sub> of the A-cluster. For more detailed information on the structure, function, and mechanism of this most interesting, but challenging enzyme, the reader is referred to recent reviews on the subject, and the references cited therein.<sup>275–277</sup>

### 7.11.5.2 Acylating Enzymes: CoA Thioesters Acting as Acyl Donors toward Noncarbon Acceptors

#### 7.11.5.2.1 Nitrogen- and oxygen-based acceptors (N- and O-acyltransferases)

N- and O-acyltransferases (EC 2.3.1.–) are involved in a tremendously diverse array of reactions both with regard to substrate diversity and physiological relevance. The functional implications of these enzymes include roles in antibiotic resistance, in transcription regulation, in the regulation of the membrane association of



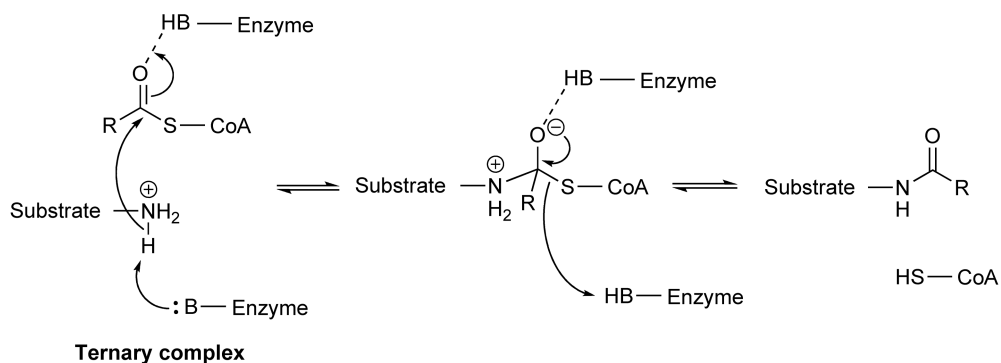
proteins, in the regulation of circadian rhythms, and in the biosynthesis and regulation of various central metabolites, to name but a few examples. Despite this diversity structural considerations allow all these enzymes to broadly be divided into four main groups. The first contains those enzymes that belong to the Gcn5-related *N*-acetyltransferase (GNAT) superfamily, so-called because of the identification of the Gcn5 transcription factor as a histone acetyltransferase (HAT), and the subsequent demonstration that histone acetylation and transcriptional regulation are directly linked.<sup>256,257</sup> This GNAT-like group is by far the largest – more than 10 000 members of this superfamily have been identified in all kingdoms – and includes enzymes such as aminoglycoside *N*-acetyltransferase, serotonin *N*-acetyltransferase (also known as arylalkylamine *N*-acetyltransferase), glucosamine-6-phosphate *N*-acetyltransferase, spermidine acetyltransferase, and several enzymes that acylate antibiotics and homoserine lactones.<sup>278</sup> Most acyltransferases (ATs) with protein substrates also belong to this group, such as protein  $\alpha$ -*N*-myristoyltransferase that transfers a myristoyl group to the N-terminal glycine of several eukaryotic proteins, as well as members of most families of the HAT enzymes mentioned above.<sup>279</sup> GNAT-like enzymes are identified by the presence of up to four sequence conserved regions (the A, B, C, and D motifs) that span more than 100 residues, although not all members of the superfamily necessarily have all four of these motifs.<sup>280</sup> In contrast, members of the second group of ATs all hold a distinct left-handed parallel  $\beta$ -helix (L $\beta$ H) structural motif in common and belong to the trimeric LpxA-like superfamily.<sup>281</sup> This group includes UDP-*N*-acetylglucosamine 3-*O*-acyltransferase, tetrahydrodipicolinate *N*-succinyltransferase, galactoside acetyltransferase, maltose *O*-acetyltransferase, xenobiotic acetyltransferase, and serine acetyltransferase. Enzymes of the third group belong to the CoA-dependent AT superfamily that includes carnitine acetyltransferase, chloramphenicol acetyltransferase, dihydrolipoamide acetyl- and succinyltransferase and choline *O*-acetyltransferase.<sup>282,283</sup> The most well-studied examples of the fourth group are homoserine acetyl- and succinyltransferase, which belong to the family of  $\alpha$ , $\beta$ -hydrolase enzymes but have their active sites at the end of a deep tunnel that have been suggested to promote catalysis of the acyl transfer reaction, rather than hydrolysis.<sup>284</sup>

Apart from this structural diversity, two distinct chemical mechanisms for the acyl transfer reaction have also been described (**Scheme 7**). In the direct-attack mechanism the acyl-CoA and acyl acceptor substrate binds to the enzyme by an ordered mechanism to form a ternary complex in which the *N*- or *O*-based nucleophile displaces the acyl group and releases CoA. This process may involve active site residues acting as general acids, general bases, both or neither – examples of all these cases have been described. Nearly all GNATs utilize this mechanism.<sup>278</sup> The other mechanistic possibility involves the formation of an acyl-enzyme intermediate by a ping-pong mechanism in which the acyl group is first transferred to either a serine or cysteine active site residue before it is transferred to the final acceptor. In such cases, a catalytic Asp-His-Ser/Cys triad similar to those found in the cysteine and serine proteases often forms the core of the catalytic mechanism. Such a mechanism has mainly been demonstrated for homoserine acetyl- and succinyltransferase and the ornithine acetyltransferase involved in clavulanic acid biosynthesis.<sup>285</sup>

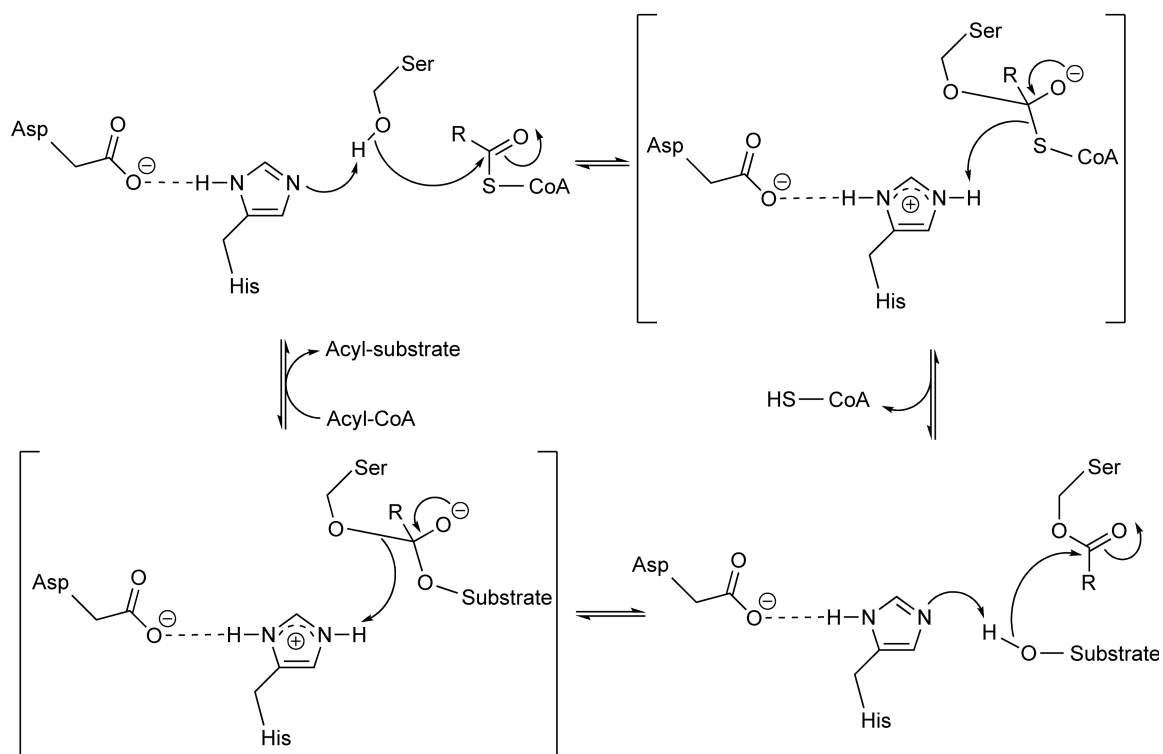
The HAT enzymes deserve special mention because of the central role they play in transcriptional activation and regulation. Two classes of HATs can be distinguished: A-type or nuclear HATs are localized in the nucleus and are mainly associated with modification of histones that are incorporated in chromatin, whereas B-type or cytosolic HATs are able to acetylate free, but not nucleosomal, histones and are therefore most likely involved in the acetylation of newly synthesized histones during their transport from the cytosol to the nucleus.<sup>286</sup> Among the A-type HATs a number of families have been described that differ with regard to origin, function, and substrate, whereas the sole-characterized representative of B-type HATs is the Hat1 protein. The best-studied HATs are mainly members of the Gcn5/PCAF family, the MYST family and the CBP/p300 family, and crystal structures of representative enzymes of these families, as well as of Hat1, have been determined.<sup>287,288</sup> These structures show that despite other differences HAT proteins generally share a conserved central core region associated with binding of acetyl-CoA and catalysis. Consequently, most HATs have been shown to make use of a direct-attack mechanism (even the MYST family member Esa1, which was previously thought to make use of an acyl-enzyme intermediate<sup>289</sup>) in which a conserved active site glutamate acts as a general base. However, recent structural and kinetic studies of p300/CBP have suggested that this protein's mechanism deviates from this general motif, and that catalysis in its case proceeds by a Theorell-Chance mechanism in which no stable ternary complex is formed as would be the case with a standard sequential mechanism.<sup>290</sup>



## Mechanism A:



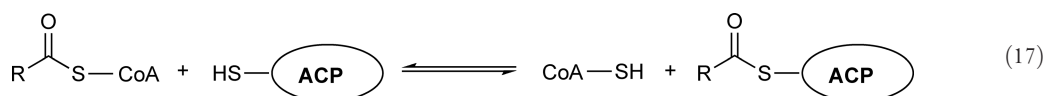
## Mechanism B:



**Scheme 7** Two proposals for the chemical mechanism of a CoA-dependent acyltransferase. Mechanism A represents a direct-attack mechanism in which a ternary complex of acyl-CoA, substrate, and enzyme is formed, followed by a nucleophilic attack of the substrate on the acyl-CoA leading to displacement of the acyl group from CoA. The acyl transfer can be aided by general acid and/or base catalysis, and may involve water molecules as intermediates in proton transfer, depending on the specific system. Mechanism B, also known as the ping-pong mechanism, proceeds through an acyl-enzyme intermediate in which an acyl-serine (or acyl-cysteine, depending on the case) conjugate is formed by transfer of the acyl group to an active site serine or cysteine residue in a manner reminiscent of the serine and cysteine proteases. The acyl group is subsequently transferred to the acceptor group of the substrate. This substrate acceptor group can be either an amine or a hydroxyl group in both mechanisms; in the proposals drawn above, only one is shown as an illustrative example in each case.

### 7.11.5.2.2 Sulfur-based acceptors (transacylases)

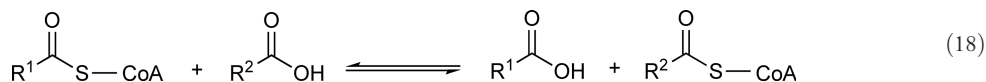
Transacylase enzymes (or *S*-acyltransferases) catalyze the exchange of acyl groups between the thioesters of CoA and the phosphopantetheine prosthetic group of ACPs (Equation (17)). Malonyl-CoA:ACP transacylase (MCAT, EC 2.3.1.39) and acetyl-CoA:ACP transacylase (ACAT, EC 2.3.1.38) are the best examples of such enzymes, and also the most extensively studied. This is especially true in the case of MCAT, which plays a central role in fatty acid and polyketide biosynthesis by providing the acyl units from which the growing carbon chain of these metabolites is built-up. The involvement of AT enzymes in loading the ACPs has also led to them being regarded as the gatekeepers of these biosynthetic pathways, as they select the subunits that will be added at each step. The basis of the substrate selectivity of ATs has therefore long been a subject of intense interest, as manipulation and modification of this selectivity could allow the production of related yet structurally distinct secondary metabolites.<sup>291</sup> Only a brief overview of the structure- and mechanism-related characteristics of AT enzymes is provided here; for a more thorough discussion of their metabolic roles and studies of their substrate selectivity the relevant chapters on fatty acid and polyketide biosynthesis in this series should be consulted.



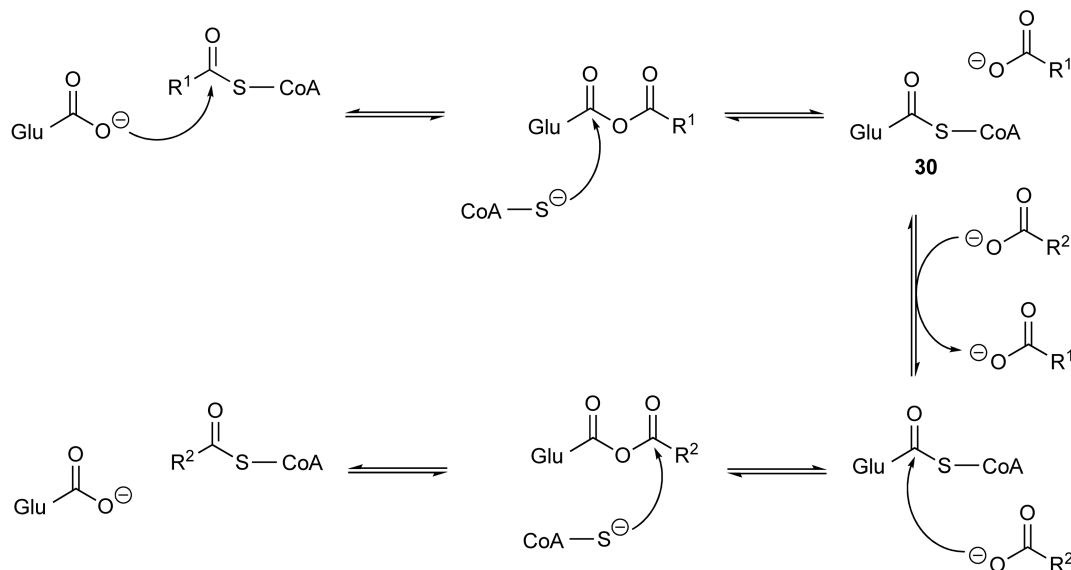
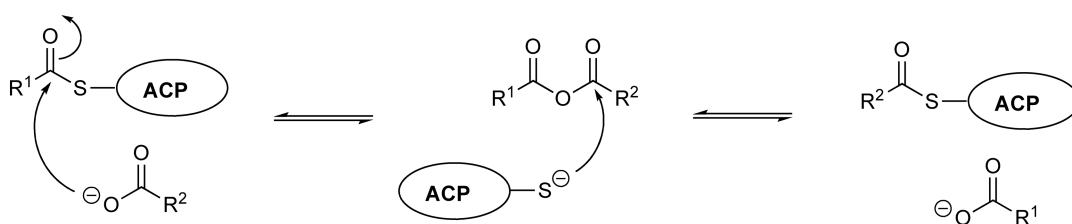
Structural analysis of the MCATs from *E. coli*, *S. coelicolor*, and *M. tuberculosis* shows that these enzymes are only distantly related to other members of the  $\alpha/\beta$  hydrolase superfamily as they have unique structures consisting of two subdomains, a large subdomain which hosts the active site and has a core that bears some structural similarity to the  $\alpha/\beta$  hydrolases, and a small subdomain which has a ferredoxin-like fold.<sup>292–294</sup> They also differ mechanistically from most other hydrolases in that acyl transfer is catalyzed through a ping-pong bi–bi mechanism using a His–Ser catalytic dyad, rather than the usual Asp–His–Ser/Cys triad depicted in **Scheme 7** (mechanism B). Instead, the His residue is stabilized by interactions with backbone carbonyl groups during proton transfer; no evidence of stabilization by interaction with Asp or Glu residues has been found. However, the mechanism still involves the formation of an acyl–serine intermediate, and the rest of the mechanism is thought to correlate with the one shown in **Scheme 7**. The structure also provides some indication of the residues that may be involved in substrate recognition; this information has been exploited in recent structure-based mutagenesis studies aimed at changing the enzyme's substrate selectivity.<sup>295</sup>

### 7.11.5.2.3 Acyl exchange reactions (CoA-transferases)

CoA-transferase enzymes (EC 2.8.3.–) catalyze the reversible transfer of CoA between an existing acyl-CoA thioester and a free carboxylic acid. No cofactors are involved in the reaction and no prior activation of the carboxylic acid is required (Equation (18)).



Three classes of CoA-transferase – as distinguished by differences in their primary sequences, their substrate specificity, and the mechanism by which acyl transfer occurs – have been described.<sup>296,297</sup> The Class I CoA-transferases are mainly involved in fatty acid metabolism, and include enzymes that transfer 3-ketoacids, short-chain fatty acids, and (*E*)-glutaconate in which acetyl-CoA or succinyl-CoA are usually the CoA donors. They make use of a well-established ping-pong mechanism in which the acyl group is transferred from the donor CoA thioester to an active site glutamate residue, forming an acyl–glutamyl anhydride as intermediate. Free CoA subsequently reacts with this anhydride to give an enzyme-bound glutamyl-CoA thioester **30** – the formation of this intermediate was confirmed in a recent crystallographic study in which incubation of the enzyme with various CoA thioesters resulted in it being trapped and structurally resolved.<sup>298</sup> The acceptor carboxylic acid is subsequently transferred to CoA by reversal of this reaction sequence (**Scheme 8**, mechanism A). The Class I enzyme succinyl-CoA:3-oxoacid CoA-transferase (SCOT) has been studied extensively as an example of how binding energy from noncovalent interactions (in this case with CoA) can lead to increases in  $k_{\text{cat}}/K_{\text{m}}$  of up to  $\sim 10^{10}$ -fold.<sup>299,300</sup> These studies also show how

**Mechanism A:****Mechanism B:****Ternary complex**

**Scheme 8** Two mechanistic proposals for the catalytic mechanism of CoA-transferases. In mechanism A, an acyl-enzyme intermediate is formed by reaction of an enzyme-bound glutamate (aspartate for Class III enzymes) with the donor acyl-CoA, followed by the formation of an enzyme-bound glutamyl- (or aspartyl-) CoA thioester intermediate. The thioester subsequently reacts with the acceptor carboxylate to give a new acyl-enzyme anhydride from which the acyl group is transferred to CoA. In Class I transferases, this process follows classical ping-pong kinetics, whereas in Class III enzymes the donor carboxylate only leaves the enzyme complex upon formation of the product (see text for details). Mechanism B represents a ternary complex mechanism as used by Class II enzymes in which a transient anhydride made up of the donor and acceptor acyl groups is formed by reaction of the acceptor carboxylate with the donor acyl-ACP. The free ACP subsequently reacts with this anhydride to complete acyl transfer.

the interactions between the enzyme and the various CoA constituent parts change during the course of catalysis, thereby influencing the formation of the various transition states and enzyme-substrate complexes. Structurally, Class I enzymes belong to the NagB/RpiA/CoA transferase-like superfamily of open  $\alpha/\beta$ -proteins.<sup>256,257</sup>

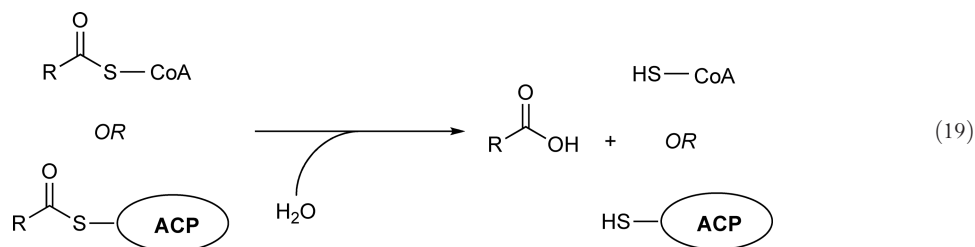
Only two examples of Class II CoA-transferases have been described; both of these are found as the  $\alpha$ -subunits of the complex citrate and citramalate lyase enzymes that also contain ACP-like  $\gamma$ -subunits with 2'-(5''-phosphoribosyl)-3'-dephospho-CoA prosthetic groups (see Section 7.11.4.1.2), and catalytic  $\beta$ -subunits which promote the retro-Claisen cleavage of citrate and citramalate, respectively (refer to Section 7.11.5.4). During the lyase reaction these CoA-transferases (EC 2.8.3.10; EC 2.8.3.11) catalyze the exchange of acetyl- and citryl (or citramalyl) groups on the thiol of the prosthetic group to ensure the continuity of the catalytic cycle.<sup>301</sup> These transferases make use of a ternary complex mechanism in which no enzyme-bound intermediates are formed; instead, acyl transfer occurs by formation of a transient anhydride consisting of both the

substrate and the product acyl groups and subsequent attack of the prosthetic group's thiol on the carbonyl of the acceptor acyl group (**Scheme 8**, mechanism B). Although the structures of the  $\alpha$ -subunits of putative citrate lyases from *S. typhimurium* (pdb: 1XR4) and *Streptococcus mutans* (pdb: 2HJ0) have been determined, no detailed analysis of these structures has been published as yet. However, the preliminary assignments indicate that these enzymes are structurally similar to Class I enzymes.<sup>256,257</sup>

The last group of Class III CoA-transferases include enzymes involved in the metabolism of oxalate, carnitine, bile acids, toluene, and other aromatic compounds, as well as the succinyl-CoA:citramalate CoA-transferase that is involved in the 3-hydroxypropionate cycle of CO<sub>2</sub>-fixation.<sup>296,297,302</sup> The formyl-CoA:oxalate CoA-transferase (FCOR, EC 2.8.3.16) enzyme is the best-characterized example of a Class III transferase as both its crystal structure and its catalytic mechanism have been determined. X-ray crystallography studies show that Class III enzymes also have open  $\alpha/\beta$ -protein domains but are not structurally homologous to the Class I enzymes, belonging to the CoA-transferase family III (CaiB/BaiF) superfamily instead.<sup>256,257,297,303,304</sup> The mechanistic studies show that similar to the Class I enzymes these transferases also make use of acyl-enzyme intermediates during catalysis (**Scheme 8**, mechanism A). The formation of such enzyme-bound anhydride intermediates during catalysis has been confirmed by crystallography; these studies also indicate that these enzymes use an aspartate as a catalytic residue instead of a glutamate.<sup>305</sup> However, in contrast to the classical ping-pong mechanism of the Class I enzymes, kinetic studies of FCOR suggest that the carboxylic acid of the donor acyl-CoA does not leave the enzyme complex until final release of the product.<sup>306</sup> The enzyme also does not need the formation of the ternary complex to occur before catalysis can commence.

#### 7.11.5.2.4 Water as acceptor (thioesterases)

Acyl-CoA thioesterase enzymes (EC 3.1.2.-), although their catalytic activity simply entails the hydrolysis of CoA and ACP thioesters to release the fatty acids and other carboxylic acids bound to them (Equation (19)), have wide and varied physiological functions that includes the regulation of fatty acid metabolism and playing a central role in the biosynthesis of polyketide and nonribosomal peptide-based metabolites (especially the macrocyclic versions) and the degradation of aromatic compounds. These enzymes are thoroughly discussed in several recent reviews as well as the relevant chapters of this series that include fatty acids, polyketides, and nonribosomal peptide biosynthesis<sup>307,308</sup> (see Chapters 1.05, 1.02, and 5.19); therefore, only a brief overview of the structural and mechanistic diversity of acyl-CoA and acyl-ACP thioesterases is provided in this section.



Thioesterases can generally be described as being one of three types. The first includes all  $\alpha/\beta$  hydrolase-like enzymes that have a catalytic triad in their active sites and transfers the acyl group transiently to the serine or cysteine residue of this triad in a fashion similar to the mechanism depicted in **Scheme 7** (mechanism B) except that a water molecule acts as the final acyl acceptor. Myristoyl-ACP-specific thioesterase, the thioesterase I domains of PKSs and NRPSs, most fatty acid acyl-CoA thioesterases and the broad specificity *E. coli* thioesterase I (*tesA* gene product), which have been shown to also have protease and lipase activity, are all examples of type I thioesterases.<sup>309–311</sup> In fact, the thioesterase/protease I/lysophospholipase L<sub>1</sub> (TAP) enzyme is one of the representative members of the SGNH hydrolase superfamily whose members all have a strictly conserved Ser-Gly-Asn-His signature sequence and residues lining the oxyanion hole with an architecture different from the other  $\alpha/\beta$  hydrolases.<sup>312</sup>

In contrast, type II thioesterases generally adopt the unique 'hotdog' fold, which derives its name from the observation that the fold consists of a central long  $\alpha$ -helix wrapped in a  $\beta$ -sheet reminiscent of a sausage in a bun. Enzymes of this type include 4-hydroxybenzoyl-CoA thioesterases, *E. coli* thioesterase II (*tesB* gene product), the independently acting type II thioesterases that have been shown to have an error-correcting

function in polyketide and nonribosomal peptide biosynthesis, the broad specificity YciA acyl-CoA thioesterases from *H. influenzae* and *E. coli*, and human thioesterase superfamily member 2.<sup>313–317</sup> Instead of the catalytic triad of the  $\alpha/\beta$  hydrolases these enzymes all have either an Asp or a Glu residue in their active sites that have been implicated in catalysis. Two mechanisms of hydrolysis have been proposed: one in which the carboxylate acts as a general base, and a second in which it acts as a nucleophile to form an acyl–enzyme anhydride intermediate that is subsequently hydrolyzed. However, it is still unclear which of these mechanistic possibilities predominate in most of the hotdog-type thioesterases. Interestingly, the structures of the two 4-hydroxybenzoyl-CoA thioesterases that have been determined – one from *Artbrobacter* and the other from *Pseudomonas* – show that although both these enzymes have the same overall fold and act on the same substrate with comparable efficiency, both their quaternary structures and substrate-binding sites are decidedly different.<sup>318,319</sup> Also, the active site carboxylate is an Asp in one case and a Glu on the other, and these are positioned opposite to one another when the active sites of the two enzymes are superimposed. Whether these enzymes are therefore examples of either convergent or divergent evolution is a question that remains to be answered.

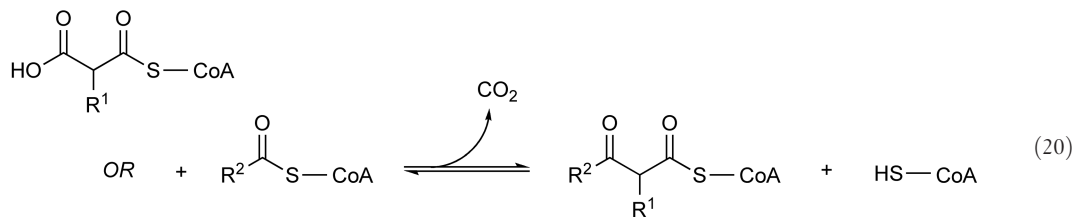
Recently, the structure of a putative third type of thioesterase from *M. tuberculosis* was determined.<sup>320</sup> This enzyme, shown to be a long-chain fatty acyl-CoA thioesterase (FcoT), has a hotdog fold similar to type II thioesterases, but does not have a carboxylate in its active site that can act as a nucleophile or general base during catalysis. Instead, it has been suggested that hydroxide ions are formed in the active site which subsequently attack the thioester; this mechanism is similar to the one described for the hydrolytic antibody D2.3, which have a similar active site architecture to that of FcoT. This difference in mechanism suggests that FcoT in fact represents a third type of thioesterase.

### 7.11.5.3 Condensing Enzymes: CoA Thioesters Acting as Enolate Nucleophiles toward Carbonyl-Based Electrophiles

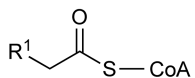
#### 7.11.5.3.1 Acyl-CoAs (or acyl-ACPs) as electrophiles: True Claisen condensations (thiolases)

The Claisen condensation – strictly the condensation of two esters with the concomitant release of one of the constituent alcohol moieties – is arguably the most important carbon–carbon bond-forming reaction in biology, as it is used in the biosynthesis of fatty acids, polyketides, and a large variety of other essential small molecule metabolites. Enzymes that catalyze this reaction all use either acyl-CoA or acyl-ACP thioesters to generate a reactive enolate nucleophile, which can be done in one of the two ways: by  $\alpha$ -decarboxylation, as in the case of reactions where malonyl thioesters are the substrates, or by direct  $\alpha$ -deprotonation of the reacting thioester. These enolates subsequently react with the carbonyl group of a second acyl-CoA or acyl-ACP thioester, giving a  $\beta$ -keto thioester and either CoA or ACP as products (Equation (20)).

#### Decarboxylating Claisen condensation



#### Nondecarboxylating Claisen condensation



**Note:**  $-\overset{\xi}{\text{S}}-\text{CoA}$  may also be  $-\overset{\xi}{\text{S}}-\text{ACP}$

Enzymes that catalyze such Claisen condensations (and some related reactions, see below) are all members of the thiolase superfamily (named for the first member of the family which was described, the degradative thiolase I from *S. cerevisiae*), as they all share a common fold despite a lack of sequence similarity and a large

diversity of substrate specificity among them.<sup>321</sup> The active sites of thiolase superfamily enzymes are defined by four catalytic loops that contain the key catalytic residues; among these a Cys residue contained in the N $\beta$ 3–N $\alpha$ 3 loop is strictly conserved. This Cys serves as the acceptor in the first acyl-transfer reaction of the mechanism shown in **Scheme 9(a)**. However, Claisen condensing enzymes differ with regard to their mechanism (decarboxylating vs. nondecarboxylating), the architecture of their active sites (i.e., the active site residues contained in the three other catalytic loops), and their biological function. Based on these considerations three different groups of condensing enzymes can be distinguished: first, the  $\beta$ -ketoacyl-ACP synthase (KAS)/PKS III group that consists of the so-called initiation enzymes – usually type III KASs (KAS III) and PKSs (PKS III). These enzymes normally use acyl-CoA thioesters as primers to catalyze the first decarboxylating condensation reaction of fatty acid and polyketide biosynthesis, although some PKS III enzymes (such as chalcone synthase) catalyze multiple condensations (all using acyl-CoAs as substrates) followed by a cyclization. The KAS III (FabH) from *E. coli*, which catalyzes the condensation of malonyl-ACP and acetyl-CoA, is the best-studied example of a KAS/PKS III-type enzyme. The second group is made up of the so-called elongation enzymes that use acyl-ACP thioesters as substrates in decarboxylating condensation reactions to elongate the acyl-ACP intermediates of fatty acid and polyketide biosynthesis. These usually are type-I and type-II  $\beta$ -ketoacyl-ACP and PKSs, and this group can therefore be referred to as KAS/PKS I/II. *Escherichia coli*'s FabB and FabF (KAS I and KAS II, respectively) are examples of such enzymes. Both KAS/PKS I/II and KAS/PKS III enzymes have (in addition to the strictly conserved Cys) an active site His residue in the C $\beta$ 2–C $\alpha$ 2 catalytic loop, but differ in the identity of the residue contained in the C $\beta$ 3–C $\alpha$ 3 loop: type III initiation enzymes have an Asn, whereas KAS/PKS I/II elongation enzymes have a His. This difference has been implicated in the observation that these enzymes have different sensitivities to known keto synthase inhibitors, such as thiolactomycin and cerulenin. The third group of Claisen condensing enzymes contains the nondecarboxylating biosynthetic thiolases that catalyze the condensation of two acetyl-CoA units to form acetoacetyl-CoA, and their degradative counterparts that catalyze the reverse reaction in the  $\beta$ -oxidation of fatty acids. In fact, both biosynthetic and degradative thiolases can catalyze either the Claisen condensation or the retro-Claisen cleavage, and often the reverse reaction is preferred – even for the so-called biosynthetic enzymes. Thiolases have significantly different active site architectures compared with the other two groups, and have Cys, Asn, and His residues in the N $\beta$ 3–N $\alpha$ 3, C $\beta$ 2–C $\alpha$ 2, and C $\beta$ 3–C $\alpha$ 3 catalytic loops, respectively. In addition they have a second Cys in the fourth C $\beta$ 4–C $\alpha$ 4 catalytic loop that acts as the base in the  $\alpha$ -deprotonation that leads to the formation of the reacting enolate, as illustrated by the mechanism of the thiolase from *Zoogloea ramigera* shown in **Scheme 9(b)**.<sup>322</sup>

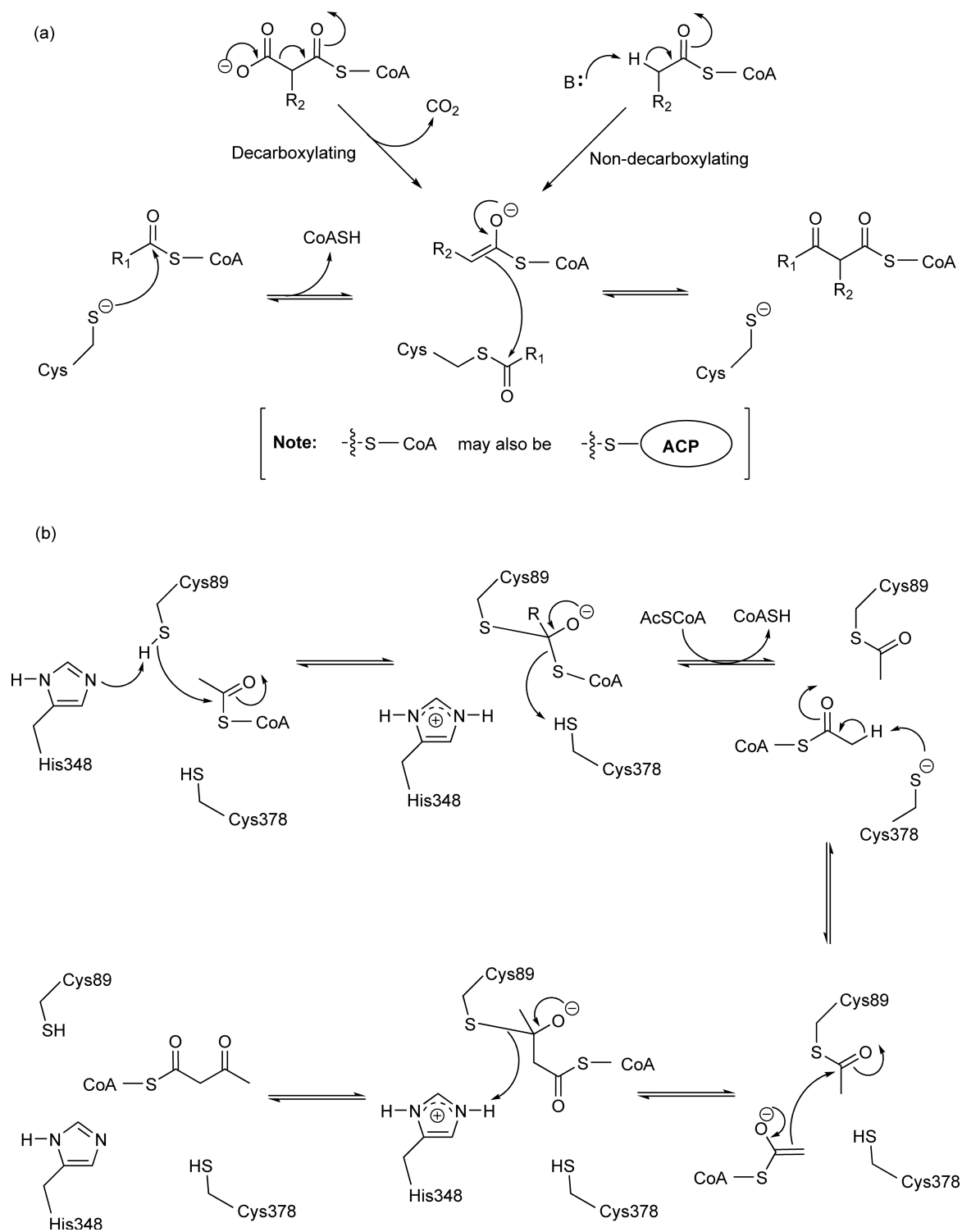
Because Claisen condensing enzymes form an integral part of the fatty acid or polyketide biosynthesis, the reader is referred to the relevant chapters of this series for an in-depth discussion of the KAS and PKS enzymes. Several excellent reviews of Claisen-type enzymes have also been published recently and provide detailed perspectives on the mechanism, inhibition, and structural aspects of the thiolase superfamily<sup>321,323</sup> (see Chapters 1.05 and 1.02).

#### 7.11.5.3.2 Ketones, aldehydes, and carboxylates as electrophiles

Apart from the true Claisen condensations discussed in the previous section in which the electrophilic reaction partner is another thioester, a number of enzymes also catalyze related Claisen-like condensations in which an acyl-CoA-based nucleophile reacts with other electrophilic carbonyl groups such as ketones, aldehydes, and the carboxylate group of carboxy-biotin. The most important examples of such enzymes are hydroxymethylglutaryl-CoA (HMG-CoA) synthase, citrate and homocitrate synthase (HCS), malate and  $\alpha$ -isopropylmalate synthase ( $\alpha$ -IPMS), and the biotin-dependent acetyl- and propionyl-CoA carboxylases.

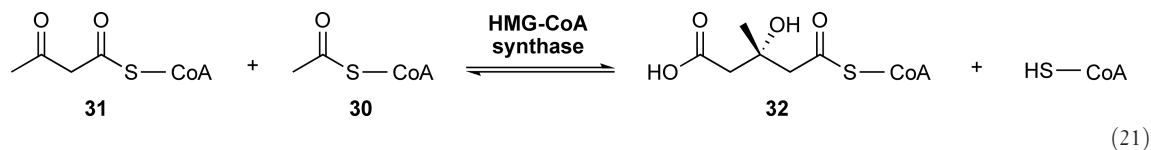
HMG-CoA synthase (EC 2.3.3.10) catalyzes the formation of (*S*)-3-hydroxy-3-methylglutaryl-CoA **32** through the condensation of acetyl-CoA **30** with acetoacetyl-CoA **31**, a key reaction in the mevalonate pathway of isopentenyl pyrophosphate (IPP) biosynthesis (Equation (21)). The enzyme is also a member of the thiolase superfamily (see above) and belongs to the KAS/PKS III group of condensing enzymes because it has, among other characteristics, the same active site residues as the other members of this group in three of the four catalytic loops (i.e., Cys, His, and Asn).<sup>321</sup> Its mechanism is similar to that of biosynthetic thiolases (as depicted in **Scheme 9(b)**), with the notable exception that instead of acting as the electrophile, the acetyl-enzyme intermediate is deprotonated by a glutamate acting as a general base to turn it into the enolate





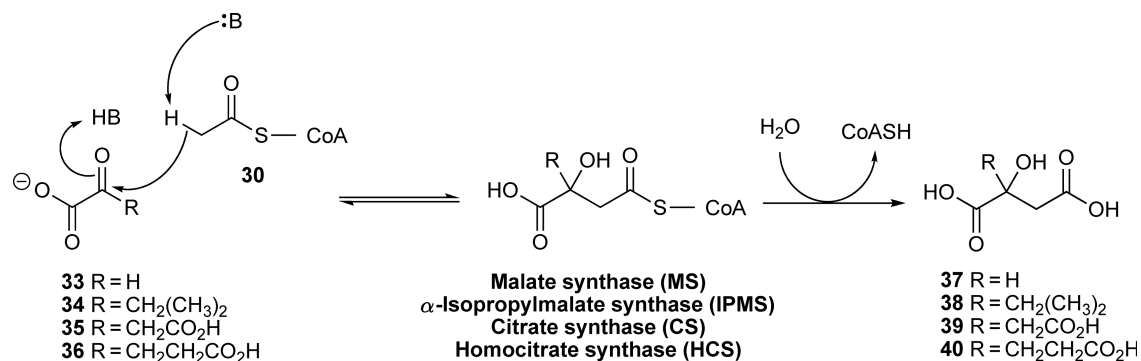
**Scheme 9** (a) The general mechanism of decarboxylating and nondecarboxylating Claisen condensing enzymes. (b) A mechanistic proposal for the thiolase enzyme from *Zoogloea ramigera* as based on its determined crystal structure.<sup>322</sup>

nucleophile; the enolate subsequently attacks the ketone of acetoacetyl-CoA. This exchanged reactivity pattern results in the formation of enzyme-bound HMG-CoA (the acyl-enzyme intermediate is not cleaved during the condensation reaction as in the case of thiolase) and consequently the last step of the reaction involves hydrolysis of the enzyme-bound adduct. Current crystal structure data suggest that the same glutamate that acts as general base in deprotonation of the acetyl-enzyme intermediate also facilitates the hydrolysis reaction through deprotonation of a water molecule acting as a nucleophile.<sup>324–326</sup>



Although malate synthase (MS, EC 2.3.3.9),<sup>327,328</sup>  $\alpha$ -IPMS (EC 2.3.3.13),<sup>329,330</sup> citrate synthase (CS, EC 2.3.3.1(*Si*); EC 2.3.3.3(*Re*)),<sup>331</sup> and HCS (EC 2.3.3.14),<sup>332</sup> all catalyze essentially the same reaction, that is, the condensation of an acetyl-CoA-based nucleophile with the carbonyl group of an  $\alpha$ -keto acid (glyoxylate **33**,  $\alpha$ -ketovalerate **34**, oxaloacetate **35**, and  $\alpha$ -ketoglutarate **36**, respectively), they do so using different mechanisms (Scheme 10). Although all of these enzymes are proposed to use a general base to deprotonate acetyl-CoA (no enzyme-bound intermediate is formed), they differ in how the carbonyl group is polarized in anticipation of the nucleophilic attack, and how the resulting oxyanion is stabilized. For CS two histidine residues play this role, but all the other enzymes use divalent cations instead: MS and IPMS are  $Mg^{2+}$ -dependent, and HCS is proposed to use  $Zn^{2+}$ . For some of these enzymes there is strong evidence that the enolization and condensation reactions are coupled, which suggests a concerted mechanism (as in the case of IPMS). Others, such as MS, clearly use stepwise mechanisms. These enzymes also catalyze the hydrolysis of the thioester bond of the Claisen condensation products, releasing the free acids **37–40** as final products of the reaction. Current knowledge suggests that they do so using a general base that abstracts a proton from a water molecule, which subsequently attacks the thioester carbonyl group. With the exception of IPMS, which follows a random bi-bi kinetic mechanism, all these enzymes follow ordered kinetic mechanisms. None of these Claisen condensing enzymes are members of the thiolase superfamily, nor do they share any structural similarity among them.<sup>333,334</sup> Instead, they each seem to have a unique fold, although both MS and IPMS have been shown to be TIM barrel enzymes.<sup>327,328,333,335</sup>

The final examples of CoA-dependent Claisen-like enzymes are the acetyl- and propionyl-CoA carboxylases (EC 6.4.1.2 and EC 6.4.1.3, respectively) in which acyl-CoA-based enolates react with the activated carboxylate bound to the biotin cofactor of these enzymes. Their enzymology is described in the chapter in this series including biotin-dependent enzymes (see Chapters 1.08 and 7.03).

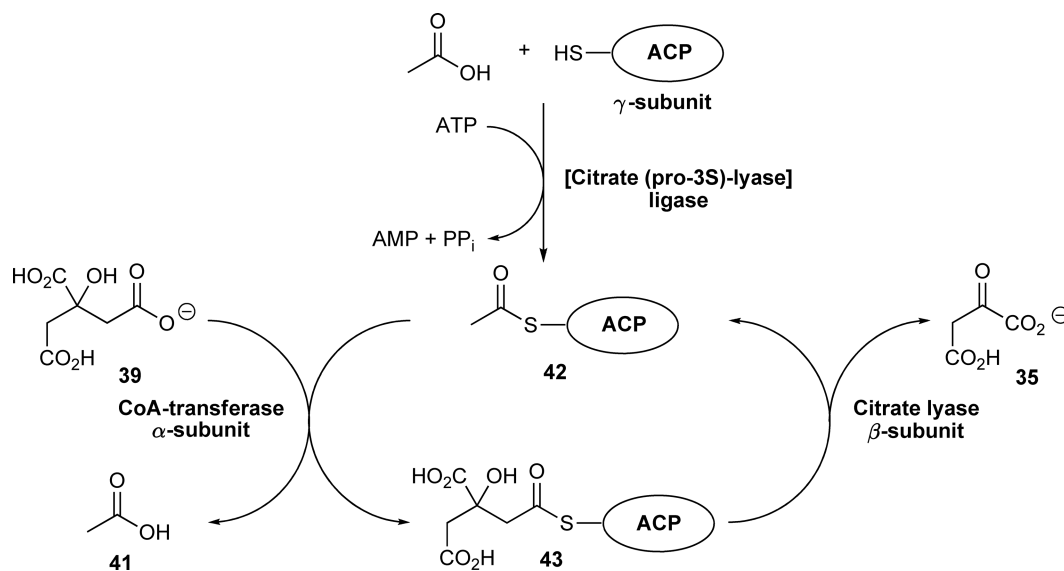


**Scheme 10** General reaction mechanism of the CoA-dependent Claisen-type condensing enzymes, malate synthase,  $\alpha$ -isopropylmalate synthase, citrate synthase, and homocitrate synthase.

### 7.11.5.4 Lyases: Using CoA to Break Carbon–Carbon bonds

In this section two enzymes that use CoA (or CoA-derived prosthetic groups) to cleave carbon–carbon bonds will be discussed: citrate lyase (EC 4.1.3.6) and ATP-citrate lyase, also called the citrate cleavage enzyme (EC 4.1.3.8, but recently transferred to 2.3.3.8). Although these are two of the few CoA-dependent activities that are labeled as carbon–carbon bond ‘lyases,’ it should be pointed out that many of the Claisen condensing reactions discussed in the previous section are reversible, and the enzymes that catalyze these reactions (most notably thiolase) are therefore also able to facilitate the cleavage of a C–C bond through similar retro-Claisen reactions. An interesting exception is HMG-CoA lyase (EC 4.1.3.4), which catalyzes the Claisen-type cleavage of HMG-CoA **32** to give acetyl-CoA **30** and oxaloacetate **35**. Although this is the reverse reaction of the condensation catalyzed by HMG-CoA synthase (see Equation (21)) there is little similarity between these two enzymes both with regard to structure and mechanism. HMG-CoA lyase is an  $Mg^{2+}$ -dependent enzyme with a mechanism similar to that of malate and IPMS. It is also structurally related to MS, as both are TIM barrel enzymes.<sup>336,337</sup>

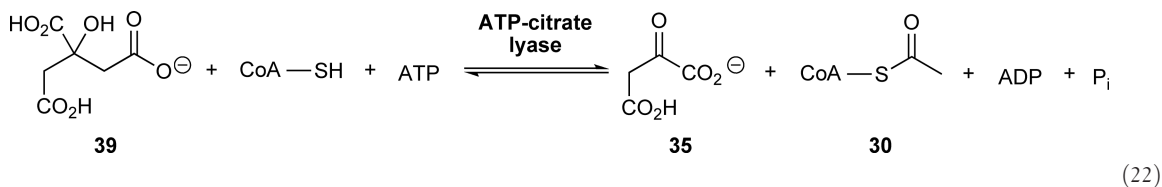
The ATP-independent citrate lyase that acts in the citrate fermentation pathways of mainly enterobacteria is one of the few enzyme activities that is dependent on the 2'-(5"-phosphoribosyl)-3'-dephospho-CoA prosthetic group for catalysis (see Section 7.11.4.1.2).<sup>205,206</sup> These enzymes are large complexes about 550 kDa in size, which consist of three different subunits in a 1:1:1 stoichiometry and with a native composition of  $\alpha_6\beta_6\gamma_6$ .<sup>338</sup> Although the  $\alpha$ - and  $\beta$ -subunits each have distinct enzymatic activities, the  $\gamma$ -subunits of this complex are noncatalytic ACPs that carry the essential prosthetic group. The lyase reaction proceeds in distinct steps: first, a separate [citrate (pro-3S)-lyase] ligase enzyme (EC 6.2.1.22) catalyzes the ATP-dependent transfer of an acetyl group to the thiol of the prosthetic group. This primes the complex for catalysis, as enzymes in which the acetyl group has been removed loses all activity. Second, the  $\alpha$ -subunit of the citrate lyase complex acts as a CoA-transferase (EC 2.8.3.10) and facilitates the exchange of the acetyl group for a citryl group. Finally, the  $\beta$ -subunit catalyzes the lyase reaction, the  $Mg^{2+}$ -dependent retro-Claisen cleavage of citryl-ACP **43** to give acetyl-ACP **42** (which reenters the catalytic cycle) and oxaloacetate **35** (Scheme 11). Trivial detail of the kinetic mechanism of the lyase reaction is known, and the active site residues that are involved in the cleavage reaction have not been identified. Only two crystal structures of citrate lyase  $\beta$ -subunits have been determined: the apo-CitE protein from *D. radiodurans* (pdb: 1SGJ) on which no detailed crystallographic



**Scheme 11** Reaction mechanism of the bacterial ATP-independent citrate lyases that are made up of three distinct subunits, each with its own function. The  $\alpha$ -subunit acts as a CoA-transferase, the  $\beta$ -subunit has the actual citrate lyase activity, and the  $\gamma$ -subunit acts as an ACP which has a unique 2'-(5"-phosphoribosyl)-3'-dephospho-CoA prosthetic group. The lyase catalytic cycle can only take place after the ACP has been primed by condensation of an acetyl group to the thiol of the prosthetic group; this is catalyzed by a separate AMP-forming ligase enzyme.

analysis has been published, and the apo-CitE protein from *M. tuberculosis*.<sup>339</sup> However, the latter protein is not associated with the usual  $\alpha$ - and  $\beta$ -subunits of a typical citrate lyase, and based on its distant sequence similarity to a malyl-CoA lyase it has been suggested that it functions as a citryl-CoA lyase, that is, catalyzing the CS reaction in reverse and forming acetyl-CoA and oxaloacetate as products.

A second ATP-dependent citrate lyase (ACL) activity is responsible for the formation of acetyl-CoA **30** and oxaloacetate **35** from citrate **39** with concomitant hydrolysis of ATP to ADP and phosphate (Equation (22)).<sup>340</sup> It is proposed to play a vital role in maintaining acetyl-CoA and oxaloacetate levels in most mammals, whereas in some bacteria it is an essential enzyme of the reductive tricarboxylic acid cycle (RTCA).



The reaction mechanism of ACL enzymes is complex and involves the formation of a phospho-enzyme intermediate by transfer of the  $\gamma$ -phosphate of ATP to a conserved histidine residue. The phospho-histidine subsequently transfers its phosphate to citrate, forming a tightly bound citryl-phosphate intermediate, which then reacts with CoA to give citryl-CoA.<sup>341</sup> Up to this point the enzyme shows great similarity in mechanism to the succinyl-CoA synthetase described in an earlier subsection (see Section 7.11.5.1.1). In fact, comparison of the primary sequence of the *Chlorobium tepidum* ACL enzyme with that of *E. coli*'s succinyl-CoA synthetase indicate up to 33% sequence identity, as well as the presence of a putative ATP grasp domain in the former similar to those found in other ADP-forming ACSs.<sup>342</sup> Also, in agreement with the reaction mechanism of these enzymes, no evidence currently exists suggesting the formation of a covalent citryl-enzyme intermediate during the ACL-catalyzed reaction as has been proposed in some earlier studies.<sup>341</sup> In the last step citryl-CoA is cleaved in a retro-Claisen reaction to give acetyl-CoA and oxaloacetate. The residues that are involved in this last step have not been unequivocally identified, despite some kinetic and site-directed mutagenesis studies of the *Chlorobium* ACL enzymes.<sup>342,343</sup> No structure of an ACL enzyme has been determined to date.

#### 7.11.5.5 Enzymes Modifying the Acyl Moiety of CoA Thioesters

By far the largest proportion of supposed acyl-CoA-dependent enzymes in fact only depend on CoA for binding; all the chemistry takes place in the acyl moiety of the thioester, and CoA itself is not chemically involved in the transformation. Enzymes that fall into this category include various oxidoreductases, hydratases, dehydratases, desaturases, and even dehalogenases, most of which are involved in fatty acid and polyketide biosynthesis, or the degradation of various aromatic compounds. **Table 4** provides a selection of some of these enzymes and the type of reaction they catalyze. For more information on their enzymology, the reader is referred to the relevant chapter in this series that deals with the chemistry and enzymology of the involved cofactors, the class of compounds concerned or the specific reaction type. The 4-chlorobenzoyl-CoA dehalogenase enzyme is an especially interesting case that is discussed at length in the chapter on catalytic promiscuity in enzymes (see Chapters 1.08 and 7.03).

#### 7.11.6 Chemical Biology Tools Based on CoA Enzymology

CoA-based enzymology has been manipulated and exploited since the cofactor's structure first became known, and especially since its biosynthetic pathway has been reconstituted. Of special interest has been the use of CoA analogues as a means to further our understanding of the mechanisms of CoA-dependent enzymes, as a tool for the discovery and application of specific probes and inhibitors, and as a basis on which to build a new

**Table 4** A selection of acyl-CoA-dependent enzymes and the type of reactions they catalyze

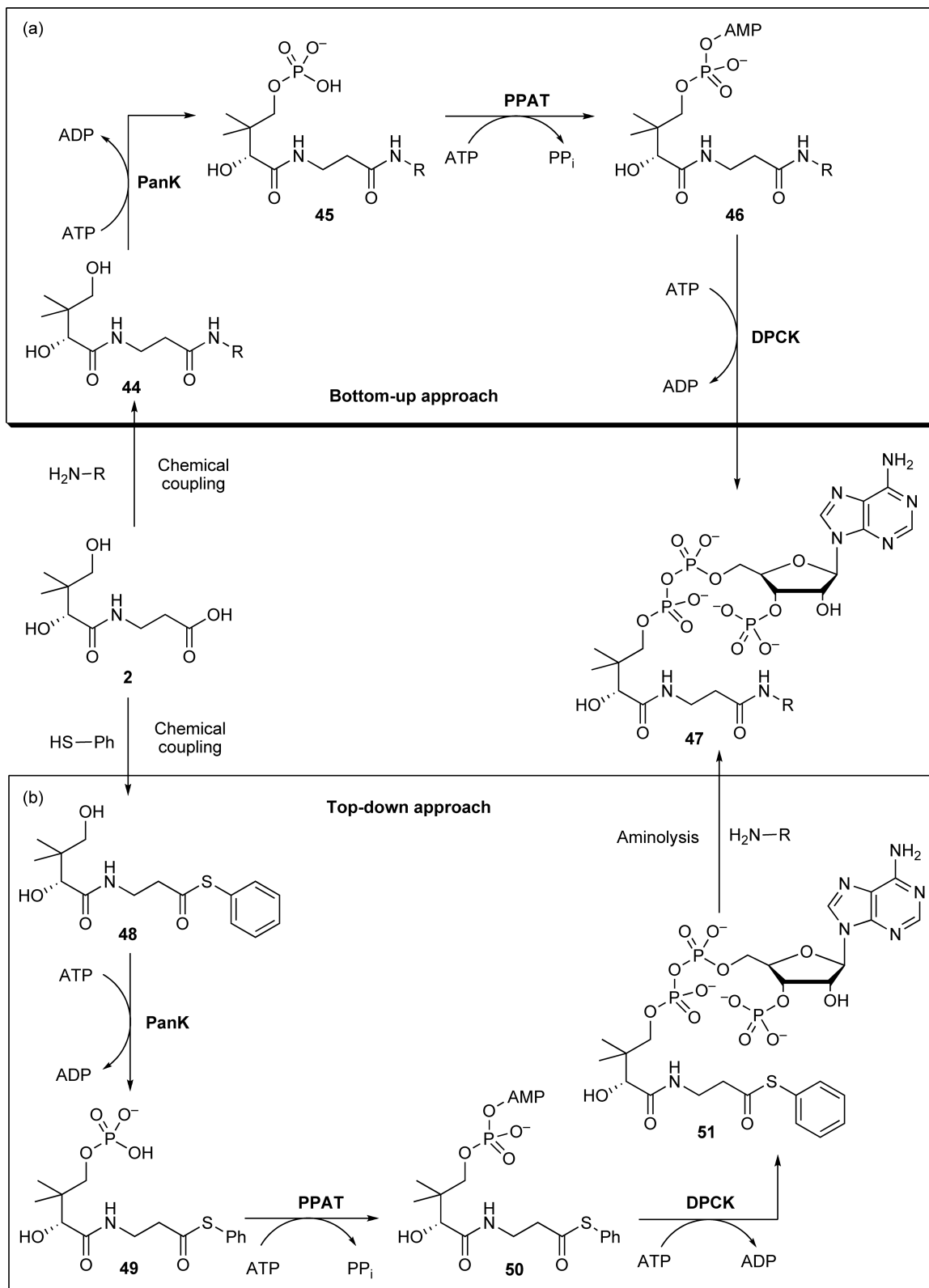
Entry	Enzyme	EC number	Reaction catalyzed
1	HMG-CoA reductase	EC 1.1.1.34	NADPH-dependent reduction of HMG-CoA to give mevalonate
2	3-Hydroxyacyl-CoA dehydrogenase	EC 1.1.1.35/36	Reversible NAD(P) <sup>+</sup> -dependent oxidation of alcohol to ketone
3	2-Enoyl-CoA reductase	EC 1.3.1.38	NADPH-dependent saturation of $\alpha,\beta$ -unsaturated acyl-CoAs
4	Acyl-CoA oxidase	EC 1.3.3.6	Flavin-dependent desaturation of saturated acyl-CoAs
5	Acyl-CoA dehydrogenases	EC 1.3.99. –	
6	Benzoyl-CoA reductase	EC 1.3.99.15	Reduction of aromatic ring
7	Stearoyl-CoA desaturase	EC 1.14.19.1	$\Delta^9$ -desaturation of stearoyl-CoA
8	4-Chlorobenzoyl-CoA dehalogenase	EC 3.8.1.7	Exchange of Cl for OH on 4-position of aromatic ring
9	Enoyl-CoA hydratase	EC 4.2.1.17	Hydration of $\alpha,\beta$ -unsaturated acyl-CoAs
10	3-Hydroxybutyryl-CoA epimerase	EC 5.1.2.3	Epimerization at 3-position through a dehydration/hydration mechanism
11	Methylmalonyl-CoA mutase	EC 5.4.99.2	Cobalamin-dependent isomerization of methylmalonyl-CoA to succinyl-CoA

biotechnology for the orthogonal labeling of proteins with reporter labels. This section will highlight the current state of development of CoA-based enzymology as a tool in chemical biology studies.

### 7.11.6.1 Preparation of CoA Analogues

The various methods for the preparation of CoA analogues have evolved drastically since the total synthesis of CoA was first published. Initial protocols were based on the chemical modification of CoA itself because of the difficulties involved in introducing the adenylate and 3'-phosphate moieties using chemical transformations. However, it became increasingly easier to prepare CoA analogues after the CoA biosynthetic enzymes were identified, cloned, and overexpressed, as this made it possible to synthetically prepare 4'-phosphopantothenamide analogues **45**, which could then be adenylated and 3'-phosphorylated using the PPAT and DCPK enzymes, respectively.<sup>133,344</sup> These initial protocols used phosphorylated substrates because it was found that the feedback inhibition PanK enzymes experienced in the presence of increasing concentrations of the CoA product made them unsuitable biocatalysts. Subsequent procedures addressed this problem by either adding the enzymes in sequence,<sup>345,346</sup> or by using the PanK enzyme from *S. aureus*, which is not influenced by feedback inhibition.<sup>114</sup> Such chemo-enzymatic procedures have since become the mainstay of preparing CoA and CoA analogues that have modified pantetheine moieties. However, few studies have been reported on the modification of CoA in its ribose and adenine moieties, most probably because such changes would make it impossible to use the CoA biosynthetic enzymes for biotransformation of the precursors, and would therefore require a significant investment in the development of new synthetic protocols.

Currently, two different strategies are generally used for the preparation of CoA analogues. The first, a bottom-up approach, requires that a pantothenamide analogue **44** be prepared synthetically. This is usually done from pantothenic acid **2**, although a modular synthesis has also been developed.<sup>347</sup> This starting material is subsequently 4'-phosphorylated (using the *E. coli* PanK-I or *S. aureus* PanK-II), adenylated (using *E. coli* PPAT), and 3'-phosphorylated (using *E. coli* DCPK) (**Scheme 12**, panel (a)).<sup>115</sup> This strategy is ideal if an analogue is required in larger amounts (e.g., for use in mechanistic or labeling studies), but may become tedious if only small amounts of many different CoAs are needed. In such cases an alternative top-down approach can be followed in which a single so-called pre-CoA **51** is prepared first, followed by its transformation into any number of CoA analogues (**Scheme 12**, panel (b)).<sup>114</sup> This is made possible by the presence of a reactive thioester in the pre-CoA which will undergo aminolysis when treated with primary and secondary amines to give the requisite CoA analogues. The pre-CoA therefore acts as a CoA synthon and allows the parallel preparation of any number of analogues simply by exchanging the amine that is used in the aminolysis reaction. This strategy



**Scheme 12** Strategies for the chemoenzymatic preparation of CoA analogues. In the bottom-up approach (panel A), pantothenamides are prepared synthetically, and subsequently enzymatically transformed into the corresponding CoA analogue by action of the PanK, PPAT, and DPCK enzymes. In the top-down approach (panel B), a pre-CoA containing a reactive thioester (preferably a phenyl thioester) is prepared first by chemical synthesis of a pantothenic acid thioester **48** and its subsequent biotransformation by the CoA biosynthetic enzymes. This pre-CoA can then act as a virtual CoA synthase, and can be converted into any number of CoA analogues by aminolysis of the thioester by primary and secondary amines.



is ideally suited to the preparation of many CoA analogues for screening or selection purposes as it can be adopted for use in multi-well plates, although it is limited with regard to the amounts of analogue that can be produced. Unfortunately, no simple or high-throughput technique is currently available for the purification of these CoA analogues, regardless of how they were prepared. In previous studies preparative HPLC, ion exchange, and solid-phase extraction (SPE) have successfully been used, although analogues are also often prepared and used *in situ* (i.e., without purification). The end user therefore has to determine whether purification is necessary considering the downstream applications for which the analogues are prepared.

### 7.11.6.2 Applications of CoA Analogues in Chemical Biology

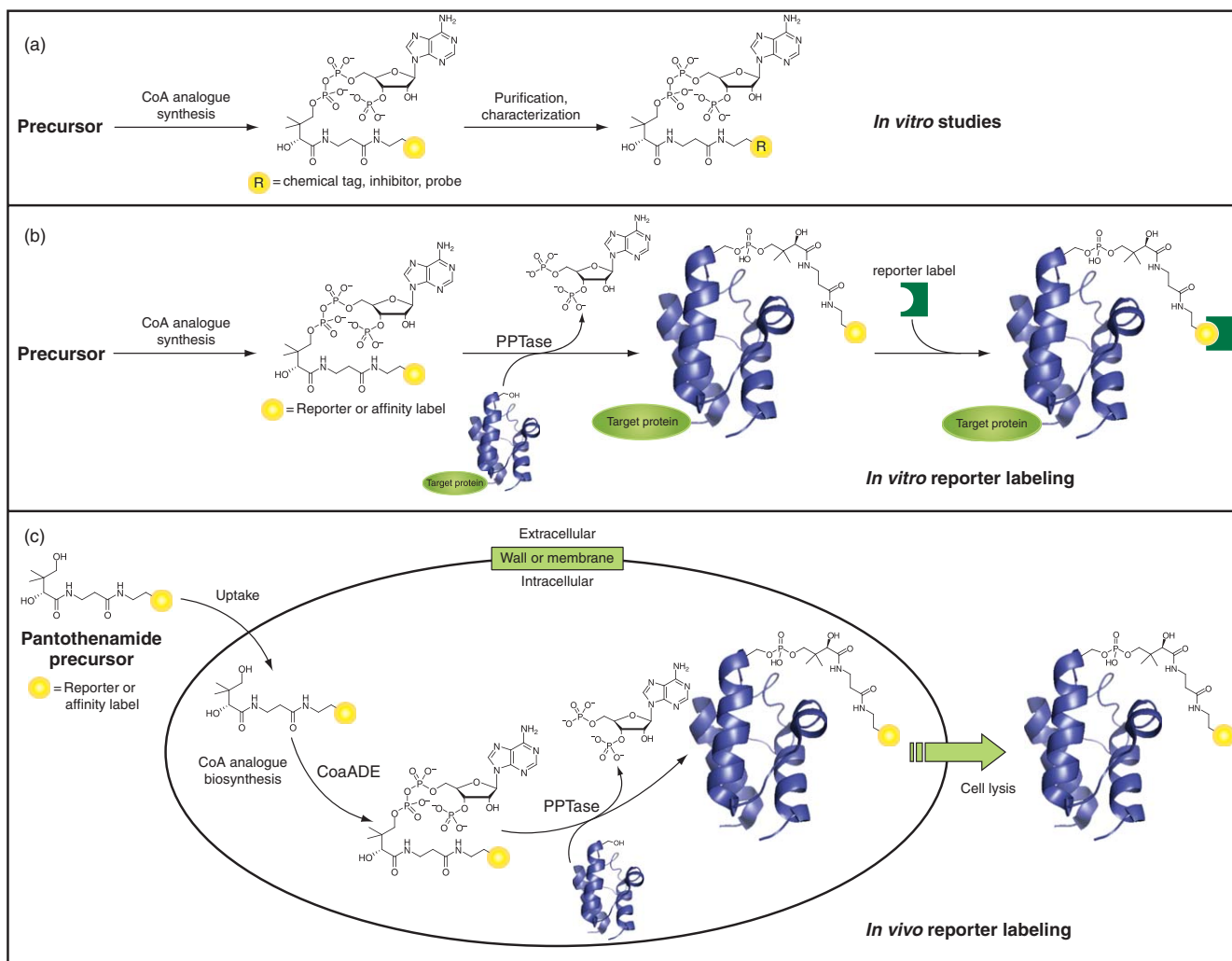
The relative ease with which most CoA analogues can currently be prepared has resulted in them finding widespread use as mechanistic and reactivity probes, as enzyme inhibitors and, much more recently, as reporter labels. In general, two different strategies for the application of CoA analogues in chemical biology studies can be distinguished. First, for *in vitro* studies, CoA analogues are prepared by one of the methods outlined in the previous section, purified and characterized (if necessary), and subsequently used for the study at hand (Figure 4, panel (a)). If the study does not require that the analogue has to be used at a specific concentration it may be prepared *in situ*, as is often done where CoA analogues serve as reporter labels (Figure 4, panel (b)). However, a different approach is used where the main aim of the study is to have the CoA analogue be available *in vivo*. In this second strategy only the pantothenamide analogue precursor is prepared by an appropriate synthetic method,<sup>348</sup> after which it is used to treat the cells in which the CoA analogue is required. This strategy exploits the known characteristics of most pantothenamides: their ability to cross cellular membranes (either by diffusion or by an unidentified transporter), and – once inside the cell – their ability to act as alternative substrates of its native CoA biosynthetic machinery. In this manner the analogue is biosynthesized and made available *in vivo*, where it most often is utilized in the alternative posttranslational modification of carrier proteins by PPTase enzymes (Figure 4, panel (c)). This *in vivo* labeling of carrier proteins has received much recent interest as a bio-orthogonal labeling technique with potential in reporter studies. The various applications of CoA analogues by either strategy are summarized in Sections 7.11.6.2.1 and 7.11.6.2.2.

#### 7.11.6.2.1 CoA analogues as mechanistic and reactivity probes, and as enzyme inhibitors

A broad range of CoA analogues have been prepared for use in studies geared toward a better understanding of the mechanism and reactivity of CoA- or CoA thioester-utilizing enzymes, or the development of inhibitors of such enzymes. In many of these cases the analogues were engineered to have specific chemical entities – such as reactive electrophilic centers – that could be exploited either in the enzyme's mechanistic elucidation or in its mechanism-based inhibition. Other CoA analogues have been designed as structural mimics of proposed transition states or of the enzymes' native substrates. An extensive review of CoA analogues and their varied uses was published in 2000, and also provides a background on the different methods for their preparation.<sup>3</sup> A selection of analogues that have been prepared since then, together with their anticipated targets and/or uses, is presented in Table 5. It is notable that the synthesis and use of CoA-based bisubstrate type inhibitors (Table 5, entry 4) have increased significantly in recent years, mainly as specific inhibitors of the p300 HAT and melatonin rhythm enzymes. Also, various activity-based substrate probes have recently been developed (Table 5, entries 1, 2, and 13) and suggests that such compounds may have significant impact as proteomic tools in future studies.

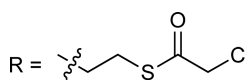
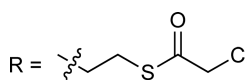
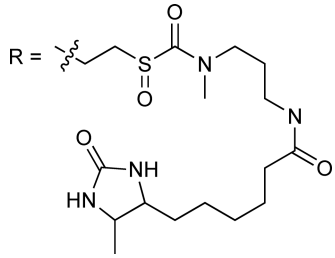
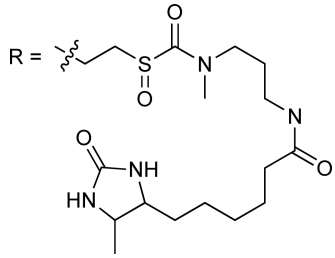
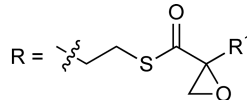
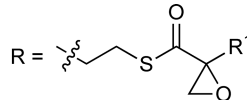
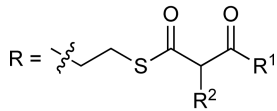
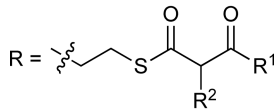
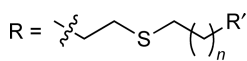
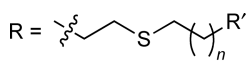
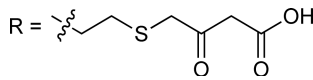
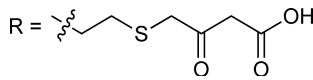
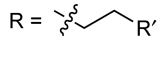
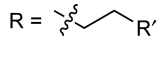
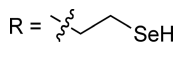
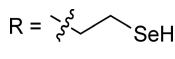
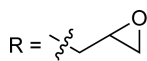
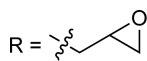
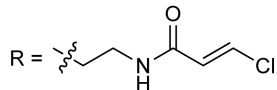
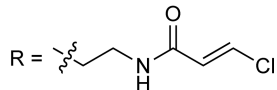
#### 7.11.6.2.2 CoA analogues as reporter labels

The finding that promiscuous PPTase enzymes will accept nearly any CoA analogue as a substrate to selectively modify carrier proteins with its phosphopantothenamide moiety<sup>369</sup> (see Section 7.11.4.1.1) has led to the development of a range of versatile biotechnological applications based on this process.<sup>370</sup> In general, most of these applications have exploited the fact that the thiol of CoA itself can easily be modified to incorporate fluorescent groups or affinity tags (such as biotin), thereby making any number of reporter-labeled CoA analogues readily available. Also, the Sfp-type PPTase from *B. subtilis* has become the enzyme of choice to



**Figure 4** Strategies for the application of CoA analogues. For *in vitro* studies appropriate analogues are prepared from a suitable precursor, purified, and characterized if necessary, and then normally used as mechanism-based probes and inhibitors (panel A). When CoA analogues are used for the *in vitro* site-specific modification of proteins by transfer of a reporter label to a carrier protein module using a PPTase enzyme, the analogue can be prepared and used *in situ* (panel B). *In vivo* reporter labeling is made possible when cells are provided with suitably modified pantothenamide precursors, which are subsequently transformed into the respective CoA analogues and transferred a carrier protein by the cell's native CoA biosynthesis (CoaADE) and PPTase enzymes. The labeled protein can be recovered from the cells by cell lysis.

**Table 5** A selection of CoA analogues that have been used as mechanistic or reactivity probes or as enzyme inhibitors since 2000

Entry	Structures	Enzyme target/Use	References
1	 <p>R = </p>	1. Substrate for the preparation of various <i>N</i> -acetyltransferase bisubstrate inhibitors 2. Activity-based substrate profiling of <i>N</i> -acetyltransferases from the GNAT superfamily	<a href="#">349</a> <a href="#">350</a>
2	 <p>R = </p>	Activity-based substrate probe of <i>N</i> -acetyltransferases	<a href="#">351</a>
3	 <p>R = </p>  <p>R = </p> <p>R<sup>1</sup> = alkyl; R<sup>2</sup> = Br, H</p>	Inhibitors of <i>N</i> -myristoyltransferase	<a href="#">352</a>
4	 <p>R = </p> <p>R' = second substrate, <i>n</i> = 0, 1</p>	Bisubstrate inhibitors of: 1. Arylalkylamine <i>N</i> -acetyltransferase 2. p300 Histone acetyltransferase 3. Spermidine/spermine acetyltransferase 4. Acetyl-CoA carboxylase 5. Aminoglycoside 6'- <i>N</i> -acetyltransferase	<a href="#">353–355</a> <a href="#">356–359</a> <a href="#">360</a> <a href="#">361</a> <a href="#">362–365</a>
5	 <p>R = </p>	Trap of intermediates of polyketide biosynthesis	<a href="#">366</a>
6	 <p>R = </p> <p>R' = alkyl, alkoxy, thioalkoxy</p>	1. Fatty acid biosynthesis 2. Enoyl-CoA hydratase	<a href="#">63, 76, 100</a> <a href="#">345</a>
7	 <p>R = </p>	Acetyl-CoA synthase mechanistic probe	<a href="#">367</a>
8	 <p>R = </p>  <p>R = </p>	Mechanism-based protein cross-linking probes selective for carrier proteins	<a href="#">368</a>

perform these carrier protein modifications because of its demonstrated permissiveness with regard to both its CoA and carrier protein substrates. The combination of these tools has allowed the labeling of both carrier proteins themselves (thereby converting them into the so-called *crypto*-ACPs), as well as a variety of other target proteins; the latter is done by fusing the protein in question to an ACP, which is subsequently modified.<sup>115</sup> This technology has been used to label proteins on the surfaces of cells, to modify target proteins *in vitro* for immobilization, and in microarray studies.<sup>371–373</sup> Recently it has been shown that Sfp and AcpS PPTases will also act on short peptide tags that contain a serine residue within a minimal recognition sequence. Proteins therefore only have to be fused to these short tags to be targeted for site-specific labeling.<sup>374,375</sup> Additionally, a variety of fluorescent- and azide- or alkyne-tagged pantothenamides have been developed to perform *in vivo* carrier protein labeling, with the latter analogues being ideally suited for ‘click’ chemistry modification of the proteins after cell lysis.<sup>64,65</sup> These biotechnological applications of CoA analogues in carrier protein-based modifications are described in detail in a recent review.<sup>184</sup> Finally, azido-tagged analogues of myristoyl-CoA analogues have been shown to act as substrates of the enzyme myristoyl-CoA transferase, and was used to selectively modify target proteins under both *in vitro* and *in vivo* conditions as part of a bio-orthogonal labeling strategy.<sup>376</sup>

### 7.11.7 Drug Development Efforts Targeting CoA Enzymology

Any attempt at an in-depth discussion of the compounds that inhibit CoA-depending enzymes will occupy the bulk of the pages of this chapter; their number is vast, especially considering the recent heightened interest in the development of inhibitors of fatty acid biosynthesis and metabolism. A recent review specifically focused on providing an account of the compounds and their enzyme targets that have been developed for this latter purpose.<sup>377</sup> This section focuses on the drug development efforts that target the CoA biosynthetic enzymes.

Although all the CoA biosynthetic enzymes are potential targets for the development of selective inhibitors – with the possible exception of PPCS and PPCDC, as these enzymes may be bypassed if an organism can salvage pantetheine (see Section 7.11.3) – only two CoA biosynthetic enzymes have received special attention in this regard: PanK and PPAT, both of which are known to play key regulatory roles in maintaining the intracellular CoA concentration. Most of the studies on PanK inhibitors have centered on the pantothenamide family of pantothenic acid analogues; these compounds have an *N*-substituted amide instead of the carboxylic acid of pantothenic acid **2**. The best-characterized examples of such pantothenamide inhibitors are those with *N*-pentyl and *N*-heptyl substituents,<sup>63,76,102</sup> although a large variety of others substituents have also been prepared.<sup>100</sup> However, strictly speaking the pantothenamides are not inhibitors of PanK but rather alternate substrates of the enzyme, often with similar or better  $k_{\text{cat}}/K_{\text{m}}$  values than that for pantothenate.<sup>63</sup> They therefore do not inhibit the enzyme’s phosphorylation activity but rather tie it up by providing a structurally equivalent 4’-hydroxyl group to phosphorylate. Even homopantothenic acid, an analogue of pantothenic acid in which the  $\beta$ -alanine moiety have been replaced with  $\gamma$ -aminobutyric acid (i.e., the insertion of a methylene group), which have been shown to reduce the levels of CoA in the livers of treated mice, acts in a similar manner.<sup>378</sup> Nonetheless, the resulting 4’-phosphopantothenamides are subsequently adenylylated by PPAT and 3’-phosphorylated by DPCK to form CoA antimetabolites that indeed act as inhibitors of CoA-dependent enzymes, especially fatty acid biosynthesis.<sup>63,76</sup> The continued development and study of pantothenamides as a class of inhibitors are therefore of considerable use and importance, but the latest studies suggest that they are unlikely to abolish PanK activity altogether. In fact, to date no compound has been described that directly inhibits the ability of a PanK enzyme to perform its ATP-dependent phosphorylation reaction.

A number of actual PPAT inhibitors have been described, although much of these have only been reported in the patent literature. The classes of compound that have been tested include bispiperidines, [(heterocycl)yl)methylene]-substituted benzoic acids and nicotinic acids, aromatic sulfonamides, dipeptides, and tetrahydro- $\beta$ -carboline derivatives. Of these, the dipeptide Fmoc-Glu-D-His-NHCH<sub>2</sub>CH(CH<sub>3</sub>)<sub>2</sub> exhibited the best inhibitory activity and selectivity with an IC<sub>50</sub> of 6 nmol l<sup>-1</sup> against the *E. coli* enzyme and no inhibition of porcine PPAT.<sup>379</sup>

A recent review provides a thorough and extensive review on all the compounds that have been prepared and tested as inhibitors of CoA biosynthesis, and especially those that were considered as potential antimalarial

agents.<sup>11</sup> Not all of these have been thoroughly characterized with regard to their enzymatic targets, and consequently many may still serve as new scaffolds for the design of specific CoA-targeting drugs.

### 7.11.8 Conclusion

More than 60 years of noteworthy studies on CoA and its dependent enzymes have had a tremendous impact on the field of enzymology. However, the fact that so much about CoA enzymology is already known does not necessarily mean there is not more to discover: for example, as this conclusion was being written a study appeared reporting on a CoA-dependent hydratase enzyme which surprisingly relies on a ketyl radical intermediate for catalysis.<sup>380</sup> However, what is even more surprising about this finding was the unprecedented observation that the enzyme does not make use of any of the usual enzyme-based radical generators – in fact, the radical is generated directly from the substrate. One can therefore only conclude that even at this seemingly advanced stage of our understanding of enzymology, both CoA and CoA-dependent enzymes have much to teach us and many secrets to reveal. Hopefully many budding enzymologists will be enticed by this prospect and join the exciting search for more of such examples.

### References

1. F. Lipmann, *Science* **1954**, *120*, 855–865.
2. G. D. Novelli; F. Lipmann, *Arch. Biochem.* **1947**, *14*, 23–27.
3. P. K. Mishra; D. G. Drueckhammer, *Chem. Rev.* **2000**, *100*, 3283–3309.
4. W. Yang; D. G. Drueckhammer, *Org. Lett.* **2000**, *2*, 4133–4136.
5. W. Yang; D. G. Drueckhammer, *J. Am. Chem. Soc.* **2001**, *123*, 11004–11009.
6. J. Baddiley; E. M. Thain; G. D. Novelli; F. Lipmann, *Nature* **1953**, *171*, 76.
7. J. G. Moffatt; H. G. Khorana, *J. Am. Chem. Soc.* **1961**, *83*, 663–675.
8. R. Leonardi; Y.-M. Zhang; C. O. Rock; S. Jackowski, *Prog. Lipid Res.* **2005**, *44*, 125–153.
9. E. Strauss; C. Kinsland; Y. Ge; F. W. McLafferty; T. P. Begley, *J. Biol. Chem.* **2001**, *276*, 13513–13516.
10. T. P. Begley; C. Kinsland; E. Strauss, *Vitam. Horm.* **2001**, *61*, 157–171.
11. C. Spry; K. Kirk; K. J. Saliba, *FEMS Microbiol. Rev.* **2008**, *32*, 56–106.
12. R. J. Williams; C. M. Lyman; G. H. Goodyear; J. H. Truesdail; D. Holaday, *J. Am. Chem. Soc.* **1933**, *55*, 2912–2927.
13. S. Lepkovsky, *Annu. Rev. Biochem.* **1940**, *9*, 383–422.
14. A. F. Morgan, *Annu. Rev. Biochem.* **1941**, *10*, 337–394.
15. R. J. Williams; H. H. Weinstock, Jr.; E. Rohrmann; J. H. Truesdail; H. K. Mitchell; C. E. Meyer, *J. Am. Chem. Soc.* **1939**, *61*, 454–457.
16. R. K. Hill; T.-H. Chan, *Biochem. Biophys. Res. Commun.* **1970**, *38*, 181–183.
17. F. Lipmann; L. C. Tuttle, *J. Biol. Chem.* **1944**, *153*, 571–582.
18. F. Lipmann, *J. Biol. Chem.* **1945**, *160*, 173–190.
19. F. Lipmann; N. O. Kaplan, *J. Biol. Chem.* **1946**, *162*, 743–744.
20. F. Lipmann, *Bacteriol. Rev.* **1953**, *17*, 1–16.
21. F. Lipmann, *Annu. Rev. Biochem.* **1984**, *53*, 1–33.
22. F. Lynen; E. Reichert, *Angew. Chem.* **1951**, *63*, 47–48.
23. F. Lynen; E. Reichert; L. Rueff, *Ann. Chem.* **1951**, *574*, 1–32.
24. T. Wieland; E. F. Moller, *Z. Physiol. Chem.* **1941**, *269*, 227–235.
25. W. K. Maas, *J. Biol. Chem.* **1952**, *198*, 23–32.
26. W. K. Maas; G. D. Novelli, *Arch. Biochem. Biophys.* **1953**, *43*, 236–238.
27. W. K. Maas; H. J. Vogel, *J. Bacteriol.* **1953**, *65*, 388–393.
28. G. M. Brown, *Physiol. Rev.* **1960**, *40*, 331–368.
29. G. W. E. Plaut; J. J. Bethel, *Annu. Rev. Biochem.* **1956**, *25*, 463–496.
30. B. C. Johnson, *Annu. Rev. Biochem.* **1955**, *24*, 419–464.
31. W. S. Pierpoint; D. E. Hughes; J. Baddiley; A. P. Mathias, *Biochem. J.* **1955**, *61*, 368–374.
32. L. Levintow; G. D. Novelli, *J. Biol. Chem.* **1954**, *207*, 761–765.
33. Y. Abiko; S. Ashida; M. Shimizu, *Biochim. Biophys. Acta* **1972**, *268*, 364–372.
34. G. B. Ward; G. M. Brown; E. E. Snell, *J. Biol. Chem.* **1955**, *213*, 869–876.
35. E. E. Snell; G. M. Brown, Pantethine and Related Forms of the Lactobacillus Bulgaricus Factor (LBF). In *Advances in Enzymology and Related Areas of Molecular Biology*; F. F. Nord, Ed.; Interscience Publishers, Inc., 1953, Vol. 14. pp 49–71. doi: 10.1002/9780470122594.ch2.
36. G. M. Brown, *J. Biol. Chem.* **1959**, *234*, 370–378.
37. G. M. Brown, *J. Am. Chem. Soc.* **1958**, *80*, 3161.
38. T. Suzuki; Y. Abiko; M. Shimizu, *J. Biochem. (Tokyo)* **1967**, *62*, 642–649.
39. Y. Abiko, *Methods Enzymol.* **1970**, *18*, 358–364.

40. G. M. Brown; J. J. Reynolds, *Annu. Rev. Biochem.* **1963**, *32*, 419–462.
41. S. D. Dunn; E. E. Snell, *J. Bacteriol.* **1979**, *140*, 805–808.
42. W. J. Song; S. Jackowski, *J. Bacteriol.* **1992**, *174*, 6411–6417.
43. S. Jackowski, Biosynthesis of Pantothenic Acid and Coenzyme A. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*; F. C. Neidhardt, R. Curtiss, C. A. Gross, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaechter, H. E. Umbarger, Eds.; American Society for Microbiology: Washington, DC, 1996; pp 687–694.
44. S. Jackowski; C. O. Rock, *J. Bacteriol.* **1981**, *148*, 926–932.
45. M. E. Webb; A. G. Smith; C. Abell, *Nat. Prod. Rep.* **2004**, *21*, 695–721.
46. S. Y. Gerdes; M. D. Scholle; M. D'Souza; A. Bernal; M. V. Bae; M. Farrell; O. V. Kurnasov; M. D. Daugherty; F. Msee; B. M. Polanuyer; J. W. Campbell; S. Anantha; K. Y. Shatalin; S. A. K. Chowdhury; M. Y. Fonstein; A. L. Osterman, *J. Bacteriol.* **2002**, *184*, 4555–4572.
47. S. Jackowski; J. H. Alix, *J. Bacteriol.* **1990**, *172*, 3842–3848.
48. D. S. Vallari; C. O. Rock, *J. Bacteriol.* **1985**, *162*, 1156–1161.
49. H. Nakamura; Z. Tamura, *J. Nutr. Sci. Vitaminol.* **1973**, *19*, 389–400.
50. D. S. Vallari; C. O. Rock, *J. Bacteriol.* **1985**, *164*, 136–142.
51. J. Reizer; A. Reizer; M. H. Saier, Jr., *Res. Microbiol.* **1990**, *141*, 1069–1072.
52. J. Stolz; N. Sauer, *J. Biol. Chem.* **1999**, *274*, 18747–18752.
53. J. Stolz; T. Caspari; A. M. Carr; N. Sauer, *Eukaryot. Cell* **2004**, *3*, 406–412.
54. P. D. Prasad; H. Wang; W. Huang; Y.-J. Fei; F. H. Leibach; L. D. Devoe; V. Ganapathy, *Arch. Biochem. Biophys.* **1999**, *366*, 95–106.
55. P. D. Prasad; H. Wang; R. Kekuda; T. Fujita; Y.-J. Fei; L. D. Devoe; F. H. Leibach; V. Ganapathy, *J. Biol. Chem.* **1998**, *273*, 7501–7506.
56. P. D. Prasad; V. Ganapathy, *Curr. Opin. Clin. Nutr. Metab. Care* **2000**, *3*, 263–266.
57. K. J. Saliba; H. A. Horner; K. Kirk, *J. Biol. Chem.* **1998**, *273*, 10190–10195.
58. K. J. Saliba; K. Kirk, *J. Biol. Chem.* **2001**, *276*, 18115–18121.
59. J. A. Craig; E. E. Snell, *J. Bacteriol.* **1951**, *61*, 283–291.
60. S. Jackowski; C. O. Rock, *J. Bacteriol.* **1984**, *158*, 115–120.
61. D. S. Vallari; S. Jackowski, *J. Bacteriol.* **1988**, *170*, 3961–3966.
62. S. Jackowski; C. O. Rock, *J. Biol. Chem.* **1984**, *259*, 1891–1895.
63. E. Strauss; T. P. Begley, *J. Biol. Chem.* **2002**, *277*, 48205–48209.
64. K. M. Clarke; A. C. Mercer; J. J. La Clair; M. D. Burkart, *J. Am. Chem. Soc.* **2005**, *127*, 11234–11235.
65. J. L. Meier; A. C. Mercer; H. Rivera, Jr.; M. D. Burkart, *J. Am. Chem. Soc.* **2006**, *128*, 12174–12184.
66. K. Yang; Y. Eyobo; L. A. Brand; D. Martynowski; D. Tomchick; E. Strauss; H. Zhang, *J. Bacteriol.* **2006**, *188*, 5532–5540.
67. D. S. Vallari; C. O. Rock, *J. Bacteriol.* **1987**, *169*, 5795–5800.
68. W.-J. Song; S. Jackowski, *J. Biol. Chem.* **1994**, *269*, 27051–27058.
69. D. S. Vallari; S. Jackowski; C. O. Rock, *J. Biol. Chem.* **1987**, *262*, 2468–2471.
70. J. E. Walker; M. Saraste; M. J. Runswick; N. J. Gay, *EMBO J.* **1982**, *1*, 945–951.
71. M. Yun; C.-G. Park; J.-Y. Kim; C. O. Rock; S. Jackowski; H.-W. Park, *J. Biol. Chem.* **2000**, *275*, 28093–28099.
72. R. A. Ivey; Y.-M. Zhang; K. G. Virga; K. Hevener; R. E. Lee; C. O. Rock; S. Jackowski; H.-W. Park, *J. Biol. Chem.* **2004**, *279*, 35622–35629.
73. S. Cheek; K. Ginalski; H. Zhang; N. V. Grishin, *BMC Struct. Biol.* **2005**, *5*, 6.
74. S. Cheek; H. Zhang; N. V. Grishin, *J. Mol. Biol.* **2002**, *320*, 855–881.
75. G. Clifton; S. R. Bryant; C. G. Skinner, *Arch. Biochem. Biophys.* **1970**, *137*, 523–528.
76. Y.-M. Zhang; M. W. Frank; K. G. Virga; R. E. Lee; C. O. Rock; S. Jackowski, *J. Biol. Chem.* **2004**, *279*, 50969–50975.
77. S. Das; P. Kumar; V. Bhor; A. Surolia; M. Vijayan, *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *D62*, 628–638.
78. S. Das; P. Kumar; V. Bhor; A. Surolia; M. Vijayan, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2005**, *F61*, 65–67.
79. T. Kato; Y. Ikenaka, Novel Pantothenate Kinase Gene and Use Thereof. PCT (WO) Patent WO2,006,104,124, 5 October 2006.
80. S. Shimizu; A. Esumi; R. Komaki; H. Yamada, *Appl. Environ. Microbiol.* **1984**, *48*, 1118–1122.
81. R. B. Calder; R. S. B. Williams; G. Ramaswamy; C. O. Rock; E. Campbell; S. E. Unkles; J. R. Kinghorn; S. Jackowski, *J. Biol. Chem.* **1999**, *274*, 2014–2020.
82. T. Kupke; P. Hernandez-Acosta; F. A. Cullianez-Macia, *J. Biol. Chem.* **2003**, *278*, 38229–38237.
83. G. B. Tilton; W. J. Wedemeyer; J. Browse; J. Ohlrogge, *Plant Mol. Biol.* **2006**, *61*, 629–642.
84. K. L. Falk; D. J. Guerra, *Arch. Biochem. Biophys.* **1993**, *301*, 424–430.
85. K. Afshar; P. Gonczy; S. DiNardo; S. A. Wasserman, *Genetics* **2001**, *157*, 1267–1276.
86. Y. Yang; Z. Wu; Y. M. Kuo; B. Zhou, *J. Inherit. Metab. Dis.* **2005**, *28*, 1055–1064.
87. Y.-M. Zhang; C. O. Rock; S. Jackowski, *J. Biol. Chem.* **2005**, *280*, 32594–32601.
88. Y. Li; Y. Chang; L. Zhang; Q. Feng; Z. Liu; Y. Zhang; J. Zuo; Y. Meng; F. Fang, *Mol. Cell. Biochem.* **2005**, *277*, 117–125.
89. B. Zhou; H. K. Westaway; B. Levinson; M. A. Johnson; J. Gitschier; S. J. Hayflick, *Nat. Genet.* **2001**, *28*, 345–349.
90. X. Ni; Y. Ma; H. Cheng; M. Jiang; K. Ying; Y. Xie; Y. Mao, *Int. J. Biochem. Cell Biol.* **2002**, *34*, 109–115.
91. C. O. Rock; M. A. Karim; Y.-M. Zhang; S. Jackowski, *Gene* **2002**, *291*, 35–43.
92. G. Ramaswamy; M. A. Karim; K. G. Murti; S. Jackowski, *J. Lipid Res.* **2004**, *45*, 17–31.
93. R. Leonardi; Y.-M. Zhang; A. Lykidis; C. O. Rock; S. Jackowski, *FEBS Lett.* **2007**, *581*, 4639–4644.
94. C. O. Rock; R. B. Calder; M. A. Karim; S. Jackowski, *J. Biol. Chem.* **2000**, *275*, 1377–1387.
95. K. Hortnagel; H. Prokisch; T. Meitinger, *Hum. Mol. Genet.* **2003**, *12*, 321–327.
96. M. A. Johnson; Y. M. Kuo; S. K. Westaway; S. M. Parker; K. H. L. Ching; J. Gitschier; S. J. Hayflick, *Ann. N. Y. Acad. Sci.* **2004**, *1012*, 282–298.
97. Y.-M. Zhang; C. O. Rock; S. Jackowski, *J. Biol. Chem.* **2006**, *281*, 107–114.
98. R. Leonardi; C. O. Rock; S. Jackowski; Y.-M. Zhang, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1494–1499.



99. A. M. Lehane; R. V. Marchetti; C. Spry; D. A. van Schalkwyk; R. Teng; K. Kirk; K. J. Saliba, *J. Biol. Chem.* **2007**, *282*, 25395–25405.
100. K. G. Virga; Y.-M. Zhang; R. Leonardi; R. A. Ivey; K. Hevener; H.-W. Park; S. Jackowski; C. O. Rock; R. E. Lee, *Bioorg. Med. Chem.* **2006**, *14*, 1007–1020.
101. A. E. Choudhry; T. L. Mandichak; J. P. Broskey; R. W. Egolf; C. Kinsland; T. P. Begley; M. A. Seefeld; T. W. Ku; J. R. Brown; M. Zalacain; K. Ratnam, *Antimicrob. Agents Chemother.* **2003**, *47*, 2051–2055.
102. R. Leonardi; S. Chohnan; Y.-M. Zhang; K. G. Virga; R. E. Lee; C. O. Rock; S. Jackowski, *J. Biol. Chem.* **2005**, *280*, 3314–3322.
103. B. S. Hong; G. Senisterra; W. M. Rabeh; M. Vedadi; R. Leonardi; Y.-M. Zhang; C. O. Rock; S. Jackowski; H.-W. Park, *J. Biol. Chem.* **2007**, *282*, 27984–27993.
104. S. B. delCardayre; K. P. Stock; G. L. Newton; R. C. Fahey; J. E. Davies, *J. Biol. Chem.* **1998**, *273*, 5744–5751.
105. B. S. Hong; M. K. Yun; Y.-M. Zhang; S. Chohnan; C. O. Rock; S. W. White; S. Jackowski; H.-W. Park; R. Leonardi, *Structure* **2006**, *14*, 1251–1261.
106. A. Osterman; R. Overbeek, *Curr. Opin. Chem. Biol.* **2003**, *7*, 238–251.
107. R. R. Yocum; T. A. Patterson, *Microorganisms and Assays for the Identification of Antibiotics Acting on the Pantothenate Kinase Encoded by the coaX Gene*. U.S. Patent 6,830,898, 14 December 2002.
108. L. A. Brand; E. Strauss, *J. Biol. Chem.* **2005**, *280*, 20185–20188.
109. K. Yang; E. Strauss; C. Huerta; H. Zhang, *Biochemistry* **2008**, *47*, 1369–1380.
110. N. I. Nicely; D. Parsonage; C. Paige; G. L. Newton; R. C. Fahey; R. Leonardi; S. Jackowski; T. C. Mallett; A. Claiborne, *Biochemistry* **2007**, *46*, 3234–3245.
111. U. Genschel, *Mol. Biol. Evol.* **2004**, *21*, 1242–1251.
112. Y. Ye; A. Osterman; R. Overbeek; A. Godzik, *Bioinformatics* **2005**, *21* (Suppl. 1), i478–i486.
113. R. Overbeek; T. Begley; R. M. Butler; J. V. Choudhuri; H. Y. Chuang; M. Cohoon; V. de Crecy-Lagard; N. Diaz; T. Disz; R. Edwards; M. Fonstein; E. D. Frank; S. Gerdes; E. M. Glass; A. Goesmann; A. Hanson; D. Iwata-Reuyl; R. Jensen; N. Jamshidi; L. Krause; M. Kubal; N. Larsen; B. Linke; A. C. McHardy; F. Meyer; H. Neuweger; G. Olsen; R. Olson; A. Osterman; V. Portnoy; G. D. Pusch; D. A. Rodionov; C. Ruckert; J. Steiner; R. Stevens; I. Thiele; O. Vassieva; Y. Ye; O. Zagnitko; V. Vonstein, *Nucleic Acids Res.* **2005**, *33*, 5691–5702.
114. M. van Wyk; E. Strauss, *Chem. Commun.* **2007**, 398–400.
115. A. S. Worthington; M. D. Burkart, *Org. Biomol. Chem.* **2006**, *4*, 44–46.
116. M. Daugherty; B. Polanuyer; M. Farrell; M. Scholle; A. Lykidis; V. De Crecy-Lagard; A. Osterman, *J. Biol. Chem.* **2002**, *277*, 21431–21439.
117. Y. Abiko, *J. Biochem. (Tokyo)* **1967**, *61*, 290–299.
118. Y. Abiko; M. Tomikawa; M. Shimizu, *J. Biochem. (Tokyo)* **1968**, *64*, 115–117.
119. Y. Abiko, *Methods Enzymol.* **1970**, *18*, 350–354.
120. E. Strauss, *Thiols, Radicals and Antibiotics: Mechanistic Studies in Coenzyme A Biosynthesis*. Ph.D. Dissertation, Cornell University, Ithaca, NY, 2003.
121. E. Strauss; T. P. Begley, *J. Am. Chem. Soc.* **2001**, *123*, 6449–6450.
122. T. Kupke, *J. Biol. Chem.* **2002**, *277*, 36137–36145.
123. T. Kupke, *Eur. J. Biochem.* **2004**, *271*, 163–172.
124. T. Kupke; W. Schwarz, *J. Biol. Chem.* **2006**, *281*, 5435–5444.
125. J. Yao; J. Patrone; G. Dotson, *Biochemistry* **2009**, *48*, 2799–2806.
126. N. Manoj; E. Strauss; T. P. Begley; S. E. Ealick, *Structure* **2003**, *11*, 927–936.
127. S. Stanitzek; M. A. Augustin; R. Huber; T. Kupke; S. Steinbacher, *Structure* **2004**, *12*, 1977–1988.
128. F. von Delft; A. Lewendon; V. Dhanaraj; T. L. Blundell; C. Abell; A. G. Smith, *Structure* **2001**, *9*, 439–450.
129. R. Zheng; J. S. Blanchard, *Biochemistry* **2001**, *40*, 12904–12912.
130. S. Wang; D. Eisenberg, *Protein Sci.* **2003**, *12*, 1097–1108.
131. S. Wang; D. Eisenberg, *Biochemistry* **2006**, *45*, 1554–1561.
132. E. Strauss; T. P. Begley, *ChemBioChem* **2005**, *6*, 284–286.
133. C. J. Stewart; W. J. Ball, Jr., *Biochemistry* **1966**, *5*, 3883–3886.
134. R. Scandurra; E. Barboni; F. Granata; B. Pensa; M. Costa, *Eur. J. Biochem.* **1974**, *49*, 1–9.
135. R. Scandurra; L. Politi; L. Santoro; V. Consalvi, *FEBS Lett.* **1987**, *212*, 79–82.
136. R. Scandurra; V. Consalvi; L. Politi; C. Gallina, *FEBS Lett.* **1988**, *231*, 192–196.
137. H. Yang; R. H. Abeles, *Biochemistry* **1987**, *26*, 4076–4081.
138. P. D. Van Poelje; E. E. Snell, *Annu. Rev. Biochem.* **1990**, *59*, 29–59.
139. T. Kupke; M. Uebele; D. Schmid; G. Jung; M. Blaesse; S. Steinbacher, *J. Biol. Chem.* **2000**, *275*, 31838–31846.
140. T. Kupke; S. Stevanovic; H. G. Sahl; F. Goetz, *J. Bacteriol.* **1992**, *174*, 5354–5361.
141. F. Majer; D. G. Schmid; K. Altena; G. Bierbaum; T. Kupke, *J. Bacteriol.* **2002**, *184*, 1234–1243.
142. M. Blaesse; T. Kupke; R. Huber; S. Steinbacher, *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *D59*, 1414–1421.
143. H.-G. Sahl; G. Bierbaum, *Annu. Rev. Microbiol.* **1998**, *52*, 41–79.
144. T. Kupke; C. Kempter; V. Gnau; G. Jung; F. Goetz, *J. Biol. Chem.* **1994**, *269*, 5653–5659.
145. T. Kupke; C. Kempter; G. Jung; F. Goetz, *J. Biol. Chem.* **1995**, *270*, 11282–11289.
146. D. G. Schmid; F. Majer; T. Kupke; G. Jung, *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1779–1784.
147. C. Kempter; T. Kupke; D. Kaiser; J. W. Metzger; G. Jung, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2104–2107.
148. M. Blaesse; T. Kupke; R. Huber; S. Steinbacher, *EMBO J.* **2000**, *19*, 6299–6310.
149. T. Kupke, *J. Biol. Chem.* **2001**, *276*, 27597–27604.
150. T. Kupke; P. Hernandez-Acosta; S. Steinbacher; F. A. Cullianez-Macia, *J. Biol. Chem.* **2001**, *276*, 19190–19196.
151. P. Hernandez-Acosta; D. G. Schmid; G. Jung; F. A. Cullianez-Macia; T. Kupke, *J. Biol. Chem.* **2002**, *277*, 20490–20498.
152. S. Steinbacher; P. Hernandez-Acosta; B. Bieseler; M. Blaesse; R. Huber; F. A. Cullianez-Macia; T. Kupke, *J. Mol. Biol.* **2003**, *327*, 193–202.
153. E. Strauss; H. Zhai; L. A. Brand; F. W. McLafferty; T. P. Begley, *Biochemistry* **2004**, *43*, 15520–15533.

154. A. J. Kresge; Q. Meng, *J. Am. Chem. Soc.* **1998**, *120*, 11830–11831.
155. S. Sklenak; Y. Apeloig; Z. Rappoport, *J. Chem. Soc., Perkin Trans. 2* **2000**, 2269–2279.
156. D. J. Aberhart; P. K. Ghoshal; J. A. Cotting; D. J. Russell, *Biochemistry* **1985**, *24*, 7178–7182.
157. E. Strauss; T. P. Begley, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 339–342.
158. D. M. Worrall; P. K. Tubbs, *Biochem. J.* **1983**, *215*, 153–157.
159. D. P. Martin; D. G. Drueckhammer, *Biochem. Biophys. Res. Commun.* **1993**, *192*, 1155–1161.
160. S. Aghajanian; D. M. Worrall, *Biochem. J.* **2002**, *365*, 13–18.
161. A. Zhyvoloup; I. Nemazany; A. Babich; G. Panasyuk; N. Pobigailo; M. Vudmaska; V. Naidenov; O. Kukhareno; S. Palchevskii; L. Savinska; G. Ovcharenko; F. Verdier; T. Valovka; T. Fenton; H. Rebholz; M.-L. Wang; P. Shepherd; G. Matsuka; V. Filonenko; I. T. Gout, *J. Biol. Chem.* **2002**, *277*, 22107–22110.
162. J. Armengaud; B. Fernandez; V. Chaumont; F. Rollin-Genetet; S. Finet; C. Marchetti; H. Myllykallio; C. Vidaud; J.-L. Pellequer; S. Gribaldo; P. Forterre; P. Gans, *J. Biol. Chem.* **2003**, *278*, 31078–31087.
163. M. Nalezkova; A. de Groot; M. Graf; P. Gans; L. Blanchard, *Protein Expr. Purif.* **2005**, *39*, 296–306.
164. A. Geerlof; A. Lewendon; W. V. Shaw, *J. Biol. Chem.* **1999**, *274*, 27105–27111.
165. T. Izard; A. Geerlof, *EMBO J.* **1999**, *18*, 2021–2030.
166. T. Izard, *J. Mol. Biol.* **2002**, *315*, 487–495.
167. T. Izard, *J. Bacteriol.* **2003**, *185*, 4074–4080.
168. H. Takahashi; E. Inagaki; Y. Fujimoto; C. Kuroishi; Y. Nodake; Y. Nakamura; F. Arisaka; K. Yutani; S. Kuramitsu; S. Yokoyama; M. Yamamoto; M. Miyano; T. H. Tahirou, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *D60*, 97–104.
169. V. K. Morris; T. Izard, *Protein Sci.* **2004**, *13*, 2547–2552.
170. J. Y. Kang; H. H. Lee; H. J. Yoon; H. S. Kim; S. W. Suh, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2006**, *F62*, 1131–1133.
171. J. Badger; J. M. Sauder; J. M. Adams; S. Antonysamy; K. Bain; M. G. Bergseid; S. G. Buchanan; M. D. Buchanan; Y. Batiyenko; J. A. Christopher; S. Emtage; A. Eroshkina; I. Feil; E. B. Furlong; K. S. Gajiwala; X. Gao; D. He; J. Hendle; A. Huber; K. Hoda; P. Kearins; C. Kissinger; B. Laubert; H. A. Lewis; J. Lin; K. Loomis; D. Lorimer; G. Louie; M. Maletic; C. D. Marsh; I. Miller; J. Molinari; H. J. Muller-Dieckmann; J. M. Newman; B. W. Noland; B. Pagarigan; F. Park; T. S. Peat; K. W. Post; S. Radojicic; A. Ramos; R. Romero; M. E. Rutter; W. E. Sanderson; K. D. Schwinn; J. Tresser; J. Winhoven; T. A. Wright; L. Wu; J. Xu; T. J. R. Harris, *Proteins* **2005**, *60*, 787–796.
172. P. Bork; L. Holm; E. V. Koonin; C. Sander, *Proteins* **1995**, *22*, 259–266.
173. M. B. Hoagland; G. D. Novelli, *J. Biol. Chem.* **1954**, *207*, 767–773.
174. P. K. Mishra; P. K. Park; D. G. Drueckhammer, *J. Bacteriol.* **2001**, *183*, 2774–2778.
175. N. O'Toole; A. R. G. Barbosa Joao; Y. Li; L.-W. Hung; A. Matte; M. Cygler, *Protein Sci.* **2003**, *12*, 327–336.
176. S. J. Eom; H. J. Ahn; H. W. Kim; S. H. Baek; S. W. Suh, *Acta Crystallogr. D Biol. Crystallogr.* **2003**, *D59*, 561–562.
177. G. Obmolova; A. Teplyakov; N. Bonander; E. Eisenstein; A. J. Howard; G. L. Gilliland, *J. Struct. Biol.* **2001**, *136*, 119–125.
178. A. Seto; K. Murayama; M. Toyama; A. Ebihara; N. Nakagawa; S. Kuramitsu; M. Shirouzu; S. Yokoyama, *Proteins* **2005**, *58*, 235–242.
179. G. E. Schulz; C. W. Muller; K. Diederichs, *J. Mol. Biol.* **1990**, *213*, 627–630.
180. S. Gerdes; R. Edwards; M. Kubal; M. Fonstein; R. Stevens; A. Osterman, *Curr. Opin. Biotechnol.* **2006**, *17*, 448–456.
181. J. I. Glass; N. Assad-Garcia; N. Alperovich; S. Yooshep; M. R. Lewis; M. Maruf; C. A. Hutchison, III; H. O. Smith; J. C. Venter, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 425–430.
182. N. T. Liberati; J. M. Urbach; S. Miyata; D. G. Lee; E. Drenkard; G. Wu; J. Villanueva; T. Wei; F. M. Ausubel, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 2833–2838.
183. C. M. Sasseti; D. H. Boyd; E. J. Rubin, *Mol. Microbiol.* **2003**, *48*, 77–84.
184. A. C. Mercer; M. D. Burkart, *Nat. Prod. Rep.* **2007**, *24*, 750–773.
185. H. Donato; N. I. Krupenko; Y. Tsybovsky; S. A. Krupenko, *J. Biol. Chem.* **2007**, *282*, 34159–34166.
186. P. Venkatasubramanian; L. Daniels; J. P. N. Rosazza, *J. Biol. Chem.* **2007**, *282*, 478–485.
187. D. E. Ehmann; A. M. Gehring; C. T. Walsh, *Biochemistry* **1999**, *38*, 6171–6177.
188. S. Guo; J. K. Bhattacharjee, *Yeast* **2004**, *21*, 1279–1288.
189. R. H. Lambalot; A. M. Gehring; R. S. Flugel; P. Zuber; M. LaCelle; M. A. Marahiel; R. Reid; C. Khosla; C. T. Walsh, *Chem. Biol.* **1996**, *3*, 923–936.
190. C. T. Walsh; A. M. Gehring; P. H. Weinreb; L. E. Quadri; R. S. Flugel, *Curr. Opin. Chem. Biol.* **1997**, *1*, 309–315.
191. K. Reuter; M. R. Mofid; M. A. Marahiel; R. Ficner, *EMBO J.* **1999**, *18*, 6823–6831.
192. K. D. Parris; L. Lin; A. Tam; R. Mathew; J. Hixon; M. Stahl; C. C. Fritz; J. Sehra; W. S. Somers, *Structure* **2000**, *8*, 883–895.
193. N. Y. Chirgadze; S. L. Briggs; K. A. McAllister; A. S. Fischl; G. Zhao, *EMBO J.* **2000**, *19*, 5281–5287.
194. R. Finking; J. Solsbacher; D. Konz; M. Schobert; A. Schaefer; D. Jahn; M. A. Marahiel, *J. Biol. Chem.* **2002**, *277*, 50293–50302.
195. R. Finking; M. R. Mofid; M. A. Marahiel, *Biochemistry* **2004**, *43*, 8946–8956.
196. M. R. Mofid; R. Finking; L. O. Essen; M. A. Marahiel, *Biochemistry* **2004**, *43*, 4128–4136.
197. A. Koglin; M. R. Mofid; F. Loehr; B. Schaefer; V. V. Rogov; M.-M. Blum; T. Mittag; M. A. Marahiel; F. Bernhard; V. Doetsch, *Science* **2006**, *312*, 273–276.
198. J. R. Lai; A. Koglin; C. T. Walsh, *Biochemistry* **2006**, *45*, 14869–14879.
199. L. E. N. Quadri; P. H. Weinreb; M. Lei; M. M. Nakano; P. Zuber; C. T. Walsh, *Biochemistry* **1998**, *37*, 1585–1595.
200. G. Bunkoczi; S. Pasta; A. Joshi; X. Wu; K. L. Kavanagh; S. Smith; U. Oppermann, *Chem. Biol.* **2007**, *14*, 1243–1253.
201. S. Chopra; S. Kumar Singh; S. Prasad Sati; A. Ranganathan; A. Sharma, *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *D58*, 179–181.
202. D. L. Daubaras; E. M. Wilson; T. Black; C. Strickland; B. M. Beyer; P. Orth, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *D60*, 773–774.
203. S. Hoenke; M. Schmid; P. Dimroth, *Biochemistry* **2000**, *39*, 13233–13240.
204. S. Hoenke; M. R. Wild; P. Dimroth, *Biochemistry* **2000**, *39*, 13223–13232.
205. K. Schneider; P. Dimroth; M. Bott, *FEBS Lett.* **2000**, *483*, 165–168.

206. K. Schneider; P. Dimroth; M. Bott, *Biochemistry* **2000**, 39, 9438–9450.
207. K. Schneider; C. N. Kastner; M. Meyer; M. Wessel; P. Dimroth; M. Bott, *J. Bacteriol.* **2002**, 184, 2439–2446.
208. E. G. Trams; H. A. Fales; A. E. Gal, *Biochem. Biophys. Res. Commun.* **1968**, 31, 973–976.
209. E. G. Trams; W. L. Stahl; J. Robinson, *Biochim. Biophys. Acta* **1968**, 163, 472–482.
210. J. E. Franklin; E. G. Trams, *Biochim. Biophys. Acta* **1971**, 230, 105–116.
211. S. Skrede, *Eur. J. Biochem.* **1973**, 38, 401–407.
212. J. L. Cartwright; L. Gasmi; D. G. Spiller; A. G. McLennan, *J. Biol. Chem.* **2000**, 275, 32925–32930.
213. L. Gasmi; A. G. McLennan, *Biochem. J.* **2001**, 357, 33–38.
214. R. Ofman; D. Speijer; R. Leen; R. J. A. Wanders, *Biochem. J.* **2006**, 393, 537–543.
215. S. R. AbdelRaheim; A. G. McLennan, *BMC Biochem.* **2002**, 3, 5.
216. A. S. Mildvan; Z. Xia; H. F. Azurmendi; V. Saraswat; P. M. Legler; M. A. Massiah; S. B. Gabelli; M. A. Bianchet; L. W. Kang; L. M. Amzel, *Arch. Biochem. Biophys.* **2005**, 433, 129–143.
217. L.-W. Kang; S. B. Gabelli; M. A. Bianchet; W. L. Xu; M. J. Bessman; L. M. Amzel, *J. Bacteriol.* **2003**, 185, 4110–4118.
218. S. Dupre; D. Cavallini, *Methods Enzymol.* **1979**, 62, 262–267.
219. C. T. Wittwer; B. W. Wyse; R. G. Hansen, *Methods Enzymol.* **1986**, 122, 36–43.
220. J. A. Calvino; R. Barcia, *J. Food Biochem.* **2002**, 26, 103–118.
221. A. Gonthier; V. Fayol; J. Viollet; D. J. Hartmann, *Food Chem.* **1998**, 63, 287–294.
222. M. Rychlik, *Analyst* **2003**, 128, 832–837.
223. B. Maras; D. Barra; S. Dupre; G. Pitari, *FEBS Lett.* **1999**, 461, 149–152.
224. G. Pitari; F. Malergue; F. Martin; J. M. Philippe; M. T. Massucci; C. Chabret; B. Maras; S. Dupre; P. Naquet; F. Galland, *FEBS Lett.* **2000**, 483, 149–154.
225. F. Martin; F. Malergue; G. Pitari; J. M. Philippe; S. Philips; C. Chabret; S. Granjeaud; M. G. Mattei; A. J. Mungall; P. Naquet; F. Galland, *Immunogenetics* **2001**, 53, 296–306.
226. R. K. Airas, *Methods Enzymol.* **1979**, 62, 267–275.
227. R. K. Airas, *Biochem. J.* **1988**, 250, 447–451.
228. R. K. Airas, *Biochemistry* **1978**, 17, 4932–4938.
229. R. C. Fahey, *Annu. Rev. Microbiol.* **2001**, 55, 333–356.
230. B. Setlow; P. Setlow, *J. Bacteriol.* **1977**, 132, 444–452.
231. C. S. Hummel; K. M. Lancaster; E. J. Crane, III, *FEMS Microbiol. Lett.* **2005**, 252, 229–234.
232. G. L. Newton; K. Arnold; M. S. Price; C. Sherrill; S. B. Delcardayre; Y. Aharonowitz; G. Cohen; J. Davies; R. C. Fahey; C. Davis, *J. Bacteriol.* **1996**, 178, 1990–1995.
233. S. B. delCardayre; J. E. Davies, *J. Biol. Chem.* **1998**, 273, 5752–5757.
234. R. D. Swerdlow; P. Setlow, *J. Bacteriol.* **1983**, 153, 475–484.
235. J. Luba; V. Charrier; A. Claiborne, *Biochemistry* **1999**, 38, 2725–2737.
236. D. R. Harris; D. E. Ward; J. M. Feasel; K. M. Lancaster; R. D. Murphy; T. C. Mallet; E. J. Crane, III, *FEBS J.* **2005**, 272, 1189–1200.
237. J. A. Boylan; C. S. Hummel; S. Benoit; J. Garcia-Lara; J. Treglown-Downey; E. J. Crane, III; F. C. Gherardini, *Mol. Microbiol.* **2006**, 59, 475–486.
238. A. Argyrou; J. S. Blanchard, *Prog. Nucleic Acid Res. Mol. Biol.* **2004**, 78, 89–142.
239. A. Claiborne; J. I. Yeh; T. C. Mallet; J. Luba; E. J. Crane, III; V. Charrier; D. Parsonage, *Biochemistry* **1999**, 38, 15407–15416.
240. T. C. Mallet; J. R. Wallen; P. A. Karplus; H. Sakai; T. Tsukihara; A. Claiborne, *Biochemistry* **2006**, 45, 11278–11289.
241. G. J. Schut; S. L. Bridger; M. W. W. Adams, *J. Bacteriol.* **2007**, 189, 4431–4441.
242. W. A. Bridger, Succinyl-CoA Synthetase. In *The Enzymes*, 3rd ed.; P. D. Boyer, Ed.; Academic Press: New York, 1974; Vol. 10, pp 581–606.
243. K. M. Knights, *Curr. Med. Chem. Immunol. Endocr. Metab. Agents* **2003**, 3, 235–244.
244. J. C. Londesborough; L. T. Webster, Jr., Fatty Acyl-CoA Synthetases. In *The Enzymes*, 3rd ed.; P. D. Boyer, Ed.; Academic Press: New York, 1974; Vol. 10, pp 469–488.
245. L. B. Sanchez; M. Y. Galperin; M. Muller, *J. Biol. Chem.* **2000**, 275, 5794–5803.
246. V. J. Starai; J. C. Escalante-Semerena, *Cell Mol. Life Sci.* **2004**, 61, 2020–2030.
247. P. A. Watkins, *Prog. Lipid Res.* **1997**, 36, 55–83.
248. P. A. Watkins, *J. Biol. Chem.* **2008**, 283, 1773–1777.
249. K. M. Knights; C. J. Drogemuller, *Curr. Drug Metab.* **2000**, 1, 49–66.
250. U. Linne; A. Schaefer; M. T. Stubbs; M. A. Marahiel, *FEBS Lett.* **2007**, 581, 905–910.
251. A. R. Horswill; J. C. Escalante-Semerena, *Biochemistry* **2002**, 41, 2379–2387.
252. K. Schuehle; J. Gescher; U. Feil; M. Paul; M. Jahn; H. Schaegger; G. Fuchs, *J. Bacteriol.* **2003**, 185, 4920–4929.
253. D. Cukovic; J. Ehling; J. A. VanZiffle; C. J. Douglas, *Biol. Chem.* **2001**, 382, 645–654.
254. P. A. Watkins; D. Maignel; Z. Jia; J. Pevsner, *J. Lipid Res.* **2007**, 48, 2736–2750.
255. E. Conti; N. P. Franks; P. Brick, *Structure* **1996**, 4, 287–298.
256. A. Andreeva; D. Howorth; J.-M. Chandonia; S. E. Brenner; T. J. P. Hubbard; C. Chothia; A. G. Murzin, *Nucleic Acids Res.* **2008**, 36, D419–D425.
257. A. G. Murzin; S. E. Brenner; T. Hubbard; C. Chothia, *J. Mol. Biol.* **1995**, 247, 536–540.
258. M. E. Fraser; M. N. G. James; W. A. Bridger; W. T. Wolodko, *J. Mol. Biol.* **2000**, 299, 1325–1339.
259. M. E. Fraser; M. N. G. James; W. A. Bridger; W. T. Wolodko, *J. Mol. Biol.* **1999**, 285, 1633–1653.
260. M. Y. Galperin; E. V. Koonin, *Protein Sci.* **1997**, 6, 2639–2643.
261. M. Musfeldt; P. Schonheit, *J. Bacteriol.* **2002**, 184, 636–644.
262. C. Ingram-Smith; S. R. Martin; K. S. Smith, *Trends Microbiol.* **2006**, 14, 249–253.
263. A. J. Wolfe, *Microbiol. Mol. Biol. Rev.* **2005**, 69, 12–50.
264. Y. Jiang; C. H. Chan; J. E. Cronan, *Biochemistry* **2006**, 45, 10008–10019.
265. N. Y. Kedishvili; G. W. Goodwin; K. M. Popov; R. A. Harris, *Methods Enzymol.* **2000**, 324, 207–218.
266. C. Stines-Chaumeil; F. Talfournier; G. Branlant, *Biochem. J.* **2006**, 395, 107–115.

267. L. B. Sanchez, *Arch. Biochem. Biophys.* **1998**, *354*, 57–64.
268. M. Chen; E. Li; S. L. Stanley, *Mol. Biochem. Parasitol.* **2004**, *137*, 201–205.
269. C. Furdul; S. W. Ragsdale, *Biochemistry* **2002**, *41*, 9921–9937.
270. S. O. Mansoorabadi; J. Seravalli; C. Furdul; V. Krymov; G. J. Gerfen; T. P. Begley; J. Melnick; S. W. Ragsdale; G. H. Reed, *Biochemistry* **2006**, *45*, 7122–7131.
271. S. W. Ragsdale, *Chem. Rev.* **2003**, *103*, 2333–2346.
272. P. A. Frey, *Annu. Rev. Biochem.* **2001**, *70*, 121–148.
273. S. W. Ragsdale, *Chem. Rev.* **2006**, *106*, 3317–3337.
274. J. Seravalli; Y. Xiao; W. Gu; S. P. Cramer; W. E. Antholine; V. Krymov; G. J. Gerfen; S. W. Ragsdale, *Biochemistry* **2004**, *43*, 3944–3955.
275. S. Ragsdale, *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 165–195.
276. S. W. Ragsdale, *J. Inorg. Biochem.* **2007**, *101*, 1657–1666.
277. E. L. Hegg, *Acc. Chem. Res.* **2004**, *37*, 775–783.
278. M. W. Vetting; L. P. S. de Carvalho; M. Yu; S. S. Hegde; S. Magnet; S. L. Roderick; J. S. Blanchard, *Arch. Biochem. Biophys.* **2005**, *433*, 212–226.
279. R. N. Dutnall; S. T. Tafrov; R. Stemglanz; V. Ramakrishnan, *Cell* **1998**, *94*, 427–438.
280. A. F. Neuwald; D. Landsman, *Trends Biochem. Sci.* **1997**, *22*, 154–155.
281. C. M. Johnson; S. L. Roderick; P. F. Cook, *Arch. Biochem. Biophys.* **2005**, *433*, 85–95.
282. A.-R. Kim; R. J. Rylett; B. H. Shilton, *Biochemistry* **2006**, *45*, 14621–14631.
283. R. R. Ramsay; J. H. Naismith, *Trends Biochem. Sci.* **2003**, *28*, 343–346.
284. I. A. Mirza; I. Nazi; M. Korczynska; G. D. Wright; A. M. Berghuis, *Biochemistry* **2005**, *44*, 15768–15773.
285. J. M. Elkins; N. J. Kershaw; C. J. Schofield, *Biochem. J.* **2005**, *385*, 565–573.
286. S. Y. Roth; J. M. Denu; C. D. Allis, *Annu. Rev. Biochem.* **2001**, *70*, 81–120.
287. M. R. Parthun, *Oncogene* **2007**, *26*, 5319–5328.
288. S. C. Hodawadekar; R. Marmorstein, *Oncogene* **2007**, *26*, 5528–5540.
289. C. E. Berndsen; B. N. Albaugh; S. Tan; J. M. Denu, *Biochemistry* **2007**, *46*, 623–629.
290. X. Liu; L. Wang; K. Zhao; P. R. Thompson; Y. Hwang; R. Marmorstein; P. A. Cole, *Nature* **2008**, *451*, 846–850.
291. N. A. Schnarr; C. Khosla, *Combinatorial Biosynthesis of Polyketides and Nonribosomal Peptides*. In *Chemical Biology*; S. L. Schreiber, T. M. Kapoor, G. Wess, Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2007; Vol. 2, pp 519–536.
292. A. T. Keatinge-Clay; A. A. Shelat; D. F. Savage; S.-C. Tsai; L. J. W. Miercke; J. D. O’Connell; C. Khosla; R. M. Stroud, *Structure* **2003**, *11*, 147–154.
293. Z. Li; Y. Huang; J. Ge; H. Fan; X. Zhou; S. Li; M. Bartlam; H. Wang; Z. Rao, *J. Mol. Biol.* **2007**, *371*, 1075–1083.
294. L. Serre; E. C. Verbree; Z. Dauter; A. R. Stuitje; Z. S. Derewenda, *J. Biol. Chem.* **1995**, *270*, 12961–12964.
295. A. T. Koppisch; C. Khosla, *Biochemistry* **2003**, *42*, 11057–11064.
296. J. Heider, *FEBS Lett.* **2001**, *509*, 345–349.
297. S. Ricagno; S. Jonsson; N. Richards; Y. Lindqvist, *EMBO J.* **2003**, *22*, 3210–3219.
298. E. S. Rangarajan; Y. Li; E. Ajamian; P. Iannuzzi; S. D. Kernaghan; M. E. Fraser; M. Cygler; A. Matte, *J. Biol. Chem.* **2005**, *280*, 42919–42928.
299. A. Whitty; C. A. Fierke; W. P. Jencks, *Biochemistry* **1995**, *34*, 11678–11689.
300. S. D. Tamman; J.-C. Rochet; M. E. Fraser, *Biochemistry* **2007**, *46*, 10852–10863.
301. P. Dimroth; H. Eggerer, *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3458–3462.
302. S. Friedmann; B. E. Alber; G. Fuchs, *J. Bacteriol.* **2006**, *188*, 6460–6468.
303. A. Gruez; V. Roig-Zamboni; C. Valencia; V. Campanacci; C. Cambillau, *J. Biol. Chem.* **2003**, *278*, 34582–34586.
304. E. S. Rangarajan; Y. Li; P. Iannuzzi; M. Cygler; A. Matte, *Biochemistry* **2005**, *44*, 5728–5738.
305. C. L. Berthold; C. G. Toyota; N. G. J. Richards; Y. Lindqvist, *J. Biol. Chem.* **2008**, *283*, 6519–6529.
306. S. Jonsson; S. Ricagno; Y. Lindqvist; N. G. J. Richards, *J. Biol. Chem.* **2004**, *279*, 36003–36012.
307. F. Kopp; M. A. Marahiel, *Nat. Prod. Rep.* **2007**, *24*, 735–749.
308. M. C. Hunt; S. E. H. Alexson, *Prog. Lipid Res.* **2002**, *41*, 99–130.
309. D. M. Lawson; U. Derewenda; L. Serre; S. Ferri; R. Szttnner; Y. Wei; E. A. Meighen; Z. S. Derewenda, *Biochemistry* **1994**, *33*, 9382–9388.
310. Y.-C. Lo; S.-C. Lin; J.-F. Shaw; Y.-C. Liaw, *J. Mol. Biol.* **2003**, *330*, 539–551.
311. C. W. I. V. Pemble; L. C. Johnson; S. J. Kridel; W. T. Lowther, *Nat. Struct. Mol. Biol.* **2007**, *14*, 704–709.
312. C. C. Akoh; G.-C. Lee; Y.-C. Liaw; T.-H. Huang; J.-F. Shaw, *Prog. Lipid Res.* **2004**, *43*, 534–552.
313. J. Li; U. Derewenda; Z. Dauter; S. Smith; Z. S. Derewenda, *Nat. Struct. Biol.* **2000**, *7*, 555–559.
314. D. Schwarzer; H. D. Mootz; U. Linne; M. A. Marahiel, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14083–14088.
315. Z. Cheng; F. Song; X. Shan; Z. Wei; Y. Wang; D. Dunaway-Mariano; W. Gong, *Biochem. Biophys. Res. Commun.* **2006**, *349*, 172–177.
316. M. A. Willis; Z. Zhuang; F. Song; A. Howard; D. Dunaway-Mariano; O. Herzberg, *Biochemistry* **2008**, *47*, 2797–2805.
317. Z. Zhuang; F. Song; H. Zhao; L. Li; J. Cao; E. Eisenstein; O. Herzberg; D. Dunaway-Mariano, *Biochemistry* **2008**, *47*, 2789–2796.
318. J. B. Thoden; H. M. Holden; Z. Zhuang; D. Dunaway-Mariano, *J. Biol. Chem.* **2002**, *277*, 27468–27476.
319. J. B. Thoden; Z. Zhuang; D. Dunaway-Mariano; H. M. Holden, *J. Biol. Chem.* **2003**, *278*, 43709–43716.
320. F. Wang; R. Langley; G. Gulten; L. Wang; J. C. Sacchettini, *Chem. Biol.* **2007**, *14*, 543–551.
321. A. M. Haapalainen; G. Merilainen; R. K. Wierenga, *Trends Biochem. Sci.* **2006**, *31*, 64–71.
322. Y. Modis; R. K. Wierenga, *J. Mol. Biol.* **2000**, *297*, 1171–1182.
323. R. J. Heath; C. O. Rock, *Nat. Prod. Rep.* **2002**, *19*, 581–596.
324. F. Pojer; J.-L. Ferrer; S. B. Richard; D. A. Nagegowda; M.-L. Chye; T. J. Noel, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11491–11496.



325. M. J. Theisen; I. Misra; D. Saadat; N. Campobasso; H. M. Miziorko; D. H. T. Harrison, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16442–16447.
326. N. Campobasso; M. Patel; I. E. Wilding; H. Kallender; M. Rosenberg; M. N. Gwynn, *J. Biol. Chem.* **2004**, *279*, 44883–44888.
327. C. V. Smith; C.-C. Huang; A. Micczak; D. G. Russell; J. C. Sacchettini; K. Hoener zu Bentrup, *J. Biol. Chem.* **2003**, *278*, 1735–1743.
328. D. M. Anstrom; S. J. Remington, *Protein Sci.* **2006**, *15*, 2002–2007.
329. L. P. S. De Carvalho; J. S. Blanchard, *Arch. Biochem. Biophys.* **2006**, *451*, 141–148.
330. L. P. S. De Carvalho; J. S. Blanchard, *Biochemistry* **2006**, *45*, 8988–8999.
331. L. C. Kurz; G. Drysdale; M. Riley; M. A. Tomar; J. Chen; R. J. M. Russell; M. J. Danson, *Biochemistry* **2000**, *39*, 2283–2296.
332. J. Qian; A. H. West; P. F. Cook, *Biochemistry* **2006**, *45*, 12136–12143.
333. J. A. Francois; C. M. Starks; S. Sivanuntakorn; H. Jiang; A. E. Ransome; J.-W. Nam; C. Z. Constantine; T. J. Kappock, *Biochemistry* **2006**, *45*, 13487–13499.
334. R. J. M. Russell; J. M. C. Ferguson; D. W. Hough; M. J. Danson; G. L. Taylor, *Biochemistry* **1997**, *36*, 9983–9994.
335. D. M. Anstrom; K. Kallio; S. J. Remington, *Protein Sci.* **2003**, *12*, 1822–1832.
336. F. Forouhar; M. Hussain; R. Farid; J. Benach; M. Abashidze; W. C. Edstrom; S. M. Vorobiev; R. Xiao; T. B. Acton; Z. Fu; J.-J. P. Kim; H. M. Miziorko; G. T. Montelione; J. F. Hunt, *J. Biol. Chem.* **2006**, *281*, 7533–7545.
337. Z. Fu; J. A. Runquist; F. Forouhar; M. Hussain; J. F. Hunt; H. M. Miziorko; J.-J. P. Kim, *J. Biol. Chem.* **2006**, *281*, 7526–7532.
338. M. Bott, *Arch. Microbiol.* **1997**, *167*, 78–88.
339. C. W. Goulding; P. M. Bowers; B. Segelke; T. Lekin; C.-Y. Kim; T. C. Terwilliger; D. Eisenberg, *J. Mol. Biol.* **2007**, *365*, 275–283.
340. L. B. Spector, Citrate Cleavage and Related Enzymes. In *The Enzymes*, 3rd ed.; P. D. Boyer, Ed.; Academic Press: New York, 1972; Vol. 7, pp 357–389.
341. T. N. C. Wells, *Eur. J. Biochem.* **1991**, *199*, 163–168.
342. W. Kim; F. R. Tabita, *J. Bacteriol.* **2006**, *188*, 6544–6552.
343. T. Kanao; T. Fukui; H. Atomi; T. Imanaka, *Eur. J. Biochem.* **2002**, *269*, 3409–3416.
344. D. P. Martin; D. G. Drueckhammer, *J. Am. Chem. Soc.* **1992**, *114*, 7287–7288.
345. M. Dai; Y. Feng; P. J. Tonge, *J. Am. Chem. Soc.* **2001**, *123*, 506–507.
346. I. Nazi; K. P. Koteva; G. D. Wright, *Anal. Biochem.* **2004**, *324*, 100–105.
347. A. L. Mandel; J. J. La Clair; M. D. Burkart, *Org. Lett.* **2004**, *6*, 4801–4803.
348. M. v. Wyk; E. Strauss, *Org. Biomol. Chem.* **2008**, *6*, 4348–4355.
349. M. Yu; M. L. B. Magalhaes; P. F. Cook; J. S. Blanchard, *Biochemistry* **2006**, *45*, 14788–14794.
350. M. Yu; L. P. Sorio de Carvalho; G. Sun; J. S. Blanchard, *J. Am. Chem. Soc.* **2006**, *128*, 15356–15357.
351. Y. Hwang; P. R. Thompson; L. Wang; L. Jiang; N. L. Kelleher; P. A. Cole, *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 7621–7624.
352. M. K. Pasha; P. Selvakumar; L. Ashakumary; M. Qureshi; F. S. Guziec, Jr.; J. R. Dimmock; R. L. Felsted; C. J. Glover; R. K. Sharma, *Int. J. Mol. Med.* **2004**, *13*, 557–563.
353. C. M. Kim; P. A. Cole, *J. Med. Chem.* **2001**, *44*, 2479–2485.
354. W. Zheng; P. A. Cole, *Bioorg. Chem.* **2003**, *31*, 398–411.
355. Y. Hwang; S. Ganguly; A. K. Ho; D. C. Klein; P. A. Cole, *Bioorg. Med. Chem.* **2007**, *15*, 2147–2155.
356. O. D. Lau; T. K. Kundu; R. E. Soccio; S. Ait-Si-Ali; E. M. Khalil; A. Vassilev; A. P. Wolffe; Y. Nakatani; R. G. Roeder; P. A. Cole, *Mol. Cell* **2000**, *5*, 589–595.
357. M. Cebrat; C. M. Kim; P. R. Thompson; M. Daugherty; P. A. Cole, *Bioorg. Med. Chem.* **2003**, *11*, 3307–3313.
358. V. Sagar; W. Zheng; P. R. Thompson; P. A. Cole, *Bioorg. Med. Chem.* **2004**, *12*, 3383–3390.
359. Y. Zheng; K. Balasubramanyam; M. Cebrat; D. Buck; F. Guidez; A. Zelent; R. M. Alani; P. A. Cole, *J. Am. Chem. Soc.* **2005**, *127*, 17182–17183.
360. S. S. Hegde; J. Chandler; M. W. Vetting; M. Yu; J. S. Blanchard, *Biochemistry* **2007**, *46*, 7187–7195.
361. K. L. Leverit; G. L. Waldrop, *Biochem. Biophys. Res. Commun.* **2002**, *291*, 1213–1217.
362. F. Gao; X. Yan; O. M. Baettig; A. M. Berghuis; K. Auclair, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6859–6862.
363. F. Gao; X. Yan; T. Shakya; O. M. Baettig; S. Ait-Mohand-Brunet; A. M. Berghuis; G. D. Wright; K. Auclair, *J. Med. Chem.* **2006**, *49*, 5273–5281.
364. M. L. B. Magalhaes; M. W. Vetting; F. Gao; L. Freiburger; K. Auclair; J. S. Blanchard, *Biochemistry* **2008**, *47*, 579–584.
365. X. Y. Feng Gao; Karine Auclair, *Chem. Eur. J.* **2009**, *15*, 2064–2070.
366. D. Spitteller; C. L. Waterman; J. B. Spencer, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 7079–7082.
367. J. Seravalli; W. Gu; A. Tam; E. Strauss; T. P. Begley; S. P. Cramer; S. W. Ragsdale, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3689–3694.
368. A. S. Worthington; H. Rivera; J. W. Torpey; M. D. Alexander; M. D. Burkart, *ACS Chem. Biol.* **2006**, *1*, 687–691.
369. A. M. Gehring; R. H. Lambalot; K. W. Vogel; D. G. Drueckhammer; C. T. Walsh, *Chem. Biol.* **1997**, *4*, 17–24.
370. J. J. La Clair; T. L. Foley; T. R. Schegg; C. M. Regan; M. D. Burkart, *Chem. Biol.* **2004**, *11*, 195–201.
371. J. Yin; F. Liu; X. Li; C. T. Walsh, *J. Am. Chem. Soc.* **2004**, *126*, 7754–7755.
372. N. George; H. Pick; H. Vogel; N. Johnsson; K. Johnsson, *J. Am. Chem. Soc.* **2004**, *126*, 8896–8897.
373. I. Sielaff; A. Arnold; G. Godin; S. Tugulu; H.-A. Klok; K. Johnsson, *ChemBioChem* **2006**, *7*, 194–202.
374. J. Yin; P. D. Straight; S. M. McLoughlin; Z. Zhou; A. J. Lin; D. E. Golan; N. L. Kelleher; R. Kolter; C. T. Walsh, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15815–15820.
375. Z. Zhou; P. Cironi; A. J. Lin; Y. Xu; S. Hrvatin; D. E. Golan; P. A. Silver; C. T. Walsh; J. Yin, *ACS Chem. Biol.* **2007**, *2*, 337–346.
376. W. P. Heal; S. R. Wickramasinghe; P. W. Bowyer; A. A. Holder; D. F. Smith; R. J. Leatherbarrow; E. W. Tate, *Chem. Commun.* **2008**, 480–482.
377. Y.-M. Zhang; S. W. White; C. O. Rock, *J. Biol. Chem.* **2006**, *281*, 17541–17544.
378. Y.-M. Zhang; S. Chohnan; K. G. Virga; R. D. Stevens; O. R. Ilkayeva; B. R. Wenner; J. R. Bain; C. B. Newgard; R. E. Lee; C. O. Rock; S. Jackowski, *Chem. Biol.* **2007**, *14*, 291–302.
379. L. Zhao; N. M. Allanson; S. P. Thomson; J. K. F. Maclean; J. J. Barker; W. U. Primrose; P. D. Tyler; A. Lewendon, *Eur. J. Med. Chem.* **2003**, *38*, 345–349.
380. J. Kim; D. J. Darley; W. Buckel; A. J. Pierik, *Nature* **2008**, *452*, 239–242.

**Biographical Sketch**

Erick Strauss was born and bred in Pretoria, South Africa, where he obtained his Bachelors in Science in Chemistry in 1997. He subsequently moved to the United States to pursue his graduate studies at Cornell University, where he worked on the biosynthesis of coenzyme A with Professor Tadhg Begley. He received a Ph.D. in chemistry and chemical biology in 2003. He returned to South Africa in the same year to accept an offer from Stellenbosch University to establish his own research group, with a focus on chemical biology, in the Department of Chemistry and Polymer Science. In 2008 he moved to the Department of Biochemistry, Stellenbosch University, where currently he is an associate professor. Erick is a recipient of the President's Award from the South African National Research Foundation, and the Rector's Award for Excellence in Teaching from Stellenbosch University. He continues to pursue his research interests in the mechanistic aspects of CoA biosynthesis, as well as the applications of CoA analogues and precursors in inhibitor and drug design – with special emphasis on compounds that could be applied to the treatment of the so-called 'neglected diseases,' such as tuberculosis, malaria, and human sleeping sickness.



## 7.12 Menaquinone/Ubiquinone Biosynthesis and Enzymology

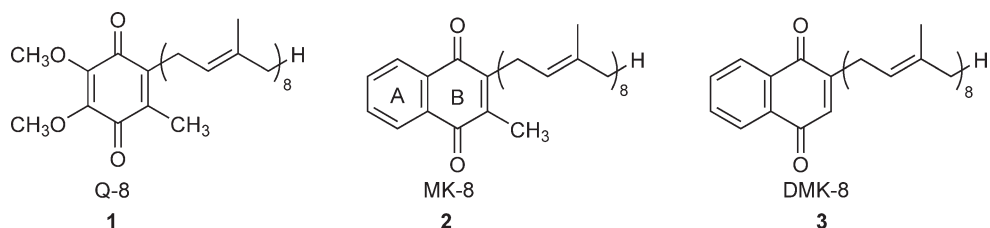
R. Meganathan, Northern Illinois University, DeKalb, IL, USA

© 2010 Elsevier Ltd. All rights reserved.

7.12.1	Introduction	411
7.12.2	Menaquinone Biosynthesis	414
7.12.2.1	o-Succinylbenzoate Pathway	414
7.12.2.1.1	Formation of isochorismate (compound (4) → (5))	415
7.12.2.1.2	Formation of succinic semialdehyde-thiamine pyrophosphate anion and Michael addition to isochorismate (compound (5) + (13) → (14))	418
7.12.2.1.3	The aromatization of SHCHC → OSB (compound (15) → (7))	419
7.12.2.1.4	Cyclization of OSB to DH2NA (compound (7) → (18))	420
7.12.2.1.5	Prenylation of DH2NA to DMK (compound (18) → (3))	421
7.12.2.1.6	Methylation of DMK to MK (compound (3) → (2))	421
7.12.2.2	Non-o-Succinylbenzoate or Futasoline Pathway	422
7.12.3	Phylloquinone Biosynthesis	425
7.12.4	Ubiquinone Biosynthesis	426
7.12.4.1	The Conversion of Chorismate to 4-HB by the CPL (Compound (4) → (11))	429
7.12.4.2	Tyrosine-4-Hydroxyphenylpyruvate Pathway (Compound (12) → (10))	430
7.12.4.3	Chorismate-Prephenate-4-Hydroxyphenylpyruvate Pathway (Compound (4) → (9) → (10))	431
7.12.4.4	Prenylation of 4-Hydroxybenzoate (Compound (11) → (49))	431
7.12.4.5	Formation of 2-Octaprenylphenol (Compound (49) → (50))	432
7.12.4.6	Hydroxylation and Methylation Reactions	433
7.12.4.6.1	Hydroxylation reactions	433
7.12.4.6.2	Methylation reactions	434
7.12.5	Organization of Q Biosynthetic Enzymes into a Complex	436
7.12.6	Comparison of Q Biosynthesis in Yeast and <i>Escherichia coli</i>	436
7.12.7	Biosynthesis of Isoprenoid Side Chain of MK and Q	439
References		440

### 7.12.1 Introduction

Bacteria contain isoprenoid quinones of the benzene and naphthalene series. According to the International Union of Pure and Applied Chemistry-International Union of Biochemistry (IUPAC-IUB) recommendations,<sup>1</sup> the benzoquinones are termed ubiquinones (Q-*n*) and the naphthoquinones are termed either menaquinones (MK-*n*) or demethylmenaquinones (DMK-*n*). The *n* refers to a side chain composed of varying numbers of five carbon isoprenoid residues, and the number of such units is indicated in the abbreviation. Thus, Q-8 (1) (Figure 1) and MK-8 (2) (Figure 1) indicate Q and MK with side chains of 40 carbon atoms, respectively. MK lacking a methyl group also occur and are described as demethylmenaquinones (DMK); again, the number of isoprenoid residues is indicated as just described. DMK-8 is a DMK with a 40-carbon side chain (3) (Figure 1). The major quinones present in *Escherichia coli* are Q-8, MK-8, and DMK-8; small amounts of Q-1–Q-7, Q-9, MK-6, MK-7, MK-9, and DMK-7 may also be present. Furthermore, the length of the isoprenoid side chain is species-specific; for example, yeast has Q-6, whereas humans and *Gluconobacter suboxydans* have Q-10.<sup>2</sup> It should be pointed out that while MK is considered a vitamin (vitamin K<sub>2</sub>), Q is not, due to the fact that vitamin K is an essential nutrient (cannot be synthesized) in mammals, whereas Q is not an essential nutrient since it can be synthesized from the aromatic amino acid tyrosine.



**Figure 1** Structures of major isoprenoid quinones found in *E. coli*. (1) Q-8, ubiquinone; (2) MK-8, menaquinone; (3) DMK-8, demethylmenaquinone. In the structure of MK, the A ring and B ring of the naphthoquinone are shown.

Most aerobic Gram-negative bacteria contain Q as the sole quinone, whereas the aerobic Gram-positive bacteria contain MK as the sole quinone. In contrast, most of the anaerobic bacteria, irrespective of whether they are Gram-positive or Gram-negative, contain MK as their sole quinone. However, Gram-negative facultative anaerobes including *E. coli* contain both Q and MK, and hence they are able to grow under both aerobic and anaerobic conditions, that is, *E. coli* can be grown aerobically with oxygen as an electron acceptor or anaerobically by fermentation.

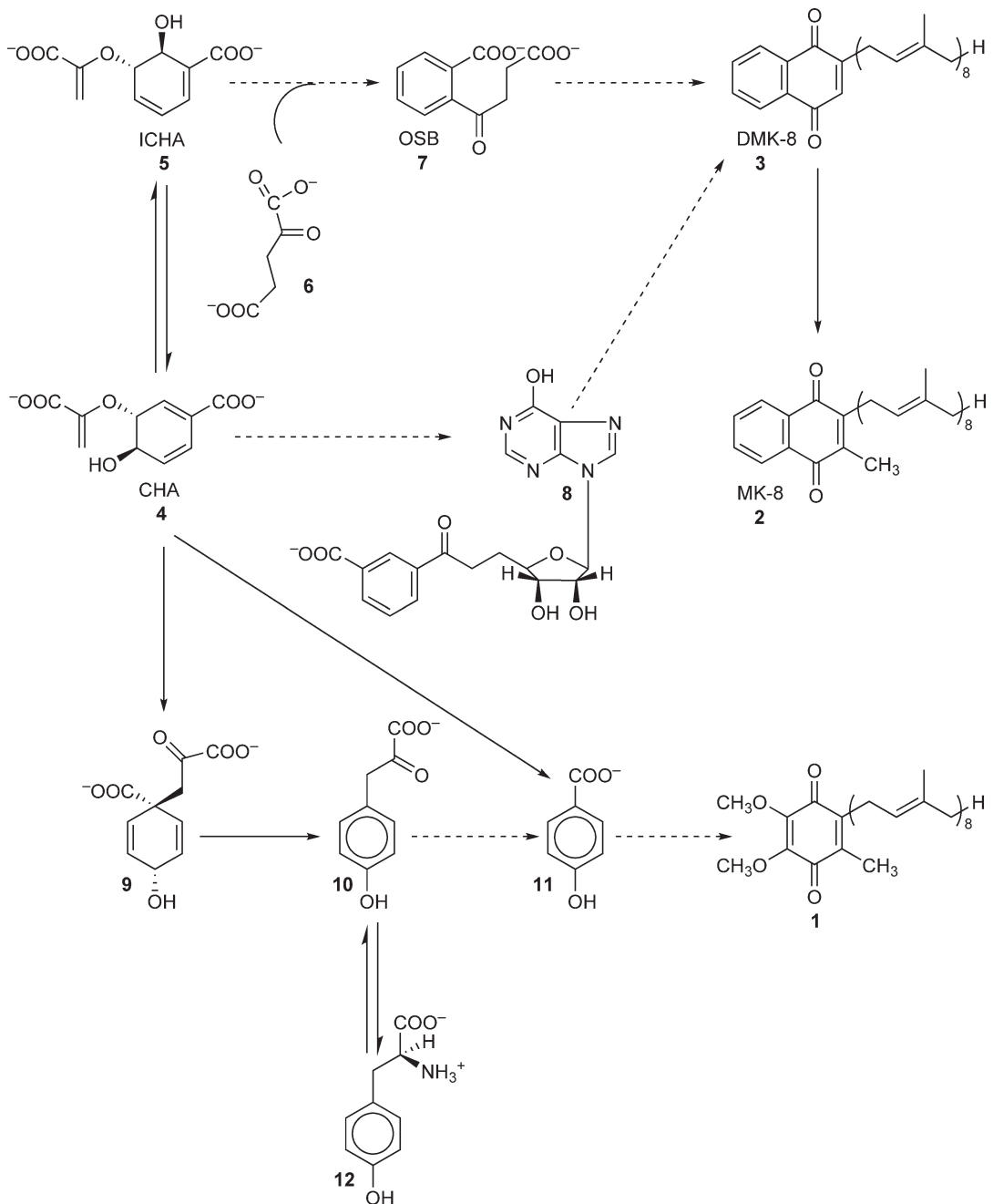
The quinones serve as links between the various dehydrogenases and the next protein component in the respiratory chain. Q functions in the aerobic respiratory chain when oxygen serves as the electron acceptor. In *E. coli*, under anaerobic conditions, when nitrate is the electron acceptor, Q is the preferred quinone. But in the absence of Q, MK can substitute for this function.<sup>3</sup> However, under anaerobic conditions when fumarate, trimethylamine *N*-oxide (TMAO), dimethylsulfoxide (DMSO), or tetrahydrothiophene 1-oxide (THTO) serve as the electron acceptors, the presence of MK is obligatory.<sup>4-7</sup> Reconstitution studies using membrane preparations from quinone-deficient strains suggest that MK is specifically involved in fumarate and DMSO reduction, while either MK or DMK can function in TMAO reduction, and that nitrate reduction requires either Q or MK.<sup>8</sup>

The benzoquinones and naphthoquinones are derived from the shikimate pathway (Figure 2), which has been described as a ‘metabolic tree with many branches.’<sup>9</sup> Most of the information concerning the biosynthesis of MK and Q was obtained with *E. coli* by using isotopic tracers, by the isolation of mutants, and by the accumulation of intermediates and enzyme assays. Due to space limitations, only a general account is given here; for more information, several comprehensive reviews should be consulted.<sup>9-16</sup>

Reviews on the reaction mechanisms of various enzymes involved in MK and Q biosynthetic pathways have been published by Begley *et al.*<sup>17</sup> and Meganathan.<sup>18</sup> Two short reviews on Q biosynthesis have appeared.<sup>19,20</sup> In most of these reviews, the work on *E. coli* and to a lesser extent on *Salmonella* predominated due to the ease with which these organisms can be manipulated. However, due to advances in technology, it has become a reality that work on other bacteria can now be carried out with ease.

Hence, driven by the economic and chemotherapeutic potential, research in certain aspects of the MK and Q biosynthesis in some organisms has moved ahead of that of *E. coli*. The MK and phyloquinone (K) biosynthetic pathways (the reactions of biosynthesis are identical except for the prenylation) are unique to bacteria and plants, respectively, and are absent in humans and animals. Hence, there is great commercial interest in discovering chemicals that will inhibit the enzymes of the pathway. Such chemicals will be of use as herbicides and chemotherapeutic agents against pathogens such as multidrug-resistant *Mycobacterium tuberculosis* and methicillin-resistant *Staphylococcus aureus* (MRSA). With the idea of designing drugs, the crystal structure of MenB from *M. tuberculosis* has been solved by two different groups<sup>21,22</sup> and that of *S. aureus* was recently reported.<sup>23</sup> It is expected that the crystal structures of enzymes will be similar if not identical in most organisms since they perform chemically identical reactions.

Both quinones play essential biochemical and physiological roles in the well-being of humans and animals. Vitamin K was discovered in 1935. Thirty years later, in 1964, it was shown to be derived from the ‘shikimate pathway.’ Another 10 years had to elapse before it was assigned a specific physiological role in blood coagulation. Vitamin K deficiency leads to the synthesis of abnormal (nonfunctional) forms of coagulation factors. Under normal conditions, specific glutamate residues in the N-terminal portion of the protein are carboxylated (posttranslational modification) to  $\gamma$ -carboxyglutamates (gla). Other proteins in bone, kidney, and



**Figure 2** Primary pathways for the formation of MK, DMK, and Q in microorganisms. (1) Q-8, ubiquinone; (2) MK-8, menaquinone; (3) DMK-8, demethylmenaquinone; (4) CHA, chorismate; (5) ICHA, isochorismate; (6) 2-ketoglutarate; (7) OSB, o-succinylbenzoate; (8) futasoline; (9) prephenate; (10) 4-hydroxyphenylpyruvate; (11) 4-hydroxybenzoate; (12) tyrosine.

spermatozoa also contain gla. Gla proteins also appear to be involved in calcium metabolism. Human deficiencies of vitamin K-dependent coagulation factors are rare. However, vitamin K-dependent deficiencies occur during the treatment with antibiotics, anticoagulants, and large doses of vitamin E.<sup>24-26</sup> A significant part of the human requirement for vitamin K is provided by the microflora of the gut. The remainder is provided by K from plant sources.<sup>26,27</sup>

In green plants, K plays a role in photosynthesis and is present in photosystem I. K has been identified as the secondary acceptor A<sub>1</sub> in photosystem I.<sup>28</sup> It should be pointed out that the structure of K is identical to that of MK except that the prenyl side chain is replaced by a phytyl side chain. Hence, it is assumed that the biosynthetic pathway in plants is identical to that in bacteria.

It is well established that Q is a component of the respiratory chain and plays a critical role in respiration and oxidative phosphorylation. It was thought that the presence of Q was confined exclusively to the inner mitochondrial membrane and its sole function was to serve as the redox component of the respiratory chain. However, this belief has been modified as it has been shown that Q is present in all cellular membranes examined. The major part of Q is present in the reduced form in human and animal tissues, and serves as an important antioxidant.<sup>29</sup> In fact, it has been shown that QH<sub>2</sub>-10, the reduced form of Q-10, efficiently scavenges free radicals and it is as effective as  $\alpha$ -tocopherol in preventing peroxidative damage to lipids, considered the best lipid-soluble antioxidant in humans.<sup>30</sup> The antioxidant and prooxidant properties of mitochondrial ubiquinone have recently been reviewed.<sup>31</sup>

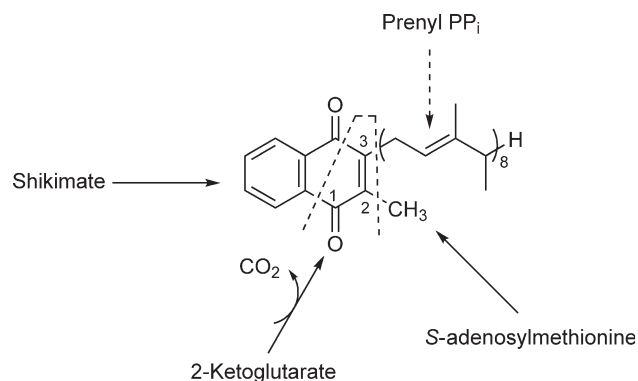
In mammals, QH<sub>2</sub> prevents lipid peroxidation in membranes and in serum low-density lipoprotein (LDL).<sup>32</sup> Mitochondrial QH<sub>2</sub> provides protection from lipid peroxidation, modification of integral membrane proteins, and DNA oxidation and strand breakage.<sup>33</sup> Packer *et al.*<sup>34</sup> have reported that QH<sub>2</sub> is able to scavenge nitric oxide radicals (NO) *in vitro*, but the biological significance of this reaction is yet to be determined. Clinically, Q-10 has been used as an orally administered therapy against a wide variety of diseases in humans. Q deficiency can be reversed by supplementation with Q-10, which is not toxic even in high doses and shows no side effects. Recently, Q-10 has gained importance because of its usefulness in the treatment of heart disease. Human LDL is protected from lipid peroxidation by Q-10, thus delaying the onset of atherosclerosis. Patients suffering from ischemic heart disease show a significantly higher LDL/Q ratio. When the diet is supplemented with Q-10, the level of QH<sub>2</sub>-10 increases within the LDL resulting in resistance to the initiation of lipid peroxidation.<sup>35</sup> Degenerative diseases and aging may result from decreased ability to maintain sufficient QH<sub>2</sub> levels.<sup>36</sup> The metabolism and functions of Q have recently been reviewed.<sup>37-39</sup>

## 7.12.2 Menaquinone Biosynthesis

Shikimate through chorismate is a major precursor of MK as well as the closely related plant quinone, K. In prokaryotes, there are two different pathways for the biosynthesis of MK. Both pathways are branches of the shikimate pathway that diverge at chorismate and converge at DMK. Among the known MK-producing bacteria, the vast majority so far examined biosynthesize MK by the well-established *o*-succinylbenzoate (OSB) pathway,<sup>18,40</sup> while a small group of organisms have been shown to produce MK by the newly discovered non-OSB pathway or futasine pathway.<sup>41</sup> The latter pathway has so far been demonstrated in *Streptomyces coelicolor* A3(2), *Helicobacter pylori*, *Campylobacter jejuni*, and *Thermus thermophilus*. Of the four genes experimentally identified in the non-OSB or futasine pathway, at least three orthologs have been identified in a number of Gram-positive and Gram-negative bacteria and archaea.<sup>41</sup>

### 7.12.2.1 *o*-Succinylbenzoate Pathway

MK biosynthesis by the OSB pathway has been elucidated on the basis of isotopic tracer experiments, isolation of mutants blocked in the various steps, isolation and identification of intermediates accumulated by the mutants, and by enzyme assays. Early isotopic tracer experiments with various bacteria established that methionine and prenyl PPI contribute to the methyl and prenyl substituents of the naphthoquinone. The early isotopic tracer studies and other work have been reviewed by Bentley and Meganathan.<sup>14</sup> In 1964, Cox and Gibson<sup>42</sup> observed that [G-<sup>14</sup>C] shikimate was incorporated into both MK and ubiquinone by *E. coli*, thus providing the first evidence for the involvement of the shikimate pathway.<sup>42</sup> Chemical degradation of the labeled isolated menaquinone (MK-8) showed that essentially all of the radioactivity was retained in the phthalic anhydride. It was concluded that "the benzene ring of the naphthoquinone (sic) portion of vitamin K<sub>2</sub> arises from shikimate in *E. coli*."<sup>42</sup> The authors further suggested that shikimate was first converted to chorismate before incorporation into MK. A more complete chemical degradation of the MK derived from



**Figure 3** Primary biosynthetic precursors of MK derived by the OSB pathway.

labeled shikimate established that all seven carbon atoms were incorporated.<sup>43</sup> The remaining three carbon atoms of the naphthoquinone ring were shown to be derived from the middle three carbons of 2-ketoglutarate with the loss of both carboxyl groups (Figure 3).<sup>44–46</sup>

These studies established the immediate precursors of MK as shikimate and the noncarboxyl carbon atoms of 2-ketoglutarate forming the naphthoquinone nucleus. The methyl and isoprenoid side chains were also shown to be derived from *S*-adenosylmethionine (SAM) and an isoprenyl alcohol pyrophosphate ester, respectively. Subsequently, it was shown that the benzenoid aromatic compound OSB (7) (Figure 4)<sup>47</sup> and the naphthalenoid aromatic compound 1,4-dihydroxy-2-naphthoate (DH2NA) (18)<sup>14,48</sup> were incorporated into the naphthoquinone ring of MK (2). This work was confirmed by the demonstration that *menB* and *menA* mutants of *E. coli* excrete OSB (7) and DH2NA (18), respectively, into the culture medium.<sup>49</sup> During a study of the biosynthesis of OSB (7) by growing cultures of *E. coli menB* it was demonstrated that carbon one of glutamate (2-ketoglutarate) was lost and consequently not incorporated into OSB (7).<sup>50</sup> The isotopic labeling pattern is summarized (Figure 3).

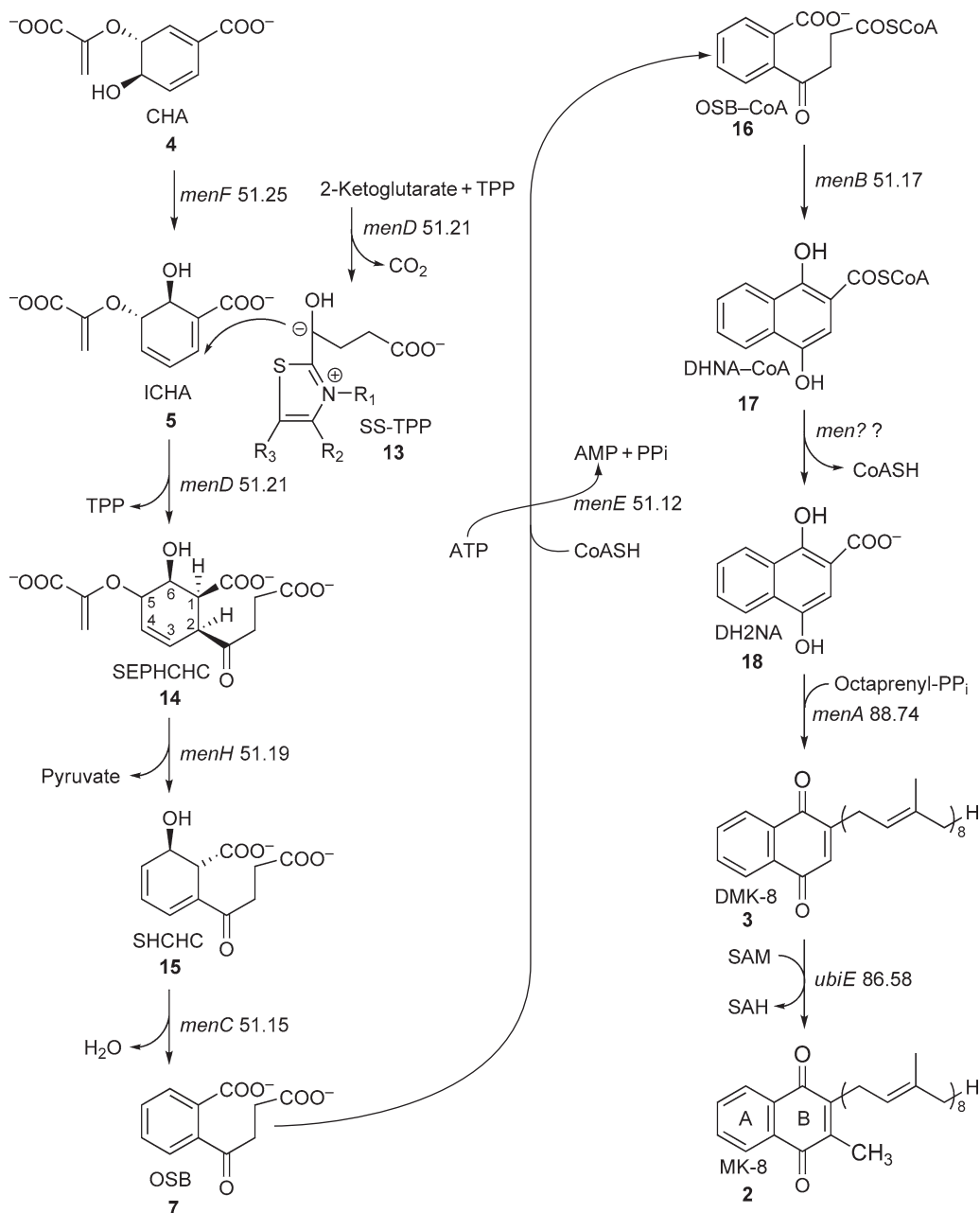
#### 7.12.2.1.1 Formation of isochorismate (4) → (5)

The first synthesis of OSB (7) from chorismate (4) and 2-ketoglutarate in the presence of thiamin pyrophosphate (TPP) by cell-free extracts of *E. coli* was obtained by Meganathan (Figure 4).<sup>51</sup> However, it had been suggested that isochorismate (5) was a much more attractive precursor than chorismate (4) on chemical grounds.<sup>10,47</sup> Evidence in support of this hypothesis was provided (Figure 4).<sup>52–54</sup>

Isochorismate is a common intermediate in the biosynthesis of the siderophore enterobactin and MK. The conversion of chorismate (4) to isochorismate (5) in enterobactin biosynthesis is mediated by the enzyme isochorismate synthase (ICS) encoded by the *entC* gene.<sup>55,56</sup> The dual role of isochorismate led to the question as to whether the *entC*-encoded ICS (EntC) was supplying the isochorismate required for both pathways. Kaiser and Leistner<sup>57</sup> reported the isolation of a Tn10 insertion in the *entC* gene that had lost simultaneously the ability to form enterobactin and MK. It is generally accepted that the *entC* gene is derepressed under iron deficiency and repressed under iron sufficiency.<sup>58,59</sup> Enterobactin is required only under aerobic conditions due to the poor solubility and the consequent unavailability of iron in the Fe<sup>3+</sup> form. When *E. coli* is grown anaerobically, iron is present in the highly soluble Fe<sup>2+</sup> form. Hence, the synthesis of enterobactin is unnecessary for the acquisition of iron by the cell under anaerobic conditions.<sup>58,59</sup>

In contrast, MK is required under anaerobic conditions.<sup>14</sup> Furthermore, when the organism is grown with fumarate, TMAO, or DMSO as electron acceptor, the presence of MK is obligatory.<sup>4,5,14</sup> When oxygen or nitrate is the electron acceptor, the aerobic quinone, ubiquinone, is used by *E. coli*.<sup>3</sup> Thus, while the conditions that favor the biosynthesis and function of Q are compatible with the biosynthesis of enterobactin, they are incompatible with the biosynthesis of MK.

These apparent contradictions raised some intriguing questions. How does *E. coli*, growing aerobically under iron deficiency when the *entC* gene is fully derepressed, prevent the synthesis of MK? Furthermore, under anaerobic conditions, how does *E. coli* prevent the synthesis of enterobactin when MK synthesis is induced? This



**Figure 4** The OSB pathway for MK biosynthesis. Each compound in the pathway is identified by its abbreviation and Arabic numeral. (4) CHA, chorismate; (5) ICHA, isochorismate; (13) SS-TPP, succinic semialdehyde-TPP,  $R_1$  = pyrimidine component of TPP,  $R_2$  =  $\text{CH}_3$ ,  $R_3$  =  $\text{CH}_2\text{CH}_2\text{OP}_2\text{O}_6^{3-}$ ; (14) SEPHCHC, 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate; (15) SHCHC, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate; (7) OSB, *o*-succinylbenzoate; (16) OSB-CoA, *o*-succinylbenzoyl-CoA; (17) DHNA-CoA, 1,4-dihydroxy-2-naphthoyl-CoA; (18) DH2NA, 1,4-dihydroxy-2-naphthoate; (3) DMK-8, demethylmenaquinone (may be initially formed as a quinol); (2) MK-8, menaquinone; SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine. The genes encoding the enzymes are shown for each reaction followed by their location on the chromosome in minutes. The gene encoding the thioesterase for the conversion of DHNA-CoA (compound (17)) to DH2NA (compound (18)) remains to be identified in *E. coli* and is shown as *men??*.

paradox might be resolved if the *entC* gene is regulated by iron in the presence of oxygen and by MK requirement in the absence of oxygen. To study the regulation of the *entC* gene, an *entC-lacZ<sup>+</sup>* operon fusion was constructed and the expression of  $\beta$ -galactosidase monitored under various conditions. It was found that



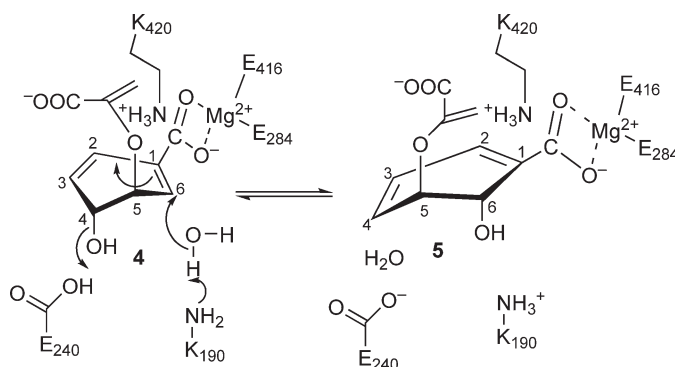
the  $\beta$ -galactosidase was fully derepressed at low concentrations of iron and repressed at high concentrations of iron under both aerobic and anaerobic growth conditions.<sup>60,61</sup>

These results raised the question as to how *E. coli* is able to synthesize MK anaerobically when growing in the presence of high concentrations of iron? How does the organism prevent the excess production of MK under iron-deficient aerobic conditions when the *entC* gene is fully derepressed? To answer these questions, anaerobic growth of an *entC::Tn5* mutant was tested on glycerol medium with TMAO, DMSO, or fumarate as electron acceptor. The mutant was able to grow at the same rate as the parent, even in the presence of high concentrations of iron. Furthermore, the mutant produced as much MK as the parent.<sup>60–62</sup> These results provided clear evidence for the presence of an alternate ICS specifically involved in MK biosynthesis. As a first step in locating and identifying the gene encoding this alternate ICS involved in MK biosynthesis, further sequencing upstream of the 5' region of the *menD* gene was carried out. An open reading frame encoding a 430-amino acid protein exhibiting about 20% amino acid identity with EntC was identified as MenF.<sup>61,63,64</sup>

The ICS (MenF) encoded by the *menF* gene has been overexpressed and purified to homogeneity. The purified enzyme had a relative  $M_r$  of 48 000<sup>63,64</sup> as determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The native  $M_r$ , as determined by gel filtration chromatography, was 98 000, thus establishing that the native enzyme is a homodimer.<sup>64</sup> The enzyme showed a requirement for  $Mg^{2+}$  for maximal activity.

It is expected that the origin of the hydroxyl groups and the mechanism of the reaction of MenF and EntC will be identical. Four different mechanisms were proposed for the EntC enzyme.<sup>65</sup> The origin of the hydroxyl could be from three possible sources: (1) molecular oxygen, (2) intramolecular transfer of the hydroxyl, or (3) the solvent  $H_2O$ . While the incorporation of molecular oxygen is possible only in the case of aerobic organisms, intramolecular transfer or the incorporation of hydroxyl from water can be carried out by both aerobes and anaerobes. Due to the reported absolute requirement of enterobactin for the chelation of iron during aerobic growth under iron deficiency,<sup>58,59</sup> one would expect the incorporation of molecular oxygen into the hydroxyl group. However, the absence of redox cofactor rules out the involvement of oxygen and evidence has been obtained demonstrating the incorporation of the C-6 hydroxyl from the solvent  $H_2O$  for EntC.<sup>65</sup> Consistent with this result is the demonstration of anaerobic biosynthesis of enterobactin in *E. coli*.<sup>66</sup>

Recently, the three-dimensional (3D) structure of MenF has been determined and the catalytic mechanism was probed by site-directed mutagenesis and biochemical studies.<sup>67,68</sup> Lys190 has been identified as the active site base that assists in the attack by water at the  $C_2$  carbon. An  $S_N2''$  reaction results in the rearrangement of 1–2, 5–6 double bonds resulting in the elimination of the  $C_4$  hydroxyl group (Figure 5). For activity, Glu240 is required and it provides acid catalysis for the elimination of the hydroxyl at  $C_4$ .<sup>67</sup> As seen from the figure, magnesium plays a central role in the mechanism consistent with the demonstrated requirement for  $Mg^{2+}$  in the reaction.<sup>63,64</sup> These findings are in complete agreement with a common mechanism proposed for the three chorismate-utilizing enzymes: anthranilate synthase (AS), 4-amino-4-deoxychorismate synthase (ADCS), and ICS.<sup>69</sup>



**Figure 5** Proposed mechanism for the isomerization of chorismate (4) to isochorismate (5) by the enzyme isochorismate synthase.

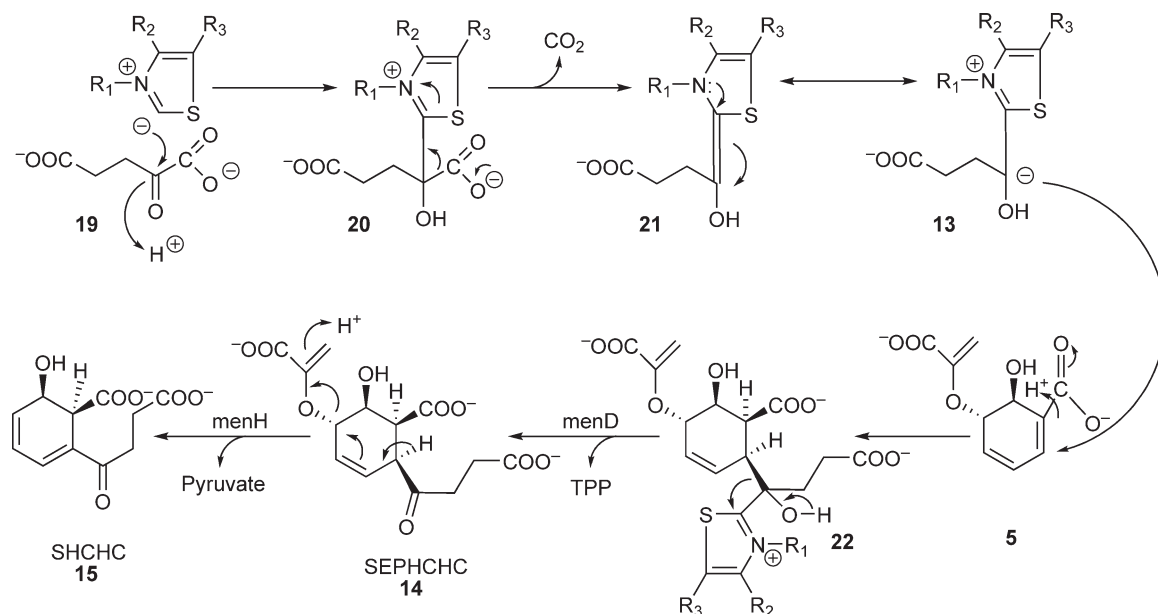
### 7.12.2.1.2 Formation of succinic semialdehyde-thiamine pyrophosphate anion and Michael addition to isochorismate (compound (5) + (13) → (14))

During the studies on the biosynthesis of OSB (7), the cell extracts of two groups of mutants designated as *menC* and *menD* blocked in the formation of OSB and requiring OSB for anaerobic growth on glycerol–fumarate medium were examined. The cell extracts of either mutant alone did not form OSB from chorismate (4) and 2-ketoglutarate in the presence of TPP. However, extracts from both mutants in combination produced OSB, and extracts of *menC* mutants accumulated an intermediate, which was converted to OSB by extracts of *menD* mutants.<sup>70</sup> The intermediate was found to be unstable and on mild acid treatment yielded OSB and succinylbenzene. On the basis of these properties and nuclear magnetic resonance data, the intermediate was identified as 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) (15).<sup>53</sup>

It has been postulated that the 2-ketoglutarate undergoes a TPP-dependent decarboxylation with the formation of succinic semialdehyde anion of TPP (13)<sup>44,51</sup> and a requirement for TPP in the reaction was shown.<sup>70</sup> The mechanism of decarboxylation is identical to that catalyzed by the first enzyme of the 2-ketoglutarate dehydrogenase (KGDH) complex (Figure 6).<sup>52,71</sup> Using a *sucA* mutant, which lacks the first enzyme of the KGDH complex, and by selective removal of the KGDH complex it was established that the 2-ketoglutarate decarboxylase (KDC) involved in OSB synthesis is a separate enzyme.<sup>15,71,72</sup>

Subsequent studies established that the succinic semialdehyde anion (13) of TPP reacted with isochorismate (5) resulting in the formation of SHCHC (15)<sup>52,53,72</sup> as had been postulated previously.<sup>10,47</sup> A mechanism for this reaction has been proposed (Figure 6).<sup>9,18,52</sup> When the complete nucleotide sequence of the *menD* gene was determined, it was discovered that both SHCHC synthase and KDC activities are encoded by a single gene.<sup>73</sup> This conclusion was further strengthened by overexpression and purification of the MenD protein and by showing that both activities were copurified during various steps of the purification process.<sup>74,75</sup>

However, recently, Guo and colleagues have shown that the formation of SHCHC from isochorismate and 2-ketoglutarate is a two-step process requiring two different enzymes. The first enzyme MenD decarboxylates 2-ketoglutarate and adds the resulting succinic semialdehyde anion of TPP (13) to isochorismate (5) resulting in the formation of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate (SEPHCHC) (14).<sup>76</sup> The mechanism of formation of succinic semialdehyde anion of TPP and its addition to isochorismate is shown



**Figure 6** Proposed mechanism of succinic semialdehyde-TPP anion (13) formation and the synthesis of SHCHC (15). Only the thiazole ring of the TPP is shown since it is the active site of the molecule. For R<sub>1</sub>, R<sub>2</sub>, and R<sub>3</sub>, see the legend of Figure 4. The reactions from (19) to SEPHCHC (14) are carried out by MenD and the conversion of SEPHCHC (14) → SHCHC (15) is by MenH.

in (Figure 6). The stereochemistry of SEPHCHC was determined and shown to be (1R,2S,5S,6S)-2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic acid.<sup>77</sup> On the basis of these results, MenD was designated as SEPHCHC synthase.<sup>76</sup>

SEPHCHC is an unstable compound, which, in mildly basic solutions, spontaneously undergoes a 2,5 elimination reaction resulting in the formation of SHCHC and pyruvate. Crystallization and a preliminary X-ray analysis of MenD have been reported.<sup>78</sup>

The *in vivo* conversion of SEPHCHC (14) to SHCHC (15) is carried out by (1R,6R)-SHCHC synthase encoded by the *menH* gene (Figure 6). Surprisingly, MenH contains a Ser-His-Asp catalytic triad, which is typical of many proteases. This triad plays a critical role in enzymatic activity since replacing any one of the three amino acids by alanine results in a dramatic decrease in catalytic activity.<sup>79</sup> The structure of MenH from the enteric pathogen *Vibrio cholerae* has been determined (Pdb# 1R3D).

### 7.12.2.1.3 The aromatization of SHCHC → OSB (compound (15) → (7))

Enzymatic removal of the elements of water from SHCHC (15) leads to the formation of the benzenoid aromatic compound OSB (7) (Figure 4). The first evidence for the presence of such an enzyme was obtained by the demonstration that cell-free extracts of a *menD* mutant converted SHCHC (designated as 'X' at the time) to OSB.<sup>70</sup> This enzyme was subsequently designated as OSB synthase.<sup>80,81</sup> The gene encoding OSB synthase was cloned and its complete nucleotide sequence was reported.<sup>81</sup>

The enzyme was overexpressed, purified to homogeneity, and its properties investigated. The enzyme required a divalent metal ion for activity like the other members of the enolase superfamily. The enzyme was shown to carry out the dehydration of SHCHC (15) to OSB (7) very efficiently with  $K_{cat}$  ( $19 \pm 1 \text{ s}^{-1}$ ) and a  $K_{cat}/K_m$  ( $1.6 \pm 0.3 \times 10^6 \text{ mol}^{-1} \text{ s}^{-1}$ ).<sup>82</sup> OSB synthase was classified as a member of the enolase superfamily. Members of this superfamily carry out reactions initiated by the abstraction of the  $\alpha$ -proton from a carboxylate anion substrate to generate a stabilized enolate anion intermediate.<sup>83</sup> As pointed out above, the reaction catalyzed by OSB synthase is a dehydration. It was proposed that the  $\alpha$ -proton of the carboxylate substrate (SHCHC) is likely abstracted by a basic catalyst (one lysine) followed by the elimination of the  $\beta$ -hydroxyl group presumably by the assistance of an acid catalyst (a second lysine).<sup>82</sup>

The structure of OSB synthase from *E. coli* in complex with  $\text{Mg}^{2+}$  and OSB was determined. It was found that contrary to the previous finding OSB synthase is the only monomeric member of the enolase superfamily. The product OSB was found to be sandwiched between Lys 133 and Lys 235 located at the ends of the second and sixth  $\beta$ -strands. In addition, one carboxylate oxygen of the substrate is coordinated to the  $\text{Mg}^{2+}$ .<sup>84</sup> Subsequently, the structure of OSB synthase from an inactive K133R mutant in complex with the substrate SHCHC was determined. It was found that contrary to the previous finding Lys 133 is the single base/acid catalyst for the dehydration with the transient  $\text{Mg}^{2+}$  coordinated enolate anion intermediate<sup>85</sup> (Figure 7).

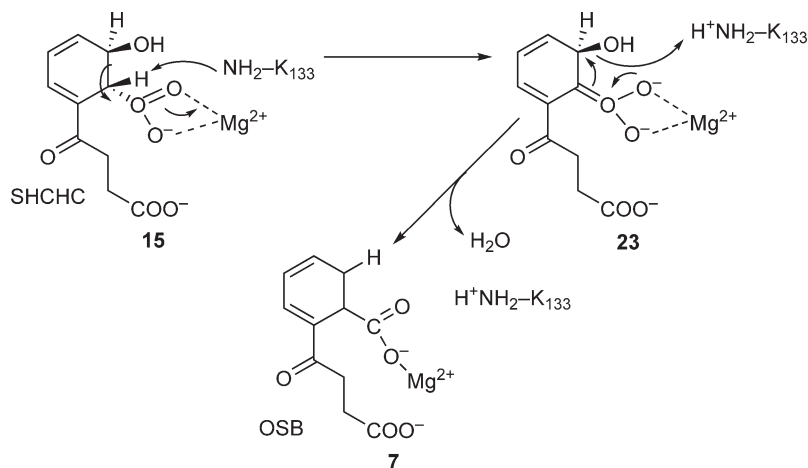


Figure 7 o-Succinylbenzoate synthase reaction mechanism.

The dehydration was shown to follow a *syn*-stereochemical course.<sup>85</sup> The mechanism and specificity of various members of the enolase superfamily including OSB synthase have been reviewed and should be consulted for further details of the reaction mechanism.<sup>86</sup>

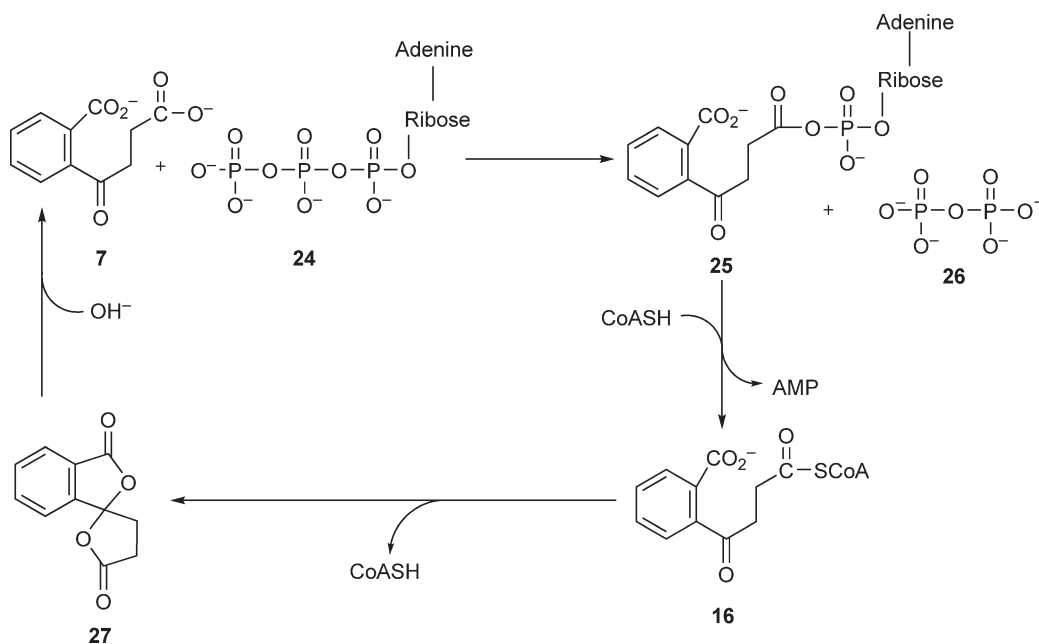
#### 7.12.2.1.4 Cyclization of OSB to DH2NA (compound (7) → (18))

The conversion of the benzenoid aromatic compound OSB (7) to the naphthalenoid aromatic compound DH2NA (18) was demonstrated by Bryant and Bentley.<sup>87</sup> The process showed an absolute requirement for ATP and coenzyme A (CoA). Hence, OSB-CoA (16) was suggested as an intermediate. Using extracts of *M. pblei*, evidence was obtained for the presence of two enzymatic activities (OSB-CoA synthetase and DH2NA synthase). The OSB-CoA (16) was found to be an unstable intermediate, which spontaneously hydrolyzed to the spirodilactone (27) form of OSB (Figure 8). Furthermore, it was shown that during the formation of OSB-CoA, ATP was hydrolyzed to AMP and PPI, which is typical of ligases forming CoA esters.<sup>88</sup>

The CoA moiety was suggested to be on the aromatic carboxyl group<sup>87,88</sup> and evidence in support of this suggestion was obtained.<sup>89,90</sup> However, in subsequent publications, it was reported that the CoA is located on the aliphatic carboxyl group.<sup>91,92</sup>

A group of *E. coli* mutants responding to DH2NA (18), but not to OSB (7), for anaerobic growth on glycerol–fumarate medium was analyzed for their ability to convert OSB to DH2NA.<sup>93</sup> None of the mutant extracts formed DH2NA. However, when the cell extracts from different mutants were mixed with each other, one of the mutant extracts complemented with the extracts of each one of the other three mutants and formed DH2NA. To identify the nature of the enzymatic defect in these mutants, the cell extracts from each one of these mutants were complemented with OSB-CoA synthetase and DH2NA synthase from *M. pblei* as described above and assayed for DH2NA formation. The single mutant whose extract was complemented by OSB-CoA synthetase and, therefore lacking this enzyme, was designated as *menE*. The other three mutants, whose extracts were complemented by DH2NA synthase, were designated as *menB*.<sup>93</sup>

The *menE* gene was cloned and sequenced.<sup>94</sup> The gene was overexpressed and the enzyme was purified to homogeneity. The purified enzyme had subunits  $M_r$  49 000 and a native  $M_r$  of 185 000. Thus the native enzyme appears to be a homotetramer. The  $K_m$  values for OSB, ATP, and CoA were 16, 73.5, and 360  $\mu\text{mol l}^{-1}$ , respectively.<sup>61</sup> By chemical inactivation and site-directed mutagenesis studies, an essential histidine residue (His341) located in the ATP-binding region has been identified as necessary for catalytic activity of the



**Figure 8** Proposed mechanism for the formation of acyl adenylate of OSB (25) and its subsequent conversion to OSB-CoA (16). The conversion of OSB-CoA (16) to spirodilactone of OSB (27) is nonenzymatic.

enzyme.<sup>95</sup> Sequence analysis combined with the fact that OSB-CoA synthetase hydrolyses ATP to AMP and PPI, and requires CoASH for the reaction earns it membership in the acyl-adenylate/thioester-forming superfamily of enzymes.<sup>96,97</sup> A mechanism for the reaction has been proposed (Figure 8).

The *menB* gene was cloned and its complete nucleotide sequence was determined.<sup>98</sup> When the gene was overexpressed and the protein was purified to homogeneity, the subunits were found to have an  $M_r$  of 32 000, whereas the native protein had an  $M_r$  of 112 000 as determined by gel filtration. Thus, the enzyme is a homotetramer.<sup>18</sup>

As discussed above, the substrate for MenB, OSB-CoA, is highly unstable. Hence, for the assay of MenB, the required OSB-CoA is generated *in vitro* by coupling the reaction with the MenE reaction.<sup>14,88</sup> However, for the coupled assays, crude cell-free extracts were always used. Surprisingly, when the overexpressed and purified MenE and MenB enzymes were used in the coupled assay, DH2NA formation was not observed. In order to determine the reasons for the lack of formation of DH2NA, small amounts of a crude cell-free extract of *E. coli* were added to the reaction mixture, and this resulted in the restoration of activity in the incubation mixture. Hence, it appeared that either a cofactor or another protein might be involved in the reaction.

On the basis of alignment and analysis of the sequence, MenB was included in the enoyl-CoA hydratase/isomerase (crotonase) superfamily.<sup>99</sup> The failure of the purified MenB to form DH2NA in the complementation assay discussed above and its membership in the enoyl-CoA hydratase superfamily (where other members form CoA esters) suggested that the product of MenB is DHNA-CoA (17) rather than DH2NA (18).<sup>18</sup> Evidence in support of this prediction has been obtained in *M. tuberculosis* where the product of MenB was identified by mass spectrometry (MS) as DHNA-CoA (17).<sup>22</sup> The crystal structure of MenB as the native enzyme and in complex with acetoacetyl-CoA and DHNA-CoA, respectively, has been reported by Truglio *et al.*<sup>22</sup> and Johnston *et al.*<sup>21</sup> The highly conserved active site of MenB contained a deep pocket lined with Asp-192, Tyr-287, and hydrophobic amino acids. Site-directed mutagenesis studies have established that Asp-192 and Tyr-287 are essential for enzymatic catalysis. On the basis of structural and mutagenesis studies, the authors have proposed a possible mechanism for cyclization of OSB-CoA to DHNA-CoA (Figure 9).<sup>22</sup> This mechanism is in complete agreement with the previous proposal for the two-step reaction, namely, the cyclization of OSB-CoA to DH2NA.<sup>17,18</sup>

On the basis of amino acid sequence homology to thioesterases, an unidentified *orf152* (*yfbB*) was postulated to carry out the conversion of DHNA-CoA to DH2NA and designated as *menH*.<sup>18,82</sup> Evidence in support of the proposal was provided by experimental demonstration of thioesterase activity of the protein.<sup>100</sup>

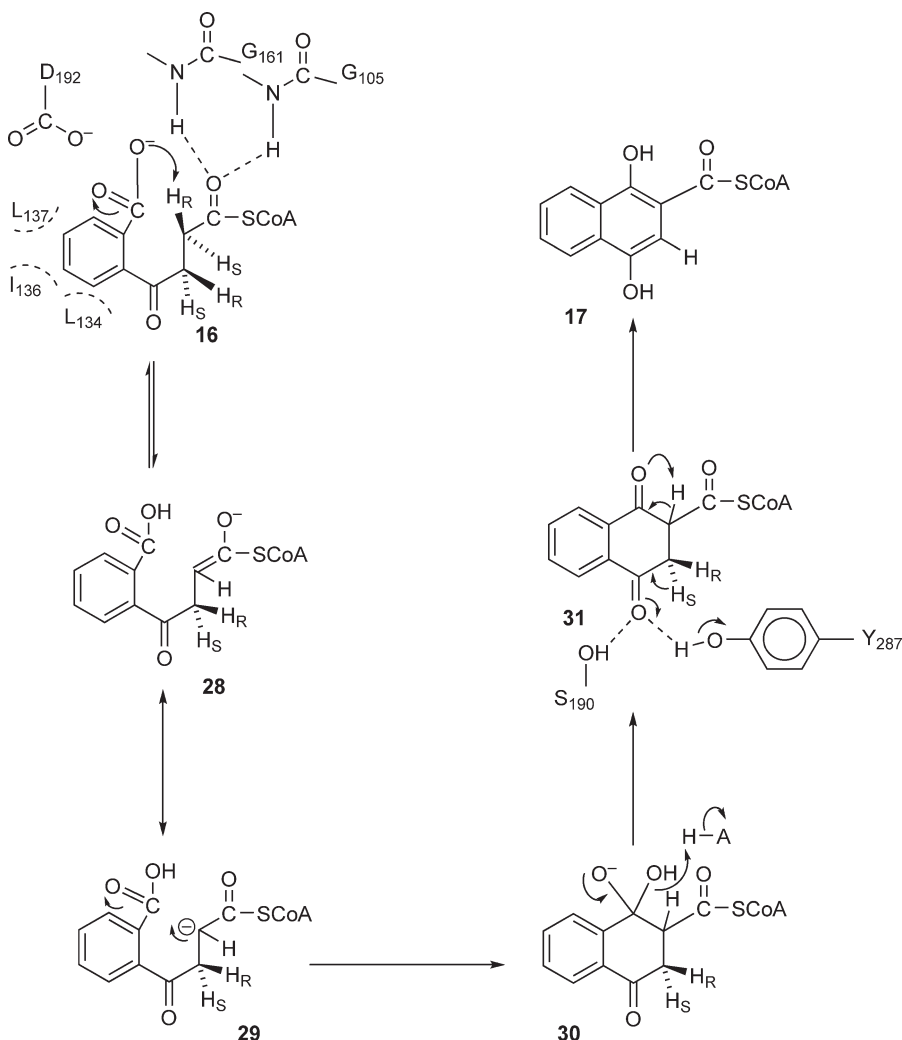
However, as discussed above in Section 7.12.2.1.2, the MenH protein has been unequivocally demonstrated to carry out the conversion of the newly discovered intermediate SEPHCHC to SHCHC and has been christened as SHCHC synthase.<sup>79</sup> Thus, the enzyme responsible for the conversion of DHNA-CoA to DH2NA still remains to be identified. In this connection, it is worth mentioning that in the cyanobacterium *Synechocystis* sp. a DHNA-CoA thioesterase belonging to the hotdog-fold family has been recently shown to be essential for K biosynthesis.<sup>101</sup>

#### 7.12.2.1.5 Prenylation of DH2NA to DMK (compound (18) → (3))

The conversion of DH2NA (18) to DMK (3) in the extracts of *E. coli* was shown by Bentley.<sup>12</sup> Shineberg and Young<sup>102</sup> were able to isolate a membrane-bound DH2NA octaprenyltransferase. The *menA* gene encoding the enzyme has been cloned.<sup>103</sup> The enzyme MenA has many features in common with 4-hydroxybenzoate octaprenyltransferase (UbiA) involved in the biosynthesis of ubiquinone. The two enzymes share a common pool of membrane-bound octaprenyl diphosphate.<sup>102</sup> The conversion of DH2NA to DMK requires replacement of the carboxyl with the isoprenoid side chain. Prenylation and decarboxylation may occur in a single active site, since symmetry experiments exclude 1,4-naphthoquinone as an intermediate.<sup>104</sup> Moreover, there has been no evidence for two separate reaction steps or enzymes. A carbocation mechanism based on the dimethylallyl tryptophan synthase reaction<sup>105</sup> has been proposed for the reaction (Figure 10).<sup>18</sup> In addition, a quinol to quinone oxidation is required in which demethylmenaquinol is a likely intermediate; the oxidation to DMK is thought to be spontaneous.

#### 7.12.2.1.6 Methylation of DMK to MK (compound (3) → (2))

DMK (3) is methylated to MK (2) by a methyltransferase, which uses SAM as the methyl donor. In experiments with whole cells, it was shown that all three hydrogen atoms of the methyl group of methionine are transferred



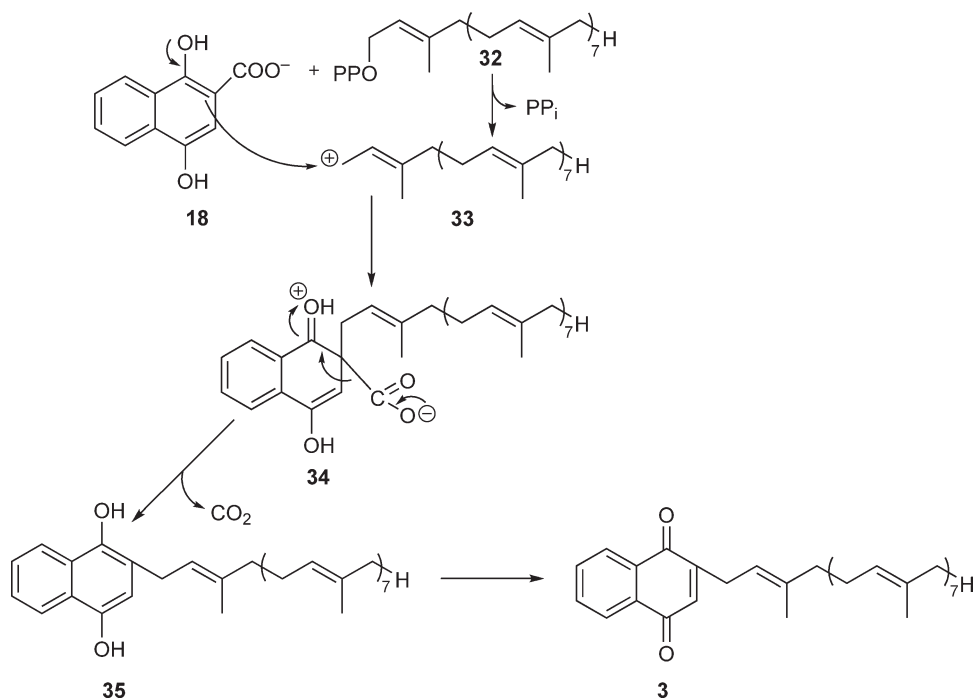
**Figure 9** Proposed mechanism for the cyclization of OSB-CoA (**16**) to DHNA-CoA (**17**) by the enzyme DHNA-CoA synthase (MenB).

to DMK.<sup>106</sup> The conversion of DMK-3 to MK-3 was demonstrated in the cell extracts using *S*-[<sup>14</sup>CH<sub>3</sub>]-adenosyl-L-methionine by Bryant and Bentley.<sup>87</sup> An *ubiA* mutant of *E. coli* was found to accumulate DMK but not MK. This mutant is believed to be defective in the methylation of DMK to MK, suggesting that this is a double mutant.<sup>107</sup> In a subsequent study, it was shown that an *ubiE* mutant blocked in the methylation of the ubiquinone biosynthetic intermediate 2-octaprenyl-6-methoxy-1,4-benzoquinol (OMB) (**53**) to 2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol (OMMB) (**54**) (Figure 13)<sup>108</sup> accumulated DMK but not MK.<sup>109</sup> Consistent with this observation is the simultaneous loss of *C*-methyltransferase activity toward both OMB and DMK and its restoration by a plasmid containing the *ubiE* gene.<sup>109,110</sup>

### 7.12.2.2 Non-*o*-Succinylbenzoate or Fualosine Pathway

The availability of the complete genomic sequence of a number of organisms made it possible to search for and annotate genes for various metabolic pathways. Surprisingly, in some of the organisms that are known to contain MK, the genes for the OSB pathway could not be identified.<sup>111</sup>



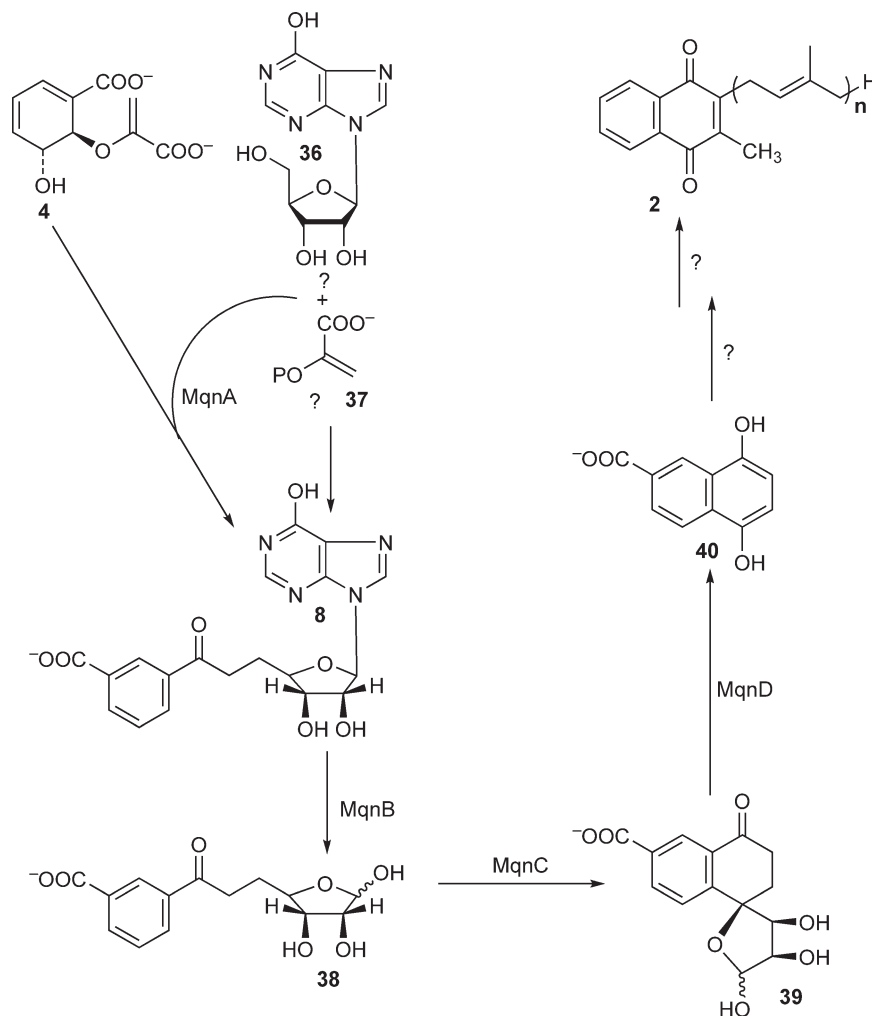


**Figure 10** Proposed mechanism for the prenylation of 1,4-dihydroxy-2-naphthoate (DH2NA) (**18**) to demethylmenaquinone (DMK-8) (**3**).

In order to discover this alternate MK biosynthetic pathway, the incorporation of <sup>13</sup>C glucose into MK by *Streptomyces* sp. was investigated. These labeling studies established that the A ring of MK was derived from the shikimate pathway. The remaining carbons of the B ring were reported to be derived from phosphoenolpyruvate (PEP) and an as yet unidentified metabolite was derived from C-5 and C-6 of glucose.<sup>111</sup> Furthermore, on the basis of the distinct labeling patterns, these authors deduced 1,4-dihydroxy-6-naphthoate (DH6NA) (**40**) (Figure 11) as an intermediate. It was further demonstrated that the synthetic DH6NA was able to restore growth and synthesis of MK in *S. coelicolor* auxotrophs.<sup>111</sup>

In a subsequent study, by chemical mutagenesis, a number of MK auxotrophs were isolated. These auxotrophs required the incorporation of either MK-4 or chorismate into the medium for growth. Furthermore, these mutants failed to respond to shikimate, thus establishing that the branch point of the pathway is at chorismate.<sup>41</sup>

To determine and delineate the intermediates, genes, and enzymes, bioinformatic methods were used. Among the bacteria whose genomes have been completely sequenced, the epsilon branch of the proteobacteria contained the organisms that lacked orthologs of the *men* genes of the OSB pathway. From these, *H. pylori*, *C. jejuni*, *T. thermophilus*, and *S. coelicolor* were selected. From the genome sequences of these four organisms, presumptive orthologous genes were identified by basic local alignment search tool (BLAST) searches as reciprocal best-fit pairs with an *e* value that was <10<sup>-10</sup>. Subsequently, the orthologs that were present in the genomes of these four organisms and absent in organisms that contained the traditional OSB pathway of MK biosynthesis, *E. coli*, *B. subtilis*, *Corynebacterium glutamicum*, and *M. tuberculosis*,<sup>41</sup> were selected for further study. This process resulted in the identification of 50 candidate genes in the genomic sequence of *S. coelicolor*. From these candidate genes, a number of membrane proteins that are involved in the transport of a wide variety of metabolites were eliminated. Finally, four genes in the genome annotated as hypothetical proteins, namely, SCO4506, SCO4326, SCO4327, and SCO4550, were left for further analysis. Knockout mutants were isolated by replacing these genes with the gene encoding thiostreptin resistance. These insertion mutants required the addition of MK-4 to the medium for growth. Cross-feeding experiments showed that each one of these mutants could be cultivated together with any of the other mutants



**Figure 11** Non-*o*-succinylbenzoate (non-OSB) or futasoline pathway for MK biosynthesis. (4) CHA, chorismate; (36) inosine; (37) C<sub>2</sub> unit, Phosphoenolpyruvate?; (8) futasoline; (38) dehypoxanthinylfutasoline (DHFL); (39) cyclic dehypoxanthinylfutasoline; (40) 1,4-dihydroxy-6-naphthoate; (2) menaquinone. In the MK structure,  $n = 9$  in *Streptomyces* and  $n = 8$  in *Thermus*. ?, uncertain or not definite.

without the addition of MK-4 to the medium. The cross-feeding experiments clearly demonstrated that these strains carried blocks in different steps of the pathway and that the intermediate excreted by one mutant is being used by the other mutants for MK biosynthesis. Each one of these mutants was grown in the medium supplemented with MK-4 and, after removal of the cells by centrifugation, the supernatant was concentrated. From the concentrate, the MK-4 was eliminated by ethyl acetate extraction and the aqueous phase was incorporated into fresh growth media. The growth of each of the four mutants in the supplemented medium was tested and the information obtained led to the ordering of the blocks in the pathway as SCO4506, SCO4327, SCO4550, and SCO4326<sup>41</sup> These open reading frames have been recently assigned with new gene/protein designations (Figure 11).<sup>112</sup>

Furthermore, the extracted intermediates were purified and the structures were determined by nuclear magnetic resonance (NMR) and MS. The first intermediate that was identified was excreted by the mutant SCO4327 (*mqnB*) and utilized by SCO4506 (*mqnA*). This compound was identified by NMR and MS as futasoline (8). Interestingly, 10 years earlier, this compound had been isolated from the culture broth of a *Streptomyces* sp. and identified.<sup>113</sup>

Since, as discussed above, chorismate (4) is the branch point of the pathway and SCO4506 (MqnA) is the enzyme responsible for the formation of futasoline (8), attempts were made to demonstrate the reaction in a cell-free system. Structural examination of futasoline indicated that the nucleoside moiety is likely derived from inosine (36). Isotopic tracer experiments established that the C-6 and C-7 positions of futasoline (8) between chorismate (4) and inosine (36) are likely derived from a C<sub>2</sub> unit derived from either pyruvate or PEP. A recombinant enzyme from *T. thermophilus* (TTHAO803) was tested for the formation of futasoline (8). However, no futasoline was formed, but instead *m*-hydroxybenzoate and 3-(1-carboxyvinyl)oxy benzoate were formed from chorismate, both in the presence and in the absence of flavin mononucleotide (FMN).

In order to demonstrate cell-free conversion of futasoline (8) to the expected hypoxanthine (36), an SCO4327 (MqnB) recombinant was prepared and assayed for enzymatic activity. However, no enzymatic activity could be detected and the failure was attributed to the inherent instability of the enzyme. Hence, these authors expressed and assayed an ortholog of SCO4327 (MqnB) from the thermophilic bacterium *T. thermophilus*. When the enzyme from this recombinant was assayed for activity, it released the expected hypoxanthine and a second product identified as dehypoxanthinylfutasoline (DHFL) (38). The latter compound supported the growth of the disruptive mutants SCO4506 (*mqnA*) and SCO4327 (*mqnB*), thus establishing DHFL as an intermediate in the pathway.<sup>41</sup> Recently, the enzyme futasoline hydrolase, responsible for the degradation of futasoline to hypoxanthine, and DHFL has been overexpressed, purified, and characterized.<sup>112</sup>

The intermediate accumulated in the medium by each of the mutants was isolated and identified by NMR and MS. As expected, the mutants SCO4550 (*mqnC*) and SCO4326 (*mqnD*) accumulated DHFL (38) and cyclic DHFL (39), respectively. However, an *in vitro* enzyme assay using an ortholog of SCO4550 (MqnC) from *T. thermophilus* (TTHA1092) failed to convert DHFL to cyclic DHFL. The lack of success in demonstrating cell-free conversions has been attributed to the lack of optimization of assay condition or requirement for additional enzymes or cofactors.<sup>41</sup>

Further studies showed that an ortholog of SCO4326 (MqnD) from *T. thermophilus* recombinant (TTHA1568) converted cyclic DHFL into DH6NA. Previous studies have established DH6NA as an intermediate in the pathway.<sup>111</sup>

### 7.12.3 Phylloquinone Biosynthesis

K is an essential component present in all oxygenic photosynthetic organisms and it is required for the functioning of photosystem I.<sup>114</sup> Structurally, K is identical to MK except that the prenyl side chain is replaced by a phytyl side chain. It was pointed out above that the pathway for the biosynthesis of K is identical to that of MK. Systematic genetic analysis in the cyanobacterium *Synechocystis* sp. has shown that K biosynthesis involves almost the same eight enzymatic steps catalyzed by the Men enzymes required for the biosynthesis of MK.<sup>18,115–118</sup> Phylloquinone-deficient mutants of *Arabidopsis* referred to as *pba* (phylloquinone absence) show impaired photosystem I activity. The affected gene is called *PHYLLLO*, and it arose by the fusion of four genes, *menF*, *menD*, *menC*, and *menH*, that are required for the biosynthesis of K in cyanobacteria and MK in prokaryotes. However, the 3' end of the *menF* module of *PHYLLLO* is truncated resulting in the loss of the C-terminal chorismate-binding domain<sup>119</sup> in the genomes of higher plants.<sup>114</sup>

The *Arabidopsis* genome contains two additional ICS genes (*ICS1* and *ICS2*). Isochorismate has been shown to be the precursor of not only K but also salicylic acid (SA), which is required for plant defense. The presence of independent *ICS1* provides an early branching point from the K biosynthetic pathway for the biosynthesis of SA under conditions of phytopathogenic attack.<sup>119,120</sup>

The truncated *PHYLLLO menF5'* module product showed homology to the N-terminal part of ICS, and hence it was important to establish whether it is functional. Hence, plants containing single and double knockouts of the genes (*ICS1* and *ICS2*) were tested. It was found that only the double knockout plants showed the expected *pba* phenotype, demonstrating that *ICS1* and *ICS2* overlap in their function for the synthesis of isochorismate required for K. This experiment in addition established that the product of the *menF5'* module is unable to complement K deficiency.<sup>114</sup> The *PHYLLLO* product was found to contain 1715 deduced amino acid residues with the four Men modules and a predicted transit peptide that directs the protein to the chloroplasts.

The fused cluster of four *men* homologs is also present in the genomes of the plants *Populus trichocarpa* and rice showing the conservation of *PHYLLO* in both monocots and dicots. Furthermore, the *PHYLLO* locus is also present in the nuclear genomes of the green algae *Chlamydomonas reinhardtii* and *Ostreococcus tauri*. The diatom alga *Tbalassiosira pseudonana* also contains a *PHYLLO* locus in which the *menH* module precedes the *menC* module. These results provide convincing evidence for the presence of fused *PHYLLO* locus in the green and red plant lineages.<sup>114</sup>

The *Arabidopsis* homologs of the bacterial MenB (At1g60550), MenA (At1g60600), MenG (At1g23360), and MenE (comprising a large gene family) have been identified by BLAST searches.<sup>114</sup>

The *menA* homolog of *Arabidopsis thaliana* gene, *AtmenA*, was recently demonstrated to be essential for K biosynthesis and accumulation of photosystem I. This gene was shown to encode 1,4-dihydroxy-2-naphthoic acid phytyltransferase, and a mutant strain on high-performance liquid chromatography analysis showed no K and contained only about 3% plastoquinone. The amounts of photosystem I and photosystem II core subunits in the mutant were significantly decreased compared to that of the wild type. Thus, the deficiency of K caused the loss of photosystem I and a partial defect of photosystem II, but did not influence the ultrastructure of the chloroplasts in young leaves.<sup>118</sup>

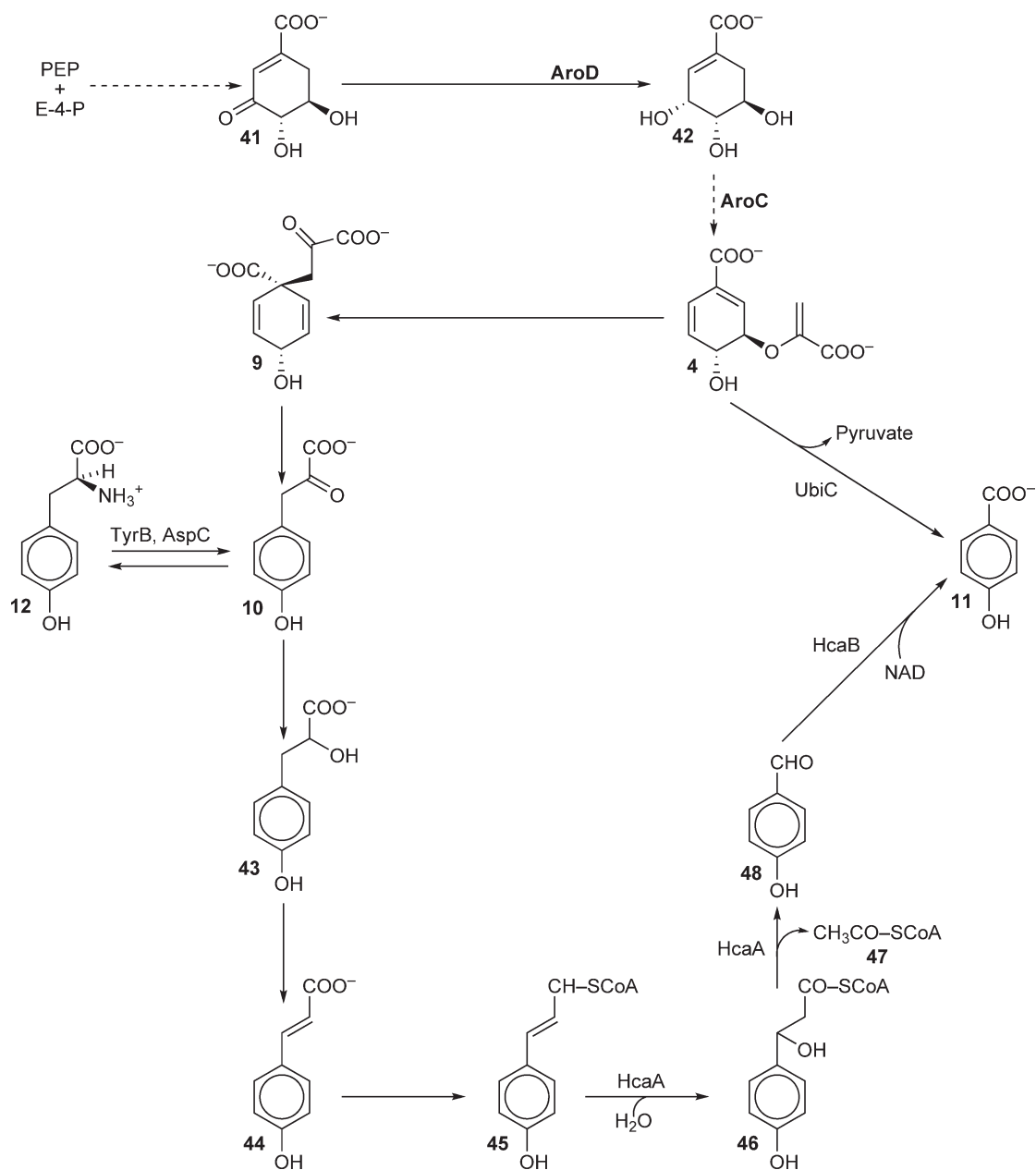
Recently, a homolog of the *menG* gene of *Arabidopsis* has been characterized and it shows sequence similarity to and functionally complements the *Synechocystis menG* mutant. An *Arabidopsis* mutant, *AtmenG*, carrying a T-DNA insertion in the gene At1g23360 is devoid of K, but contains an increased amount of demethyl 2-phytyl-1,4-naphthoquinone. K in wild type and *AtmenG* predominantly localize to photosystem I. The loss of K methylation has been shown to affect photosystem I stability or turnover and decrease in the function of photosystem I complex.<sup>121</sup> As pointed out before, a DHNA-CoA thioesterase belonging to the hotdog-fold family has been demonstrated in the cyanobacterium *Synechocystis* sp. to be essential for phyloquinone biosynthesis.<sup>101</sup>

## 7.12.4 Ubiquinone Biosynthesis

Ubiquinone (coenzyme Q) is a component of the aerobic respiratory chain. Hence, it is present in animals, plants, and aerobic and facultative Gram-negative bacteria. However, as pointed out previously, Q is absent in Gram-positive bacteria and archaea, where it is replaced by MK in the respiratory chain.<sup>2</sup> The ubiquinone biosynthetic pathway was elucidated largely due to the work of Gibson, Cox, Young, and colleagues.<sup>122,123</sup> In 1964, it was observed by Cox and Gibson<sup>42</sup> that [G-<sup>14</sup>C]-shikimate (**42**) (**Figure 12**) was incorporated into ubiquinone, thus establishing that the quinone was derived from the shikimate pathway. In their studies, Cox and Gibson used *E. coli* and *Aerobacter aerogenes* (*Klebsiella pneumoniae*) to study the role of shikimate in the biosynthesis of Q and MK.<sup>124</sup> It was found that both organisms converted shikimate to Q and MK, establishing that both compounds are derived from shikimate. When they tested an *aroC* mutant of *Aerobacter* (*Klebsiella*) for MK and Q biosynthesis, the mutant as expected failed to form MK; however, surprisingly, the same mutant formed Q. On the basis of this result, Cox and Gibson suggested that *Klebsiella* might be forming Q from the amino acid tyrosine.<sup>124</sup> In contrast, *E. coli* forms Q directly only from chorismate by the chorismate pyruvate lyase (CPL) reaction. In the following years, *E. coli* became the organism of choice for studies on Q biosynthesis in prokaryotes and *S. cerevisiae* for eukaryotic microorganisms.<sup>18,40,62,125,126</sup>

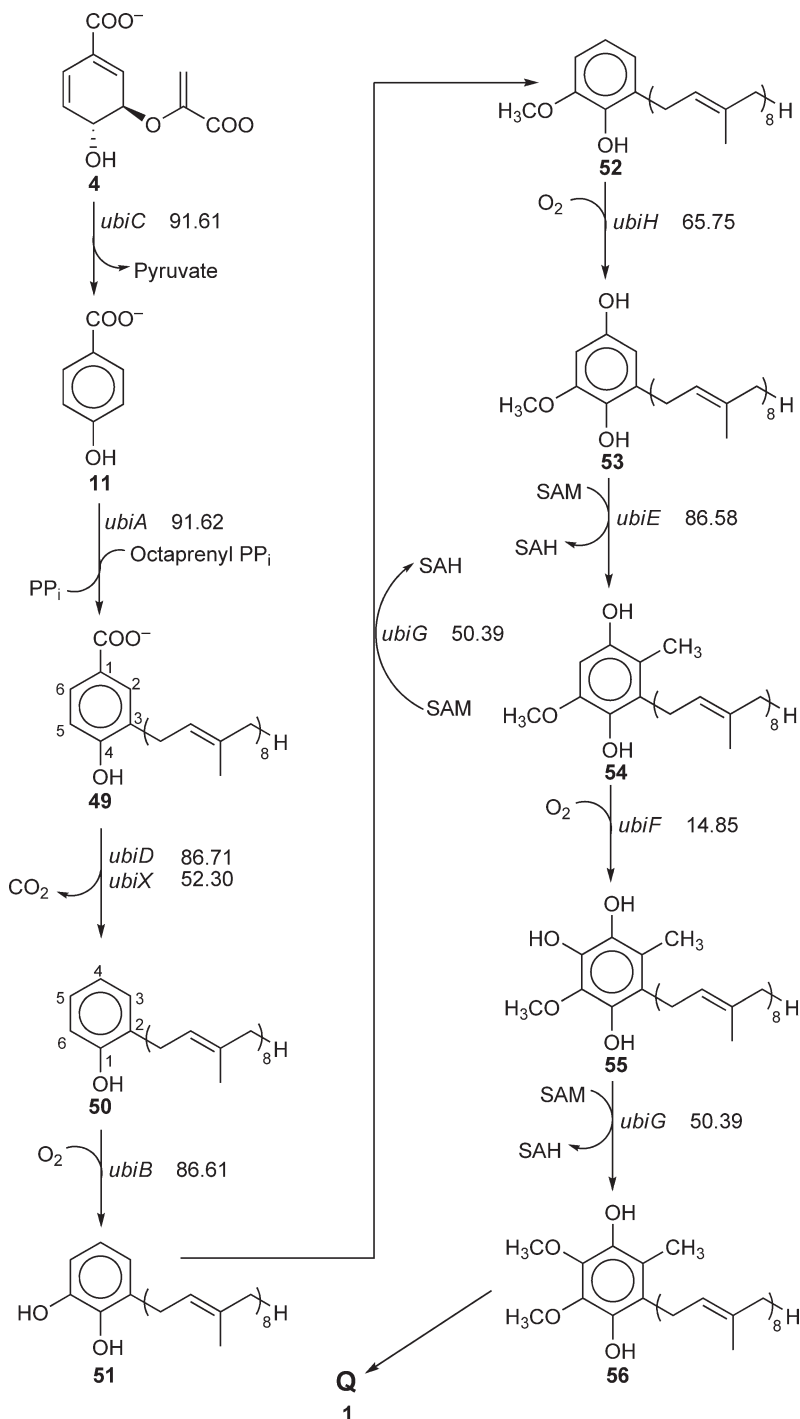
On the basis of studies with these organisms, it was generally accepted that while prokaryotes form the 4-hydroxybenzoate (4-HB) required for Q biosynthesis from chorismate, eukaryotic microorganisms can form 4-HB either from chorismate by the CPL reaction or from the amino acid tyrosine.<sup>19,127,128</sup> In contrast, higher eukaryotes, which lack the shikimate pathway, can form 4-HB only from the essential amino acid tyrosine.

Gibson and colleagues reasoned that since ubiquinone is required for aerobic electron transport, mutants deficient in its biosynthesis would grow fermentatively on glucose, but not aerobically on oxidizable substrates such as malate or succinate as the sole source of carbon and energy. Mutagenized cultures of *E. coli* were screened for the desired phenotype, and potential mutants were analyzed for the presence or absence of ubiquinone.<sup>122,123</sup> Using this procedure, a number of mutants were isolated and it was found that these mutants accumulated sufficient amounts of intermediates so that their structure could be determined by MS and magnetic resonance spectrometry.<sup>122,129</sup>



**Figure 12** Reactions leading to the formation of 4-hydroxybenzoate. (41) 3-dehydroshikimate; (42) shikimate; (4) CHA, chorismate; (9) prephenate; (10) 4-hydroxyphenylpyruvate; (12) tyrosine; (43) 4-hydroxyphenyllactate; (44) 4-hydroxycinnamate or 4-coumarate; (45) 4-coumaroyl-CoA; (46)  $\beta$ -hydroxythioester of 4-coumaroyl-CoA; (47) acetyl-CoA. (48) 4-hydroxybenzaldehyde (4-HBA); (11) 4-hydroxybenzoate (4-HB).

The biosynthesis of the quinonoid ring and the various ring modification reactions in *E. coli* have been reviewed.<sup>20,62,125,130,131</sup> A mechanistic perspective on the various reactions has been provided.<sup>17,18</sup> As pointed out in the introduction, in *E. coli* and *Salmonella* the first committed step in the biosynthesis of Q is the formation of 4-HB (11) (Figure 13) from chorismate by the cytoplasmic enzyme chorismate lyase. The 4-HB formed is attached to the membrane-bound octaprenyl diphosphate by a membrane-bound octaprenyltransferase. For the subsequent reactions, all the substrates and enzymes are in a membrane-bound complex.



**Figure 13** Ubiquinone biosynthetic pathway. Each compound in the pathway is identified by an Arabic numeral. Under anaerobic conditions, there are alternate hydroxylases for the three enzymes incorporating molecular oxygen (UbiB, UbiH, and UbiF). It should be noted that in compound (49), the chemical numbering system locates the prenyl side chain at the C-3 carbon; in compound (50) and subsequent intermediates, the prenyl side chain is assigned to C-2. Compounds (53), (54), and (55) are drawn in the quinol form. Some authors draw these structures in the quinone form. For other abbreviations, see the legend of Figure 4. The chemical names for the intermediates of the pathway are as follows: (4) chorismate; (11) 4-hydroxybenzoate; (49) 3-octaprenyl-4-hydroxybenzoate; (50) 2-octaprenylphenol; (51) 2-octaprenyl-6-hydroxyphenol; (52) 2-octaprenyl-6-methoxyphenol; (53) 2-octaprenyl-6-methoxy-1,4-benzoquinol; (54) 2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol; (55) 2-octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinol; (56) ubiquinol; (1) Q, ubiquinone. The conversion of (56) to (1) is thought to be nonenzymatic.



However, recent studies with bacteria other than *E. coli* and the availability and analysis of genomic sequences from many prokaryotic and eukaryotic microorganisms have established that there are three different routes to the soluble intermediate 4-HB (Figure 12). The three routes to 4-HB are (1) the CPL reaction, (2) the tyrosine-4-hydroxyphenylpyruvate (THP) pathway, and (3) the chorismate-prephenate-4-hydroxyphenylpyruvate (CPHP) pathway.

#### 7.12.4.1 The Conversion of Chorismate to 4-HB by the CPL (Compound (4) → (11))

The elimination of pyruvate from chorismate (4) results in the formation of 4-HB (11). This aromatizing reaction is the first committed step in Q biosynthesis in *E. coli* and *Salmonella enterica* and is catalyzed by the CPL encoded by the *ubiC* gene (Figure 14).

The *ubiC* gene from *E. coli* has been cloned; the enzyme was overexpressed and purified to homogeneity. UbiC is a monomer of 165 amino acids from which the N-terminal methionine is posttranslationally removed resulting in the functional enzyme. The enzyme has a molecular mass (MW) of 18.645 kDa and functions as a monomer. The  $K_m$  was reported to be around 6–10  $\mu\text{mol l}^{-1}$ .<sup>132,133</sup> The purified enzyme failed to accept isochorismate as a substrate but did convert 4-amino-4-deoxychorismate to 4-aminobenzoate.<sup>132</sup> Thus, it appears that the enzyme is unable to distinguish between the hydroxyl group and the amino group at the C-4 position. Walsh *et al.*<sup>65</sup> have proposed a 1,2-elimination of the elements of pyruvate for the aromatization in a manner similar to that of the AS reaction. The C<sub>4</sub>-H of chorismate is abstracted by the enzyme and loss of the C-3-enolpyruvyl group then results in the formation of 4-HB. It has been reported that the enzyme was inhibited by 4-HB but not pyruvate.<sup>133</sup> In a subsequent study, Holden *et al.*<sup>134</sup> circumvented the rapid influence of product inhibition on the initial reaction rate by using progress curve analysis of stopped-flow kinetic measurements. Under these conditions the  $K_m$  increased by about threefold to 29  $\mu\text{mol l}^{-1}$ .<sup>134</sup> The enzyme releases the pyruvate quickly and retains the 4-HB with a 10-fold higher affinity ( $K_p = 2.1 \mu\text{mol l}^{-1}$ ).<sup>135</sup>

The crystal structure of UbiC protein has been solved. The wild-type enzyme tended to aggregate and precipitate even in the presence of reducing agents and salt. To circumvent this problem, two surface-accessible cysteines at sequence positions 14 and 81 were converted by site-directed mutagenesis into serine. This mutant enzyme C14S/C81S, designated as CCSS, showed greatly improved solubility and stability with minimal effect on the catalytic properties.<sup>134</sup> The crystal structure of the enzyme from the double mutant at 1.4 Å and the wild-type enzyme at 2.0 Å in complex with the product 4-HB was determined. The core of the chorismate lyase consisted of a six-stranded antiparallel  $\beta$ -sheet without spanning helices and novel connectivity. The product, 4-HB, was shown to be bound in an internal cavity behind two flaps, which completely cover and shield the product from the solvent. Three hydrogen bonds link the product to the internally charged side chains of Arg76 and Glu155, and two additional hydrogen bonds link it to the flap atoms 34N and 114N. These five hydrogen bonds play a direct role in binding the product. There are three additional hydrogen bonds that link the flaps together and further enhance product retention.<sup>136</sup>

To further clarify and understand the unusual ligand binding and the mechanism of the reaction, additional structures of mutant enzymes, enzyme–inhibitor complexes and mutant enzyme–inhibitor complexes were studied. When a high-resolution crystal structure (1.0 Å) of the enzyme–substrate complex was examined, a substrate-sized internal cavity was found behind the flaps near the product-binding site. The crystal structure (2.4 Å) of the enzyme complexed with the inhibitor vanillate showed that the flaps were partly opened when compared to the product-bound enzyme.

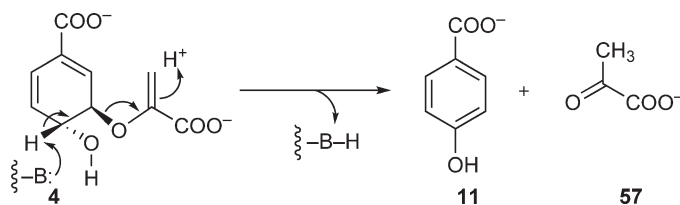


Figure 14 Proposed mechanism for chorismate lyase reaction.

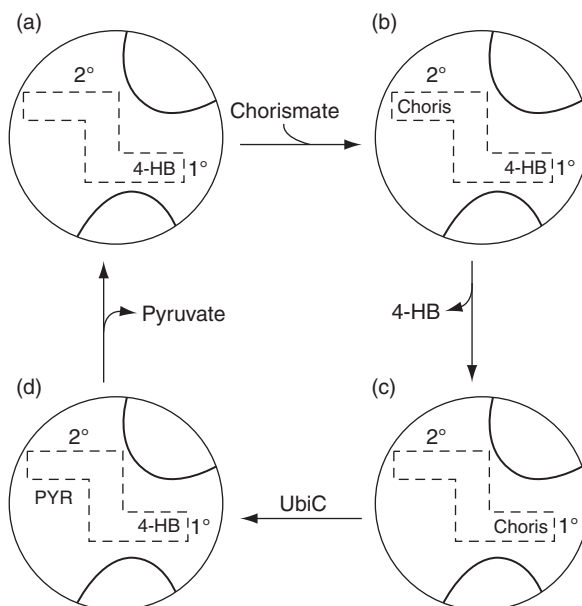
An active site mutant enzyme G90A complexed with the product was examined at a resolution of 2.0 Å. It was found that the presence of the additional methyl group in the mutant enzyme resulted in the enlargement of the 4-HB-binding pocket by about 1 Å. However, all the eight hydrogen bonds involved in product binding in the wild-type enzyme are maintained. When the kinetic properties of the mutant enzyme were compared with the wild-type enzyme, it was found that the product inhibition increased by about 40%. The wild-type enzyme had a  $K_p$  value of  $1.5 (\pm 0.2) \mu\text{mol l}^{-1}$  versus  $0.9 (\pm 0.1) \mu\text{mol l}^{-1}$  for the mutant. The increase in product inhibition in the mutant is attributed to the presence of the additional methyl group acquired in the conversion of glycine to alanine. The  $K_m$  values did not change, whereas the  $K_{\text{cat}}$  value of the mutant decreased from  $1.4 (\pm 0.2) \text{s}^{-1}$  to  $0.9 (\pm 0.4) \text{s}^{-1}$ .

When the G90A mutant enzyme was bound with the inhibitor vanillate, the structure at 1.9 Å showed two vanillate molecules. One of the vanillate molecules occupied the product site normally occupied by 4-HB and the second molecule of vanillate occupied an adjacent site or cavity. The two sites were found to be connected by a tunnel that is open partly on both ends. The product-binding site was designated as the primary ligand site and the adjacent site (additional site) where the second vanillate molecule binds was designated as the secondary ligand site.<sup>136</sup> On the basis of structural studies summarized above in combination with molecular modeling, molecular dynamics, and binding measurements with inhibitors, a model has been proposed to account for catalytic, product binding, and product release mechanisms (Figure 15).

It has been proposed that the enzyme operates by a two-site or tunnel mechanism.<sup>136</sup> According to this mechanism, the enzyme contains bound 4-HB in the primary site (designated as primary ligand site 1°). When the substrate binds to the second site (designated as secondary ligand site 2°), it promotes the release of the product from the primary site. As the product 4-HB is released from the primary site, the substrate chorismate moves to the primary site. In the primary site, the substrate is unstable and it is rapidly converted to the products 4-HB and pyruvate. Since pyruvate is small, it exits rapidly from the primary site, whereas the 4-HB is retained in the bound state and the process is repeated in a cyclic manner (Figure 15).

#### 7.12.4.2 Tyrosine-4-Hydroxyphenylpyruvate Pathway (Compound (12) → (10))

In animal cells, 4-HB is formed from the essential amino acid tyrosine. A pathway for the conversion of tyrosine to 4-HB was proposed by Booth *et al.*<sup>137</sup> following urinary excretion studies on animals administered with various phenolic acids. Based on these studies and the incorporation of radiolabeled phenolic acids into Q by



**Figure 15** Proposed two-site mechanism for the chorismate lyase catalytic cycle.

liver and yeast, the following pathway was proposed: tyrosine  $\rightarrow$  4-hydroxyphenylpyruvate  $\rightarrow$  4-hydroxyphenyllactate  $\rightarrow$  4-hydroxyphenylcinnamate  $\rightarrow$  4-HB.

The *in vitro* conversion of tyrosine (**12**) (Figure 12) to 4-hydroxyphenylpyruvate (**10**) and its subsequent reduction to 4-hydroxyphenyllactate (**43**) has been shown in the extracts of rat liver and yeast. Reactions involved in the conversion of 4-hydroxyphenyllactate (**43**) to 4-HB (**11**) are not known.<sup>18,125</sup> However, it has been suggested that 4-hydroxyphenyllactate is converted to 4-hydroxycinnamate (**44**), followed by  $\beta$ -oxidation as the CoA derivative, resulting in the formation of 4-HB<sup>24</sup> (see Figure 12, (**44**)  $\rightarrow$  (**45**)  $\rightarrow$  (**46**)  $\rightarrow$  (**48**)  $\rightarrow$  (**11**)). Evidence in support of  $\beta$ -oxidation has been obtained in tissue cultures of the plant *Litbospermum erythrorhizon*.<sup>138</sup>

It has been shown that three different species of *Klebsiella*, namely, *K. pneumoniae* subsp. *pneumoniae*, *K. aerogenes*, and *K. oxytoca*, can convert tyrosine to 4-HB.<sup>124,139</sup> It is very likely that many other bacteria and lower eukaryotes like many of the fungi have the ability to convert tyrosine to 4-HB through the TP pathway.

#### 7.12.4.3 Chorismate-Prephenate-4-Hydroxyphenylpyruvate Pathway (Compound (**4**) $\rightarrow$ (**9**) $\rightarrow$ (**10**))

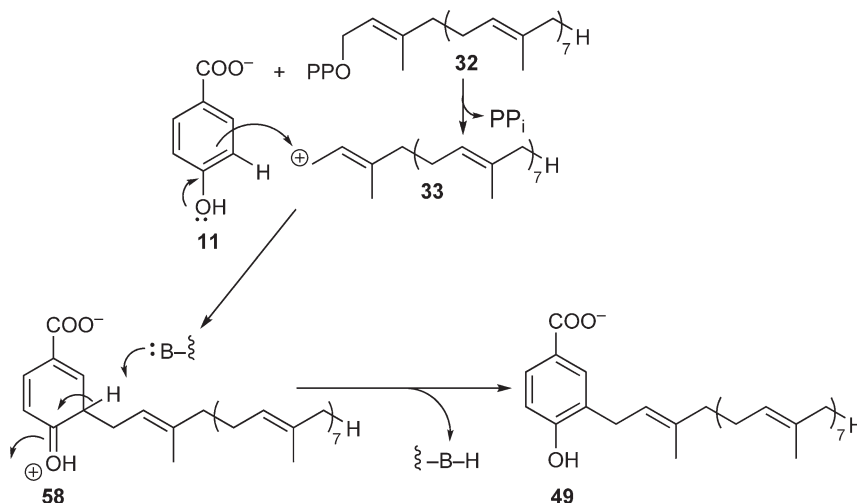
4-Hydroxyphenylpyruvate is a common intermediate in both the biosynthesis and the catabolism of tyrosine (Figure 12). Hence, it is logical to expect that many genera of bacteria and lower eukaryotes to synthesize 4-HB by converting chorismate (**4**)  $\rightarrow$  prephenate (**9**)  $\rightarrow$  4-hydroxyphenylpyruvate (**10**) and the subsequent reactions as shown in Figure 12. For these studies, a *K. oxytoca*  $\Delta$ *aroD* mutant was used as the parent strain for the isolation of mutants blocked in the conversion of tyrosine (**12**) to 4-HB (**11**). The ideal choice would have been a  $\Delta$ *ubiC* mutant blocked in the direct conversion of chorismate (**4**) to 4-HB (**11**). However, such a mutant strain is unsuitable for these studies due to the inherent highly unstable nature of chorismate, which spontaneously breaks down to 4-HB and pyruvate.<sup>65,140</sup> Mutants blocked in the second reaction (*ubiA*) were deemed unsuitable since it involves the attachment of the isoprenoid side chain, which will be required for the prenylation of the 4-HB (Figure 13, (**11**)  $\rightarrow$  (**49**)).

Hence, as discussed above, in an *aroD* mutant of *Klebsiella* when the conversion of tyrosine (**12**) to 4-hydroxyphenylpyruvate (**10**) was blocked (*tyrB*, *aspC*), the mutant formed **Q**, when the medium was supplemented with 4-hydroxyphenylpyruvate or 4-hydroxyphenyllactate.<sup>19,139</sup> Furthermore, analysis of the genomes of prokaryotes and lower eukaryotes revealed that a number of bacteria such as *Xanthomonas oryzae*, *Bradyrhizobium japonicum*, *Nitrobacter winogradskyi*, *Agrobacterium tumefaciens*, *Paracoccus denitrificans*, *Chromobacterium violaceum*, and *Legionella pneumophila*, and fungi such as *S. cerevisiae*, *Schizosaccharomyces pombe*, *Neurospora crassa*, and *Eremothecium gossypii* lack the gene encoding CPL (homolog of *ubiC*). It is likely that those prokaryotes and lower eukaryotes that lack CPL use the CPHP pathway. For a complete listing of the organisms, the SEED database should be consulted.<sup>141</sup>

#### 7.12.4.4 Prenylation of 4-Hydroxybenzoate (Compound (**11**) $\rightarrow$ (**49**))

The prenylation of 4-HB (**11**) to 3-octaprenyl-4-hydroxybenzoate (**49**) (Figure 13) is carried out by the enzyme 4-hydroxybenzoate octaprenyltransferase encoded by the *ubiA* gene. The enzyme is membrane bound and requires octaprenyl diphosphate and  $Mg^{2+}$ .<sup>142</sup> In addition to octaprenyl diphosphate, the enzyme could incorporate geranyl, farnesyl, phytyl, or solanesyl diphosphate as a side chain precursor.<sup>107,143</sup> This lack of specificity also extends to the aromatic substrate; thus, 4-aminobenzoate can replace 4-HB as a substrate.<sup>107</sup> Recently, it has been shown that the enzyme accepts a wide variety of benzoic acid derivatives as substrates. As already mentioned, replacing the C-4 hydroxyl with an amino group did not affect reactivity. However, replacing the hydroxyl with a methoxy group was not tolerated. Compounds substituted at C-5 with OH,  $NH_2$ , Cl, or  $CO-CH_3$  groups were used as substrates by the enzyme. Similarly, compounds with hydroxyl groups at C-4, C-5, and C-6 or hydroxyl group at C-4, C-6 and methyl group at C-5 were used as substrates.<sup>144</sup>

The prenyl transfer reactions are electrophilic substitution reactions. The reaction mechanism probably includes a carbocation<sup>145</sup> (Figure 16), evidence for which comes from studies on the related enzyme dimethyltryptophan synthase.<sup>146</sup>



**Figure 16** Proposed mechanism for the prenylation of 4-hydroxybenzoate (11) to 3-octaprenyl-4-hydroxybenzoate (49).

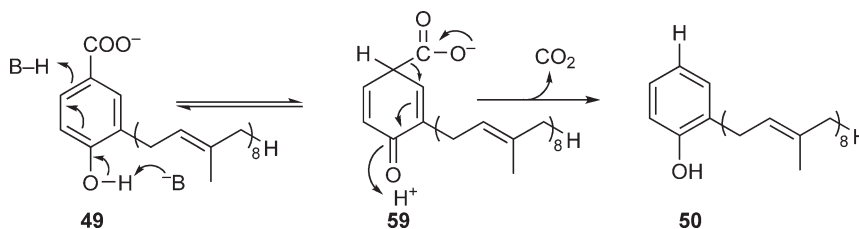
#### 7.12.4.5 Formation of 2-Octaprenylphenol (Compound (49) → (50))

The conversion of 3-octaprenyl-4-hydroxybenzoate (49) to 2-octaprenylphenol (50) was demonstrated by Cox *et al.*<sup>147</sup> The enzyme responsible for this conversion was named 3-octaprenyl-4-hydroxybenzoate decarboxylase. The presence of a decarboxylase was also observed by El Hachimi *et al.*<sup>107</sup> This enzyme activity was absent in *ubiD* mutants.<sup>147</sup>

When cell extracts were prepared using a French press, centrifuged at 30 000 ×g, and the supernatant further centrifuged at 150 000 ×g for 3 h, most of the activity remained in the soluble fraction, establishing that the enzyme separated from the membrane. A 24-fold purified preparation of the enzyme was obtained. The MW of the enzyme was reported to be  $M_r$  340 000.<sup>148</sup> For optimal activity, the enzyme required  $Mn^{2+}$ , washed membranes or an extract of phospholipids, and an unidentified heat stable factor of MW <10 000. The reaction was strongly stimulated by dithiothreitol and methanol. Since the substrate of the enzyme 3-octaprenyl-4-hydroxybenzoate (49) is membrane bound and the enzyme is stimulated by phospholipid, it has been suggested that the enzyme normally functions in association with the cytoplasmic membrane *in vivo*.<sup>148</sup> A reaction mechanism has been suggested (Figure 17).<sup>17,18</sup>

A number of *ubiD* mutants studied formed about 20% of the wild-type levels of Q<sub>8</sub>, indicating that either the mutants are leaky or there is an alternate enzyme for the reaction. However, the significance of any alternate carboxy-lyase in the wild-type strains has been questioned.<sup>148</sup>

An alternate 3-octaprenyl-4-hydroxybenzoate decarboxylase encoded by the *ubiX* gene has been described in *S. typhimurium*, which carries out the same reaction as the *ubiD*-encoded enzyme.<sup>149</sup> An *ubiX* gene showing 70% homology to the *S. typhimurium* gene has been identified in *E. coli*.<sup>150,151</sup> The *orf*s encoding the two enzymes UbiD and UbiX have been identified from *E. coli*.<sup>152</sup> Recently, a report has appeared suggesting that both UbiD and UbiX are required for the decarboxylation of 3-octaprenyl-4-hydroxybenzoate particularly during the logarithmic phase of growth.<sup>153</sup>



**Figure 17** Proposed mechanism for 3-octaprenyl-4-hydroxybenzoate decarboxylase reaction.

It has been reported that several *E. coli* strains, including the enterohemorrhagic O157:H7, contain, in addition to UbiX, a second paralog designated as Pad1. The amino acid sequence of this paralog was reported to have a 52% identity to UbiX and a slightly higher identity to *S. cerevisiae* phenylacrylic acid decarboxylase Pad1. The exact biochemical role of *E. coli* Pad1 remains to be determined.<sup>154</sup>

#### 7.12.4.6 Hydroxylation and Methylation Reactions

In the subsequent steps of the pathway, the 2-octaprenylphenol (**50**) undergoes three hydroxylation reactions alternating with three methylation reactions resulting in the formation of ubiquinol (**56**) and then Q (**1**). For convenience, the hydroxylation reactions are considered together and this will be followed by a description of the three methylation reactions.

##### 7.12.4.6.1 Hydroxylation reactions

Three flavin-linked monooxygenases are involved in the three hydroxylation reactions of the pathway with three hydroxyl groups being introduced at positions C-6, C-4, and C-5 of the benzene nucleus, respectively. The three reactions are (1) compound (**50**) → (**51**), (2) compound (**52**) → (**53**), and (3) compound (**54**) → (**55**). Mutants blocked in each of these hydroxylation reactions were isolated and designated as *ubiB*, *ubiH*, and *ubiF*, respectively.

Consistent with their metabolic block, *ubiB* mutants accumulate 2-octaprenylphenol (**50**).<sup>147,155</sup> However, the predicted product of the UbiB reaction, 2-octaprenyl-6-hydroxyphenol (**51**), has never been isolated and characterized, and it may not occur as a free intermediate.<sup>155,156</sup>

As part of the *E. coli* genome project, when the sequence was annotated, *ubiB* was considered identical to that of *fre* and *luxG*.<sup>157</sup> In subsequent studies, an *orf* previously designated *yigR* was identified as *ubiB*. An insertion mutant was isolated and shown to accumulate the expected intermediate in the pathway, 2-octaprenylphenol (**50**). As mentioned above, the expected product of the reaction 2-octaprenyl-6-hydroxyphenol (**51**) could not be isolated.<sup>158</sup>

Mutants blocked in the methylation of 2-octaprenyl-6-hydroxyphenol (**51**) to 2-octaprenyl-6-methoxyphenol (**52**) have been isolated (*ubiG::kan*) (see Section 7.12.4.6.2). However, surprisingly, these mutants also failed to accumulate the expected intermediate before the block 2-octaprenyl-6-hydroxyphenol (**51**), thus supporting the suggestion that it may not occur as a free intermediate.<sup>155,156</sup>

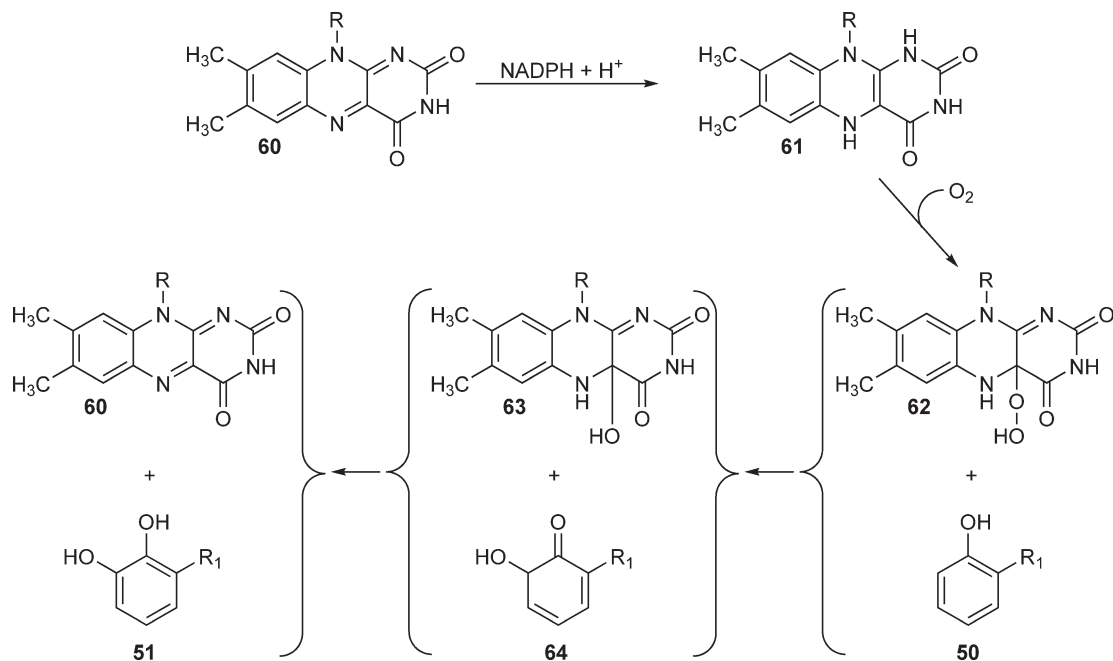
Mutants unable to convert (**52**) to (**53**) have been isolated with the gene being designated as *ubiH*.<sup>155</sup> The *ubiH* gene is identical to the *visB* gene and confers a photosensitive phenotype due to the accumulation of (**52**).<sup>159</sup>

The final hydroxylation in Q biosynthesis is the conversion of (**54**) to (**55**) and mutants blocked in the reaction were isolated and characterized. As expected, these mutants, designated as *ubiF*, accumulated (**54**), which was isolated and identified.<sup>160</sup> The *ubiF* gene was identified as *orf391* and the product accumulated by insertion mutants in this *orf* was found to be (**54**).<sup>161</sup>

Under aerobic conditions, the origin of the oxygen atoms of Q was determined by <sup>18</sup>O labeling experiments. Cultures were grown on the oxidizable carbon source succinate, under strictly aerobic conditions in a defined atmosphere of <sup>18</sup>O<sub>2</sub>. Q was isolated from these cultures and subjected to mass spectral analysis. The spectrum showed several prominent peaks with *m/z* values differing from that of normal Q by +6 establishing that <sup>18</sup>O had been incorporated. Furthermore, it was demonstrated that <sup>18</sup>O was incorporated at positions 4, 5, and 6.<sup>156</sup>

The nature of the hydroxylation reactions discussed above has been investigated. A *bemA* mutant defective in the biosynthesis of cytochromes was able to convert 2-octaprenyl-[<sup>14</sup>C] phenol (**50**) to <sup>14</sup>C-labeled Q-8, ruling out the involvement of the cytochrome P-450 monooxygenase system and suggesting the involvement of flavin-linked monooxygenases in these reactions.<sup>162</sup> A mechanism analogous to that proposed for the flavin-dependent tyrosine hydroxylase<sup>163</sup> has been suggested.<sup>17,18,40</sup> (Figure 18)

When grown anaerobically, with glycerol as a carbon source and fumarate as an electron acceptor, *E. coli* forms considerable quantities of Q (50–70% of the level found in aerobically grown cells). Mutants blocked in the various nonhydroxylating reactions of the pathway, such as *ubiA*, *ubiD*, and *ubiE*, remain Q deficient under both aerobic and anaerobic conditions, establishing that the same genes and enzymes participate under both aerobic and anaerobic conditions.<sup>164</sup>



**Figure 18** Proposed mechanism for 2-octaprenylphenol monooxygenase reaction.

In contrast, the three groups of mutants blocked in the three oxygenases discussed above, *ubiB*, *ubiH*, and *ubiF*, were able to synthesize Q under anaerobic conditions providing evidence that specific hydroxylases are involved in the anaerobic pathway.<sup>164</sup> These hydroxylases likely derive the hydroxyl groups from the solvent H<sub>2</sub>O similar to EntC and MenF reactions discussed above.

#### 7.12.4.6.2 Methylation reactions

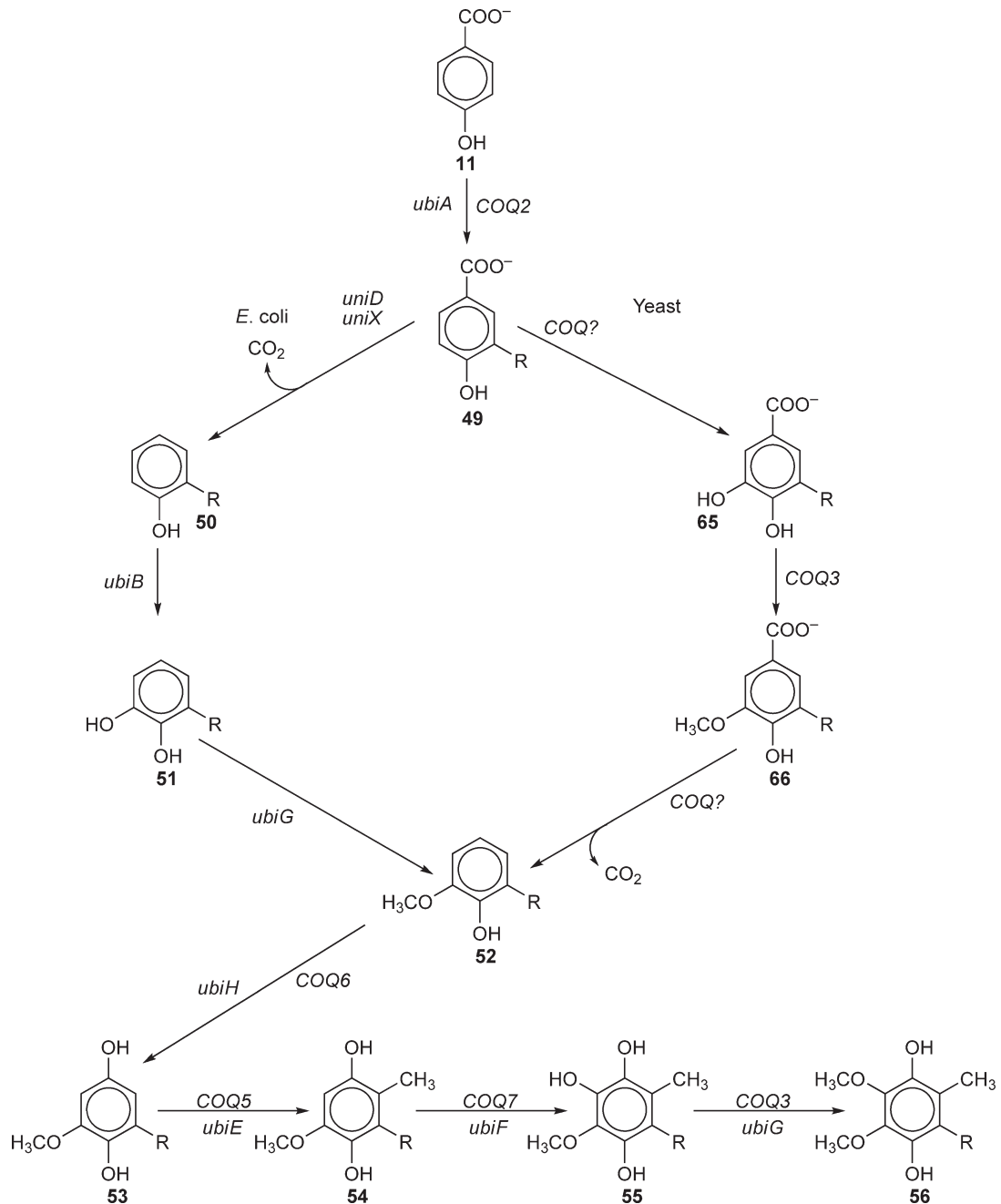
Two methylations on O and one on C involved in the pathway are (1) compound (51) → (52) (i.e., O-methylation), (2) compound (53) → (54) (i.e., C-methylation), and (3) compound (55) → (56) (i.e., O-methylation).

The methylation steps alternate with the three hydroxylations described above introducing methyl groups at 6-OH, at the ring C-3, and at the 5-OH group, respectively. The three methyl groups are derived from methionine, with SAM being the actual methyl donor.<sup>165</sup>

The C-methylase responsible for the methylation of ring C-3 is encoded by the *ubiE* gene. Mutants blocked in the methylation accumulate the substrate of the enzyme, 2-octaprenyl-6-methoxy-1,4-benzoquinol (53).<sup>160</sup> The UbiE enzyme is nonspecific and carries out the methylation of the MK intermediate, DMK (3) → MK (2), in addition to its role in the methylation of (53) → (54)<sup>110</sup> (discussed in MK biosynthesis 7.12.2.16).

During the screening for mutants blocked in the O-methylation reactions, mutants blocked in the methylation of 6-OH were not obtained. However, mutants blocked in the methylation of the 5-OH were isolated, designated as *ubiG*, and shown to accumulate compound (55), which was isolated and characterized.<sup>129</sup> Furthermore, the *ubiG* mutants were leaky and formed about 10% of the wild-type levels of Q.<sup>164</sup> In subsequent studies, it was reported that the O-methylase encoded by the *ubiG* gene is nonspecific and that it carries out the methylation of both 6-OH and 5-OH groups.<sup>166</sup> This lack of specificity also extends to the presence of other groups on the benzoquinone ring; the enzyme in addition methylates 3,4-dihydroxy-5-hexaprenylbenzoquinol (65) (Figure 19) to 3-methoxy-4-hydroxy-5-hexaprenylbenzoquinol (66). The reported leakiness of the *ubiG* mutant likely allowed sufficient intermediate (51) to be methylated at the 6-OH resulting in the formation of (52), which was subsequently converted to (55) and methylated at the 5-OH resulting in the formation of (56) and Q. A *ubiG::kan* mutant has been isolated. However, surprisingly, this mutant failed to accumulate the expected intermediate before the block





**Figure 19** Comparison of Q biosynthesis in *E. coli* and yeast. *E. coli* and yeast contain Q-8 and Q-6, respectively. The side chain of the compounds is represented by the letter 'R'. 'R' for *E. coli* is octaprenyl and for yeast it is hexaprenyl. The names of the intermediates for *E. coli* are shown in the legend of **Figure 13**. For yeast, (65) is 3,4-dihydroxy-5-hexaprenylbenzoquinol and (66) is 3-methoxy-4-hydroxy-5-hexaprenylbenzoquinol. Other compounds in the pathway should contain the prefix 'hexaprenyl' in place of 'octaprenyl.'

2-octaprenyl-6-hydroxyphenol (51). Two possible reasons have been advanced for the failure to detect compound (51). First, as mentioned above, compound (51) may not occur as a free intermediate.<sup>155,156</sup> Second, it has been suggested that the compound may be highly reactive due to the presence of the catechol moiety and hence degraded.<sup>166</sup>

### 7.12.5 Organization of Q Biosynthetic Enzymes into a Complex

As discussed above, not all the enzymes involved in Q biosynthesis have been studied in cell-free extracts. Among the enzymes studied, CPL (UbiC) is a cytoplasmic enzyme, whereas 4-hydroxybenzoate octaprenyltransferase (UbiA) is firmly membrane bound. Two other enzymes that have been studied, 3-octaprenyl-4-hydroxybenzoate carboxy-lyase and 2-octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinol methyltransferase, are considered to be normally associated with the membrane.<sup>148,167</sup> The association of enzymes with membrane is supported by the isolation of a 2-octaprenyl [U-<sup>14</sup>C] phenol (**50**) charged enzyme complex of  $M_r$   $2 \times 10^6$  containing at least 12 proteins ranging from 40 000 to 80 000  $M_r$  from cells grown anaerobically on glycerol/fumarate medium in the presence of 4-hydroxy[U-<sup>14</sup>C] benzoate. When this complex was incubated with SAM, NADH, NADPH, Mg<sup>2+</sup>, and a cytoplasmic enzyme of  $M_r$  of about 20 000 (probably a methyltransferase)<sup>168</sup> in the presence of oxygen, all of the <sup>14</sup>C-labeled phenol was converted to Q.<sup>169</sup> This complex, therefore, contains the oxygen-dependent Q-8 biosynthetic apparatus. In anaerobically grown cells, this apparatus, which is charged with 2-octaprenylphenol, may be kept in a standby position. When oxygen becomes available, Q-8 biosynthesis can be effectively turned on.<sup>168,169</sup> Since this complex was isolated without detergent treatment, it was thought that it had broken from the membrane as a distinct and native domain. This complex contains in addition to a high level of 2-octaprenylphenol and low levels of Q, phospholipid, and other membrane proteins.<sup>168,169</sup>

Based on studies with a thiol-sensitive mutant (ICS16), it was reported that there is genetic evidence for interaction between UbiX and UbiG proteins.<sup>170</sup> The ICS16 mutant had point mutations resulting in change of a single amino acid in UbiX (S98R) and UbiG (L132Q) when compared to the sequence of the same two proteins in *E. coli* K<sub>12</sub>. Complementation of this mutant with either *ubiX* from *E. coli* K<sub>12</sub> strain (*ubiX* K<sub>12</sub>) or *ubiG* K<sub>12</sub> restored the wild-type phenotype. In contrast, while an *ubiG* insertion mutant was rescued by complementation by *ubiG* K<sub>12</sub>, it was not rescued by *ubiX* K<sub>12</sub>.<sup>170</sup> Rescue of Q-deficient phenotypes can be achieved by levels of Q that are significantly lower than that present in the wild-type strains.<sup>170</sup> These studies were cited as providing supporting evidence for the polypeptide complex described by Knoell<sup>168</sup> discussed above.

### 7.12.6 Comparison of Q Biosynthesis in Yeast and *Escherichia coli*

As pointed out above, *E. coli* synthesizes 4-HB directly by the CPL encoded by the *ubiC* gene. However, analysis of the genomic sequence of the yeast *S. cerevisiae* has shown the absence of a homolog for this gene. It has been demonstrated previously that mutants of yeast blocked in the shikimate pathway are able to synthesize Q from the aromatic amino acid tyrosine.<sup>24,128</sup>

Since, unlike higher eukaryotes, yeast and other lower eukaryotes have the shikimate pathway, they can synthesize Q from chorismate (**4**) by the prephenate (**9**) → 4-hydroxyphenylpyruvate pathway (**10**) or by the tyrosine (**12**) → 4-hydroxyphenylpyruvate pathway (**10**) (Figure 12), similar to those prokaryotes that lack the gene encoding CPL (the homolog of *E. coli ubiC*).

*S. cerevisiae* mutants blocked in the conversion of 4-HB to Q (Figure 19) are unable to grow on oxidizable carbon sources. An extensive genetic analysis of these mutants led to the identification of the various steps involved in the conversion of 4-HB to Q. Nine complementation groups designated (COQ1–COQ9) were identified among the mutants unable to form Q.<sup>21,171</sup> These mutants proved to be a very valuable resource for the subsequent molecular studies in yeast and in the identification of homologous genes in higher eukaryotes by complementation.

Furthermore, the studies in yeast have been greatly aided by the knowledge of the *E. coli* pathway and the availability of well-characterized *ubi* mutants for complementation studies. A comparative diagram of the Q biosynthetic pathway in yeast and *E. coli* is outlined in Figure 19. For details of the cofactors, reactions, and reaction mechanisms, see Figures 12–18. The Q biosynthetic pathway in yeast and other eukaryotes diverge from that of *E. coli* after the formation of 3-polyprenyl-4-hydroxybenzoate (**49**). In yeast, this intermediate accumulates in fourfold higher concentration than the final product of the pathway.<sup>172</sup> As pointed out above (Figure 19), in *E. coli*, the biosynthesis of Q involves decarboxylation of 3-polyprenyl-4-hydroxybenzoate

followed by hydroxylation and *O*-methylation. In yeast, the common intermediate (49) undergoes hydroxylation and *O*-methylation before decarboxylation (Figure 19). The three reactions, described above in different order in yeast and *E. coli*, converge with the formation of 2-polyprenyl-6-methoxyphenol (52).

Of the nine complementation groups, *coq2*–*coq8* are involved in the synthesis of the benzoquinone, while *coq1* participates in the synthesis of the prenyl side chain. The isopentenyl diphosphate (IPP), required for the biosynthesis of the side chain precursor, hexaprenyl diphosphate (HexPP), for the yeast Q-6 unlike that of *E. coli*, is derived from the mevalonate pathway.<sup>29,173,174</sup>

As already mentioned above, the *COQ1* gene encodes the HexPP synthase.<sup>175</sup> To determine the effect of the length of the prenyl side chain of Q on its function, polyprenyl diphosphate synthase genes encoding prenyl diphosphates ranging in length from C<sub>25</sub> to C<sub>50</sub> were expressed in the *coq1* mutant. The expression of the genes required the mitochondrial leader sequence. Analysis of the Q composition revealed that the mutant produced Q-5 to Q-10 depending on the complementing polyprenyl diphosphate synthase gene. The strains carrying Q with different side chain lengths grew as well as the wild-type strain on the oxidizable carbon source glycerol. These results clearly demonstrated that the side chain length has no influence on the *in vivo* function of Q.<sup>176–178</sup>

The enzyme 4-hydroxybenzoate hexaprenyl transferase (Figure 19, (11) → (49)) is encoded by the *COQ2* gene.<sup>175</sup> The protein shows a 31.4% identity and 67.9% similarity to the UbiA. Furthermore, *COQ2* can complement *ubiA* mutants.<sup>179</sup> The 4-hydroxybenzoate polyprenyl transferase is nonspecific and as discussed above can incorporate prenyl diphosphates ranging in length from C<sub>25</sub> to C<sub>50</sub> synthesized *in vivo* by complementing the *coq1* mutant by the respective genes.

Homologs of *COQ2* have been isolated from the plant species *A. thaliana*<sup>180</sup> and *Oryza sativa*,<sup>181</sup> the fission yeast *S. pombe*<sup>182</sup> as well as *Homo sapiens*.<sup>183</sup> The polyprenyl transferases irrespective of the organism are indifferent to the chain length of the polyprenyl diphosphate.<sup>18,19,178,184</sup>

The *COQ3* gene encodes a methyltransferase having 40% identity to the UbiG of *E. coli*, and both enzymes contain the four motifs common to SAM-dependent methyltransferases. The yeast and the *E. coli* enzymes have been studied using synthetic farnesylated substrate analogs. It was found that both the enzymes carry out the three *O*-methylation reactions (Figure 19), namely, (51) → (52), (55) → (56), and (65) → (66).

It was further demonstrated that the *ubiG* gene can restore respiration in the yeast *coq3* mutant, provided that the *ubiG* is modified to contain the mitochondrial leader sequence.<sup>166,185</sup> Thus, it appears that both the yeast and the *E. coli* enzymes tolerate substitutions on the benzene ring.

By the functional complementation of the *COQ3* mutants, homologs have been identified in humans,<sup>130</sup> *A. thaliana*,<sup>186</sup> and the rat.<sup>187</sup> These homologous proteins, irrespective of the organism from which they are derived, contain the four motifs common to methyltransferases utilizing SAM as the methyl donor.<sup>126,188</sup>

The *COQ5* enzyme carries out the C-methylation of (53) → (54) and shows 44% identity to UbiE of *E. coli*. Both *Coq5* and UbiE sequences contain the motifs common to SAM-dependent methyltransferases. *E. coli ubiE* was able to complement a *coq5* mutant when fused with a fragment encoding the mitochondrial leader sequence expressed from the yeast *CYCl* promoter. The functional complementation of the *coq5* mutant by the *ubiE* gene provides conclusive evidence that the two gene products carry out the same reaction.<sup>189,190</sup>

In a subsequent study, a collection of *coq5* mutant strains was characterized. All the five mutant strains were rescued by the yeast *COQ5* gene. In contrast, only two of the mutant strains were complemented by *E. coli ubiE*. The two mutants that were complemented by *E. coli ubiE* contained point mutations within or near the conserved SAM-binding methyltransferase motifs.

When the mitochondrial proteins were fractionated and analyzed, it was found that the *Coq5* protein is associated with the inner mitochondrial membrane on the matrix side.<sup>191</sup> Furthermore, the two point mutants described above contained the same amount of inactive polypeptides as the wild-type *Coq5P*.

In contrast, the mutants that were not complemented by the *E. coli ubiE* contained greatly reduced levels of not only *Coq5* polypeptides but also *Coq3* and *Coq4* polypeptides. An analysis of these mutants showed that one of them was a null mutant (*coq5*); the second mutant can be considered a null since it contained a stop codon at position 93 (*coq5-4*). However, the third mutant had a single amino acid substitution at the fourth amino acid from the C-terminus (G304D). On the basis of these results, it was concluded that at least two functions are carried out by *Coq5P*. In addition to carrying out the C-methylation, it is also involved in the

stabilization of Coq3 and Coq4 peptides, thus providing supporting evidence for the existence of a complex involved in Q biosynthesis.<sup>126,191</sup>

The promoter region of COQ5 contains consensus binding sites for transcription factors Ger1, Mig1, Rtg1/2/3, and Hap2/3/4.<sup>192,193</sup> Consistent with the role of Q in aerobic metabolism, the expression of COQ5 is inducible by the oxidizable carbon sources glycerol and oleic acid compared to that of the fermentable carbon source glucose.<sup>192,193</sup>

The COQ6 gene product shows 38% identity and 51% similarity to that of the *ubiH*-encoded oxygenase of *E. coli*. COQ6 is considered a nonessential gene. However, mutants blocked in the gene are impaired in growth on oxidizable carbon sources. The Coq6 protein is associated with the inner mitochondrial membrane on the matrix side.<sup>194</sup> The Coq6 protein contains three conserved motifs that are common features of a large family of flavin adenine dinucleotide (FAD)-binding aromatic hydroxylases.<sup>195</sup> The three motifs are adenosine diphosphate (ADP) binding, FAD/nicotinamide adenine dinucleotide phosphate (NAD(P)H) binding, and a motif that binds the ribityl moiety of FAD. On the basis of the presence of these motifs, the Coq6p has been considered a flavin-dependent monooxygenase presumed to introduce hydroxyl groups to either 4-hydroxy-3-polyprenylbenzoic acid (49) or 6-methoxy-2-polyprenylphenol (52) or both.<sup>194,196</sup>

The COQ7 gene has been characterized biochemically and a *coq7* mutant accumulates 2-hexaprenyl-4-hydroxybenzoate (49) and 2-hexaprenyl-3-methyl-6-methoxy-1,4-benzoquinone (54). Hence, it is presumed to correspond to the *ubiF* gene of *E. coli*.<sup>161</sup> The COQ7 gene was also isolated as the *CAT5* gene that is involved in the release of gluconeogenic genes from repression by glucose.<sup>197</sup> However, database searches using the amino acid sequence of the COQ7 gene as query revealed no similarity to any of the known *ubi*-encoded amino acid sequences of *E. coli*.<sup>187</sup> The *ubiF* gene of *E. coli* has been identified and shown to catalyze the conversion of (54) → (55).<sup>198</sup> Furthermore, *ubiF* mutants were found to accumulate (54). Surprisingly, UbiF shows 33% identity and 48% similarity to UbiH and 29% identity and 42% similarity to COQ6-encoded protein but not to COQ7-encoded protein.<sup>161</sup>

In spite of lack of homology, *E. coli* UbiF was able to functionally complement very efficiently a *coq7* point mutant (E194K) and, to a lesser extent, a null mutant. Recent mitochondrial fractionation studies have shown that Coq7 protein is peripherally associated with the inner mitochondrial membrane on the matrix side.<sup>196</sup>

Mutants of yeast lacking the protein encoded by the COQ8 gene are unable to synthesize Q. Functionally, COQ8 is conserved in a wide variety of organisms representing both prokaryotes and eukaryotes.<sup>199,200</sup> A *coq8* deletion mutant has been shown to be functionally complemented by homologs from *Arabidopsis* and *Schizosaccharomyces*.<sup>186,201</sup>

*E. coli ubiB*, which is considered a homolog of COQ8, encodes the first monooxygenase involved in the biosynthesis of Q. It has been shown that *menB* mutants accumulate the intermediate preceding the block, namely, 2-octaprenylphenol (50).<sup>185</sup> As pointed out before, UbiB shows no sequence homology to the other two monooxygenases of the pathway, namely, UbiF and UbiH. However, both Coq8p and UbiB belong to the same large super family of proteins that contain motifs common to eukaryotic-type protein kinases.<sup>199,202</sup>

As discussed above, a specific enzymatic function can be attributed to Coq1, Coq2, Coq3, Coq5, Coq6, and Coq7 proteins, but it is not known what exact role the other Coq proteins play.<sup>130,196</sup>

Genetic and biochemical studies have shown that *E. coli* mutants blocked in the Q biosynthetic pathway tend to accumulate the biosynthetic intermediate immediately preceding the genetic block. In contrast, yeast mutants blocked in any of the genes from *coq3* to *coq9* accumulate the very same intermediate 3-hexaprenyl-4-hydroxybenzoic acid (49), the product of the Coq2 enzyme. Further studies with these mutants showed that when there is a mutation in any *coq* gene, the levels of Coq3, Coq4, Coq6, Coq7, and Coq9 polypeptides are decreased. Biochemical analysis of mitochondrial proteins established that while Coq2 is the sole integral membrane protein, the other Coq polypeptides are located within the mitochondrial matrix and are peripherally associated with the inner membrane.<sup>126,203</sup>

By size exclusion chromatography, it was shown that Coq3, Coq4, Coq6, and Coq7 proteins elute in a high molecular mass (MW) complex of >700 kDa with *O*-methyltransferase activity and DMQ6 (55). Coq10p was shown to be eluted together with Coq4p by gel filtration, but not with Coq3p and Coq5p. Evidence for a putative Q-biosynthetic complex of MW of >1 MDa was provided by native gel electrophoresis. The complex contained Coq3, Coq4, Coq7, and Coq9 proteins as constituents. Coq7p seems to stabilize other Coq proteins.<sup>126</sup>

Coimmunoprecipitation was used to demonstrate the physical interaction between the Coq3, Coq4,<sup>204</sup> and Coq9 proteins with that of the Coq4, Coq5, Coq6, and Coq7 proteins.<sup>126,203</sup>

In a recent study, by two-dimensional (2D)-Blue Native (BN)/sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), it was demonstrated that Coq3p, Coq5p, and Coq9p are present in a high MW complex of about 1.3 MDa,<sup>199</sup> which is probably the same complex as that described previously as 1 MDa.<sup>199</sup>

Analysis of the 2D-BN/SDS-PAGE separation profiles of Coq2, Coq8, and Coq10 showed that these proteins formed a complex of 500 kDa and are not part of the 1.3 MDa Q-biosynthetic complex. These authors suggested that these proteins represent either an independent second complex or, alternatively, an intermediate form of the 1.3 MDa complex.<sup>199</sup> However, it should be pointed out that none of the three proteins are directly involved in the conversion of (49) to Q. As pointed out above, Coq2p catalyzes the formation of (49), which is subsequently used as the substrate for Q biosynthesis.<sup>175</sup> Consistent with this is the fact that (49) is accumulated by every mutant of the Q-biosynthetic complex (coq3–coq9).<sup>126,171</sup> It has been suggested that Coq2 enzyme activity is independent of the Q-biosynthetic complex. Coq10 has been shown to be an integral membrane protein that is not involved in Q synthesis. Further,  $\Delta$ coq10 mutants contain the same level of Q as the wild type.<sup>126,205</sup>

### 7.12.7 Biosynthesis of Isoprenoid Side Chain of MK and Q

*E. coli* and other Gram-negative bacteria synthesize the IPP and dimethylallyl diphosphate (DMAPP) by the mevalonate-independent pathway, also known as the nonmevalonate pathway (other names are deoxyxylulose phosphate or methylerythritol phosphate pathway). In contrast, Gram-positive bacteria and eukaryotes, including yeast, synthesize the side chain precursors by the mevalonate pathway. Interestingly, Streptomyces possess both the mevalonate and nonmevalonate pathways. These pathways are the subject of Chapters 1.12, 1.13, 1.14, 1.22.

## Acknowledgments

Research in the author's laboratory over the years was supported by Public Health Service Grants from the National Institutes of Health.

### Abbreviations

<b>ADP</b>	Adenosine diphosphate
<b>AMP</b>	Adenosine monophosphate
<b>ATP</b>	Adenosine triphosphate
<b>BLAST</b>	Basic Local Alignment Search Tool
<b>CoA or CoASH</b>	Coenzyme A
<b>FAD</b>	Flavin adenine dinucleotide
<b>FMN</b>	Flavin mononucleotide
<b>IUPAC-IUB</b>	International Union of Pure and Applied Chemistry-International Union of Biochemistry
<b>NAD</b>	Nicotinamide adenine dinucleotide
<b>NADH</b>	Nicotinamide adenine dinucleotide (Reduced)
<b>NADP</b>	Nicotinamide adenine dinucleotide phosphate
<b>NADPH</b>	Nicotinamide adenine dinucleotide phosphate (Reduced)
<b>NMR</b>	Nuclear magnetic resonance
<b>SEED</b>	None available
<b>TPP</b>	Thiamine pyrophosphate



## References

1. IUPAC-IUB, *Biochim. Biophys. Acta* **1965**, *107*, 5–10.
2. M. D. Collins; D. Jones, *Microbiol. Rev.* **1981**, *45*, 316–354.
3. A. Kroger, Phosphorylative Electron Transport with Fumarate and Nitrate as Terminal Hydrogen Acceptors. In *Microbial Energetics*; B.A. Haddock, W.A. Hamilton, Eds.; Cambridge University Press: Cambridge, 1977; pp 61–93.
4. J. R. Guest, *J. Gen. Microbiol.* **1979**, *115*, 259–271.
5. R. Meganathan, *FEMS Microbiol. Lett.* **1984**, *24*, 57–62.
6. R. Meganathan; J. Schrementi, *J. Bacteriol.* **1987**, *169*, 2862–2865.
7. L. Miguel; R. Meganathan, *Curr. Microbiol.* **1991**, *22*, 109–115.
8. U. Wissenbach; A. Kroger; G. Uden, *Arch. Microbiol.* **1990**, *154*, 60–66.
9. R. Bentley, *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 307–384.
10. E. Haslam, *The Shikimate Pathway*; John Wiley & Sons, Inc.: New York, 1974.
11. R. Bentley, *Biosynthesis* **1975**, *3*, 181–246.
12. R. Bentley, *Pure Appl. Chem.* **1975**, *41*, 47–68.
13. U. Weiss; J. M. Edwards, *The Biosynthesis of Aromatic Compounds*; John Wiley & Sons, Inc.: New York, 1980.
14. R. Bentley; R. Meganathan, *Microbiol. Rev.* **1982**, *46*, 241–280.
15. R. Bentley; R. Meganathan, *J. Nat. Prod.* **1983**, *46*, 44–59.
16. E. Haslam, *Shikimic Acid. Metabolism and Metabolites*; John Wiley & Sons, Inc.: New York, 1996.
17. T. P. Begley; C. Kinsland; S. Taylor; M. Tandon; R. Nicewonger; M. Wu; H. J. Chiu; N. Kelleher; N. Campobasso; Y. Zhang, *Top. Curr. Chem.* **1998**, *195*, 93–142.
18. R. Meganathan, *Vitam. Horm.* **2001**, *61*, 173–218.
19. R. Meganathan, *FEMS Microbiol. Lett.* **2001**, *203*, 131–139.
20. B. Soballe; R. K. Poole, *Microbiology* **1999**, *145*, 1817–1830.
21. J. M. Johnston; V. L. Arcus; E. N. Baker, *Acta Crystallogr. D: Biol. Crystallogr.* **2005**, *61*, 1199–1206.
22. J. J. Truglio; K. Theis; Y. Feng; R. Gajda; C. Machutta; P. J. Tonge; C. Kisker, *J. Biol. Chem.* **2003**, *278*, 42352–42360.
23. V. Ulaganathan; M. F. Agacan; L. Buetow; L. B. Tulloch; W. N. Hunter, *Acta Crystallogr. F: Struct. Biol. Cryst. Commun.* **2007**, *63*, 908–913.
24. R. E. Olson, *Annu. Rev. Nutr.* **1984**, *4*, 281–337.
25. J. W. Suttie, Vitamin-K. In *Fat Soluble Vitamins*; A. T. Diplock, Ed.; Technomic Publishing: Lancaster, PA, 1985; pp 225–311.
26. J. W. Suttie, *Annu. Rev. Nutr.* **1995**, *15*, 399–417.
27. F. Fernandez; M. D. Collins; M. J. Hill, *Biochem. Soc. Trans.* **1985**, *13*, 223–224.
28. S. W. Snyder; R. R. Rustandi; J. Biggins; J. R. Norris; M. C. Thurnauer, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9895–9896.
29. J. Grunler; J. Ericsson; G. Dallner, *Biochim. Biophys. Acta – Lipids Lipid Metab.* **1994**, *1212*, 259–277.
30. B. Frei; M. C. Kim; B. N. Ames, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 4879–4883.
31. A. M. James; R. A. J. Smith; M. P. Murphy, *Annual Meeting of the Oxygen Club of California (OCC)*, Elsevier Science Inc, Santa Barbara, CA, 2004; pp 47–56.
32. L. Ernster; P. Forsmarkandree, In *Proceedings of the 7th International Symposium on the Biomedical and Clinical Aspects of Coenzyme Q*; Springer-Verlag: Copenhagen, Denmark, 1992; pp S60–S65.
33. L. Ernster; G. Dallner, *Biochimica et Biophysica Acta.* **1995**, *1271*, 195–204.
34. L. Packer; V. Kagan; E. Cadenas, In *Antioxidant Activity of Coenzyme Q and Ubiquinols: Reactions with Tocopheroxyl and Nitric Oxide Radicals*; Proceedings of the First Conference of the International Coenzyme Q-10 Association, Boston, MA, USA, 1998, Abstract; pp 22–24.
35. D. Mohr; V. W. Bowry; R. Stocker, *Biochim. Biophys. Acta* **1992**, *1126*, 247–254.
36. G. Lenaz, *Biochim. Biophys. Acta – Bioenerg.* **1998**, *1366*, 53–67.
37. M. Turunen; J. Olsson; G. Dallner, *Biochim. Biophys. Acta – Biomembr.* **2004**, *1660*, 171–199.
38. C. M. Quinzii; S. DiMauro; M. Hirano, *Neurochem. Res.* **2007**, *32*, 723–727.
39. F. L. Crane, *Protoplasma* **2000**, *213*, 127–133.
40. R. Meganathan; O. Kwon, Biosynthesis of Menaquinone (Vitamin K<sub>2</sub>) and Ubiquinone (Coenzyme Q). In *EcoSal-Escherichia coli and Salmonella: Cellular and Molecular Biology*; A. Böck, R. Curtiss, III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, D. Ussery, Eds.; American Society for Microbiology Press: Washington, DC, 2008. <http://www.ecosal.org>.
41. T. Hiratsuka; K. Furihata; J. Ishikawa; H. Yamashita; N. Itoh; H. Seto; T. Dairi, *Science* **2008**, *321*, 1670–1673.
42. G. B. Cox; F. Gibson, *Biochim. Biophys. Acta* **1964**, *93*, 204–206.
43. I. M. Campbell; C. J. Coscia; M. Kelsey; R. Bentley, *Biochem. Biophys. Res. Commun.* **1967**, *28*, 25–29.
44. I. M. Campbell, *Tetrahedron Lett.* **1969**, *1969*, 4777–4780.
45. D. J. Robins; I. M. Campbell; R. Bentley, *Biochem. Biophys. Res. Commun.* **1970**, *39*, 1081–1086.
46. D. J. Robins; R. Bentley, *J. Chem. Soc. Chem. Commun.* **1972**, 232–233.
47. P. Dansette; R. Azerad, *Biochem. Biophys. Res. Commun.* **1970**, *40*, 1090–1095.
48. D. J. Robins; R. B. Yee; R. Bentley, *J. Bacteriol.* **1973**, *116*, 965–971.
49. I. G. Young, *Biochemistry* **1975**, *14*, 399–406.
50. R. Meganathan; R. Bentley, *Biochemistry* **1981**, *20*, 5336–5340.
51. R. Meganathan, *J. Biol. Chem.* **1981**, *256*, 9386–9388.
52. R. Bentley; R. Meganathan, The Biosynthesis of Isoprenoid Quinones Ubiquinone and Menaquinone. In *Escherichia coli and Salmonella Typhimurium Cellular and Molecular Biology*; F. C. Neidhardt, J. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, H. W. Umbarger, Eds.; American Society for Microbiology: Washington, DC, 1987; pp 512–520.
53. G. T. Emmons; I. M. Campbell; R. Bentley, *Biochem. Biophys. Res. Commun.* **1985**, *131*, 956–960.
54. A. Weische; E. Leistner, *Acta Agron. Acad. Sci. Hung.* **1985**, *34*, 115.

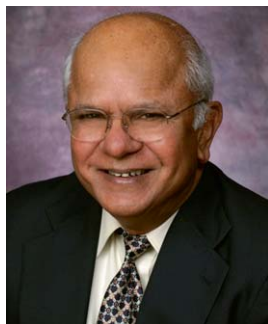


55. B. A. Ozenberger; T. J. Brickman; M. A. McIntosh, *J. Bacteriol.* **1989**, *171*, 775–783.
56. M. K. Tummuru; T. J. Brickman; M. A. McIntosh, *J. Biol. Chem.* **1989**, *264*, 20547–20551.
57. A. Kaiser; E. Leistner, *Arch. Biochem. Biophys.* **1990**, *276*, 331–335.
58. M. Guerinot, *Annu. Rev. Microbiol.* **1994**, *48*, 743–772.
59. J. B. Neilands, *Microbiol. Sci.* **1984**, *1*, 9–14.
60. O. Kwon; M. E. S. Hudspeth; R. Meganathan, *Abstr. Gen. Meet. Am. Soc. Microbiol.* **1993**, *93*, 272.
61. O. Kwon; D. K. Bhattacharyya; R. Meganathan, *J. Bacteriol.* **1996**, *178*, 6778–6781.
62. R. Meganathan, Biosynthesis of the Isoprenoid Quinones Menaquinone (Vitamin K<sub>2</sub>) and Ubiquinone (Coenzyme Q). In *Escherichia coli and Salmonella: Cellular and Molecular Biology*; F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, H. E. Umbarger, Eds.; American Society for Microbiology: Washington, DC, 1996; pp 642–656.
63. R. Daruwala; O. Kwon; R. Meganathan; M. E. Hudspeth, *FEMS Microbiol. Lett.* **1996**, *140*, 159–163.
64. R. Daruwala; D. K. Bhattacharyya; O. Kwon; R. Meganathan, *J. Bacteriol.* **1997**, *179*, 3133–3138.
65. C. T. Walsh; J. Liu; F. Rusnak; M. Sakaitani, *Chem. Rev.* **1990**, *90*, 1105–1129.
66. O. Kwon; M. E. Hudspeth; R. Meganathan, *J. Bacteriol.* **1996**, *178*, 3252–3259.
67. S. Kolappan; J. Zwahlen; R. Zhou; J. J. Truglio; P. J. Tonge; C. Kisker, *Biochemistry* **2007**, *46*, 946–953.
68. J. F. Parsons; K. M. Shi; J. E. Ladner, *Acta Crystallogr. D: Biol. Crystallogr.* **2008**, *64*, 607–610.
69. Z. He; K. D. Stigers Lavoie; P. A. Bartlett; M. D. Toney, *J. Am. Chem. Soc.* **2004**, *126*, 2378–2385.
70. R. Meganathan; R. Bentley, *J. Bacteriol.* **1983**, *153*, 739–746.
71. M. G. Marley; R. Meganathan; R. Bentley, *Biochemistry* **1986**, *25*, 1304–1307.
72. A. Weische; E. Leistner, *Tetrahedron Lett.* **1985**, *26*, 1487–1490.
73. C. Palaniappan; V. Sharma; M. E. Hudspeth; R. Meganathan, *J. Bacteriol.* **1992**, *174*, 8111–8118.
74. O. Kwon; D. K. Bhattacharyya; C. Palaniappan; R. Meganathan, *Abstr. Gen. Meet. Am. Soc. Microbiol.* **1997**, 356.
75. M. Bhasin; J. L. Billinsky; D. R. Palmer, *Biochemistry* **2003**, *42*, 13496–13504.
76. M. Jiang; Y. Cao; Z. F. Guo; M. J. Chen; X. L. Chen; Z. H. Guo, *Biochemistry* **2007**, *46*, 10979–10989.
77. M. Jiang; M. Chen; Y. Cao; Y. Yang; K. H. Sze; X. Chen; Z. Guo, *Org. Lett.* **2007**, *9*, 4765–4767.
78. E. A. L. Sieminska; A. Macova; D. R. J. Palmer; D. A. R. Sanders, *Acta Crystallogr. F: Struct. Biol. Cryst. Commun.* **2005**, *61*, 489–492.
79. M. Jiang; X. L. Chen; Z. F. Guo; Y. Cao; M. J. Chen; Z. H. Guo, *Biochemistry* **2008**, *47*, 3426–3434.
80. J. L. Popp, *J. Bacteriol.* **1989**, *171*, 4349–4354.
81. V. Sharma; R. Meganathan; M. E. Hudspeth, *J. Bacteriol.* **1993**, *175*, 4917–4921.
82. D. R. Palmer; J. B. Garrett; V. Sharma; R. Meganathan; P. C. Babbitt; J. A. Gerlt, *Biochemistry* **1999**, *38*, 4252–4258.
83. P. C. Babbitt; M. S. Hasson; J. E. Wedekind; D. R. Palmer; W. C. Barrett; G. H. Reed; I. Rayment; D. Ringe; G. L. Kenyon; J. A. Gerlt, *Biochemistry* **1996**, *35*, 16489–16501.
84. T. B. Thompson; J. B. Garrett; E. A. Taylor; R. Meganathan; J. A. Gerlt; I. Rayment, *Biochemistry* **2000**, *39*, 10662–10676.
85. V. A. Klenchin; E. A. Taylor Ringia; J. A. Gerlt; I. Rayment, *Biochemistry* **2003**, *42*, 14427–14433.
86. J. A. Gerlt; P. C. Babbitt; I. Rayment, *Arch. Biochem. Biophys.* **2005**, *433*, 59–70.
87. R. W. Bryant, Jr.; R. Bentley, *Biochemistry* **1976**, *15*, 4792–4796.
88. R. Meganathan; R. Bentley, *J. Bacteriol.* **1979**, *140*, 92–98.
89. L. Heide; S. Arendt; E. Leistner, *J. Biol. Chem.* **1982**, *257*, 7396–7400.
90. L. Heide; E. Leistner, *FEBS Lett.* **1981**, *128*, 201–204.
91. R. Kolkman; E. Leistner, *Z. Naturforsch. C, J. Biosci.* **1987**, *42*, 1207–1214.
92. R. Kolkman; E. Leistner, *Z. Naturforsch. C, J. Biosci.* **1987**, *42*, 542–552.
93. D. J. Shaw; J. R. Guest; R. Meganathan; R. Bentley, *J. Bacteriol.* **1982**, *152*, 1132–1137.
94. V. Sharma; M. E. S. Hudspeth; R. Meganathan, *Gene* **1996**, *168*, 43–48.
95. D. K. Bhattacharyya; O. Kwon; R. Meganathan, *J. Bacteriol.* **1997**, *179*, 6061–6065.
96. K. H. Chang; D. Dunaway-Mariano, *Biochemistry* **1996**, *35*, 13478–13484.
97. K. H. Chang; H. Xiang; D. Dunaway-Mariano, *Biochemistry* **1997**, *36*, 15650–15659.
98. V. Sharma; K. Suvarna; R. Meganathan; M. E. Hudspeth, *J. Bacteriol.* **1992**, *174*, 5057–5062.
99. W. J. Wu; V. E. Anderson; D. P. Raleigh; P. J. Tonge, *Biochemistry* **1997**, *36*, 2211–2220.
100. E. Kuznetsova; M. Proudfoot; S. A. Sanders; J. Reinking; A. Savchenko; C. H. Arrowsmith; A. M. Edwards; A. F. Yakunin, *FEMS Microbiol. Rev.* **2005**, *29*, 263–279.
101. J. R. Widhalm; C. van Oostende; F. Furt; G. J. C. Basset, *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5599–5603.
102. B. Shineberg; I. G. Young, *Biochemistry* **1976**, *15*, 2754–2758.
103. K. Suvarna; D. Stevenson; R. Meganathan; M. E. Hudspeth, *J. Bacteriol.* **1998**, *180*, 2782–2787.
104. R. M. Baldwin; C. D. Snyder; H. Rapoport, *Biochemistry* **1974**, *13*, 1523–1530.
105. J. C. Gebler; A. B. Woodside; C. D. Poulter, *J. Am. Chem. Soc.* **1992**, *114*, 7354–7360.
106. L. M. Jackman; I. G. O'Brien; G. B. Cox; F. Gibson, *Biochim. Biophys. Acta* **1967**, *141*, 1–7.
107. Z. El-Hachimi; O. Samuel; R. Azerad, *Biochimie* **1974**, *56*, 1239–1247.
108. I. G. Young; L. M. McCann; P. Strooban; F. Gibson, *J. Bacteriol.* **1971**, *105*, 769–778.
109. U. Wissenbach; T. Ternes; G. Uden, *Arch. Microbiol.* **1992**, *158*, 68–73.
110. P. T. Lee; A. Y. Hsu; H. T. Ha; C. F. Clarke, *J. Bacteriol.* **1997**, *179*, 1748–1754.
111. H. Seto; Y. Jinnai; T. Hiratsuka; M. Fukawa; K. Furihata; N. Itoh; T. Dairi, *J. Am. Chem. Soc.* **2008**, *130*, 5614–5615.
112. T. Hiratsuka; N. Itoh; H. Seto; T. Dairi, *Biosci. Biotechnol. Biochem.* **2009**, *73*, 1137–1141.
113. N. Hosokawa; H. Naganawa; T. Kasahara; S. Hattori; M. Hamada; T. Takeuchi; S. Yamamoto; K. S. Tsuchiya; M. Hori, *Chem. Pharm. Bull.* **1999**, *47*, 1032–1034.
114. J. Gross; W. K. Cho; L. Lezhneva; J. Falk; K. Krupinska; K. Shinozaki; M. Seki; R. G. Herrmann; J. Meurer, *J. Biol. Chem.* **2006**, *281*, 17189–17196.
115. T. W. Johnson; G. Z. Shen; B. Zybailov; D. Kolling; R. Reategui; S. Beauparlant; I. R. Vassiliev; D. A. Bryant; A. D. Jones; J. H. Golbeck; P. R. Chitnis, *J. Biol. Chem.* **2000**, *275*, 8523–8530.

116. T. W. Johnson; S. Naithani; C. Stewart; B. Zybailov; A. D. Jones; J. H. Golbeck; P. R. Chitnis, *Biochim. Biophys. Acta – Bioenerg.* **2003**, *1557*, 67–76.
117. Y. Sakuragi; B. Zybailov; G. Z. Shen; A. D. Jones; P. R. Chitnis; A. van der Est; R. Bittl; S. Zech; D. Stehlik; J. H. Golbeck; D. A. Bryant, *Biochemistry* **2002**, *41*, 394–405.
118. H. Shimada; R. Ohno; M. Shibata; I. Ikegami; K. Onai; M. Ohto; K. Takamiya, *Plant J.* **2005**, *41*, 627–637.
119. M. C. Wildermuth; J. Dewdney; G. Wu; F. M. Ausubel, *Nature* **2001**, *414*, 562–565.
120. P. Brodersen; F. G. Malinovsky; K. Hematy; M. A. Newman; J. Mundy, *Plant Physiol.* **2005**, *138*, 1037–1045.
121. A. Lohmann; M. A. Schottler; C. Brehelin; F. Kessler; R. Bock; E. B. Cahoon; P. Dormann, *J. Biol. Chem.* **2006**, *281*, 40461–40472.
122. F. Gibson, *Biochem. Soc. Trans.* **1973**, *1*, 317–326.
123. F. Gibson; I. G. Young, *Methods Enzymol.* **1978**, *53*, 600–609.
124. G. B. Cox; F. Gibson, *Biochem. J.* **1966**, *100*, 1–6.
125. T. Jonassen; C. Clarke, Genetic Analysis of Coenzyme Q Biosynthesis. In *Coenzyme Q: Molecular Mechanisms in Health and Disease*; V. E. Kagan, J. Quinn, Eds.; CRC Press: Boca Raton, FL, 2001; pp 185–208.
126. U. C. Tran; C. F. Clarke, *Mitochondrion* **2007**, *7*, S62–S71.
127. R. E. Olson; H. Rudney, *Vitam. Horm. – Adv. Res. Appl.* **1983**, *40*, 1–43.
128. J. F. Pennock; D. R. Threlfall, Biosynthesis of Ubiquinone and Related Compounds. In *Biosynthesis of Isoprenoid Compounds*; J. W. Porter, S. L. Spurgeon, Eds.; John Wiley and Sons: New York, 1983; Vol. 2, pp 191–303.
129. P. Stroobant; I. G. Young; F. Gibson, *J. Bacteriol.* **1972**, *109*, 134–139.
130. T. Jonassen; C. F. Clarke, *J. Biol. Chem.* **2000**, *275*, 12381–12387.
131. M. Kawamukai, *J. Biosci. Bioeng.* **2002**, *94*, 511–517.
132. B. P. Nichols; J. M. Green, *J. Bacteriol.* **1992**, *174*, 5309–5316.
133. M. Siebert; K. Severin; L. Heide, *Microbiology* **1994**, *140*, 897–904.
134. M. J. Holden; M. P. Mayhew; D. T. Gallagher; V. L. Vilker, *Biochim. Biophys. Acta (BBA) – Prot. Str. Mol. Enzymol.* **2002**, *1594*, 160–167.
135. D. T. Gallagher; M. Mayhew; M. J. Holden; A. Howard; K. J. Kim; V. L. Vilker, *Proteins* **2001**, *44*, 304–311.
136. N. Smith; A. E. Roitberg; E. Rivera; A. Howard; M. J. Holden; M. Mayhew; S. Kaistha; D. T. Gallagher, *Arch. Biochem. Biophys.* **2006**, *445*, 72–80.
137. A. N. Booth; M. S. Masri; D. J. Robbins; O. H. Emerson; F. T. Jones; F. DeEds, *J. Biol. Chem.* **1960**, *235*, 2649–2652.
138. R. Loscher; L. Heide, *Plant Physiol.* **1994**, *106*, 271–279.
139. Y. Ranganathan; R. Meganathan, *Abstr. Gen. Meet. Am. Soc. Microbiol.* **2007**, *107*, 389.
140. J. Liu; N. Quinn; G. A. Berchtold; C. T. Walsh, *Biochemistry* **1990**, *29*, 1417–1425.
141. R. Overbeek; T. Begley; R. M. Butler; J. V. Choudhuri; H. Y. Chuang; M. Cohoon; V. de Crecy-Lagard; N. Diaz; T. Disz; R. Edwards; M. Fonstein; E. D. Frank; S. Gerdes; E. M. Glass; A. Goesmann; A. Hanson; D. Iwata-Reuyl; R. Jensen; N. Jamshidi; L. Krause; M. Kubal; N. Larsen; B. Linke; A. C. McHardy; F. Meyer; H. Neuweger; G. Olsen; R. Olson; A. Osterman; V. Portnoy; G. D. Pusch; D. A. Rodionov; C. Ruckert; J. Steiner; R. Stevens; I. Thiele; O. Vassieva; Y. Ye; O. Zagnitko; V. Vonstein, *Nucleic Acids Res.* **2005**, *33*, 5691–5702.
142. I. G. Young; R. A. Leppik; J. A. Hamilton; F. Gibson, *J. Bacteriol.* **1972**, *110*, 18–25.
143. M. Melzer; L. Heide, *Biochim. Biophys. Acta – Lipids Lipid Metab.* **1994**, *1212*, 93–102.
144. L. Wessjohann; B. Sontag, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1697–1699.
145. R. A. Gibbs, Prenyl Transfer and the Enzymes of Terpenoid and Steroid Biosynthesis. In *Comprehensive Biological Catalysis*; M. Sinnott, Ed.; Academic Press: San Diego, CA, 1998; Vol. I, pp 31–118.
146. J. C. Gebler; A. B. Woodside; D. C. Poulter, *J. Am. Chem. Soc.* **1992**, *114*, 7354–7360.
147. G. B. Cox; I. G. Young; L. M. McCann; F. Gibson, *J. Bacteriol.* **1969**, *99*, 450–458.
148. R. A. Leppik; I. G. Young; F. Gibson, *Biochim. Biophys. Acta.* **1976**, *436*, 800–810.
149. B. J. Howlett; J. Bar-Tana, *J. Bacteriol.* **1980**, *143*, 644–651.
150. M. L. Nonet; C. C. Marvel; D. R. Tolan, *J. Biol. Chem.* **1987**, *262*, 12209–12217.
151. H. Zeng; I. Snavely; P. Zamorano; G. T. Javor, *J. Bacteriol.* **1998**, *180*, 3681–3685.
152. H. Zhang; G. T. Javor, *FEMS Microbiol. Lett.* **2003**, *223*, 67–72.
153. M. Gulmezian; K. R. Hyman; B. N. Marbois; C. F. Clarke; G. T. Javor, *Arch. Biochem. Biophys.* **2007**, *467*, 144–153.
154. E. S. Rangarajan; Y. Li; P. Iannuzzi; A. Tocilj; L. W. Hung; A. Matte; M. Cygler, *Protein Sci.* **2004**, *13*, 3006–3016.
155. I. G. Young; P. Stroobant; C. G. Macdonald; F. Gibson, *J. Bacteriol.* **1973**, *114*, 42–52.
156. K. Alexander; I. G. Young, *Biochemistry* **1978**, *17*, 4745–4750.
157. D. L. Daniels; G. Plunkett; V. Burland; F. R. Blattner, *Science* **1992**, *257*, 771–778.
158. W. W. Poon; D. E. Davis; H. T. Ha; T. Jonassen; P. N. Rather; C. F. Clarke, *J. Bacteriol.* **2000**, *182*, 5139–5146.
159. K. Nakahigashi; K. Miyamoto; K. Nishimura; H. Inokuchi, *J. Bacteriol.* **1992**, *174*, 7352–7359.
160. I. G. Young; L. McCann; P. Stroobant; F. Gibson, *J. Bacteriol.* **1971**, *105*, 769–778.
161. O. Kwon; A. Kotsakis; R. Meganathan, *FEMS Microbiol. Lett.* **2000**, *186*, 157–161.
162. H.-E. Knoell, *FEMS Microbiol. Lett.* **1981**, *10*, 63–65.
163. C. S. J. Walpole; R. Wrigglesworth; M. I. Page; A. Williams, Eds., *Enzyme Mechanisms*; Royal Society of Chemistry: London, Illus., 1987; pp xviii, 506–533, 550p.
164. K. Alexander; I. G. Young, *Biochemistry* **1978**, *17*, 4750–4755.
165. L. M. Jackman; I. G. O'Brien; G. B. Cox; F. Gibson, *Biochim. Biophys. Acta* **1967**, *141*, 1–7.
166. A. Y. Hsu; W. W. Poon; J. A. Shepherd; D. C. Myles; C. F. Clarke, *Biochemistry* **1996**, *35*, 9797–9806.
167. R. A. Leppik; P. Stroobant; B. Shineberg; I. G. Young; F. Gibson, *Biochim. Biophys. Acta.* **1976**, *428*, 146–156.
168. H. E. Knoell, *Biochem. Biophys. Res. Commun.* **1979**, *91*, 919–925.
169. H.-E. Knoell, *FEMS Microbiol. Lett.* **1981**, *10*, 59–62.
170. M. Gulmezian; H. Zhang; G. T. Javor; C. F. Clarke, *J. Bacteriol.* **2006**, *188*, 6435–6439.
171. A. Tzagoloff; C. L. Dieckmann, *Microbiological Rev.* **1990**, *54*, 211–225.

172. W. W. Poon; D. E. Davis; H. T. Ha; T. Jonassen; P. N. Rather; C. F. Clarke, *J. Bacteriol.* **2000**, *182*, 5139–5146.
173. M. Rohmer, *Nat. Prod. Rep.* **1999**, *16*, 565–574.
174. M. Rohmer, A Mevalonate-Independent Route to Isopentenyl Diphosphate. In *Comprehensive Natural Products Chemistry: Isoprenoids including Carotenoids and Steroids*; D. E. cane, Ed.; Elsevier: Oxford, UK, 1999; Vol. 2, pp 45–67.
175. M. N. Ashby; P. A. Edwards, *J. Biol. Chem.* **1990**, *265*, 13157–13164.
176. K. Okada; T. Kainou; H. Matsuda; M. Kawamukai, *FEBS Lett.* **1998**, *431*, 241–244.
177. K. Okada; T. Kainou; K. Tanaka; T. Nakagawa; H. Matsuda; M. Kawamukai, *Eur. J. Biochem.* **1998**, *255*, 52–59.
178. K. Okada; K. Suzuki; Y. Kamiya; X. F. Zhu; S. Fujisaki; Y. Nishimura; T. Nishino; T. Nakagawa; M. Kawamukai; H. Matsuda, *Biochim. Biophys. Acta – Lipids Lipid Metab.* **1996**, *1302*, 217–223.
179. K. Suzuki; M. Ueda; M. Yuasa; T. Nakagawa; M. Kawamukai; H. Matsuda, *Biosci. Biotechnol. Biochem.* **1994**, *58*, 1814–1819.
180. K. Okada; K. Ohara; K. Yazaki; K. Nozaki; N. Uchida; M. Kawamukai; H. Nojiri; H. Yamane, *Plant Mol. Biol.* **2004**, *55*, 567–577.
181. K. Ohara; K. Yamamoto; M. Hamamoto; K. Sasaki; K. Yazaki, *Plant Cell Physiol.* **2006**, *47*, 581–590.
182. N. Uchida; K. Suzuki; R. Saiki; T. Kainou; K. Tanaka; H. Matsuda; M. Kawamukai, *J. Bacteriol.* **2000**, *182*, 6933–6939.
183. M. Forsgren; A. Attersand; S. Lake; J. Grunler; E. Swiezewska; G. Dallner; C. L. Climent, *Biochem. J.* **2004**, *382*, 519–526.
184. P. Gin; C. F. Clarke, *J. Biol. Chem.* **2005**, *280*, 2676–2681.
185. W. W. Poon; D. E. Davis; H. T. Ha; T. Jonassen; P. N. Rather; C. F. Clarke, *J. Bacteriol.* **2000**, *182*, 5139–5146.
186. M. H. Avelange-Macherel; J. Joyard, *Plant J.* **1998**, *14*, 203–213.
187. B. N. Marbois; C. F. Clarke, *J. Biol. Chem.* **1996**, *271*, 2995–3004.
188. R. M. Kagan; S. Clarke, *Arch. Biochem. Biophys.* **1994**, *310*, 417–427.
189. R. J. Barkovich; A. Shtanko; J. A. Shepherd; P. T. Lee; D. C. Myles; A. Tzagoloff; C. F. Clarke, *J. Biol. Chem.* **1997**, *272*, 9182–9188.
190. E. Dibrov; K. M. Robinson; B. D. Lemire, *J. Biol. Chem.* **1997**, *272*, 9175–9181.
191. S. W. Baba; G. I. Belogradov; J. C. Lee; P. T. Lee; J. Strahan; J. N. Shepherd; C. F. Clarke, *J. Biol. Chem.* **2004**, *279*, 10052–10059.
192. R. A. Hagerman; P. J. Trotter; R. A. Willis, In *The Regulation of COQ5 Gene Expression by Energy Source*, Proceedings of the 2nd Conference of the International Coenzyme Q10 Association, Taylor & Francis Ltd.: Frankfurt, Germany, 2000; pp 485–490.
193. R. A. Hagerman; R. A. Willis, *Biochim. Biophys. Acta – Gene Struct. Expr.* **2002**, *1578*, 51–58.
194. P. Gin; A. Y. Hsu; S. C. Rothman; T. Jonassen; P. T. Lee; A. Tzagoloff; C. F. Clarke, *J. Biol. Chem.* **2003**, *278*, 25308–25316.
195. B. A. Palfey; D. P. Ballou; V. Massey, In *Active Oxygen in Biochemistry*; J. S. Valentine, C. S. Foote, A. Greenberg, J. F. Liebman, Eds.; Blackie Academic and Professional: London, UK, 1995; pp 37–83.
196. U. C. Tran; B. Marbois; P. Gin; M. Gulmezian; T. Jonassen; C. F. Clarke, *J. Biol. Chem.* **2006**, *281*, 16401–16409.
197. T. Jonassen; M. Proft; F. Randez-Gil; J. R. Schultz; B. N. Marbois; K. D. Entian; C. F. Clarke, *J. Biol. Chem.* **1998**, *273*, 3351–3357.
198. O. Kwon; A. Kotsakis; R. Meganathan, *Abstr. Gen. Meet. Am. Soc. Microbiol.* **1998**, *98*, 347.
199. A. Tauche; U. Krause-Buchholz; G. Rodel, *FEMS Yeast Res.* **2008**, *8*, 1263–1275.
200. T. Q. Do; A. Y. Hsu; T. Jonassen; P. T. Lee; C. F. Clarke, *J. Biol. Chem.* **2001**, *276*, 18161–18168.
201. N. Bonnefoy; M. Kermorgant; P. BrivetChevillotte; G. Dujardin, *Mol. Gen. Genet.* **1996**, *251*, 204–210.
202. C. J. Leonard; L. Aravind; E. V. Koonin, *Genome Res.* **1998**, *8*, 1038–1047.
203. E. J. Hsieh; P. Gin; M. Gulmezian; U. C. Tran; R. Saiki; B. N. Marbois; C. F. Clarke, *Arch. Biochem. Biophys.* **2007**, *463*, 19–26.
204. B. Marbois; P. Gin; K. F. Faull; W. W. Poon; P. T. Lee; J. Strahan; J. N. Shepherd; C. F. Clarke, *J. Biol. Chem.* **2005**, *280*, 20231–20238.
205. M. H. Barros; A. Johnson; P. Gin; B. N. Marbois; C. F. Clarke; A. Tzagoloff, *J. Biol. Chem.* **2005**, *280*, 42627–42635.

### Biographical Sketch



Professor Meganathan received his B.Sc. (Agri.) degree from the Agriculture College (now Tamil Nadu Agricultural University, TNAU), Coimbatore, India, and his M.S. and Ph.D. degrees from Oklahoma State University, Stillwater, OK. He carried out his postdoctoral studies at the Department of Bacteriology, University of Wisconsin, and the Department of Microbiology, the University of Rochester Medical School. Subsequently, he was a research

associate in the laboratory of Professor Ronald Bentley at the University of Pittsburgh. In Professor Bentley's laboratory, studies on the biosynthesis of the isoprenoid quinone menaquinone (vitamin K<sub>2</sub>) was initiated and continued in subsequent years. In addition, during this period he investigated the biosynthesis of the earthy smelling compound Geosmin and that of methylisoborneol in Streptomyces. By isotopic labeling techniques it was shown that these compounds are isoprenoid metabolites that are not produced by nondifferentiating mutants. In 1982, he joined the faculty at Northern Illinois University and became a full professor in 1991. In 2002 he was selected and appointed as a presidential research professor and since 2006 he has been a distinguished research professor. He was elected as a Fellow of the American Academy of Microbiology in 1998.

## 7.13 Biosynthesis of Heme and Vitamin B<sub>12</sub>

**Gunhild Layer and Dieter Jahn**, Technical University of Braunschweig, Braunschweig, Germany

**Evelyne Deery, Andrew D. Lawrence, and Martin J. Warren**, University of Kent, Kent, UK

© 2010 Elsevier Ltd. All rights reserved.

7.13.1	Tetrapyrroles – The Pigments of Life	446
7.13.2	Tetrapyrrole Biosynthetic Pathways	446
7.13.3	Formation of 5-Aminolevulinic Acid	448
7.13.3.1	5-Aminolevulinic Acid Synthase	448
7.13.3.2	Glutamyl-tRNA Reductase and Glutamate-1-Semialdehyde-2,1-Aminomutase	449
7.13.4	Conversion of 5-Aminolevulinic Acid into Uroporphyrinogen III	453
7.13.4.1	Porphobilinogen Synthase	453
7.13.4.2	Porphobilinogen Deaminase	456
7.13.4.3	Uroporphyrinogen III Synthase	456
7.13.5	Tetrapyrrole Biosynthesis Branch Point – Uroporphyrinogen III	459
7.13.6	Conversion of Uroporphyrinogen III into Protoheme	459
7.13.6.1	Uroporphyrinogen III Decarboxylase	459
7.13.6.2	Coproporphyrinogen III Oxidase	462
7.13.6.3	Protoporphyrinogen IX Oxidase	463
7.13.6.4	Ferrochelatase	463
7.13.7	The Biosynthesis of Adenosylcobalamin (Vitamin B <sub>12</sub> )	465
7.13.8	A Note on Nomenclature	465
7.13.9	Overview of Cobalamin Biosynthesis	469
7.13.10	Uroporphyrinogen III to Precorrin-2	469
7.13.11	The Aerobic Pathway	474
7.13.11.1	Precorrin-3A Synthesis, C20 Methylation	474
7.13.11.2	Precorrin-3B Synthesis	474
7.13.11.3	Precorrin-4 Synthesis, Ring Contraction, and C17 Methylation	476
7.13.11.4	Precorrin-5 Synthesis, C11 Methylation	477
7.13.11.5	Precorrin-6 Synthesis, C1 Methylation, and Reduction	477
7.13.11.6	Precorrin-8 Synthesis, Methylation at C5, C15, and Decarboxylation	477
7.13.11.7	Hydrogenobyric Acid Synthesis	478
7.13.11.8	Hydrogenobyric Acid <i>a,c</i> -Diamide Synthesis	479
7.13.11.9	Cobalt Insertion	479
7.13.11.10	Cobalt Reduction	479
7.13.11.11	Adenosylation	480
7.13.11.12	Adenosylcobyrinic Acid Synthesis	481
7.13.12	The Anaerobic Pathway – From Precorrin-2 to Adenosylcobyrinic Acid	481
7.13.12.1	Sirohydrochlorin Synthesis	481
7.13.12.2	Cobalt Insertion	481
7.13.12.3	Methylation at C20	483
7.13.12.4	Ring Contraction and C17 Methylation	484
7.13.12.5	Methylation at C11 and Deacylation	484
7.13.12.6	Methylation at C1	484
7.13.12.7	Methylations at C5, C15, and Decarboxylation	485
7.13.12.8	Cobyric Acid Synthesis	485
7.13.12.9	Amidation of the Macrocycle	485
7.13.12.10	The Anaerobic Pathway and Adenosylation	486
7.13.13	Biosynthesis of Cobinamide (Phosphate)	486
7.13.13.1	Synthesis of ( <i>R</i> )-1-Amino-2-Propanol	486



7.13.13.2	Attachment of ( <i>R</i> )-1-Amino-2-Propanol <i>O</i> -2-Phosphate	486
<b>7.13.14</b>	<b>Nucleotide Loop Assembly</b>	489
7.13.14.1	Synthesis of Adenosyl-GDP-Cobinamide	489
7.13.14.2	Synthesis of 5,6-Dimethylbenzimidazole	489
7.13.14.3	Synthesis of $\alpha$ -Ribazole	491
7.13.14.4	Synthesis of Adenosylcobalamin	492
<b>References</b>		493

### 7.13.1 Tetrapyrroles – The Pigments of Life

Tetrapyrroles are a class of molecules that add color to the world around us. They are the most abundant pigments found in nature, fulfill widely diverse biochemical functions, and are essential in most known living organisms. Well-known members of this group of molecules are hemes, which give blood its red color, and the chlorophylls, which are responsible for the green color of plants, algae, and some bacteria. Other naturally occurring tetrapyrroles are cobalamin (vitamin B<sub>12</sub>), siroheme, coenzyme F<sub>430</sub>, heme *d*<sub>1</sub>, and the bilins.<sup>1–5</sup>

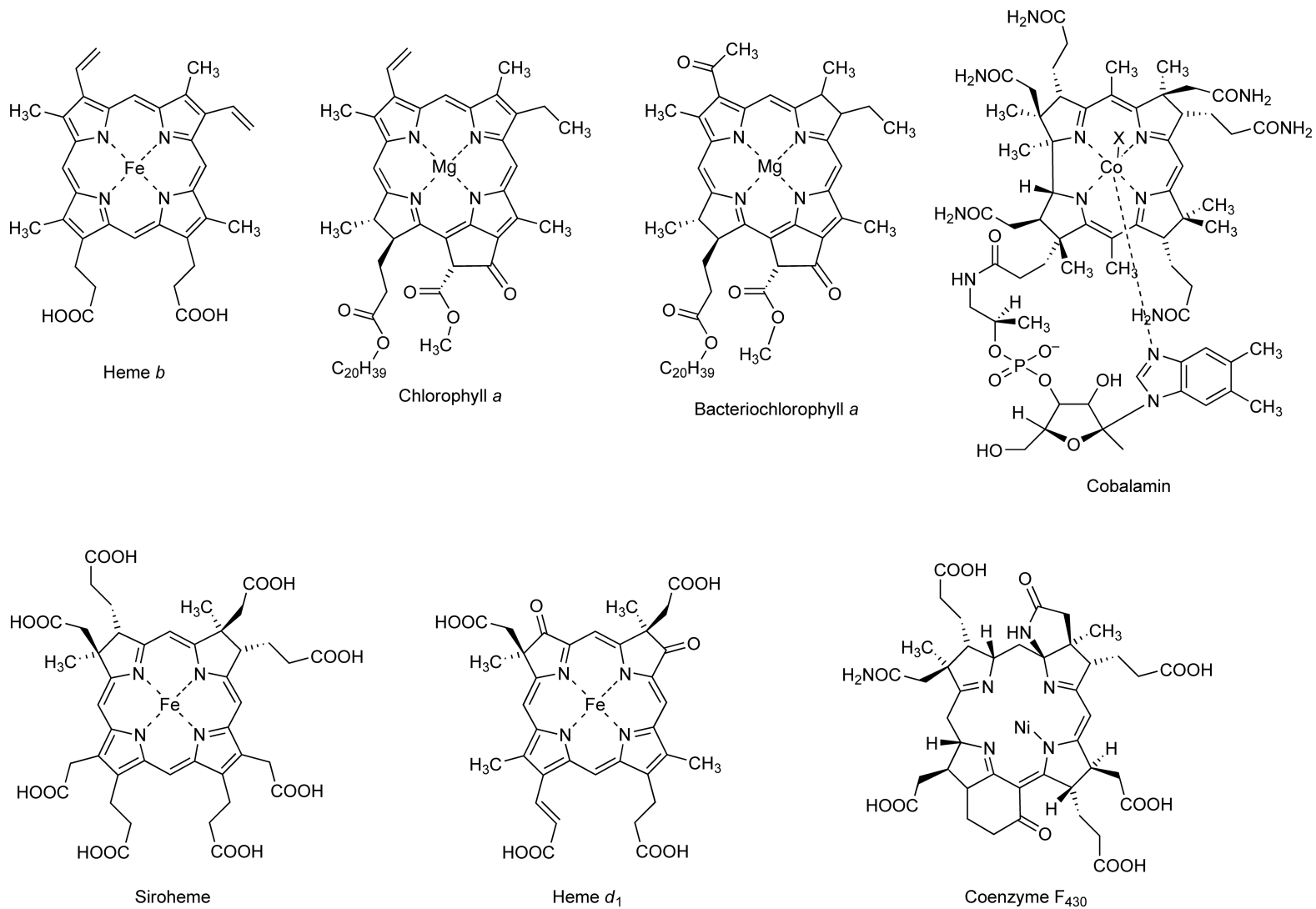
As suggested by the name, tetrapyrroles consist of four pyrrole-derived compounds either linked in linear or cyclic fashion via methine bridges. The only exceptions are the corrinoids, whose ring system lacks one bridge carbon between the first and fourth pyrrole ring. All linear tetrapyrroles result from the cleavage of cyclic tetrapyrroles and therefore contain only three methine bridges. In general, the pyrrole rings of all cyclic tetrapyrroles are denoted A–D in a clockwise direction. An important property of the cyclic tetrapyrroles is their ability to chelate diverse metal ions such as Fe, Mg, Co, or Ni. The naturally occurring tetrapyrroles differ in the oxidation state of their ring system, the nature of the chelated metal ion, and the ring substituents.<sup>6</sup> True porphyrins like hemes possess a macrocycle with a system of fully conjugated double bonds. Tetrapyrroles with variations in the basic porphyrin structure include more reduced cyclic tetrapyrroles such as chlorophylls (a chlorin), bacteriochlorophylls (chlorins or bacteriochlorins), siroheme and heme *d*<sub>1</sub> (isobacteriochlorins), and coenzyme F<sub>430</sub>. Finally, tetrapyrroles like cobalamin with their ring-contracted framework belong to a separate class that are called corrinoids. The structures of some important representatives of naturally occurring tetrapyrroles are shown in [Figure 1](#).

Iron-containing heme, as a prosthetic group of hemoglobin, transports molecular oxygen and carbon dioxide through the cardiovascular system. Heme also functions as a prosthetic group of many enzymes such as catalases, peroxidases, and cytochrome P-450 enzymes.<sup>7</sup> Furthermore, heme plays an important role in respiration. Heme-containing cytochromes are integral parts of various electron transfer chains. Heme proteins can also serve as sensors for diatomic gases such as O<sub>2</sub>, CO, and NO.<sup>8</sup> Magnesium-chelating chlorophylls and bacteriochlorophylls are essential constituents of the photosynthetic apparatus of plants and bacteria, trapping sunlight and helping to convert its energy into chemical energy.<sup>9</sup> The greenish cofactor siroheme is involved in reactions during assimilatory nitrite reduction or sulfite reduction.<sup>10,11</sup> The green heme *d*<sub>1</sub> is also an iron-containing tetrapyrrole, which serves as a prosthetic group exclusively in the cytochrome *cd*<sub>1</sub> nitrite reductases of some bacteria during the denitrification process.<sup>12</sup> The yellow nickel-containing coenzyme F<sub>430</sub> is the cofactor of methyl-coenzyme M reductase involved in methane formation.<sup>4</sup> The most complex known tetrapyrrole is the cobalt-containing cobalamin (vitamin B<sub>12</sub>). Cobalamin and its derivatives serve as cofactors in a number of enzymes that catalyze methyl transfer reactions or radical-based rearrangements.<sup>1</sup> In contrast to the cyclic tetrapyrroles, the linear tetrapyrroles do not contain a tightly bound metal. They serve as chromophoric photoreceptors in cyanobacterial and higher plant light-harvesting systems.<sup>13</sup> Moreover, light-responsive regulatory systems including the phytochrome family utilize linear tetrapyrroles as cofactors.<sup>14</sup>

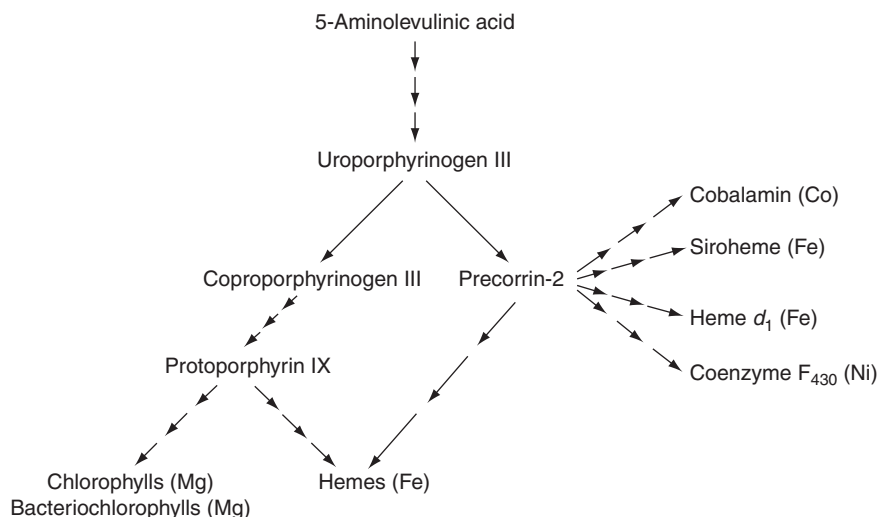
### 7.13.2 Tetrapyrrole Biosynthetic Pathways

The shared structural core of tetrapyrroles intuitively implies a common biosynthetic pathway. This is indeed the case. All tetrapyrroles derive from a common precursor molecule, 5-aminolevulinic acid (ALA), and the following biosynthetic steps up to uroporphyrinogen III are all conserved. The major branching point occurs





**Figure 1** Important representatives of naturally occurring tetrapyrroles.



**Figure 2** Overview of the divergent tetrapyrrole biosynthetic pathways. All tetrapyrroles derive from the common precursor 5-aminolevulinic acid. The major branching point of tetrapyrrole biosynthesis is uroporphyrinogen III. Coproporphyrinogen III is used to synthesize hemes and chlorophylls, whereas precorrin-2 gives rise to cobalamin, siroheme, heme  $d_1$ , and coenzyme F<sub>430</sub>. In archaea and some bacteria, precorrin-2 is also the precursor for heme.

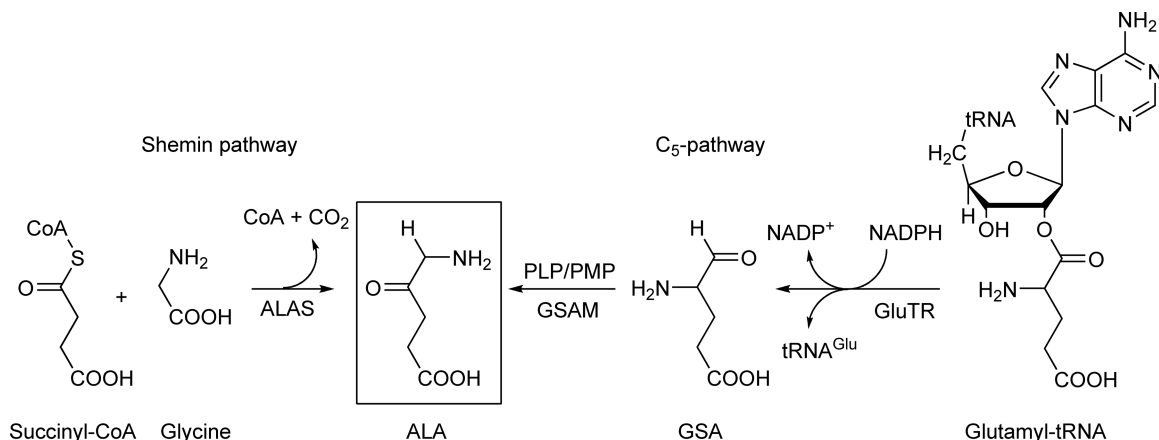
after the formation of uroporphyrinogen III, which represents the first cyclic intermediate. It is converted into either coproporphyrinogen III or precorrin-2. Coproporphyrinogen III is used to synthesize hemes, chlorophylls, and bacteriochlorophylls, while precorrin-2 gives rise to siroheme, heme  $d_1$ , coenzyme F<sub>430</sub>, and cobalamin. In archaea and some bacteria precorrin-2 is also the precursor for heme.<sup>15,16</sup> An overview of the divergent biochemical pathways of tetrapyrroles is shown in **Figure 2**. In eukaryotes, tetrapyrrole biosynthesis is limited to the formation of heme, siroheme, chlorophyll, and bilins. Prokaryotes are additionally able to synthesize the more complicated tetrapyrroles such as cobalamin, coenzyme F<sub>430</sub>, and heme  $d_1$ .

### 7.13.3 Formation of 5-Aminolevulinic Acid

The biosynthesis of all tetrapyrroles begins with the formation of ALA, which is the sole source of all carbon and nitrogen atoms required for the formation of the tetrapyrrole macrocycle. In nature, this small molecule is synthesized by two alternative, unrelated biosynthetic routes (**Figure 3**). In mammals, fungi, and the  $\alpha$ -group of proteobacteria ALA is formed via the condensation of succinyl-CoA and glycine, the so-called ‘Shemin pathway’ named after one of the two discoverers, and is catalyzed by 5-aminolevulinic acid synthase (ALAS).<sup>17,18</sup> In plants, archaea, and most bacteria, ALA formation starts from the C5-skeleton of glutamate and proceeds via the so-called ‘C5-pathway’.<sup>19,20</sup> Here, two enzymes are involved in ALA formation. NADPH-dependent glutamyl-tRNA reductase (GluTR) catalyzes the reduction of glutamyl-tRNA to glutamate-1-semialdehyde (GSA), which is subsequently converted by pyridoxal 5'-phosphate (PLP)-dependent glutamate-1-semialdehyde-2,1-aminomutase (GSAM) into ALA by a transamination reaction.<sup>21</sup>

#### 7.13.3.1 5-Aminolevulinic Acid Synthase

ALAS (EC 2.3.1.37) uses PLP as a cofactor and catalyzes the one step condensation of glycine and succinyl-CoA to give rise to ALA, carbon dioxide, and free coenzyme A (**Figure 3**).<sup>18</sup> ALAS belongs to the  $\alpha$ -oxoamine synthase subfamily of PLP-dependent enzymes that typically catalyze the condensation of a carboxylic acid coenzyme A thioester and an amino acid with the concomitant decarboxylation of the latter.<sup>22,23</sup> Since the first description of ALAS activities in 1958 by Neuberger and coworkers<sup>17</sup> and by Shemin and coworkers,<sup>18</sup> ALAS

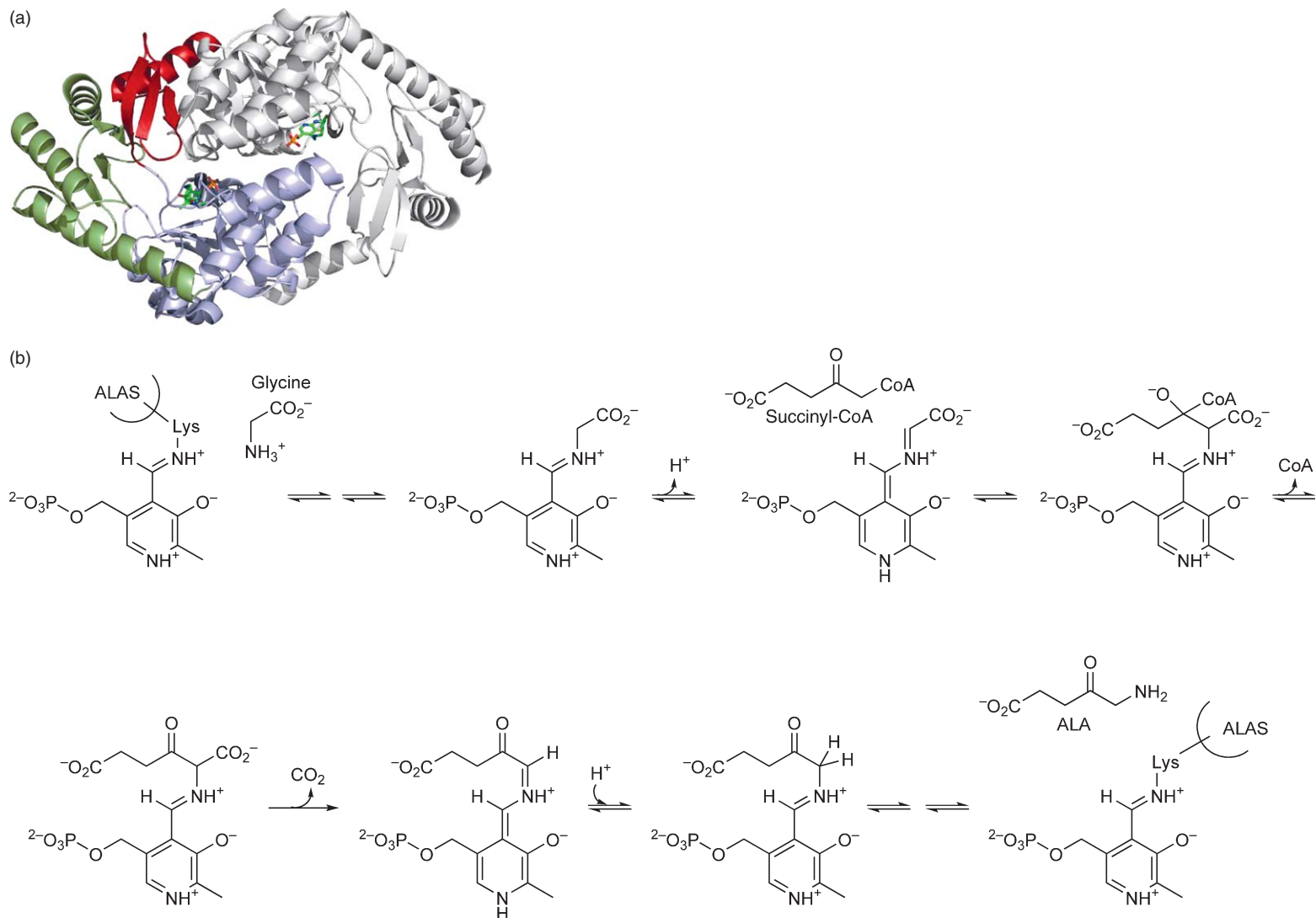


**Figure 3** Formation of 5-aminolevulinic acid. ALA is either synthesized from succinyl-CoA and glycine (Shemin pathway) or from the C<sub>5</sub>-skeleton of glutamate (C<sub>5</sub>-pathway).

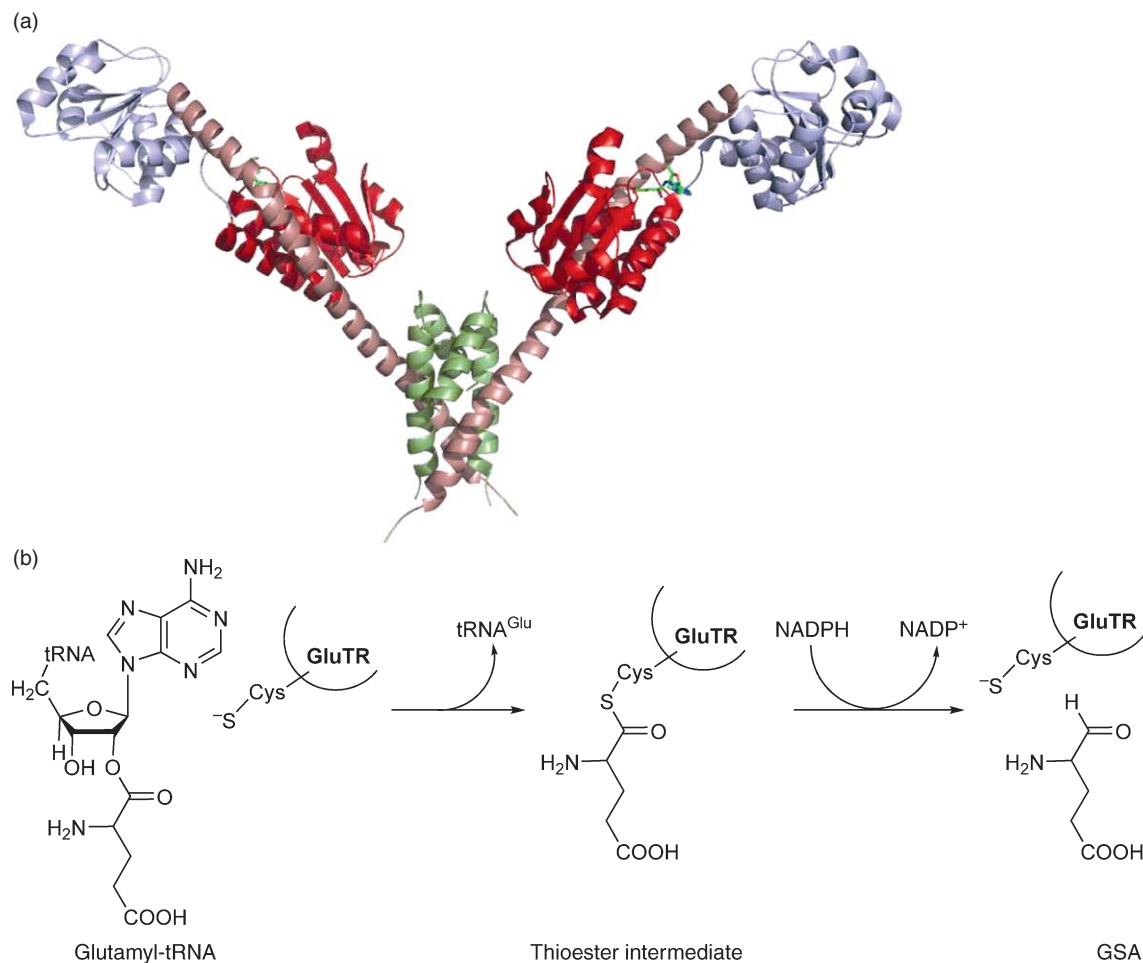
proteins have been purified and characterized from various eukaryotic and bacterial sources.<sup>3</sup> In mammals there are two ALAS isoforms, which are encoded by two different genes. One of them fulfills a housekeeping function (ALAS1/ALAS-H) whereas the other (ALAS2/ALAS-E) is required to sustain the extra need for heme in the hemoglobin of erythrocytes.<sup>24</sup> In contrast to that, most  $\alpha$ -proteobacteria use only one form of ALAS with the exception of *Rhodobacter sphaeroides*, which also contains two isoforms of the protein.<sup>25</sup> The first crystal structure of an ALAS was solved for the enzyme from *Rhodobacter capsulatus* including the PLP cofactor and the bound substrates glycine and succinyl-CoA.<sup>26</sup> The structure shows functional ALAS as a tightly interlocked homodimer in which each monomer consists of three domains: an N-terminal domain, the central catalytic domain, and a C-terminal domain (**Figure 4(a)**). In the catalytic domain of each monomer a PLP molecule is covalently bound to a conserved active-site lysine residue via a Schiff base (internal aldimine). During the catalytic cycle of ALAS (**Figure 4(b)**) the incoming substrate glycine first replaces the lysine residue of the enzyme in a transaldimination reaction forming the external aldimine. Stereospecific deprotonation of the PLP-bound glycine leads to the formation of a quinonoid intermediate, which then condenses with the second substrate succinyl-CoA. After the release of the coenzyme A group the resulting  $\alpha$ -amino- $\beta$ -ketoaldipate aldimine is decarboxylated to form the second quinonoid intermediate. Final protonation of this intermediate with the release of ALA from the active site and reformation of the internal PLP-lysine aldimine completes the catalytic cycle.<sup>27–29</sup> The release of ALA is the slowest and therefore the rate-limiting step of the reaction.

### 7.13.3.2 Glutamyl-tRNA Reductase and Glutamate-1-Semialdehyde-2,1-Aminomutase

During the C<sub>5</sub>-pathway for ALA formation GluTR (EC 6.1.1.17) catalyzes the NADPH-dependent reduction of glutamyl-tRNA to GSA, which in turn is converted by GSAM (EC 5.4.3.8) into ALA (**Figure 3**). The initial substrate, glutamyl-tRNA, is provided by glutamyl-tRNA synthetase (GluRS) and is used for both protein and tetrapyrrole biosynthesis.<sup>21</sup> GluTR proteins were purified from various bacterial and plant sources and biochemically characterized.<sup>30–32</sup> The recognition of the substrate glutamyl-tRNA and the cofactor NADPH by GluTR was studied. For substrate recognition by GluTR the unique overall shape of the tRNA<sup>Glu</sup> molecule was found essential.<sup>33</sup> Studies using various different NADPH analogues revealed that the reduced dinucleotide cofactor is recognized by all its major determinants.<sup>30</sup> The crystal structure of *Methanopyrus kandleri* GluTR was obtained in complex with the inhibitor glutamycin, a synthetic glutamate analogue.<sup>34</sup> The *M. kandleri* GluTR structure revealed an unusual extended V-shaped dimer in which each monomer consists of three distinct domains (**Figure 5(a)**). At the N-terminus the catalytic domain is found followed by an NADPH-binding domain. The third domain functions in dimerization of two monomers. These three domains are arranged along an extended curved ‘spinal’ helix. The inhibitor glutamycin was found bound in the active site,



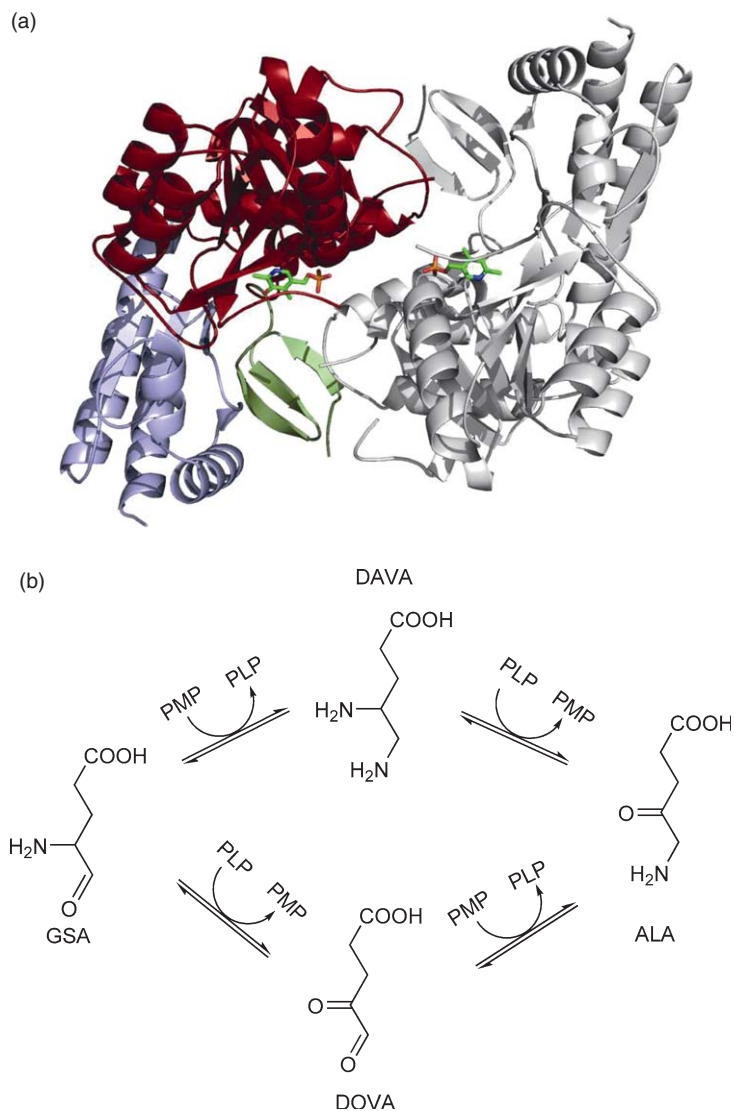
**Figure 4** Structure and reaction mechanism of ALAS. (a) Crystal structure of *Rhodobacter capsulatus* ALAS. Each subunit of the dimeric protein consists of an N-terminal domain (red), the central catalytic domain (blue), and a C-terminal domain (green). Each monomer contains a covalently attached PLP (shown in sticks representation). This and all other protein structure representations were generated using PyMOL (<http://pymol.sourceforge.net/>). (b) The ALAS reaction proceeds via formation of a quinonoid intermediate, release of CoA and  $\text{CO}_2$ , and formation of a second quinonoid intermediate before ALA is released from the enzyme.



**Figure 5** Structure and reaction mechanism of GluTR. (a) Crystal structure of *Methanopyrus kandleri* GluTR in complex with the inhibitor glutamycin. Each subunit of the dimeric protein consists of a catalytic domain (red), an NADPH binding domain (blue), and a dimerization domain (green). These three domains are arranged along an extended curved helix (violet). (b) During the GluTR reaction a thioester intermediate is formed between a conserved cysteine residue of GluTR and the substrate glutamate.

a deep pocket of the catalytic domain, and was coordinated by highly conserved amino acid residues. A strictly conserved cysteine residue within the active site acts as the catalytically active nucleophile, which attacks the  $\alpha$ -carbonyl group of tRNA-bound glutamate forming an enzyme-bound thioester intermediate. The following hydride transfer from NADPH to the thioester-bound glutamate produces GSA (**Figure 5(b)**). In the absence of NADPH the reactive thioester intermediate is hydrolyzed and free glutamate is released.<sup>30,31</sup>

In the following reaction step GSA is converted into ALA by the action of GSAM in a PLP-dependent transamination reaction. GSAM belongs to a large family of structurally related, PLP-dependent enzymes, which includes synthases, racemases, decarboxylases, mutases, and aminotransferases. GSAM functions as a mutase catalyzing the exchange of amino and oxo groups in GSA. GSAM proteins from different bacterial and plant sources have been purified and characterized.<sup>35–38</sup> They were mainly found in a dimeric form. For the reaction mechanism by which GSAM catalyzes the transamination of GSA there are two possible routes (**Figure 6(b)**). In one of the possible mechanisms the reaction starts with GSAM-bound pyridoxamine 5'-phosphate (PMP). Here, the PMP reacts with the aldehyde of GSA to yield the reaction intermediates 4,5-diaminovalerate (DAVA) and PLP. In the second step PMP is regenerated by abstraction of the C4 amino group of DAVA to give the final reaction product ALA. In the alternative reaction mechanism the initial



**Figure 6** Structure and reaction mechanism of GSAM. (a) Crystal structure of *Thermosynechococcus elongatus*. Each subunit of the dimeric protein consists of a small N-terminal domain (green), the central catalytic domain (red), and a C-terminal domain (blue). Each monomer contains a covalently attached PLP (sticks representation). (b) Two possible routes for the GSAM reaction with either 4,5-diaminovalerate (DAVA) or 4,5-dioxovalerate (DOVA) as reaction intermediates.

cofactor is PLP and the reaction proceeds via the intermediates 4,5-dioxovalerate (DOVA) and PMP. All GSAMs are capable of catalyzing both reactions, although kinetic studies favored the first mechanism.<sup>37,39</sup> Crystal structures for GSAM have been determined for the *Synechococcus* (GSAM<sub>Syn</sub>) and the *Thermosynechococcus elongatus* (GSAM<sub>Tel</sub>) proteins.<sup>40,41</sup> The GSAM<sub>Syn</sub> structure revealed an asymmetric homodimer in which one of the monomers contained the cofactor in its PLP form covalently attached to a lysine residue. The other monomer contained the noncovalently bound PMP. The PLP-containing monomer adopted an ‘open’ conformation, whereas the PMP-containing subunit was observed to be in a ‘closed’ conformation. In contrast to that, the GSAM<sub>Tel</sub> structure revealed a perfectly symmetric homodimer consisting of two PLP-binding monomers adopting an open conformation (Figure 6(a)). Consequently, the two different crystal structures leave the question for the true reaction mechanism still open. Only crystals containing reaction intermediates or inhibitors will allow the true mechanism to be established.



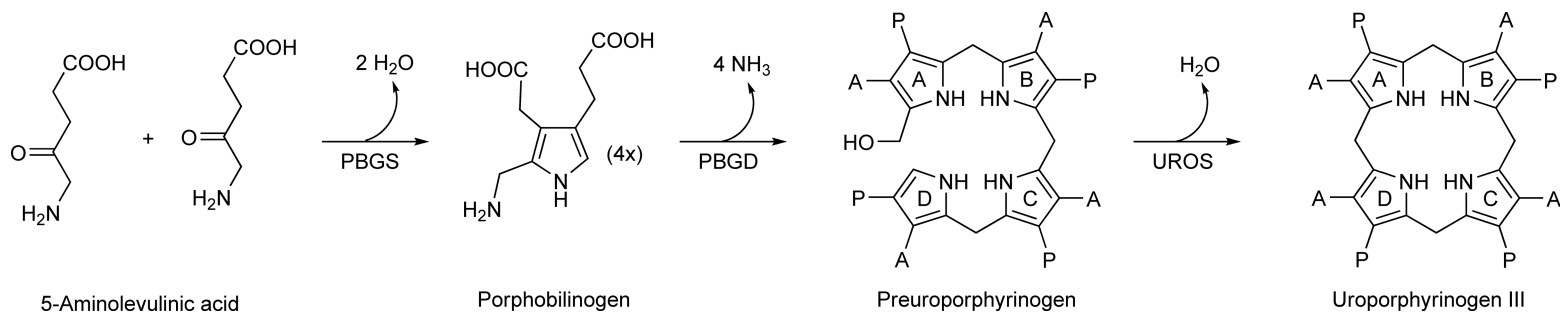
The tRNA-dependent formation of ALA in plants and most bacteria requires the concerted action of the two participating enzymes GluTR and GSAM. The reaction intermediate GSA is notoriously unstable in solution and therefore a direct channeling of GSA from GluTR to GSAM has been proposed. Indeed, an *in silico* model complex between GluTR and GSAM showed a striking degree of surface complementarity between the two proteins.<sup>34</sup> Additionally, in this model complex the proposed active site entrance of GSAM was positioned opposite the proposed active site exit of GluTR thus allowing the direct channeling of GSA from one enzyme to the other without being exposed to the surrounding medium. The existence of the proposed GluTR–GSAM complex was biochemically proven for the enzymes from *Escherichia coli* and *Chlamydomonas reinhardtii*.<sup>42,43</sup>

### 7.13.4 Conversion of 5-Aminolevulinic Acid into Uroporphyrinogen III

Uroporphyrinogen III, the first cyclic intermediate of the tetrapyrrole biosynthetic pathway, is formed from eight molecules of ALA in three consecutive enzymatic steps (Figure 7). In the first step porphobilinogen synthase (PBGs) catalyzes the asymmetric condensation of two molecules of ALA to yield the pyrrole derivative porphobilinogen (PBG). Next, four molecules of porphobilinogen are linked together by the enzyme porphobilinogen deaminase (PBGD) to yield the linear tetrapyrrole preuroporphyrinogen (1-hydroxymethylbilane). Finally, uroporphyrinogen III synthase (UROS) converts preuroporphyrinogen into uroporphyrinogen III under inversion of the D ring, which creates an asymmetry found in all cyclic tetrapyrroles.

#### 7.13.4.1 Porphobilinogen Synthase

PBGs (EC 4.2.1.24), also known as 5-aminolevulinic acid dehydratase, catalyzes the asymmetric condensation of two ALA molecules to form the pyrrole derivative porphobilinogen. PBGS proteins from prokaryotes and eukaryotes share a high degree of sequence similarity and have been purified from various sources.<sup>44–48</sup> The majority of PBGS enzymes are homooctameric proteins. However, the various proteins differ in their metal ion utilization.<sup>49</sup> Zinc-, magnesium-, and zinc/magnesium-dependent PBGS enzymes have been described. For example, human and yeast PBGS proteins contain two zinc ions. One zinc ion is bound in the active site and ligated by three highly conserved cysteine residues. The second zinc ion is bound near the active site and probably has structural roles. Magnesium-dependent PBGS like the *P. aeruginosa* enzyme do not contain any zinc ions. The mixed zinc/magnesium-dependent enzymes like *E. coli* PBGS contain the active site zinc and the magnesium ion. However, it was possible to change the magnesium-dependent *P. aeruginosa* PBGS into a zinc-dependent enzyme by stepwise introduction of cysteine residues into the active site.<sup>50</sup> Additionally, there is also evidence that some PBGS proteins are completely metal independent.<sup>51</sup> Various X-ray crystal structures of PBGS have been determined.<sup>52–54</sup> These structures showed that the fundamental structural unit, independent of metal dependence, is an asymmetric homodimer (Figure 8(a)). Four of these dimers form the final octameric structure. Each subunit adopts the classical TIM barrel fold in which the active site is located in an opening formed by the loops connecting the C-terminal ends of the parallel  $\beta$ -strands of the barrel with their following  $\alpha$ -helices. Dimer formation is assured by the N-terminal arm of each monomer, which wraps around the TIM barrel domain of the second monomer. In the active site cavity two strictly conserved lysine residues are responsible for initial substrate binding. Single-turnover experiments have shown that the ALA molecule that ultimately forms the ‘propionate’ group of the product PBG binds first to the so-called ‘P’-site. Subsequently, the ‘acetate’ side chain forming ALA molecule binds to the ‘A’-site.<sup>55</sup> The catalytic mechanism of PBGS has been elucidated by enzymatic and structural studies using different inhibitors of the protein.<sup>56–58</sup> After the initial Schiff base formation between the two ALA molecules and the active site lysine residues the reaction proceeds via intersubstrate C–C bond formation between C3 of the A-site ALA and C4 of the P-site ALA. Subsequently, the C–N bond between the two substrate molecules is formed and final deprotonation of the former P-site ALA yields the reaction product PBG (Figure 8(b)).



**Figure 7** Eight ALA molecules are converted into uroporphyrinogen III by three consecutive enzymatic steps via porphobilinogen and preuroporphyrinogen. These reactions are catalyzed by PBGS, PBGD, and UROS. A = acetate, P = propionate.

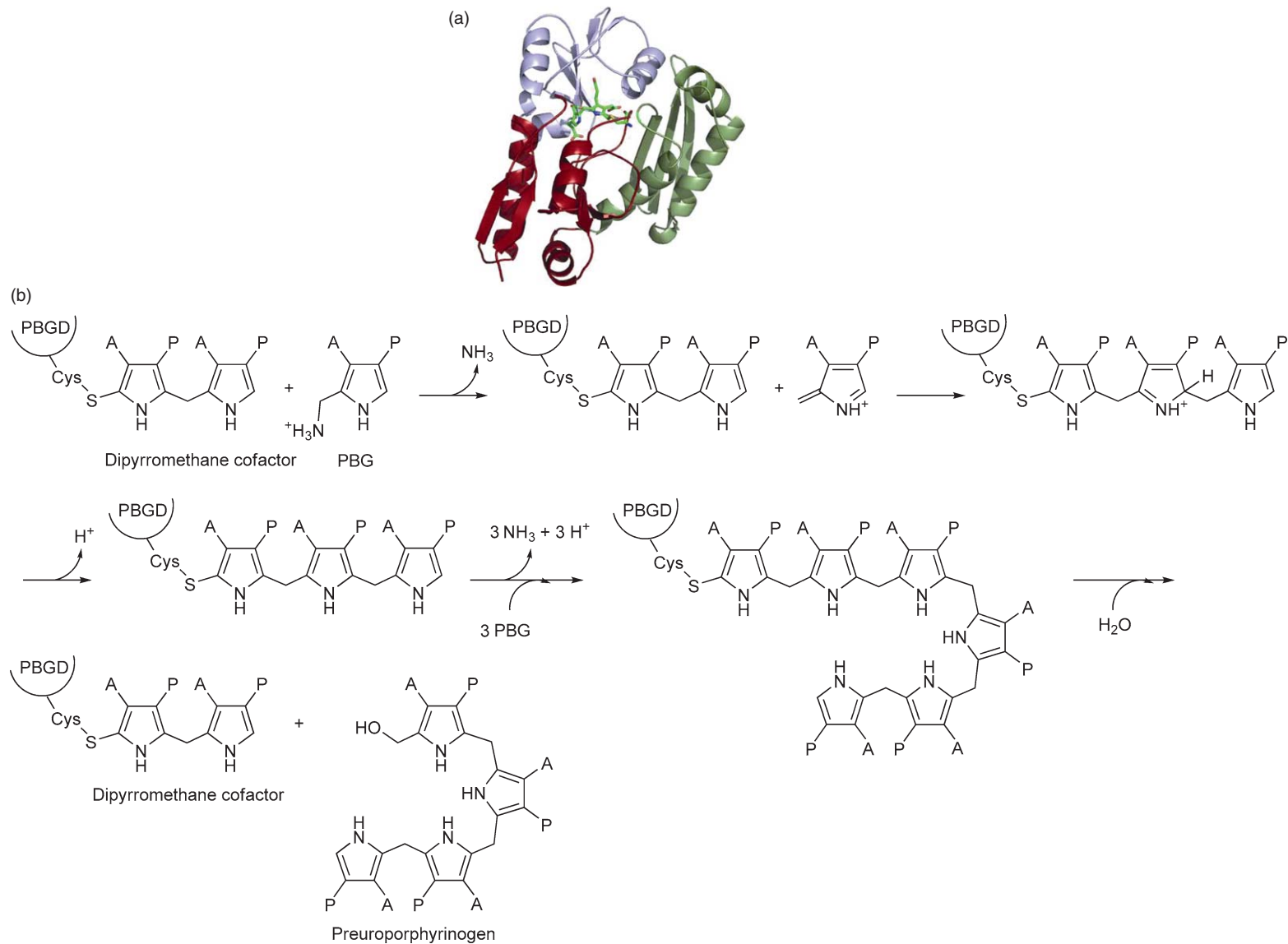


### 7.13.4.2 Porphobilinogen Deaminase

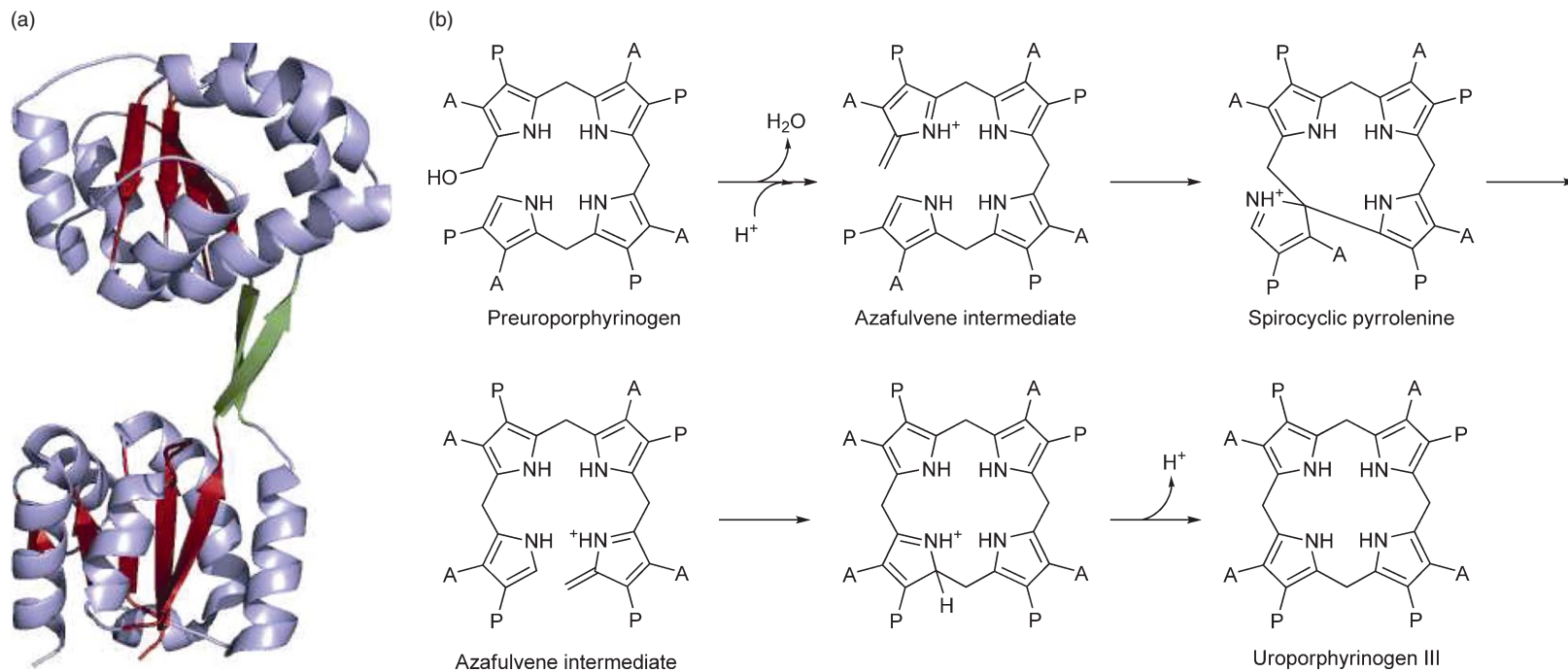
PBGD (EC 4.3.1.8), also known as hydroxymethylbilane synthase, catalyzes the consecutive condensation of four porphobilinogen molecules to form the linear tetrapyrrole preuroporphyrinogen (1-hydroxymethylbilane). PBGD proteins have been purified from various sources and biochemically characterized.<sup>59–62</sup> PBGD is a monomeric protein containing a unique dipyrromethane cofactor. This cofactor consists of two molecules of the substrate PBG and remains covalently attached to the protein via a thioether linkage to a conserved cysteine residue during the protein's lifetime.<sup>63–66</sup> The crystal structure of PBGD from *E. coli* has been solved and revealed a monomeric protein, which consists of three approximately equal-sized domains.<sup>67</sup> The dipyrromethane cofactor is positioned at the opening of the active site cleft between domains I and II (**Figure 9(a)**). At the beginning of the PBGD catalytic cycle the dipyrromethane cofactor serves as a primer for the sequential polymerization of the four following PBG molecules (**Figure 9(b)**). It has been shown that the pyrrole forming ring A of the final cyclic tetrapyrrole arrives first at the enzyme followed by rings B, C, and D.<sup>68</sup> A conserved aspartate residue is involved in catalysis putatively by deaminating PBG.<sup>69</sup> At present, it is not clear how the enzyme exactly deals with the growing reaction intermediates in order to position them correctly in the active site for the new incoming PBG molecules to be attached. It has been observed that PBGD undergoes conformational changes during the tetramerization of PBG, which could be either due to the enzyme 'pulling' the growing chain of pyrroles through the active site or to a reorganization of the oligopyrrole as it is assembled. The catalytic cycle of PBGD ends with the cleavage of the linear tetrapyrrole preuroporphyrinogen from the cofactor dipyrromethane.

### 7.13.4.3 Uroporphyrinogen III Synthase

UROS (EC 4.2.1.75) catalyzes the cyclization of the linear tetrapyrrole preuroporphyrinogen under inversion of the D ring to yield the asymmetric cyclic tetrapyrrole uroporphyrinogen III.<sup>70,71</sup> In the absence of UROS preuroporphyrinogen cyclizes nonenzymatically without inversion of the D ring to the nonphysiological uroporphyrinogen I, which constitutes a toxic dead end product since it cannot be further metabolized. This implies that PBGD and UROS have to act in a highly coordinated fashion to achieve uroporphyrinogen III formation. Indeed, in many bacteria the genes encoding PBGD (*bemC*) and UROS (*bemD*) are organized in an operon resulting in coexpression of both genes.<sup>72,73</sup> Additionally, complex formation between the two proteins to allow a direct transfer of preuroporphyrinogen from PBGD to UROS has been proposed. UROS proteins from different organisms have been purified and characterized.<sup>74–77</sup> Crystal structures have been determined for the human and the *Thermus thermophilus* enzymes.<sup>78,79</sup> Biochemical as well as structural studies showed that UROS is a monomeric protein (**Figure 10(a)**). The enzyme consists of two  $\alpha\beta\alpha$  sandwich domains that are connected with each other by a long two-stranded  $\beta$ -sheet. The human UROS structure shows an 'open' conformation in which the two domains are separated by a large cleft whereas the *T. thermophilus* enzyme showed a more 'closed' conformation. Conformational changes might be necessary in order to correctly accommodate and orientate the tetrapyrrole substrate and to allow catalysis to proceed. Currently, it is believed that catalysis involves formation of a spirocyclic pyrrolenine intermediate. The proposed reaction mechanism has been confirmed *inter alia* by <sup>13</sup>C-NMR experiments and by enzymatic studies using synthetic spirocyclic intermediates that act as inhibitors for UROS.<sup>80–84</sup> The first step of the UROS catalytic cycle consists of a rearrangement of ring A resulting in a positive charge on the pyrrole nitrogen, loss of the A ring hydroxyl group as a water molecule, and formation of the first azafulvene intermediate (**Figure 10(b)**). Subsequently, the A ring azafulvene attacks the substituted  $\alpha$  position of the D ring pyrrole after bond rearrangement on the latter to yield the spirocyclic pyrrolenine intermediate. This spirocyclic intermediate collapses by bond breakage between the C and the D ring with the formation of another azafulvene intermediate on ring C. In order to allow an attack of the C ring azafulvene on the D ring the enzyme has to orientate the D ring such that its free  $\alpha$ -position is in proximity to the C ring azafulvene. Finally, after the attack of the C ring azafulvene on the D ring the reaction product uroporphyrinogen III is formed by deprotonation and bond rearrangements.



**Figure 9** Structure and reaction mechanism of PBGD. (a) Crystal structure of *Escherichia coli* PBGD. Monomeric PBGD consists of three approximately equal-sized domains (red, blue, and green). The PBGD structure contains the covalently bound dipyrromethane cofactor (sticks representation). (b) The PBGD reaction proceeds via formation of an azafulvene intermediate. A = acetate, P = propionate.



**Figure 10** Structure and reaction mechanism of UROS. (a) Crystal structure of human UROS. Monomeric UROS consists of two similar domains, which are connected to each other by a two-stranded  $\beta$ -sheet. (b) The proposed reaction mechanism for UROS involves formation of an azafulvene intermediate, a spirocyclic pyrrolenine intermediate, and a second azafulvene intermediate before the product uroporphyrinogen III is released. A = acetate, P = propionate.



### 7.13.5 Tetrapyrrole Biosynthesis Branch Point – Uroporphyrinogen III

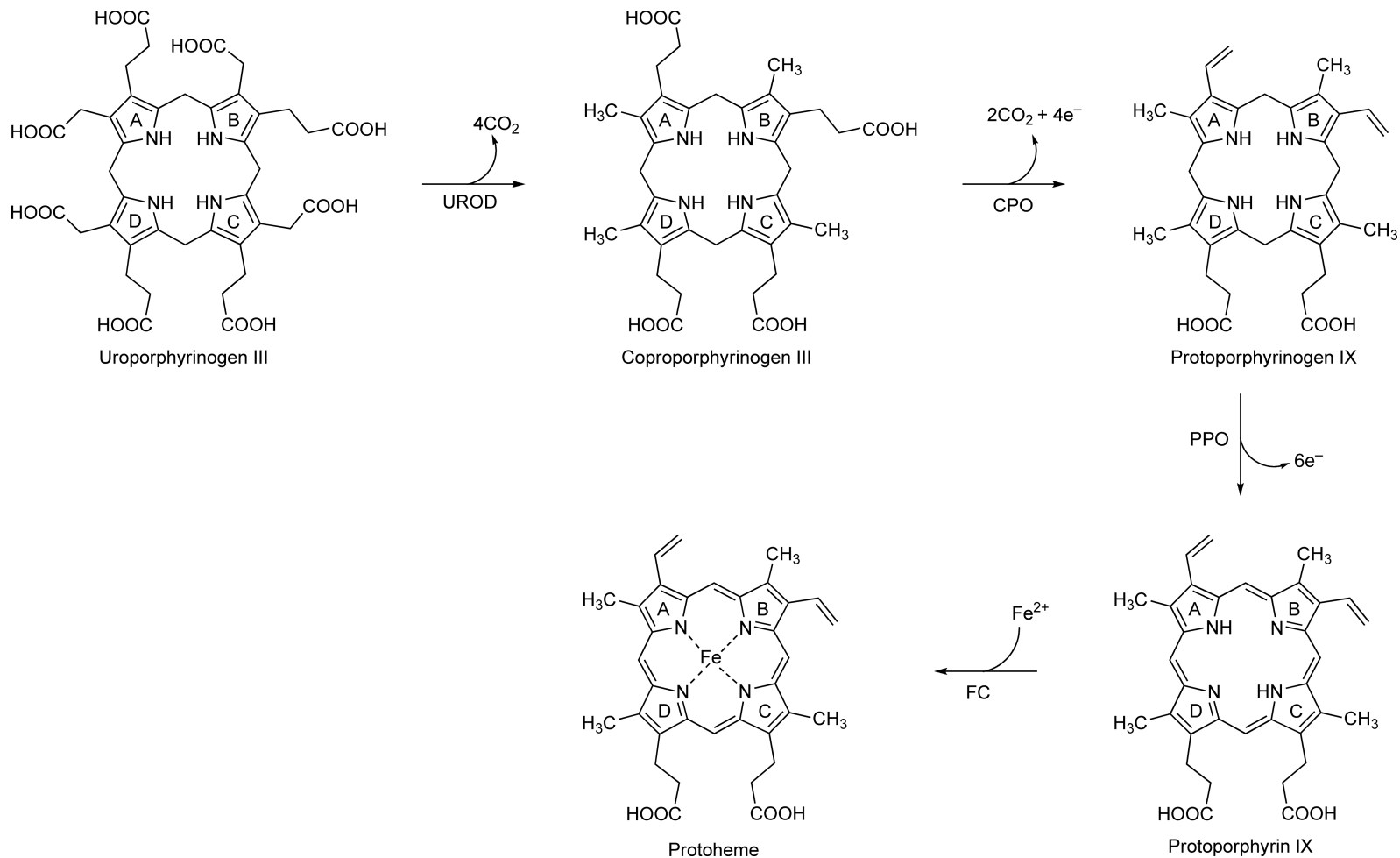
Uroporphyrinogen III, the first cyclic tetrapyrrole of the pathway, constitutes the last common precursor for all naturally occurring tetrapyrroles. Uroporphyrinogen III is converted either into hemes and chlorophylls or into cobalamin, siroheme, heme *d*<sub>1</sub>, and coenzyme F<sub>430</sub>. For heme and chlorophyll biosyntheses the next branching point occurs at the stage of protoporphyrin IX. Iron insertion into protoporphyrin IX leads to the formation of hemes whereas magnesium insertion leads to the biosynthesis of chlorophylls. On the path toward cobalamin, siroheme, heme *d*<sub>1</sub>, and coenzyme F<sub>430</sub> uroporphyrinogen III is first converted into precorrin-2 and at this point the different biosyntheses diverge toward their final products (Figure 2).

### 7.13.6 Conversion of Uroporphyrinogen III into Protoheme

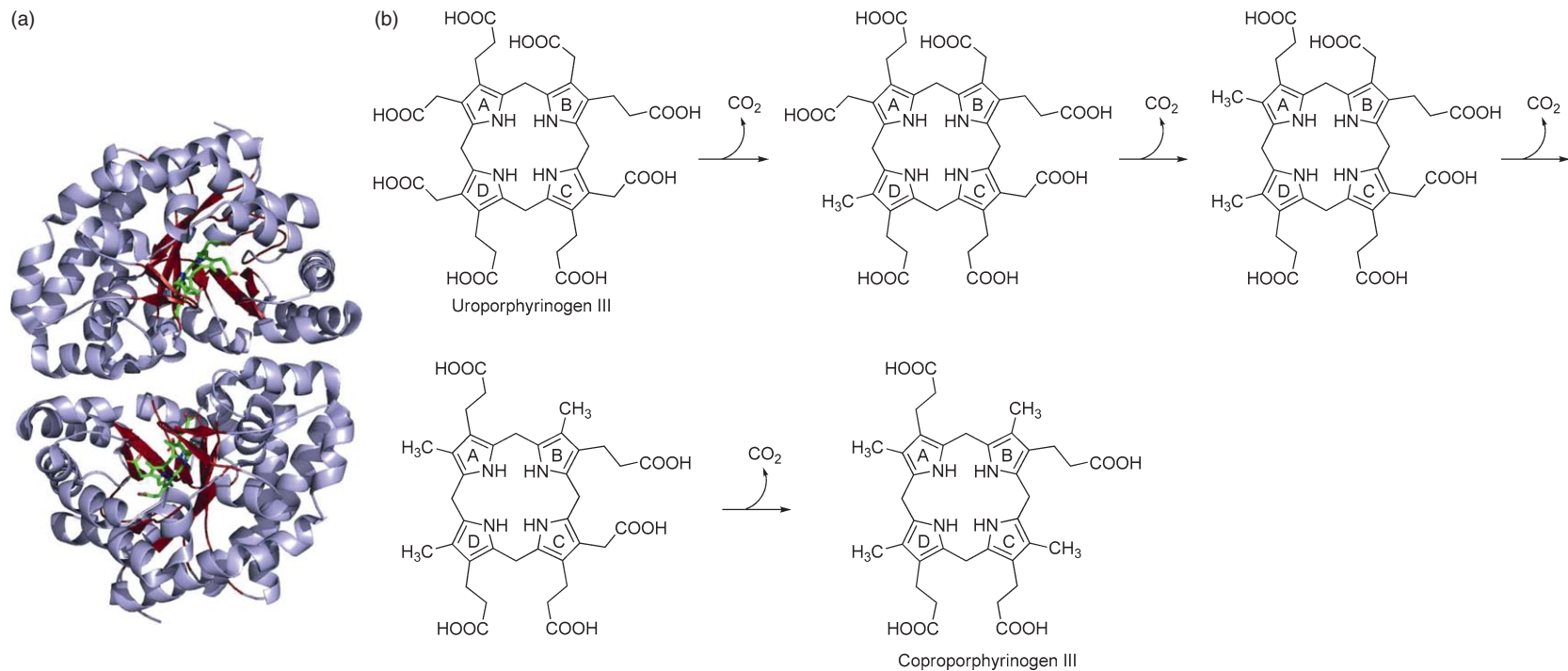
The transformation of uroporphyrinogen III into protoheme requires four consecutive enzymatic steps in which the side chains of the tetrapyrrole are modified, the oxidation state of the macrocycle is changed, and, finally, iron is inserted (Figure 11). First, the four acetate side chains of uroporphyrinogen III are decarboxylated to yield the corresponding methyl groups of coproporphyrinogen III. This reaction is catalyzed by uroporphyrinogen III decarboxylase (UROD). Next, coproporphyrinogen III oxidase (CPO) catalyzes the oxidative decarboxylation of the A and B ring propionate side chains of coproporphyrinogen III with the formation of vinyl groups at these positions. Protoporphyrinogen IX oxidase (PPO) converts protoporphyrinogen IX into protoporphyrin IX. This reaction step requires the elimination of six electrons in order to obtain the system of completely conjugated double bonds in protoporphyrin IX. Finally, ferrochelatase (FC) is responsible for the insertion of iron into the protoporphyrin IX macrocycle to yield the final product protoheme.

#### 7.13.6.1 Uroporphyrinogen III Decarboxylase

UROD (EC 4.1.1.37) catalyzes the conversion of uroporphyrinogen III into coproporphyrinogen III involving the decarboxylation of four acetate side chains to methyl groups.<sup>85</sup> UROD proteins have been isolated from various sources.<sup>86–89</sup> If physiological concentrations of the substrate uroporphyrinogen III are present the four decarboxylations proceed in an ordered manner beginning with the acetate side chain on ring D followed by those of rings A, B, and finally C (Figure 12(b)). Under conditions of high substrate concentrations the decarboxylation reactions take place randomly.<sup>90</sup> Up to now, no direct evidence has been found for the requirement of a cofactor or a prosthetic group, making UROD unique among the enzyme class of decarboxylases. Crystal structures are available for the human, tobacco, and *Bacillus subtilis* enzymes.<sup>91–93</sup> These structures show a dimeric protein in which each subunit consists of a ( $\beta\alpha$ )<sub>8</sub> barrel (Figure 12(a)). The active site cleft contains several invariant polar residues, which are probably involved in binding the carboxyl groups of the substrate. Furthermore, a highly conserved aspartate residue was proposed to be involved in catalysis. The structure of human UROD, in complex with the reaction product coproporphyrinogen III, revealed that the porphyrinogen is bound to the protein in a domed conformation that places the central NH groups of the pyrrole rings in optimal hydrogen bonding distance to this conserved aspartate residue.<sup>94</sup> The two monomers of the UROD dimer are in a head-to-head orientation with the active site clefts at the dimer interface resulting in one large active site that is well shielded from the surrounding solvent. Based on the dimeric architecture of UROD one potential reaction mechanism was proposed in which decarboxylation of the D ring occurs in the active site of one monomer, while the following decarboxylations of rings A, B, and C take place in the active site of the second monomer. Such a mechanism avoids the otherwise necessary 180° flipping of the substrate within a single active site.<sup>92</sup> Another proposal for a potential mechanism is based on the observation of the domed conformation of the porphyrinogen in the human UROD structure. Here, the decarboxylations take place in one single active site via rotation of the reaction intermediates by 90°, respectively, assuming that there is little specific interaction between the tetrapyrrole side chains and the amino acid residues of the active site.<sup>94</sup>



**Figure 11** Uroporphyrinogen III is converted into protoheme by four consecutive enzymatic steps. Decarboxylation of uroporphyrinogen III gives rise to coproporphyrinogen III, which in turn is oxidatively decarboxylated to protoporphyrinogen IX. Oxidation of protoporphyrinogen IX produces protoporphyrin IX and final iron insertion gives rise to protoheme. These four reactions steps are catalyzed by UROD, CPO, PPO, and FC, respectively.



**Figure 12** Structure and enzymatic reaction of UROD. (a) Crystal structure of human UROD in complex with the reaction product coproporphyrin III. Each subunit of the dimeric protein consists of a  $(\beta\alpha)_8$  barrel. (b) During the UROD reaction the acetate side chains of uroporphyrin III are decarboxylated to yield the corresponding methyl groups. The decarboxylations proceed in an ordered manner beginning with ring D followed by rings A, B, and finally C.

### 7.13.6.2 Coproporphyrinogen III Oxidase

CPO (EC 1.3.3.3) catalyzes the oxidative decarboxylation of coproporphyrinogen III to protoporphyrinogen IX. During this reaction the propionate side chains on rings A and B of the tetrapyrrole are converted into the corresponding vinyl groups with the release of CO<sub>2</sub> and the elimination of four electrons.<sup>95</sup> Two structurally unrelated proteins are found in nature for the oxidative decarboxylation of coproporphyrinogen III. In mammals and a few bacteria the oxygen-dependent CPO HemF is responsible for the conversion of coproporphyrinogen III into protoporphyrinogen IX, whereas in most bacteria the oxygen-independent CPO HemN catalyzes this reaction.<sup>96</sup>

HemF is a dimeric, membrane-associated protein and requires molecular oxygen as the terminal electron acceptor for catalysis.<sup>97–101</sup> Early studies have shown that the propionate side chain on ring A is decarboxylated prior to that on ring B with the monovinyl intermediate harderoporphyrinogen.<sup>102</sup> Crystal structures are available for the human and yeast proteins.<sup>103,104</sup> Both structures confirmed the dimeric state of HemF (**Figure 13(a)**). Each monomer shows an unusual fold consisting of a large flat seven-stranded  $\beta$ -sheet flanked on both sides by  $\alpha$ -helices. In the homodimer the two subunits are inclined by 40° to each other. In the yeast structure an ‘open’ and a ‘closed’ conformation of the active site was observed. In the closed form the active site size corresponds to the size of the substrate. Amino acid residues putatively involved in substrate binding and catalysis were assigned by modeling the substrate into the structure. In the human CPO structure a citrate molecule was found bound in the active site, which also allowed the assignment of several amino acid residues potentially involved in catalysis. Both HemF crystal structures were found not to contain any cofactor, prosthetic group, or metal ion. In contrast, the recombinant mouse HemF was described to contain copper and the *E. coli* HemF was shown to be stimulated by manganese.<sup>97,105</sup> These contradictory observations have raised questions as to how molecular oxygen is utilized as the terminal electron acceptor. Overall, the exact reaction mechanism by which HemF catalyzes coproporphyrinogen III decarboxylation remains to be determined.

The oxygen-independent CPO HemN catalyzes the two decarboxylation reactions employing an alternative terminal electron acceptor. The enzyme from *E. coli* is a monomeric protein and contains an oxygen-labile [4Fe–4S] cluster, which is coordinated by three highly conserved cysteine residues. Further, HemN requires *S*-adenosyl-L-methionine (SAM), NADH, and some components of an *E. coli* cell-free extract for catalysis.<sup>106</sup> HemN was identified as belonging to a class of so-called ‘radical SAM’ enzymes.<sup>107</sup> All these enzymes contain an iron–sulfur cluster and use SAM during catalysis for the generation of catalytically active radicals. The catalytic cycle of all radical SAM enzymes starts with the one electron reduction of the iron–sulfur cluster by an external electron donor like reduced flavodoxin or ferredoxin. The reduced iron–sulfur cluster then transfers the electron to SAM inducing the homolytic cleavage of SAM to methionine and a highly reactive 5′-deoxyadenosyl radical. This radical then abstracts a hydrogen atom from either the enzyme’s substrate or from a glycine residue on a partner protein to generate substrate or protein-localized radicals, respectively. After these common initial steps the final reactions are different for each radical SAM enzyme. The SAM molecule is used either as a cofactor and is regenerated at the end of each catalytic cycle or it is used as a cosubstrate with the release of methionine and 5′-deoxyadenosine as side products of the catalyzed reaction.<sup>108</sup> For HemN it has been observed that methionine is produced as a reaction side product with a 2:1 (methionine/protoporphyrinogen IX) ratio indicating that each of the two decarboxylation reactions requires the cleavage of one SAM molecule.<sup>109</sup> For HemN it has been proposed that the 5′-deoxyadenosyl radical abstracts stereospecifically the pro-*S*-hydrogen atom from the  $\beta$ -C atom of the substrate propionate side chain resulting in a substrate radical. Indeed, formation of such a substrate radical was observed for the HemN reaction by EPR spectroscopy. Using deuterium-labeled substrates this radical was characterized.<sup>110</sup> After formation of the substrate radical the decarboxylation reaction is then completed by elimination of an electron and CO<sub>2</sub> (**Figure 13(c)**). The final electron acceptor for the HemN-catalyzed reaction has not been identified yet, but explains the requirement for a cell-free extract in HemN activity assays. For *E. coli* HemN the crystal structure is available.<sup>111</sup> The monomeric enzyme consists of two distinct domains (**Figure 13(b)**). The N-terminal domain, which contains the active site, is characterized by a curved, 12-stranded  $\beta$ -sheet surrounded by  $\alpha$ -helices. The N-terminal  $(\beta/\alpha)_6$  repeat shows some resemblance to TIM barrel domains although in HemN the  $\beta$ -strands are less strongly inclined relative to the barrel axis and the curvature of

the  $\beta$ -sheet is not as tight as in real TIM barrels. The C-terminal domain consists of a bundle of four  $\alpha$ -helices and a small  $\beta$ -sheet. The iron–sulfur cluster is located deep inside the active site near the C-terminal ends of the parallel  $\beta$ -strands of the  $(\beta/\alpha)_6$  repeat. Interestingly, there were two SAM molecules bound in the HemN structure, one of them coordinating an iron atom of the iron–sulfur cluster. The other was bound in close proximity to the first. The two bound SAM molecules correspond well with the observation that two SAM molecules are cleaved during the HemN reaction.

### 7.13.6.3 Protoporphyrinogen IX Oxidase

The penultimate step of heme biosynthesis is catalyzed by PPO (EC1.3.3.4) and consists of the oxidation of the colorless protoporphyrinogen IX to the red colored protoporphyrin IX.<sup>95</sup> During this reaction six electrons have to be eliminated. In the absence of PPO, protoporphyrinogen IX slowly auto-oxidizes to protoporphyrin IX, however, *in vivo* this reaction is enzyme catalyzed. There exist at least two unrelated proteins in nature to catalyze the protoporphyrinogen IX oxidation, an oxygen-dependent PPO that uses molecular oxygen as the terminal electron acceptor for the reaction and an oxygen-independent PPO that uses alternative electron acceptors. Oxygen-dependent PPO has been described as a soluble monomer for the *B. subtilis* protein,<sup>112</sup> as a membrane-associated monomer in bovine<sup>113</sup> or as a homodimer in *Mycococcus xanthus* and humans.<sup>114,115</sup> In each case PPO contains a noncovalently bound FAD as a cofactor. Crystal structures are available for the PPOs from tobacco and *M. xanthus*.<sup>116,117</sup> The tobacco PPO structure revealed a dimeric protein in which each subunit is composed of three domains, a membrane-, an FAD-, and a substrate-binding domain (Figure 14(a)). Based on modeling the substrate into the structure a potential reaction mechanism has been proposed in which all hydride abstractions occur from C20 after imine–enamine tautomerizations (Figure 14(b)). It was also suggested that the reaction product protoporphyrin IX is transported through a channel in PPO directly to FC without exposure to the surrounding medium. Indeed, it was possible to dock the structure of tobacco PPO with that of human FC *in silico*.<sup>117</sup>

Much less is known about the oxygen-independent PPO. In some bacteria like *E. coli* there is no homologue to the oxygen-dependent enzyme; however there exists a gene (*bemG*) that encodes a 21 kDa protein.<sup>118</sup> An *E. coli bemG* mutant was deficient in PPO activity but could be complemented with clones encoding the oxygen-dependent PPO from tobacco and mouse.<sup>119,120</sup> The oxygen-independent PPO couples protoporphyrinogen IX oxidation to the respiratory chain and for this several different terminal electron acceptors can be used.<sup>96</sup>

### 7.13.6.4 Ferrochelatase

The last step of heme biosynthesis, the insertion of ferrous iron into protoporphyrin IX, is catalyzed by FC (EC 4.99.1.1).<sup>121</sup> This enzyme has been purified and biochemically and structurally characterized from various sources. The *B. subtilis* FC was found a soluble, monomeric protein,<sup>122</sup> whereas the enzymes from eukaryotes are membrane-associated dimers.<sup>123,124</sup> FC requires no cofactors for activity. However, in the human protein and some other eukaryotic FCs an iron–sulfur cluster was found in the C-terminal part of the protein.<sup>124</sup> FC can use a range of different porphyrins as substrates.<sup>125</sup> It was observed that the side chains on rings A and B can be varied to a certain degree, whereas the propionate side chains on rings C and D appear to be essential for enzyme activity probably by sustaining the correct orientation of the substrate within the active site. FC can also use other divalent metal ions, like Co<sup>2+</sup> or Zn<sup>2+</sup>, but ferric iron is not a substrate.<sup>126</sup> Crystal structures for FC are available for the human, yeast, and *B. subtilis* enzymes.<sup>122,124,127</sup> The structure of the *B. subtilis* monomer and the subunits of the dimeric eukaryotic FCs are similar. The FC monomer is composed of two domains each containing a four-stranded parallel  $\beta$ -sheet and five or six  $\alpha$ -helices. Recently, human FC has been cocrystallized with the substrate protoporphyrin IX.<sup>128</sup> The substrate was observed to bind deep in an enclosed pocket (Figure 15). Other structures containing bound metal ions (Mg<sup>2+</sup>, Ni<sup>2+</sup>, or Co<sup>2+</sup>) helped to define the metal-binding site of the enzyme. A potential reaction mechanism was proposed for the iron insertion into protoporphyrin IX in which the enzyme causes a distortion of the porphyrin ring to a nonplanar conformation, which allows the insertion of the metal ion.<sup>129</sup>

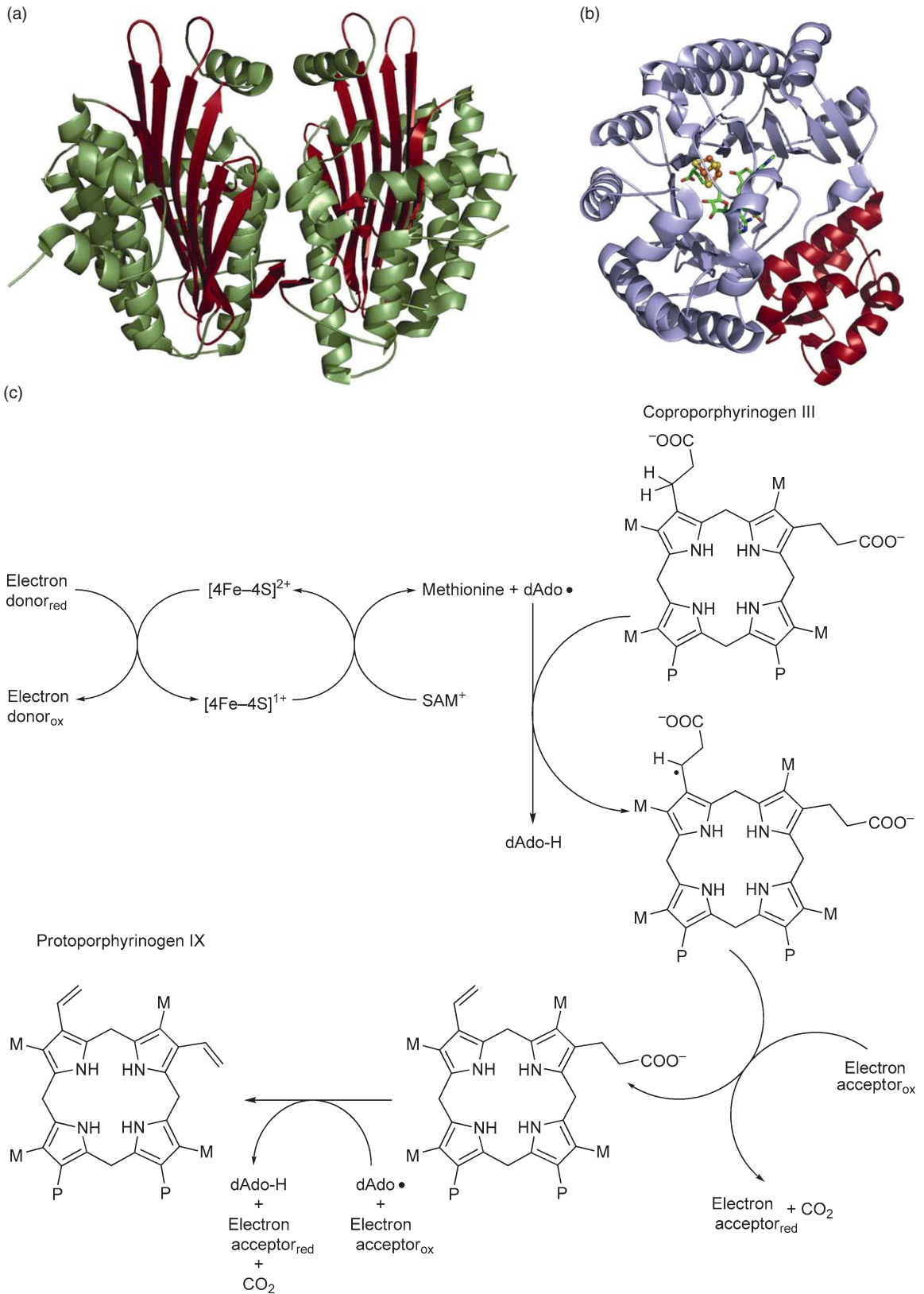


Figure 13



### 7.13.7 The Biosynthesis of Adenosylcobalamin (Vitamin B<sub>12</sub>)

Vitamin B<sub>12</sub> represents one of the most complex small molecules found in nature. It consists of a tetrapyrrole-derived cobalt-containing corrin ring, which is attached to a lower nucleotide with an unusual base called dimethylbenzimidazole (**Figure 16**). The cobalt ion is coordinated with the pyrrole-derived nitrogens of the corrin and with the nitrogen of the lower base. The upper axial ligand for the cobalt can be water (aquacobalamin), a methyl group (methylcobalamin), adenosine (adenosylcobalamin), or cyanide (vitamin B<sub>12</sub> or cyanocobalamin). The biological forms of cobalamin are methylcobalamin and adenosylcobalamin and it is the properties of the cobalt–carbon bond in these species that are exploited by living systems to enhance a range of metabolic processes.<sup>1</sup>

The complete biosynthesis of cobalamin requires somewhere around 30 enzyme-mediated steps, which compares favorably to the total chemical synthesis of the compound where, for instance, 37 steps are involved in the photochemical route to synthetic cobyrinic acid,<sup>130</sup> an intermediate that represents only the corrin component of the final cobalamin molecule. For this reason vitamin B<sub>12</sub> is produced commercially by bacterial culture and is the most expensive of the water-soluble vitamins to produce because of the comparatively low yields obtained.

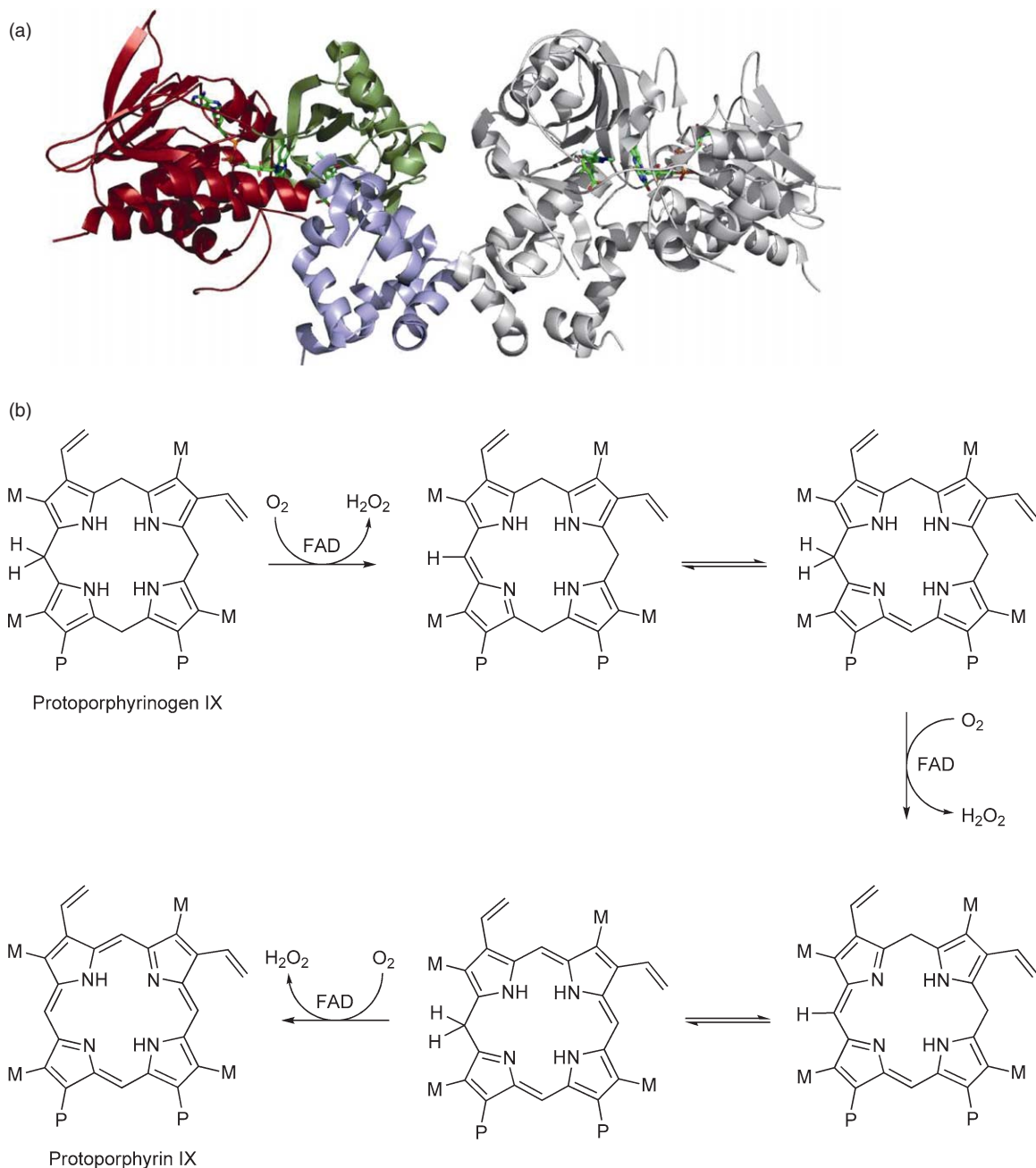
The cobalamin branch of the modified tetrapyrrole pathway diverges from the main porphyrin route at uroporphyrinogen III and shares some commonality with the pathway for siroheme (**Figure 17**). This iron-containing isobacteriochlorin is constructed from uroporphyrinogen III in three steps as outlined in **Figure 17**. First, uroporphyrinogen III is methylated at positions 2 and 7 to generate precorrin-2. Second, precorrin-2 undergoes a dehydrogenation to give sirohydrochlorin (factor II) and finally sirohydrochlorin is chelated with iron to give siroheme. The reason why the siroheme pathway is closely integrated with cobalamin biosynthesis is that the two intermediates of the siroheme pathway are also privy to the two known cobalamin biosynthetic pathways. Cobalamin biosynthesis is mediated along either an aerobic pathway from precorrin-2 or from sirohydrochlorin along an anaerobic pathway (**Figure 17**). The two pathways are described in more detail below.

### 7.13.8 A Note on Nomenclature

Over the years there has been a tendency to be less precise over the naming of the many compounds associated with modified tetrapyrrole biosynthesis. Perhaps the most common mistake is to refer to the biologically active forms of vitamin B<sub>12</sub>, adenosylcobalamin and methylcobalamin, as vitamin B<sub>12</sub>. Adenosylcobalamin is the coenzyme form of vitamin B<sub>12</sub> and it is used in several rearrangement reactions whereas methylcobalamin is the cofactor form of the vitamin and is involved in methyltransfer reactions. Vitamin B<sub>12</sub> refers to cyanocobalamin, which is a nonphysiological form of cobalamin that is generated as it is extracted from biological sources. Moreover, in biological systems other forms of cobalamin are found where different side chains and bases are incorporated into the molecule. For instance, norcobalamin is a form of cobalamin where ethanolamine is found in place of the aminopropanol linker, and pseudocobalamin is where adenine replaces the dimethylbenzimidazole base (**Figure 16**).

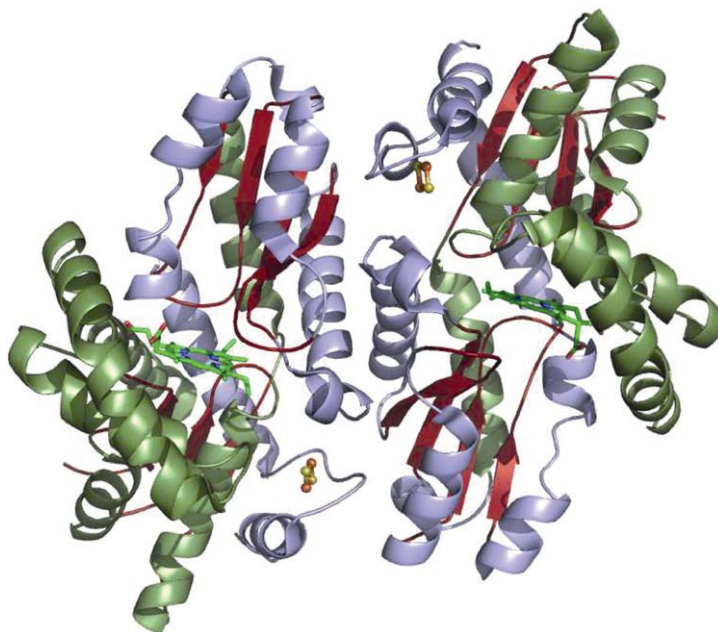
---

**Figure 13** Structures of oxygen-dependent and oxygen-independent CPOs and reaction mechanism for oxygen-independent CPO. (a) Crystal structure of human oxygen-dependent CPO. Each subunit of the dimeric protein shows an unusual fold consisting of a large flat seven-stranded  $\beta$ -sheet flanked on both sides by  $\alpha$ -helices. (b) Crystal structure of *Escherichia coli* oxygen-independent CPO HemN in complex with two SAM molecules. The monomeric protein consists of two domains. The larger N-terminal catalytic domain consists of a curved, 12-stranded  $\beta$ -sheet surrounded by  $\alpha$ -helices (blue). In this domain the iron–sulfur cluster of HemN and the two SAM molecules are bound (ball and stick representation). The C-terminal domain (red) consists of a bundle of four  $\alpha$ -helices and a small  $\beta$ -sheet. (c) Proposed reaction mechanism for the oxygen-independent decarboxylation of coproporphyrinogen III via a radical mechanism. The HemN reaction cycle starts with the formation of a highly reactive 5'-deoxyadenosyl radical (dAdo•), which abstracts the pro-S-hydrogen atom from the  $\beta$ -C atom of the substrate propionate side chain. After formation of the substrate radical CO<sub>2</sub> is released and a terminal electron acceptor is required to take up the remaining electron. Two SAM molecules are required for the two decarboxylation reactions. M = methyl, P = propionate.

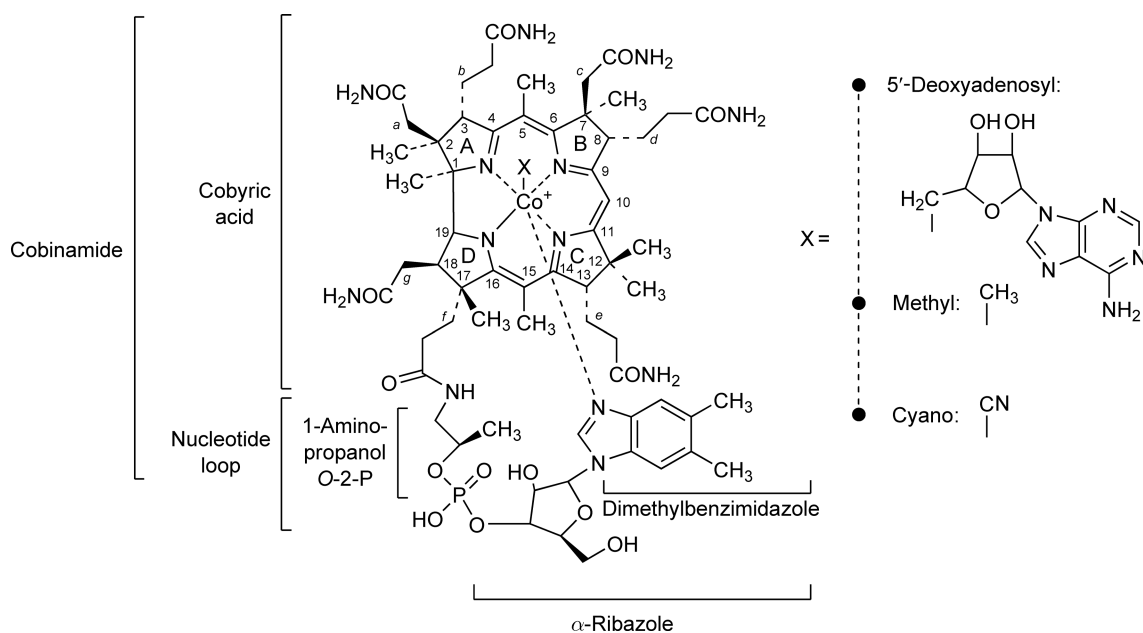


**Figure 14** Structure and reaction mechanism of oxygen-dependent PPO. (a) Crystal structure of tobacco PPO in complex with the cofactor FAD and a phenyl-pyrazol inhibitor. Each subunit of the dimeric protein consists of three domains, a membrane- (blue), an FAD- (red), and a substrate-binding (green) domain. (b) Proposed reaction mechanism for oxygen-dependent PPO. All hydride abstractions occur from C<sub>20</sub> after imine–enamine tautomerizations. M = methyl, P = propionate.

In terms of biosynthesis, some of the more easily-isolated compounds have been given formal names such as hydrogenobyric acid, cobyric acid, and cobinamide. However, as many intermediates proved technically challenging to isolate they were given predictive names that referred to the number of methyl groups that had been attached to the macrocyclic template during the biosynthesis.<sup>131</sup> Therefore, precorrin-2 refers to an early

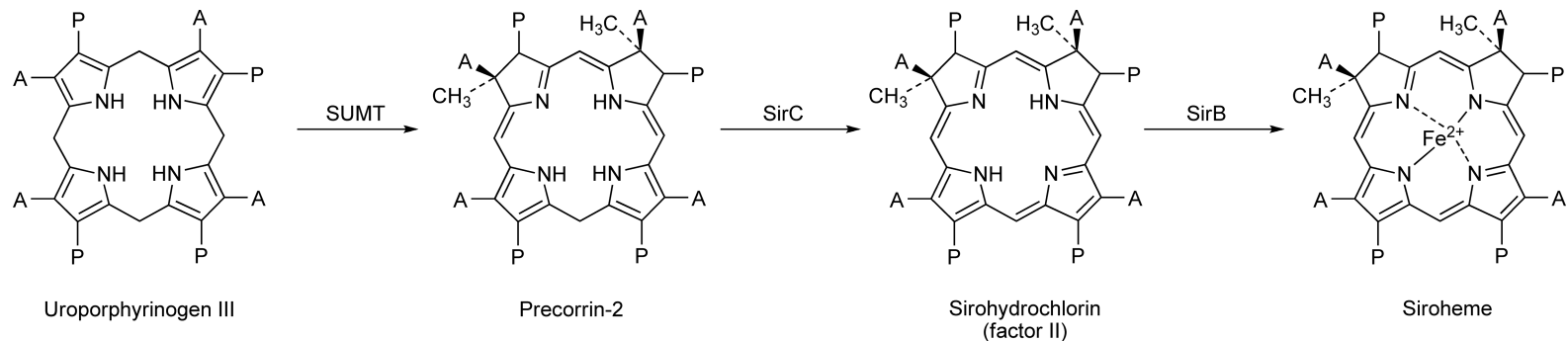


**Figure 15** Crystal structure of human ferrochelatase in complex with the substrate protoporphyrin IX. Each subunit of the dimeric protein consists of two domains composed of a four-stranded  $\beta$ -sheet and five or six  $\alpha$ -helices. The substrate protoporphyrin IX (sticks representation) binds in an enclosed active site pocket.

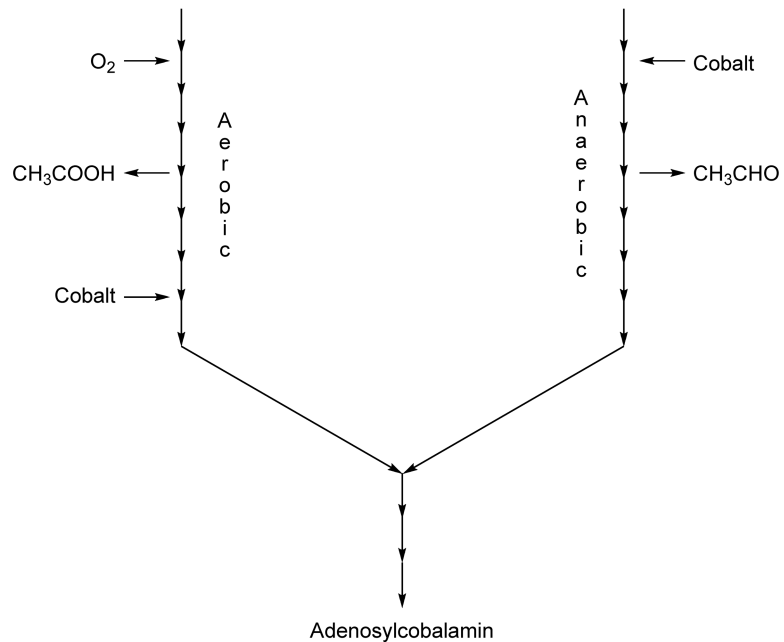


**Figure 16** Structure of cobalamin. The figure highlights some of the names given to the components of the molecule. The numbering of the corrin framework is given as in the lettering of the side chains.

corrinoid intermediate with two methyl groups attached to the ring periphery. If more than one intermediate is found with the same number of methyl groups, then the compounds are further prefixed with A, B, and so on. Thus, precorrin-3A and 3B refer to two intermediates that both contain three added methyl groups yet have quite different structures.



A = CH<sub>2</sub>CO<sub>2</sub>H  
 P = CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H



**Figure 17** The transformation of uroporphyrinogen III into siroheme and how it integrates with cobalamin biosynthesis. Precorrin-2 is the entry point for the aerobic pathway whereas sirohydrochlorin (factor II) is the entry point for the anaerobic pathway.

The isolation of the genetic software for the two cobalamin pathways added another level of confusion.<sup>132–136</sup> The genes for the aerobic pathway were the first to be reported and were given the prefix *cob* (for cobalamin biosynthesis). For the anaerobic pathway, the genes for cobinamide biosynthesis were termed *cbi* yet the genes required for the transformation of cobinamide into adenosylcobalamin were also termed *cob*. Thus there are some genes in the aerobic pathway that have the same names as genes in the anaerobic pathway but actually encode enzymes that have very different activities (see, e.g., *cobA*, *cobC*, and *cobD*).

For the purposes of this review, the aerobic pathway will be exemplified by the research largely undertaken on *Pseudomonas denitrificans*. This is an organism that has been used as a good commercial source of the vitamin. However, its name is not a precise taxonomic description of the organism,<sup>136</sup> as it is clearly not a pseudomonad. In reality, the organism is likely to be a member of the alpha subgroup of the purple bacteria.<sup>137</sup> Although much is known about how the pathway operates in this organism, little is known about what it is made for or how its synthesis is controlled and regulated. In contrast, the anaerobic pathway will be exemplified by research largely undertaken on *Salmonella enterica*. In this organism, there is less detail on the intermediates but significantly more is known about its regulation and requirements.<sup>138</sup> A comparison of the genetic requirements of the two pathways is shown in **Table 1**.

### 7.13.9 Overview of Cobalamin Biosynthesis

The transformation of uroporphyrinogen III into adenosylcobalamin requires the following events:

- The addition of eight SAM-derived methyl groups to the periphery of the corrin ring structure
- Insertion of cobalt
- Ring contraction with extrusion of the methylated C20 carbon as either acetic acid or acetaldehyde
- Decarboxylation of the acetic acid side chain on ring C
- Amidation of six of the eight acidic side chains
- Adenylation of the cobalt to provide the upper axial ligand
- Attachment of aminopropanol (phosphate) to the propionate side chain on ring D
- Synthesis and attachment of the lower nucleotide loop

A summary of these events is outlined in **Figure 18**. Although there are two distinct pathways for cobalamin biosynthesis, there are many similarities between them and the two pathways are evolutionarily related. The major differences are associated with the process of ring contraction and the method for cobalt insertion. The relatedness of some of the enzymes associated with the two pathways is highlighted in **Table 1**. The evolution of complex metabolic pathways raises many interesting issues, some of which are discussed elsewhere.<sup>139</sup>

### 7.13.10 Uroporphyrinogen III to Precorrin-2

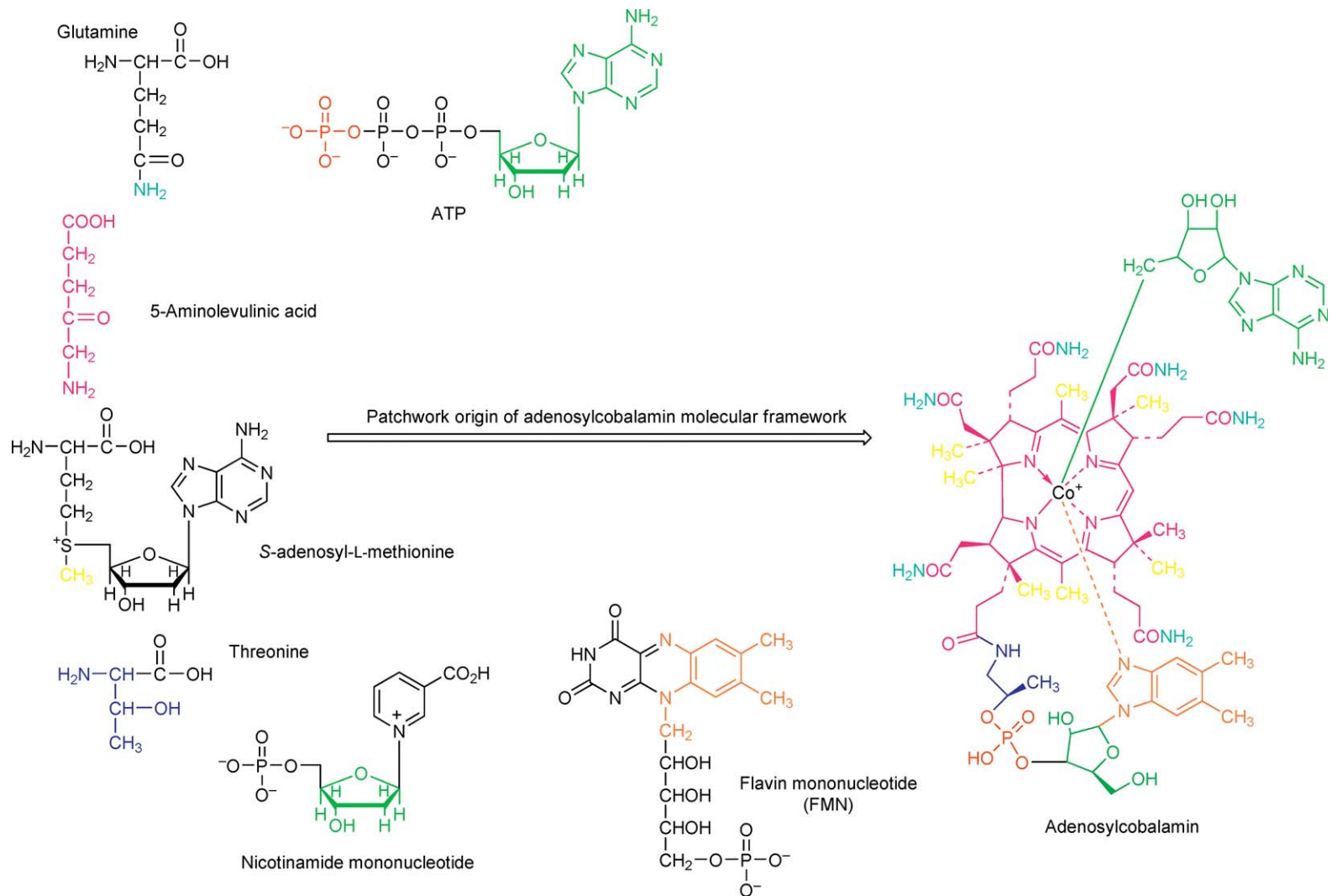
The synthesis of precorrin-2 requires the addition of two methyl groups to the C2 and C7 positions of the uroporphyrinogen III template (**Figure 17**). This reaction is catalyzed by an enzyme called *S*-adenosyl-L-methionine uroporphyrinogen III methyltransferase (SUMT). The enzyme was first isolated to homogeneity from *P. denitrificans*, where kinetic analysis revealed that the enzyme methylates initially at C2 to generate precorrin-1 and subsequently at C7 to produce precorrin-2.<sup>140</sup> The methyl donor is SAM, which is converted into *S*-adenosyl-L-homocysteine (SAH). The enzyme was found to have a  $K_m$  for uroporphyrinogen III of 1.0 and 6.3  $\mu\text{mol l}^{-1}$  for SAM, and a relatively slow  $k_{\text{cat}}$  of 38  $\text{h}^{-1}$ .<sup>140</sup> As with many transmethyases, SUMT displays strong feedback inhibition to SAH, which acts as a powerful competitive inhibitor to SAM, with a  $K_i$  of 0.32  $\mu\text{mol l}^{-1}$ . A further interesting kinetic anomaly of SUMT is that the enzyme is inhibited by the substrate uroporphyrinogen III. This inhibition is observed at physiological levels of uroporphyrinogen III suggesting that the enzyme may be inhibited *in vivo*, thereby implying a regulatory role for the enzyme.<sup>140</sup> An investigation of a SUMT from a methanogen, *Methanobacterium ivanovii*, demonstrated that this substrate inhibition is not present in all such enzymes and is likely dependent on whether uroporphyrinogen represents the main branchpoint intermediate.<sup>141</sup>

**Table 1** Comparison of the genetic and enzymatic requirements for the two cobalamin biosynthetic pathways

<i>Aerobic pathway Pseudomonas denitrificans</i>				<i>Anaerobic pathway Salmonella enterica (LT2)</i>		
<i>Step/intermediate</i>	<i>Enzyme name (EC)</i>	<i>Gene name</i>	<i>% id</i>	<i>Gene name</i>	<i>Enzyme name (EC)</i>	<i>Intermediate/step</i>
<b>Uroporphyrinogen III</b>				<b>Uroporphyrinogen III</b>		
Methylation at C2, C7 ↓	SUMT 2.1.1.107	<i>cobA</i>	43%	<i>cysG</i>	2.1.1.107 1.3.1.76	↓ Methylation at C2, C7 ↓
<b>Dihydrosirohydrochlorin (precorrin-2)</b>				<b>Sirohydrochlorin</b>		
↓				<i>cbiK</i>	4.99.1.3	↓ Cobalt insertion
↓				<b>Cobalt sirohydrochlorin</b>		
Methylation at C20 ↓	SP2MT 2.1.1.130	<i>cobI</i>	29%	<i>cbiL</i>	2.1.1.151	↓ Methylation at C20 ↓
<b>Precorrin-3A</b>				<b>Cobalt-factor III</b>		
Monooxygenation ↓	1.14.13.83	<i>cobG</i>				↓
<b>Precorrin-3B</b>						↓
Methylation at C17 ↓	2.1.1.131	<i>cobJ</i>	40%	<i>cbiH</i>	2.1.1.131	↓ Methylation at C17
<b>Precorrin-4</b>				<b>Cobalt-precorrin-4</b>		
Methylation at C11 ↓	2.1.1.133	<i>cobM</i>	41%	<i>cbiF</i>	2.1.1.133	↓ Methylation at C11
<b>Precorrin-5</b>				<b>Cobalt-precorrin-5a</b>		
Methylation at C1 ↓	2.1.1.152	<i>cobF</i>				↓
Acetic acid extrusion ↓						↓
<b>Precorrin-6A</b>				<i>cbiG</i>	-	↓ Lactone opening ↓ Acetaldehyde extrusion
↓						
↓				<b>Cobalt-precorrin-5B</b>		
↓				<i>cbiD</i>	2.1.1.-	↓ Methylation at C1
↓				<b>Cobalt-precorrin-6A</b>		
C18:C19 reduction ↓	1.3.1.54	<i>cobK</i>	25%	<i>cbiJ</i>	1.3.1.54	↓ C18:C19 reduction
<b>Precorrin-6B</b>				<b>Cobalt-precorrin-6B</b>		
Methylation at C5, C15 ↓	2.1.1.132	<i>cobL</i>	31%	<i>cbiE</i>	2.1.1.132	↓ Methylation at C5 ↓
Decarboxylation ↓						
↓				<b>Cobalt-precorrin-7</b>		
↓			32%	<i>cbiT</i>	2.1.1.-	↓ Methylation at C15 ↓ Decarboxylation
↓						



<b>Precorrin-8</b>					<b>Cobalt-precorrin-8</b>	
Methyl rearrangement ↓	5.4.1.2	<i>cobH</i>	33%	<i>cbiC</i>	5.4.1.2	↓ Methyl rearrangement
<b>Hydrogenobyric acid</b>				<b>Cobyric acid</b>		
<i>a,c</i> -Amidation ↓	6.3.5.9	<i>cobB</i>	34%	<i>cbiA</i>	6.3.1.-	↓ <i>a,c</i> -Amidation
<b>Hydrogenobyric acid <i>a,c</i>-diamide</b>				<b>Cobyric acid <i>a,c</i>-diamide</b>		
Cobalt insertion ↓	6.6.1.2	<i>cobNST</i>				↓
<b>Cobyric acid <i>a,c</i>-diamide</b>						↓
Cobalt reduction ↓	1.14.13.3	<i>cobR</i>	-	<i>fldA</i>		↓ Cobalt reduction
<b>Cob(II)yrinic acid <i>a,c</i>-diamide</b>				<b>Cob(II)yrinic acid <i>a,c</i>-diamide</b>		
Adenosylation ↓	2.5.1.17	<i>cobO</i>	42%	<i>cobA</i>	2.5.1.17	↓ Adenosylation
<b>Adenosylcobyric acid <i>a,c</i>-diamide</b>				<b>Adenosylcobyric acid <i>a,c</i>-diamide</b>		
<i>b,d,e,g</i> -Amidation ↓	6.3.5.10	<i>cobQ</i>	44%	<i>cbiP</i>	6.3.5.10	↓ <i>b,d,e,g</i> -Amidation
<b>Adenosylcobyric acid</b>				<b>Adenosylcobyric acid</b>		
<i>f</i> -side chain attachment ↓	6.3.1.10	<i>cobD</i>	34%	<i>cbiB</i>	6.3.1.10	↓ <i>f</i> -side chain attachment
<b>Adenosylcobinamide (phosphate)</b>				<b>Adenosylcobinamide phosphate</b>		
Phosphorylation ↓	2.7.1.156	<i>cobP</i>	39%	<i>cobU</i>	2.7.1.156	↓ Phosphorylation
GTP addition ↓	2.7.7.62				2.7.7.62	↓ GTP addition
<b>Adenosyl-GDP-cobinamide (phosphate)</b>				<b>Adenosyl-GDP-cobinamide phosphate</b>		
$\alpha$ -Ribazole attachment ↓		<i>cobV</i>	34%	<i>cobS</i>	2.7.8.26	↓ $\alpha$ -Ribazole attachment
<b>Adenosylcobalamin 5'-phosphate</b>				<b>Adenosylcobalamin 5'-phosphate</b>		
Dephosphorylation ↓		?		<i>cobC</i>		↓ Dephosphorylation
<b>Adenosylcobalamin</b>				<b>Adenosylcobalamin</b>		

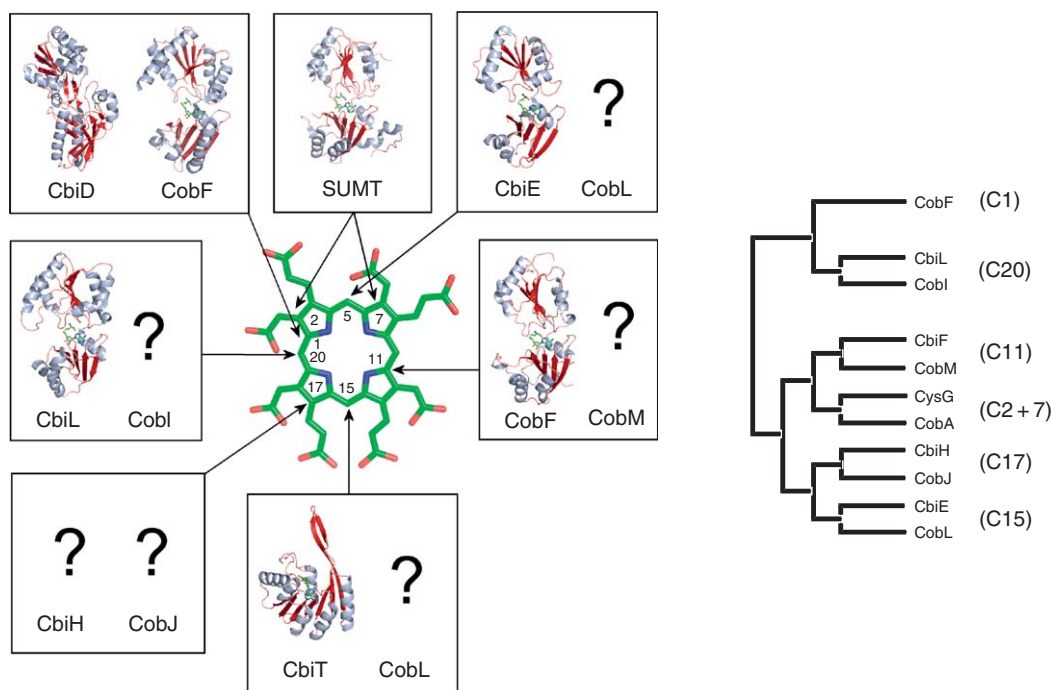


**Figure 18** Origin of the molecular framework of cobalamin. Adenosylcobalamin is derived from aminolevulinic acid, SAM, cobalt, glutamine, threonine, ATP, NaMN, and FMN. The diagram is color coded for easy interpretation.

Gene sequencing of all the cobalamin biosynthetic genes in *P. denitrificans* revealed that SUMT was encoded by *cobA*.<sup>133</sup> The encoded protein contains 259 amino acid with an  $M_r$  of 29 k. Sequence comparisons revealed that SUMT has varying degrees of similarity with five other enzymes in the *P. denitrificans* cobalamin pathway, including CobI, CobJ, CobM, CobF, and CobL. These five enzymes are also methyl transferases that belong to the same class (class III) of SAM-dependent methyltransferases, which share a similar overall topology (where the active site is formed between two  $\alpha\beta\alpha$  domains) (Figure 19).<sup>142</sup> The sequence similarity between these enzymes indicates that they have evolved from a common ancestor. The crystal structure of the *P. denitrificans* SUMT was solved revealing that the protein had crystallized with SAH bound in the active site.<sup>143</sup> As a class III methyltransferase,<sup>142</sup> the enzyme structure was similar to other cobalamin biosynthetic methyltransferases as well as to the C-terminal region of CysG,<sup>144</sup> which is also an SUMT. Site-directed mutagenesis of *cobA*, in particular D47N and L49A variants, generated enzyme forms that made largely precorrin-1, the monomethylated (C2) precursor of precorrin-2.<sup>143</sup> There would not appear to be any conserved amino acid residue within SUMT that acts as a nucleophile in the methylation process. Catalysis would thus appear to be enhanced by approximation where the tetrapyrrole substrate is positioned in close proximity to the electrophilic methyl donor, SAM.

In terms of biotechnology and the enhancement of cobalamin production within industrial bacterial strains,<sup>145</sup> it has been shown that replacement of the *P. denitrificans* SUMT, which demonstrates substrate inhibition, with an equivalent enzyme from an archaean,<sup>141</sup> which does not display substrate inhibition, leads to a significant increase in cobalamin production.<sup>146</sup> Therefore, in many bacteria, SUMT plays an important role in metabolic control and flux along the corrin pathway.

The synthesis of precorrin-2 represents the point at which the aerobic and anaerobic pathways for cobalamin biosynthesis divide (Figure 17; Table 1).<sup>147</sup> They rejoin somewhere around the synthesis of (adenosyl)cobyric acid.<sup>147</sup> For the purposes of this review, we will describe separately how the two pathways arrive at this intermediate.



**Figure 19** Cobalamin biosynthetic methyltransferases. The cobalamin biosynthetic methyltransferases share a high degree of sequence similarity indicating that they have evolved from a common ancestor, and this relationship is shown in the dendrogram. The orthologous enzymes of the aerobic and anaerobic pathways have the highest degree of similarity. The structures of the enzymes that have been determined so far, together with their site of methylation. The two enzymes that do not belong to this family of canonical structures are CbiD and CbiT.

### 7.13.11 The Aerobic Pathway

The aerobic pathway differs from the anaerobic pathway in that the tetrapyrrole derived macrocycle is ring contracted and then has cobalt inserted. The aerobic nature of the pathway belies a requirement for molecular oxygen to help in the contraction process.

#### 7.13.11.1 Precorrin-3A Synthesis, C20 Methylation

The aerobic pathway is initiated by the addition of a methyl group to C20 of the tetrapyrrole template resulting in the synthesis of precorrin-3A (Figure 20). This reaction is catalyzed by an enzyme called SAM: precorrin-2 methyltransferase (SP2MT), an enzyme that is encoded by *cobI* within *P. denitrificans*.<sup>148</sup> The enzyme generates a trimethylated dipyrrocorphin called precorrin-3A as well as SAH.<sup>148,149</sup> It has an  $M_r$  of 26 k but in its native state it exists as a homodimer.<sup>148</sup> The primary sequence of the enzyme suggests that it belongs to the class III methyltransferases although no crystal structure of the enzyme has yet been elucidated (Figure 19).

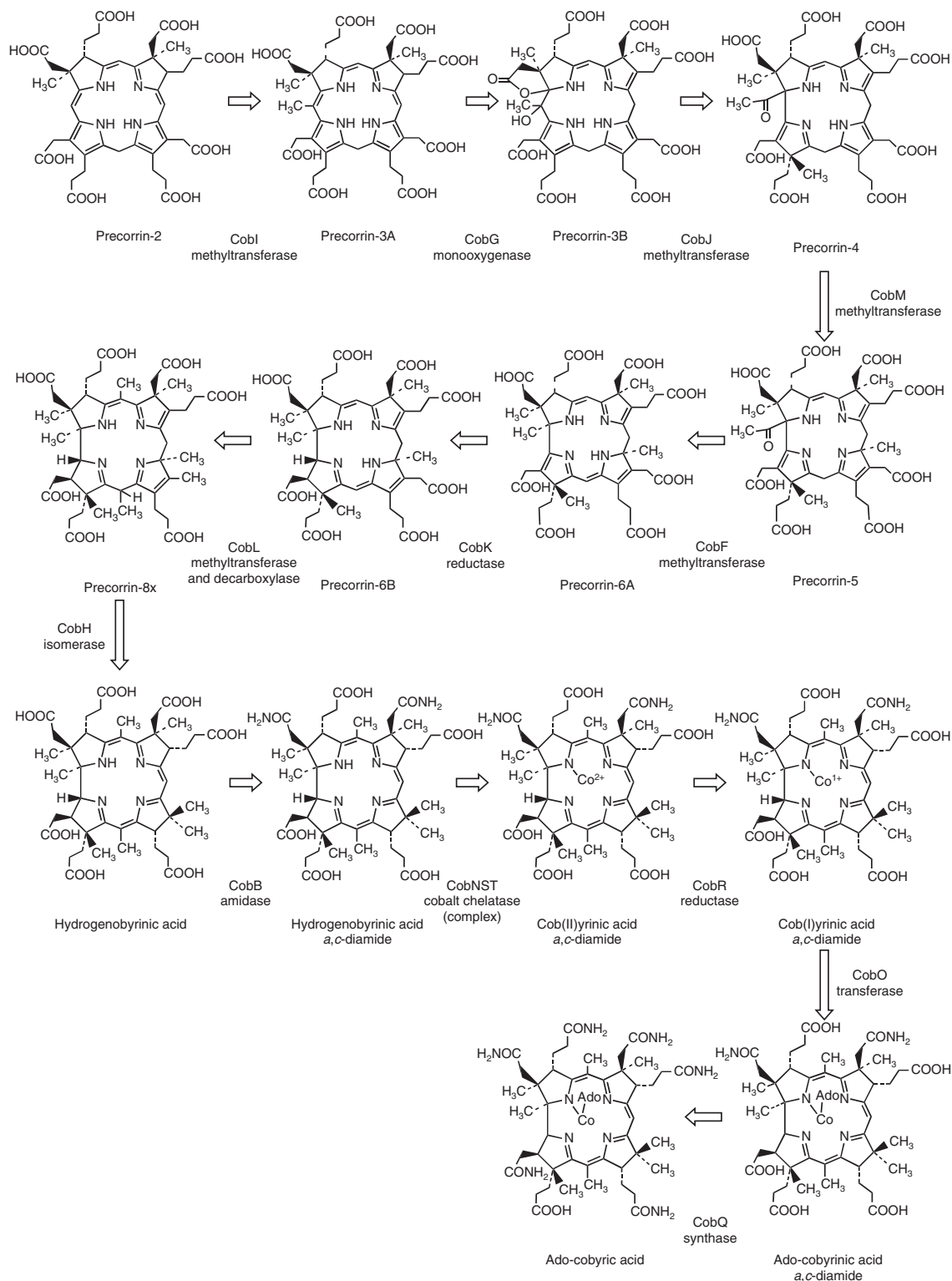
#### 7.13.11.2 Precorrin-3B Synthesis

It is one of the apparent oddities of cobalamin biosynthesis that the C20 position is methylated only for this position to be subsequently lost during the ring contraction process. The reason for this apparently futile process was explained by Eschenmoser with some classical chemical studies that were performed as he, with Woodward, pioneered the total chemical synthesis of vitamin B<sub>12</sub>. He demonstrated that there were many routes to ring contraction, which all revealed a dependency on the presence of a tertiary alcohol at C20.<sup>150</sup> In biological systems, this initial derivization of C20 is achieved by methylation. The generation of an alcohol on C20 is catalyzed by CobG, a monooxygenase that oversees not only the hydroxylation event but also the formation of a gamma lactone on ring A.<sup>151,152</sup> The product of this transformation is precorrin-3B (Figure 20), a compound that was first found to accumulate in incubations of precorrin-3A with cell-free protein preparations of *P. denitrificans* overproducing CobG.<sup>151,153</sup> The structure of precorrin-3B was confirmed by rigorous NMR analysis as well as by mass spectrometry and FT-IR.<sup>151,153</sup>

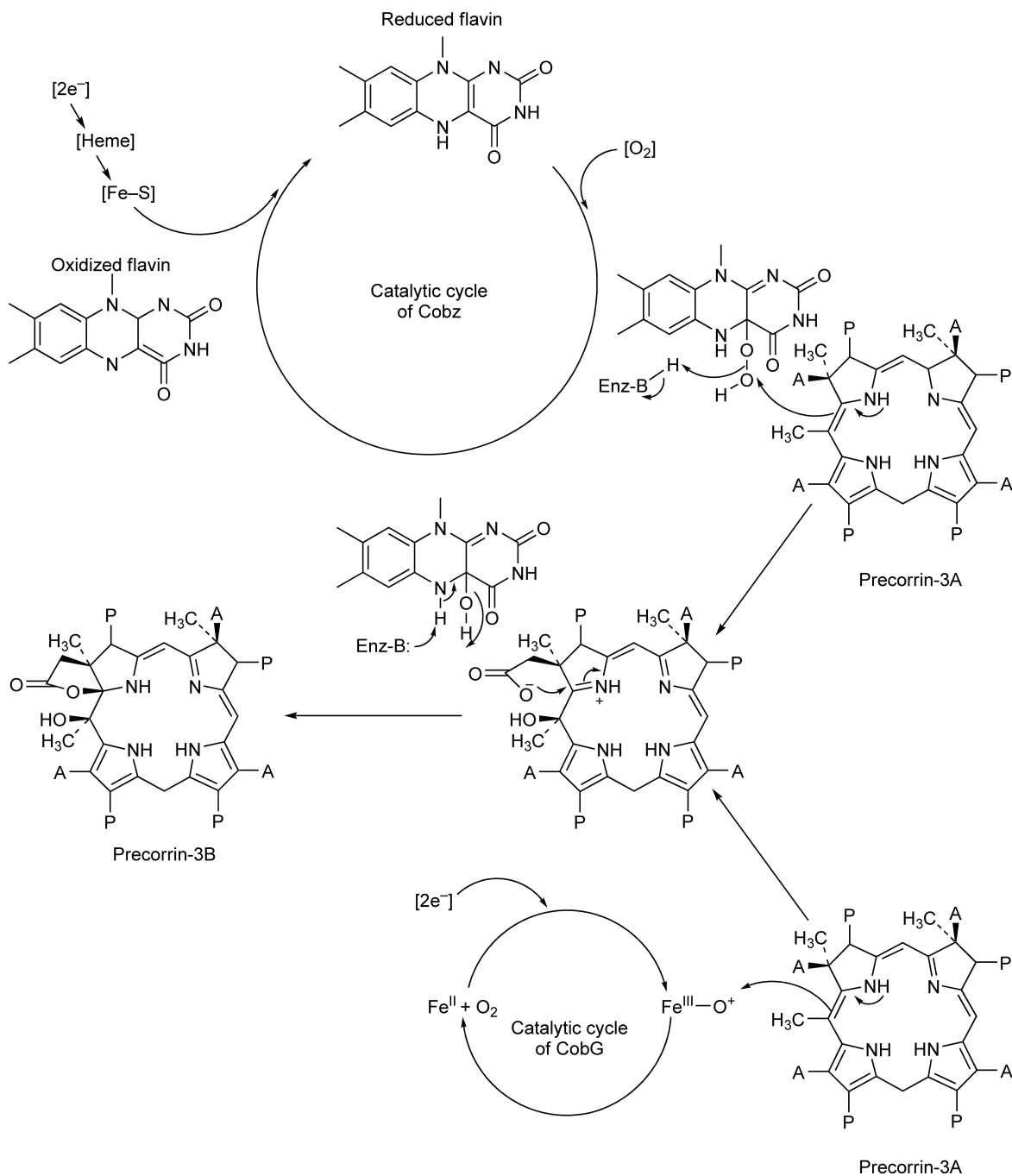
The mechanism of this reaction catalyzed by CobG is unclear. CobG is a 46 kDa protein that displays some sequence similarity to sulphite reductase, including the presence of four conserved cysteine residues that constitute an Fe<sub>4</sub>-S<sub>4</sub> center in the latter.<sup>151,152</sup> Indeed, purified CobG appears brown-green in solution and has an ultraviolet-visible spectrum consistent with the presence of such a redox group. Further analysis confirmed the presence of 4 mol of Fe and 4 mol of S per mol of purified protein.<sup>152</sup> The demonstration that CobG is a monooxygenase was confirmed by a number of elegant NMR experiments. Here, the course of the oxygen introduced at C20 was traced to molecular oxygen by incubating <sup>13</sup>C-enriched precorrin-3A with <sup>18</sup>O-labeled O<sub>2</sub> in the presence of CobG and CobJ.<sup>154</sup> This resulted in the synthesis of precorrin-4 (Figure 20), where clear evidence of an isotopic shift of <sup>18</sup>O on the <sup>13</sup>C chemical shift of C20 was observed.<sup>154</sup>

The absolute stereochemistry of precorrin-3B was deduced by NMR and revealed that the lactone is on the upper face, *cis* to the C20 hydroxyl, and the C20 methyl is on the lower face.<sup>155</sup> Mechanistically, it has been suggested that CobG may contain a nonheme iron, as well as an Fe<sub>4</sub>-S<sub>4</sub> center, which could form an Fe<sup>III</sup>-O<sup>+</sup> species, resulting in the insertion of a hydroxyl group followed by gamma lactone formation from the upper face of the imine (Figure 21).<sup>154</sup> Confirmation of such a mechanism awaits further characterization of CobG.

In some bacteria that operate the aerobic pathway, CobG is replaced by another enzyme called CobZ.<sup>156</sup> This enzyme is isofunctional with CobG but is very different in its composition. CobZ consists of two distinct regions. The N-terminal region contains a flavin (FAD) whereas the C-terminal region contains a heme group and is membrane embedded. The two regions are joined together by a linker section that contains two Fe<sub>4</sub>-S<sub>4</sub> centers. It has been suggested that the flavin acts as the cofactor for oxygenation by forming a peroxy intermediate that subsequently hydroxylates the C20 position (Figure 20).<sup>156</sup>



**Figure 20** The aerobic pathway for cobalamin biosynthesis. The steps leading to the transformation of precorrin-2 into adenosylcobyrinic acid are shown.



**Figure 21** Mechanism for the synthesis of precorrin-3B from precorrin-3A. In nature there are at least two distinct enzymes that can catalyze the synthesis of the hydroxyl gamma lactone derivative of precorrin-3A. In *Pseudomonas denitrificans* this reaction is catalyzed by CobG, which likely utilizes a bound nonheme iron to activate molecular oxygen. In *R. capsulatus*, this reaction is mediated by CobZ, which harnesses the oxygen-binding ability of a flavin.

### 7.13.11.3 Precorrin-4 Synthesis, Ring Contraction, and C17 Methylation

As an intermediate containing a tertiary alcohol at C20, precorrin-3B is now primed and ready for ring contraction. The enzyme that mediates the actual contraction is CobJ, a protein that, based on its sequence similarity, belongs to the class III methyltransferases (Figure 19). CobJ catalyzes the ring contraction and also



methylates the tetrapyrrole-derived template at C17 generating precorrin-4 (**Figure 20**), which was first isolated in its oxidized form as factor IV.<sup>157</sup> It is not known whether ring contraction occurs prior to methylation although CobJ does not work in the absence of SAM and it is likely that some concerted mechanism is in place.<sup>151,152</sup> The ring contraction sees the opening of the lactone ring and leaves a methyl ketone at C1 (**Figure 20**). The methylation at C17 is also notable for the fact that the methyl group is added from above the plane of the macrocycle, as opposed to methyl groups that are found at C1, C2, C7, and C12, which are found on the lower side of the macrocycle in the final corrin molecule. This tells us that the substrate for CobJ must fit into the active site of the enzyme in the opposite orientation to, for instance, the way uroporphyrinogen III is accommodated into the active site of SUMT. This likely reflects the inverse orientation of the acetate and propionate side chains in ring D, where C17 methylation occurs, in comparison to the other pyrrole rings of the template. The product of the CobJ reaction is precorrin-4, although it has only been isolated in its oxidized form (factor IV) as a methyl ester derivative.<sup>158</sup>

#### 7.13.11.4 Precorrin-5 Synthesis, C11 Methylation

Precorrin-4 acts as the substrate for another class III methyl transferase (**Figure 19**), CobM, which methylates at C11 to generate precorrin-5 (**Figure 20**).<sup>159</sup> It was a surprise when this intermediate was first identified as the corrin macrocycle does not contain a methyl group at C11 but does have one at C12. In fact, as will be described later, the methyl group added to C11 is subsequently rearranged to C12. The reason for methylation at C11 is to allow the decarboxylation of the acetate side chain attached to C12 prior to its methylation but it also plays a role in permitting the function of CobF (see Section 7.13.11.5).

#### 7.13.11.5 Precorrin-6 Synthesis, C1 Methylation, and Reduction

Precorrin-5 is a highly unstable yellow compound<sup>159</sup> that is acted upon by CobF, the fifth of the seven methyl transferases that are required in the construction of the corrin macrocycle (**Figure 19**).<sup>152</sup> As with CobJ, which contracts the macrocycle and methylates it, CobF is also a multifunctional enzyme in that it initially has to remove the methyl ketone function from C1 and then methylate the position (**Figure 20**).<sup>152</sup> The extruded attachment at C1 is lost as acetic acid and the elimination of this unit presumably generates a transient intermediate that is quickly methylated. The previous methylation at C11 allows for this concerted process in permitting the deacylation of precorrin-5 and the return of the electron density to C1. As with the apparent futile methylation at C20, the seemingly bizarre methylation at C11 is underpinned by sound chemical logic. There is little mechanistic information available on CobF although a crystal structure of the enzyme is now available. This confirms that CobF belongs to the same class III family as SUMT (**Figure 19**) but does not present any detail on how the enzyme is able to catalyze the deacylation process. The product of the CobF reaction is precorrin-6A (**Figure 20**). The oxidative synthesis of precorrin-3B saw a change in the oxidation state of the macrocycle, which is maintained through the synthesis of precorrin-6A.<sup>160</sup> The next step sees the intermediate return to the oxidation level of a hexahydroporphyrin by the NADPH-dependent reduction of precorrin-6A to precorrin-6B in a reaction catalyzed by CobK (**Figure 20**).<sup>160,161</sup> A mechanistic study revealed that the enzyme reduces the double bond between C18 and C19 of the macrocycle, stereospecifically transferring H<sup>R</sup> of the NADPH cofactor to C19.<sup>162</sup>

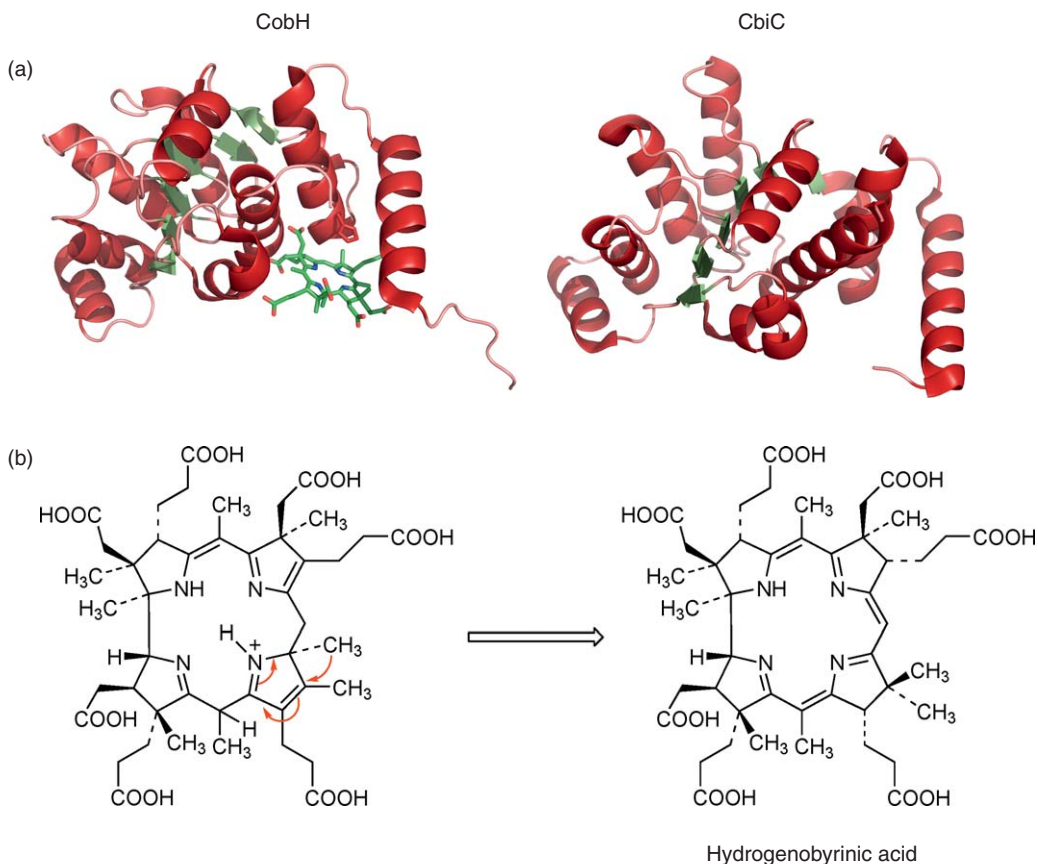
#### 7.13.11.6 Precorrin-8 Synthesis, Methylation at C5, C15, and Decarboxylation

The next steps are catalyzed by another multifunctional enzyme called cobL. This enzyme is responsible for the final two methylations at C5 and C15 as well as the decarboxylation of the acetate side chain attached to C12.<sup>163</sup> The enzyme oversees the transformation of precorrin-6B into precorrin-8 (**Figure 20**).<sup>164</sup> Sequence analysis of CobL, a 50 kDa protein, indicates that the N-terminal half of the protein displays similarity to the class III methyltransferases (**Figure 19**).<sup>132</sup> On this basis it was assumed that the N-terminal region of the protein performed the methylations in a similar way to how SUMT is able to mediate the *bis*-methylation at C2 and C7, whereas the C-terminal region was responsible for the decarboxylation.<sup>163</sup> However, in the anaerobic pathway the orthologous transformation is catalyzed by two separate enzymes that align with the N- and

C-termini of CobL.<sup>135</sup> Crystal structures of these enzymes of the anaerobic pathway revealed that both were methyltransferases, where the protein that aligns with the C-terminal region of CobL is a class II enzyme and not a member of the canonical B<sub>12</sub> biosynthetic methyltransferase.<sup>165</sup> The outcome of this is that the N-terminal region of CobL is likely to methylate precorrin-6B at C5 to give precorrin-7 and this is then methylated at C5 and decarboxylated to give precorrin-8.

### 7.13.11.7 Hydrogenobyric Acid Synthesis

CobH catalyzes the methyl group migration from C11 to C12 in a probable 1,5 sigmatropic rearrangement to generate hydrogenobyric acid (Figure 20).<sup>166</sup> This compound was first identified from cells of *Rhodospseudomonas spheroides* that had been grown in the absence of cobalt, indicating that this intermediate must be close to the point of cobalt insertion.<sup>167</sup> CobH is a comparatively small enzyme with a molecular mass of about 23 kDa. The protein binds its product tightly. The enzyme from *P. denitrificans* has a  $K_m$  of  $0.91 \mu\text{mol l}^{-1}$  for precorrin-8X and a  $V_{max}$  of  $230 \text{ nmol h}^{-1} \text{ mg}^{-1}$ .<sup>166</sup> The product of the reaction, hydrogenobyric acid, acts as a competitive inhibitor of the enzyme with a  $K_i$  of  $0.17 \mu\text{mol l}^{-1}$ . Indeed, CobH was first identified as a hydrogenobyric acid binding protein prior to its discovery as the enzyme responsible for the methyl group migration within precorrin-8.<sup>168</sup> The enzyme has been crystallized and the structure of the protein, bound with hydrogenobyric acid, has been solved (Figure 22).<sup>169</sup> The structure reveals that CobH exists as a homodimer, where each subunit contributes to a set of active sites. A conserved histidine residue resides close to the point of



**Figure 22** The structures and mechanism of CobH and CbiC. (a) The structures of CobH and CbiC are shown as indicated. The enzyme catalyzes the migration of a methyl group from C11 of the corrin ring to C12. The substrate-bound (hydrogenobyric acid – stick representation) structure of CobH indicates that an invariant histidine residue, located above ring C of the substrate, plays an important role in this process. (b) Proposed mechanism for the 1,5-sigmatropic rearrangement.

methyl group migration in the enzyme–product structure, suggesting that it may play a role as a general acid in the protonation of the ring C nitrogen (Figure 22).<sup>169</sup> CobH seemingly represents one of the few examples of an enzyme-catalyzed sigmatropic rearrangement.<sup>169</sup>

#### 7.13.11.8 Hydrogenobyric Acid *a,c*-Diamide Synthesis

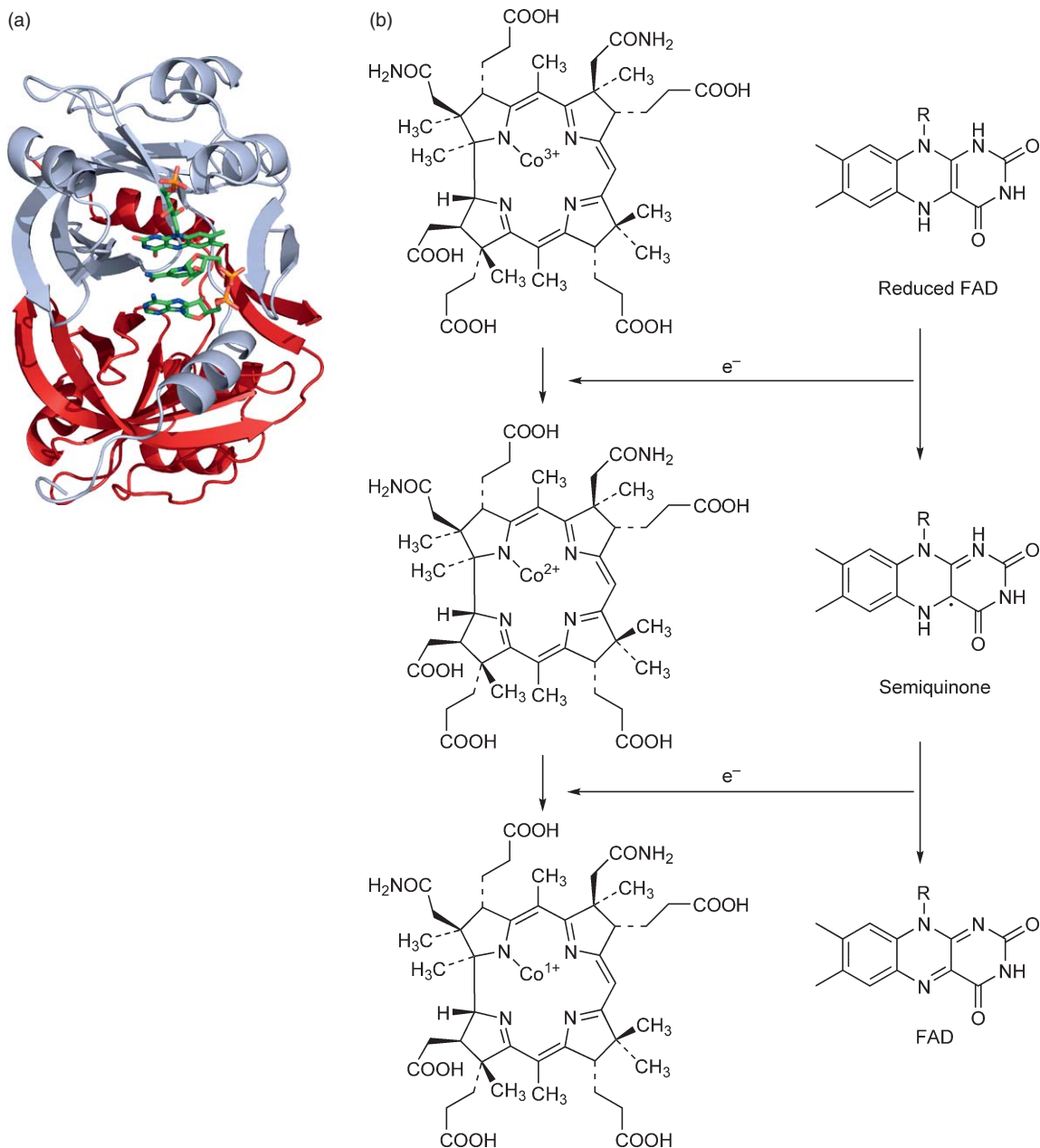
Hydrogenobyric acid is next amidated by an enzyme that amidates the acetic acid side chain on rings A and B (side chains *a* and *c*) (Figure 20).<sup>170</sup> The enzyme, cobyrinic *a,c*-diamide synthetase, is encoded by *cobB* and exists as a homodimer with a subunit molecular mass of 45 kDa.<sup>132,170</sup> The preferred amino donor was found to be glutamine with a  $K_m$  of 20.3  $\mu\text{mol l}^{-1}$ , while the  $K_m$  for hydrogenobyric acid was 0.41  $\mu\text{mol l}^{-1}$ . The reaction is ATP dependent. Sequence analysis of the cobyrinic acid *a,c*-diamide synthetase and the subsequent pathway amidase (cobyrinic acid synthetase, encoded by *CobQ* – see Section 7.13.11.12) suggest that they both belong to a novel class of ATP-dependent amidoligases, which also includes dethiobiotin synthetase of biotin synthesis.<sup>171</sup> This family of enzymes contain a number of conserved cysteine and histidine residues.

#### 7.13.11.9 Cobalt Insertion

Hydrogenobyric acid *a,c*-diamide acts as the substrate for cobalt insertion (Figure 20). Cobalt chelation is mediated by an enzyme system made up of three subunits, CobN, CobS, and CobT.<sup>172</sup> Of these, CobN is the largest with a subunit molecular mass of around 140 kDa. CobS and CobT are somewhat smaller, with masses of 37 and 70 kDa, respectively, although they form a complex together with a native mass in excess of 450 kDa.<sup>132,172</sup> The chelation requires the presence of the substrates,  $\text{Co}^{2+}$  and hydrogenobyric acid *a,c*-diamide, the three protein subunits, CobN and CobS-T, and also ATP. The  $K_m$  for hydrogenobyric acid *a,c*-diamide was found to be 0.085  $\mu\text{mol l}^{-1}$ , whereas the  $K_m$  values for  $\text{Co}^{2+}$  and ATP were 220 and 4.2  $\mu\text{mol l}^{-1}$ , respectively.<sup>172</sup> The product of the reaction is cob(II)yrinic acid *a,c*-diamide (Figure 20). Binding studies indicated that hydrogenobyric acid binds to the large CobN subunit, whereas sequence data suggests that CobS-T houses an ATP-binding site. Sequence comparisons revealed a similarity between the cobaltochelate subunits of the aerobic cobalamin biosynthetic pathway and the magnesium chelate system employed in chlorophyll/bacteriochlorophyll synthesis.<sup>173</sup> In the latter, magnesium is inserted into protoporphyrin IX by the subunits, ChlH, I and D, which are comparable to CobN, S, and T. Again ATP is required for the chelation process but why ATP is required remains unclear, although it is apparent that a large number of ATP molecules ( $\sim 15$ ) are hydrolyzed for metal insertion.<sup>174</sup> Overall, there is little information available on how the cobaltochelate system operates. These three-component chelatascs have been designated the class I chelatascs, to differentiate them from the single subunit chelatascs (type II) found on the anaerobic and heme pathways and the multifunctional enzymes (type III) found associated with siroheme synthesis.<sup>175</sup>

#### 7.13.11.10 Cobalt Reduction

The insertion of cobalt to produce the first corrin intermediate leads to the addition of an upper axial ligand for the metal ion (Figure 20). This process is initiated by a flavoprotein that was originally purified from a crude cell extract of *P. denitrificans*.<sup>176</sup> The enzyme was shown to catalyze an NADH-dependent reduction of Co(II) to Co(I) for a range of substrates including cobyrinic acid *a,c*-diamide and cobyrinic acid. The protein was shown to contain flavin mononucleotide (FMN) as a cofactor but although the enzyme had been purified and its N-terminus sequenced, the peptide sequence did not correspond to any of the known cobalamin biosynthetic genes. It was not until much later that the gene for the enzyme was found and isolated from *Brucella melitensis*.<sup>177</sup> This allowed the protein to be produced recombinantly. In this organism the cofactor for the enzyme was FAD and the large amounts of protein isolated allowed the enzyme to be crystallized. The resulting structure revealed that the protein exists as a homodimer, with each dimer contributing to two active sites (Figure 23). Kinetic and EPR analysis demonstrated that, as expected for a single electron reduction, the enzyme proceeds via a semiquinone form.<sup>177</sup>



**Figure 23** The structure and reaction mechanism of a corrin reductase, CobR. (a) The enzyme is made from two monomers (blue helices/purple sheets and green helices/orange sheets), although the flavin (stick representation) is only observed in one. The figure shows NADH modeled in a bent configuration into the active site below the flavin. (b) The reduction of the cobalt ion is mediated via single electron steps via reduced FAD.

#### 7.13.11.11 Adenylation

The cobalt(I) species that is generated is a very powerful nucleophile and a highly unstable entity. It is required as a substrate for the adenylation, which adds the upper adenosyl ligand to the centrally chelated cobalt ion. The enzyme responsible for this process is encoded by *cobO*.<sup>178</sup> The enzyme catalyzes the adenylation transfer from ATP to a range of corrins, including cob(I)yrinic *a,c*-diamide, cob(I)yrinic acid, and cob(I)inamide, generating the cobalt(III) corrinoid and triphosphate.<sup>179</sup> The *P. denitrificans* CobO is a homodimer with a

subunit molecular mass of 28 kDa. There are at least three types of adenosyltransferase in nature – the CobO type described above, PduO, and EutN.<sup>180–182</sup> Sequence and structural information indicate that these enzymes have arisen by convergent evolution. More detail on the mechanism of the CobO type enzyme is given in Section 7.13.12.10.

#### 7.13.11.12 Adenosylcobyrinic Acid Synthesis

Adenosylcobyrinic acid *a,c*-diamide is subject to amidation of its acetic acid side chains *b*, *d*, *e*, and *g* in a reaction catalyzed by the enzyme cobyrinic acid synthetase (**Figure 20**).<sup>183</sup> The enzyme is encoded by *cobQ*<sup>178</sup> and shares sequence similarly with the amidoligase that amidates the *a* and *c* side chains.<sup>170</sup> Cobyrinic acid synthetase has a subunit molecular mass of 57 kDa and exists in its native state as a homodimer. The enzyme-catalyzed reaction is ATP dependent, where glutamine is the preferred amide donor with a  $K_m$  of 45  $\mu\text{mol l}^{-1}$ . The  $K_m$  for the corrinoid substrate is around 1  $\mu\text{mol l}^{-1}$ . During the course of the reaction tri, tetra, and penta-amidated intermediates were detected, which were homogenous when detected by HPLC.<sup>183</sup> This suggests that the amidation reactions follow a specific sequence, although the sequence has not been determined (although it has now been elucidated for the orthologous enzyme of the anaerobic pathway – see Section 7.13.12.9). The biosynthesis of adenosylcobyrinic acid represents the point at which the two distinct cobalamin biosynthetic pathways converge.

### 7.13.12 The Anaerobic Pathway – From Precorrin-2 to Adenosylcobyrinic Acid

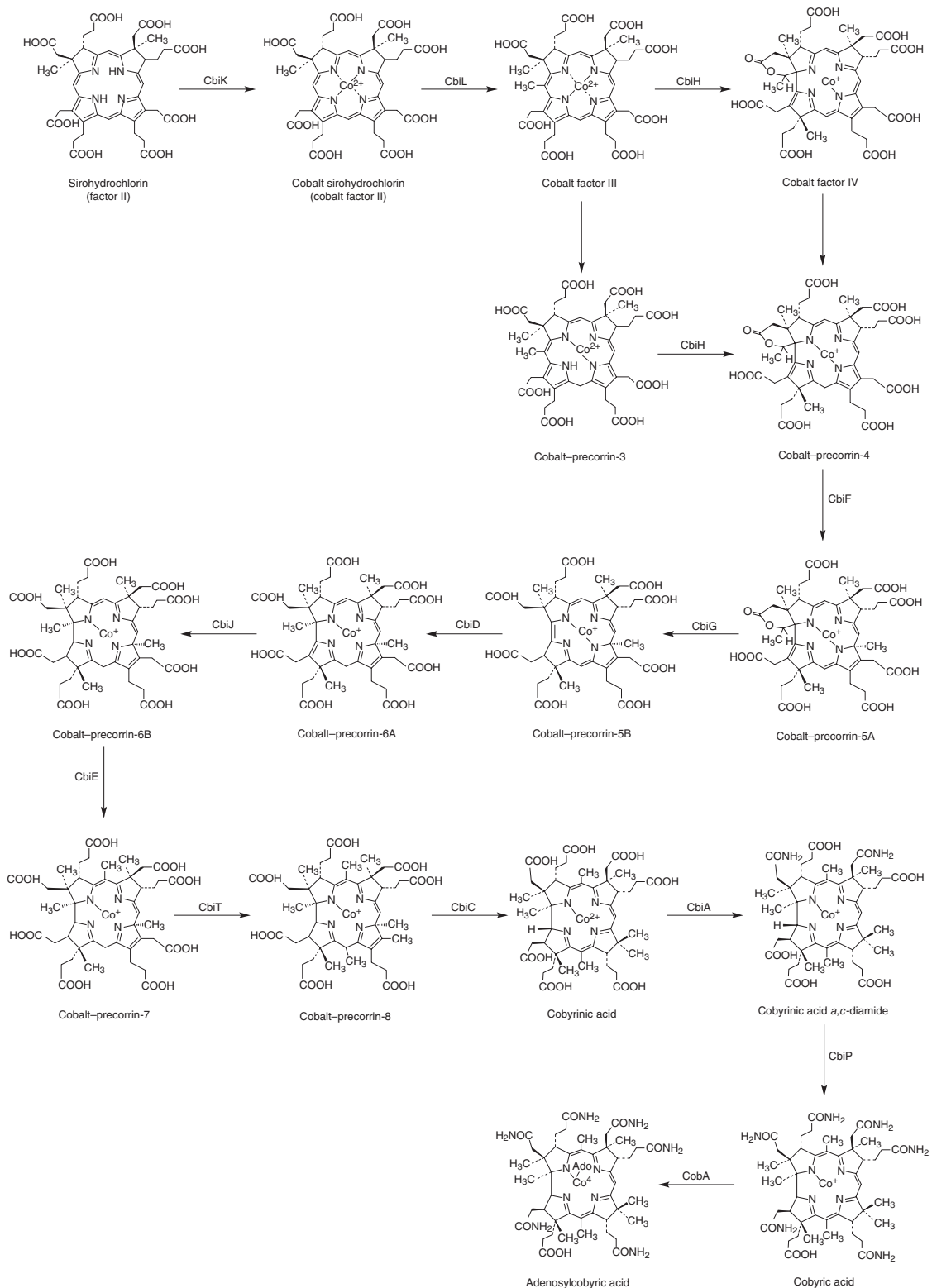
In contrast to the aerobic pathway, where overproducing strains and isolated enzymes have allowed the *ex vivo* synthesis of a complete range of cobalamin biosynthetic intermediates,<sup>152,184–186</sup> the anaerobic pathway remains poorly defined. This is largely due to the very low yield of highly unstable intermediates that are produced when the pathway is studied *in vitro*. However, recently, some progress has been made in characterizing some of the pathway intermediates and this is starting to provide some clarity on the processes involved.<sup>187–190</sup>

#### 7.13.12.1 Sirohydrochlorin Synthesis

From precorrin-2, the likely first step in the anaerobic pathway is the oxidation of this intermediate to generate the isobacteriochlorin factor II (also known as sirohydrochlorin) (**Figure 17**). This dehydrogenation can be catalyzed either by a precorrin-2 dehydrogenase, such as that encoded by *sirC*, or by the multifunctional enzyme CysG (which also has SUMT and FC activity).<sup>191,192</sup> The SirC dehydrogenase uses NAD<sup>+</sup> as a cofactor and has a structure similar to the N-terminal region of CysG and Met8p.<sup>193,194</sup> The latter two enzymes are primarily involved in siroheme synthesis and within this molecular framework both CysG and Met8p also have FC activity. However, the FC activity is absent in SirC.<sup>191</sup> Orthologues of SirC have recently been found in the archaea too.

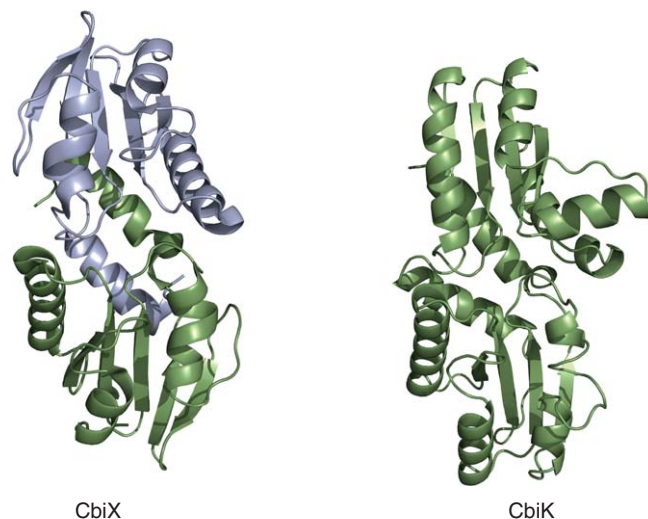
#### 7.13.12.2 Cobalt Insertion

Factor II acts as the substrate for cobalt chelation along the anaerobic route (**Figure 24**). This early insertion of cobalt is the hallmark of the anaerobic pathway and is catalyzed by a number of cobaltochelataes, that are referred to as CbiK, CbiX<sup>L</sup>, or CbiX<sup>S</sup>.<sup>175,195,196</sup> These enzymes are class II chelataes in that they are either monomeric or homodimeric enzymes that insert metal ions without a requirement for ATP. It was originally proposed that these enzymes inserted the metal ion into precorrin-2, but it appears now that factor II is the preferred substrate.<sup>195</sup> The product of the reaction is cobalt(II)–factor II. CbiK has been crystallized and its structure determined (**Figure 25**).<sup>197</sup> The topology of CbiK was unexpectedly found to be similar to that of protoporphyrin IX FC,<sup>122</sup> indicating that the enzymes had evolved from a common ancestor. The level of sequence identity between these chelataes is about 10%. The structure of CbiK revealed that the active site is formed between two domains, which are very similar. Within the active site are two histidine residues that are thought to play a role in deprotonation of the tetrapyrrole substrate and metal binding and insertion



**Figure 24** The anaerobic pathway for cobalamin biosynthesis. The steps leading to the transformation of sirohydrochlorin into adenosylcobyrinic acid are shown.





**Figure 25** The structures of the anaerobic cobaltochelataes CbiK and CbiX. CbiK is a monomeric enzyme derived from a single polypeptide chain (helices in blue, purple sheets). CbiX is a dimeric enzyme, where each monomer (green helices/orange sheets) corresponds to one of the domains observed in CbiK. It has been suggested therefore that CbiK arose from a gene duplication and fusion event of *cbiX*.

(Figure 25).<sup>197</sup> CbiK also displays some similarity to CbiX<sup>L</sup>, which is a cobaltochelatae first identified in *Bacillus megaterium*. Here, the C-terminal region of the protein contains an extension that has a natural polyhistidine region and also houses an Fe<sub>4</sub>-S<sub>4</sub> center.<sup>195</sup> The role of the redox cluster is unknown and the histidine-rich region is thought to act as a region to bind excess cobalt and deliver it to the active site. Alternatively the region may act in some regulatory fashion and perhaps even act as a transcriptional regulator. The CbiX<sup>S</sup> proteins were first identified in members of the archaea. The proteins are typically very small and contain around 120 amino acids.<sup>175</sup> They align with both the N-terminal and C-terminal regions of CbiX<sup>L</sup> (and to a lesser extent CbiK), suggesting that these larger proteins have arisen from a process of gene duplication and fusion. A crystal structure of CbiX<sup>S</sup> from *Archaeoglobus fulgidus* has shown that CbiX<sup>S</sup> does indeed contain the same topology as CbiK (Figure 25).<sup>198</sup>

### 7.13.12.3 Methylation at C20

Cobalt-factor II acts as the substrate for the C20 methyltransferase, CbiL (Figure 24).<sup>199</sup> The product of the reaction is cobalt(II)-factor III.<sup>200</sup> Cobalt-precorrin-2 also acts as a substrate but the transformation is much more efficient with the oxidized substrate. The enzyme has a high specificity for the presence of the cobalt ion, as cobalt-free analogues are not accepted as substrates.<sup>199</sup> The enzymes from *Chlorobium tepidum* and *Methanothermobacter thermoautotrophicus* have been crystallized and their structures determined, revealing as expected that they belonged to class III of cobalamin biosynthetic methyltransferases (Figure 19).<sup>200,201</sup> EPR analysis of CbiL from *M. thermoautotrophicus* has shown that the protein donates a fifth ligand to the cobalt ion of the substrate, explaining the specificity of the enzyme for a metal-containing substrate.<sup>200</sup>

The anaerobic biosynthesis of cobalt-factor III has so far proceeded with a net oxidation of the starting hexahydroporphyrin substrate. There is some debate still as to whether this is the correct oxidation state.<sup>188,202,203</sup> It may be that cobalt-factor III is reduced back to cobalt-precorrin-3 for an oxidative ring contraction process, or cobalt-factor III undergoes a nonoxidative contraction. In either case, the oxidative state of the next ring-contracted intermediate should be equivalent to that of precorrin-4 of the aerobic pathway (Figure 24). This is one of the gray areas of the anaerobic cobalamin biosynthetic pathways.

#### 7.13.12.4 Ring Contraction and C17 Methylation

The ring contraction step along the anaerobic pathway is mediated by CbiH, a methyltransferase that is homologous to CobJ of the aerobic route (Figure 19). Experiments have shown that feeding either cobalt-factor III or cobalt-precorrin-3 into cell extracts containing overproduced CbiH results in the accumulation of a tetramethylated intermediate, which after esterification and extraction, has been identified as cobalt-factor IV.<sup>204,205</sup> The yields of this compound are comparatively low and the compound is oxidized with respect to precorrin-4 and thus the true oxidation state of the pathway intermediate is thought to be cobalt-precorrin-4 (Figure 24). CbiH catalyzes the methylation of C17 and also the formation of a delta lactone ring with the acetic acid side chain on ring A and the C1 position.

#### 7.13.12.5 Methylation at C11 and Deacylation

Cobalt-precorrin-5A is formed next by the methylation of the corrin framework at C11 (Figure 24),<sup>189</sup> in a reaction that is analogous to the synthesis of precorrin-5 by CobM (Figure 19). Along the anaerobic pathway, the synthesis of the tetramethylated intermediate is catalyzed by CbiF, which was the first of the cobalamin biosynthetic enzymes to have its structure determined and demonstrated that the cobalamin biosynthetic methyltransferases represent a unique class (III) of methyltransferase (Figure 19).<sup>206</sup> Apart from the presence of a cobalt ion, the major difference between the aerobic and anaerobic pathways at this stage is that cobalt-precorrin-5A still carries the delta lactone ring that is formed as a result of the ring contraction. Recently, it has been shown that this lactone is opened and the C2 fragment is extruded by the action of CbiG.<sup>189</sup> This conclusion was reached when it was found that cell lysates containing CbiF and CbiG, incubated with cobalt-precorrin-4, resulted in the formation of cobalt-precorrin-5B, a deacylated intermediate, whereas incubation in the absence of CbiG generated only cobalt-precorrin-5A (Figure 24). The C2 fragment, representing the original methylated C20 position, is lost as acetaldehyde.<sup>207</sup> Thus, CbiG is involved in both lactone ring opening and deacylation. Little is known about the mechanism of CbiG, although the C-terminus of the protein does display some similarity to CobE, a protein of unknown function from the aerobic pathway. It has been proposed, based on the function of CbiG, that CobE may play a role in the deacylation of precorrin-5, which is known to be an unstable intermediate.<sup>188</sup>

The mechanism, therefore, for ring contraction and extrusion of the two-carbon fragment, representing the methylated C20 position, are quite distinct between the aerobic and anaerobic pathways. Earlier research on *Propionibacterium sbermanii* had revealed that during cobalamin biosynthesis in this organism, which operates an anaerobic pathway, substantial exchange takes place with the carbonyl oxygen on the ring A acetate side chain and the medium.<sup>208,209</sup> This would be consistent with the opening of the delta lactone on cobalt-precorrin-5A as catalyzed by CbiG.<sup>189</sup> In contrast, along the aerobic pathway, there is no oxygen exchange between this carbonyl group and the medium up to either precorrin-5<sup>210</sup> or hydrogenobyric acid.<sup>155,211</sup> Thus, along the aerobic pathway, the extruded two-carbon fragment is lost as acetic acid by hydrolysis of precorrin-5 involving external water during the reaction catalyzed by CobF. However, along the anaerobic pathway, the structure of cobalt-factor IV and cobalt-factor VA reveals the presence of a proton on the C20 carbon within the delta lactone, indicating that the extruded two-carbon fragment must be lost as acetaldehyde.<sup>189,205</sup> This was subsequently confirmed by trapping the released fragment.<sup>207</sup>

#### 7.13.12.6 Methylation at C1

Along the aerobic pathway, methylation at C1 is catalyzed by CobF,<sup>152,159,177</sup> another of the canonical B<sub>12</sub> biosynthetic methyltransferases (Figure 19). However, no homologue to this enzyme is found in the anaerobic pathway and thus it had been proposed that a protein called CbiD may take on this role instead (Figure 24).<sup>202</sup> No direct evidence that CbiD is the C1 methyltransferase has been obtained, but researchers have found that a recombinant *E. coli* strain containing 12 *S. enterica* genes required for cobalamin synthesis was able to make cobyrinic acid *a,c*-diamide, but when CbiD was missing a nonphysiological 1-*des*-methyl-cobyrinic acid *a,c*-diamide was formed in its place.<sup>190</sup> This provides strong circumstantial evidence that CbiD is the C1 methyltransferase and is likely to catalyze the synthesis of cobalt-precorrin-6A (Figure 24), a compound that

has yet to be isolated. Intriguingly, the synthesis of cobyrinic acid *a,c*-diamide also required the presence of the amidating enzymes CbiA and CbiP, which suggests that perhaps some macromolecular enzyme complex is required for the biosynthesis of cobalamin and the absence of the components may prevent the complex from functioning normally.<sup>190</sup>

The presence of an orthologue to CobK, the precorrin-6X reductase, in the form of CbiJ in the anaerobic pathway suggests that cobalt–precorrin-6A is reduced to cobalt–precorrin-6B (Figure 24).<sup>188,202</sup> However, until the elusive cobalt–precorrin-6A is isolated, this part of the pathway remains hypothetical.

#### 7.13.12.7 Methylations at C5, C15, and Decarboxylation

By analogy with the aerobic pathway, cobalt–precorrin-6B would be methylated next at positions C5 and C15 and have its ring C acetic acid side chain decarboxylated. Along the aerobic pathway this is catalyzed by a multifunctional enzyme called CobL.<sup>163</sup> On the anaerobic pathway orthologues of CobL are found as two discrete enzymes termed CbiE and CbiT.<sup>135</sup> Sequence analysis of CbiE suggested that it was a canonical cobalamin biosynthetic enzyme, which was confirmed when a structure of the enzyme was determined by X-ray crystallography. It was initially suggested that CbiE catalyzed both the methylations at C5 and C15 in a similar fashion to how SUMT is able to perform its methylations at C2 and C7. However, when the structure of CbiT was deduced and also shown to have a methyltransferase topology, albeit a slightly different fold to the normal canonical cobalamin biosynthetic enzyme, it was proposed that CbiE may act as the C5 methyltransferase to generate cobalt–precorrin-7 and then CbiT would catalyze the decarboxylation and methylation at C15 to produce cobalt–precorrin-8 (Figure 24).<sup>165</sup> Evidence that CbiT does indeed act as both the C15 methyltransferase and a decarboxylase came from experiments where incubations were set up with a number of biosynthetic enzymes and cobalt–precorrin-3.<sup>187</sup> Incubation with CbiH, F, G, and T resulted in the synthesis of a number of nonphysiological but novel cobalt corrinoids. These novel compounds were found to be either methylated at C15 or methylated at C15 and decarboxylated.<sup>187</sup> No methylation was found without decarboxylation suggesting that CbiT was first a methylase and then a decarboxylase.

#### 7.13.12.8 Cobyric Acid Synthesis

Cobalt–precorrin-8 is likely to be converted into cobyrinic acid by the action of CbiC (Figure 24). The enzyme has strong sequence similarity to its aerobic homologue CobH. A structure for CbiC has also been deduced (Figure 22) but it does not provide any further mechanistic insight into this proposed 1,5-sigmatropic rearrangement.<sup>212</sup>

#### 7.13.12.9 Amidation of the Macrocycle

Cobyric acid acts as the substrate for the first of two amidases, CbiA, which amidates the *a* and *c* acetic acid side chains (Figure 24). As with CobB, CbiA uses glutamine as its preferred amino donor and requires ATP as an energy source.<sup>213</sup> Kinetic analysis has shown that the hydrolysis of glutamine and the synthesis of cobyrinic acid *a,c*-diamide are uncoupled. It was also shown that the amino groups are transferred in the order of *c* first and then *a*, allowing a model to be proposed whereby phosphinylated intermediates are formed.<sup>213</sup>

Similar mechanistic studies have also been carried out on the second amidase enzyme, CbiP, which transfers four amido groups to side chains *b*, *d*, *e*, and *g* in the synthesis of cobyrinic acid (Figure 24).<sup>214</sup> As with CbiA, the reaction catalyzed by CbiP was found to be dissociative whereby the partially amidated intermediates were released into solution. The order of amidation was found to be *e* followed sequentially by *d*, *b*, and *g*. Isotope exchange experiments support the theory that CbiP catalyzes the reaction by using ATP to activate the carboxyl groups by formation of a phosphinylated intermediate.<sup>214</sup> Interestingly, a point mutation of D146N of the *S. enterica* CbiP led to partial randomization of the amidation order.<sup>215</sup>

### 7.13.12.10 The Anaerobic Pathway and Adenosylation

Along the aerobic pathway, analysis of cobalamin-producing mutant strains coupled with the study of the kinetics of purified enzymes led to the deduction that cobyrinic acid *a,c*-diamide is the intermediate that becomes adenosylated.<sup>185</sup> However, it is not known which intermediate is adenosylated along the anaerobic pathway, although the enzyme responsible for the adenosylation (CobA) has been well characterized.<sup>216</sup>

The first step in the adenosylation process is the reduction of the cobalt(II) corrinoid to a Co(I) species. No specific enzyme has been found in *S. enterica* and the reduction is thought to be mediated by reduced flavodoxin A (Figure 24).<sup>217,218</sup> The Co(I) species is then acted upon by the adenosyltransferase (CobA). In *S. enterica*, CobA has been shown to be responsible for the *de novo* synthesis of the corrin ring as well as in the salvage pathway.<sup>182,219</sup> The enzyme thus has a broad substrate specificity. By analogy with the aerobic pathway it has been suggested that cobyrinic acid *a,c*-diamide is the physiological substrate for the point of adenosylation along the anaerobic pathway, but there is no experimental data to support this idea.<sup>188</sup> Mechanistic studies have shown that the enzymes require the 2'-OH group of ATP for function and yields inorganic triphosphate as its reaction by-product.<sup>220</sup>

The structures of CobA and CobA with a number of substrates have been determined (Figure 26).<sup>221</sup> They show that the 196 amino acid protein exists as a homodimer, with an alpha/beta structure built from a six-stranded parallel  $\beta$ -sheet. Significantly, ATP is bound to the P-loop in the opposite orientation to all other nucleotide hydrolases, to accommodate the need to transfer an adenosyl group rather than the gamma phosphate. The corrin ring is bound in such a way that the cobalt ion is too far away from the C5' of the ribose to allow direct nucleophilic attack, suggesting that a conformational change is required upon reduction of the cobalt ion.<sup>220</sup> Spectroscopic and computational studies have suggested how the highly energetically unfavorable Co(II) to Co(I) transition may be stabilized.<sup>222</sup> It has been suggested that the corrin substrate forms a four coordinate species upon binding to CobA in the presence of ATP. This activated square planar arrangement would help increase the reduction potential of the corrinoid and allow the transfer of one electron from the flavodoxin A. This arrangement would have to bring the Co(I) supernucleophile in closer proximity to the C5' of the ribose and thereby permit adenosylation with release of triphosphate.<sup>222</sup>

### 7.13.13 Biosynthesis of Cobinamide (Phosphate)

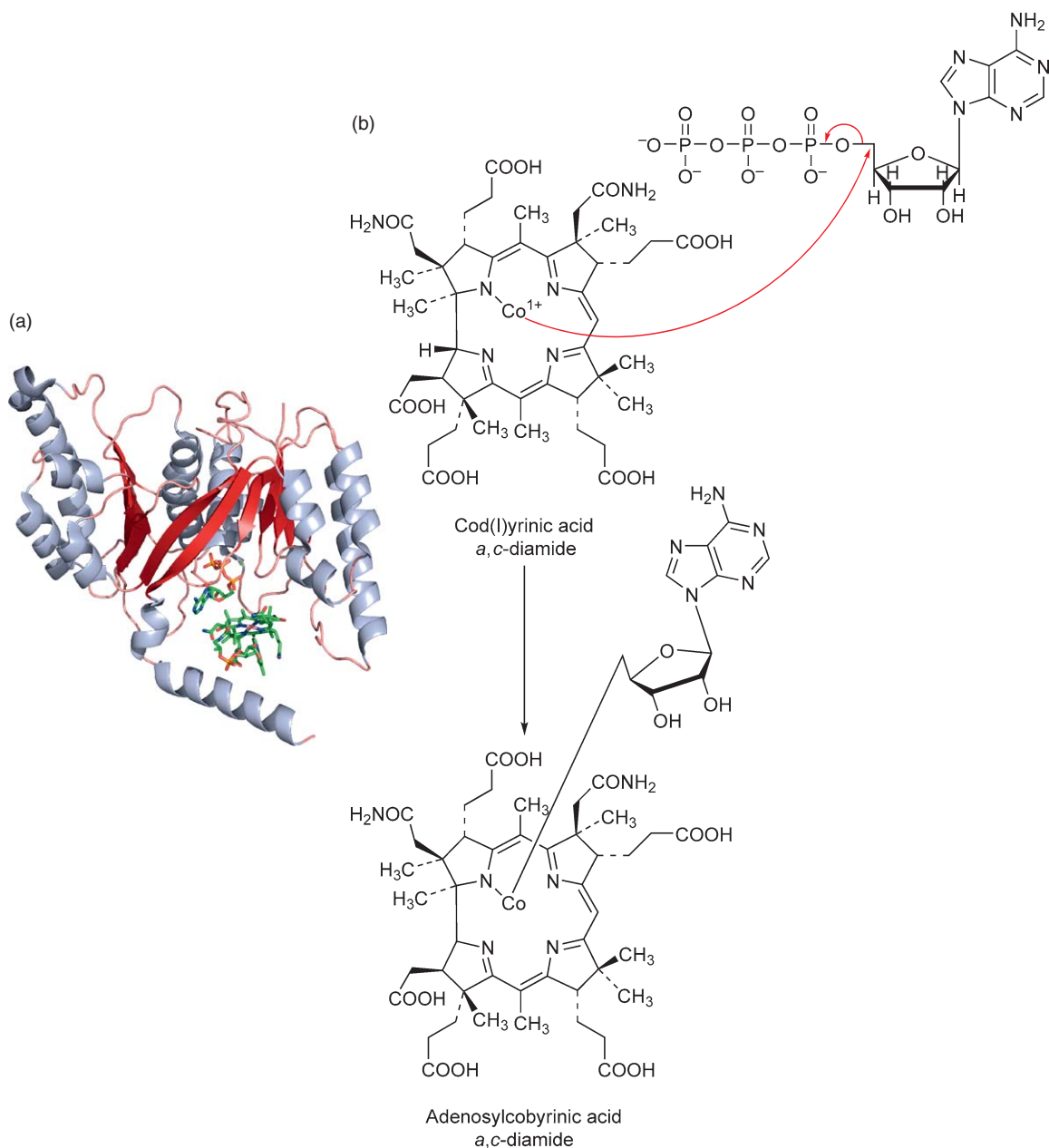
The biosynthesis of cobinamide phosphate requires the attachment of an aminopropanol linker arm to the one remaining carboxylic acid side chain of the original tetrapyrrole template (Figure 27).

#### 7.13.13.1 Synthesis of (*R*)-1-Amino-2-Propanol

The (*R*)-1-amino-2-propanol linker is known to be derived from threonine.<sup>223</sup> In *S. enterica*, CobD was found to be an enzyme with L-threonine *O*-3-phosphate decarboxylase activity, which generates (*R*)-1-amino-2-propanol phosphate.<sup>224</sup> The enzyme is a pyridoxal phosphate requiring enzyme and the structure of the protein has been determined by X-ray crystallography (Figure 28).<sup>225</sup> The structure of CobD was found to be highly similar to the aspartate aminotransferase family of enzymes. Structures of CobD with substrate and product bound have allowed a detailed mechanism for the enzyme to be proposed, whereby the external aldimine is directed toward decarboxylation rather than aminotransfer.<sup>226</sup> Threonine phosphate, itself, is synthesized from L-threonine by the action of a kinase, which is encoded by *pduX*.<sup>227</sup> The *pduX* is housed within the propanediol utilization operon<sup>228</sup> rather than the cobalamin biosynthetic operon for reasons that are not clear.

#### 7.13.13.2 Attachment of (*R*)-1-Amino-2-Propanol *O*-2-Phosphate

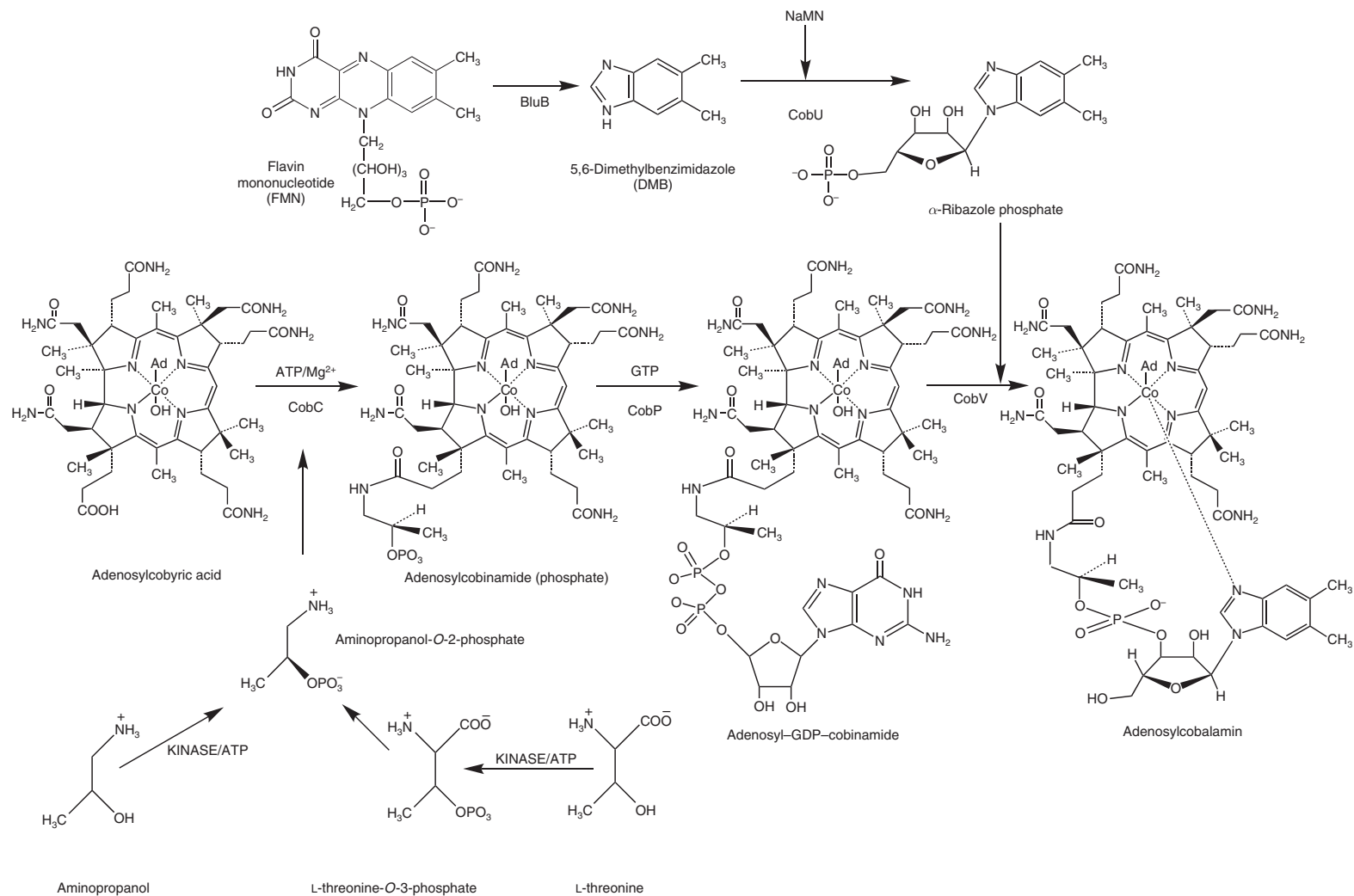
The (*R*)-1-amino-2-propanol *O*-2-phosphate is incorporated into the corrin ring by the action of either the *S. enterica* CbiB<sup>229</sup> or the *P. denitrificans* CobD.<sup>185</sup> In the case of the latter, cobinamide synthesis was found to be catalyzed by two proteinaceous components that were termed  $\alpha$  and  $\beta$ .<sup>185</sup> Protein  $\alpha$  was purified to homogeneity and shown to consist of a 38 kDa polypeptide whose N-terminal sequence did not match any of the



**Figure 26** The structure and mechanism of the adenosyltransferase, CobA. (a) The structure of CobA, which reveals how the corrinoid substrate is bound in proximity to ATP. (b) The mechanism involves the use of the  $\text{Co}^+$  supernucleophile, which attacks the C5 carbon of the ribose of ATP, releasing triphosphate.

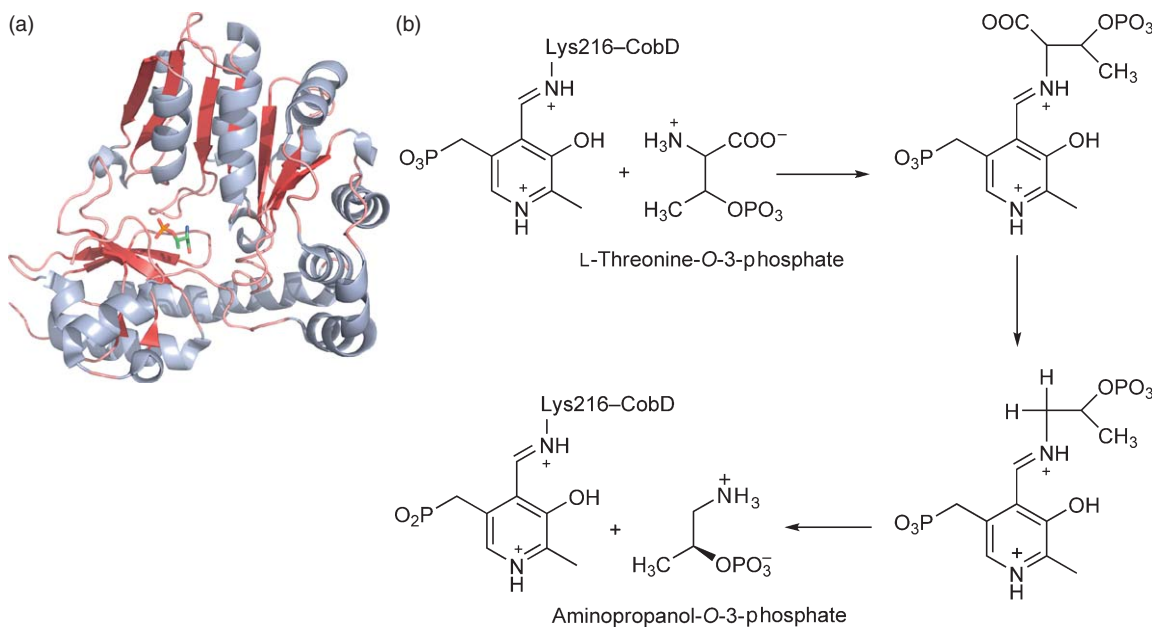
known encoded cobalamin biosynthetic enzymes. The  $\beta$  component is thought to be composed of the *P. denitrificans* CobD (cobinamide synthesis) and CobC (threonine phosphate decarboxylase). Together  $\alpha$  and  $\beta$  were shown to catalyze the attachment of (*R*)-1-aminopropanol to adenosylcobyrinic acid in the presence of ATP<sup>185</sup> (even if aminopropanol phosphate is likely to be the true substrate).

Although the *P. denitrificans* CobD and *S. enterica* CbiB share a high degree of sequence similarity, in contrast to CobD, CbiB is found as an integral membrane protein and is largely insoluble.<sup>229</sup> It seems likely that CbiB activates adenosylcobyrinic acid by phosphorylation to favor the attachment of the aminopropanol phosphate



**Figure 27** Transformation of adenosylcobyrinic acid into adenosylcobalamin. The pathway shown in this figure is highlighted with the aerobic pathway enzymes. The corresponding enzymes of the anaerobic pathway are shown in [Table 1](#). It is likely that the  $\alpha$ -ribazole is synthesized as  $\alpha$ -ribazole phosphate and that the phosphate is removed from the final coenzyme form of cobalamin.





**Figure 28** The structure and mechanism of threonine phosphate decarboxylase, CobD. (a) This enzyme of the anaerobic pathway identifies the active site by the presence of its pyridoxal phosphate cofactor (stick representation), which is attached to an active site lysine. (b) The mechanism of the enzyme involves the binding of the threonine phosphate to the pyridoxal phosphate cofactor via a Schiff base. Following decarboxylation, the product is released.

linker and to generate adenosylcobinamide phosphate (Figure 27). It will also incorporate ethanolamine phosphate in place of aminopropanol phosphate to generate norcobinamide phosphate.<sup>229</sup>

## 7.13.14 Nucleotide Loop Assembly

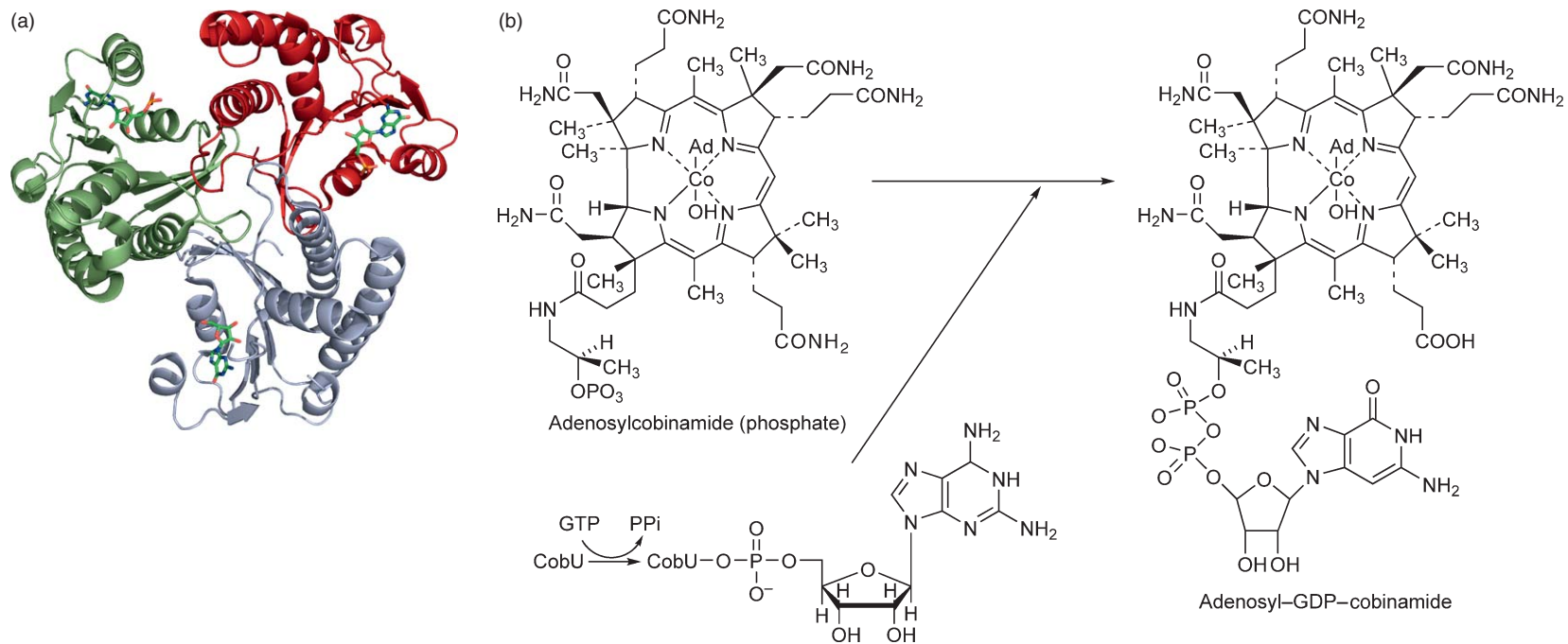
### 7.13.14.1 Synthesis of Adenosyl-GDP-Cobinamide

The lower nucleotide assembly process is initiated by the transfer of a guanosyl monophosphate moiety (GMP) to the adenosylcobinamide phosphate intermediate (Figure 27). The enzyme responsible for this reaction was first described in detail in *P. denitrificans* as CobP, a comparatively small homodimeric enzyme with a subunit molecular mass of 20 kDa.<sup>230</sup> In fact CobP was found to have two activities: firstly it is able to phosphorylate adenosylcobinamide and convert it to adenosylcobinamide phosphate. This is probably part of a salvage pathway to reconstruct broken-down cobinamides. The second function of CobP is to transfer the GMP moiety of GTP to adenosylcobinamide phosphate.<sup>230</sup> In *S. enterica* the equivalent enzyme is termed CobU, and like CobP this enzyme has both adenosylcobinamide kinase as well as nucleotidyltransferase activity.<sup>231</sup> Under anaerobic conditions, CobU can recognize cobinamide rather than adenosylcobinamide as a substrate.<sup>232</sup>

Much more detail is known about CobU than CobP as the crystal structure has been elucidated in a range of nucleotide substrate-bound forms (Figure 29).<sup>233–235</sup> The transferase activity progresses via a covalently linked enzyme-guanylyl intermediate, which is attached to His46. Upon binding the GMP moiety to the enzyme, the protein undergoes a large conformational change to bring the histidine residue closer to the P-loop motif to allow efficient transfer to the amino phosphate moiety of the corrin substrate.<sup>233–235</sup>

### 7.13.14.2 Synthesis of 5,6-Dimethylbenzimidazole

The nucleotide loop assembly requires the synthesis and incorporation of the lower base, which is 5,6-dimethylbenzimidazole (DMB). It had been known for some time that this unusual base could be made by one of two routes – an aerobic route where the base was derived from flavin<sup>236</sup> or an anaerobic pathway

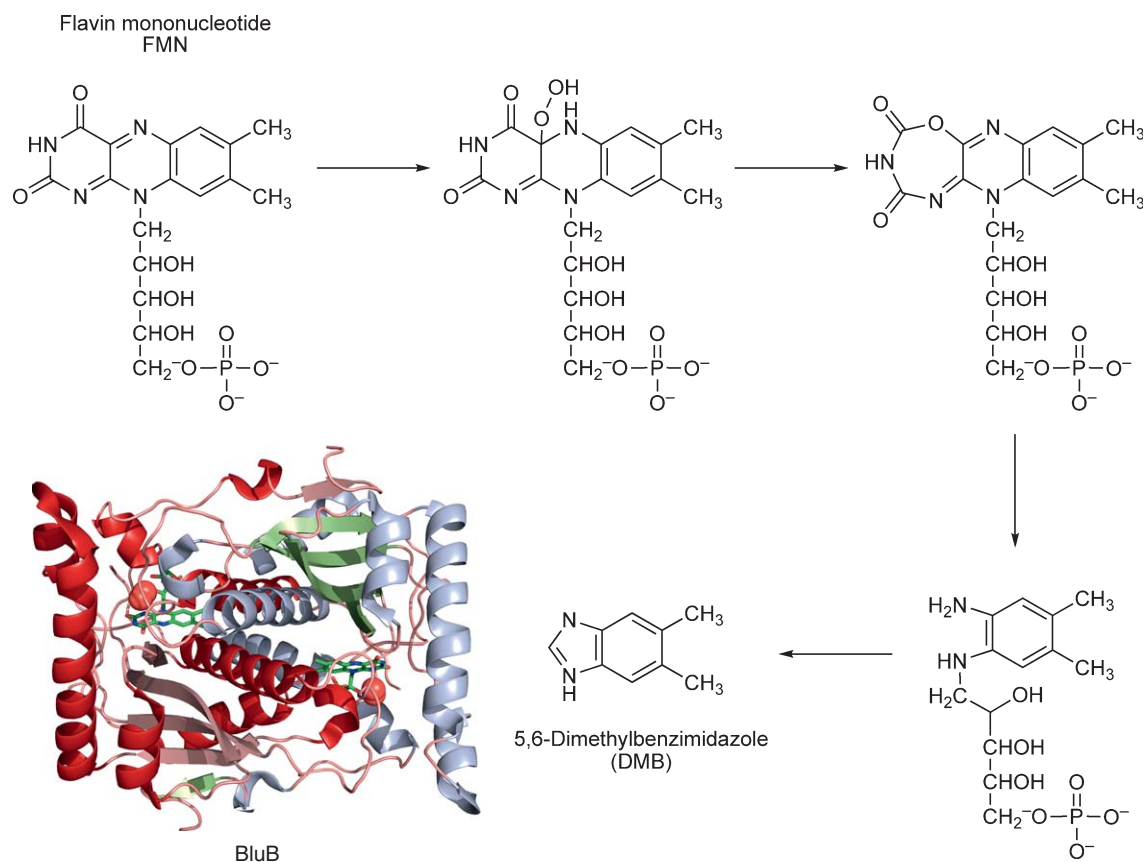


**Figure 29** The structure and mechanism of GMP addition to give adenosyl-GDP-cobinamide. (a) The structure of the anaerobic pathway enzyme, CobU, is shown. The enzyme exists as a trimer, and in the figure the subunits are differentially colored (blue, green, purple). In this representation GTP (stick representation) is bound in the active site. (b) The mechanism involves the covalent addition of GMP to CobU prior to the transfer of the nucleotide to adenosylcobinamide phosphate.

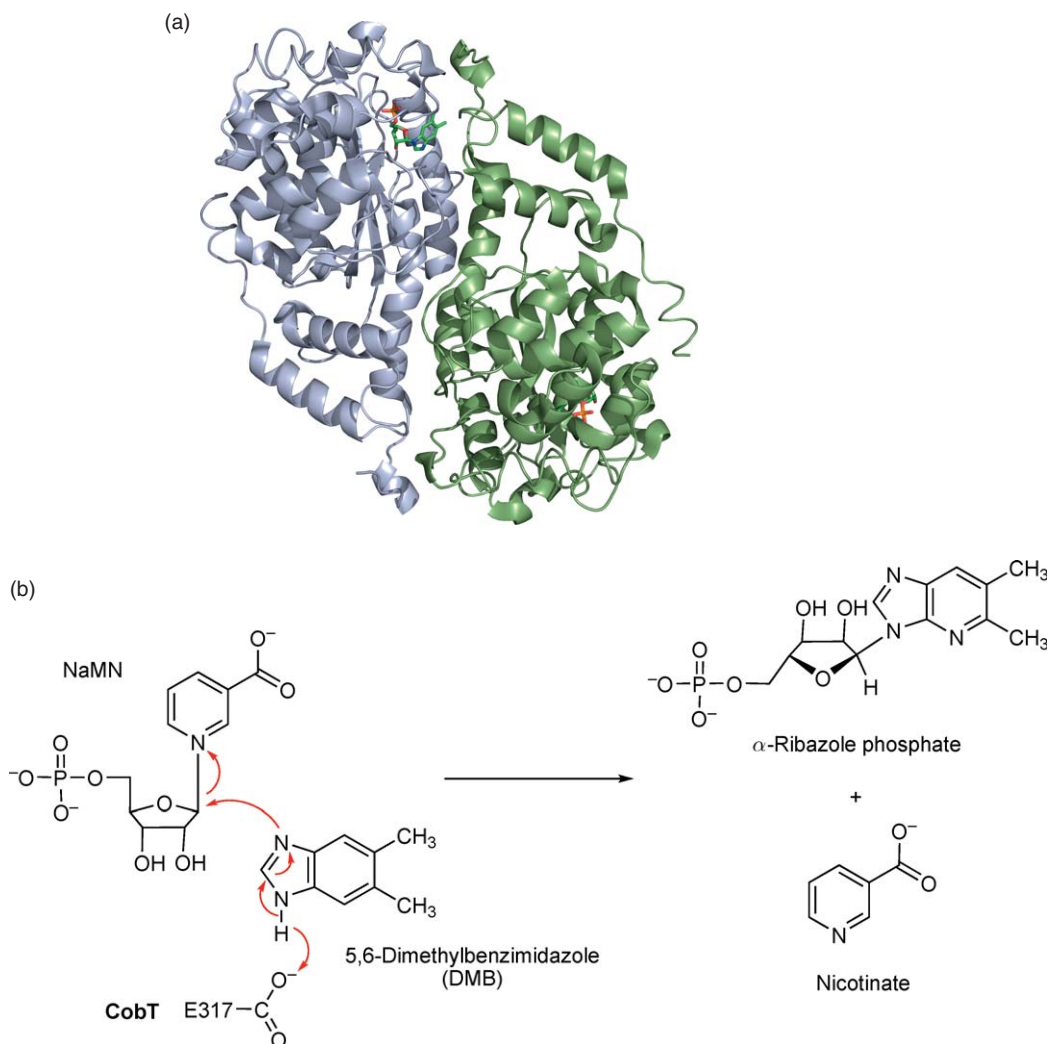
involving the incorporation of erythrose, glycine, formate, and methionine.<sup>237</sup> However, little is known about how this anaerobic biosynthesis is pieced together. In contrast, significant progress has recently been made in how DMB is made from riboflavin (**Figure 27**). In this transformation the C2 of DMB is derived from the C1 of the ribose of FMN. Molecular genetics helped identify the gene required for this process as *bluB*.<sup>238</sup> This encodes a protein with similarity to the NADH/FMN-dependent nitroreductase family. Research on the overproduced enzyme established that the enzyme in the presence of molecular oxygen is able to catalyze the cannibalism of the reduced flavin into DMB, with the release of D-erythrose 4-phosphate.<sup>239,240</sup> The crystal structure of BluB reveals that it has a restricted binding pocket for the reduced flavin, where molecular oxygen binds over the reduced flavin (**Figure 30**).<sup>240</sup> It has been suggested that the mechanism of this reaction may proceed via ring expanded and diaminobenzene intermediates.<sup>241</sup> These are a remarkable set of reactions to be accommodated within a single enzyme active site.

### 7.13.14.3 Synthesis of $\alpha$ -Ribazole

DMB is converted into a nucleoside to give  $\alpha$ -ribazole. This reaction is catalyzed by the *P. denitrificans* CobU<sup>242</sup> or *S. enterica* CobT (**Figure 27**).<sup>243</sup> The activation sees the formation of an *N*- $\alpha$ -glycosidic bond between ribose of  $\beta$ -nicotinate mononucleotide (NaMN) and DMB with the release of nicotinate. The resulting product,  $\alpha$ -ribazole 5'-phosphate is subsequently hydrolyzed by an uncharacterized 5'-phosphatase to give  $\alpha$ -ribazole.



**Figure 30** Mechanism and structure of BluB. The proposed mechanism by Ealick and Begley for the transformation of reduced FMN into dimethylbenzimidazole is shown. All these reactions are performed by a comparatively small homodimeric enzyme called BluB. In the figure, the monomers are colored green helices/orange sheets and blue helices/purple sheets. An oxygen-trapped structure of the enzyme, where the oxygen (ball) is positioned above the isoalloxazine ring (stick representation) of the flavin is shown.



**Figure 31** The structure and mechanism of  $\alpha$ -ribazole synthesis. (a) The structure of the anaerobic pathway  $\alpha$ -ribazole synthase, CobT, is shown. The two subunits of the homodimeric enzyme are differentially colored (green helices/yellow sheets and blue helices/purple sheets). The active site is marked by the binding of the  $\alpha$ -ribazole (stick representation). (b) The mechanism involves the attachment of DMB to the ribose moiety of NaMN, generating  $\alpha$ -ribazole and nicotinic acid.

In *S. enterica*, CobT can use  $\text{NAD}^+$  in place of NaMN as a substrate for the phosphoribosyltransferase activity.<sup>244</sup> The resulting product,  $\alpha$ -DMB adenine dinucleotide, is then cleaved by an unknown enzyme to give AMP and  $\alpha$ -ribazole 5'-phosphate. Detail on the mechanism of CobT has come from structural studies of this dimeric enzyme, which has a subunit molecular mass of 35 kDa (Figure 31).<sup>245–247</sup> Complexes of the enzyme with NaMN, and NaMN and DMB have provided insight into how the substrates are bound within the active site and how catalysis is promoted by proximity.

#### 7.13.14.4 Synthesis of Adenosylcobalamin

The final part in the construction of adenosylcobalamin synthesis involves the transfer of  $\alpha$ -ribazole 5'-phosphate to adenosyl-GDP-cobinamide with displacement of the GDP moiety to generate adenosylcobalamin 5'-phosphate (Figure 27). The enzyme that catalyzes this reaction is CobV in *P. denitrificans*.<sup>242</sup> and CobS in *S. enterica*.<sup>229,248</sup> The enzyme from *P. denitrificans* is able to accept a range of substrates including

$\alpha$ -ribazole and GDP-cobinamide and is called cobalamin (5'-phosphate) synthase.<sup>242</sup> CobV appears to form a complex with a number of other proteins. In *S. enterica* CobS would appear to be an integral membrane protein and has only been produced in small quantities.<sup>229,248</sup>

In *S. enterica*, it has recently been shown that the adenosylcobalamin 5'-phosphate is acted upon by a phosphatase called CobC.<sup>249</sup> This enzyme generates adenosylcobalamin and is therefore technically the final enzyme in cobalamin biosynthesis.<sup>249</sup>

## References

1. R. Banerjee; S. W. Ragsdale, *Annu. Rev. Biochem.* **2003**, *72*, 209–247.
2. A. R. Battersby, *Nat. Prod. Rep.* **2000**, *17*, 507–526.
3. P. M. Shoolingin-Jordan; K. M. Cheung, *The Biosynthesis of Heme*; Elsevier: Pergamon, 1999.
4. R. K. Thauer; L. G. Bonacker, *Ciba Found. Symp.* **1994**, *180*, 210–222; discussion 222–227.
5. M. J. Warren; A. G. Smith, *Tetrapyrroles: Birth, Life and Death*; Landes Biosciences, Springer Science and Business Media: New York, 2008.
6. N. Frankenberg; J. Moser; D. Jahn, *Appl. Microbiol. Biotechnol.* **2003**, *63*, 115–127.
7. I. Matsunaga; Y. Shiro, *Curr. Opin. Chem. Biol.* **2004**, *8*, 127–132.
8. K. R. Rodgers, *Curr. Opin. Chem. Biol.* **1999**, *3*, 158–167.
9. D. W. Bollivar, *Photosynth. Res.* **2006**, *90*, 173–194.
10. G. Fritz; O. Einsle; M. Rudolf; A. Schiffer; P. M. Kroneck, *J. Mol. Microbiol. Biotechnol.* **2005**, *10*, 223–233.
11. D. J. Richardson; N. J. Watmough, *Curr. Opin. Chem. Biol.* **1999**, *3*, 207–219.
12. C. K. Chang, *Ciba Found. Symp.* **1994**, *180*, 228–238; discussion 238–246.
13. N. Frankenberg; J. C. Lagarias, *The Porphyrin Handbook*; Elsevier: Amsterdam, 2003.
14. M. A. van der Horst; J. Key; K. J. Hellingwerf, *Trends Microbiol.* **2007**, *15*, 554–562.
15. B. Buchenau; J. Kahnt; I. U. Heinemann; D. Jahn; R. K. Thauer, *J. Bacteriol.* **2006**, *188*, 8666–8668.
16. T. Ishida; L. Yu; H. Akutsu; K. Ozawa; S. Kawanishi; A. Seto; T. Inubushi; S. Sano, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4853–4858.
17. K. D. Gibson; W. G. Laver; A. Neuberger, *Biochem. J.* **1958**, *70*, 71–81.
18. G. Kikuchi; A. Kumar; P. Talmage; D. Shemin, *J. Biol. Chem.* **1958**, *233*, 1214–1219.
19. Y. J. Avissar; J. G. Ormerod; S. I. Beale, *Arch. Microbiol.* **1989**, *151*, 513–519.
20. S. I. Beale; P. A. Castelfranco, *Biochem. Biophys. Res. Commun.* **1973**, *52*, 143–149.
21. D. Jahn; E. Verkamp; D. Soll, *Trends Biochem. Sci.* **1992**, *17*, 215–218.
22. P. K. Mehta; P. Christen, *Adv. Enzymol. Relat. Areas Mol. Biol.* **2000**, *74*, 129–184.
23. G. Schneider; H. Kack; Y. Lindqvist, *Structure* **2000**, *8*, R1–R6.
24. G. C. Ferreira, *Iron Metabolism*; Wiley-VCH: Weinheim, 1999.
25. E. L. Bolt; L. Kryszak; J. Zeilstra-Ryalls; P. M. Shoolingin-Jordan; M. J. Warren, *Eur. J. Biochem.* **1999**, *265*, 290–299.
26. I. Astner; J. O. Schulze; J. van den Heuvel; D. Jahn; W. D. Schubert; D. W. Heinz, *EMBO J.* **2005**, *24*, 3166–3177.
27. G. A. Hunter; G. C. Ferreira, *Biochemistry* **1999**, *38*, 12526.
28. G. A. Hunter; G. C. Ferreira, *J. Biol. Chem.* **1999**, *274*, 12222–12228.
29. J. Zhang; G. C. Ferreira, *J. Biol. Chem.* **2002**, *277*, 44660–44669.
30. J. Moser; S. Lorenz; C. Hubschwerlen; A. Rompf; D. Jahn, *J. Biol. Chem.* **1999**, *274*, 30679–30685.
31. S. Schauer; S. Chaturvedi; L. Randau; J. Moser; M. Kitabatake; S. Lorenz; E. Verkamp; W. D. Schubert; T. Nakayashiki; M. Murai; K. Wall; H. U. Thomann; D. W. Heinz; H. Inokuchi; D. Soll; D. Jahn, *J. Biol. Chem.* **2002**, *277*, 48657–48663.
32. U. C. Vothknecht; C. G. Kannangara; D. von Wettstein, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9287–9291.
33. L. Randau; S. Schauer; A. Ambrogelly; J. C. Salazar; J. Moser; S. Sekine; S. Yokoyama; D. Soll; D. Jahn, *J. Biol. Chem.* **2004**, *279*, 34931–34937.
34. J. Moser; W. D. Schubert; V. Beier; I. Bringemeier; D. Jahn; D. W. Heinz, *EMBO J.* **2001**, *20*, 6583–6590.
35. H. C. Friedmann; M. E. Duban; A. Valasinas; B. Frydman, *Biochem. Biophys. Res. Commun.* **1992**, *185*, 60–68.
36. L. L. Ilag; D. Jahn, *Biochemistry* **1992**, *31*, 7143–7151.
37. C. E. Pugh; J. L. Harwood; R. A. John, *J. Biol. Chem.* **1992**, *267*, 1584–1588.
38. M. A. Smith; B. Grimm; C. G. Kannangara; D. von Wettstein, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9775–9779.
39. M. A. Smith; C. G. Kannangara; B. Grimm; D. von Wettstein, *Eur. J. Biochem.* **1991**, *202*, 749–757.
40. M. Hennig; B. Grimm; R. Contestabile; R. A. John; J. N. Jansonius, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4866–4871.
41. J. O. Schulze; W. D. Schubert; J. Moser; D. Jahn; D. W. Heinz, *J. Mol. Biol.* **2006**, *358*, 1212–1220.
42. C. Luer; S. Schauer; K. Mobius; J. Schulze; W. D. Schubert; D. W. Heinz; D. Jahn; J. Moser, *J. Biol. Chem.* **2005**, *280*, 18568–18572.
43. L. A. Nogaj; S. I. Beale, *J. Biol. Chem.* **2005**, *280*, 24301–24307.
44. P. M. Anderson; R. J. Desnick, *J. Biol. Chem.* **1979**, *254*, 6924–6930.
45. N. Frankenberg; D. W. Heinz; D. Jahn, *Biochemistry* **1999**, *38*, 13968–13975.
46. K. D. Gibson; A. Neuberger; J. J. Scott, *Biochem. J.* **1955**, *61*, 618–629.
47. W. Liedgens; C. Lutz; H. A. Schneider, *Eur. J. Biochem.* **1983**, *135*, 75–79.
48. P. Spencer; P. M. Jordan, *Biochem. J.* **1993**, *290* (Pt. 1), 279–287.
49. E. K. Jaffe, *Acta Crystallogr. D Biol. Crystallogr.* **2000**, *56*, 115–128.
50. F. Frere; H. Reents; W. D. Schubert; D. W. Heinz; D. Jahn, *J. Mol. Biol.* **2005**, *345*, 1059–1070.
51. D. W. Bollivar; C. Clauson; R. Lighthall; S. Forbes; B. Kokona; R. Fairman; L. Kundrat; E. K. Jaffe, *BMC Biochem.* **2004**, *5*, 17.



52. P. T. Erskine; E. Norton; J. B. Cooper; R. Lambert; A. Coker; G. Lewis; P. Spencer; M. Sarwar; S. P. Wood; M. J. Warren; P. M. Shoolingin-Jordan, *Biochemistry* **1999**, *38*, 4266–4276.
53. P. T. Erskine; N. Senior; S. Awan; R. Lambert; G. Lewis; I. J. Tickle; M. Sarwar; P. Spencer; P. Thomas; M. J. Warren; P. M. Shoolingin-Jordan; S. P. Wood; J. B. Cooper, *Nat. Struct. Biol.* **1997**, *4*, 1025–1031.
54. N. Frankenberger; P. T. Erskine; J. B. Cooper; P. M. Shoolingin-Jordan; D. Jahn; D. W. Heinz, *J. Mol. Biol.* **1999**, *289*, 591–602.
55. P. M. Jordan; P. N. Gibbs, *Biochem. J.* **1985**, *227*, 1015–1020.
56. F. Frere; W. D. Schubert; F. Stauffer; N. Frankenberger; R. Neier; D. Jahn; D. W. Heinz, *J. Mol. Biol.* **2002**, *320*, 237–247.
57. E. K. Jaffe; J. Kervinen; J. Martins; F. Stauffer; R. Neier; A. Wlodawer; A. Zdanov, *J. Biol. Chem.* **2002**, *277*, 19792–19799.
58. J. Kervinen; E. K. Jaffe; F. Stauffer; R. Neier; A. Wlodawer; A. Zdanov, *Biochemistry* **2001**, *40*, 8227–8236.
59. E. Fujino; T. Fujino; S. Karita; T. Kimura; K. Sakka; K. Ohmiya, *J. Biosci. Bioeng.* **1999**, *87*, 535–537.
60. R. M. Jones; P. M. Jordan, *Biochem. J.* **1994**, *299* (Pt. 3), 895–902.
61. P. M. Jordan; S. D. Thomas; M. J. Warren, *Biochem. J.* **1988**, *254*, 427–435.
62. V. A. Nagaraj; S. R. Arumugam; B. Gopalakrishnan; Y. S. Jyothsna; P. N. Rangarajan; G. Padmanaban, *J. Biol. Chem.* **2008**, *283*, 437–444.
63. S. J. Awan; G. Siligardi; P. M. Shoolingin-Jordan; M. J. Warren, *Biochemistry* **1997**, *36*, 9273–9282.
64. P. M. Jordan; M. J. Warren, *FEBS Lett.* **1987**, *225*, 87–92.
65. P. M. Shoolingin-Jordan; M. J. Warren; S. J. Awan, *Biochem. J.* **1996**, *316* (Pt. 2), 373–376.
66. P. M. Shoolingin-Jordan; M. J. Warren; S. J. Awan, *Methods Enzymol.* **1997**, *281*, 317–327.
67. G. V. Louie; P. D. Brownlie; R. Lambert; J. B. Cooper; T. L. Blundell; S. P. Wood; M. J. Warren; S. C. Woodcock; P. M. Jordan, *Nature* **1992**, *359*, 33–39.
68. P. M. Jordan, *Ciba Found. Symp.* **1994**, *180*, 70–89; discussion 89–96.
69. S. C. Woodcock; P. M. Jordan, *Biochemistry* **1994**, *33*, 2688–2695.
70. A. R. Battersby; C. J. R. Fookes; K. E. Gustafsonpotter; G. W. J. Matcham; E. McDonald, *J. Chem. Soc. Chem. Commun.* **1979**, 1155–1158.
71. P. M. Jordan; G. Burton; H. Nordlov; M. M. Schneider; L. Pryde; A. I. Scott, *J. Chem. Soc. Chem. Commun.* **1979**, 204–205.
72. M. Hansson; L. Rutberg; I. Schroder; L. Hederstedt, *J. Bacteriol.* **1991**, *173*, 2590–2599.
73. C. D. Mohr; S. K. Sonstebly; V. Deretic, *Mol. Gen. Genet.* **1994**, *242*, 177–184.
74. A. F. Alwan; B. I. Mgbeje; P. M. Jordan, *Biochem. J.* **1989**, *264*, 397–402.
75. G. J. Hart; A. R. Battersby, *Biochem. J.* **1985**, *232*, 151–160.
76. M. Kohashi; R. P. Clement; J. Tse; W. N. Piper, *Biochem. J.* **1984**, *220*, 755–765.
77. S. F. Tsai; D. F. Bishop; R. J. Desnick, *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7049–7053.
78. M. A. Mathews; H. L. Schubert; F. G. Whitby; K. J. Alexander; K. Schadick; H. A. Bergonia; J. D. Phillips; C. P. Hill, *EMBO J.* **2001**, *20*, 5832–5839.
79. P. H. Rehse; T. Kitao; T. H. Tahirov, *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *61*, 913–919.
80. A. R. Battersby; C. J. R. Fookes; G. W. J. Matcham; P. S. Pandey, *Angew. Chem. Int. Ed. Engl.* **1981**, *20*, 293–295.
81. A. R. Battersby; C. J. R. Fookes; E. McDonald; M. J. Meegan, *J. Chem. Soc. Chem. Commun.* **1978**, 185–186.
82. M. A. Cassidy; N. Crockett; F. J. Leeper; A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1991**, 384–386.
83. C. Pichon; B. P. Atshaves; N. J. Stolowich; A. I. Scott, *Bioorg. Med. Chem.* **1994**, *2*, 153–168.
84. W. M. Stark; G. J. Hart; A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1986**, 465–467.
85. P. M. Shoolingin-Jordan, *The Porphyrin Handbook*; Elsevier: Amsterdam, 2003; Vol. 12.
86. F. Felix; N. Brouillet, *Eur. J. Biochem.* **1990**, *188*, 393–403.
87. R. M. Jones; P. M. Jordan, *Biochem. J.* **1993**, *293* (Pt. 3), 703–712.
88. A. G. Roberts; G. H. Elder, *Methods Enzymol.* **1997**, *281*, 349–355.
89. Y. Seki; S. Kawanishi; S. Sano, *Methods Enzymol.* **1986**, *123*, 415–421.
90. J. Luo; C. K. Lim, *Biochem. J.* **1993**, *289* (Pt. 2), 529–532.
91. J. Fan; Q. Liu; Q. Hao; M. Teng; L. Niu, *J. Bacteriol.* **2007**, *189*, 3573–3580.
92. B. M. Martins; B. Grimm; H. P. Mock; R. Huber; A. Messerschmidt, *J. Biol. Chem.* **2001**, *276*, 44108–44116.
93. F. G. Whitby; J. D. Phillips; J. P. Kushner; C. P. Hill, *EMBO J.* **1998**, *17*, 2463–2471.
94. J. D. Phillips; F. G. Whitby; J. P. Kushner; C. P. Hill, *EMBO J.* **2003**, *22*, 6225–6233.
95. M. Akhtar, *The Porphyrin Handbook*; Elsevier: New York, 2003; Vol. 12; vol. 12.
96. H. A. Dailey, *Biochem. Soc. Trans.* **2002**, *30*, 590–595.
97. D. Breckau; E. Mahlitz; A. Sauerwald; G. Layer; D. Jahn, *J. Biol. Chem.* **2003**, *278*, 46625–46631.
98. P. Labbe, *Methods Enzymol.* **1997**, *281*, 367–378.
99. S. Macieira; B. M. Martins; R. Huber, *FEMS Microbiol. Lett.* **2003**, *226*, 31–37.
100. A. E. Medlock; H. A. Dailey, *J. Biol. Chem.* **1996**, *271*, 32507–32510.
101. T. Yoshinaga, *Methods Enzymol.* **1997**, *281*, 355–367.
102. G. H. Elder; J. O. Evans; J. R. Jackson; A. H. Jackson, *Biochem. J.* **1978**, *169*, 215–223.
103. D. S. Lee; E. Flachsova; M. Bodnarova; B. Demeler; P. Martasek; C. S. Raman, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 14232–14237.
104. J. D. Phillips; F. G. Whitby; C. A. Warby; P. Labbe; C. Yang; J. W. Pflugrath; J. D. Ferrara; H. Robinson; J. P. Kushner; C. P. Hill, *J. Biol. Chem.* **2004**, *279*, 38960–38968.
105. H. Kohno; T. Furukawa; R. Tokunaga; S. Taketani; T. Yoshinaga, *Biochim. Biophys. Acta* **1996**, *1292*, 156–162.
106. G. Layer; K. Verfurth; E. Mahlitz; D. Jahn, *J. Biol. Chem.* **2002**, *277*, 34136–34142.
107. H. J. Sofia; G. Chen; B. G. Hetzler; J. F. Reyes-Spindola; N. E. Miller, *Nucleic Acids Res.* **2001**, *29*, 1097–1106.
108. S. C. Wang; P. A. Frey, *Trends Biochem. Sci.* **2007**, *1774*, 286–296.
109. G. Layer; K. Grage; T. Teschner; V. Schunemann; D. Breckau; A. Masoumi; M. Jahn; P. Heathcote; A. X. Trautwein; D. Jahn, *J. Biol. Chem.* **2005**, *280*, 29038–29046.
110. G. Layer; A. J. Pierik; M. Trost; S. E. Rigby; H. K. Leech; K. Grage; D. Breckau; I. Astner; L. Jansch; P. Heathcote; M. J. Warren; D. W. Heinz; D. Jahn, *J. Biol. Chem.* **2006**, *281*, 15727–15734.



111. G. Layer; J. Moser; D. W. Heinz; D. Jahn; W. D. Schubert, *EMBO J.* **2003**, *22*, 6214–6224.
112. A. V. Corrigan; K. B. Siziba; M. H. Maneli; E. G. Shephard; M. Ziman; T. A. Dailey; H. A. Dailey; R. E. Kirsch; P. N. Meissner, *Arch. Biochem. Biophys.* **1998**, *358*, 251–256.
113. L. J. Siepker; M. Ford; R. de Kock; S. Kramer, *Biochim. Biophys. Acta* **1987**, *913*, 349–358.
114. H. A. Dailey; T. A. Dailey, *J. Biol. Chem.* **1996**, *271*, 8714–8718.
115. T. A. Dailey; H. A. Dailey, *Protein Sci.* **1996**, *5*, 98–105.
116. H. R. Corradi; A. V. Corrigan; E. Boix; C. G. Mohan; E. D. Sturrock; P. N. Meissner; K. R. Acharya, *J. Biol. Chem.* **2006**, *281*, 38625–38633.
117. M. Koch; C. Breithaupt; R. Kiefersauer; J. Freigang; R. Huber; A. Messerschmidt, *EMBO J.* **2004**, *23*, 1720–1728.
118. A. Sasarman; J. Letowski; G. Czaika; V. Ramirez; M. A. Nead; J. M. Jacobs; R. Morais, *Can. J. Microbiol.* **1993**, *39*, 1155–1161.
119. T. A. Dailey; H. A. Dailey; P. Meissner; A. R. Prasad, *Arch. Biochem. Biophys.* **1995**, *324*, 379–384.
120. I. Lermontova; E. Kruse; H. P. Mock; B. Grimm, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8895–8900.
121. H. A. Dailey; T. A. Dailey, *The Porphyrin Handbook*; 2003; Vol. 12.
122. S. Al-Karadaghi; M. Hansson; S. Nikonov; B. Jonsson; L. Hederstedt, *Structure* **1997**, *5*, 1501–1510.
123. E. Grzybowska; M. Gora; D. Plochocka; J. Rytka, *Arch. Biochem. Biophys.* **2002**, *398*, 170–178.
124. C. K. Wu; H. A. Dailey; J. P. Rose; A. Burden; V. M. Sellers; B. C. Wang, *Nat. Struct. Biol.* **2001**, *8*, 156–160.
125. H. A. Dailey; C. S. Jones; S. W. Karr, *Biochim. Biophys. Acta* **1989**, *999*, 7–11.
126. S. Shipovskov; T. Karlberg; M. Fodje; M. D. Hansson; G. C. Ferreira; M. Hansson; C. T. Reimann; S. Al-Karadaghi, *J. Mol. Biol.* **2005**, *352*, 1081–1090.
127. T. Karlberg; D. Lecerof; M. Gora; G. Silvegren; R. Labbe-Bois; M. Hansson; S. Al-Karadaghi, *Biochemistry* **2002**, *41*, 13499–13506.
128. A. Medlock; L. Swartz; T. A. Dailey; H. A. Dailey; W. N. Lanzilotta, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1789–1793.
129. D. Lecerof; M. N. Fodje; R. Alvarez Leon; U. Olsson; A. Hansson; E. Sigfridsson; U. Ryde; M. Hansson; S. Al-Karadaghi, *J. Biol. Inorg. Chem.* **2003**, *8*, 452–458.
130. A. Eschenmoser; C. E. Wintner, *Science* **1977**, *196*, 1410–1420.
131. H. C. Uzar; A. R. Battersby; T. A. Carpenter; F. J. Leeper, *J. Chem. Soc. Perkin Trans.* **1987**, *1*, 1689–1696.
132. J. Crouzet; B. Cameron; L. Cauchois; S. Rigault; M. C. Rouyez; F. Blanche; D. Thibaut; L. Debussche, *J. Bacteriol.* **1990**, *172*, 5980–5990.
133. J. Crouzet; L. Cauchois; F. Blanche; L. Debussche; D. Thibaut; M. C. Rouyez; S. Rigault; J. F. Mayaux; B. Cameron, *J. Bacteriol.* **1990**, *172*, 5968–5979.
134. R. M. Jeter; J. R. Roth, *J. Bacteriol.* **1987**, *169*, 3189–3198.
135. J. R. Roth; J. G. Lawrence; M. Rubenfield; S. Kieffer-Higgins; G. M. Church, *J. Bacteriol.* **1993**, *175*, 3303–3316.
136. B. Cameron; K. Briggs; S. Pridmore; G. Brefort; J. Crouzet, *J. Bacteriol.* **1989**, *171*, 547–557.
137. E. Raux; H. L. Schubert; M. J. Warren, *Cell Mol. Life Sci.* **2000**, *57*, 1880–1893.
138. J. R. Roth; J. G. Lawrence; T. A. Bobik, *Annu. Rev. Microbiol.* **1996**, *50*, 137–181.
139. G. L. Holliday; J. M. Thornton; A. Marquet; A. G. Smith; F. Rebeille; R. Mendel; H. L. Schubert; A. D. Lawrence; M. J. Warren, *Nat. Prod. Rep.* **2007**, *24*, 972–987.
140. F. Blanche; L. Debussche; D. Thibaut; J. Crouzet; B. Cameron, *J. Bacteriol.* **1989**, *171*, 4222–4231.
141. F. Blanche; C. Robin; M. Couder; D. Faucher; L. Cauchois; B. Cameron; J. Crouzet, *J. Bacteriol.* **1991**, *173*, 4637–4645.
142. H. L. Schubert; R. M. Blumenthal; X. Cheng, *Trends Biochem. Sci.* **2003**, *28*, 329–335.
143. J. Vevodova; R. M. Graham; E. Raux; H. L. Schubert; D. I. Roper; A. A. Brindley; A. Ian Scott; C. A. Roessner; N. P. Stamford; M. Elizabeth Stroupe; E. D. Getzoff; M. J. Warren; K. S. Wilson, *J. Mol. Biol.* **2004**, *344*, 419–433.
144. M. E. Stroupe; H. K. Leech; D. S. Daniels; M. J. Warren; E. D. Getzoff, *Nat. Struct. Biol.* **2003**, *10*, 1064–1073.
145. J. H. Martens; H. B. Karg; M. J. Warren; D. Jahn, *Appl. Microbiol. Biotechnol.* **2002**, *58*, 275–285.
146. F. Blanche; B. Cameron; J. Crouzet; L. Debussche; S. Levy-Schil; D. Thibaut, Rhône-Poulenc Biochimie. Eur. Patent 0516647 B1, 1998.
147. F. Blanche; D. Thibaut; L. Debussche; R. Hertle; F. Zipfel; G. Muller, *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1651–1653.
148. D. Thibaut; M. Couder; J. Crouzet; L. Debussche; B. Cameron; F. Blanche, *J. Bacteriol.* **1990**, *172*, 6245–6251.
149. M. J. Warren; C. A. Roessner; S. Ozaki; N. J. Stolowich; P. J. Santander; A. I. Scott, *Biochemistry* **1992**, *31*, 603–609.
150. A. Eschenmoser, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 5–39.
151. A. I. Scott; C. A. Roessner; N. J. Stolowich; J. B. Spencer; C. Min; S. I. Ozaki, *FEBS Lett.* **1993**, *331*, 105–108.
152. L. Debussche; D. Thibaut; B. Cameron; J. Crouzet; F. Blanche, *J. Bacteriol.* **1993**, *175*, 7430–7440.
153. L. Debussche; D. Thibaut; M. Danzer; F. Debu; D. Frechet; F. Herman; F. Blanche; M. Vuilhorgne, *J. Chem. Soc. Chem. Commun.* **1993**, 1100–1103.
154. J. B. Spencer; N. J. Stolowich; C. A. Roessner; C. H. Min; A. I. Scott, *J. Am. Chem. Soc.* **1993**, *115*, 11610–11611.
155. N. J. Stolowich; J. J. Wang; J. B. Spencer; P. J. Santander; C. A. Roessner; A. I. Scott, *J. Am. Chem. Soc.* **1996**, *118*, 1657–1662.
156. H. M. McGoldrick; C. A. Roessner; E. Raux; A. D. Lawrence; K. J. McLean; A. W. Munro; S. Santabarbara; S. E. Rigby; P. Heathcote; A. I. Scott; M. J. Warren, *J. Biol. Chem.* **2005**, *280*, 1086–1094.
157. D. Thibaut; L. Debussche; D. Frechet; F. Herman; M. Vuilhorgne; F. Blanche, *J. Chem. Soc. Chem. Commun.* **1993**, 513–515.
158. A. I. D. Alanine; K. Ichinose; D. Thibaut; L. Debussche; N. P. J. Stamford; F. J. Leeper; F. Blanche; A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1994**, 193–196.
159. C. H. Min; B. P. Atshaves; C. A. Roessner; N. J. Stolowich; J. B. Spencer; A. I. Scott, *J. Am. Chem. Soc.* **1993**, *115*, 10380–10381.
160. F. Blanche; D. Thibaut; A. Famechon; L. Debussche; B. Cameron; J. Crouzet, *J. Bacteriol.* **1992**, *174*, 1036–1042.
161. D. Thibaut; F. Kiuchi; L. Debussche; F. J. Leeper; F. Blanche; A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1992**, 139–141.
162. F. Kiuchi; D. Thibaut; L. Debussche; F. J. Leeper; F. Blanche; A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1992**, 306–308.
163. F. Blanche; A. Famechon; D. Thibaut; L. Debussche; B. Cameron; J. Crouzet, *J. Bacteriol.* **1992**, *174*, 1050–1052.

164. D. Thibaut; F. Kiuchi; L. Debussche; F. Blanche; M. Kodera; F. J. Leeper; A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1992**, 982–985.
165. J. P. Keller; P. M. Smith; J. Benach; D. Christendat; G. T. deTitta; J. F. Hunt, *Structure* **2002**, *10*, 1475–1487.
166. D. Thibaut; M. Couder; A. Famechon; L. Debussche; B. Cameron; J. Crouzet; F. Blanche, *J. Bacteriol.* **1992**, *174*, 1043–1049.
167. B. Dresow; G. Schlingmann; L. Ernst; V. B. Koppenhagen, *J. Biol. Chem.* **1980**, *255*, 7637–7644.
168. F. Blanche; D. Thibaut; D. Frechet; M. Vuilhorgne; J. Crouzet; B. Cameron; K. Hlineny; U. Traubeberhard; M. Zboron; G. Muller, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 884–886.
169. L. W. Shipman; D. Li; C. A. Roessner; A. I. Scott; J. C. Sacchettini, *Structure* **2001**, *9*, 587–596.
170. L. Debussche; D. Thibaut; B. Cameron; J. Crouzet; F. Blanche, *J. Bacteriol.* **1990**, *172*, 6239–6244.
171. M. Y. Galperin; N. V. Grishin, *Proteins-Struct. Funct. Genet.* **2000**, *41*, 238–247.
172. L. Debussche; M. Couder; D. Thibaut; B. Cameron; J. Crouzet; F. Blanche, *J. Bacteriol.* **1992**, *174*, 7445–7451.
173. L. C. Gibson; R. D. Willows; C. G. Kannangara; D. von Wettstein; C. N. Hunter, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1941–1944.
174. J. D. Reid; C. N. Hunter, *J. Biol. Chem.* **2004**, *279*, 26893–26899.
175. A. A. Brindley; E. Raux; H. K. Leech; H. L. Schubert; M. J. Warren, *J. Biol. Chem.* **2003**, *278*, 22388–22395.
176. F. Blanche; L. Maton; L. Debussche; D. Thibaut, *J. Bacteriol.* **1992**, *174*, 7452–7454.
177. A. D. Lawrence; E. Deery; K. J. McLean; A. W. Munro; R. W. Pickersgill; S. E. Rigby; M. J. Warren, *J. Biol. Chem.* **2008**, *283*, 10813–10821.
178. J. Crouzet; S. Levyschil; B. Cameron; L. Cauchois; S. Rigault; M. C. Rouyez; F. Blanche; L. Debussche; D. Thibaut, *J. Bacteriol.* **1991**, *173*, 6074–6087.
179. L. Debussche; M. Couder; D. Thibaut; B. Cameron; J. Crouzet; F. Blanche, *J. Bacteriol.* **1991**, *173*, 6300–6302.
180. N. R. Buan; S. J. Suh; J. C. Escalante-Semerena, *J. Bacteriol.* **2004**, *186*, 5708–5714.
181. C. L. Johnson; E. Pechonick; S. D. Park; G. D. Havemann; N. A. Leal; T. A. Bobik, *J. Bacteriol.* **2001**, *183*, 1577–1584.
182. S. J. Suh; J. C. Escalante-Semerena, *Gene* **1993**, *129*, 93–97.
183. F. Blanche; M. Couder; L. Debussche; D. Thibaut; B. Cameron; J. Crouzet, *J. Bacteriol.* **1991**, *173*, 6046–6051.
184. C. A. Roessner; A. I. Scott, *Chem. Biol.* **1996**, *3*, 325–330.
185. F. Blanche; B. Cameron; J. Crouzet; L. Debussche; D. Thibaut; M. Vuilhorgne; F. J. Leeper; A. R. Battersby, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 383–411.
186. A. R. Battersby, *Science* **1994**, *264*, 1551–1557.
187. P. J. Santander; Y. Kajiwara; H. J. Williams; A. I. Scott, *Bioorg. Med. Chem.* **2006**, *14*, 724–731.
188. C. A. Roessner; A. I. Scott, *J. Bacteriol.* **2006**, *188*, 7331–7334.
189. Y. Kajiwara; P. J. Santander; C. A. Roessner; L. M. Perez; A. I. Scott, *J. Am. Chem. Soc.* **2006**, *128*, 9971–9978.
190. C. A. Roessner; H. J. Williams; A. I. Scott, *J. Biol. Chem.* **2005**, *280*, 16748–16753.
191. E. Raux; H. K. Leech; R. Beck; H. L. Schubert; P. J. Santander; C. A. Roessner; A. I. Scott; J. H. Martens; D. Jahn; C. Thermes; A. Rambach; M. J. Warren, *Biochem. J.* **2003**, *370*, 505–516.
192. J. B. Spencer; N. J. Stolowich; C. A. Roessner; A. I. Scott, *FEBS Lett.* **1993**, *335*, 57–60.
193. E. Raux; T. McVeigh; S. E. Peters; T. Leustek; M. J. Warren, *Biochem. J.* **1999**, *338*, 701–708.
194. M. J. Warren; E. L. Bolt; C. A. Roessner; A. I. Scott; J. B. Spencer; S. C. Woodcock, *Biochem. J.* **1994**, *302* (Pt. 3), 837–844.
195. H. K. Leech; E. Raux; K. J. McLean; A. W. Munro; N. J. Robinson; G. P. Borrelly; M. Malten; D. Jahn; S. E. Rigby; P. Heathcote; M. J. Warren, *J. Biol. Chem.* **2003**, *278*, 41900–41907.
196. E. Raux; C. Thermes; P. Heathcote; A. Rambach; M. J. Warren, *J. Bacteriol.* **1997**, *179*, 3202–3212.
197. H. L. Schubert; E. Raux; K. S. Wilson; M. J. Warren, *Biochemistry* **1999**, *38*, 10660–10669.
198. J. Yin; L. X. Xu; M. M. Cherney; E. Raux-Deery; A. A. Bindley; A. Savchenko; J. R. Walker; M. E. Cuff; M. J. Warren; M. N. James, *J. Struct. Funct. Genomics* **2006**, *7*, 37–50.
199. P. Spencer; N. J. Stolowich; L. W. Sumner; A. I. Scott, *Biochemistry* **1998**, *37*, 14917–14927.
200. S. Frank; E. Deery; A. A. Brindley; H. K. Leech; A. Lawrence; P. Heathcote; H. L. Schubert; K. Brocklehurst; S. E. Rigby; M. J. Warren; R. W. Pickersgill, *J. Biol. Chem.* **2007**, *282*, 23957–23969.
201. K. Wada; J. Harada; Y. Yaeda; H. Tamiaki; H. Oh-Oka; K. Fukuyama, *FEBS J.* **2007**, *274*, 563–573.
202. M. J. Warren; E. Raux; H. L. Schubert; J. C. Escalante-Semerena, *Nat. Prod. Rep.* **2002**, *19*, 390–412.
203. A. I. Scott; C. A. Roessner, *Biochem. Soc. Trans.* **2002**, *30*, 613–620.
204. P. J. Santander; C. A. Roessner; N. J. Stolowich; M. T. Holderman; A. I. Scott, *Chem. Biol.* **1997**, *4*, 659–666.
205. A. I. Scott; N. J. Stolowich; J. Wang; O. Gawatz; E. Fridrich; G. Muller, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14316–14319.
206. H. L. Schubert; K. S. Wilson; E. Raux; S. C. Woodcock; M. J. Warren, *Nat. Struct. Biol.* **1998**, *5*, 585–592.
207. J. Wang; N. J. Stolowich; P. J. Santander; J. H. Park; A. I. Scott, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14320–14322.
208. A. I. Scott; N. J. Stolowich; B. P. Atshaves; P. Karuso; M. J. Warren; H. J. Williams; M. Kajiwara; K. Kurumaya; T. Okazaki, *J. Am. Chem. Soc.* **1991**, *113*, 9891–9893.
209. R. A. Vishwakarma; S. Balachandran; A. I. D. Alanine; N. P. J. Stamford; F. Kiuchi; F. J. Leeper; A. R. Battersby, *J. Chem. Soc. Perkin Trans.* **1993**, *1*, 2893–2899.
210. J. B. Spencer; N. J. Stolowich; P. J. Santander; C. Pichon; M. Kajiwara; S. Tokiwa; K. Takatori; A. I. Scott, *J. Am. Chem. Soc.* **1994**, *116*, 4991–4992.
211. Y. F. Li; A. I. D. Alanine; R. A. Vishwakarma; S. Balachandran; F. J. Leeper; A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1994**, 2507–2508.
212. Y. Xue; Z. Wei; X. Li; W. Gong, *J. Struct. Biol.* **2006**, *153*, 307–311.
213. V. Fresquet; L. Williams; F. M. Raushel, *Biochemistry* **2004**, *43*, 10619–10627.
214. L. Williams; V. Fresquet; P. J. Santander; F. M. Raushel, *J. Am. Chem. Soc.* **2007**, *129*, 294–295.
215. V. Fresquet; L. Williams; F. M. Raushel, *Biochemistry* **2007**, *46*, 13983–13993.
216. J. C. Escalante-Semerena, *J. Bacteriol.* **2007**, *189*, 4555–4560.
217. M. V. Fonseca; J. C. Escalante-Semerena, *J. Biol. Chem.* **2001**, *276*, 32101–32108.
218. M. V. Fonseca; J. C. Escalante-Semerena, *J. Bacteriol.* **2000**, *182*, 4304–4309.

219. S. Suh; J. C. Escalante-Semerena, *J. Bacteriol.* **1995**, *177*, 921–925.
220. M. V. Fonseca; N. R. Buan; A. R. Horswill; I. Rayment; J. C. Escalante-Semerena, *J. Biol. Chem.* **2002**, *277*, 33127–33131.
221. C. B. Bauer; M. V. Fonseca; H. M. Holden; J. B. Thoden; T. B. Thompson; J. C. Escalante-Semerena; I. Rayment, *Biochemistry* **2001**, *40*, 361–374.
222. T. A. Stich; N. R. Buan; J. C. Escalante-Semerena; T. C. Brunold, *J. Am. Chem. Soc.* **2005**, *127*, 8710–8719.
223. H. C. Friedmann; L. M. Cagen, *Annu. Rev. Microbiol.* **1970**, *24*, 159–208.
224. K. R. Brushaber; G. A. O'Toole; J. C. Escalante-Semerena, *J. Biol. Chem.* **1998**, *273*, 2684–2691.
225. C. G. Cheong; C. B. Bauer; K. R. Brushaber; J. C. Escalante-Semerena; I. Rayment, *Biochemistry* **2002**, *41*, 4798–4808.
226. C. G. Cheong; J. C. Escalante-Semerena; I. Rayment, *Biochemistry* **2002**, *41*, 9079–9089.
227. C. Fan; T. A. Bobik, *J. Biol. Chem.* **2008**, *283*, 11322–11329.
228. T. A. Bobik; G. D. Havemann; R. J. Busch; D. S. Williams; H. C. Aldrich, *J. Bacteriol.* **1999**, *181*, 5967–5975.
229. C. L. Zayas; K. Claas; J. C. Escalante-Semerena, *J. Bacteriol.* **2007**, *189*, 7697–7708.
230. F. Blanche; L. Debussche; A. Famechon; D. Thibaut; B. Cameron; J. Crouzet, *J. Bacteriol.* **1991**, *173*, 6052–6057.
231. G. A. O'Toole; J. C. Escalante-Semerena, *J. Biol. Chem.* **1995**, *270*, 23560–23569.
232. G. A. O'Toole; J. C. Escalante-Semerena, *J. Bacteriol.* **1993**, *175*, 6328–6336.
233. M. G. Thomas; T. B. Thompson; I. Rayment; J. C. Escalante-Semerena, *J. Biol. Chem.* **2000**, *275*, 27576–27586.
234. T. B. Thompson; M. G. Thomas; J. C. Escalante-Semerena; I. Rayment, *Biochemistry* **1998**, *37*, 7686–7695.
235. T. B. Thompson; M. G. Thomas; J. C. Escalante-Semerena; I. Rayment, *Biochemistry* **1999**, *38*, 12995–13005.
236. P. Renz, *FEBS Lett.* **1970**, *6*, 187–189.
237. J. R. Vogt; P. Renz, *Eur. J. Biochem.* **1988**, *171*, 655–659.
238. G. R. Campbell; M. E. Taga; K. Mistry; J. Lloret; P. J. Anderson; J. R. Roth; G. C. Walker, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4634–4639.
239. M. J. Gray; J. C. Escalante-Semerena, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2921–2926.
240. M. E. Taga; N. A. Larsen; A. R. Howard-Jones; C. T. Walsh; G. C. Walker, *Nature* **2007**, *446*, 449–453.
241. S. E. Ealick; T. P. Begley, *Nature* **2007**, *446*, 387–388.
242. B. Cameron; F. Blanche; M. C. Rouyez; D. Bisch; A. Famechon; M. Couder; L. Cauchois; D. Thibaut; L. Debussche; J. Crouzet, *J. Bacteriol.* **1991**, *173*, 6066–6073.
243. J. R. Trzebiatowski; G. A. O'Toole; J. C. Escalante-Semerena, *J. Bacteriol.* **1994**, *176*, 3568–3575.
244. L. A. Maggio-Hall; J. C. Escalante-Semerena, *Microbiology* **2003**, *149*, 983–990.
245. C. G. Cheong; J. C. Escalante-Semerena; I. Rayment, *Biochemistry* **1999**, *38*, 16125–16135.
246. C. G. Cheong; J. C. Escalante-Semerena; I. Rayment, *J. Biol. Chem.* **2001**, *276*, 37612–37620.
247. C. G. Cheong; J. C. Escalante-Semerena; I. Rayment, *J. Biol. Chem.* **2002**, *277*, 41120–41127.
248. L. A. Maggio-Hall; K. R. Claas; J. C. Escalante-Semerena, *Microbiology* **2004**, *150*, 1385–1395.
249. C. L. Zayas; J. C. Escalante-Semerena, *J. Bacteriol.* **2007**, *189*, 2210–2218.

### Biographical Sketches



Gunhild Layer studied chemistry at Freiburg University, Germany. She did her Ph.D. work in the group of Dieter Jahn at Braunschweig University. She was a postdoc with Marc Fontecave at the CEA Grenoble, France, before she returned to the Institute of Microbiology at Braunschweig University as leader of a Junior Research Group. Her research interests include tetrapyrrole biosynthesis and iron–sulfur cluster biosynthesis.



Dieter Jahn studied biology and chemistry and did his Ph.D. work at Marburg University, Germany. He was a postdoc with Dieter Söll at Yale University before he returned to the Max-Planck Institute for Microbiology in Marburg. He was an associate professor of biochemistry at Freiburg University before he got a chair in microbiology at Braunschweig University. Besides tetrapyrrole biosynthesis his research focuses on the systems biology of bacterial infections and biotechnological production processes.



Evelyne Deery is a postdoctoral associate at the University of Kent. She started her career as a technician with Dr. Alain Rambach at CHROMagar in Paris, before moving to do a Ph.D. with Martin Warren at University College London. She has over 18 years experience on the molecular genetics of cobalamin biosynthesis.



Andrew D. Lawrence is a postdoctoral associate at the University of Kent. He completed his Ph.D. with Martin Warren at Queen Mary, University of London in 2006 and is currently working on the structural and functional characterization of cobalamin biosynthetic enzymes.



Martin J. Warren is a BBSRC professorial fellow and professor of biochemistry at the University of Kent. He completed his Ph.D. with Professor Peter Shoolingin-Jordan before doing a postdoc stint with Professor Ian Scott at Texas A&M University. He has had academic appointments at both University College London and Queen Mary, University of London, before taking up his current position at the University of Kent.

## 7.14 Cobalamin Coenzymes in Enzymology

Perry Allen Frey, University of Wisconsin-Madison, Madison, WI, USA

© 2010 Elsevier Ltd. All rights reserved.

---

7.14.1	<b>Introduction</b>	502
7.14.2	<b>Early Theories of B<sub>12</sub> Function</b>	503
7.14.2.1	Enzymatic Reactions of Adenosylcobalamin (Coenzyme B <sub>12</sub> )	503
7.14.2.2	Methylcobalamin	505
7.14.3	<b>Cobalamin Chemistry</b>	506
7.14.3.1	Oxidation States	506
7.14.3.2	Chemical Reactions	506
7.14.3.3	Reactions of Alkylcobalamins	507
7.14.4	<b>Adenosylcobalamin in Enzymatic Catalysis</b>	509
7.14.4.1	<b>Diol Dehydrases</b>	509
7.14.4.1.1	Stereochemistry	509
7.14.4.1.2	Suicide inactivation	511
7.14.4.1.3	Hydrogen transfer by adenosylcobalamin	511
7.14.4.1.4	A free radical intermediate	512
7.14.4.1.5	Subunit composition and structure	514
7.14.4.1.6	Spanning the distance	515
7.14.4.1.7	The roles of K <sup>+</sup>	516
7.14.4.1.8	Isomerization mechanism	517
7.14.4.2	<b>Ethanolamine Ammonia-Lyase</b>	518
7.14.4.2.1	Stereochemistry	519
7.14.4.2.2	Participation of coenzyme B <sub>12</sub> in hydrogen transfer	519
7.14.4.2.3	Suicide inactivation	520
7.14.4.2.4	Radical intermediates	520
7.14.4.2.5	Structure	521
7.14.4.2.6	Reaction mechanism	521
7.14.4.3	<b>Lysine 5,6-Aminomutase</b>	523
7.14.4.3.1	Reaction and molecular properties	524
7.14.4.3.2	Molecular structure	524
7.14.4.3.3	Reaction mechanism	525
7.14.4.3.4	Suicide inactivation by substrates	525
7.14.4.4	<b>Methylmalonyl-CoA Mutase</b>	527
7.14.4.4.1	Reaction and molecular composition	527
7.14.4.4.2	Structure of methylmalonyl-CoA mutase	528
7.14.4.4.3	The radical mechanism of methylmalonyl-CoA mutase	528
7.14.4.4.4	Cleavage of the Co–C5' bond	530
7.14.4.5	<b>Glutamate and Methylene-glutarate Mutases</b>	531
7.14.4.5.1	Function of adenosylcobalamin	531
7.14.4.5.2	Alternative and suicidal substrates	532
7.14.4.5.3	Molecular properties and structure	532
7.14.4.5.4	Reaction mechanism	534
7.14.4.5.5	2-Methylene-glutarate mutase	535
7.14.4.6	<b>Ribonucleoside Triphosphate Reductase</b>	535
7.14.4.6.1	Discovery and properties	535
7.14.4.6.2	Reaction mechanism	536

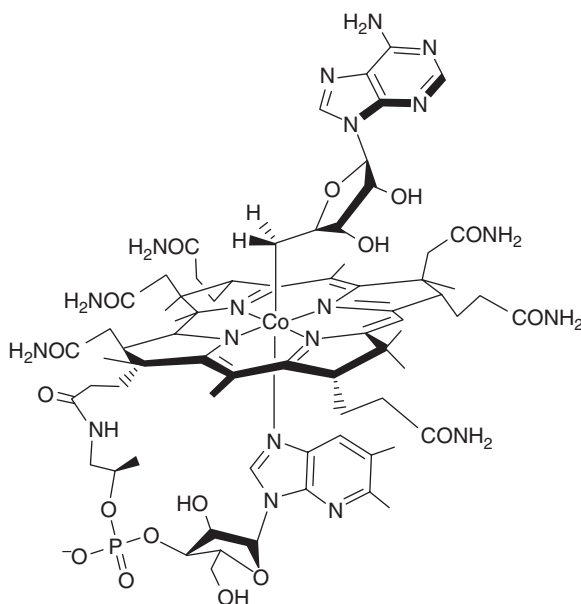
---



7.14.5	Methylcobalamin in Enzymatic Catalysis	538
7.14.5.1	Methionine Synthase	538
7.14.5.1.1	Molecular properties	538
7.14.5.1.2	Reaction mechanism	538
7.14.5.1.3	Modular function	539
7.14.6	Conclusion	540
References		542

### 7.14.1 Introduction

Vitamin B<sub>12</sub>, the antipernicious anemia factor, is a cobalt(III) complex of any of several macrocyclic tetradentate ligand systems known as cobamides. One biologically active derivative of the vitamin is 5'-deoxyadenosylcobalamin (adenosylcobalamin), also known as coenzyme B<sub>12</sub>, a ruby red compound, shown in **Figure 1**. Another full name no longer in use is 5,6-dimethylbenzimidazolyl cobamide coenzyme, in which coenzyme refers to the 5'-deoxyadenosyl derivative of 5,6-dimethylbenzimidazolyl cobamide. Cobamide refers to the tetradentate macrocyclic ligand in **Figure 1**. Variants differ in the identities of the lower axial base, which is most commonly 5,6-dimethylbenzimidazole but may also be benzimidazole, methylbenzimidazole, or adenine. In the last case, the basic macrocycle is known as adenyl cobamide. The most common forms with 5,6-dimethylbenzimidazole are known as cobalamins. The commercial vitamin is cyanocobalamin and differs from adenosylcobalamin in **Figure 1** by the presence of cyanide in place of 5'-deoxyadenosyl as the upper axial ligand to cobalt(III). The vitamin is converted into biologically active forms *in vivo* through enzymatic alkylation of cobalt by the 5'-deoxyadenosyl moiety of adenosine triphosphate (ATP) to form adenosylcobalamin, or by the methyl group of either 5-methyltetrahydrofolate or S-adenosylmethionine (SAM) to form methylcobalamin.



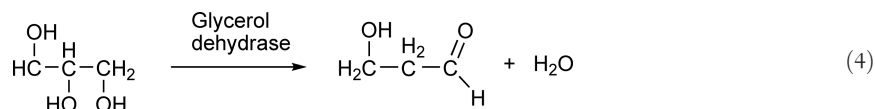
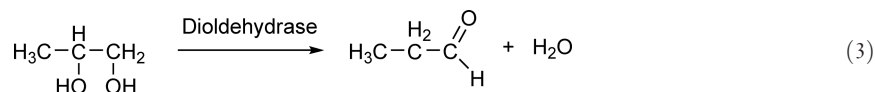
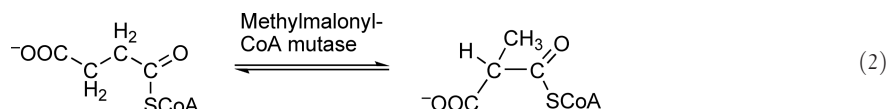
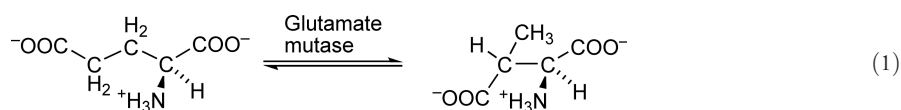
**Figure 1** Structure of adenosylcobalamin.

## 7.14.2 Early Theories of B<sub>12</sub> Function

### 7.14.2.1 Enzymatic Reactions of Adenosylcobalamin (Coenzyme B<sub>12</sub>)

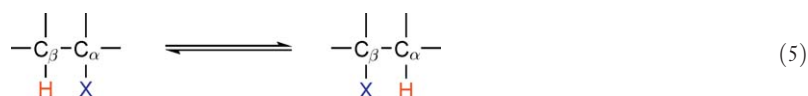
The state of knowledge in the field of coenzyme B<sub>12</sub> biochemistry in the early 1960s, closely following the discovery of the Co–C5' bond in the crystal structure of adenosylcobalamin,<sup>1</sup> was one of the flourishing ideas tempered by mystery. Such unstable bonds had not been expected to exist in aqueous solutions under physiological conditions, and we shall see that they are just stable enough. In this period, confusion reigned, the function of coenzyme B<sub>12</sub> was far from obvious, and creative ideas and theories flourished. The theories set the stage for fascinating and surprising experimentation, which contradicted and falsified most of the theories and led eventually to a unified understanding of the function of adenosylcobalamin. The molecule proved to be well fitted to its role in biochemistry.

The enzymatic reactions in which adenosylcobalamin was known to serve as coenzyme by 1964 were those of glutamate mutase,<sup>2</sup> methylmalonyl-CoA mutase (MCM),<sup>3,4</sup> propane-1,2-diol hydro-lyase (dioldehydrase, DDH),<sup>5</sup> and glycerol hydro-lyase (glycerol dehydrase),<sup>6,7</sup> shown in Equations (1)–(4).



The mutases and dehydrases appeared to have little in common from a chemical standpoint.

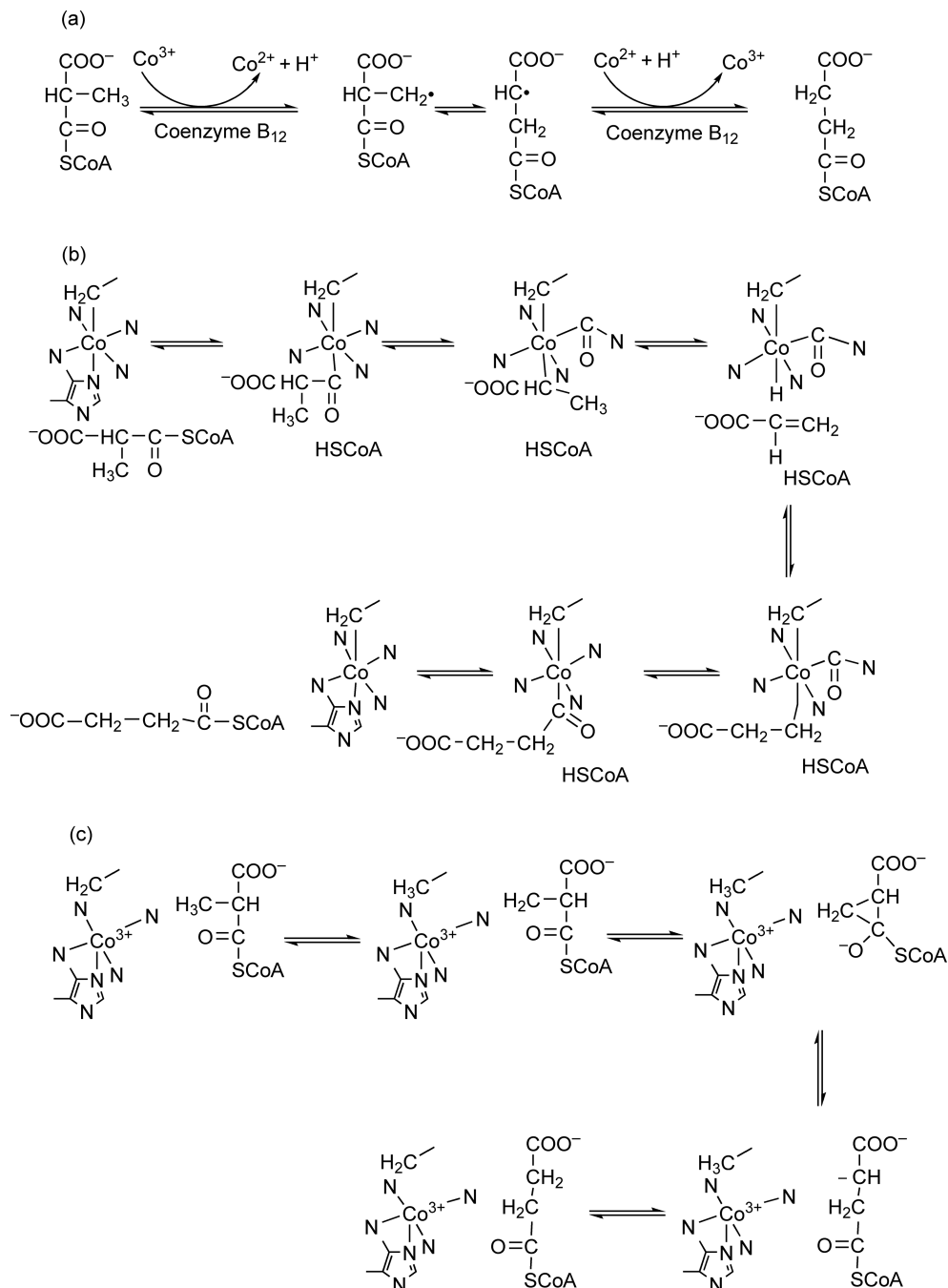
Research into most of these reactions was fundamentally descriptive. It had been determined that the atoms in the molecules were preserved in the products; that is, no hydrogen from water was incorporated into any of the products.<sup>8–11</sup> This would change with elucidation of the properties of class II ribonucleotide reductases.<sup>12</sup> The stereochemistry of each transformation was under investigation.<sup>13,14</sup> Mass spectrometric experiments showed that in the reaction of MCM, the COCoA moiety migrated from C2 to C3 of succinyl-CoA,<sup>15,16</sup> moreover intramolecularly within the same molecule and not in an intermolecular rearrangement between two molecules.<sup>17</sup> A corollary to intramolecular acyl-CoA or glycol migration, and of the retention of substrate hydrogen, was the cross-migration of a hydrogen atom, as in Equation (5).



This pattern of rearrangement could be adapted to the reactions of glycol dehydrases by postulating the analogous cross-migration of an HO group and H to form initially the aldehyde hydrate, which would undergo dehydration to the free aldehyde. Further experiments verified the paradigm of Equation (5) for adenosylcobalamin-dependent isomerization and dehydration reactions.

While most early experiments did not directly address the function of adenosylcobalamin, this question was at the forefront of thought and theorization, and a number of theories were put forward based on the knowledge of metallic and organometallic chemistry available at the time. Some of the theories addressed specific reactions and others were put forward as general principles for all coenzyme B<sub>12</sub>-dependent reactions.

Before the crystal structure of adenosylcobalamin became known, coenzyme B<sub>12</sub> had been described as an adenine derivative of vitamin B<sub>12</sub>.<sup>2</sup> Taking into consideration the cobaltic nature of coenzyme B<sub>12</sub>, the mechanism in **Figure 2(a)** for the reaction of MCM was put forward in 1960.<sup>18</sup> This was a proton-coupled electron transfer and free radical mechanism, in which the Co<sup>3+</sup> of coenzyme B<sub>12</sub> was postulated to potentiate the removal of an electron and proton from the substrate to form a carbon-centered substrate radical and Co<sup>2+</sup>. The substrate radical would undergo rearrangement to the product-related radical, which would then accept an



**Figure 2** Early mechanistic theories of the action of adenosylcobalamin.

electron from  $\text{Co}^{2+}$ , together with a proton to generate the product. A reasonable chemical precedent for the radical rearrangement was put forward. This mechanism did not explain how the proton released in the first step and reacquired in the last step would be processed. However, the radical rearrangement foreshadowed the modern mechanism of isomerization. The lack of hydrogen exchange in  $\text{D}_2\text{O}$  cast doubt on the mechanism. Furthermore, the conventional wisdom of the time held that a free radical could not be an enzymatic intermediate because of the impossibility of an enzyme controlling the reactivity of such a reactive species. It is now known that the conventional wisdom was incorrect, and a mechanism for removing an electron and proton – a hydrogen atom – from an alkyl group through the action of coenzyme  $\text{B}_{12}$  is available.

The mechanism in **Figure 2(b)** was suggested based on excellent precedents for migrations of CO groups in organometallic chemistry.<sup>19</sup> In this mechanism, the isomerization of methylmalonyl-CoA could be explained by binding the acyl-CoA moiety to the lower axial position of adenosylcobalamin in place of the dimethylbenzimidazole. Subsequent insertion of the acyl-C=O group as CO into the macrocyclic ring between cobalt and a nitrogen ligand led to a new Co-alkyl bond, which could undergo rearrangement and reinsertion of CO to form the product. In this mechanism, the Co-C5 bond to the adenosyl moiety either did not react or played a secondary role, and the coenzyme A (CoA) had to be retained in the active site, a plausible rationale.

In another mechanistic principle put forward, adenosylcobalamin served as a biological Grignard reagent, wherein the linkage between the 5'-deoxyadenosyl moiety and  $\text{Co}^{3+}$  was viewed as essentially ionic, and the nucleoside functioned as the 5'-carbanion.<sup>20</sup> The diamagnetic character of adenosylcobalamin was cited in support of the ionic nature of the interaction between cobalt and C5' of the nucleoside. The mechanism in **Figure 2(c)** was suggested for the action of coenzyme  $\text{B}_{12}$  in the reaction of MCM. The first chemical step was pictured as a ligand exchange between the methyl group of methylmalonyl-CoA and the 5'-deoxyadenosyl carbanion, with transfer of a proton to give 5'-deoxyadenosine and the  $\text{Co}^{3+}$ -stabilized homoenolate of methylmalonyl-CoA. Precedent for a homoenolate intermediate in a cyclic molecule was cited. The carbanionic homoenolate could undergo an internal isomerization to the C3-enolate of succinyl-CoA, and ligand exchange with 5'-deoxyadenosine regenerated coenzyme  $\text{B}_{12}$  with the formation of succinyl-CoA.

The mechanisms in **Figure 2** stimulated interest and served as working hypotheses; however, none of them survived for very long as information about the actions of enzymes became available. Spectrophotometric and chemical observations of the interactions of propane-1,2-diol with the suicide inactivator glycolaldehyde indicated that the Co-C5' bond in adenosylcobalamin was cleaved in the course of catalysis.<sup>21,22</sup> This ruled out the mechanistic principle in **Figure 2(b)**, at least for DDH.

The elegance of the macrocyclic structure of cobalamin, with its extensive stereochemistry linked to cobalt chemistry, inspired researchers to look for a direct role of the cobamide ring in catalysis. Despite extensive experimentation, no evidence for a chemical interaction between a substrate intermediate and cobalt in coenzyme  $\text{B}_{12}$  has ever been found, essentially excluding the carbonyl migration and Grignard mechanisms in **Figures 2(b)** and **2(c)**. The only evidence of electron transfer between a substrate and cobalt in coenzyme  $\text{B}_{12}$  led to suicide inactivation,<sup>23</sup> inconsistent with the mechanistic principle in **Figure 2(a)**. The radical intermediates in **Figure 2(a)** and the radical isomerization mechanism are sustained by recent experiments.<sup>24,25</sup> However, the mechanism of radical formation has not been found to involve proton-coupled electron transfer.

### 7.14.2.2 Methylcobalamin

Early work in fractionated extracts of both *Escherichia coli* and liver indicated the participation of a vitamin  $\text{B}_{12}$ -containing protein in the reaction of homocysteine with 5-methyltetrahydrofolate to form methionine and tetrahydrofolate.<sup>26–30</sup> The reaction was stimulated by SAM, although SAM was not the stoichiometric methyl donor. Methylcobalamin was not the primary methyl donor; however, it could serve as a methyl donor in the absence of 5-methyltetrahydrofolate. It was suggested that methylcobalamin could be a catalytic intermediate in methyl transfer from 5-methyltetrahydrofolate. The role of SAM remained obscure until recent years, when it was found to preserve enzyme activity by maintaining cobalamin as methylcobalamin in the  $\text{B}_{12}$ -protein.

*Escherichia coli* was found to have two enzymes that produced methionine from homocysteine and 5-methyltetrahydrofolate, one of which required vitamin  $\text{B}_{12}$  and SAM and another that did not require either coenzyme. The  $\text{B}_{12}$ -requiring enzyme was found in a methionine auxotroph, in which vitamin  $\text{B}_{12}$  relieved the growth requirement for methionine. Most strains of *E. coli* did not require either methionine or vitamin  $\text{B}_{12}$  for

growth, and the involvement of B<sub>12</sub> was confusing for a time, until it was found that the B<sub>12</sub>-dependent system was produced only during growth in a medium containing vitamin B<sub>12</sub>. The B<sub>12</sub>-independent enzyme was not found in mammals.

### 7.14.3 Cobalamin Chemistry

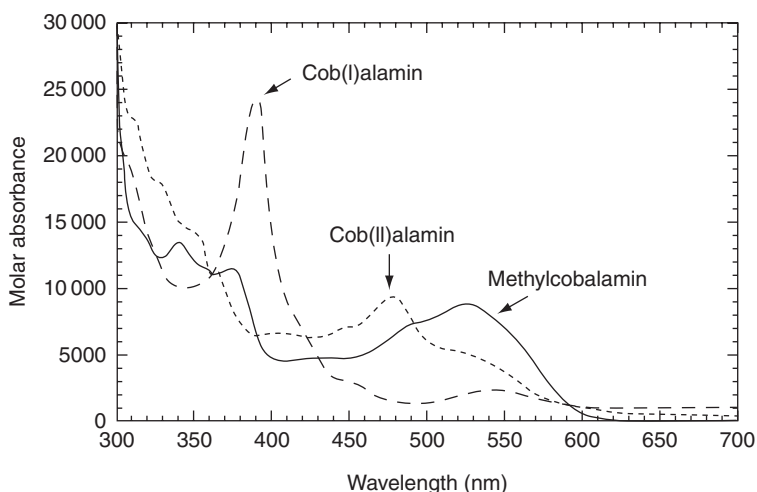
Salient features of cobalamin chemistry form the basis for biological reactions in which they participate. The cyano group in vitamin B<sub>12</sub> is introduced during purification. The vitamin is biosynthesized in bacteria as adenosylcobalamin. Its ready decomposition product is aquocobalamin, with H<sub>2</sub>O as the upper axial ligand. Upon ionization ( $pK_a = 7.8$ ), the upper axial ligand becomes HO<sup>-</sup> in hydroxocobalamin.<sup>31</sup> At low pH, the lower axial ligand becomes protonated ( $pK_a = 2.9$ ) and undergoes dissociation to the form known as 'base-off'. In bacteria and animals, aquo- or hydroxocobalamin can be adenylylated by ATP through the action of adenosylcobalamin synthetases, or they can be methylated to methylcobalamin at the active sites of cobalamin-dependent methyltransferases.

#### 7.14.3.1 Oxidation States

Three oxidation states of cobalt in free cobalamins are 3+, 2+, and 1+. Cob(III)alamins (3+), including alkylcobalamins, are red in color. Historically, aquocobalamin is known as vitamin B<sub>12a</sub> and hydroxocobalamin as vitamin B<sub>12b</sub>. Cob(II)alamin (2+) is yellow, lacks an upper axial ligand, and is historically known as vitamin B<sub>12r</sub>, for reduced cobalamin. Cob(I)alamin (1+) is gray-green, lacks an upper axial ligand to cobalt, and is historically known as vitamin B<sub>12s</sub>. The absorption spectra of methylcob(III)alamin, cob(II)alamin, and cob(I)alamin are shown in **Figure 3**.

#### 7.14.3.2 Chemical Reactions

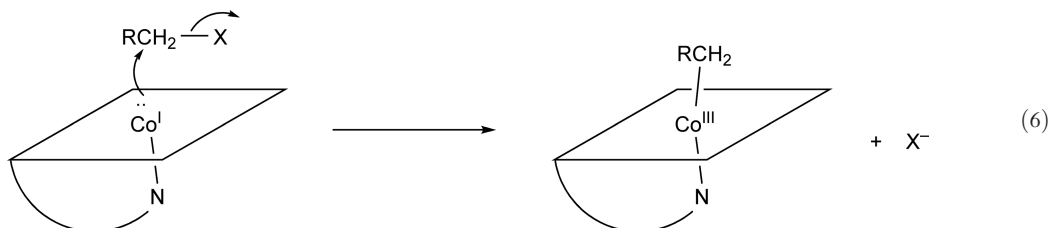
Cob(III)alamins undergo ligand exchange reactions typical of cobalt(III) complexes, most of which are not directly relevant to biological function. One-electron reduction to cob(II)alamin and two-electron reduction to cob(I)alamin are biologically relevant. The standard reduction potentials are +200 and -610 mV, respectively, for the base-on Cob(III)/Cob(II) and Cob(II)/Cob(I) couples, referenced to the normal hydrogen electrode.<sup>31,32</sup>



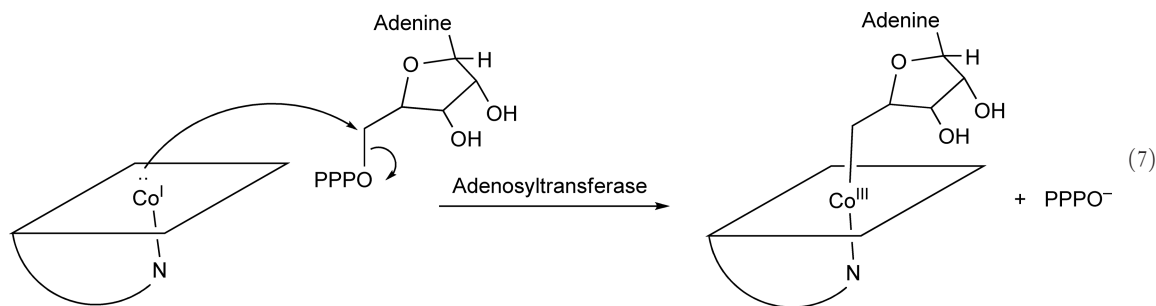
**Figure 3** Absorption spectra of cobalamins in three oxidation states. The spectra are of (—) methylcob(III)alamin, (---) cob(I)alamin, and (- - -) cob(II)alamin. Reproduced from R. G. Matthews, Cobalamin-Dependent Methionine Synthase. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; Wiley-Interscience: New York, 1999; pp 681–706, **Figure 5**.

Cob(II)alamin is an intermediate in many adenosylcobalamin-dependent enzymatic reactions. The Co–C5' bond to the adenosyl group in **Figure 1** undergoes reversible homolytic scission to cob(II)alamin and the 5'-deoxyadenosyl radical as a step in the catalytic process.

Cob(I)alamin is a powerful nucleophile, 40 000 times as reactive as the thiolate anion.<sup>33</sup> Its intrinsic chemical reactivity as a nucleophile makes it an important species in the nonenzymatic synthesis of alkylcobalamins. Cob(I)alamin is easily alkylated at ambient temperatures by primary alkyl halides, alkenes, and alkynes. Secondary and tertiary alkyl halides either do not react or form unstable products because of steric crowding with the macrocyclic cobamide. Nonenzymatic alkylation by an alkyl halide is illustrated in Equation (6), where X is Cl, Br, or I.



The nucleophilic reactivity of cob(I)alamin is also critically important in the biosynthesis of adenosylcobalamin and methylcobalamin. Adenosylcobalamin synthetases catalyze the 5'-adenosylation of cob(I)alamin by ATP according to Equation (7), where PPPO is triphosphate.



### 7.14.3.3 Reactions of Alkylcobalamins

Physicochemical properties of cobalt–carbon (Co–C) bonds in alkylcobalamins are subjects of continuing investigations dating from the discovery of the crystal structure of adenosylcobalamin.<sup>1</sup> Issues of interest include the strength and chemical reactivities of Co–C bonds, both of which vary with the structures of the alkyl groups. Strengths of Co–C bonds in alkylcobalamins listed in **Table 1** as bond dissociation enthalpies  $\Delta H$ , referring to homolytic scission to cob(II)alamin and an alkyl radical, range from 39 to 24 kcal mol<sup>-1</sup>, the strongest being methylcobalamin and the weakest being 3',4'-anhydroadenosylcobalamin, in which the 5'-deoxyadenosyl moiety

**Table 1** Bond dissociation energies for alkylcobalamins

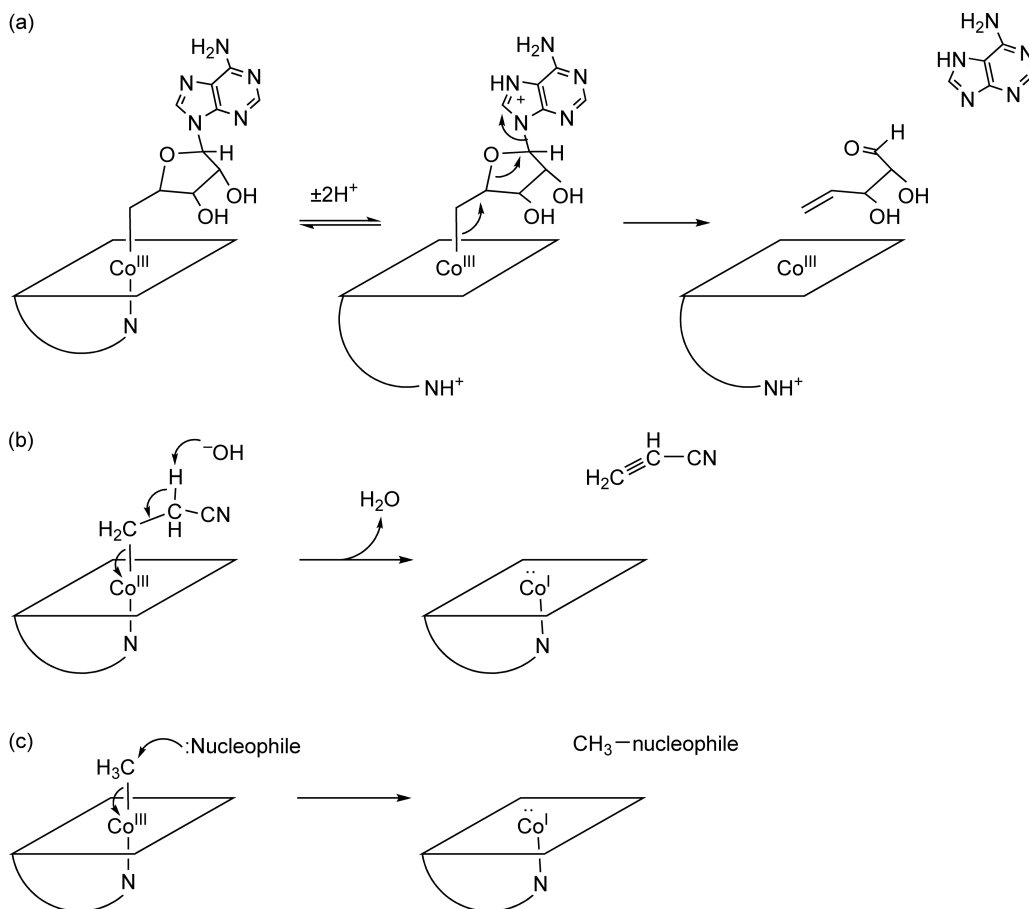
Alkylcobalamin	$\Delta H$ (kcal mol <sup>-1</sup> )	Reference(s)
Methylcobalamin	36 (hexacoordinate, base-on)	34
	39	35
	42 (pentacoordinate, base-off)	35
Adenosylcobalamin	31	35, 36
2',5'-Dideoxyadenosylcobalamin	32	37
Neopentylcobalamin	28	38
Benzylcobalamin	27	39
3',4'-Anhydroadenosylcobalamin	24	40



is dehydrated between C3' and C4'.<sup>40</sup> Homolytic scission of the Co–C5' bond in adenosylcobalamin is regarded as the first chemical step in the actions of adenosylcobalamin-dependent enzymes.<sup>24</sup>

Bond dissociation energies can be determined through measurements of the kinetics of thermal cleavage as a function of temperature<sup>41,42</sup> or by laser-induced photoacoustic calorimetry.<sup>35,37</sup> The weakness of the Co–C bond in isopentylcobalamin (Table 1) is attributed to steric strain in the ground state, owing to the steric requirements of the methyl groups in the isopentyl moiety, and to the stability of the tertiary isopentyl radical formed upon homolytic scission of the Co–C bond. The weakness of the Co–C bond in 3',4'-anhydroadenosyl cobalamin (Table 1) is attributed to the stability of the allylic radical 3',4'-anhydroadenosine-5'-yl formed upon thermal cleavage. The stability of the latter radical allows it to be observed by electron paramagnetic resonance (EPR) as a transient intermediate in adenosylcobalamin-dependent enzymatic reactions, as well as in SAM-dependent lysine 2,3-aminomutase, a radical SAM enzyme.<sup>43,44</sup>

The Co–C bonds in alkylcobalamins are subject to heterolytic as well as homolytic cleavages. Adenosylcobalamin undergoes acid-catalyzed hydrolysis to aquocobalamin, adenine, and  $\Delta$ 4,5-*erythro*-dihydroxypenten-1-al, as shown in Figure 4(a).<sup>45</sup> The reaction is thought to proceed by protonation of the adenine ring and departure of the nucleoside products with the electrons of the Co–C5' bond, directly generating cob(III)alamin, which quickly captures a water molecule to form aquocobalamin. The biological role of this type of cleavage is not known.



**Figure 4** Mechanisms of Co–C bond cleavage and formation in alkylcobalamins. The Co–C bonds in alkylcobalamins are weak and subject to homolysis. However, they also undergo polar reactions under certain circumstances. (a) Adenosylcobalamin undergoes acid-catalyzed hydrolysis to cob(III)alamin, adenine, and 2,3-dihydroxypentenal. Similarly, acid hydrolysis cleaves hydroxyethylcobalamin to cob(III)alamin and ethylene. (b) Cyanoethylcobalamin undergoes base-catalyzed cleavage to cob(I)alamin and acrylonitrile. (c) Nucleophiles are methylated by methylcobalamin to form cob(I)alamin.

Heterolytic cleavages in the opposite sense, in which cob(I)alamin departs with the electrons of the Co–C bond, are also known, as shown in **Figures 4(b) and 4(c)**. In **Figure 4(b)**, 2-cyanoethylcobalamin is deprotonated by a base to expel cob(I)alamin and generate acrylonitrile in a reaction that might be concerted or stepwise through a carbanionic intermediate.<sup>46</sup> The reaction is written as concerted in **Figure 4(b)**. More commonly, and of biological significance, methylcobalamin is an effective methylating agent, in which a nucleophile displaces cob(I)alamin from the methyl group, as shown in **Figure 4(c)**. This process occurs in the actions of methionine synthase and methylcoenzyme M synthase, among other methyltransferase enzymes.<sup>47,48</sup>

Homolytic scission of the Co–C5' bond in adenosylcobalamin is brought about thermally at high temperatures (>80 °C), by photolysis, or by the actions of adenosylcobalamin-dependent enzymes. The fate of the 5'-deoxyadenosyl radical depends on the reaction conditions. In an enzymatic site, the radical most often abstracts a hydrogen atom from a substrate to initiate a radical rearrangement process in the substrate. In the absence of a substrate or/and enzyme, the 5'-deoxyadenosyl radical faces alternative fates. In anoxic conditions, photolysis of adenosylcobalamin to the transient 5'-deoxyadenosyl radical ends in cyclization to 5',8-*cyclo*-adenosine, as shown in **Figure 5**, in what must be a multistep process.<sup>49</sup> In the presence of oxygen gas, the 5'-deoxyadenosyl radical reacts with oxygen to form adenosine-5'-aldehyde.<sup>50</sup>

The mechanism of aldehyde formation proceeds in steps that involve cob(II)alamin and cob(III)alamin in significant ways.<sup>51</sup> The initial reaction of the 5'-deoxyadenosyl radical with molecular oxygen leads directly to the 5'-peroxyadenosyl radical (**Figure 5**), which accepts an electron from cob(II)alamin to form cob(III)alamin and 5'-peroxyadenosine. Ligation of cob(III)alamin by 5'-peroxyadenosine leads to 5'-peroxyadenosylcobalamin, which decomposes to adenosine-5-aldehyde and aquocob(III)alamin. Thus, cob(III)alamin efficiently catalyzes the transformation of 5'-peroxyadenosine into adenosine-5'-aldehyde. In the absence of cob(III)alamin, 5'-peroxyadenosine undergoes spontaneous transformation into adenosine-5-aldehyde at a very slow rate, but the presence of aquocob(III)alamin increases the rate by more than 200-fold.<sup>51</sup>

## 7.14.4 Adenosylcobalamin in Enzymatic Catalysis

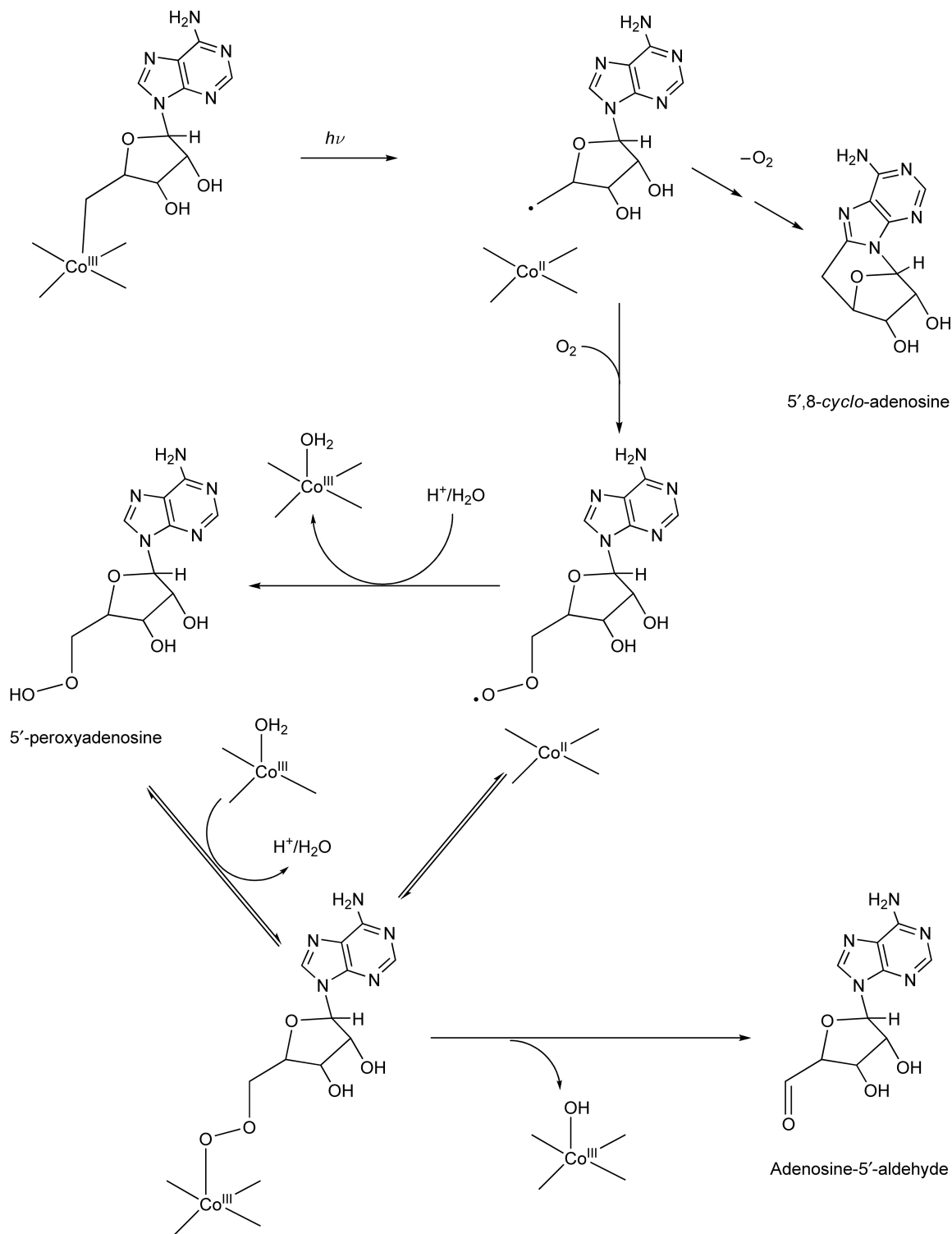
Two main classes of adenosylcobalamin-activated enzymes function by facilitating the homolytic scission of the Co–C5' to cob(II)alamin and the 5'-deoxyadenosyl radical. The resultant 5'-deoxyadenosyl radical initiates catalysis by abstraction of a hydrogen atom, either from a substrate in the case of the class of enzymes that catalyze radical isomerizations, or by abstraction of a hydrogen atom from Cys408-β-SH in the active site of ribonucleotide reductase II. The resultant enzymatic thiyl radical initiates the reduction mechanism by abstraction of a hydrogen atom from the ribonucleotide substrate. We shall begin with the isomerization/elimination reactions of adenosylcobalamin.

### 7.14.4.1 Diol Dehydrases

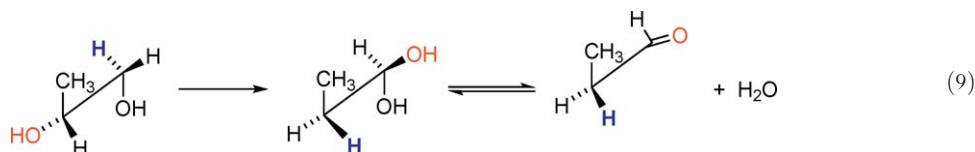
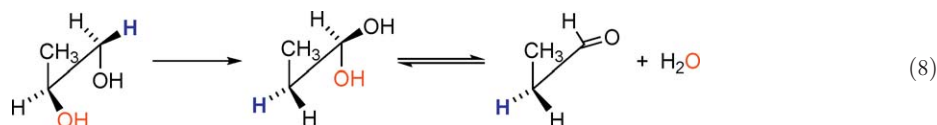
Growth of certain bacteria in media containing propane-1,2-diol induces the production of DDH, which catalyzes the production of propionaldehyde from propane-1,2-diol.<sup>5</sup> The purified enzyme, originally from *Klebsiella pneumoniae* (formerly *Aerobacter aerogenes*) ATCC 8724, is very active and also accepts ethylene glycol as a substrate, leading to acetaldehyde. Growth of *Klebsiella* or *Lactobacillus* in the presence of glycerol induces glycerol dehydrase, which catalyzes the production of 3-hydroxypropionaldehyde.<sup>6,7</sup> These transformations are difficult in non-enzymatic conditions but not impossible in principle. The strong acid conditions of the pinacol rearrangement could lead to the dehydration of vicinal diols to the corresponding aldehydes. However, strong acid is incompatible with biological conditions. Both enzymes require a cobalamin coenzyme and a monovalent cation such as NH<sub>4</sub><sup>+</sup> or K<sup>+</sup>. Because of the more intensive functional analysis of DDH, this enzyme will be the focus of discussion.

#### 7.14.4.1.1 Stereochemistry

DDH acts on both (*R*)- and (*S*)-propane-1,2-diol with a slight preference for the (*S*)-enantiomer. Despite the apparent lack of stereospecificity for enantiomers, finer details of stereochemistry, employing <sup>2</sup>H- and <sup>18</sup>O-labeling, reveal impressive stereochemistry with respect to H- and O-transfer, as illustrated in Equations (8) and (9).



**Figure 5** Fates of adenosylcobalamin upon homolytic cleavage of the Co-C bond. In the absence of O<sub>2</sub>, the photolytic cleavage of the Co-C bond leads to 5',8-cyclo-adenosine, shown in the upper right. This must proceed by a multistep radical and radical quenching mechanism that has not yet been determined. The Co-C bond is subject to thermal or photolytic cleavages to cob(II)alamin and the 5'-deoxyadenosyl radical. In the absence of molecular oxygen, the 5'-deoxyadenosyl radical undergoes cyclization to 5',8-cyclo-adenosine (top). In the presence of molecular oxygen, the radical and cob(II)alamin react to form 5'-peroxyadenosine and cob(II)alamin (center). Cob(III)alamin catalyzes the isomerization of 5'-peroxyadenosine into adenosine-5'-aldehyde, presumably through formation of 5'-peroxyadenosylcobalamin as an intermediate (bottom).



Thus, the 1-*pro-R* hydrogen in (*S*)-propane-1,2-diol migrates to the 2-*pro-R* position in propionaldehyde with *inversion* of configuration at C2, as in Equation (8). And the 1-*pro-S* hydrogen in (*R*)-propane-1,2-diol migrates to the 2-*pro-S* position of propionaldehyde with *inversion* of configuration, as in Equation (9).<sup>52–54</sup> Moreover, the C2 oxygens of both (*R*)- and (*S*)-propane-1,2-diols are transferred, to C1 of propionaldehyde derived from (*S*-propane)-1,2-diol and to water in the reaction of (*R*-propane)-1,2-diol.<sup>55</sup> These fates of <sup>18</sup>O are best explained by migration from C2 of the substrates to the *pro-R* or *pro-S* positions at C1 of propionaldehyde hydrate produced as a transient, enzyme-bound intermediate. Stereospecific dehydration of propionaldehyde hydrate at the active site leads to retention of <sup>18</sup>O in the reaction of (*R*)-[2-<sup>18</sup>O]propane-1,2-diol and loss of <sup>18</sup>O in the reaction of the enantiomer. The stereochemistry of <sup>18</sup>O-transfer and elimination proved the existence of propionaldehyde hydrate as an intermediate and validated the nature of the reaction of DDH as an isomerization following the pattern of Equation (5).

#### 7.14.4.1.2 Suicide inactivation

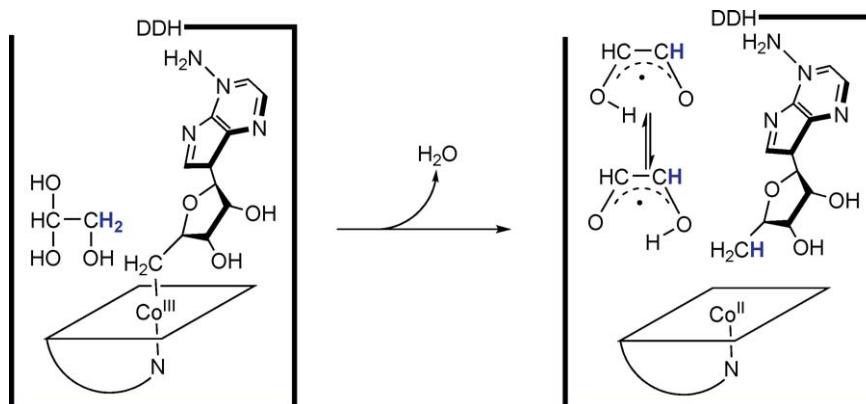
The early observation that addition of propane-1,2-diol to the complex of DDH and adenosylcobalamin led to the transient formation of cob(II)alamin provided the first evidence for the direct chemical participation of coenzyme B<sub>12</sub> in catalysis.<sup>21</sup> In the same study, another reaction leading to cleavage of coenzyme B<sub>12</sub> by DDH was suicide inactivation by glycolaldehyde.

Hydrated glycolaldehyde is an analogue of the substrate propane-1,2-diol, and it reacts with the complex of DDH and adenosylcobalamin to form cob(II)alamin, 5'-deoxyadenosine, and the glycolaldehyde radical shown in Figure 6.<sup>22,56,57</sup> The reaction proceeds by direct hydrogen transfer from glycolaldehyde to 5-deoxyadenosine, as proven by reaction of [2-<sup>3</sup>H]glycolaldehyde to form 5'-deoxy-[<sup>3</sup>H]adenosine (P. A. Frey, unpublished results). The formation of cob(II)alamin and 5'-deoxyadenosine is observed spectrophotometrically and in <sup>14</sup>C-tracer experiments,<sup>21,22</sup> and the glycolaldehyde radical is observed by EPR.<sup>56,57</sup> The carbon atoms of glycolaldehyde as well as 5'-deoxyadenosine and cob(II)alamin remain tightly bound to the enzyme in the inactivated complex. The highly delocalized electron in the glycolaldehyde radical lends it such stability that it is neither able to react further nor to revert to glycolaldehyde and adenosylcobalamin. The resulting complex is inactive.

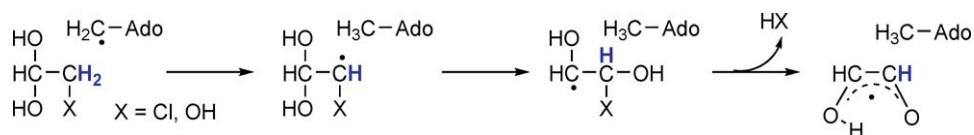
Chloroacetaldehyde also rapidly inactivates DDH with production of cob(II)alamin, 5'-deoxyadenosine, and the glycolaldehyde radical.<sup>58–61</sup> The two reactions vary in that the reaction of chloroacetaldehyde leads to the elimination of HCl, whereas elimination of water accompanies the reaction of glycolaldehyde. A mechanism for suicide inactivation by chloroacetaldehyde or glycolaldehyde is given in Figure 7. The first three steps are likely to be similar to the catalytic mechanism by which DDH acts on substrates.

#### 7.14.4.1.3 Hydrogen transfer by adenosylcobalamin

All adenosylcobalamin-dependent reactions entail hydrogen transfer, and the coenzyme facilitates this process. Reaction of [1-<sup>3</sup>H]propane-1,2-diol with DDH and adenosylcobalamin to produce [<sup>3</sup>H]propionaldehyde also leads to the incorporation of tritium into the coenzyme at C5' of the adenosyl moiety.<sup>62,63</sup> The pathway of tritium transfer from [1-<sup>3</sup>H]propane-1,2-diol to [<sup>3</sup>H]propionaldehyde can be either intermolecular or intramolecular; that is, tritium in a given molecule of [<sup>3</sup>H]propionaldehyde might be derived from either the precursor or a different molecule of substrate. This situation implies the formation of an intermediate with a pool of hydrogens subject to transfer into the product. Such a pool is the methyl group in 5'-deoxyadenosine. Reaction of propane-1,2-diol with DDH and [5'-<sup>3</sup>H]adenosylcobalamin leads to [<sup>3</sup>H]propionaldehyde and



**Figure 6** Suicide inactivation of DDH by glycolaldehyde. Products of the suicide inactivation of DDH by glycolaldehyde are cob(II)alamin, 5'-deoxyadenosine, and the glycolaldehyde radical.



**Figure 7** A mechanism for suicide inactivation of DDH. Both glycolaldehyde and fluoracetaldehyde react with DDH to form the inactive complex in **Figure 6**. In a mechanism shown here, in common for the two suicide inactivators, the 5'-deoxyadenosyl radical from homolytic scission of the Co–C5' bond abstracts a hydrogen atom from C1 of the hydrated inactivator. Isomerization by migration of the C2-hydroxyl group to C1 leads to an intermediate that eliminates either H<sub>2</sub>O or HCl to form the stable glycolaldehyde radical.

unlabeled adenosylcobalamin. Thus, C5' of adenosylcobalamin mediates hydrogen transfer in the reaction of DDH, and it does so with kinetic competence.<sup>64</sup>

Reaction of propionaldehyde with the complex of DDH and [5'-<sup>3</sup>H]adenosylcobalamin leads to [<sup>3</sup>H]propionaldehyde and unlabeled adenosylcobalamin, implying reversibility in the action of DDH. However, the reversal of the overall reaction cannot be detected, so that at least one step must be practically irreversible. Late steps in the mechanism must be reversible to account for tritium exchange between propionaldehyde and [5'-<sup>3</sup>H]adenosylcobalamin.

The deuterium kinetic isotope effects in the reactions of (*R*)- and (*S*)-[1-<sup>2</sup>H]propanediol of 12 and 10, respectively, prove that a hydrogen transfer step limits the rate.<sup>53</sup> The spectrophotometric observation of cob(II)alamin in the steady state indicates that the second hydrogen transfer, from 5'-deoxyadenosine to an intermediate, is rate limiting. The transfer of tritium from [1-<sup>3</sup>H]propane-1,2-diol proceeds with discrimination against tritium by a factor of 20, and tritium transfer from [5'-<sup>3</sup>H]adenosylcobalamin to propionaldehyde takes place with discrimination by a factor of 125 against tritium, corresponding to a tritium kinetic isotope effect of 28 ( $k_H/k_T$ ).<sup>64</sup> The rate at which tritium derived from [1-<sup>3</sup>H]propane-1,2-diol appears in the propionaldehyde corresponds to the rates at which it is transferred to adenosylcobalamin and then to propionaldehyde; that is, mediation of hydrogen transfer by adenosylcobalamin is both chemically and kinetically competent. The large deuterium and tritium kinetic isotope effects indicate quantum mechanical tunneling by hydrogen in this and other coenzyme B<sub>12</sub>-dependent reactions.

Mediation of hydrogen transfer by adenosylcobalamin is a property of all other isomerases that follow the pattern of Equation (5). The role of adenosylcobalamin in the action of class II ribonucleotide reductases is very closely related.

#### 7.14.4.1.4 A free radical intermediate

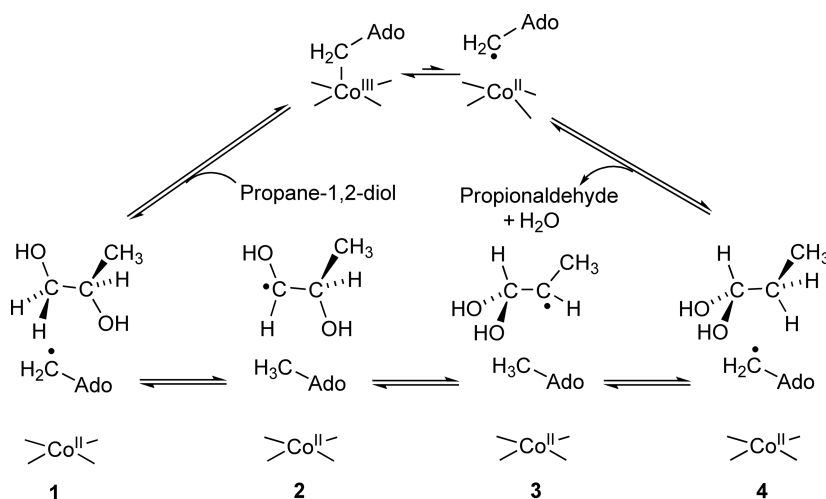
Observations of paramagnetic species in the reactions of ethanolamine ammonia-lyase (EAL), an adenosylcobalamin-dependent deamination,<sup>65</sup> and glycerol dehydrase<sup>66</sup> inspired searches for free radical intermediates. Such intermediates would be compatible with the observation of cob(II)alamin in the reaction of DDH.

Addition of propane-1,2-diol to the complex of DDH, adenosylcobalamin, and  $K^+$  led to the appearance of a paramagnetic signal in the EPR spectrum.<sup>61</sup> The signal was described in two parts: a feature at  $g=2.2$  attributed to cob(II)alamin and one at  $g=2.0$  attributed to a radical. The two paramagnetic species were broadened by spin coupling, and the distance between them was estimated at  $9 \text{ \AA}$ .<sup>67,68</sup> The radical species was shown to be kinetically competent to serve as an intermediate.<sup>69</sup> This distance was incompatible with earlier suggestions of covalent linkages between radical species and cob(II)alamin.<sup>61</sup>

The structure of the radical species contributing to the transient EPR signal is now known to be propane-1,2-diol-1-yl resulting from the abstraction of a hydrogen atom from C1 of the substrate, that is, the substrate-related radical in complex **2** of **Figure 8**. EPR spectroscopic proof of this structure is the broadening of the signal by  $^{13}\text{C}$  when  $[1-^{13}\text{C}]$ propane-1,2-diol is the substrate and narrowing of the signal by  $^2\text{H}$  when  $[1-^2\text{H}_2]$ propane-1,2-diol is the substrate.<sup>70,71</sup> The effect of  $\text{D}_2\text{O}$  on the spectrum proves that the radical contains a single solvent exchangeable proton, supporting the presence of OH as a substituent on C1, the locus of the unpaired electron.<sup>72</sup> Thus, available evidence does not indicate that the high field doublet is a radical anion or ketyl radical such as in other dehydratases.<sup>73,74</sup>

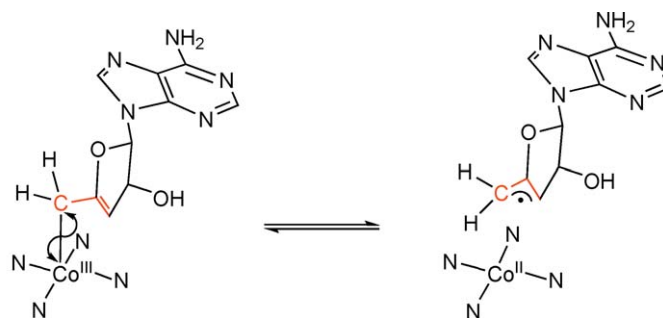
Support for the 5'-deoxyadenosyl radical as an intermediate is provided by the reaction of an analogue of adenosylcobalamin, 3',4'-anhydroadenosylcobalamin (anhydroadenosylcobalamin). The analogue functions as a coenzyme in place of adenosylcobalamin, albeit at a very slow rate.<sup>75</sup> In the steady state, freeze-quenched samples of DDH with anhydroadenosylcobalamin display radical triplet EPR spectra.<sup>44</sup> The effects of deuterium and  $^{13}\text{C}$ -labeling in the anhydroadenosyl moiety prove that the allylic anhydroadenosyl radical shown in **Figure 9** is a component of the triplet system and engaged in strong spin coupling with cob(II)alamin at a distance of  $3.7 \text{ \AA}$ . Co–C bond cleavage with this analogue is substrate-independent, supporting the preliminary equilibrium cleavage mechanism in **Figure 8**. The ready observation of the allylic analogue of the 5-deoxyadenosyl radical by EPR spectroscopy is attributed to the weakness of the Co–C5' bond in the analogue (see **Table 1**).

The mechanism in **Figure 8** accounts for available information. At the active site of DDH, adenosylcobalamin exists with a small equilibrium amount of the 5'-deoxyadenosyl radical in complex **1**. Abstraction of a hydrogen atom from C1 of the substrate by the nucleoside radical leads to 5'-deoxyadenosine and the substrate-related radical in complex **2**. Radical rearrangement to the product-related radical in complex **3** is followed by abstraction of a hydrogen atom from the methyl group to regenerate the Co–C bond and form propionaldehyde hydrate in complex **4**. Dehydration is followed by dissociation of propionaldehyde.



**Figure 8** An overall mechanism for the dehydration of propane-1,2-diol by DDH. The Co–C bond of adenosylcobalamin is homolytically cleaved in a preliminary, highly unfavorable equilibrium to the 5'-deoxyadenosyl radical and cob(II)alamin. Hydrogen atom abstraction from the substrate, isomerization of the substrate radical, and hydrogen abstraction from 5-deoxyadenosine by the product radical complete the rearrangement. Elimination of  $\text{H}_2\text{O}$  from propionaldehyde hydrate completes the dehydration process.





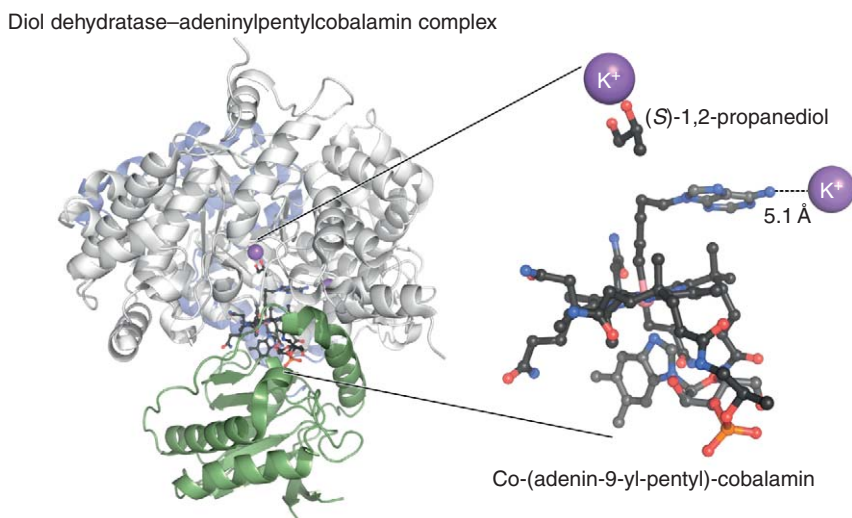
**Figure 9** Cleavage of 3',4'-anhydroadenosylcobalamin at the active site of DDH. Cleavage of the Co–C bond leads to cob(II)alamin and the allylic analogue of the 5'-deoxyadenosyl radical. The resulting triplet radical is observed and characterized by EPR spectroscopy.<sup>44</sup>

#### 7.14.4.1.5 Subunit composition and structure

DDH consists of three subunits  $\alpha$ ,  $\beta$ , and  $\gamma$  in a dimer of trimers  $(\alpha\beta\gamma)_2$ . The subunits  $\alpha$ ,  $\beta$ , and  $\gamma$  in *K. pneumoniae* are encoded by the genes *pddA*, *pddB*, and *pddC*, respectively.<sup>76</sup> The calculated subunit molecular weights are  $\alpha$ , 60 379;  $\beta$ , 24 401; and  $\gamma$ , 19 489.

Crystal structures of DDH show the three subunits and associated binding interfaces, the cobalamin-binding site, the binding site for 5'-deoxyadenosine, and the substrate-binding site. The large subunit  $\alpha$  is an  $(\alpha/\beta)_8$  barrel, a triose phosphate isomerase (TIM) barrel, and incorporates one side of the binding locus for cobalamin and 5'-deoxyadenosine. The smaller subunits  $\beta$  and  $\lambda$  bind at the base of subunit  $\alpha$ , with interfaces between each other and  $\alpha$ . Cobalamins bind at the interface of  $\alpha$ - and  $\gamma$ -subunits. The 5,6-dimethylbenzimidazole moiety of cobalamin remains as the lower axial ligand to cobalt. **Figure 10** shows a ribbon diagram of DDH with adeninylpentylcobalamin at the cobalamin-binding site and propane-1,2-diol at the substrate-binding site. The amino acid sequences of DDH and glycerol dehydrases are homologous, and the structures are very similar.<sup>70</sup>

In three salient features of available DDH structures, the distance between cobalt and propane-1,2-diol is 9 Å, one  $K^+$  is solvated by the vicinal hydroxyl groups of propane-1,2-diol, and another  $K^+$  binds very near the



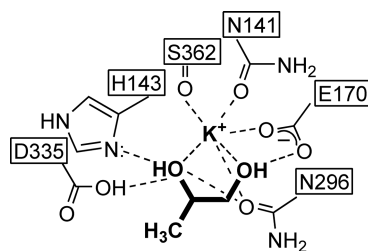
**Figure 10** A crystal structure of DDH with adeninylpentylcobalamin and (S)-propane-1,2-diol. On the left is a ribbon diagram of the three chains with adeninylpentylcobalamin and (S)-propane-1,2-diol. The  $\alpha$ -subunit is shown in gray,  $\beta$ -subunit in blue, and  $\gamma$ -subunit in green. On the right is a ball and stick expansion of adeninylpentylcobalamin and the substrate. Note the potassium ions, one in 2.6 Å association with (S)-propane-1,2-diol and the other 5.1 Å from the adenine ring. The illustration is generated from PDF 1eex.<sup>77</sup>

adenine ring of the 5'-deoxyadenosyl moiety (Figure 10). Both propane-1,2-diol and  $K^+$  are deeply buried within the TIM barrel ( $\alpha$ -subunit). The distance separating cobalt and propane-1,2-diol is in accord with EPR data showing weak spin coupling between cob(II)alamin and the substrate-based radical intermediate.<sup>67,68</sup> The spectroscopically measured distance between the glycolaldehyde radical and  $Co^{2+}$  in the glycolaldehyde inactivation complex is approximately 11 Å.<sup>56</sup> Correspondence of these distances indicates that the structure of DDH does not change dramatically in this regard in the transition from the Michaelis complex to the radical intermediate. The structure essentially rules out the possibility of direct covalent bonding between cobalamin and the substrate or a catalytic intermediate in the course of catalysis.

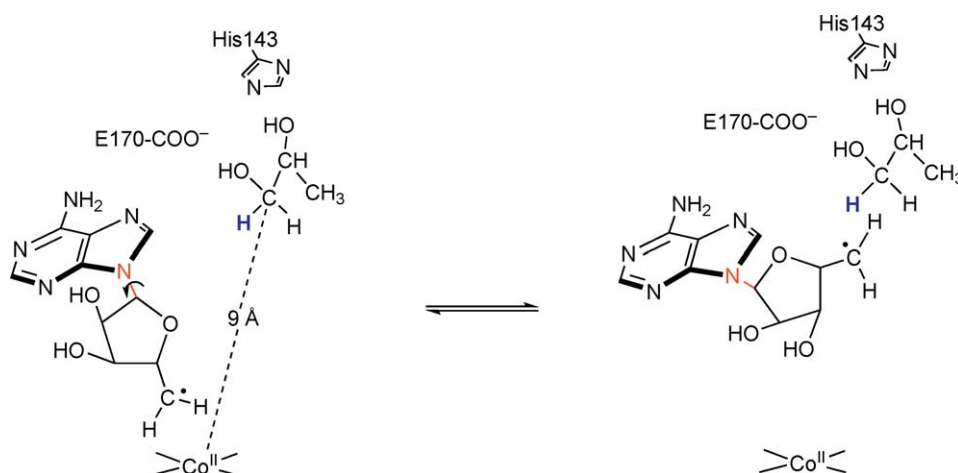
In the substrate-binding site illustrated in Figure 11, the  $\gamma$ -carboxyl group of Glu170 $\alpha$  is hydrogen bonded to C1(OH), and the  $\beta$ -carboxyl group of Asp335 $\alpha$  and N $\epsilon$ 2 of His143 $\alpha$  are hydrogen bonded to C2(OH). Mutation of any of these residues seriously decreases enzymatic activity.<sup>78</sup> These amino acid residues very likely participate in catalysis.

#### 7.14.4.1.6 Spanning the distance

A structural problem with the mechanism in Figure 8 is that C5' of the 5'-deoxyadenosyl radical must somehow come into contact with propane-1,2-diol, which is shown by the structure to be 9 Å from cobalt in cob(II)alamin. The illustration in Figure 12 offers a simple model by which C5' might traverse this distance upon cleavage from cobalt. Torsion about the *N*-glycosyl bond in the 5'-deoxyadenosyl radical would translocate the radical center just the distance required to contact C1 of propane-1,2-diol.<sup>70</sup> Then hydrogen



**Figure 11** Contacts to potassium ion at the substrate-binding site of DDH. The contacts shown are found in the crystal structure. The ionization states of Asp335, His143, and Glu170 are reasonable based on mutagenesis evidence, necessary function, and overall electrostatic neutrality, but they have not been chemically proven.



**Figure 12** A mechanism for translocation of the radical center in DDH. In DDH, the substrate-binding site is separated by 9 Å from Co in cobalamin. The radical center in the 5'-deoxyadenosyl radical arising from cleavage of adenosylcobalamin must be translocated to the substrate. Shown here is a mechanism by which this could occur through torsion about the *N*-glycosyl bond.

abstraction and isomerization can take place, and upon regeneration of the 5'-deoxyadenosyl radical, retro-rotation would restore it in position to re-form the Co-C5' bond.

#### 7.14.4.1.7 The roles of K<sup>+</sup>

Catalysis by DDH and suicide inactivation by either glycolaldehyde or chloroacetaldehyde require the presence of a monovalent cation such as K<sup>+</sup>. Both of the hydroxyl groups of propane-1,2-diol are within 2.6 Å of K<sup>+</sup> in the structure (Figure 10), and in the absence of the substrate K<sup>+</sup> binds two additional water molecules. The association of K<sup>+</sup> with the substrate, and with the adenine ring of the 5'-deoxyadenosyl moiety, together with the requirement for a monovalent cation in catalysis, focuses attention on the role of K<sup>+</sup>.

The cation requirement in the function of DDH can be satisfied by K<sup>+</sup>, NH<sub>4</sub><sup>+</sup>, and Tl<sup>+</sup>, which have similar though not identical ionic radii.<sup>79</sup> Rb<sup>+</sup>, Cs<sup>+</sup>, and Na<sup>+</sup> are less effective and Li<sup>+</sup> is practically ineffective. It seems that a monovalent cation similar in size to ammonium or potassium ion satisfies the cation requirement, whereas larger or smaller monovalent cations are less or not at all effective. The association of K<sup>+</sup> with propane-1,2-diol in the structure of DDH raises the question whether K<sup>+</sup> participates directly in catalysis, and if so in what capacity.

Suggestions that K<sup>+</sup> functions as a Lewis acid in the migration of C1(OH) to C2 have been under consideration.<sup>80,81</sup> Computational modeling indicated a modest lowering by K<sup>+</sup> of activation energy for OH migration.<sup>70</sup> To examine the question whether the monovalent cation remained bound to radical intermediates in catalysis, EPR experiments were undertaken to determine whether a magnetic interaction between Tl<sup>+</sup> and a radical at the active site could be detected. The magnetic nuclei of <sup>205,203</sup>Tl<sup>+</sup> might be expected to engage in nuclear hyperfine coupling with a radical ligated at 2.5–2.6 Å, the distance found in the association of K<sup>+</sup> to propane-1,2-diol. However, no interaction could be detected in the EPR spectra of either the glycolaldehyde radical or the substrate-based radical at the active site of DDH when Tl<sup>+</sup> was substituted for K<sup>+</sup>.<sup>58,72</sup> These negative results did not rule out a weak interaction, nor did they implicate the monovalent cation in catalysis of the migration of C1(OH).

Early experiments implicated the transformation of adenosylcobalamin into cob(II)alamin upon combining DDH with a substrate or suicide inactivator.<sup>21,22,56,59–61</sup> It has been proposed that the binding of propane-1,2-diol to DDH induces a conformational change that weakens the Co-C5' bond.<sup>70</sup> However, cleavage of the Co-C5' bond in 3',4'-anhydroadenosylcobalamin upon binding to DDH proceeded in both the presence and absence of the substrate and was attributed to the weakness of the Co-C5' bond in this analogue of adenosylcobalamin.<sup>44</sup> The latter experiments did not sustain the hypothesis that binding of propane-1,2-diol weakened the Co-C5' bond, although they did not rule out such a role. In all the foregoing experiments, K<sup>+</sup> was present and required, and its specific role was never explained.

One conceivable catalytic role of a monovalent cation is to weaken the Co-C5' bond in adenosylcobalamin upon binding to DDH. Catalysis requires that the Co-C5' bond be weakened in order to produce enough 5'-deoxyadenosyl radical to lead to hydrogen abstraction at a rate that is consistent with the observed value of  $k_{\text{cat}}$ , which is 370 s<sup>-1</sup>.<sup>70</sup> Given the enthalpy of 31 kcal mol<sup>-1</sup> for cleavage of Co-C5' bond, and assuming a modest entropy change for a simple bond dissociation without separation, the equilibrium concentration of 5'-deoxyadenosyl radical in 1 μmol l<sup>-1</sup> DDH would have to react with a rate constant >10<sup>16</sup> s<sup>-1</sup> to account for the observed rate. This rate constant exceeds bond vibrational frequencies by several orders of magnitude, so that the Co-C5' bond must be weakened. Perhaps the monovalent cation plays a role in weakening the Co-C5' bond of adenosylcobalamin in DDH.

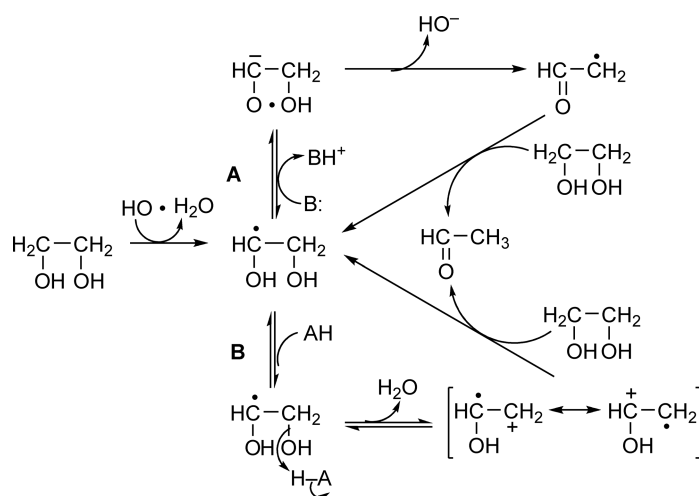
To test whether K<sup>+</sup> binding to DDH weakens the Co-C5' bond, advantage was taken of the oxygen sensitivity of the complex of adenosylcobalamin with DDH. The activated complex is subject to inactivation by oxygen in air in the absence of the substrate in a process leading to the cleavage of adenosylcobalamin to cob(III)alamin.<sup>22</sup> The 5'-deoxyadenosyl moiety of the coenzyme is transformed into 5'-peroxyadenosine in this process.<sup>79</sup> This is the product expected from oxygen quenching of the 5'-deoxyadenosyl radical. Thus, oxygen inactivation of DDH results from homolytic cleavage of adenosylcobalamin at the active site and capture of the 5'-deoxyadenosyl radical by molecular oxygen. In the presence of the substrate, molecular oxygen cannot compete for the 5'-deoxyadenosyl radical, and catalysis occurs instead of inactivation. Experiments show that DDH binds adenosylcobalamin in the absence of a monovalent cation, albeit with somewhat decreased affinity. Furthermore, the complex of adenosylcobalamin and DDH in the absence of both K<sup>+</sup> and propane-1,2-diol is stable toward molecular oxygen.<sup>79</sup> Thus, K<sup>+</sup> or a monovalent cation of similar size is essential for the

oxygen-dependent inactivation of DDH and the cleavage of adenosylcobalamin. Therefore, the presence of  $K^+$  is essential for weakening the Co–C5' bond in DDH. The detailed mechanism by which  $K^+$  weakens the Co–C5' bond is not known, but one binding site for  $K^+$  is adjacent to the adenine ring of the 5'-deoxyadenosyl moiety and might make an electrostatic contribution to the structure of the active enzyme–coenzyme complex (Figure 10).

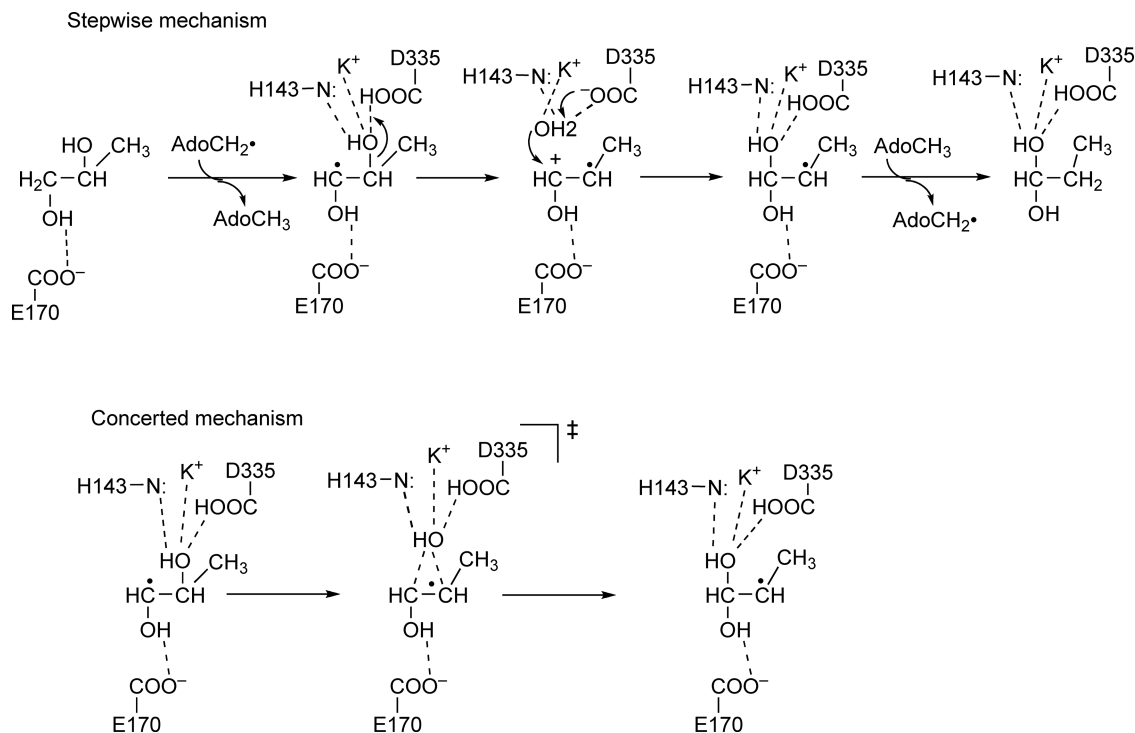
#### 7.14.4.1.8 Isomerization mechanism

The isomerization step in the mechanism of Figure 8 links two metastable radical intermediates, propane-1,2-diol-1-yl and propane-1,1-diol-2-yl. The isomerization mechanism remains unclear. Chemical models give assurance that the hydroxyl group  $\alpha$ - to the radical center can be eliminated. Chemically generated hydroxyl radicals react with ethylene glycol by mechanisms shown in Figure 13 to produce acetaldehyde. Hydrogen abstraction from ethylene glycol by  $HO\cdot$  gives the radical ethane-1,2-diol-1-yl, which can react by base catalysis in mechanism B to form a ketyl radical.<sup>82</sup> Elimination of  $HO^-$  gives acetaldehyde-2-yl, which can be recycled with acetaldehyde formation as shown. In the acid-catalyzed reaction, mechanism A, an acid catalyzes the dehydration of ethane-1,2-diol-1-yl to the resonance-delocalized radical cation, which can be recycled to acetaldehyde as shown.<sup>83,84</sup> These models show that a glycol radical can eliminate water under acid catalysis or hydroxide with base catalysis to form precursors of acetaldehyde.

At the enzymatic site, definitive evidence for the formation of propane-1,2-diol-1-yl supports the acid-catalyzed mechanism B through an intermediate oxycation radical in Figure 13 for elimination of water.<sup>71,72</sup> A base-catalyzed mechanism through an intermediate ketyl radical (A in Figure 13) is not ruled out. Experiments employing site-directed and kinetic analysis give important information. Asp335 and Glu170 are essential for activity.<sup>70,78</sup> Position 143 $\alpha$ , with histidine in the wild-type enzyme, is remarkably tolerant of mutation. While H143 $\alpha$ A-DDH is inactive, H143 $\alpha$ K- and H143 $\alpha$ Q-DDH are active, displaying 8 and 34% of wild-type activity, respectively.<sup>70,85</sup> The 34% activity upon substitution with the nonbasic glutamine is especially significant, in that glutamine is a conservative substitution for the neutral, conjugate base form of histidine and not of the protonated imidazolium form. Available experimental evidence supports the stepwise mechanism of isomerization in Figure 14. In this mechanism, water is eliminated from the spectroscopically observed substrate-related radical, in a step that is catalyzed by Asp335 $\alpha$ , assisted by hydrogen bonding from His143 $\alpha$ . The resulting water molecule is held in place by hydrogen bonding and can undergo readdition to either C2 or C1 of the resonance-delocalized oxycarbocationic intermediate. The process may be facilitated by the potassium ion solvated by propane-1,2-diol at the active site.<sup>70,86</sup>



**Figure 13** Chemical models for the radical isomerization process in the action of DDH. At the top is a base-catalyzed mechanism by way of a ketyl radical as the key intermediate. At the bottom is an acid-catalyzed mechanism by way of a radical cationic intermediate.



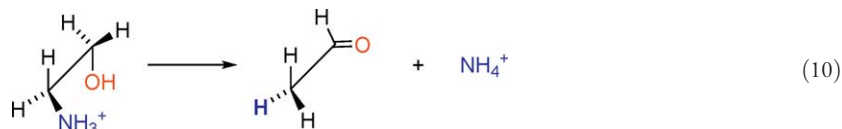
**Figure 14** Mechanisms for the migration of C1(OH) of propane-1,2-diol in DDH. Stepwise and concerted mechanisms both require acid catalysis for the migration of OH from C2 to C1 of propane-1,2-diol. In the stepwise mechanism, the elimination of OH from C1 is catalyzed by proton transfer to form a discrete water molecule and oxyanion radical intermediate. Addition of the water molecule to C1 with proton transfer completes the rearrangement. In the concerted mechanism, water is not a discrete intermediate but the migrating oxygen remains partially bonded to C1 and C2 as well as hydrogen bonded in a transition state. His143, Asp335, and potassium ion are in contact with the migrating OH group.

An interesting and potentially useful observation in the mutational work is that the substitution of leucine for His143 $\alpha$  decreased the activity to 0.2% of wild type.<sup>70</sup> This significant activity allowed the deuterium kinetic isotope effect for the reaction of [1-<sup>2</sup>H<sub>2</sub>]propane-1,2-diol to be measured. The value obtained turned out to be 2 for <sup>D</sup>*k*<sub>cat</sub>, about one-fifth to one-sixth of that for wild-type enzyme. Substitution of alanine gave about 1.5% activity and a deuterium kinetic isotope effect of 5–6. It appears that the hydrogen transfer is not solely rate limiting in these variants.

Computational efforts support a range of detailed mechanisms, all of which involve Lewis acid catalysis or/and hydrogen bonding.<sup>86–90</sup> Recent theoretical work supports a concerted mechanism of oxygen migration shown in **Figure 14**. The computational studies are burdened by the necessity to simulate the nonvacuous condensed medium of an active site. The concerted mechanism provides a satisfying rationale for the direct transfer of the hydroxyl group without equilibration with water in the medium, although this could also be accommodated in the stepwise mechanism through tight binding of a discretely formed water molecule. In any case, in the concerted mechanism, the distances separating the hydroxyl group in flight and C1 and C2 are only slightly less than the van der Waals contact distances, implying a concerted but very loose or dissociated transition state. Whether the condensed phase of a tightly packed active site is accurately simulated in computations remains to be proven.

#### 7.14.4.2 Ethanolamine Ammonia-Lyase

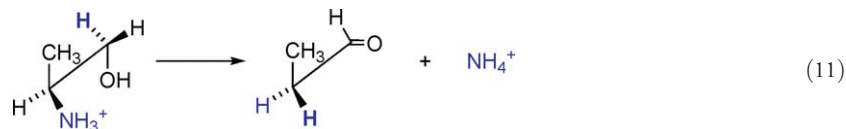
Growth of several species of bacteria induces the production of EAL, also known as ethanolamine deaminase, which catalyzes the adenosylcobalamin-dependent deamination of ethanolamine to acetaldehyde and ammonium ion according to Equation (10).



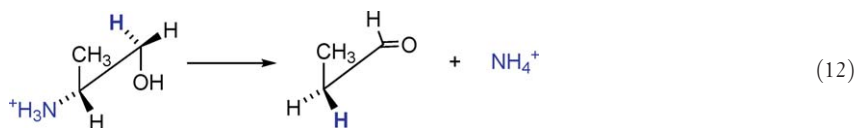
The reaction is analogous to that of DDH, with the substitution of  $\text{NH}_3^+$  in place of  $\text{OH}$  at C2 of the substrate and the elimination of  $\text{NH}_4^+$  instead of  $\text{H}_2\text{O}$ . Unlike DDH, the activity of EAL does not require  $\text{K}^+$ . Reaction of  $[1\text{-}^2\text{H}]$ ethanolamine proceeds with a deuterium kinetic isotope effect of 6.8, and deuterium transfer to C2 occurs without exchange with the medium.<sup>91</sup> Oxygen at C1 of the substrate is retained in the product, and adenosyl-C5' of adenosylcobalamin mediates hydrogen transfer from C1 of the substrate to C2 of the product as in DDH.<sup>92</sup>

#### 7.14.4.2.1 Stereochemistry

Hydrogen transfer to C2 of acetaldehyde by EAL is nonstereospecific, implying the formation of a torsionally symmetrized intermediate.<sup>93</sup> (*R*)- and (*S*)-2-aminopropanol are also substrates of EAL and are deaminated to propionaldehyde. These transformations might appear similar to the action of DDH on the enantiomers of propane-1,2-diol; however, the stereochemistry of hydrogen transfer by EAL is unlike that by DDH. In the reaction of (*S*)-2-aminopropanol, hydrogen at the 1-*pro-S* position is transferred to the 2-*pro-S* position in propionaldehyde, with *retention* of configuration at C2, as in Equation (11).



In the reaction of (*R*)-2-aminopropanol, hydrogen at the 1-*pro-S* position is transferred to the 2-*pro-S* position in propionaldehyde, with *inversion* of configuration at C2, as in Equation (12).<sup>94</sup>



This unexpected result can be explained if a planar intermediate arising from abstraction of hydrogen from C1 becomes oriented into a favored conformation before the transfer of hydrogen to C2. This could occur in a trigonal intermediate such as a free radical.

#### 7.14.4.2.2 Participation of coenzyme B<sub>12</sub> in hydrogen transfer

Experiments with  $[5'\text{-}^3\text{H}]$ adenosylcobalamin show that C5' of coenzyme B<sub>12</sub> mediates hydrogen transfer in deamination by EAL.<sup>92,95,96</sup> Reactions of ethanolamine, (*S*)-2-aminopropanol, or (*R*)-2-aminopropanol with EAL activated by  $[5'\text{-}^3\text{H}]$ adenosylcobalamin lead to tritiated acetaldehyde or propionaldehyde. (*R*)- and (*S*)-2-aminopropanol are both poor substrates relative to ethanolamine.

(*S*)-2-aminopropanol facilitates mechanistic analysis and reveals details that are masked in the reactions of ethanolamine or (*R*)-2-aminopropanol.<sup>95</sup> Tritium exchange data prove reversibility in the reaction of (*S*)-2-aminopropanol, whereas reaction of ethanolamine reveals no sign of reversibility. Reaction of (*S*)-2-aminopropanol with the complex of EAL and  $[5'\text{-}^3\text{H}]$ adenosylcobalamin produces both (*S*)-2-amino- $[^3\text{H}]$ propanol and  $[^3\text{H}]$ propionaldehyde, and reaction of the same complex with propionaldehyde in the presence of  $\text{NH}_4^+$  also produces both (*S*)-2-amino- $[^3\text{H}]$ propanol and  $[^3\text{H}]$ propionaldehyde. Appearance of tritium in 2-aminopropanol with propionaldehyde and  $\text{NH}_4^+$  proves reversibility. Partitioning of tritium between (*S*)-2-amino- $[^3\text{H}]$ propanol and  $[^3\text{H}]$ propionaldehyde favors 2-amino- $[^3\text{H}]$ propanol when (*S*)-2-aminopropanol initiates the exchange, whereas partitioning favors  $[^3\text{H}]$ propionaldehyde when propionaldehyde and  $\text{NH}_4^+$  initiate the exchange. The partitioning ratios differ by a factor of 6, and the difference implies the likelihood of at least two intermediates in which exchange of tritium takes place. These results prove to be revealing when interpreted in conjunction with  $^{15}\text{N}$ -kinetic isotope effects and the mechanism of the reaction.



Reversibility in the reaction of (*S*)-2-aminopropanol allowed the identification of 5'-deoxyadenosine as a transiently and reversibly formed intermediate. Reaction of this substrate led to both cob(II)alamin and 5-deoxyadenosine in the steady state. Experiments with [ $5'$ - $^2\text{H}_2$ ]adenosylcobalamin and (*S*)-2-amino[ $1$ - $^2\text{H}_2$ ]propanol produced 5'-deoxy- $[\text{}^2\text{H}_3]$ adenosine in the steady state, proving direct hydrogen transfer from the substrate to C5' of coenzyme B $_{12}$ .<sup>97</sup>

#### 7.14.4.2.3 Suicide inactivation

Like DDH, reaction of the complex of EAL and adenosylcobalamin with glycolaldehyde leads to the formation of an inactive complex containing cob(II)alamin, 5'-deoxyadenosine, and the glycolaldehyde radical.<sup>56</sup> The inactivation mechanism is likely to be similar to that of DDH, shown in **Figure 7**, where X = NH $_3^+$ .

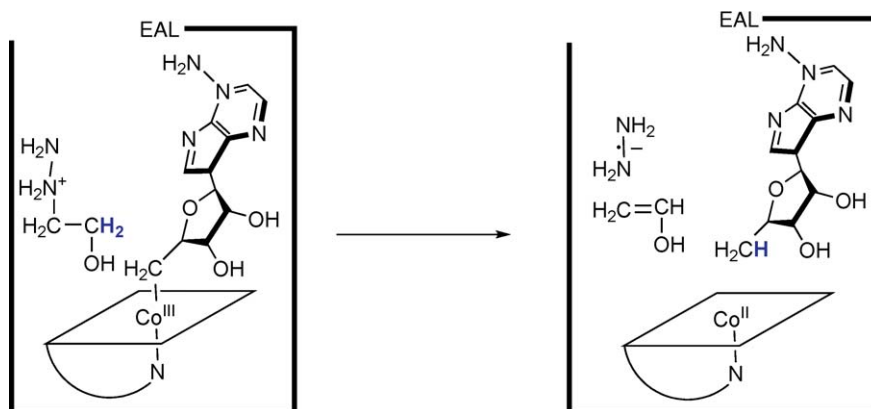
Suicide inactivation that might be unique to EAL would be the reaction with hydroxyethylhydrazine (HEH). The complex of EAL and coenzyme B $_{12}$  reacts rapidly with HEH to produce a complex of EAL with cob(II)alamin, 5'-deoxyadenosine, and the hydrazine radical cation shown in **Figure 15**.<sup>98</sup> Analysis of the EPR spectra for the effects of  $^{15}\text{N}$  in HEH and D $_2\text{O}$  as the solvent allowed the definitive identification of the radical cation in the inactive complex. EAL in the complex could be partially reactivated by precipitation in acidic conditions, implying that the protein portions of the enzyme might be intact, with the inactivation being due to blockage of the active site by cob(II)alamin.

Transient kinetics by spectrophotometry and rapid freeze-quench EPR showed the appearance of cob(II)alamin and the radical cation at the same rate, about  $3\text{ s}^{-1}$ .<sup>99</sup> Further studies of deuterium partitioning between [ $1$ - $^2\text{H}_2$ ]HEH and 5'-deoxyadenosine were consistent with direct and reversible deuterium transfer between HEH and the 5'-deoxyadenosyl radical. Reversibility was proven by the appearance of multiply deuterated 5'-deoxyadenosine at all stages of the reaction.

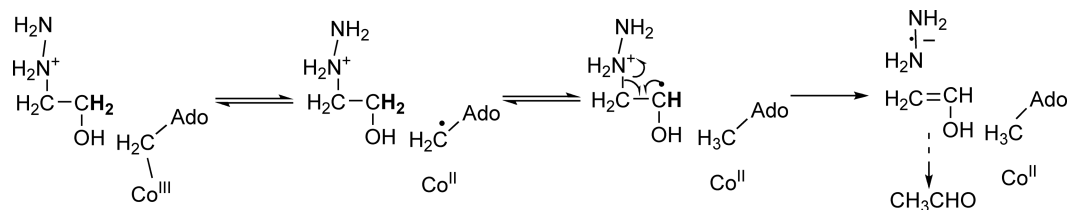
The mechanism in **Figure 16** explains the available information. Initial homolytic dissociation of the Co–C5 bond in coenzyme B $_{12}$  and abstraction of hydrogen from C1 of HEH by the 5'-deoxyadenosyl radical are thought to be similar to early steps in the catalytic deamination of a substrate by EAL. Elimination of the hydrazine radical cation in the last step in **Figure 16** is unlike the normal deamination process in that it involves homolytic cleavage of the C–N bond in HEH. This mode of cleavage appears to occur because of the intrinsic stability of the hydrazine radical cation.

#### 7.14.4.2.4 Radical intermediates

Cob(II)alamin and free radical intermediates are observed as transient intermediates in the reactions of EAL–coenzyme B $_{12}$  with ethanolamine, (*S*)-2-aminopropanol, and (*R*)-2-aminopropanol. Cob(II)alamin is



**Figure 15** Suicide inactivation of EAL by hydroxyethylhydrazide (HEH). The complex of EAL with adenosylcobalamin reacts with the substrate analogue HEH and is transformed into an inactive complex composed of EAL, cob(II)alamin, 5-deoxyadenosine, *enol*-acetaldehyde or acetaldehyde, and the hydrazine radical cation. The radical cation is so stable that further reaction does not occur.



**Figure 16** A mechanism for suicide inactivation of EAL by hydroxyethylhydrazide (HEH). HEH reacts initially as a substrate analogue through the first two steps. The HEH-related radical eliminates the hydrazine radical cation in a practically irreversible process. The first two steps are reversible, as shown in deuterium-partitioning experiments.

observed by stopped-flow spectrophotometry, and the free radicals are observed by rapid-mix freeze-quench continuous-wave EPR and pulsed EPR methods.

As in the reactions of all coenzyme  $\text{B}_{12}$ -dependent enzymes, the observable free radicals display spin–spin coupling with low-spin  $\text{Co}^{2+}$  in cob(II)alamin and are quantitatively characterized as diradicals or radical triplet species. As in DDH, this coupling is weak in EAL, and quantitative analysis indicates a separation of 9–11 Å between the free radical and  $\text{Co}^{2+}$ .<sup>100,101</sup> Isotope-edited EPR experiments employing  $^{13}\text{C}$ - or  $^2\text{H}$ -labeled (*S*)-2-aminopropanol as the substrate prove that the unpaired electron resides on C1 of the free radical; that is, it is the substrate-related radical.<sup>100,102,103</sup> This structure would arise from abstraction of a hydrogen atom from C1 by the 5'-deoxyadenosyl radical.

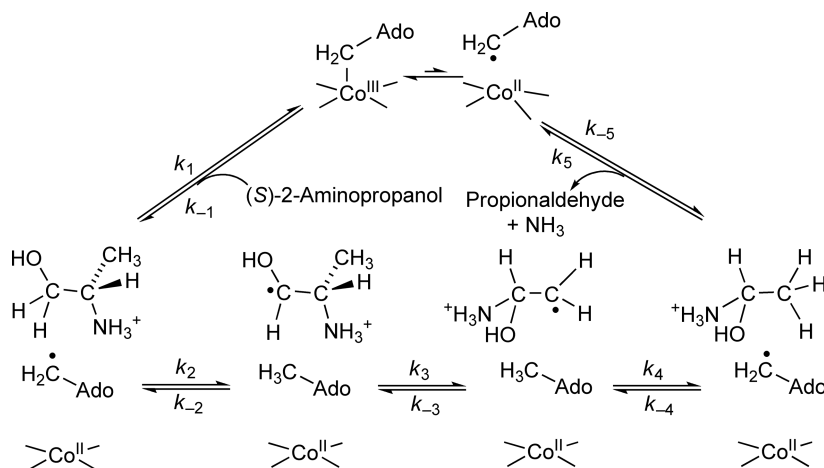
The structure of the free radical observed in the steady state of the reaction of ethanolamine was initially assigned as a product-related radical arising from initial abstraction of a hydrogen atom from C1 followed by rearrangement to the product-related radical 1-aminoethan-1-ol-2-yl.<sup>104</sup> It is now known to be the substrate-related radical 2-aminoethan-1-ol-1-yl, arising from abstraction of a hydrogen atom from C1 of the substrate.<sup>105,106</sup> This radical component of the triplet system has been analyzed by electron spin echo envelope modulation (ESEEM) spectroscopy and placed at 3.2 Å separation from the methyl group of 5'-deoxyadenosine.<sup>104,107</sup> The distance separating C1 of the substrate-related radical derived from (*S*)-2-aminopropanol was found by analogous methods to be 3.2 Å.<sup>102</sup> The results of  $^{13}\text{C}$ -ENDOR (electron nuclear double resonance) experiments indicated a separation of 3.4 Å.<sup>108</sup> These distances corresponded essentially to van der Waals contacts and indicated the importance of enzyme-binding interactions in holding the radical intermediate in contact with the methyl group of 5'-deoxyadenosine. Very close contacts would not only prevent the radical intermediate from escaping the active site, but also protect it from adventitious reactions with the protein and diffusible reducing or oxidizing agents. Such control of the radical intermediate could explain how it survives side reactions.<sup>109</sup>

#### 7.14.4.2.5 Structure

The three-dimensional structure of EAL is not available, but a structural model for the large subunit EutB has been derived based on amino acid sequence analogies with the cobalamin-binding subunits of DDH and glycerol dehydrase.<sup>110</sup> The model indicates an  $(\alpha/\beta)_8$  barrel similar to the structures of other adenosylcobalamin-dependent enzymes. The detailed model directs attention to Arg160, which overlaps the site of  $\text{K}^+$  binding to propane-1,2-diol in DDH and might explain the absence of a  $\text{K}^+$  requirement by EAL. Accordingly, mutation of R160 to isoleucine or glutamate leads to variants that do not assemble with other subunits. The variants R160A- and R160K-EAL assemble and can function. R160K-EAL displays 0.55% of wild-type activity. R160A-EAL is inactive but can be rescued by guanidinium ion to 2.3% of wild-type activity.<sup>111</sup>

#### 7.14.4.2.6 Reaction mechanism

The currently accepted chemical mechanism of action of EAL, shown in **Figure 17** for the reaction of (*S*)-2-aminopropanol, follows the pattern of other coenzyme  $\text{B}_{12}$ -dependent isomerases. The chemical mechanism begins by homolytic scission of the  $\text{Co}-\text{C}5'$  bond in adenosylcobalamin with the substrate bound at its site; translocation of the ribosyl moiety of the 5'-deoxyadenosyl radical by torsion about the *N*-ribosyl linkage to place the 5'-methylene radical in contact with C1 of the substrate, in analogy with DDH;<sup>70</sup> abstraction of C1-hydrogen from the substrate by the 5'-deoxyadenosyl radical to form 5'-deoxyadenosine and the

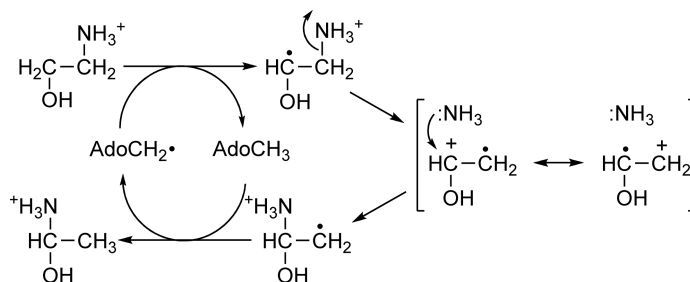


**Figure 17** A mechanism for the action of EAL on (*S*)-2-aminopropanol. Shown is a general mechanism applicable to all substrates and illustrated with (*S*)-2-aminopropanol because of available information about the relative values of rate constants,  $k_{-2} > k_3$  and  $k_4 > k_{-3}$ . The physicochemical basis for these relationships is explained in the text. While the overall mechanism appears similar to that for DDH, the details of the central isomerization ( $k_3$ ,  $k_{-3}$ ) are likely to be different (see **Figure 18**).

substrate-related radical; isomerization to the product-related radical; hydrogen abstraction from the methyl group of 5'-deoxyadenosine to form the aminal and regenerate coenzyme B<sub>12</sub>; concluding with elimination of ammonia from the aminal and release of the product propionaldehyde.

In the case of (*S*)-2-aminopropanol, all steps have been found to be reversible, as shown by the fact that exchange of tritium in [<sup>5'-<sup>3</sup>H</sup>]adenosylcobalamin bound to EAL with propionaldehyde is accompanied by the formation of (*S*)-2-amino-<sup>[3</sup>H]propanol as well as <sup>[3</sup>H]propionaldehyde.<sup>95</sup> The observation of preferential tritium transfer to residual substrate when (*S*)-2-aminopropanol was added to the complex of EAL and [<sup>5'-<sup>3</sup>H</sup>]adenosylcobalamin, and preferential transfer to the product when propionaldehyde is added to the complex, proves that more than one intermediate is engaged in tritium exchange with C5' of coenzyme B<sub>12</sub>. This was originally advanced as evidence compatible with the participation of the substrate alkylcobalamin as an intermediate.<sup>95</sup> However, the distance of the substrate-binding site from cobalt in cob(II)alamin precludes covalent bonding between them.

More recent experiments account for differential tritium exchange in the foregoing experiments in terms of the mechanism in **Figure 17**. The <sup>15</sup>N-kinetic isotope effects for the action of EAL on [<sup>15</sup>N]ethanolamine, (*R*)-[2-<sup>15</sup>N]aminopropanol, and (*S*)-[2-<sup>15</sup>N]aminopropanol are all normal, that is not inverse.<sup>112</sup> The values for [<sup>15</sup>N]ethanolamine and (*R*)-[2-<sup>15</sup>N]aminopropanol are small, 0.17 and 0.12%, respectively, consistent with the fact that hydrogen transfer is largely rate limiting for these substrates. The <sup>15</sup>N-kinetic isotope effect in the reaction of (*S*)-[2-<sup>15</sup>N]aminopropanol is 0.55%, nearly 5 times the value for (*R*)-[2-<sup>15</sup>N]aminopropanol. This is explained on the basis that hydrogen transfer is less rate limiting and isomerization more rate limiting for the reaction of (*S*)-[2-<sup>15</sup>N]aminopropanol than for the other substrates. This interpretation is reinforced by the fact that the steady-state intensity of the EPR signal in the reaction of (*S*)-2-aminopropanol is more intense than in the reactions of (*R*)-2-aminopropanol or ethanolamine.<sup>112</sup> These properties of (*S*)-2-aminopropanol as a substrate for EAL lead directly to the differential tritium exchange alluded to above in terms of the mechanism in **Figure 17**. Inasmuch as the isomerization step is significantly rate limiting for (*S*)-2-aminopropanol,  $k_3$  and  $k_{-3}$  in **Figure 17**, the product- and substrate-related intermediates in **Figure 17** can be the two hydrogen transfer intermediates required by the differential tritium transfer results observed earlier.<sup>95</sup> Thus, propionaldehyde reacts to exchange tritium with [<sup>5'-<sup>3</sup>H</sup>]adenosylcobalamin more rapidly than it produces tritiated substrate by reversal, and (*S*)-2-aminopropanol exchanges tritium with the coenzyme more rapidly than it produces propionaldehyde.



**Figure 18** A mechanism for radical isomerization in the reaction of EAL. In this mechanism, the initial substrate-related free radical intermediate undergoes elimination of ammonia to form an intermediate oxycation radical, in which the unpaired electron and positive charge are mutually delocalized over C1 and C2. The ammonia is retained in the active site and may undergo addition to either C2 in the forward direction or to C1 in reversal.

The mechanism of isomerization between the substrate- and product-related free radicals in **Figure 17** remains uncertain. The mechanism should be related to but different from that of DDH. Because the species of ethanolamine that binds to EAL is the protonated, cationic form,<sup>113</sup> the reaction of EAL requires migration of an  $-\text{NH}_3^+$  group, not of an  $-\text{OH}$  group as in DDH. Thus, the reaction of DDH is regarded as requiring acid catalysis to assist OH elimination or migration; however,  $-\text{NH}_3^+$  is a good leaving group and elimination or migration of this group is not subject to acid catalysis. Once eliminated as  $:\text{NH}_3$ , it is a weak base and subject to protonation, and can accept a hydrogen bond, but protonation cannot precede elimination. A reasonable mechanism for isomerization through elimination and addition of  $\text{NH}_3$  is shown in **Figure 18**.

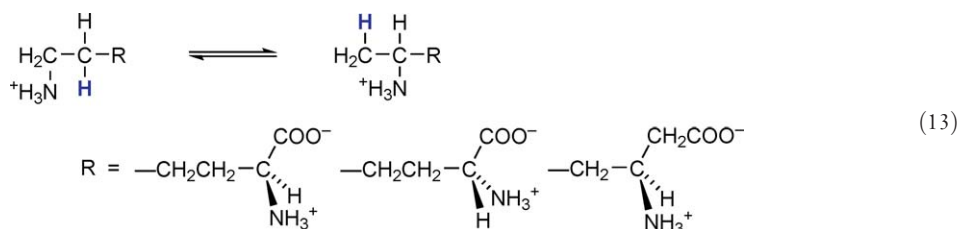
#### 7.14.4.3 Lysine 5,6-Aminomutase

The metabolism of lysine in anaerobic bacteria differs dramatically from that in eukaryotes. At least three different pathways of lysine degradation have been identified in various eukaryotes. These pathways have in common the fact that they are very long and require many enzymes to degrade lysine into acetyl-CoA and succinyl-CoA for processing through the tricarboxylic acid cycle. Catabolic degradation of molecules into biosynthetic building blocks and energy-producing molecules requires the presence of chemical functional groups that can guide the evolution of enzymes that catalyze appropriate chemistry to useful end products. Lysine has sufficient functional groups for this purpose, but they are not positioned in the molecule in such a way as to facilitate degradation in a few steps.

Anaerobic bacteria overcome this problem by moving the functional groups, and they do so through the actions of aminomutases.<sup>114</sup> The first enzyme in anaerobic catabolism of lysine is lysine 2,3-aminomutase, which produces  $\beta$ -lysine. This initial transformation sets the stage for production of acetyl-CoA. Dehydrogenation at C3 of  $\beta$ -lysine and activation of the carboxylate group produce the corresponding  $\beta$ -ketoacyl-CoA, a natural substrate for a  $\beta$ -ketothiolase and production of acetyl-CoA. Lysine 2,3-aminomutase is not  $\text{B}_{12}$ -dependent but rather is a [4Fe-4S] enzyme that requires SAM acting in an analogous capacity to adenosylcobalamin.<sup>115</sup> In the second step of anaerobic lysine catabolism, the  $\varepsilon$ -amino group of (*S*)- $\beta$ -lysine is moved to the adjacent C5 position by the action of lysine 5,6-aminomutase (5,6-LAM), a coenzyme  $\text{B}_{12}$ - and pyridoxal-5'-phosphate (PLP)-dependent enzyme. The resultant 3,5-diaminohexanoic acid upon conversion to the corresponding 3,5-diketohexanoyl-CoA is set up for  $\beta$ -ketothiolase-catalyzed breakdown to three molecules of acetyl-CoA. *Clostridia*, the original source of 5,6-LAM, have not been proven to carry out the complete degradation to acetyl-CoA, but the system of enzymes exists in most anaerobes and is likely to function in at least some of them. *Clostridia* produce acetate and butyrate by fermentation of (*S*)- or (*R*)-lysine. Another function of lysine 2,3-aminomutase is the production of (*S*)- $\beta$ -lysine for the biosynthesis of antibiotics.<sup>116</sup>

### 7.14.4.3.1 Reaction and molecular properties

5,6-LAM is not specific for (*S*)- $\beta$ -lysine, although it is the best of the known substrates, and it accepts (*R*)-lysine and (*S*)-lysine as well. The reaction is depicted in Equation (13), which emphasizes the relationship with other coenzyme B<sub>12</sub>-dependent isomerases as well as lack of specificity for the regio- and stereoisomers of lysine.

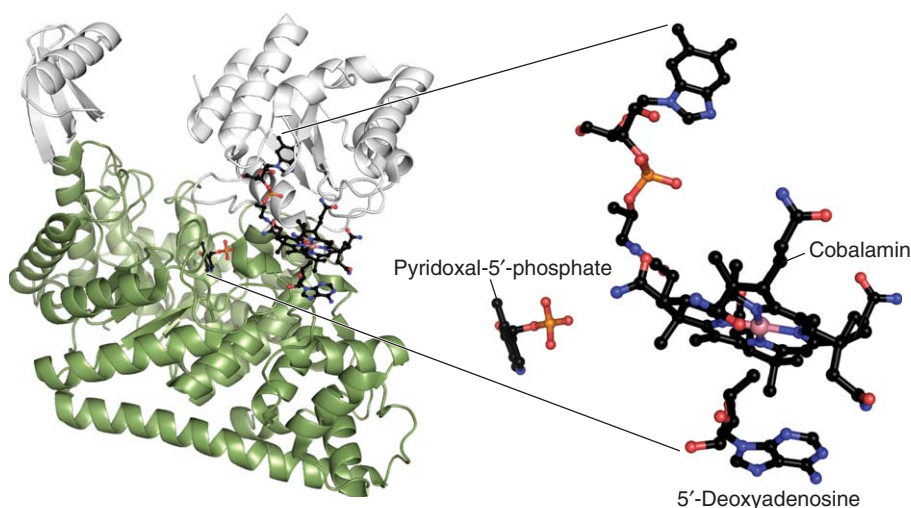


The mechanism of 1,2-hydrogen migration is similar to other coenzyme B<sub>12</sub>-dependent isomerases.<sup>117</sup>

The genes encoding 5,6-LAM have been cloned from *Clostridium sticklandii* and *Clostridium gingivale* and expressed in *E. coli*.<sup>23,118</sup> The enzyme comprises two subunits,  $\alpha$  and  $\beta$ , of molecular weights 57 and 29 kDa associated as a dimer of heterologous dimers ( $\alpha_2\beta_2$ ). EPR experiments with [<sup>15</sup>N-dimethylbenzimidazole]-adenosylcobalamin demonstrate base-off binding, as seen with other B<sub>12</sub>-dependent enzymes such as methionine synthase, MCM, and glutamate mutase and unlike DDH, EAL, and coenzyme B<sub>12</sub>-dependent ribonucleoside triphosphate reductase (RTPR). Borohydride reduction of the internal PLP aldimine followed by fragmentation to peptides and high-performance liquid chromatography (HPLC)-mass spectrometric analysis shows PLP to be bound to Lys144 $\beta$ .<sup>119</sup> Mutation of Lys144 $\beta$  to glutamine leads to an inactive protein.

### 7.14.4.3.2 Molecular structure

A crystal structure of 5,6-LAM with cyanocobalamin and PLP bound but no substrate or inhibitor reveals interesting features that are relevant to the mechanism of action.<sup>120</sup> The crystal structure is illustrated in Figure 19. The smaller  $\beta$ -subunit is a Rossmann domain that binds PLP as the internal aldimine through the C4'-carboxaldehyde group of Lys144 $\beta$ . The 2-methyl-3-hydroxy-5-phosphohydroxymethylpyridine moiety of PLP binds to the  $\alpha$ -subunit, a TIM barrel, effectively forming a cross-link between subunits. The  $\beta$ -subunit also binds cobalamin, but at a distance of 25 Å from the PLP site. The 5,6-dimethylbenzimidazole tail is buried within a pocket similar to other 'base-off' B<sub>12</sub> enzymes. As predicted from the aforementioned EPR results, the lower axial 5,6-dimethylbenzimidazole is not ligated to cobalt but is displaced by His133 $\beta$ .



**Figure 19** A crystal structure of 5,6-LAM with cob(II)alamin and PLP. Shown on the left is a ribbon diagram of the subunits of 5,6-LAM with cobalamin and PLP bound at the interface of subunits. The green ribbon is the TIM barrel ( $\alpha$ )-subunit and the gray ribbon is the Rossmann ( $\beta$ )-subunit. The illustration is generated from PDF 1xrs.<sup>120</sup>



Similar to other B<sub>12</sub>-dependent enzymes, the  $\alpha$ -subunit is a TIM barrel, and it contains most of the contact points for binding PLP, other than Lys144 $\beta$ , which forms the internal aldimine linkage to the substrate. The cross-linking interaction of PLP with the two subunits appears to represent a lock holding the subunits in an orientation unlike other coenzyme B<sub>12</sub>-dependent enzymes, in which the two coenzymes adenosylcobalamin and PLP are held apart. Upon binding a substrate as an external aldimine to PLP, the linkage of PLP to the  $\beta$ -subunit would be released, and the subunits would be free to rotate into an orientation with the  $\beta$ -Rossmann domain directly over the  $\alpha$ -TIM barrel, more like the orientations in glutamate and MCMs. In such an orientation, the adenosylcobalamin would be brought into the base of the TIM barrel and into proximity with the substrate PLP aldimine.<sup>120</sup>

#### 7.14.4.3.3 Reaction mechanism

The action of 5,6-LAM proceeds with deuterium kinetic isotope effects on  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_{\text{m}}$  in reactions of deuterated (*R*)- and (*S*)-lysine ranging from 8 to 10 for  $k_{\text{H}}/k_{\text{D}}$ , showing that hydrogen transfer is rate limiting in the mechanism.<sup>121</sup> The enzyme functions nearly as well on 4-oxa-(*RS*)-lysine as on (*RS*)-lysine. Unlike other coenzyme B<sub>12</sub>-dependent enzymes, no evidence for cob(II)alamin formation upon addition of a substrate to the activated enzyme has ever been observed.

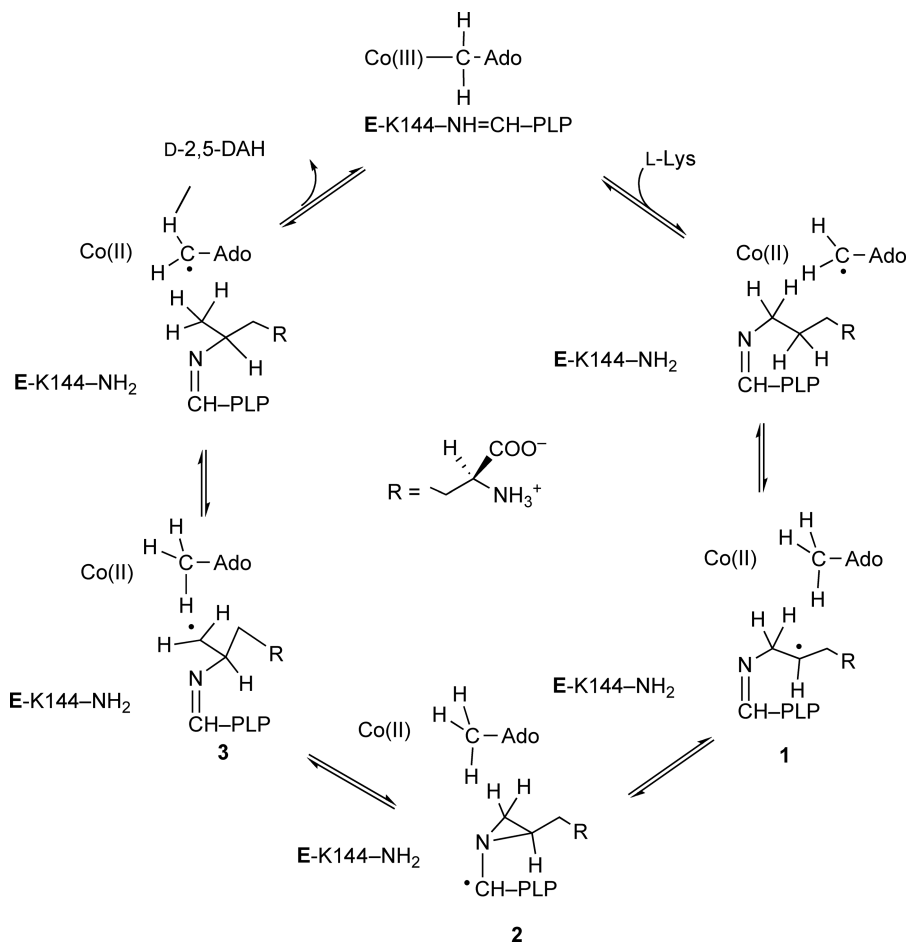
The hypothetical chemical mechanism shown in **Figure 20** is under consideration. The well-established mechanism of the action of lysine 2,3-aminomutase inspires that for 5,6-LAM in **Figure 20**.<sup>115,116</sup> The mechanism differs from that of lysine 2,3-aminomutase in that adenosylcobalamin and not SAM serves as the source of the 5'-deoxyadenosyl radical. And PLP forms the external aldimine with the  $\epsilon$ -amino group of the substrate and not with the  $\alpha$ -amino group. Unlike the reaction of lysine 2,3-aminomutase, none of the substrate-derived free radical intermediates in the reaction of 5,6-LAM are sufficiently stable to be observed by EPR spectroscopy. For this reason, no cob(II)alamin is observed in the steady state. Nonenzymatic reactions of radicals analogous in structure to those shown in **Figure 20** are counterparts to the radical isomerization mechanism shown.<sup>122</sup> The mechanism is further supported by theoretical calculations showing that the structure of radical **3** in **Figure 20** most likely serves as the intermediate connecting the substrate-related radical **2** with the product-related radical **3**.<sup>123</sup>

Reactions of 4-thia-(*R*)- and 4-thia-(*S*)-lysine with 5,6-LAM provide EPR spectroscopic evidence in support of the mechanism in **Figure 20**.<sup>124</sup> The 4-thia-analogues of lysine provide the potential to stabilize a radical intermediate with electron spin at C5, through the electronic stabilizing effect of sulfur in place of C4. The 4-thia-analogues serve as adenosylcobalamin- and PLP-dependent suicide inhibitors of 5,6-LAM. The inhibited state of the enzyme contains cob(II)alamin, 5'-deoxyadenosine, and an organic free radical. Isotope-edited EPR experiments employing <sup>2</sup>H-labeled and <sup>13</sup>C-labeled 4-thia analogues of lysine prove that the unpaired electron in the organic free radical resides on the C5,C6-locus of the inhibitor. The radical observed is assigned as the 4-thia analogue of radical **2** in **Figure 20**. The structures of the (*R*)- and (*S*)-radicals are shown in **Figure 21**. The observation of these radicals and of cob(II)alamin support the chemical mechanism in **Figure 20**.

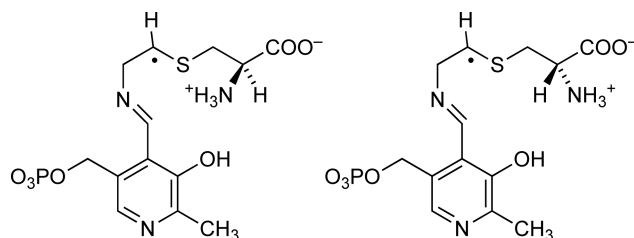
#### 7.14.4.3.4 Suicide inactivation by substrates

Most coenzyme B<sub>12</sub>-dependent enzymes lose activity during prolonged reactions with substrates. 5,6-LAM is especially susceptible to substrate-dependent suicide inactivation. Based on extensive biochemical analysis, the chemical course of this process involves proton-coupled electron transfer leading to quenching of the putative substrate radical intermediates.<sup>23</sup> Lysine-dependent suicide inactivation displays the following properties: (1) Lysine-dependent inactivation requires adenosylcobalamin and PLP in addition to the substrate. (2) (*R*)- and (*S*)-lysine and (*S*)- $\beta$ -lysine inactivate 5,6-LAM with values of  $k_{\text{inact}}$  of 0.7, 0.6, and  $> 18 \text{ min}^{-1}$ , respectively, at 37 °C. (*R*)- and (*S*)-Lysine inactivate after approximately 1000 and 200 turnovers, respectively. (3) Inactivation is accompanied by the formation of cob(III)alamin and 5'-deoxyadenosine. (4) The inactivation rate is the same under anaerobic as under aerobic conditions. (5) Substrates are not consumed in suicide inactivation. (6) Inactivation in <sup>3</sup>H<sub>2</sub>O leads to tritiated substrate and product, but not to tritiated 5'-deoxyadenosine. These facts are accounted for by the mechanism outlined in **Figure 22**. Catalytic intermediates **2** and **3** in **Figure 20** occasionally undergo electron transfer from cob(II)alamin to a free radical center in either **2** or **3**, which becomes carbanionic and is immediately quenched by a proton donor in equilibrium



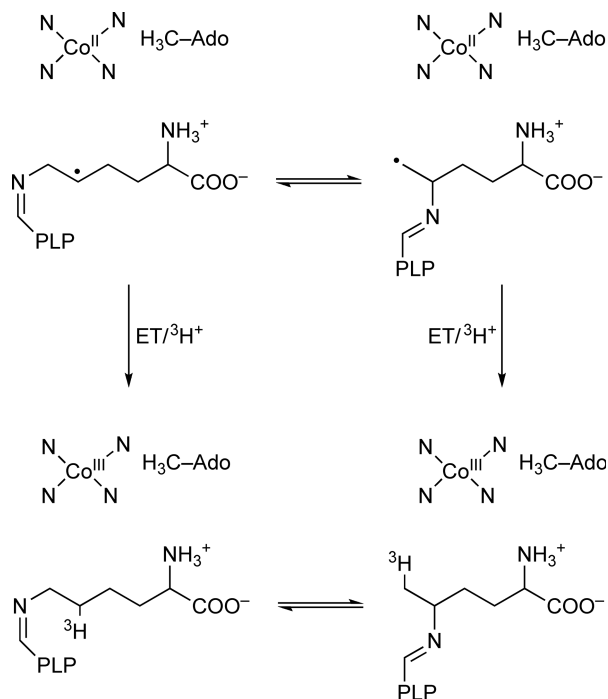


**Figure 20** A chemical mechanism for the action of 5,6-LAM on (S)-lysine. In this mechanism, the 5'-deoxyadenosyl radical from coenzyme B<sub>12</sub> abstracts hydrogen from C5 of the lysyl side chain in the internal PLP aldimine. The resultant free radical **1** undergoes isomerization by internal cyclization to the azacyclopropyl carbonyl radical **2**, which opens to the primary C6 radical **3**. Hydrogen abstraction from the methyl group of 5'-deoxyadenosine and release of 2,5-diaminohexanoate completes the mechanistic cycle.



**Figure 21** Free radicals in reactions of 5,6-LAM with 4-thia-analogues of lysine. Free radicals are generated in the reactions of the suicide inhibitors 4-thia-(S)- and 4-thia-(R)-lysine with 5,6-LAM. The structure on the left is assigned to the radical derived from 4-thia-(S)-lysine and that on the right to the radical from 4-thia-(R)-lysine.<sup>124</sup>

with the medium. This accounts for tritium incorporation, regeneration of the substrate or product with each suicide event, and the production of cob(III)alamin and 5'-deoxyadenosine without the intervention of molecular oxygen. This mode of suicide inactivation also supports the assignment of the mechanism in [Figure 20](#).



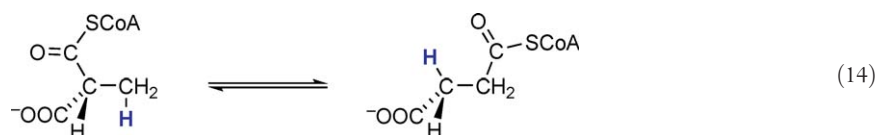
**Figure 22** Substrate-dependent suicide inactivation of 5,6-LAM by coupled electron and proton transfer.

#### 7.14.4.4 Methylmalonyl-CoA Mutase

##### 7.14.4.4.1 Reaction and molecular composition

Enzymes catalyzing carbon skeleton rearrangements were the first to be recognized as coenzyme B<sub>12</sub>-dependent. The first was glutamate mutase (GM), discovered in 1960,<sup>2</sup> and the nature of the skeletal rearrangement in the reaction of GM inspired the identification of coenzyme B<sub>12</sub> as the coenzyme for the apparently similar reaction of MCM.<sup>3,4</sup>

Both GM and MCM catalyze the migration of a carbon fragment to an adjacent carbon accompanied by the cross-migration of a hydrogen atom, following the pattern of Equation (5). The reaction of MCM is shown in Equation (14).

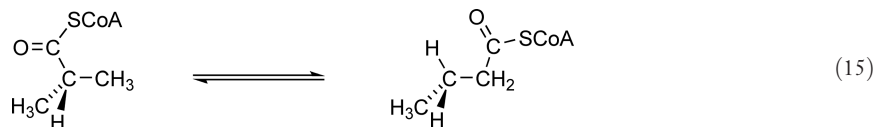


Carbon-14 labeling experiments show that the reaction proceeds with migration of the acyl-CoA moiety and not the carboxylate group.<sup>15</sup> Hydrogen transfer proceeds with overall retention of configuration, as shown in Equation (14).<sup>13</sup> As in all coenzyme B<sub>12</sub>-dependent isomerization reactions, hydrogen transfer proceeds with exchange with the C5'-methylene hydrogens of adenosylcobalamin.<sup>125</sup>

MCM plays an essential role in propionate metabolism. Propionate and propionyl-CoA are intermediates in the catabolism of leucine and isoleucine and are further metabolized by carboxylation of propionyl-CoA to methylmalonyl-CoA. Isomerization to succinyl-CoA feeds the carbon chain into the tricarboxylic acid pathway of oxidative metabolism. For this reason, MCM is an important enzyme in bacterial and mammalian metabolism. It is one of the two vitamin B<sub>12</sub>-dependent enzymes known to be important in human metabolism.

MCMs are heterodimeric enzymes comprising larger and smaller subunits. In *Propionibacterium shermanii*, MCM is encoded by genes *mutA* (80 kDa) and *mutB* (69 kDa) and can be expressed and constituted in *E. coli*.<sup>126,127</sup> The larger subunit binds adenosylcobalamin. The human and bacterial enzymes are homologous.

The reaction of isobutyryl-CoA mutase was found to be similar in principle (Equation (15)). The genes and heterodimeric subunit composition and catalytic properties have been described.<sup>128</sup>



#### 7.14.4.4.2 Structure of methylmalonyl-CoA mutase

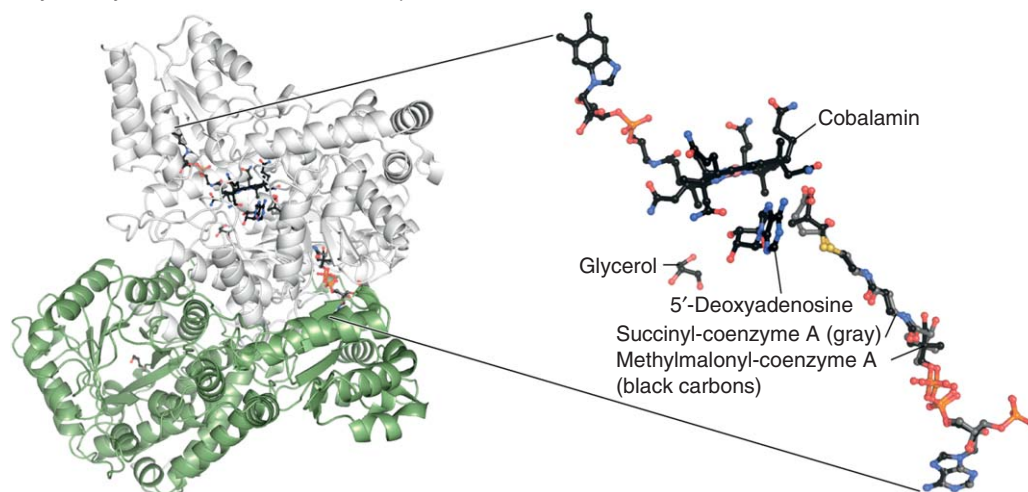
The first coenzyme B<sub>12</sub>-dependent isomerase to be crystallized and structurally analyzed by X-ray crystallography is MCM.<sup>129,130</sup> Several crystal structures of MCM are available, and one is illustrated in **Figure 23**. Like GM and 5,6-LAM, the larger subunit is a TIM barrel. Also like GM, MCM, and methionine synthase, adenosylcobalamin is bound in a ‘base-off’ mode.

Structures of MCM demonstrate most clearly an important principle in the action of enzymes that catalyze reaction with free radical intermediates. The structures with substrates or substrate analogues show that the reacting portions of the substrate and of adenosylcobalamin are deeply buried within the structure. In the structures of complexes formed between MCM and CoA compounds, the structure of CoA is extended through the protein from the surface to the buried active site. In the absence of a CoA derivative, the TIM barrel is partially open, exposing the CoA-binding site, and it closes around CoA upon binding. This feature makes the radical intermediates inaccessible to external reducing or oxidizing agents that could quench free radical intermediates. It is important for enzymes to prevent undesirable side reactions.<sup>109</sup> In cases of enzymes catalyzing free radical reactions, sequestration of radical intermediates is a means to control the highly reactive intermediates. Another feature of enzymes catalyzing radical reactions is the constraint placed on radical intermediates, by enzymatic binding interactions, with regard to the contacts they make within active sites.<sup>131</sup>

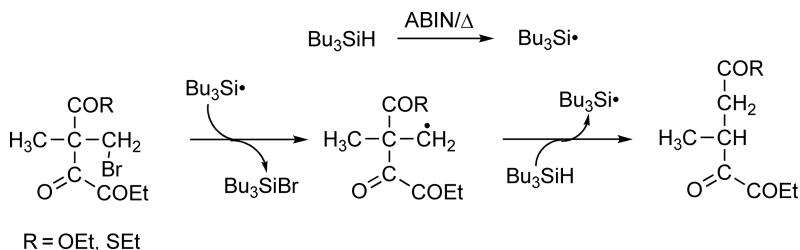
#### 7.14.4.4.3 The radical mechanism of methylmalonyl-CoA mutase

The observation of MCM-catalyzed hydrogen exchange between substrates and C5' of adenosylcobalamin predated the earliest evidence of radicals in coenzyme B<sub>12</sub>-dependent reactions.<sup>125</sup> Early evidence for the participation of free radicals included the observation of a breakdown in stereochemical specificity in the reaction of the substrate analogue ethylmalonyl-CoA.<sup>132</sup> The weakness of the Co–C5' bond and observation of

Methylmalonyl-CoA mutase–substrate complex



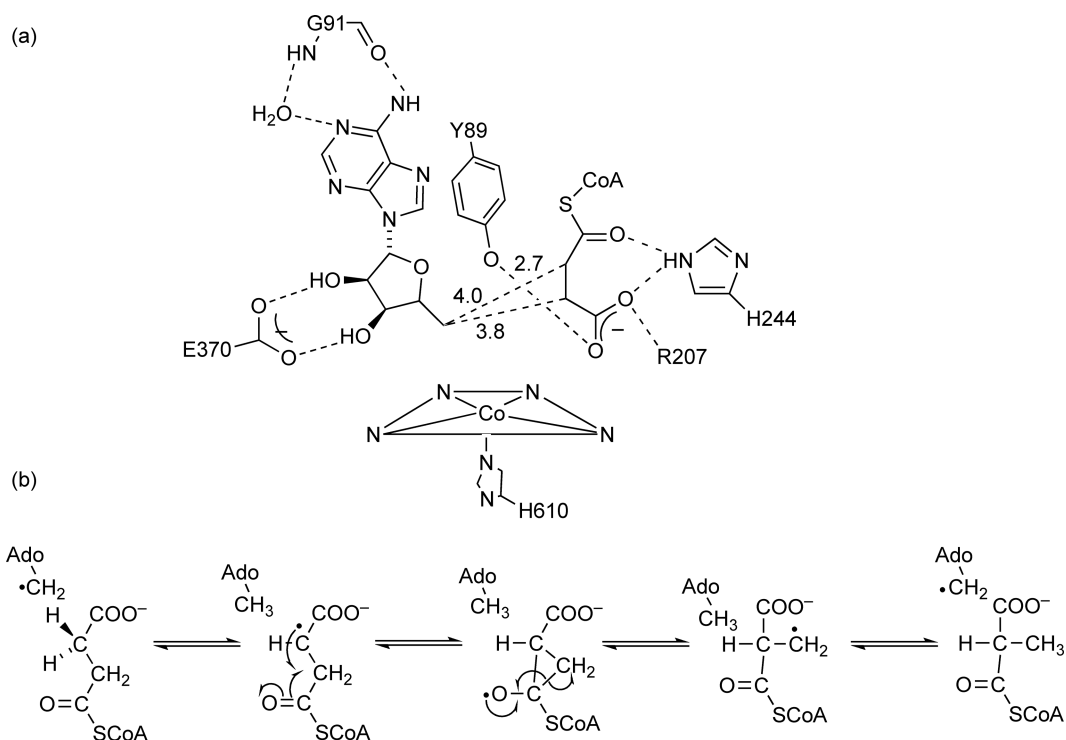
**Figure 23** A crystal structure of MCM from *Propionibacterium shermanii*. Shown on the left is a ribbon diagram of the two subunits of MCM, with cob(II)alamin, 5'-deoxyadenosine, and succinyl- or malonyl-CoA bound to MutA (gray). MutB is shown in green. The illustration was generated from PDF 4req.<sup>130</sup>



**Figure 24** A chemical model for the radical isomerization step in the action of MCM. Abstraction of bromine from 2-methyl-2-bromomethylmalonate diester by tributylsilyl radical generates the analogue of the substrate-related free radical in the action of MCM. The free radical rapidly undergoes isomerization to the 2-methylsuccinyl radical diester, which is quenched by tributylsilane. Radical isomerization presumably proceeds by way of the oxycyclopropyl radical.<sup>133</sup>

radical intermediates in the reactions of DDH also focused attention on radical mechanisms for MCM. The chemical isomerization of 2-methyl-(2-bromomethyl)malonyl diesters to 2-methylsuccinyl diesters under radical initiating conditions, as shown in **Figure 24**, established a chemical counterpart for MCM-catalyzed radical isomerization.<sup>133</sup> This mechanism, featuring the oxycyclopropyl radical as the key intermediate, has been accepted as the most likely route for radical isomerization in the reaction of MCM.

With the foregoing information in hand, it became reasonable to write the mechanism in **Figure 25(b)** occurring within the active site contacts in **Figure 25(a)**. The 5'-deoxyadenosyl radical from coenzyme B<sub>12</sub>



**Figure 25** Active site contacts and a mechanism for the action of MCM. (a) Mechanism-related active site contacts between MCM, cobalamin, 5-deoxyadenosine, and succinyl-CoA are shown, with selected distances in angstrom. Tyr89 and His244 appear to contribute hydrogen bonding to the substrate, and mutation of these residues decreases activity. Other contacts to the vicinal OH groups and adenine ring of 5-deoxyadenosine and the acyl-carbonyl and carboxylate groups of succinyl-CoA are also shown. Data obtained from PDB 4req.<sup>130</sup> (b) A mechanism for the action of MCM. The 5'-deoxyadenosyl radical from coenzyme B<sub>12</sub> initiates the reaction by abstraction of hydrogen from the methyl group of methylmalonyl-CoA. The resultant free radical undergoes isomerization by internal cyclization to the oxycyclopropyl radical, which opens to the succinyl-CoA radical, as in the chemical counterpart of **Figure 24**. Hydrogen abstraction from the methyl group of 5'-deoxyadenosine completes the mechanistic cycle.

could abstract a hydrogen atom from the methyl group of methylmalonyl-CoA to form 5'-deoxyadenosine and the methylene radical of the substrate. Isomerization of this radical by the mechanism in **Figure 24** to the succinyl-CoA radical shown, followed by hydrogen abstraction from the methyl group of 5'-deoxyadenosine would lead to succinyl-CoA and regeneration of adenosylcobalamin.

Evidence for the formation of paramagnetic species in the action of MCM includes early experiments employing EPR spectroscopy.<sup>134–136</sup> Very clear proof of the structure of the succinyl-CoA radical in **Figure 25** as the observable intermediate through isotope-edited EPR spectroscopy consolidates the evidence for the mechanism.<sup>25</sup> This structure is the organic free radical component of the radical triplet species arising from spin coupling between low-spin  $\text{Co}^{2+}$  in cob(II)alamin and the organic radical observed in the EPR spectrum. The distance between the paramagnetic species is 6.0 Å.

The tritium isotope effects in the action of MCM are large, typical of adenosylcobalamin-dependent enzymes. The large kinetic isotope effects are attributed to quantum mechanical tunneling of hydrogen. A detailed theoretical analysis shows that multidimensional hydrogen tunneling accounts for the observed kinetic isotope effects.<sup>137</sup> Conservative mutations in Tyr89 and His244 of MCM, shown in **Figure 25(a)** to be in contact with the substrate, dramatically decrease both the catalytic activity and the tritium kinetic isotope effects.<sup>138–140</sup> With wild-type MCM, tritium partitioning from  $[5\text{'-}^3\text{H}]$ adenosylcobalamin to methylmalonyl-CoA and succinyl-CoA is identical regardless of whether tritium is mobilized by addition of the substrate or product, indicating that radical isomerization does not participate in rate limitation.<sup>141</sup> However, with Y89F-MCM, the partitioning depends on whether mobilization is induced by methylmalonyl-CoA or succinyl-CoA, indicating significant rate limitation in the radical isomerization step.<sup>138</sup> The structures of H144A- and Y89F-MCM are identical to wild-type MCM except for the cavities created by the absence of functional groups. All evidence indicates that His244 and Tyr89 assist the radical isomerization step. Mutations of these residues also increase the sensitivity of MCM to suicide inactivation, presumably by decreasing the protective effects of the native structure.

The significance of hydrogen tunneling in enzymatic reactions is a current subject in mechanism studies. That tunneling occurs is not under debate, whether enzymes are evolved to engage in hydrogen tunneling is debatable, and the adenosylcobalamin-dependent enzymes are excellent examples of hydrogen tunneling. A careful nonenzymatic study of hydrogen abstraction by the 5'-deoxyadenosyl and by 8-methoxyl-5'-deoxyadenosyl radicals shows that hydrogen tunneling occurs in nonenzymatic counterparts of the reactions of DDH and MCM, with quantitative agreement between the nonenzymatic and MCM-catalyzed processes.<sup>142</sup> It is possible that tunneling is incidental to the chemical mechanisms, at least in the cases of DDH and MCM, and not to the evolution of these two enzymes.

It has been proposed for the reactions of MCM and GM that the methylene radicals in the mechanisms are stabilized by magnetic interactions with low-spin  $\text{Co}^{2+}$  in cob(II)alamin.<sup>143</sup> To date, the importance of such stabilization has not been established. The amount of stabilization has not allowed methylene radical intermediates such as the 5'-deoxyadenosyl radical to be observed by EPR spectroscopy. The only observable radicals to date have been the most chemically stable radicals in the mechanisms, so magnetic stabilization could not have led to leveling of the stabilities of the radical intermediates. The rates and activation parameters measured for the model radical isomerizations in **Figure 24** were compatible with enzymatic rates.<sup>133</sup> The results showed that participation of the cobalamin portion of the coenzyme would not be required to explain substrate isomerization. The energy of the magnetic interaction between cob(II)alamin and the free radical at the active site of MCM is 4000 G ( $0.37\text{ cm}^{-1}$ ) or  $1\text{ cal mol}^{-1}$  (G. H. Reed, personal communication). This amount of energy would not be an important contribution to stabilization in a radical.

#### 7.14.4.4.4 Cleavage of the Co–C5' bond

The means by which enzymes facilitate the homolytic cleavage of the Co–C5' bond has been addressed in detail in studies of MCM. This process can be kinetically monitored by the spectral change from Co(III) in coenzyme  $\text{B}_{12}$  to that of cob(II)alamin. This spectral change depends on the addition of the substrate to the complex of MCM and adenosylcobalamin. The kinetic barrier to bond cleavage is lowered by  $17\text{ kcal mol}^{-1}$ .<sup>144</sup> Moreover, the rate of this change displays a kinetic isotope effect of  $\geq 20$  ( $k_{\text{H}}/k_{\text{D}}$ ) when the deuterated substrate is employed.<sup>145</sup> It was concluded that Co–C5' bond cleavage and hydrogen abstraction from the substrate are kinetically coupled. This effect has been reported for other coenzyme  $\text{B}_{12}$ -dependent reactions as well.

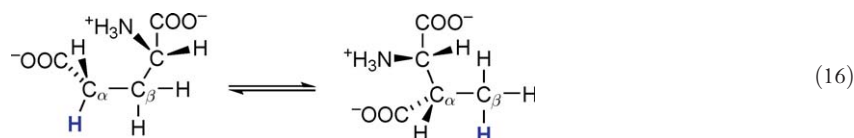
Stereochemical evidence of the MCM-induced cleavage of the Co–C5' is available through the synthesis of stereoselectively deuterated [5'-<sup>2</sup>H]adenosylcobalamin.<sup>146</sup> (5'*R*)-[5'-<sup>2</sup>H] and (5'*S*)-[5'-<sup>2</sup>H]adenosylcobalamin can be distinguished by H-NMR (nuclear magnetic resonance) analysis through the differences in proton chemical shifts and 5'/4' coupling constants. Samples of (5'*R*)-[5'-<sup>2</sup>H] and (5'*S*)-[5'-<sup>2</sup>H]adenosylcobalamin were partially epimerized upon incubation with MCM in the absence of a substrate, proving that the Co–C5' bond had been cleaved.

There has been much discussion of the possible *trans* effect of replacing 5,6-dimethylbenzimidazole with His610 as the lower axial cobalamin ligand in MCM. The *trans* effect has been eliminated as an important factor in Co–C5' cleavage by showing that a derivative of adenosylcobalamin with guanosine diphosphate (GDP) in place of the 5,6-benzimidazole ribotide functions well as the coenzyme.<sup>147</sup> This derivative of coenzyme B<sub>12</sub> binds to MCM in a His610-off manner. Therefore, the His610 as the lower axial ligand could not be an important factor in Co–C5' cleavage.

Analysis of the electronic structures of adenosylcobalamin and cob(II)alamin in free and MCM-bound states by magnetic circular dichroism spectroscopy shows that the cleaved cob(II)alamin is stabilized by interactions with MCM.<sup>148</sup> No evidence for the destabilization of adenosylcobalamin through binding to MCM could be found. The Co–C5' bond should be weakened through enzymatic stabilization of cob(II)alamin, just as it is weakened nonenzymatically in alkylcobalamins through electronic stabilization of the incipient alkyl radicals. Referring again to **Table 1**, the stabilities of the benzyl- and 3',4'-anhydro-5'-deoxyadenosyl radicals lead to weakening of the Co–C5' bonds in benzyl- and 3',4'-anhydroadenosylcobalamins. Enzymatic stabilization of incipient cob(II)alamin should as well weaken the Co–C5' bond.

#### 7.14.4.5 Glutamate and Methylene-glutarate Mutases

GM catalyzes the interconversion of (*S*)-glutamate (glutamate) and (2*S*)-*threo*-3-methylaspartate (3-methylaspartate) according to Equation (16).



The discovery of coenzyme B<sub>12</sub> is linked to the discovery of GM in connection with the fermentation of glutamate in *Clostridia*. Coenzyme B<sub>12</sub> is the light-sensitive cofactor required to activate GM.<sup>2,149</sup> The hydrogen transfer and stereochemistry are as shown in Equation (16);<sup>14</sup> that is, the glycyl moiety migrates from C3 to C2 of glutamate and hydrogen migrates in the opposite direction between the same carbons, in the pattern of Equation (5).

##### 7.14.4.5.1 Function of adenosylcobalamin

The reaction of GM shares many properties in common with the other coenzyme B<sub>12</sub> enzymes. Carbon-5' of adenosylcobalamin undergoes hydrogen isotope exchange with the substrates.<sup>150</sup> The rate constants governing the 3-methylaspartate-induced loss of tritium from [5'-<sup>3</sup>H]adenosylcobalamin and for the appearance of tritium in 3-methylaspartate are the same, implicating C5' in mediating hydrogen transfer by direct reaction with a substrate.<sup>151</sup> Cleavage of the Co–C5' bond and hydrogen transfer from the substrate are kinetically coupled, as shown by a large deuterium kinetic isotope effect on the appearance of cob(II)alamin with deuterated glutamate or 3-methylaspartate as substrate.<sup>152</sup> A substrate-based free radical can be detected by EPR spectroscopy in the steady state of the reaction of glutamate.<sup>153</sup> This radical is part of a spin triplet system in which the 4-glutamyl radical (glutamate-4-yl) is spin coupled with cob(II)alamin.<sup>154</sup> Spectral simulations indicate a distance of 6.6 Å separating the radical and cob(II)alamin, similar to the corresponding distance in MCM. Perturbations in the circular dichroism and magnetic circular dichroism spectra of cob(II)alamin bound to GD indicate stabilizing interactions with the enzyme, unlike passive interactions with adenosylcobalamin.<sup>155</sup> As in MCM, the evidence indicates stabilization of cob(II)alamin in the cleaved coenzyme and no destabilization of adenosylcobalamin.



### 7.14.4.5.2 Alternative and suicidal substrates

GM catalyzes the reaction of (*S*)-2-hydroxyglutarate to (*S*)-*threo*-3-methylmalate at a slow rate.<sup>156</sup> In this substrate, the 2-hydroxyl group replaces the 2-amino group of glutamate. The overall reaction proceeds at 1% of the rate with glutamate ( $k_{\text{cat}} = 0.05 \text{ s}^{-1}$  compared with  $5.6 \text{ s}^{-1}$ ). Cleavage of the Co–C5' bond upon addition of (*S*)-2-hydroxyglutarate is much faster, similar to the rates with glutamate or 3-methylaspartate. Available information suggests that the overall rate might be limited by the rearrangement step.

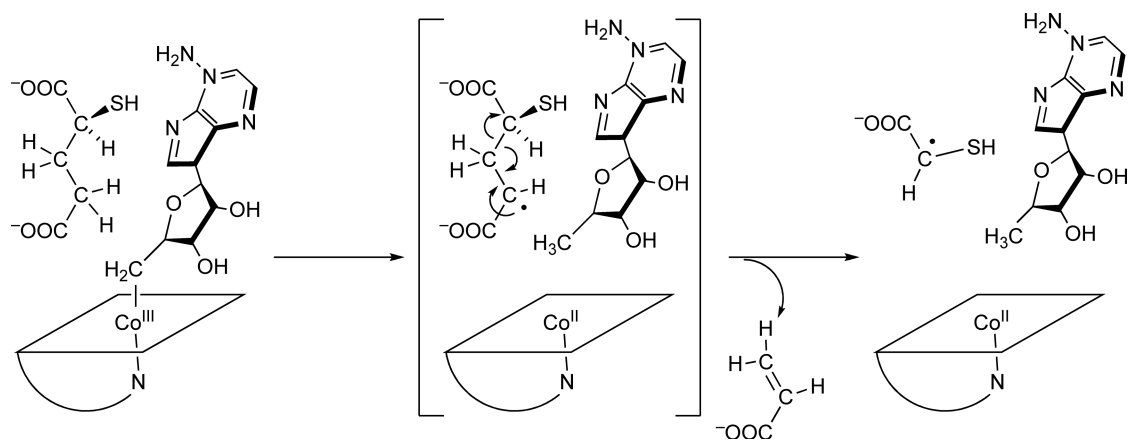
The complex of GM with adenosylcobalamin reacts with 2-thiolglutarate to form cob(II)alamin, 5'-deoxyadenosine, and the thioglycolyl radical, as illustrated in **Figure 26**.<sup>157</sup> It appears that 2-thiolglutarate reacts initially as a substrate to form the radical 2-thiolglutarate-4-yl, cob(II)alamin, and 5'-deoxyadenosine. Elimination of acrylate from the radical leads to the thioglycolyl radical, which is too stable to react further. The cleavage and fate of 2-thiolglutarate in its reaction with GM will be cited in a later section as evidence in support of a chemical mechanism of action of GM.

2-Ketoglutarate reacts with the complex of GM and adenosylcobalamin.<sup>158</sup> In the reaction, C4(H) of 2-ketoglutarate undergoes exchange with the C5'-methylene hydrogens of adenosylcobalamin, which is itself reversibly cleaved to cob(II)alamin. No rearrangement of the carbon skeleton in 2-ketoglutarate occurs. It appears that 2-ketoglutarate induces cleavage of the Co–C5' bond to generate the 5'-deoxyadenosyl radical transiently and reversibly. Hydrogen exchange presumably occurs between C4 of 2-ketoglutarate and the 5'-deoxyadenosyl radical.

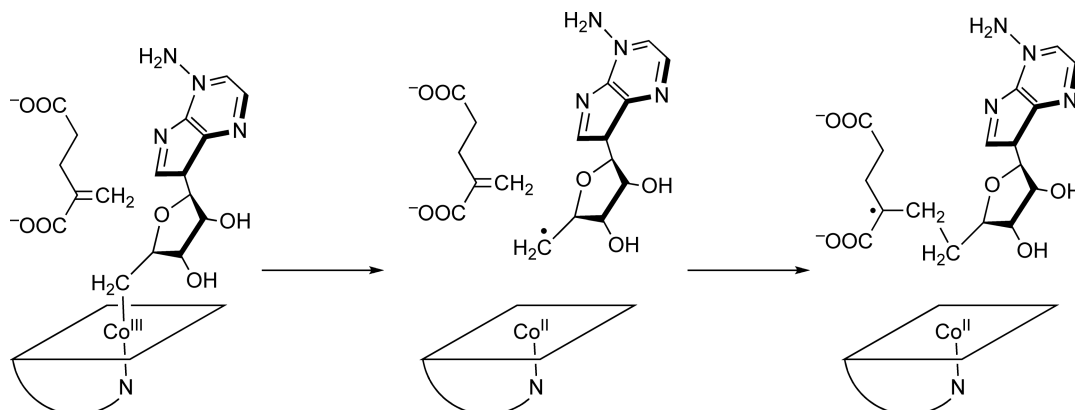
2-Methyleneglutarate, with a CH<sub>2</sub> group in place of the 2-keto-oxygen in 2-ketoglutarate, is the substrate for methyleneglutarate mutase (MGM), a coenzyme B<sub>12</sub>-dependent enzyme that catalyzes an isomerization thought to be related to that of GM. 2-Methyleneglutarate inactivates the complex of GM and coenzyme B<sub>12</sub> in a time-dependent manner.<sup>159</sup> Inactivation proceeds with Co–C5' cleavage of adenosylcobalamin to form cob(II)alamin and a radical centered at C2 of 2-methyleneglutarate. The reaction appears to follow the course in **Figure 27**, in which the reaction ends by addition of the 5'-deoxyadenosyl radical to the methylene group. Further uncharacterized reaction might involve electron transfer from cob(II)alamin, perhaps in analogy with the suicide inactivation of 5,6-LAM described in an earlier section.

### 7.14.4.5.3 Molecular properties and structure

Genes encoding GM have been cloned from *Clostridium cochlearium*, expressed in *E. coli*, and the subunits purified and assembled into active enzyme.<sup>153</sup> GM is composed of two subunits, E a dimer of 53 kDa and S a monomer of 15 kDa, which assemble into a heterotetramer. Like other coenzyme B<sub>12</sub> enzymes, the smaller



**Figure 26** Thioglycolyl radical formation at the active site of GM. 2-Thiolglutarate reacts with GM to form the thioglycolyl radical, cob(II)alamin, and 5'-deoxyadenosine. The chemical transformations shown here might be the mechanism by which the thioglycolyl radical is generated. Steps leading to the thioglycolyl radical might be analogous to the catalytic mechanism in the reaction of glutamate.

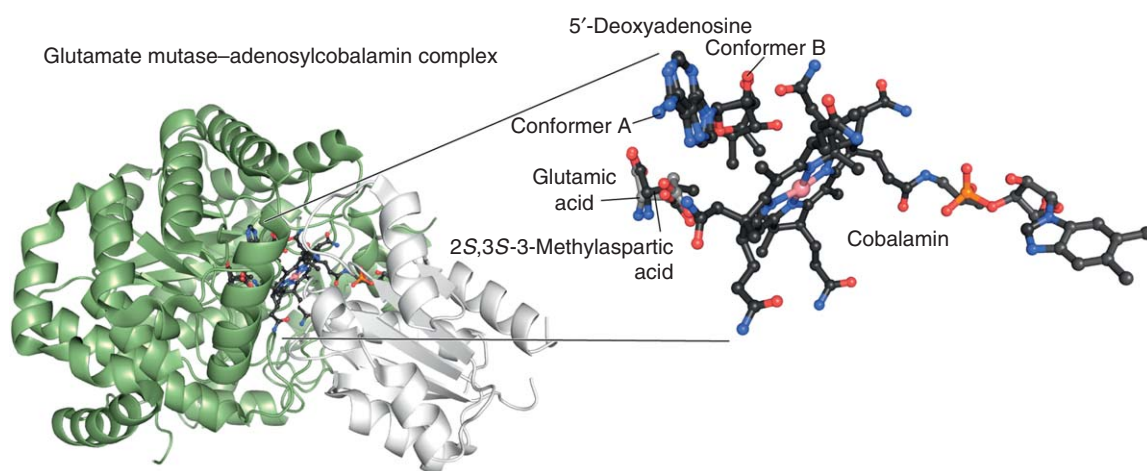


**Figure 27** Suicide inactivation of GM by 2-methyleneglutarate. The chemical steps shown here appear to be involved in the suicide inactivation of GM by 2-methyleneglutarate. The Co–C5' cleavage is likely to be the first chemical step in the reaction of glutamate as a substrate.

subunit S binds adenosylcobalamin. The similar genes from *Clostridium tetanomorphum* have also been cloned, expressed in *E. coli*, and the subunits MutE and MutS purified.<sup>160</sup>

A number of structures are available from NMR and X-ray crystallographic analysis. One structure illustrated in **Figure 28** shows the assembled subunits. The assembly of subunits is analogous to other coenzyme B<sub>12</sub> enzymes but structurally unique. Also shown in **Figure 28** are the ligands cob(II)alamin, 5'-deoxyadenosine, and glutamate as they are oriented in the active center. The positioning of glutamate relative to cob(II)alamin is consistent with the spectroscopic results.

The structure in **Figure 28** contains information potentially pertaining to translocation of the methylene group of the 5-deoxyadenosyl radical from Co<sup>2+</sup> of cobalamin to the substrate 6 Å away. The structure shows two conformations of the ribosyl ring in 5-deoxyadenosine, the C2'-*endo* and the C3'-*endo* conformations.<sup>161</sup> The two conformations project the 5'-methyl group 3.1 and 4.5 Å, respectively, from Co<sup>2+</sup>. The orientation in the C3'-*endo* conformation, if replicated in the 5-deoxyadenosyl radical, would allow the methylene radical to

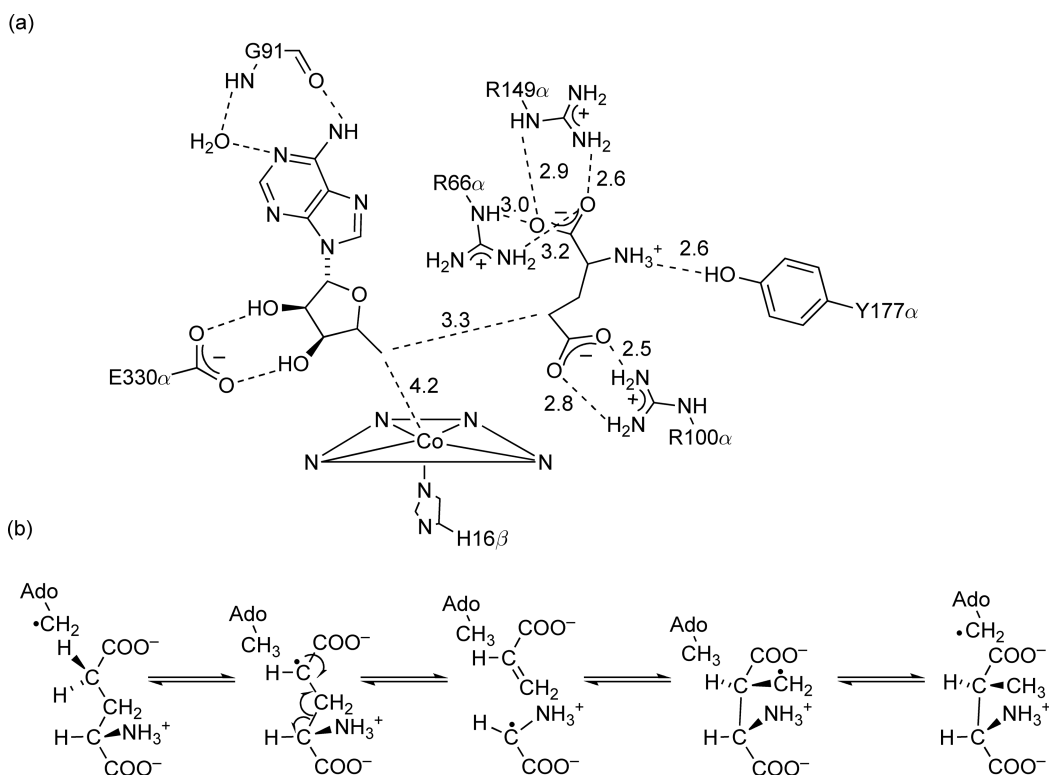


**Figure 28** A crystal structure of GM from *Clostridium cochlearium* with cob(II)alamin, 5'-deoxyadenosine, and glutamate. On the left is a ribbon diagram of the associated E (green) and S (gray) subunits with cob(II)alamin at the subunit interface. Note the extension of the 5,6-benzimidazole tail in the 'base-off' conformation. Glutamate is positioned approximately 7 Å from cobalt and near 5'-deoxyadenosine. Also shown are the 2-*endo* and 3-*endo* conformations of 5-deoxyadenosine, which are thought to be relevant to the mechanism. The illustration was generated from PDF 1i9c.<sup>161</sup>

contact the substrate hydrogen in the abstraction step. Thus, the 9 Å distance can be spanned by *N*-ribosyl torsion in DDH and EAL, and the 6 Å translocation can be accomplished by a C2'-*endo* to C3'-*endo* conformational change in GM and perhaps MCM.

#### 7.14.4.5.4 Reaction mechanism

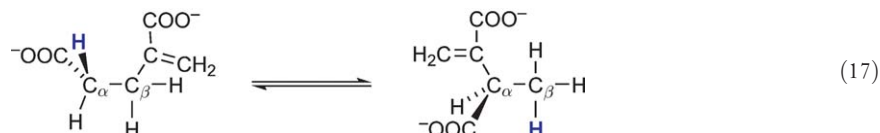
Unlike DDH, 5,6-LAM, and MCM, for which chemical examples of radical isomerizations are available, there is no nonenzymatic model reaction for the isomerization of the 4-glutamyl radical identified as the intermediate observable by EPR spectroscopy at the active site of GM.<sup>154</sup> The accepted mechanism is the radical fragmentation/addition mechanism illustrated in **Figure 29**, together with active site contacts for glutamate. In this mechanism, the initial 4-glutamyl radical undergoes fragmentation to acrylate and the 2-glycyl radical. Readdition of the 2-glycyl radical to C2 of acrylate generates the 3-methylene radical form of 3-methylaspartate, which abstracts a hydrogen atom from 5'-deoxyadenosine to complete the process.<sup>162,163</sup> This mechanism has much in its favor, including the stability of the initial 4-glutamyl radical, with spin delocalization into the 5-carboxylate group, and the stability of the 2-glycyl radical, which is stabilized by the captodative effects of the 1-carboxyl group and the 2-amino group. The highest energy radicals are the 3-methylene radical of 3-methylaspartate and the 5'-deoxyadenosyl radical. Powerful experimental support for the mechanism includes the observation that whereas neither acrylate nor glycine inhibits very effectively, the two molecules inhibit synergistically. Further, binding of both acrylate and glycine to GM elicits the formation of an EPR spectrum related to that from addition of glutamate.<sup>162</sup> Rapid mix-quench kinetics documents the formation of 5'-deoxyadenosine at the same rate as Co–C5 bond cleavage and the formation of an intermediate tentatively assigned to the acrylate/2-glycyl radical state in **Figure 29**.<sup>164</sup> The mechanism also accounts for the reaction of 2-thiolglutarate with GM.<sup>157</sup> A theoretical analysis by *ab initio* molecular orbital calculations validates the radical elimination–addition mechanism for the reaction of GM.<sup>165</sup>



**Figure 29** Active site contacts and a fragmentation/recombination mechanism for GM. (a) Salient features of the active site include the binding of glutamate through hydrogen-bonded electrostatic contacts with three arginine residues. (b) Experimental and theoretical support for the fragmentation and addition mechanism is presented in the text.

#### 7.14.4.5.5 2-Methyleneglutarate mutase

MGM catalyzes the carbon skeletal isomerization shown in Equation (17).

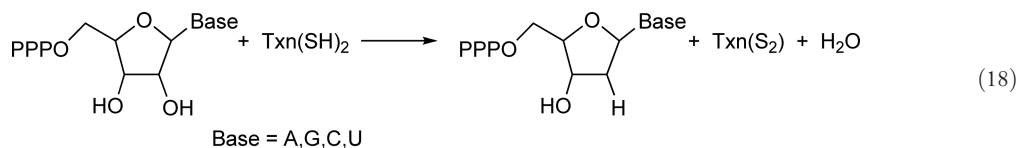


Like glutamate, 2-methyleneglutarate can be fermented in *Clostridia*, albeit by a different pathway,<sup>163</sup> and the reaction of MGM is the first step. Less information about MGM is available than other coenzyme B<sub>12</sub>-dependent enzymes. The stereochemistry, purification, and molecular characterization are available.<sup>166,167</sup> A substrate-based free radical has been observed by EPR spectroscopy.<sup>168</sup> The reaction mechanism is not known, although it likely follows the generic pathway of hydrogen abstraction by the 5'-deoxyadenosyl radical to form a substrate-derived radical, which rearranges to a product-related radical that is quenched by hydrogen transfer from 5'-deoxyadenosine. The mechanism of radical isomerization remains to be proven.<sup>169</sup>

### 7.14.4.6 Ribonucleoside Triphosphate Reductase

#### 7.14.4.6.1 Discovery and properties

The adenosylcobalamin-dependent RTPR catalyzes the reduction of ribonucleoside triphosphates (NTPs) to deoxyribonucleoside triphosphates (dNTPs) for DNA biosynthesis according to Equation (18).



The natural reducing agent thioredoxin is recycled by thioredoxin reductase and NADPH. RTPR also accepts dihydrolipoate or dithiothreitol as reductants. In addition to an NTP, adenosylcobalamin, and reducing agent, the action of RTPR in Equation (18) requires an allosteric activator, a dNTP. The allosteric activation profile is complex,<sup>170</sup> with the reduction of a particular NTP being activated by a complementary dNTP. However, studies are facilitated by the use of 2'-deoxyguanosine-5'-triphosphate (dGTP) as a universal activator.

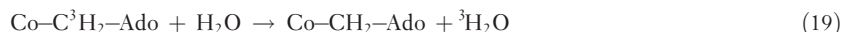
The reaction of RTPR is fundamentally different from other coenzyme B<sub>12</sub>-dependent reactions, in that it is a redox process proceeding with net reduction of the substrate, whereas the others are isomerases or isomerase/lyases. Moreover, the reduction of C2'(OH) to C2'(H) proceeds with incorporation of solvent hydrogen, unlike the other coenzyme B<sub>12</sub> reactions. RTPR shares in common with other coenzyme B<sub>12</sub>-dependent reactions the role of adenosylcobalamin, which initiates the radical mechanism. It might also share with DDH the mechanism by which C2'(OH) is eliminated.

Ribonucleotide reductases are grouped in three classes, I, II, and III, discovered in that order and differing in coenzyme requirements. They have been extensively reviewed.<sup>171-175</sup> The class I reductases have a tyrosyl radical, as the radical initiator, and a di-iron cofactor that generates the tyrosyl radical in an O<sub>2</sub>-dependent process. In class II reductases, adenosylcobalamin is the radical initiator. The class III reductases are glycy radical enzymes that use the glycy radical as the initiator. In each case, the radical initiator abstracts a hydrogen atom from a cysteinyl-SH group in the polypeptide chain to generate the cysteinyl-thiyl radical, which then reacts directly with the substrate.

RTPR from *Lactobacillus leichmanii* is the most intensively studied class II reductase and the one originally discovered and characterized.<sup>12,176</sup> It comprises a single 81.9 kDa polypeptide chain of 739 amino acids.<sup>177,178</sup> RTPR was the first enzyme displaying allosteric activation in a single chain to be discovered. While RTPR exhibits allosteric activation, it does not display cooperative kinetics. A structure is available showing the essential Cys408 nearby the cobalamin.<sup>179</sup>

### 7.14.4.6.2 Reaction mechanism

A central role for adenosylcobalamin was demonstrated in the 1960s by the observation of RTPR-catalyzed exchange of tritium in  $[5' \text{-}^3\text{H}]$ adenosylcobalamin with protons of water, according to Equation (19), in the absence of a substrate and in the presence of the allosteric activator dGTP.<sup>180,181</sup>



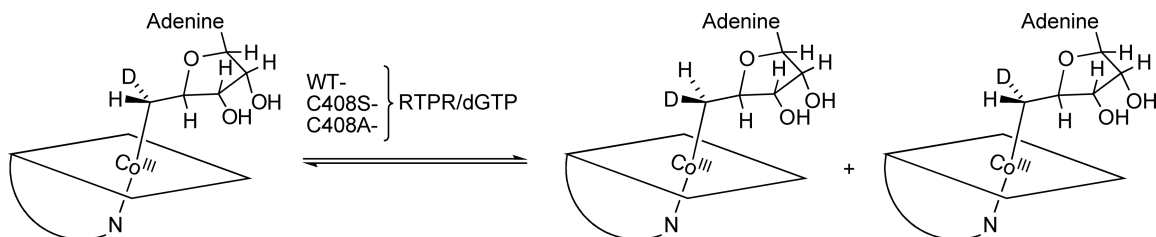
Exchange required cleavage of the Co–C5 bond, and this process required allosteric activation. Cleavage of this bond to cob(II)alamin was observed by EPR spectroscopy.<sup>182</sup>

Class I and II ribonucleotide reductases were discovered in the 1960s, but no rational chemical mechanism for the displacement of C2'(OH) by a hydrogen could be written until the 1980s. The first mechanistic clue was the observation of small but significant primary tritium kinetic isotope effects in the reductions of 3'-tritiated ribonucleotides by class I and II reductases.<sup>183,184</sup> The isotope effects proved that the mechanism had to involve transient cleavage of the C3'–H bond, as in the action of DDH in which the C–H bond adjacent to the departing OH must be broken. The 3'-tritium kinetic isotope effects established a mechanistic link to the DDH mechanism. More detailed studies of class II RTPR showed that, while adenosylcobalamin was required, there was no tritium exchange between 3'-tritiated NTPs and the coenzyme.<sup>185</sup> Experiments under single-turnover conditions proved that adenosylcobalamin could not be engaged in mediating hydrogen transfer but was somehow involved in initiating the reaction.

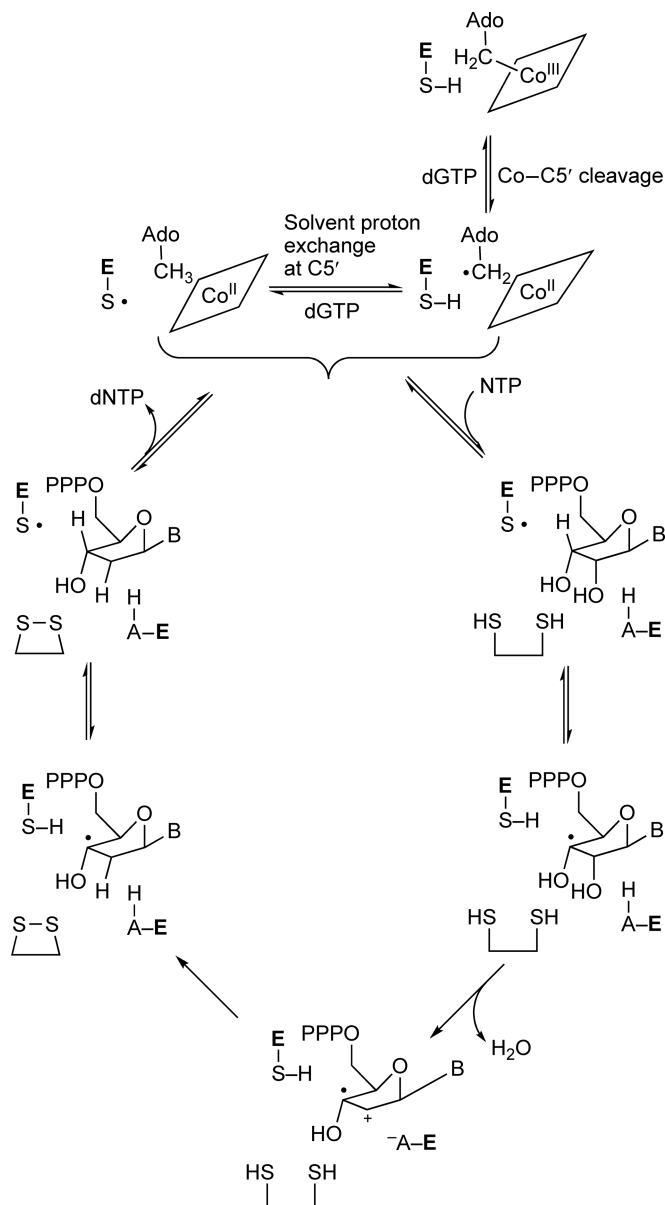
The observation of a kinetically competent cysteinyl-thiyl free radical intermediate implicated Cys408 in a critical mechanistic role.<sup>186</sup> As in other coenzyme B<sub>12</sub>-dependent enzymes, the unpaired electron of the radical was spin coupled to Co<sup>2+</sup> of cob(II)alamin. The structure of the thiyl radical was rigorously proven by isotope-edited EPR analysis, and spectral simulations showed that the thiyl radical resided at 5–7 Å from Co<sup>2+</sup> of cob(II)alamin. The emergence of the radical triplet signal required adenosylcobalamin and dGTP in addition to RTPR.

In addition to ribonucleotide reduction and exchange of tritium between water and the 5'-methylene group of adenosylcobalamin, RTPR also catalyzes the C5'-epimerization of (5'R)-[5'-<sup>2</sup>H<sub>1</sub>]adenosylcobalamin according to **Figure 30**.<sup>187</sup> Epimerization requires the presence of the allosteric activator dGTP and does not occur in its absence. Epimerization requires reversible cleavage of the Co–C5' bond. Mutation of the active site Cys408 completely inactivates RTPR as a catalyst for reduction of NTPs and as a catalyst for exchange of tritium in [5'-<sup>3</sup>H]adenosylcobalamin with water. However, the active site variants C408S- and C408A-RTPRs retain dGTP-dependent C5'-epimerase activity ( $k_{\text{cat}} = 5, 0.3, \text{ and } 0.4 \text{ min}^{-1}$  for wt-, C408A-, and C408S-RTPR, respectively). It seems clear that the allosteric activator dGTP stimulates Co–C5' cleavage, and this does not require the active site Cys408, which becomes the thiyl radical after abstraction of the hydrogen atom from the 3-SH group by the 5'-deoxyadenosyl radical.

Available information is consistent with the mechanism shown in **Figure 31**. Generation of the Cys408-thiyl radical at the top of the scheme proceeds by cleavage of the Co–C5' bond and is allosterically activated by dGTP. This step accounts for epimerization of (5'R)-[5'-<sup>2</sup>H<sub>1</sub>]adenosylcobalamin. Reversible abstraction of the thiol hydrogen atom from Cys408 generates the thiyl radical and accounts for solvent proton exchange at C5'.



**Figure 30** RTPR catalysis of epimerization at C5' of 5'-monodeuterioadenosylcobalamin. The isotopomers of [5'-<sup>2</sup>H<sub>1</sub>]adenosylcobalamin can be distinguished by chemical shift values and 4'–5' coupling constants. Starting with (5R)-[5'-<sup>2</sup>H<sub>1</sub>]adenosylcobalamin, and in the absence of a substrate, RTPR catalyzes epimerization to the two isotopomers in the presence but not the absence of the allosteric activator dGTP. The inactive, mutated C408S- and C408A-RTPRs also catalyze the dGTP-dependent epimerization.



**Figure 31** A chemical mechanism for the reactions of RTPR. The steps in this mechanism are discussed in the text. The upper steps in the absence of a substrate explain solvent hydrogen and epimerization at C5' of adenosylcobalamin and require the allosteric activator dGTP. The lower catalytic cycle shows steps in a mechanism for reduction of C2'(OH) of substrates.

Binding of NTP allows the thiyl radical access to substrate C3'(H), which it abstracts to initiate reduction. Elimination of water by the 3'-radical, followed by or concurrent with reduction by an enzymatic dithiol, leads to the 3'-radical of the product dNTP. Abstraction of the thiol hydrogen atom from Cys408 by the product radical regenerates the thiyl radical and produces dNTP. The radical cationic species following dehydration shown in **Figure 31** is a hypothetical intermediate.

Class II RTPR has in common with other coenzyme B<sub>12</sub>-dependent enzymes the use of the 5'-deoxyadenosyl radical from adenosylcobalamin as a radical initiator. The elimination of water from the 3'-radical intermediate might in part be analogous to the elimination of water in the reaction of DDH. The variant function of RTPR can be viewed as foreshadowing the wider functions of the 5'-deoxyadenosyl radical in other radical reactions, in particular the radical SAM enzymes.<sup>116</sup>



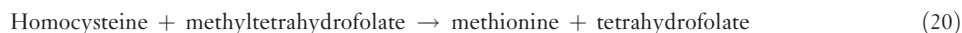
## 7.14.5 Methylcobalamin in Enzymatic Catalysis

A number of enzymes engaged in methyl transfer reactions are purified as cobalamin-binding proteins or enzymes that can use methylcobalamin as a substrate. In some and perhaps all of these reactions, methylcobalamin serves as a methyl group donor. These reactions are unlike the generic SAM-dependent methyltransferases; yet SAM often plays essential and diverse roles in the reactions of cobalamin-dependent methyltransferases. Enzymes involved in methionine biosynthesis, methanogenesis, acetogenesis, and methylation of non-nucleophilic centers are among those known or thought to require methylcobalamin in mediating methyl transfer.<sup>47,48,116,188</sup> This is a large and rapidly developing field that holds surprises for the future. A complete discussion is beyond the scope of this chapter, and focus will be directed to the oldest and best researched example, cobalamin-dependent methionine synthase, the function of which might foreshadow the roles of methylcobalamin in other methyltransferases.

Two mechanistic aspects of the cobalamin-dependent methyltransferases distinguish them from adenosylcobalamin-dependent enzymes. No methylcobalamin synthetases are known. Methylcobalamin is produced at enzymatic sites as an intermediate in the primary overall reaction. Second, methylcobalamin does not function as a coenzyme in the sense that it assists in catalysis; to date, it is instead a catalytic intermediate in its enzymatic reactions.

### 7.14.5.1 Methionine Synthase

There are two classes of methionine synthases, the cobalamin-dependent and cobalamin-independent enzymes. Both synthases catalyze the same reaction, Equation (20), the methylation of (*S*)-homocysteine (homocysteine) by *N*<sup>5</sup>-methyltetrahydrofolate (methyltetrahydrofolate).



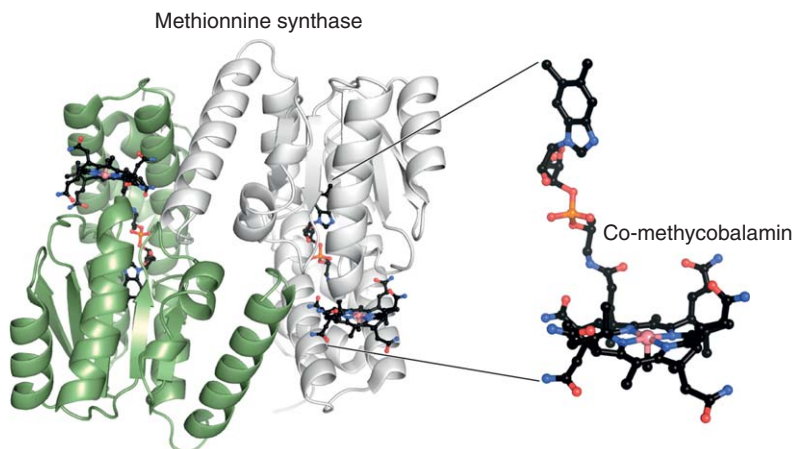
The reaction as written can appear to be a straightforward transmethylation; however, there are chemical complications because of the low reactivity of methyltetrahydrofolate as an alkylating agent. A historical perspective on cobalamin-dependent methionine synthase is available.<sup>47</sup> The cobalamin-dependent methionine synthase is the ‘other’ human (mammalian) cobalamin enzyme, the methylcobalamin-dependent counterpart to the adenosylcobalamin-dependent mammalian MCM.

#### 7.14.5.1.1 Molecular properties

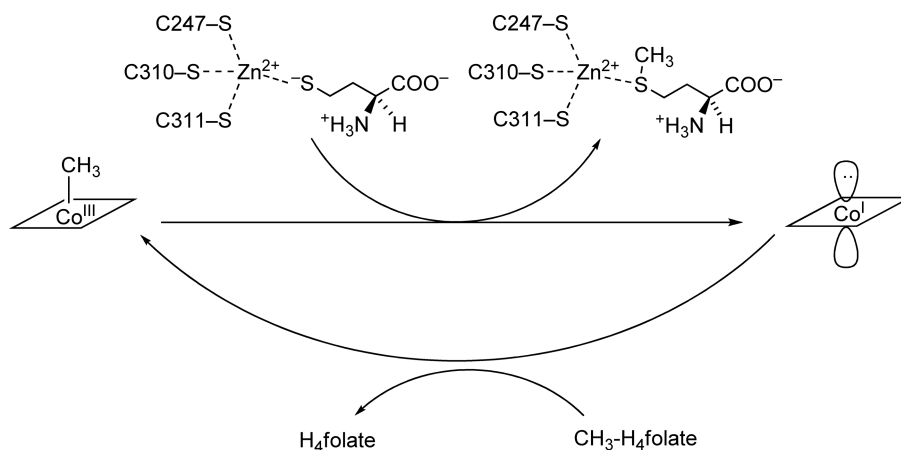
The most detailed structural and functional research to date is that on the cobalamin-dependent enzyme from *E. coli*, MetH, which is similar to the mammalian enzyme.<sup>47</sup> The MetH purified from the cloned gene is a very large monomeric protein of 136 kDa.<sup>189</sup> Biochemical and structural studies reveal that the protein comprises four major domains known as modules, which are programmed to carry out specific catalytic and regulatory functions.<sup>190,191</sup> The modules are the SAM-binding domain (MetH<sup>897–1227</sup>), cobalamin (Cob)-binding domain (MetH<sup>650–896</sup>), the homocysteine (Hcys)-binding domain (MetH<sup>2–353</sup>), and the Hcys- and tetrahydrofolate (Fol)-binding fragment (MetH<sup>2–649</sup>). A crystal structure of intact MetH is not available, but crystal structures of domains are available.<sup>192–194</sup> A structure of Cob is shown in **Figure 32**, showing the ‘base-off’ binding of cobalamin first found in this structure and later in MCM, GM, and 5,6-LAM.<sup>120,129,161</sup> Adenosylcobalamin is bound with the 5,6-dimethylbenzimidazole in the ‘base-on’ mode in DDH, EAL, and RTPR.<sup>80,179,195</sup> The modules of MetH were obtained for kinetic, mechanistic, and structural investigation either by limited proteolysis of MetH or by bacterial expression of *metH* gene segments and purification of the module or fragment.<sup>196,197</sup>

#### 7.14.5.1.2 Reaction mechanism

The methylation mechanism in **Figure 33** describes the known events in the transformation of homocysteine into methionine. The Hcys module is a zinc metallodomain and binds homocysteine with sulfur ligated to Zn<sup>2+</sup>, as proven by careful analytical measurements.<sup>198,199</sup> Ligation to Zn<sup>2+</sup> lowers the p*K*<sub>a</sub> value of the sulfhydryl group in homocysteine enough to transform it into the highly nucleophilic thiolate anion, which



**Figure 32** Crystal structure of the cobalamin-binding domain in MetH. A structure of the cobalamin-binding module Cob of MetH from *Escherichia coli* is shown with cobalamin bound in the ‘base-off’ mode. This binding mode was first observed in the structure shown here. The module crystallizes as a dimer (green and gray). This illustration was generated from PDF 1bmt.<sup>192</sup>



**Figure 33** Role of  $Zn^{2+}$  in methyl transfer catalyzed by Meth. The methyl-accepting substrate homocysteine is ligated to  $Zn^{2+}$  in the HCys module of MetH. The other ligands to  $Zn^{2+}$  are the cysteine residues indicated in the figure.

is an excellent methyl acceptor. Cob(I)alamin mediates methyl transfer from methyltetrahydrofolate to homocysteine through the intermediate formation of methylcobalamin. Cob(I)alamin at the active site is the kinetically competent methyl acceptor from methyltetrahydrofolate and methyl donor to homocysteine.<sup>200</sup>

Evidence relating to the exact roles of SAM and methylcobalamin in the action of MetH were confusing for many years. It was clear that the generic methyl donor SAM was required for activity, but methyltetrahydrofolate was the better methyl donor to homocysteine. The first evidence for the participation of cob(I)alamin came with the observation that propyl iodide inactivated the enzyme by alkylation of cob(I)alamin to propylcobalamin.<sup>201</sup> The propylated enzyme was reactivated by light. A confusing aspect was that the propylated enzyme could still catalyze methyl transfer from methylcobalamin to homocysteine.

#### 7.14.5.1.3 Modular function

Detailed investigations of the functions and interactions of MetH modules clarify the early confusing results. SAM and the SAM module maintain cobalamin in the Cob module as methylcobalamin. This is an important function. The reaction in **Figure 33** begins with methyl transfer from methylcobalamin to homocysteine to

form methionine and cob(I)alamin. As long as cob(I)alamin is remethylated by methyltetrahydrofolate, the reaction proceeds. However, cob(I)alamin is extremely reactive with oxidizing agents and is occasionally trapped as cob(II)alamin, which cannot be methylated by methyltetrahydrofolate. This event calls the SAM module into play to remethylate it to methylcobalamin. This accounts for activation by SAM. Methylcobalamin can serve as methyl donor by interacting directly with the HCys module, which normally binds to the Cob module to methylate homocysteine.

Modular function follows a definite course through four structural states involving the four mobile modules and a 'helical cap'. The center of action is the Cob module. In state 1, the helical cap covers cob(I)alamin in Cob, protecting it from oxidation and preventing its methylation by SAM and the SAM module. The cap remains in place whenever cob(I)alamin is in Cob and no other module is bound to Cob. In state 2, the helical cap is displaced by the HCys module, which then catalyzes methylation of homocysteine. In state 3, the HCys module is displaced from Cob by the Fol module, which catalyzes the methylation of cob(I)alamin by methyltetrahydrofolate. The system cycles merrily through states 1, 2, and 3 to carry out the steps in the mechanism of **Figure 33**. However, adventitious oxidation of cob(I)alamin to cob(II)alamin signals the SAM module to bind to the Cob module, and in cooperation with reduced flavodoxin it catalyzes the reduction of cob(II)alamin to cob(I)alamin and methylation by SAM to reactivate the enzyme. The enzyme then returns to state 2 and continues cycling productively.

The modular function of MetH is an elegant example of signaling in a closed system. It rationalizes all of the formerly confusing information about MetH. Other SAM and cobalamin-dependent methylases, such as the radical SAM methylases, might function in analogous ways.

### 7.14.6 Conclusion

The elegant structure of adenosylcobalamin, with all of its stereochemistry and with its long biosynthetic pathway, led investigators to expect elaborate chemical functions by cobalt and the cobalamin macrocycle in the coenzyme B<sub>12</sub>-dependent reactions. Many hypothetical functions for the macrocycle or/and for cobalt were put forward. None of them have survived experimental testing. It has become clear that the sole function of the elegant cobalamin structure has been the provision of a water-soluble ligand system that could make the Co–C bond sufficiently stable to survive in the cell, while being susceptible to reversible cleavage through binding interactions with cognate enzymes. The resultant 5'-deoxyadenosyl radical has been found to do the chemical work of initiating radical mechanisms, with the cobalamin system limited to spectator status.

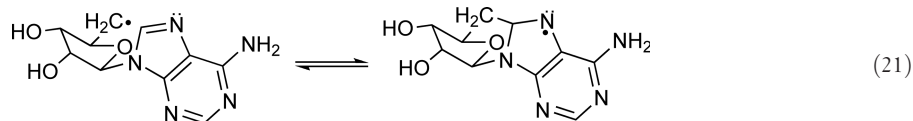
It is likely that specific cobalamin–protein and adenosyl–protein interactions serve to weaken the Co–C bond, through stabilization of the cleaved state, thereby facilitating formation of the 5'-deoxyadenosyl radical. A clear understanding of this process will require much more structural work in the future.

For many years, adenosylcobalamin was regarded as 'the' biological source of the 5'-deoxyadenosyl radical and the most important nonoxygen source of free radicals in biocatalysis. Given the long biosynthetic pathway to adenosylcobalamin, it seemed unlikely that the unique radical chemistry of coenzyme B<sub>12</sub> could have appeared by incremental evolution, guided solely by the survival benefits of 5'-deoxyadenosyl-5'-yl-initiated free radical chemistry in the anaerobic world. It seemed that there should have been a predecessor to adenosylcobalamin that would provide the 5'-deoxyadenosyl radical and serve as the chemical context within which adenosylcobalamin could evolve as a superior coenzyme. It seemed that a handful of SAM-dependent enzymes such as lysine 2,3-aminomutase could be the remnants of such a primordial world of radical chemistry.<sup>202</sup>

As the field of SAM-dependent radical mechanisms developed, the radical SAM superfamily became recognized.<sup>116,203</sup> The radical SAM enzymes were found to employ a different chemical strategy for generating the 5'-deoxyadenosyl, reductive cleavage of SAM by the unique [4Fe–4S] clusters in members of this superfamily. To date, more than 2800 members of this superfamily have been found, and more than 40 distinct biological reactions have been attributed to them.<sup>116</sup>

The increasing ubiquity of the 5'-deoxyadenosyl radical as an initiator of radical mechanisms in enzymology inspires the question whether it possesses special properties that allow it, a primary free radical of high energy, to serve in this capacity. In this connection, it seems worthwhile to consider whether it might be

internally stabilized. One could envision stabilization by delocalization of the unpaired electron through reversible, internal cyclization to C8 of the adenine ring, as in Equation (21).



In the cyclic form, the electron is widely delocalized through the adenine ring. This cyclic radical should be one intermediate in the formation of 5,8-*cyclo*-adenosine upon photolysis of adenosylcobalamin. Reversible formation of this species could allow C5' in the open form to serve as the hydrogen abstracting species in catalysis.

That the radical SAM enzymes greatly outnumber the adenosylcobalamin enzymes in no way diminishes the elegance of the chemistry carried out by the two families of radical catalysts. Much more radical chemistry remains to be discovered.

The roles of methylcobalamin in biological processes are still being discovered and are bound to become increasingly diverse. The story of methionine synthase and MetH has been enlightening and elegant. However, it is very likely the beginning of the story of methylcobalamin. The other methylcobalamin enzymes are involved in many different chemical processes, from methylation of non-nucleophilic atoms to reduction in methanogenesis. The elucidation of the roles of methylcobalamin will usher in new chemistry for the field of vitamin B<sub>12</sub>.

## Acknowledgment

The author is grateful to the National Institute of Diabetes and Kidney and Digestive Diseases of the USPHS for its continuous support of his research. The author thanks H. Adam Steinberg for generating the molecular graphics in [Figures 10, 19, 23, 28, and 31](#).

## Abbreviations

<b>CoA</b>	coenzyme A
<b>DDH</b>	diolehydrase, propane-1,2-diol hydro-lyase
<b>dGTP</b>	2'-deoxyguanosine-5'-triphosphate
<b>dNTP</b>	2'-deoxyribonucleoside-5'-triphosphate
<b>EAL</b>	ethanolamine ammonia-lyase
<b>ENDOR</b>	electron nuclear double resonance
<b>EPR</b>	electron paramagnetic resonance
<b>ESEEM</b>	electron spin envelope modulation
<b>EutB</b>	large subunit of EAL
<b>GM</b>	glutamate mutase
<b>HEH</b>	hydroxyethylhydrazine
<b>HPLC</b>	high-performance liquid chromatography
<b>5,6-LAM</b>	lysine 5,6-aminomutase
<b>MCM</b>	methylmalonyl-CoA mutase
<b>MGM</b>	methyleneglutarate mutase
<b>NMR</b>	nuclear magnetic resonance
<b>NTP</b>	ribonucleoside-5'-triphosphate
<b>PLP</b>	pyridoxal-5'-phosphate
<b>RTPR</b>	ribonucleoside-5'-triphosphate reductase
<b>SAM</b>	S-adenosyl-L-methionine
<b>TIM</b>	triose phosphate isomerase

**Nomenclature**

<b>adenosylcobalamin</b>	5'-deoxyadenosylcobalamin, coenzyme B <sub>12</sub>
<b>aquocobalamin</b>	cob(III)alamin hydrate, vitamin B <sub>12a</sub>
<b>Arg</b>	arginine
<b>Asp</b>	aspartate
<b>cyanocobalamin</b>	cob(III)alamin cyanide, vitamin B <sub>12</sub>
<b>Cys</b>	cysteine
<b>EutB</b>	large subunit of EAL
<b>Glu</b>	glutamate
<b>Hcys</b>	homocysteine
<b>His</b>	histidine
<b>hydroxocobalamin</b>	cob(III)alamin hydroxide, vitamin B <sub>12b</sub>
<b>Lys</b>	lysine
<b>MethH</b>	cobalamin-dependent methionine synthase from <i>E. coli</i>
<b>methyltetrahydrofolate</b>	N <sup>5</sup> -methyltetrahydrofolate
<b>MutE</b>	larger subunit of GM
<b>MutS</b>	smaller subunit of GM
<b>Txn(S<sub>2</sub>)</b>	oxidized thioredoxin
<b>Txn(SH)<sub>2</sub></b>	reduced thioredoxin
<b>Tyr</b>	tyrosine
<b>vitamin B<sub>12r</sub></b>	cob(II)alamin
<b>vitamin B<sub>12s</sub></b>	cob(I)alamin

**References**

1. P. G. Lenhart; D. C. Hodgkin, *Nature* **1961**, 192, 937–938.
2. H. A. Barker; R. D. Smyth; H. Weissbach; J. L. Toohey; J. N. Ladd; B. E. Volcani, *J. Biol. Chem.* **1960**, 235, 480–488.
3. S. Gurani; S. P. Mistry; B. C. Johnson, *Biochim. Biophys. Acta* **1960**, 38, 187.
4. J. R. Stern; D. L. Friedman, *Biochem. Biophys. Res. Commun.* **1960**, 2, 82.
5. R. H. Abeles; H. A. Lee, Jr., *J. Biol. Chem.* **1961**, 236, 2347–2350.
6. R. H. Abeles; A. M. Brownstein; C. H. Randles, *Biochim. Biophys. Acta* **1960**, 41, 530–531.
7. K. L. Smiley; M. Sobolov, *Arch. Biochem. Biophys.* **1962**, 97, 538–543.
8. A. M. Brownstein; R. H. Abeles, *J. Biol. Chem.* **1961**, 236, 1199.
9. P. Overath; G. M. Kellerman; F. Lynen; H. P. Fritz; H. J. Keller, *Biochem. Z.* **1962**, 335, 500–518.
10. H. A. Barker; V. Rooze; F. Suzuki; A. A. Iodice, *J. Biol. Chem.* **1964**, 239, 3260–3266.
11. J. D. Erfle; J. M. Clark, Jr.; R. F. Nystrom; B. C. Johnson, *J. Biol. Chem.* **1964**, 239, 1920–1924.
12. R. L. Blakely; H. A. Barker, *Biochem. Biophys. Res. Commun.* **1964**, 16, 391–397.
13. M. Sprecher; M. J. Clark; D. B. Sprinson, *J. Biol. Chem.* **1966**, 241, 872–877.
14. M. Sprecher; R. L. Switzer; D. B. Sprinson, *J. Biol. Chem.* **1966**, 241, 864–867.
15. H. Eggerer; P. Overath; F. Lynen; E. R. Stadtman, *J. Am. Chem. Soc.* **1960**, 82, 2643–2644.
16. C. S. Hegre; S. J. Miller; M. D. Lane, *Biochim. Biophys. Acta* **1962**, 56, 538–544.
17. R. W. Kellermeyer; H. G. Wood, *Biochemistry* **1962**, 1, 1124–1131.
18. H. Eggerer; E. R. Stadtman; P. Overath; F. Lynen, *Biochem. Z.* **1960**, 333, 1–9.
19. H. W. Whitlock, *Ann. N. Y. Acad. Sci.* **1964**, 112, 721–727.
20. L. L. Ingraham, *Ann. N. Y. Acad. Sci.* **1964**, 112, 713–720.
21. R. H. Abeles; H. A. Lee, Jr., *Ann. N. Y. Acad. Sci.* **1964**, 112, 695–702.
22. O. W. Wagner; H. A. Lee, Jr.; P. A. Frey; R. H. Abeles, *J. Biol. Chem.* **1966**, 241, 1751–1762.
23. K.-H. Tang; C. H. Chang; P. A. Frey, *Biochemistry* **2001**, 40, 5190–5199.
24. J. Halpern, *Science* **1985**, 227, 869–875.
25. S. O. Mansoorabadi; R. Padmakumar; N. Fazliddinova; M. Vlasie; R. Banerjee; G. H. Reed, *Biochemistry* **2005**, 44, 3153–3158.
26. J. M. Buchanan; H. L. Elford; R. E. Loughlin; B. M. McDougall; S. Rosenthal, *Ann. N. Y. Acad. Sci.* **1964**, 112, 756–773.
27. J. R. Guest; S. Friedman; M. J. Dilworth; D. D. Woods, *Ann. N. Y. Acad. Sci.* **1964**, 112, 774–790.
28. H. W. Dickerman; B. G. Redfield; J. G. Bieri; H. Weissbach, *Ann. N. Y. Acad. Sci.* **1964**, 112, 791–798.
29. R. E. Loughlin; H. L. Elford; J. M. Buchanan, *J. Biol. Chem.* **1964**, 239, 2888–2895.
30. H. Weissbach; B. G. Redfield; H. Dickerman; N. Brot, *J. Biol. Chem.* **1965**, 240, 856–862.

31. D. Lexa; J. M. Savéant, *Acc. Chem. Res.* **1983**, *16*, 235–243.
32. B. Kräutler, B<sub>12</sub> Electrochemistry and Organometallic Electrochemical Synthesis. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; Wiley-Interscience: New York, 1999; pp 315–339.
33. G. N. Schrauzer; E. Deutsch, *J. Am. Chem. Soc.* **1969**, *91*, 3341–3350.
34. B. D. Martin; R. G. Finke, *J. Am. Chem. Soc.* **1990**, *112*, 2419–2420.
35. L. Luo; G. Li; H. Chen; S. W. Fu; S. Y. Zhang, *J. Chem. Soc. Dalton Trans.* **1998**, 2103–2107.
36. B. P. Hay; R. G. Finke, *J. Am. Chem. Soc.* **1986**, *108*, 4820–4829.
37. H. Chen; G. Li; F. F. Zhang; L. Sun; H. L. Chen; S. Y. Zhang, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2003**, *59*, 2767–2774.
38. K. L. Brown; D. R. Evans; S. Chang; D. N. Jacobsen, *Inorg. Chem.* **1996**, *35*, 217–222.
39. K. L. Brown; D. R. Brooks; D. Behnke; D. N. Jacobsen, *J. Biol. Chem.* **1991**, *266*, 6737–6741.
40. O. Th. Magnusson; P. A. Frey, *J. Am. Chem. Soc.* **2000**, *122*, 8807–8813.
41. R. G. Finke; B. P. Hay, *Inorg. Chem.* **1984**, *23*, 3041–3043.
42. J. Halpern; S.-H. Kim; T. W. Leung, *J. Am. Chem. Soc.* **1984**, *106*, 8317–8319.
43. O. Th. Magnusson; G. H. Reed; P. A. Frey, *Biochemistry* **2001**, *40*, 7773–7782.
44. S. O. Mansoorabadi; O. Th. Magnusson; R. R. Poyner; P. A. Frey; G. H. Reed, *Biochemistry* **2006**, *45*, 14362–14370.
45. H. P. C. Hogenkamp; H. A. Barker, *J. Biol. Chem.* **1961**, *236*, 3097–3101.
46. R. Barnett; H. P. Hogenkamp; R. H. Abeles, *J. Biol. Chem.* **1966**, *241*, 1483–1486.
47. R. G. Matthews, Cobalamin-Dependent Methionine Synthase. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; Wiley-Interscience: New York, 1999; pp 681–706.
48. S. W. Ragsdale, The Acetogenic Corrinoid Proteins. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; Wiley-Interscience: New York, 1999; pp 633–653.
49. H. P. C. Hogenkamp, *J. Biol. Chem.* **1963**, *238*, 477–480.
50. H. P. C. Hogenkamp; J. N. Ladd; H. A. Barker, *J. Biol. Chem.* **1962**, *237*, 1950–1952.
51. P. A. Schwartz; P. A. Frey, *Biochemistry* **2007**, *46*, 7284–7292.
52. P. A. Frey; G. L. Karabatsos; R. H. Abeles, *Biochem. Biophys. Res. Commun.* **1965**, *18*, 551–556.
53. B. Zagalak; P. A. Frey; G. L. Karabatsos; R. H. Abeles, *J. Biol. Chem.* **1966**, *241*, 3028–3035.
54. G. J. Karabatsos; J. S. Fleming; N. Hsi; R. H. Abeles, *J. Am. Chem. Soc.* **1966**, *88*, 849–850.
55. J. Rétey; A. Umani-Ronchi; J. Sebl; D. Arigoni, *Experientia* **1966**, *22*, 502–503.
56. A. Abend; V. Bandarian; G. H. Reed; P. A. Frey, *Biochemistry* **2000**, *39*, 6250–6257.
57. G. M. Sandala; D. M. Smith; M. L. Cooate; L. Radom, *J. Am. Chem. Soc.* **2004**, *126*, 12206–12207.
58. P. Schwartz; R. LoBrutto; G. H. Reed; P. A. Frey, *Helv. Chim. Acta* **2003**, *86*, 3764–3775.
59. J. Valinsky; R. H. Abeles; A. S. Mildvan, *J. Biol. Chem.* **1974**, *249*, 2751–2755.
60. T. H. Finlay; J. Valinsky; K. Sato; R. H. Abeles, *J. Biol. Chem.* **1972**, *247*, 4197–4207.
61. T. H. Finlay; J. Valinsky; A. S. Mildvan; R. H. Abeles, *J. Biol. Chem.* **1973**, *248*, 1285–1290.
62. P. A. Frey; R. H. Abeles, *J. Biol. Chem.* **1966**, *241*, 2732–2733.
63. P. A. Frey; M. K. Essenberg; R. H. Abeles, *J. Biol. Chem.* **1967**, *242*, 5369–5377.
64. M. K. Essenberg; P. A. Frey; R. H. Abeles, *J. Am. Chem. Soc.* **1971**, *93*, 1242–1251.
65. B. Babior; D. C. Gould, *Biochem. Biophys. Res. Commun.* **1969**, *34*, 441–447.
66. S. A. Cockle; H. A. O. Hill; R. J. P. Williams; J. P. Davies; M. A. Foster, *J. Am. Chem. Soc.* **1972**, *94*, 275–277.
67. K. L. Schepler; W. R. Dunham; R. H. Sands; J. A. Fee; R. H. Abeles, *Biochim. Biophys. Acta* **1975**, *397*, 510–518.
68. J. F. Boas; P. R. Hicks; J. R. Pilbrow; T. D. Smith, *J. Chem. Soc. Faraday Trans. 2* **1978**, *74*, 417–431.
69. J. E. Valinsky; R. H. Abeles; J. A. Fee, *J. Am. Chem. Soc.* **1974**, *96*, 4709–4710.
70. T. Toraya, *Chem. Rev.* **2003**, *103*, 2095–2127.
71. M. Yamanishi; H. Murakami; T. Toraya, *Biochemistry* **2005**, *44*, 2113–2118.
72. P. A. Schwartz; R. LoBrutto; G. H. Reed; P. A. Frey, *Protein Sci.* **2007**, *16*, 1157–1164.
73. W. Buckel, *FEBS Lett.* **1996**, *389*, 20–24.
74. J. Kim; D. J. Darley; W. Buckel; A. J. Pierik, *Nature* **2008**, *452*, 239–242.
75. O. Th. Magnusson; P. A. Frey, *Biochemistry* **2002**, *41*, 1695–1702.
76. T. Tobimatsu; M. Azuma; S. Hayashi; K. Nishimoto; T. Toraya, *Biosci. Biotechnol. Biochem.* **1998**, *62*, 1774–1777.
77. N. Shibata; J. Masuda; T. Toraya; Y. Morimoto; N. N. Yasuoka, *Structure* **2000**, *8*, 775–788.
78. M. Kawata; K. Kinoshita; S. Takahashi; K. Ogura; N. Komoto; M. Yamanishi; T. Tobimatsu; T. Toraya, *J. Biol. Chem.* **2006**, *281*, 18327–18334.
79. P. A. Schwartz; P. A. Frey, *Biochemistry* **2007**, *46*, 7293–7301.
80. N. Shibata; J. Masuda; T. Tobimatsu; T. Toraya; K. Suto; Y. Morimoto; N. Yasuoka, *Structure* **1999**, *7*, 997–1008.
81. T. Toraya; K. Yoshizawa; M. Eda; T. Yamabe, *J. Biochem.* **1999**, *126*, 650–654.
82. K. M. Bansal; M. Grätzel; A. Henglein; E. Janata, *J. Phys. Chem.* **1973**, *77*, 16–19.
83. C. Walling; R. A. Johnson, *J. Am. Chem. Soc.* **1975**, *97*, 2405.
84. R. Livingston; H. Zeldes, *J. Am. Chem. Soc.* **1966**, *88*, 4333–4336.
85. K. Kinoshita; M. Kawata; K. Ogura; A. Yamasaki; T. Watanabe; N. Komoto; N. Hieda; M. Yamanishi; T. Tobimatsu; T. Toraya, *Biochemistry* **2008**, *47*, 3162–3173.
86. T. Kamachi; T. Toraya; K. Yoshizawa, *J. Am. Chem. Soc.* **2004**, *126*, 16207–16216.
87. B. Golding; L. Radom, *J. Am. Chem. Soc.* **1976**, *98*, 6331–6338.
88. P. George; J. P. Glusker; C. W. Bock, *J. Am. Chem. Soc.* **1997**, *119*, 7065–7074.
89. P. George; P. E. M. Siegbahn; J. P. Glusker; C. W. Bock, *J. Phys. Chem. B* **1999**, *103*, 7531–7541.
90. D. M. Smith; B. T. Golding; L. Radom, *J. Am. Chem. Soc.* **1999**, *121*, 1383–1384.
91. B. M. Babior, *J. Biol. Chem.* **1969**, *244*, 449–456.
92. B. M. Babior, *Biochim. Biophys. Acta* **1968**, *167*, 456–458.
93. J. Rétey; C. J. Suckling; D. Arigoni; B. M. Babior, *J. Biol. Chem.* **1974**, *249*, 6359–6360.
94. P. Dizioli; H. Haas; J. Rétey; S. W. Graves; B. M. Babior, *Eur. J. Biochem.* **1980**, *106*, 211–224.



95. T. J. Carty; B. M. Babior; R. H. Abeles, *J. Biol. Chem.* **1974**, *249*, 1683–1688.
96. K. Sato; E. Krodel; B. M. Babior; R. H. Abeles, *J. Biol. Chem.* **1976**, *251*, 3734–3737.
97. B. M. Babior; T. J. Carty; R. H. Abeles, *J. Biol. Chem.* **1974**, *249*, 1689–1695.
98. V. Bandarian; G. H. Reed, *Biochemistry* **1999**, *38*, 12394–12402.
99. V. Bandarian; R. R. Poyner; G. H. Reed, *Biochemistry* **1999**, *38*, 12403–12407.
100. V. Bandarian; G. H. Reed, *Biochemistry* **2002**, *41*, 8580–8588.
101. J. M. Canfield; K. Warncke, *J. Phys. Chem. B.* **2005**, *109*, 3053–3064.
102. K. Warncke; A. S. Utada, *J. Am. Chem. Soc.* **2001**, *123*, 8564–8572.
103. B. M. Babior; T. H. Moss; W. H. Orme-Johnson; H. Beinert, *J. Biol. Chem.* **1974**, *249*, 4537–4544.
104. K. Warncke; J. M. Canfield, *J. Am. Chem. Soc.* **2004**, *126*, 5930–5931.
105. G. Bender; R. R. Poyner; G. H. Reed, *Biochemistry* **2008**, *47*, 11360–11366.
106. K. Warncke; J. C. Schmidt; S. C. Ke, *J. Am. Chem. Soc.* **2008**, *130*, 6055.
107. K. Warncke, *Biochemistry* **2005**, *44*, 3184–3193.
108. R. LoBrutto; V. Bandarian; O. Th. Magnusson; X. Chen; V. L. Schreamm; G. H. Reed, *Biochemistry* **2001**, *40*, 9–14.
109. J. Rétey, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 355–361.
110. L. Sun; K. Warncke, *Proteins* **2006**, *64*, 308–319.
111. L. Sun; O. A. Groover; J. M. Canfield; K. Warncke, *Biochemistry* **2008**, *47*, 5523–5535.
112. R. R. Poyner; M. A. Anderson; V. Bandarian; W. W. Cleland; G. H. Reed, *J. Am. Chem. Soc.* **2006**, *128*, 7120–7121.
113. G. H. Reed, Private communication, 2008.
114. T. C. Stadtman, *Adv. Enzymol.* **1973**, *28*, 413–448.
115. P. A. Frey; O. Th. Magnusson, *Chem. Rev.* **2003**, *103*, 2129–2148.
116. P. A. Frey; A. D. Hegeman; F. J. Ruzicka, *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 63–88.
117. C. G. Morley; T. C. Stadtman, *Biochemistry* **1971**, *10*, 2325–2329.
118. C. H. Chang; P. A. Frey, *J. Biol. Chem.* **2000**, *275*, 106–114.
119. K.-H. Tang; A. Harms; P. A. Frey, *Biochemistry* **2002**, *41*, 8767–8776.
120. F. Berkovitch; E. Behshad; K.-H. Tang; E. A. Enns; P. A. Frey; C. L. Drennan, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15870–15875.
121. K.-H. Tang; A. D. Casarez; W. Wu; P. A. Frey, *Arch. Biochem. Biophys.* **2003**, *418*, 49–54.
122. O. Han; P. A. Frey, *J. Am. Chem. Soc.* **1990**, *112*, 8982–8983.
123. G. M. Sandala; D. M. Smith; L. Radom, *J. Am. Chem. Soc.* **2006**, *128*, 16004–16005.
124. K.-H. Tang; S. O. Mansoorabadi; G. H. Reed; P. A. Frey, (2009), *Biochemistry*, In press.
125. G. J. Cardinale; R. H. Abeles, *Biochim. Biophys. Acta* **1967**, *132*, 517–518.
126. E. N. Marsh; N. McKie; N. K. Davis; P. F. Leadlay, *Biochem. J.* **1989**, *260*, 345–352.
127. N. McKie; N. H. Keep; M. L. Patchett; P. F. Leadlay, *Biochem. J.* **1990**, *269*, 293–298.
128. A. Ratnatilleke; J. W. Vrijbloed; J. A. Robinson, *J. Biol. Chem.* **1999**, *274*, 31679–31685.
129. F. Mancia; N. H. Keep; A. Nakagasawa; P. R. Leadlay; S. McSweeney; B. Rasmussen; P. Bosecke; O. Diat; P. R. Evans, *Structure* **1997**, *4*, 339–350.
130. F. Mancia; P. R. Evans, *Structure* **1998**, *6*, 711–720.
131. N. S. Lees; D. Chen; C. J. Walsby; E. Behshad; P. A. Frey; B. M. Hoffman, *J. Am. Chem. Soc.* **2006**, *128*, 10145–10154.
132. J. Rétey; E. H. Smith; B. Zagalak, *Eur. J. Biochem.* **1978**, *83*, 437–451.
133. S. Wollowitz; J. Halpern, *J. Am. Chem. Soc.* **1988**, *110*, 3112–3120.
134. Y. Zhao; A. Abend; M. Kunz; P. Such; J. Rétey, *Eur. J. Biochem.* **1994**, *225*, 891–896.
135. A. Abend; V. Ilich; J. Rétey, *Eur. J. Biochem.* **1997**, *249*, 180–186.
136. R. Padmakumar; R. Banerjee, *J. Biol. Chem.* **1995**, *270*, 9295–9300.
137. A. Dybala-Defratyka; P. Paneth; R. Banerjee; G. G. Truhlar, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10774–10779.
138. N. H. Thomä; T. W. Meier; P. R. Evans; P. F. Leadlay, *Biochemistry* **1998**, *37*, 14386–14393.
139. N. H. Thomä; P. R. Evans; P. F. Leadlay, *Biochemistry* **2000**, *39*, 9213–9221.
140. N. Maiti; L. Widjaja; R. Banerjee, *J. Biol. Chem.* **1999**, *274*, 32733–32737.
141. T. W. Meier; N. H. Thomä; P. F. Leadlay, *Biochemistry* **1996**, *35*, 11791–11796.
142. K. M. Doll; R. R. Bender; R. G. Finke, *J. Am. Chem. Soc.* **2003**, *125*, 10877–10884.
143. W. Buckel; C. Kratky; B. T. Golding, *Chemistry* **2005**, *12*, 352–362.
144. S. Chowdbury; R. Banerjee, *Biochemistry* **2000**, *39*, 7998–8006.
145. R. Padmakumar; R. Padmakumar; R. Banerjee, *Biochemistry* **1997**, *36*, 3713–3718.
146. A. Gaudemer; J. Zylber; N. Zylber; M. Baran-Marszac; W. E. Hull; M. Fountoulakis; A. König; K. Wölflle; J. Rétey, *Eur. J. Biochem.* **1981**, *119*, 279–285.
147. S. Chowdbury; M. G. Thomas; J. C. Escalante-Semerena; R. Banerjee, *J. Biol. Chem.* **2001**, *276*, 1015–1019.
148. A. J. Brooks; M. Vlasie; R. Banerjee; T. C. Brunold, *J. Am. Chem. Soc.* **2005**, *127*, 16522–16528.
149. H. A. Barker; H. Weissbach; R. D. Smyth, *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 1093–1097.
150. R. L. Switzer; B. G. Baltimore; H. A. Barker, *J. Biol. Chem.* **1969**, *244*, 5263–5288.
151. E. N. Marsh, *Biochemistry* **1995**, *34*, 7542–7547.
152. E. N. Marsh; D. P. Ballou, *Biochemistry* **1998**, *37*, 11864–11872.
153. O. Zelder; B. Beatrix; U. Leutcher; W. Buckel, *Eur. J. Biochem.* **1994**, *226*, 577–585.
154. H. Bothe; D. J. Darley; S. P. Albracht; G. J. Gerfen; B. T. Golding; W. Buckel, *Biochemistry* **1998**, *37*, 4105–4113.
155. A. J. Brooks; C. C. Fox; E. N. Marsh; M. Vlasie; R. Banerjee; T. C. Brunold, *Biochemistry* **2005**, *44*, 15167–15181.
156. I. Roymoulik; N. Moon; W. R. Dunham; D. P. Ballou; E. N. Marsh, *Biochemistry* **2000**, *39*, 10340–10346.
157. M. Yoon; A. Patwardhan; C. Oiao; S. O. Mansoorabadi; A. L. Menefee; G. H. Reed; E. N. Marsh, *Biochemistry* **2006**, *45*, 11650–11657.
158. I. Roymoulik; H. P. Chen; E. N. Marsh, *J. Biol. Chem.* **1999**, *274*, 11619–11622.
159. M. S. Huhta; D. Ciceri; B. T. Golding; E. N. Marsh, *Biochemistry* **2002**, *41*, 3200–3206.

160. D. E. Holloway; E. N. Marsh, *J. Biol. Chem.* **1994**, *269*, 20425–20430.
161. K. Gruber; R. Reitzer; C. Kratky, *Angew. Chem. Int. Ed.* **2001**, *40*, 3377–3380.
162. B. Beatrix; O. Zelder; F. K. Kroll; G. Örtlygsson; B. T. Golding; W. Buckel, *Angew. Chem. Int. Ed.* **1995**, *34*, 2398–2401.
163. W. Buckel; G. Bröker; H. Bothe; A. J. Pierik; B. T. Golding, Glutamate Mutase and 2-Methyleneglutarate Mutase. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; Wiley-Interscience: New York, 1999; pp 757–781.
164. H. W. Chih; E. N. Marsh, *Biochemistry* **1999**, *38*, 13684–13691.
165. S. D. Wetmore; D. M. Smith; B. T. Golding; L. Radom, *J. Am. Chem. Soc.* **2001**, *123*, 7963–7972.
166. C. Michael; G. Hartrampf; W. Buckel, *Eur. J. Biochem.* **1989**, *184*, 103–107.
167. G. Hartrampf; W. Buckel, *Eur. J. Biochem.* **1986**, *156*, 301–304.
168. A. J. Pierik; D. Ceceri; R. F. Lopez; F. Kroll; G. Bröker; B. Beatrix; W. Buckel; B. T. Golding, *Biochemistry* **2005**, *44*, 10541–10551.
169. M. Newcomb; N. Miranda, *J. Am. Chem. Soc.* **2003**, *125*, 4080–4086.
170. P. Reichard; A. Jordan, *Annu. Rev. Biochem.* **1998**, *98*, 72–98.
171. P. Nordlund; P. Reichard, *Annu. Rev. Biochem.* **2006**, *75*, 681–706.
172. J. Stubbe, *Curr. Opin. Struct. Biol.* **2000**, *10*, 731–736.
173. M. Sahlin; B. M. Sjöberg, *Subcell. Biochem.* **2000**, *35*, 405–443.
174. M. Fontecave, *Cell. Mol. Life Sci.* **1998**, *54*, 684–695.
175. M. Fontecave; M. Mulliez, Ribonucleotide Reductases. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; Wiley-Interscience: New York, 1999; pp 731–756.
176. R. L. Blakley; R. K. Ghambeer; P. F. Nixon; E. Vitols, *Biochem. Biophys. Res. Commun.* **1965**, *20*, 439–445.
177. D. Panagou; M. D. Orr; J. R. Dunstone; R. L. Blakley, *Biochemistry* **1972**, *11*, 2378–2388.
178. S. J. Booker; J. Stubbe, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8352–8356.
179. M. D. Sintchak; G. Arjara; B. A. Kellogg; J. Stubbe; C. L. Drennan, *Nat. Struct. Biol.* **2002**, *9*, 293–300.
180. W. S. Beck; R. H. Abeles; W. G. Robinson, *Biochem. Biophys. Res. Commun.* **1966**, *25*, 421–425.
181. H. P. C. Hogenkamp; R. K. Ghambeer; C. Brownson; R. L. Blakley; E. Vitols, *J. Biol. Chem.* **1968**, *243*, 799–808.
182. W. H. Orme-Johnson; H. Beinert; R. L. Blakley, *J. Biol. Chem.* **1974**, *249*, 2338–2343.
183. J. Stubbe; D. Ackles, *J. Biol. Chem.* **1980**, *255*, 8027–8030.
184. J. Stubbe; G. Smith; R. L. Blakley, *J. Biol. Chem.* **1983**, *258*, 1619–1624.
185. G. W. Ashley; G. Harris; J. Stubbe, *J. Biol. Chem.* **1986**, *261*, 3958–3964.
186. S. Licht; G. J. Gerfen; J. Stubbe, *Science* **1996**, *271*, 477–481.
187. D. Chen; A. Abend; J. Stubbe; P. A. Frey, *Biochemistry* **2003**, *42*, 4578–4584.
188. K. Sauer; R. K. Thauer, The Role of Corrinoids in Methanogenesis. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; Wiley-Interscience: New York, 1999; pp 655–679.
189. T. J. Drummond; R. R. Orgorzalek Loo; R. G. Matthews, *Biochemistry* **1993**, *32*, 9282–9289.
190. V. Bandarian; M. L. Ludwig; R. G. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 8156–8163.
191. J. C. Evans; D. P. Huddler; M. T. Hilgers; G. Romanchuk; R. G. Matthews; M. L. Ludwig, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3729–3736.
192. C. L. Drennan; S. Huang; J. T. Drummond; R. G. Matthews; M. L. Ludwig, *Science* **1994**, *266*, 1669–1674.
193. M. M. Dixon; S. Huang; R. G. Matthews; M. Ludwig, *Structure* **1996**, *4*, 1263–1275.
194. V. Bamdarian; K. A. Partridge; B. W. Lennon; D. P. Huddler; R. G. Matthews; M. L. Ludwig, *Nat. Struct. Biol.* **2002**, *9*, 53–56.
195. A. Abend; V. Bandarian; R. Nitsche; E. Stupperich; J. Rétey; G. H. Reed, *Arch. Biochem. Biophys.* **1999**, *370*, 138–141.
196. R. V. Banerjee; N. L. Johnson; J. K. Sobeski; P. Datta; R. G. Matthews, *J. Biol. Chem.* **1989**, *264*, 13888–13895.
197. C. W. Goulding; D. Postigo; R. G. Matthews, *Biochemistry* **1997**, *36*, 8082–8091.
198. C. W. Goulding; R. G. Matthews, *Biochemistry* **1997**, *36*, 15749–15757.
199. J. T. Jarrett; C. Y. Choi; R. G. Matthews, *Biochemistry* **1997**, *36*, 15739–15748.
200. R. V. Banerjee; V. Frasca; D. P. Ballou; R. G. Matthews, *Biochemistry* **1990**, *29*, 11101–11109.
201. R. T. Taylor; H. Weissbach, *J. Biol. Chem.* **1967**, *242*, 1509–1516.
202. P. A. Frey; M. L. Moss, *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, 571–577.
203. H. Sofia; G. Chen; B. G. Hetzler; J. F. Reyes-Spindola; N. E. Miller, *Nucleic Acids Res.* **2001**, *29*, 1097–1106.

**Biographical Sketch**

Perry Allen Frey was born in Plain City, OH, USA on 14 November 1935. After graduating from the local high school, he served 2 years in the US Army and then studied chemistry at The Ohio State University, graduating in 1959 with a BS degree in chemistry. He worked as a chemist for the US Public Health Service from 1960 to 1964. He was a predoctoral fellow of the National Institutes of Health (NIH) from 1964 to 1967 and received the Ph.D. degree in biochemistry from Brandeis University in 1968. He was a postdoctoral fellow of the NIH at Harvard University in 1968. In 1969, he became an assistant professor of chemistry at The Ohio State University where he rose to professor of chemistry. In 1981, he moved to the University of Wisconsin-Madison as professor of biochemistry, and was the Robert H. Abeles Professor from 1993 to 2008 and is currently Emeritus Professor. He is a member of the American Association for the Advancement of Science (AAAS), the American Chemical Society, the American Society for Biochemistry and Molecular Biology, and the Protein Society. He is a member of the National Academy of Sciences and a Fellow of the American Academy of Arts and Sciences and a Fellow of the AAAS.

## 7.15 Thiamin Biosynthesis

Tadhg P. Begley and Steven E. Ealick, Cornell University, Ithaca, NY, USA

© 2010 Elsevier Ltd. All rights reserved.

7.15.1	Introduction	547
7.15.2	Thiamin Biosynthesis in <i>Bacillus subtilis</i>	547
7.15.3	Thiamin Phosphate Synthase	548
7.15.4	Biosynthesis of the Thiazole Moiety of Thiamin in Bacteria	549
7.15.5	Protein Thiocarboxylates as Sulfide Carriers in Thiamin Biosynthesis	550
7.15.6	Formation of the Glycine Imine	551
7.15.7	Formation of the Pyrimidine Moiety of Thiamin	551
7.15.8	Kinases Involved in Thiamin Biosynthesis	553
7.15.9	Thiamin Salvage	553
7.15.10	Chemoenzymatic Synthesis of Thiamin Pyrophosphate	554
7.15.11	Thiamin Biosynthesis in <i>Saccharomyces cerevisiae</i>	555
7.15.12	Formation of the Thiamin Thiazole in Yeast	555
7.15.13	Formation of the Thiamin Pyrimidine in Yeast	556
7.15.14	Conclusions	556
References		557

### 7.15.1 Introduction

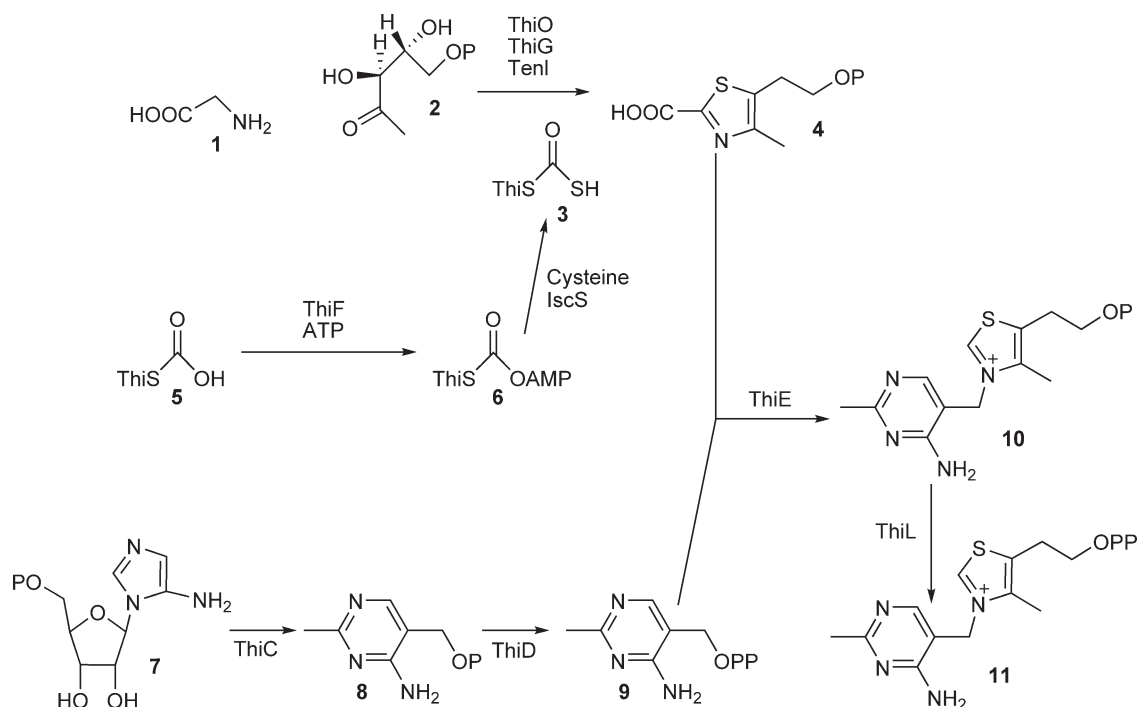
Thiamin pyrophosphate (**11**) (**Figure 1**) is an essential cofactor in all forms of life and it plays a key role in carbohydrate and amino acid metabolism by stabilizing acyl carbanion biosynthons. The mechanistic enzymology of thiamin pyrophosphate-dependent enzymes is described in detail in the chapter by Frank Jordan.<sup>1</sup> Here, we will review recent progress on the biosynthesis of thiamin pyrophosphate in bacteria and *Saccharomyces cerevisiae* with an emphasis on some of the novel organic chemistry that has emerged from these studies. Recent reviews describing the regulation of the pathway,<sup>2,3</sup> the identification of biosynthetic precursors,<sup>4</sup> and the structural biology of the pathway<sup>5–7</sup> have been published.

### 7.15.2 Thiamin Biosynthesis in *Bacillus subtilis*

The thiamin biosynthetic pathway in *B. subtilis* is representative of the major thiamin biosynthetic pathway in bacteria and has now been extensively characterized. All of the biosynthetic genes have been identified, and each of the enzymatic activities has been reconstituted. Mechanistic and structural studies have progressed greatly.<sup>6–9</sup>

The biosynthetic pathway is outlined in **Figure 1**. The thiazole **4** is formed by an oxidative condensation of glycine (**1**), deoxy-D-xylulose-5-phosphate (DXP, **2**), and a sulfide carrier protein with a thiocarboxylate at its carboxy terminus (ThiS-COSH, **3**). The pyrimidine phosphate **8** is formed by a deep-seated rearrangement of aminoimidazole ribotide (AIR, **7**). It is then pyrophosphorylated and used to alkylate the thiazole to give **10**. A final phosphorylation completes the biosynthesis. The entire pathway from glycine, DXP, cysteine, AIR, and ATP has now been reconstituted using purified enzymes.

Thiamin pyrophosphate is the active form of thiamin and is used as an enzymatic cofactor. Two other physiologically active forms of thiamin have been identified: ATP thiamin, which was recently identified as a stress metabolite in *Escherichia coli*,<sup>10,11</sup> and thiamin triphosphate,<sup>12–14</sup> the exact physiological function of which is still unclear.

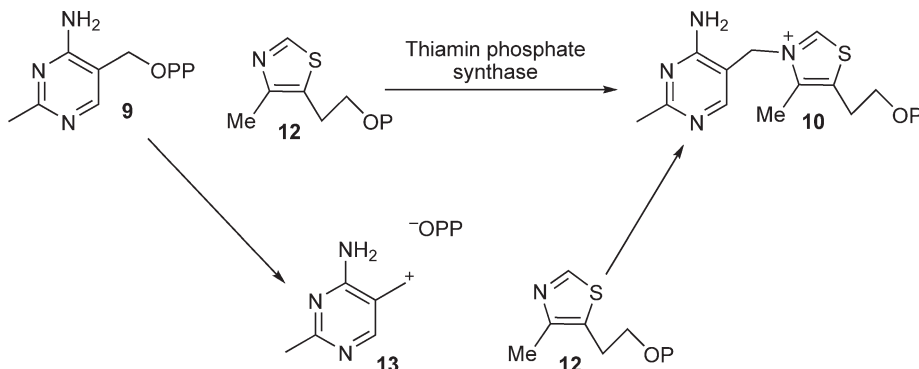


**Figure 1** Thiamin biosynthesis in *Bacillus subtilis* as representative of the major bacterial thiamin biosynthetic pathway.

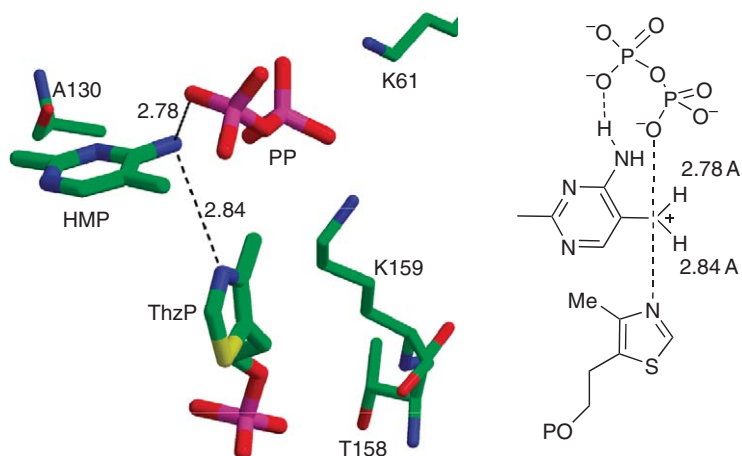
### 7.15.3 Thiamin Phosphate Synthase

Thiamin phosphate synthase catalyzes the coupling of the pyrimidine and the thiazole moieties of thiamin (**9** and **4** in **Figure 1**, **9** and **12** in **Figure 2**). This reaction proceeds by a dissociative mechanism involving initial formation of a pyrimidine carbocation (**13**), which then adds to the thiazole **12**. This mechanistic proposal was initially supported by the observation of a positional isotope effect and a large retardation of the reaction rate when the pyrimidine methyl group was replaced with carbocation-destabilizing substituents.<sup>15</sup> Recently, it has been possible to measure the intrinsic rate of carbocation formation at the enzyme active site ( $0.4 \text{ s}^{-1}$ ).<sup>16</sup> A remarkable structure in which the pyrimidine carbocation **13** is trapped between the pyrophosphate leaving group and the thiazole nucleophile has been reported (**Figure 3**).<sup>17</sup>

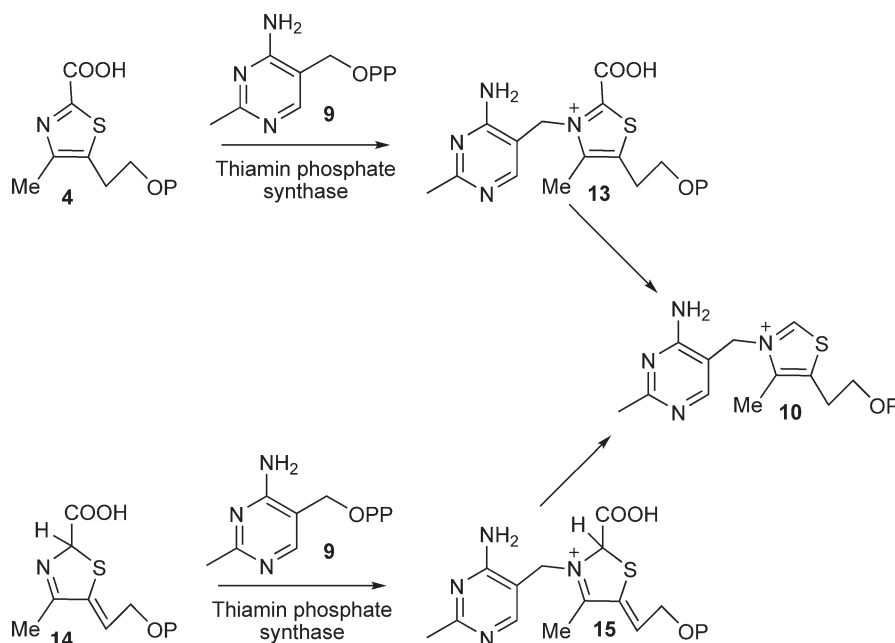
Thiamin phosphate synthase has broader substrate tolerance than initially assumed and catalyzes the formation of thiamin phosphate not only from **4** and **12** but also from **14** (**Figure 4**, T. P. Begley, unpublished data). Thiazole **14** is the product of the bacterial thiazole synthase.<sup>18</sup> It is tautomerized to **4** in a reaction



**Figure 2** Mechanism of thiamin phosphate synthase.



**Figure 3** Structure of the pyrimidine carbocation **13** trapped at the active site of thiamin phosphate synthase. The C–O and C–N bonds in the substrate and product are 1.44 and 1.57 Å, respectively.



**Figure 4** Alternative substrates of thiamin phosphate synthase.

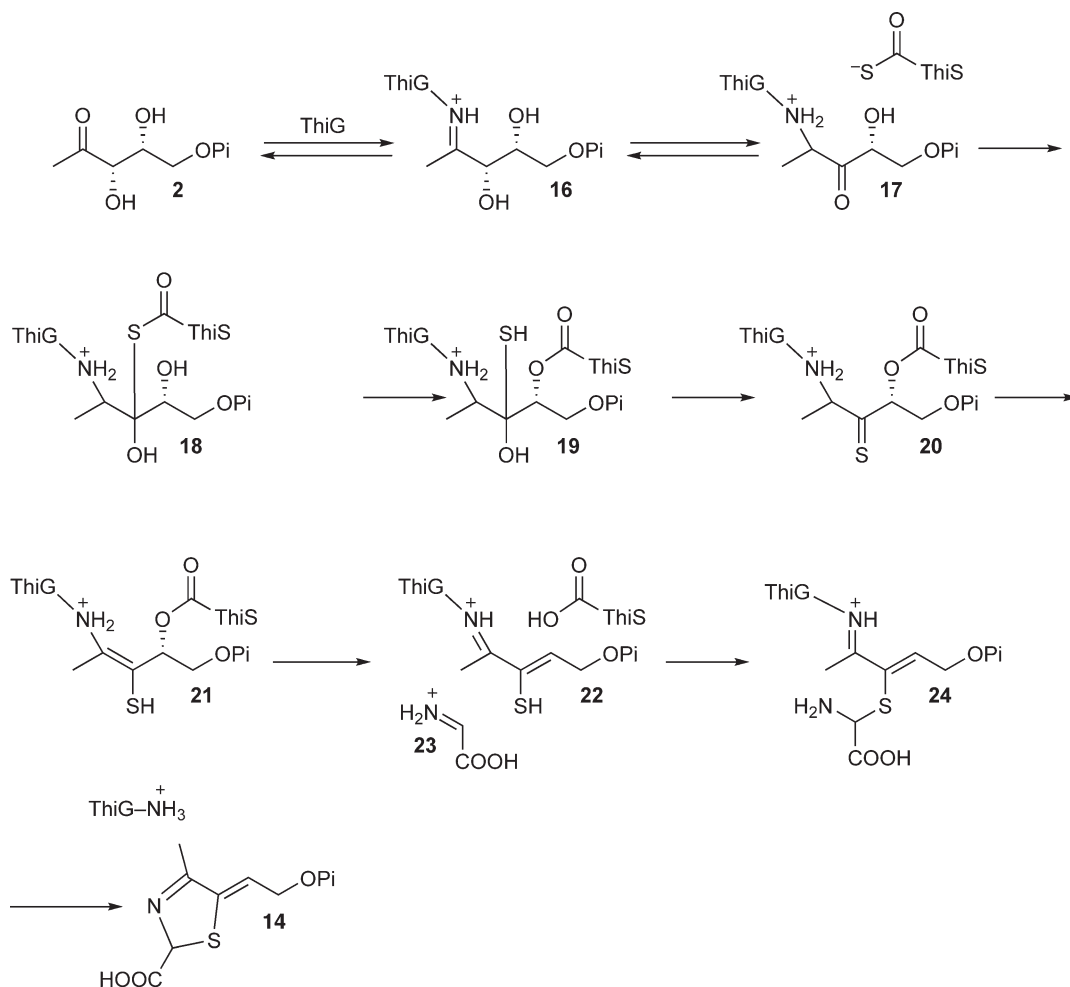
catalyzed by TenI (T. P. Begley, unpublished data). Thiazole **4** is the biosynthetic substrate of thiamin phosphate synthase in *B. subtilis*. The initially formed carboxythiamin phosphate (**13**) undergoes decarboxylation via the well-characterized thiamin ylide.

TenI is absent in *E. coli* and many other bacteria. In these systems, thiamin phosphate synthase catalyzes the formation of **15**, which then undergoes decarboxylation (T. P. Begley, unpublished data).

#### 7.15.4 Biosynthesis of the Thiazole Moiety of Thiamin in Bacteria

The *B. subtilis* thiazole synthase (ThiG) catalyzes the condensation of DXP (**2**), glycine imine (**23**), and ThiS-COSH (**3**) to give the thiazole tautomer **14**. The current mechanistic hypothesis for this complex reaction is shown in **Figure 5**. DXP forms an imine with lysine 96, which then tautomerizes to ketone **17**. Addition of ThiS-COSH





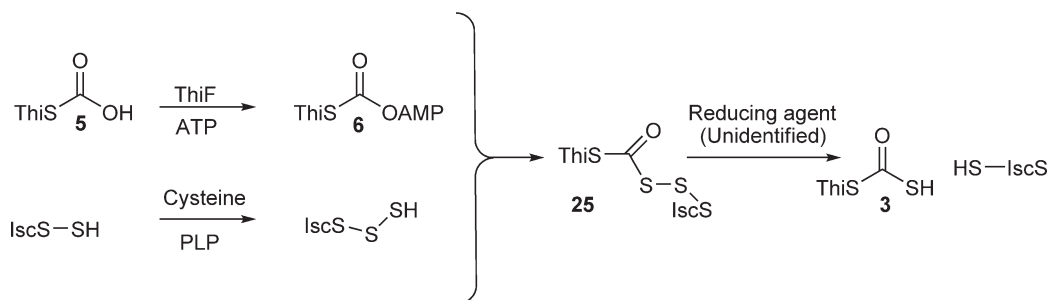
**Figure 5** Mechanistic proposal for the thiazole synthase (ThiG)-catalyzed reaction.

gives **18**. An O/S acyl shift followed by loss of water gives **20**. Tautomerization followed by elimination of ThiS-COOH gives **22**. Addition of **22** to the glycine imine **23** followed by transimination completes the reaction.

This mechanistic proposal has extensive experimental support: Imine **16** has been trapped and localized,<sup>19</sup> and enzyme-catalyzed hydrogen/deuterium (H/D) exchange at C3 and C4 of DXP as well as carbonyl oxygen exchange has been detected.<sup>20</sup> Acyl transfer to the C4 hydroxyl (**18** to **19**) as well as transfer of oxygen from DXP to nascent ThiS-COOH (**21** to **22**) has been demonstrated.<sup>20,21</sup> Intermediate **22** has been trapped and detected by mass spectrometric analysis and the reaction product **14** has been fully characterized.<sup>20</sup> Thiazole synthase complexed with ThiS-COOH has been structurally characterized and the DXP imine **16** has been modeled into the active site revealing key catalytic residues.<sup>22</sup>

### 7.15.5 Protein Thiocarboxylates as Sulfide Carriers in Thiamin Biosynthesis

The enzymology of ThiS-COSH formation is shown in **Figure 6**. Adenylation of ThiS-COOH (**5**) is catalyzed by ThiF. The resulting acyl-adenylated **6** reacts with the IscS persulfide, a species also involved in iron-sulfur cluster biosynthesis, to form the mixed acyl disulfide **25**. This is then reduced to give ThiS-COSH (**3**).<sup>23,24</sup> The enzymology of this reduction has not yet been determined. The ThiFS complex has been structurally characterized.<sup>25,26</sup>



**Figure 6** Mechanism of formation of ThiS-COSH (3), the sulfide donor in thiazole biosynthesis.

In *E. coli*, IscS-S-SH donates sulfur to ThiI and the resulting ThiI-persulfide functions as the sulfur donor for ThiS-COSH formation.<sup>27,28</sup> The biochemical logic of this extra step is unknown.

The use of ThiS-like proteins in biosynthetic sulfide transfer reactions is an emerging motif in natural product biosynthesis. Sequence analysis indicates that ThiS-like sequences are widespread and such proteins have now been experimentally characterized in the cysteine,<sup>29–31</sup> molybdopterin,<sup>32–36</sup> and thioquinolobactin<sup>37</sup> biosynthetic pathways.

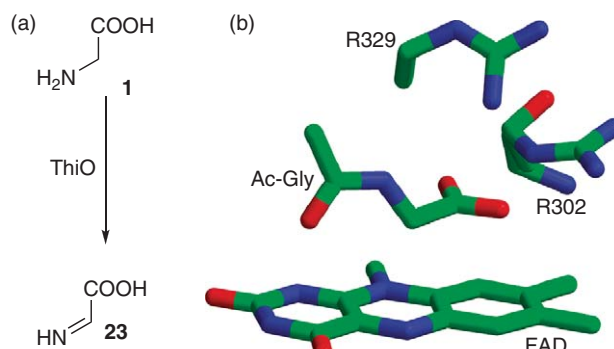
### 7.15.6 Formation of the Glycine Imine

Glycine oxidase (ThiO) catalyzes the oxidation of glycine (1) to the glycine imine 23. The ThiO structure, with *N*-acetylglycine bound at the active site, has been determined. This structure, as well as the efficient utilization of cyclopropyl glycine as a substrate, supports a hydride transfer mechanism (Figure 7).<sup>38</sup>

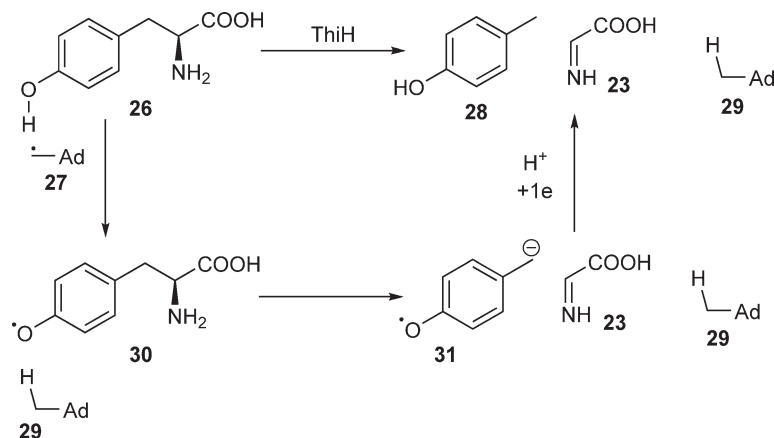
In *E. coli*, ThiH catalyzes the formation of the glycine imine 23 from tyrosine (26). ThiH is an oxygen-sensitive radical *S*-adenosyl-L-methionine (SAM) enzyme. Its activity has been reconstituted and the mechanism outlined in Figure 8 has been proposed.<sup>39–43</sup> It is unclear why *E. coli* adopts such a complex route to the glycine imine when oxidation of glycine using nicotinamide adenine dinucleotide (NAD) would accomplish the same transformation.

### 7.15.7 Formation of the Pyrimidine Moiety of Thiamin

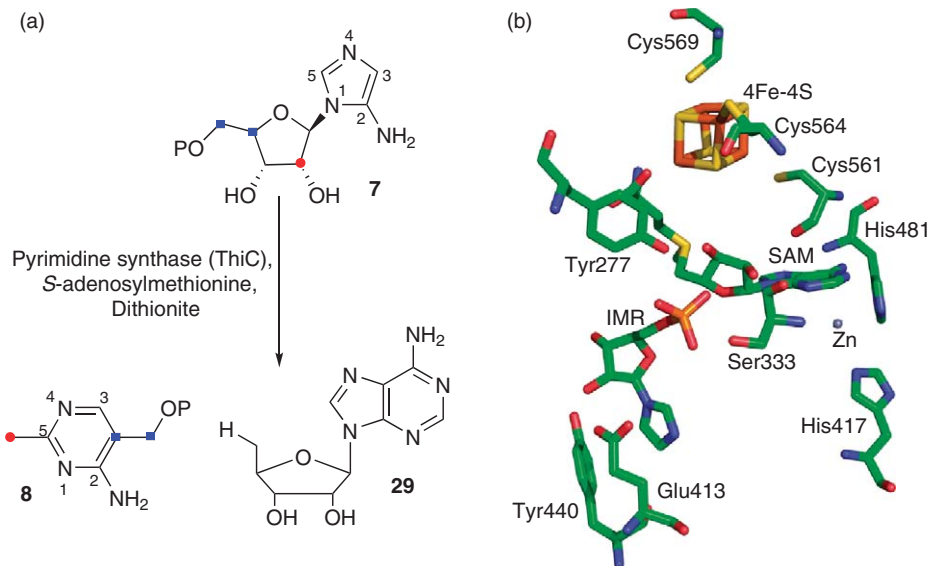
The pyrimidine moiety of thiamin is formed by a remarkable rearrangement of aminoimidazole ribotide (AIR) (7). This reaction is different from any of the known chemical or biosynthetic pathways to pyrimidines and as far as we know is the most complex unresolved rearrangement in primary metabolism.



**Figure 7** (a) Reaction catalyzed by glycine oxidase (ThiO). (b) Active site structure of glycine oxidase showing the C $\alpha$  hydrogen of *N*-acetylglycine positioned for hydride transfer to the flavin cofactor.



**Figure 8** Mechanistic proposal for the ThiH-catalyzed formation of the glycine imine **23**.



**Figure 9** (a) The pyrimidine synthase-catalyzed reaction showing that three of the pyrimidine carbons (indicated with colored dots) are derived from the ribose of AIR (**7**). (b) A model for the pyrimidine synthase active site showing the [Fe<sub>4</sub>-S<sub>4</sub>] cluster, SAM, and desamino-AIR (IMR).

The pyrimidine synthase-catalyzed reaction has recently been reconstituted in a defined biochemical system consisting of pyrimidine synthase (ThiC), SAM, and dithionite.<sup>44–46</sup> ThiC is a radical SAM enzyme and catalyzes the production of HMP-P (**8**) and deoxyadenosine (**29**) (Figure 9(a)). The reaction mechanism has not yet been determined.

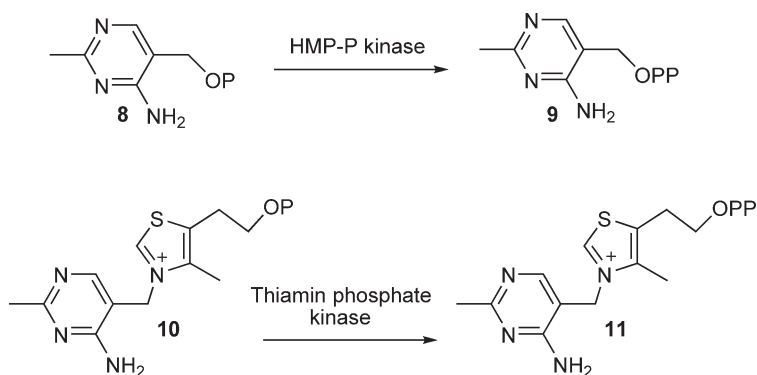
A partial structure of ThiC has been reported.<sup>44</sup> Although this structure is missing the carboxy-terminal Fe/S cluster-binding region, it shows high similarity to biotin synthase.<sup>47</sup> This facilitated the generation of a model for the missing domain. The resulting model for the active site of ThiC is shown in Figure 9(b). Based on this model, the iron–sulfur cluster is suitably positioned to reduce SAM to generate the adenosyl radical and this radical is reasonably positioned to abstract a hydrogen atom from AIR triggering the rearrangement cascade.

### 7.15.8 Kinases Involved in Thiamin Biosynthesis

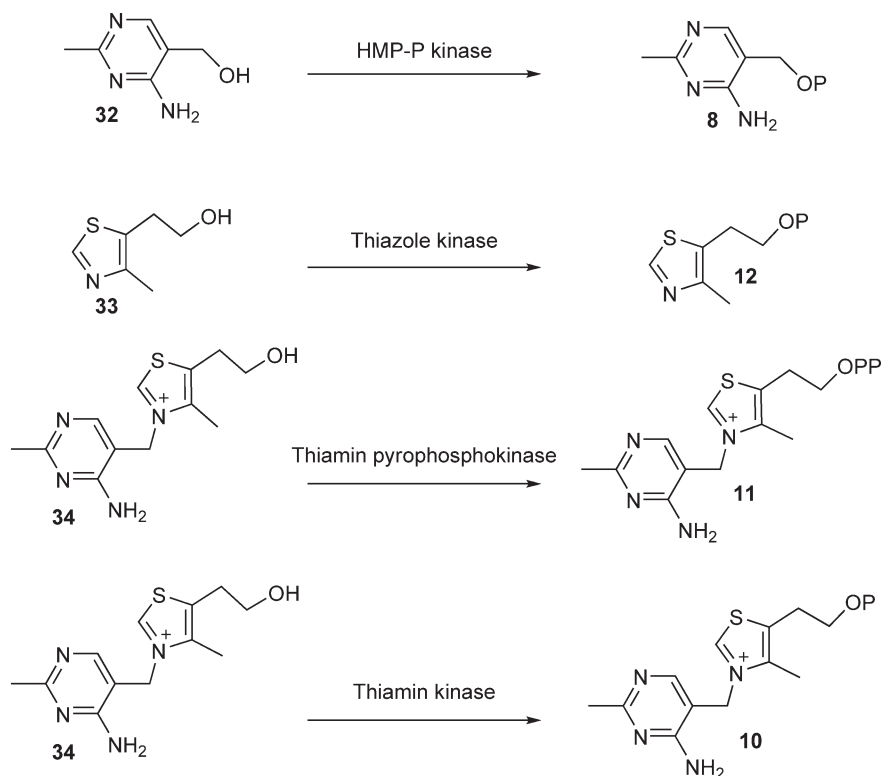
HMP-P kinase (ThiD) and thiamin phosphate kinase (ThiL) have been biochemically and structurally characterized (Figure 10). Both follow the expected in-line mechanism for phosphate transfer from ATP.<sup>48,49</sup>

### 7.15.9 Thiamin Salvage

The thiazole and pyrimidine heterocycles (33 and 32), as well as thiamin alcohol, can be salvaged from the bacterial growth medium and the kinases required for conversion of these precursors to biosynthetic intermediates have all been characterized (Figure 11).<sup>50-54</sup> HMP-P kinase catalyzes the iterative phosphorylation of



**Figure 10** Kinases involved in *de novo* biosynthesis of thiamin pyrophosphate.



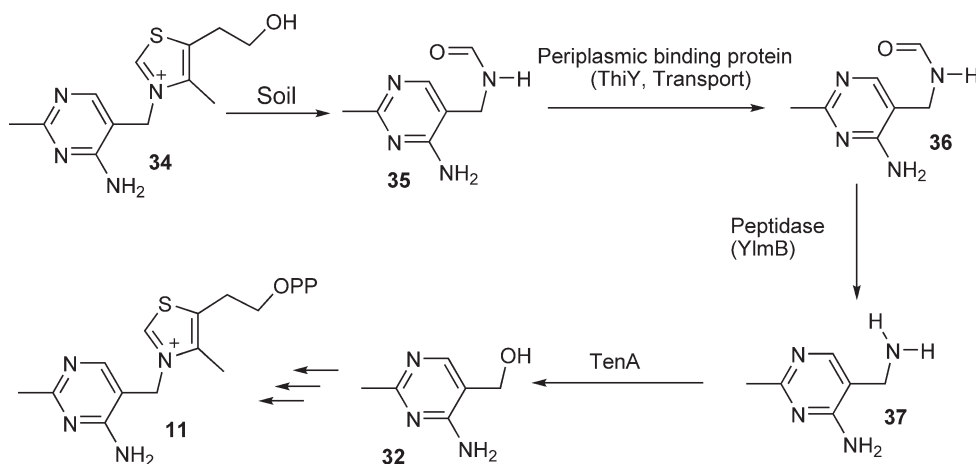
**Figure 11** Kinases involved in thiamin salvage reactions.

HMP, forming first HMP-P (8, salvage activity) and then HMP-PP (9, biosynthetic activity).<sup>51</sup> Structural studies have provided an explanation for this novel dual kinase activity.<sup>48</sup> The ABC transporter involved in thiamin uptake has also been identified and the periplasmic thiamin-binding protein has been structurally characterized.<sup>55,56</sup>

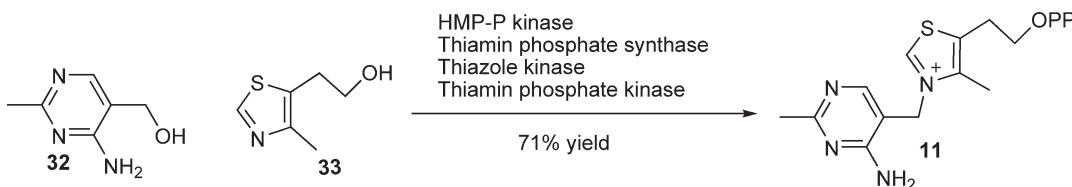
Thiamin is a reactive molecule and undergoes decomposition of the thiazolium moiety under mild basic reaction conditions. Recently, a new pathway for the salvage of a thiamin degradation product, formed in the presence of sterilized soil, was identified in *Bacillus balodurans* (Figure 12).<sup>57</sup> In this pathway, formyl amino pyrimidine (35), generated by soil-catalyzed hydrolysis of thiamin, is transported into the cell, deformylated to 37 and hydrolyzed to HMP (32). This is then incorporated into the thiamin biosynthetic pathway. TenA has been structurally and mechanistically characterized.<sup>58,59</sup> It is mechanistically similar to thiaminase I<sup>60</sup> and catalyzes the substitution reaction by an addition–elimination mechanism. TenA is widely distributed in all three kingdoms of life, suggesting that this salvage pathway is also widely used.

### 7.15.10 Chemoenzymatic Synthesis of Thiamin Pyrophosphate

While the thiazole and pyrimidine heterocycles of thiamin can be readily synthesized chemically, the thiazole/pyrimidine coupling and the final pyrophosphorylation are difficult reactions. In contrast, the enzymatic syntheses of the thiazole and the pyrimidine are complex, but the coupling and phosphorylation chemistry is straightforward. This synthetic/biosynthetic complementarity has been utilized in the development of an efficient chemoenzymatic synthesis of thiamin pyrophosphate (Figure 13).<sup>61</sup> This route has been used for the preparation of isotopically labeled thiamin pyrophosphate to characterize the thiamin radical found in pyruvate:ferredoxin oxidoreductase<sup>62</sup> and is the preferred method to synthesize isotopically labeled forms of thiamin pyrophosphate.



**Figure 12** Salvage of HMP (32) from soil-degraded thiamin.



**Figure 13** Chemoenzymatic synthesis of thiamin pyrophosphate.

### 7.15.11 Thiamin Biosynthesis in *Saccharomyces cerevisiae*

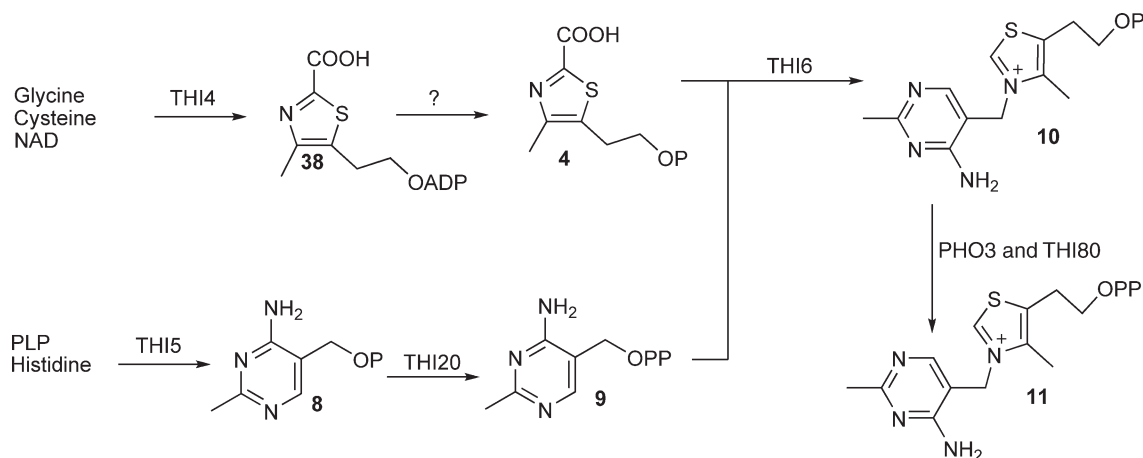
The biosynthesis of thiamin pyrophosphate in *S. cerevisiae* is outlined in **Figure 14**.<sup>63,64</sup> The pyrimidine phosphate **8** is phosphorylated by THI20, which has an associated TenA activity.<sup>65</sup> Hydrolysis of **38** gives thiazole **4**. The responsible enzyme has not yet been identified. Thiazole **4** and pyrimidine **9** are coupled to give thiamin phosphate (**10**). This reaction is catalyzed by THI6, which also has an associated thiazole kinase activity.<sup>66</sup> The formation of thiamin pyrophosphate from thiamin phosphate is unusual, involving first a dephosphorylation catalyzed by PHO3<sup>67</sup> followed by pyrophosphorylation catalyzed by thiamin pyrophosphokinase (THI80).<sup>53,54</sup> The chemistry of the thiazole and pyrimidine formation is quite different from that found in bacteria and will be described below.

### 7.15.12 Formation of the Thiamin Thiazole in Yeast

The thiazole moiety of thiamin in *S. cerevisiae* is formed from NAD (**39**) and glycine (**1**) in a reaction catalyzed by the THI4 protein.<sup>68–70</sup> Remarkably, the thiazole sulfur is derived from the Cys205 of THI4p (T. P. Begley, unpublished data). The thiazole-forming reaction has now been fully reconstituted and a structure of the THI4 protein has been reported.<sup>71</sup> A mechanistic proposal is outlined in **Figure 15**.

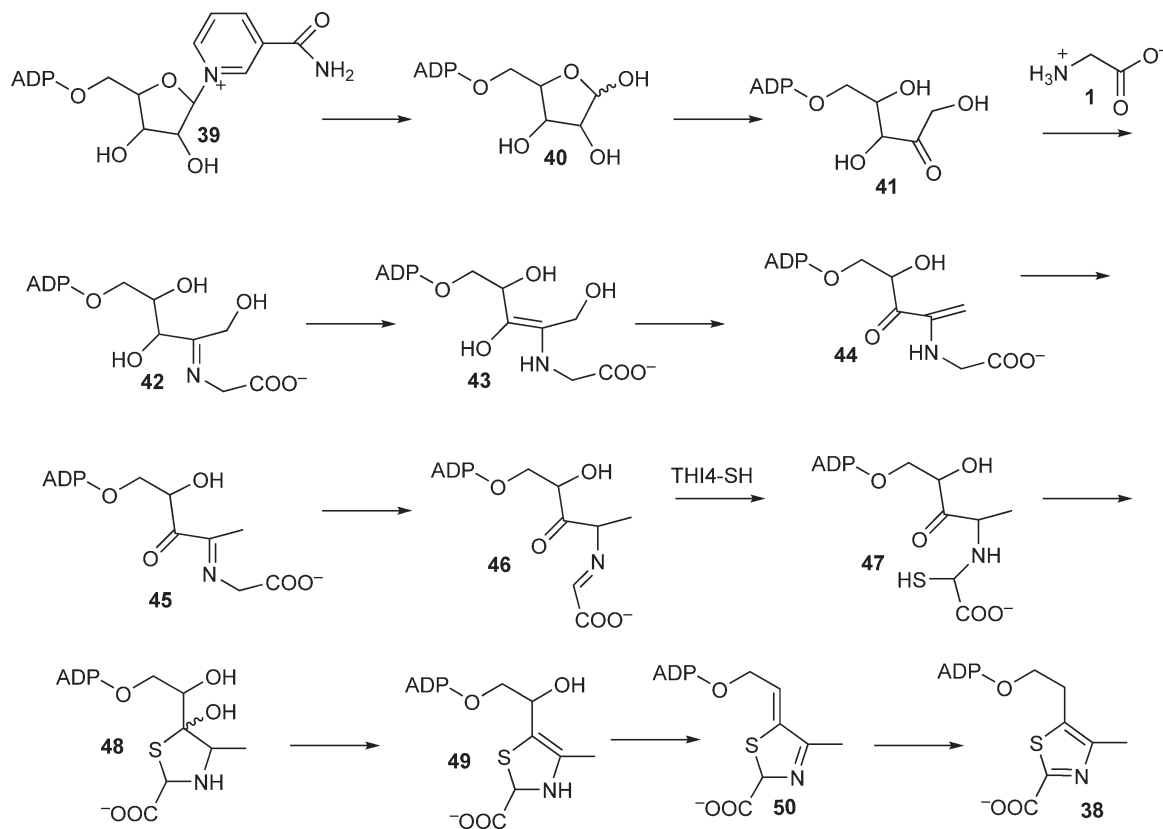
In this mechanism, cleavage of the N-glycosyl bond of NAD (**39**) gives **40**. Ring opening, tautomerization, and imine formation with glycine results in **42**. This is then converted to **46** by a tautomerization/elimination reaction sequence. Addition of the thiol from Cys205 of THI4 followed by elimination of the resulting adduct gives **47** with the concomitant conversion of Cys205 to a dehydroalanine residue. Cyclization, double dehydration, and a final tautomerization complete the formation of the thiazole **38**.

This mechanistic proposal is supported by a crystal structure of THI4 with **38** and dehydroalanine at the active site (S. Ealick, unpublished data), by the identification of **38**, **46**, and **50** as tightly bound metabolites released from the overexpressed protein by heat denaturation,<sup>69</sup> and by the observation of the formation of **40**, **41**, and **46** in the Cys204A mutant.<sup>68</sup> Isolation of the enzyme from an *E. coli* overexpression strain, grown under conditions of iron depletion, gave active protein after the addition of Fe(II). Mass spectrometry of this protein, before and after thiazole formation, revealed that the thiazole sulfur was derived from Cys205 of THI4 (T. P. Begley, unpublished data). THI4 is a remarkable biosynthetic protein that serves as a sulfide donor for the product it produces. It has not yet been established if THI4 is a single-turnover enzyme or whether as yet unidentified enzymes in *S. cerevisiae* recycle it.



**Figure 14** Thiamin pyrophosphate biosynthesis in *Saccharomyces cerevisiae*. The nudix hydrolase required for the removal of AMP from **38** has not yet been identified.





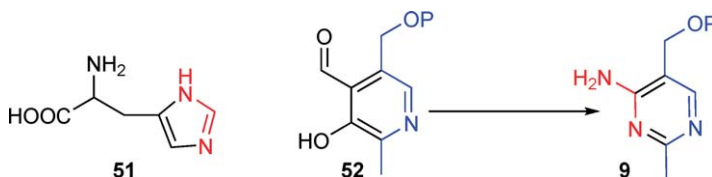
**Figure 15** Mechanism of thiazole formation in *Saccharomyces cerevisiae*.

### 7.15.13 Formation of the Thiamin Pyrimidine in Yeast

The thiamin pyrimidine **9** in yeast is also formed by a remarkable reaction. Labeling studies have identified histidine (**51**) and PLP (**52**) as precursors (**Figure 16**).<sup>72</sup> *THI5* encodes the pyrimidine synthase<sup>73</sup> but this remarkable reaction has not yet been reconstituted and the exact enzyme substrates have not been identified.

### 7.15.14 Conclusions

Our understanding of the enzymology of thiamin biosynthesis in bacteria is now at an advanced level: the complete biosynthetic pathway in *B. subtilis* has been reconstituted and all of the required enzymes have been structurally and mechanistically characterized. Thiazole formation in yeast is also relatively well understood. The formation of the thiamin pyrimidine in yeast is still poorly understood: apart from the identification of PLP and histidine as precursors, nothing is known about the mechanistic enzymology of this complex process and



**Figure 16** Isotopic labeling studies reveal the origin of the atoms of HMP in *Saccharomyces cerevisiae*.

the reaction has not been reconstituted. In addition to the bacterial and yeast pathways described here, comparative genome analysis of thiamin biosynthesis<sup>74,75</sup> shows that some organisms use the bacterial thiazole pathway and the eukaryotic pyrimidine pathway, some use the bacterial pyrimidine pathway and the eukaryotic thiazole pathway, and some lack one or both of these pathways, suggesting that additional thiamin biosynthetic chemistry still remains to be discovered.

## References

1. F. Jordan, *Nat. Prod. Rep.* **2003**, 20 (2), 184–201.
2. W. Winkler; A. Nahvi; R. R. Breaker, *Nature* **2002**, 419 (6910), 952–956.
3. T. M. Henkin, *Genes Dev.* **2008**, 22 (24), 3383–3390.
4. I. D. Spenser; R. L. White, *Angew. Chem. Int. Ed. Engl.* **1997**, 36 (10), 1032–1046.
5. E. Settembre; T. P. Begley; S. E. Ealick, *Curr. Opin. Struct. Biol.* **2003**, 13 (6), 739–747.
6. C. T. Jurgenson; T. P. Begley; S. E. Ealick, *Annu. Rev. Biochem.* **2009**, 78, 569–603.
7. C. T. Jurgenson; S. E. Ealick; T. P. Begley, Biosynthesis of Thiamin Pyrophosphate. In *EcoSal—Escherichia Coli and Salmonella: Cellular and Molecular Biology*; A. Böck, R. Curtiss III, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, D. Ussery, Eds.; ASM Press: Washington, DC. <http://www.ecosal.org>
8. T. P. Begley, *Nat. Prod. Rep.* **2006**, 23 (1), 15–25.
9. T. P. Begley; A. Chatterjee; J. W. Hanes; A. Hazra; S. E. Ealick, *Curr. Opin. Chem. Biol.* **2008**, 12 (2), 118–125.
10. L. Bettendorff; B. Wirtzfeld; A. F. Makarchikov; G. Mazzucchelli; M. Frederich; T. Gigliobianco; M. Gangolf; E. De Pauw; L. Angenot; P. Wins, *Nat. Chem. Biol.* **2007**, 3 (4), 211–212.
11. A. F. Makarchikov; A. Brans; L. Bettendorff, *BMC Biochem.* **2007**, 8, 17.
12. L. Bettendorff; H. A. Kolb; E. Schoffeniels, *J. Membr. Biol.* **1993**, 136 (3), 281–288.
13. L. Bettendorff; M. Peeters; C. Jouan; P. Wins; E. Schoffeniels, *Anal. Biochem.* **1991**, 198 (1), 52–59.
14. L. Bettendorff; M. Peeters; P. Wins; E. Schoffeniels, *J. Neurochem.* **1993**, 60 (2), 423–434.
15. J. J. Reddick; R. Nicewonger; T. P. Begley, *Biochemistry* **2001**, 40 (34), 10095–10102.
16. J. W. Hanes; S. E. Ealick; T. P. Begley, *J. Am. Chem. Soc.* **2007**, 129 (16), 4860–4861.
17. D. H. Peapus; H. J. Chiu; N. Campobasso; J. J. Reddick; T. P. Begley; S. E. Ealick, *Biochemistry* **2001**, 40 (34), 10103–10114.
18. A. Hazra; A. Chatterjee; T. P. Begley, *J. Am. Chem. Soc.* **2009**, 131 (9), 3225–3229.
19. P. C. Dorrestein; H. Zhai; S. V. Taylor; F. W. McLafferty; T. P. Begley, *J. Am. Chem. Soc.* **2004**, 126 (10), 3091–3096.
20. P. C. Dorrestein; H. Zhai; F. W. McLafferty; T. P. Begley, *Chem. Biol.* **2004**, 11 (10), 1373–1381.
21. A. Chatterjee; X. Han; F. W. McLafferty; T. P. Begley, *Angew. Chem. Int. Ed. Engl.* **2006**, 45 (21), 3507–3510.
22. E. C. Settembre; P. C. Dorrestein; H. Zhai; A. Chatterjee; F. W. McLafferty; T. P. Begley; S. E. Ealick, *Biochemistry* **2004**, 43 (37), 11647–11657.
23. S. V. Taylor; N. L. Kelleher; C. Kinsland; H.-J. Chiu; C. A. Costello; A. D. Backstrom; F. W. McLafferty; T. P. Begley, *J. Biol. Chem.* **1998**, 273 (26), 16555–16560.
24. C. T. Lauhon; R. Kambampati, *J. Biol. Chem.* **2000**, 275 (26), 20096–20103.
25. C. Lehmann; T. P. Begley; S. E. Ealick, *Biochemistry* **2006**, 45 (1), 11–19.
26. D. M. Duda; H. Walden; J. Sfoudouris; B. A. Schulman, *J. Mol. Biol.* **2005**, 349 (4), 774–786.
27. E. G. Mueller, *Nat. Chem. Biol.* **2006**, 2 (4), 185–194.
28. C. M. Wright; G. D. Christman; A. M. Snellinger; M. V. Johnston; E. G. Mueller, *Chem. Commun. (Cambridge, UK)* **2006**, (29), 3104–3106.
29. K. E. Burns; S. Baumgart; P. C. Dorrestein; H. Zhai; F. W. McLafferty; T. P. Begley, *J. Am. Chem. Soc.* **2005**, 127 (33), 11602–11603.
30. C. T. Jurgenson; K. E. Burns; T. P. Begley; S. E. Ealick, *Biochemistry* **2008**, 47 (39), 10354–10364.
31. S. E. O’Leary; C. T. Jurgenson; S. E. Ealick; T. P. Begley, *Biochemistry* **2008**, 47 (44), 11606–11615.
32. J. N. Daniels; M. M. Wuebbens; K. V. Rajagopalan; H. Schindelin, *Biochemistry* **2008**, 47 (2), 615–626.
33. M. M. Wuebbens; K. V. Rajagopalan, *J. Biol. Chem.* **2003**, 278 (16), 14523–14532.
34. S. Leimkuhler; K. V. Rajagopalan, *J. Biol. Chem.* **2001**, 276 (25), 22024–22031.
35. S. Leimkuhler; M. M. Wuebbens; K. V. Rajagopalan, *J. Biol. Chem.* **2001**, 276 (37), 34695–34701.
36. M. J. Rudolph; M. M. Wuebbens; K. V. Rajagopalan; H. Schindelin, *Nat. Struct. Biol.* **2001**, 8 (1), 42–46.
37. A. M. Godert; M. Jin; F. W. McLafferty; T. P. Begley, *J. Bacteriol.* **2007**, 189 (7), 2941–2944.
38. E. C. Settembre; P. C. Dorrestein; J. H. Park; A. M. Augustine; T. P. Begley; S. E. Ealick, *Biochemistry* **2003**, 42 (10), 2971–2981.
39. M. Kriek; F. Martins; M. R. Challand; A. Croft; P. L. Roach, *Angew. Chem. Int. Ed. Engl.* **2007**, 46 (48), 9223–9226.
40. M. Kriek; F. Martins; R. Leonardi; S. A. Fairhurst; D. J. Lowe; P. L. Roach, *J. Biol. Chem.* **2007**, 282 (24), 17413–17423.
41. R. Leonardi; S. A. Fairhurst; M. Kriek; D. J. Lowe; P. L. Roach, *FEBS Lett.* **2003**, 539 (1–3), 95–99.
42. R. Leonardi; L. Roach Peter, *J. Biol. Chem.* **2004**, 279 (17), 17054–17062.
43. N. C. Martinez-Gomez; M. Robers; D. M. Downs, *J. Biol. Chem.* **2004**, 279 (39), 40505–40510.
44. A. Chatterjee; Y. Li; Y. Zhang; T. L. Grove; M. Lee; C. Krebs; S. J. Booker; T. P. Begley; S. E. Ealick, *Nat. Chem. Biol.* **2008**, 4 (12), 758–765.
45. N. C. Martinez-Gomez; D. M. Downs, *Biochemistry* **2008**, 47 (35), 9054–9056.
46. N. C. Martinez-Gomez; R. R. Poyner; S. O. Mansoorabadi; G. H. Reed; D. M. Downs, *Biochemistry* **2009**, 48 (2), 217–219.
47. F. Berkovitch; Y. Nicolet; J. T. Wan; J. T. Jarrett; C. L. Drennan, *Science* **2004**, 303 (5654), 76–79.
48. G. Cheng; E. M. Bennett; T. P. Begley; S. E. Ealick, *Structure* **2002**, 10 (2), 225–235.
49. K. M. McCulloch; C. Kinsland; T. P. Begley; S. E. Ealick, *Biochemistry* **2008**, 47 (12), 3810–3821.

50. N. Campobasso; I. I. Mathews; T. P. Begley; S. E. Ealick, *Biochemistry* **2000**, 39 (27), 7868–7877.
51. J. J. Reddick; C. Kinsland; R. Nicewonger; T. Christian; D. M. Downs; M. E. Winkler; T. P. Begley, *Tetrahedron* **1998**, 54 (52), 15983–15991.
52. J. Melnick; E. Lis; J. H. Park; C. Kinsland; H. Mori; T. Baba; J. Perkins; G. Schyns; O. Vassieva; A. Osterman; T. P. Begley, *J. Bacteriol.* **2004**, 186 (11), 3660–3662.
53. L. J. Baker; J. A. Dorocke; R. A. Harris; D. E. Timm, *Structure* **2001**, 9 (6), 539–546.
54. D. E. Timm; J. Liu; L. J. Baker; R. A. Harris, *J. Mol. Biol.* **2001**, 310 (1), 195–204.
55. E. V. Soriano; K. R. Rajashankar; J. W. Hanes; S. Bale; T. P. Begley; S. E. Ealick, *Biochemistry* **2008**, 47 (5), 1346–1357.
56. Y. Devedjiev; Y. Surendranath; U. Derewenda; A. Gabrys; D. R. Cooper; R.-g. Zhang; L. Lezondra; A. Joachimiak; Z. S. Derewenda, *J. Mol. Biol.* **2004**, 343 (2), 395–406.
57. A. H. Jenkins; G. Schyns; S. Potot; G. Sun; T. P. Begley, *Nat. Chem. Biol.* **2007**, 3 (8), 492–497.
58. A. L. Jenkins; Y. Zhang; S. E. Ealick; T. P. Begley, *Bioorg. Chem.* **2007**, 31 (1), 29–32.
59. A. V. Toms; A. L. Haas; J. H. Park; T. P. Begley; S. E. Ealick, *Biochemistry* **2005**, 44 (7), 2319–2329.
60. N. Campobasso; C. A. Costello; C. Kinsland; T. P. Begley; S. E. Ealick, *Biochemistry* **1998**, 37 (45), 15981–15989.
61. S. Melnick Jonathan; K. I. Sprinz; J. Reddick Jason; C. Kinsland; P. Begley Tadhg, *Bioorg. Med. Chem. Lett.* **2003**, 13 (22), 4139–4141.
62. O. Mansoorabadi Steven; J. Seravalli; C. Furdui; V. Krymov; J. Gerfen Gary; P. Begley Tadhg; J. Melnick; W. Ragsdale Stephen; H. Reed George, *Biochemistry* **2006**, 45 (23), 7122–7131.
63. K. Nosaka, *Appl. Microbiol. Biotechnol.* **2006**, 72 (1), 30–40.
64. E. Kowalska; A. Kozik, *Cell. Mol. Biol. Lett.* **2008**, 13 (2), 271–282.
65. A. L. Haas; N. P. Laun; T. P. Begley, *Bioorg. Chem.* **2005**, 33 (4), 338–344.
66. K. Nosaka; H. Nishimura; Y. Kawasaki; T. Tsujihara; A. Iwashima, *J. Biol. Chem.* **1994**, 269 (48), 30510–30516.
67. K. Nosaka, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1990**, 1037 (2), 147–154.
68. A. Chatterjee; C. T. Jurgenson; F. C. Schroeder; S. E. Ealick; T. P. Begley, *J. Am. Chem. Soc.* **2007**, 129 (10), 2914–2922.
69. A. Chatterjee; C. T. Jurgenson; F. C. Schroeder; S. E. Ealick; T. P. Begley, *J. Am. Chem. Soc.* **2006**, 128 (22), 7158–7159.
70. A. Chatterjee; F. C. Schroeder; C. T. Jurgenson; S. E. Ealick; T. P. Begley, *J. Am. Chem. Soc.* **2008**, 130 (34), 11394–11398.
71. C. T. Jurgenson; A. Chatterjee; T. P. Begley; S. E. Ealick, *Biochemistry* **2006**, 45 (37), 11061–11070.
72. K. Himmeldirk; B. G. Sayer; I. D. Spenser, *J. Am. Chem. Soc.* **1998**, 120 (15), 3581–3589.
73. R. Wightman; P. A. Meacock, *Microbiology* **2003**, 149 (Pt. 6), 1447–1460.
74. The SEED Data Base for Comparative Genome Analysis. <http://theseed.uchicago.edu/FIG/subsys.cgi>
75. D. A. Rodionov; A. G. Vitreschak; A. A. Mironov; M. S. Gelfand, *J. Biol. Chem.* **2002**, 277 (50), 48949–48959.

### Biographical Sketches



Tadhg Begley obtained his B.Sc. from The National University of Ireland in Cork in 1977 and his Ph.D. from the California Institute of Technology (P. Dervan) in 1982. He carried out postdoctoral studies at the University of Geneva (W. Oppolzer) and at MIT (C. Walsh). After 23 years in the Cornell Chemistry Department, he recently moved to Texas A&M University where he is the Derek H. R. Barton Professor of Chemistry. Begley's research is focused on the mechanistic enzymology of complex organic transformations, particularly those found in the vitamin biosynthetic pathways.



Steven Ealick is the William T. Miller Professor of Chemistry and Chemical Biology at Cornell University. He received his B.S. degree in chemistry from Oklahoma State University in 1972 and his Ph.D. degree in physical chemistry from the University of Oklahoma in 1976. He is also the Director of the Northeast Collaborative Access Team at the Advanced Photon Source, Argonne National Laboratory. His research focuses on protein structure and function using X-ray crystallography, including enzyme mechanism, structure-based drug design, and protein evolution. He has published widely in the areas of thiamin biosynthesis, purine biosynthesis, polyamine biosynthesis, and cytokine structure and function. He is also interested in the development of instrumentation for synchrotron beamlines used for protein crystallography.

## 7.16 Thiamin Enzymology

Frank Jordan, Rutgers, The State University of New Jersey, Newark, NJ, USA

© 2010 Elsevier Ltd. All rights reserved.

<b>7.16.1</b>	<b>Introduction</b>	562
<b>7.16.2</b>	<b>Detection of Thiamin Diphosphate-Related Intermediates and Their Kinetic Fates, and the Information Gained from Such Data</b>	562
7.16.2.1	Thiamin Diphosphate-Related Intermediates Prior to Substrate Addition	562
7.16.2.1.1	The 4'-aminopyrimidine form of thiamin diphosphate	567
7.16.2.1.2	The 1',4'-iminopyrimidine form of thiamin diphosphate <sup>31–35</sup>	567
7.16.2.1.3	Determination of $pK_a$ for the enzyme-bound APH <sup>+</sup> form <sup>35</sup>	568
7.16.2.1.4	The C2-carbanion/ylide/carbene	569
<b>7.16.2.2</b>	<b>Thiamin-Bound Intermediates with Substrate or Substrate Analog Present</b>	569
7.16.2.2.1	The Michaelis–Menten complex	569
7.16.2.2.2	The covalent substrate–thiamin diphosphate predecarboxylation complex (LThDP and analogs)	570
7.16.2.2.3	The first postdecarboxylation intermediate: The enamine/C2 $\alpha$ -carbanion	575
7.16.2.2.4	The second postdecarboxylation intermediate, the product–ThDP complex (HEThDP and HBThDP)	576
7.16.2.2.5	2-Acetylthiamin diphosphate, the two-electron oxidation product of the enamine	577
7.16.2.2.6	The C2 $\alpha$ -hydroxyethylideneThDP radical, the one-electron oxidation product of the enamine	578
<b>7.16.3</b>	<b>Determination of Rate-Limiting Steps and Microscopic Rate Constants on ThDP Enzymes</b>	579
<b>7.16.4</b>	<b>Function and Dynamics of Mobile Loops on YPDC and E1ec</b>	580
7.16.4.1	Evidence on YPDC	580
7.16.4.2	Evidence on E1ec	582
7.16.4.2.1	Structural evidence pointing to mobile loops	582
7.16.4.2.2	Studies on the inner active center mobile loop of E1ec <sup>46,99,100</sup>	584
<b>7.16.5</b>	<b>Structure–Function Studies in ThDP-Initiated Multienzyme Complex Reactions at the E1 (ThDP-Dependent) Component</b>	587
7.16.5.1	Structural Evidence	587
7.16.5.1.1	The amino terminal region of E1ec interacts with E2ec	588
7.16.5.1.2	Region of E2ec interacting with E1ec	589
<b>7.16.6</b>	<b>Probing ThDP Enzyme Mechanisms with Coenzyme Analogs</b>	590
<b>7.16.7</b>	<b>Active Center Communication in ThDP Enzymes</b>	590
7.16.7.1	The Proton Wire Mechanism	590
7.16.7.2	Spectroscopic Evidence for Nonequivalence of Active Sites	591
7.16.7.3	Kinetic Evidence for Nonequivalence of Active Sites	591
<b>7.16.8</b>	<b>Continuing Fascination with Thiamin Enzymes as Paradigm for Enzymatic Solvent Effects</b>	593
7.16.8.1	Evidence from Observation of the Enamine on YPDC	593
7.16.8.2	Evidence from Solvent Effects on Decarboxylation Rate Constants in Model Compounds	593
7.16.8.3	Evidence from Acceleration of the Reaction Step Leading to the Enamine on E1ec	594
7.16.8.4	Spectroscopic Observations Suggesting an Apolar Environment	594
<b>7.16.9</b>	<b>Perspective for Future Studies</b>	594
<b>References</b>		595

### 7.16.1 Introduction

There is no biosynthetic pathway for thiamin or vitamin B1 in mammalian cells, hence, it must be ingested in the diet, although there are now several pathways known in other cell types.<sup>1</sup> Of the phosphorylated analogs, there are biosynthetic enzymes known for interconversion of thiamin with its mono- (thiamin monophosphate (ThMP)), di- (thiamin diphosphate (ThDP)), and triphosphate (thiamin triphosphate (ThTP)) analogs. While the diphosphate ThDP is a well-known coenzyme (synthesized enzymatically for examples from ThC via a kinase called thiamin pyrophosphokinase transferring the  $\beta$ - $\gamma$ -diphosphate of ATP to the  $\beta$ -hydroxyethyl hydroxyl group), evidence for the role of the ThTP has also been cited for many years, and has more recently received greater recognition primarily due to the work of Bettendorff and colleagues.<sup>2</sup>

In writing this chapter, the author is keenly aware of the number of other excellent reviews available in the literature<sup>3–12</sup> and therefore will concentrate on issues addressed over more than 30 years in his laboratories related to the coenzyme activity of ThDP, and reviewed by the author as well.<sup>13–15</sup> Biosynthetic issues related to thiamin are addressed by the distinguished editor of this volume.

The chemistry and enzymology of ThDP is intimately dependent on three chemical moieties comprising the coenzyme: a thiazolium ring, a 4-aminopyrimidine ring, and the diphosphate side chain (**Figure 1**). From the large number of high-resolution X-ray structures available for the past 16 years, starting with the structures of transketolase<sup>16</sup> (TK), pyruvate oxidase<sup>17</sup> (POX) from *Lactobacillus plantarum*, and pyruvate decarboxylase from the yeast *Saccharomyces cerevisiae* (YPDC),<sup>18,19</sup> it has become clear that the diphosphate serves to bind the cofactor to the protein (see **Table 1** for representative reactions discussed in this chapter). This is achieved via electrostatic bonds of the  $\alpha$  and  $\beta$  phosphoryl group negative charges with the required  $Mg^{2+}$  or  $Ca^{2+}$ , the divalent metal serving as an anchor in a highly tailored environment with a universally conserved GlyAspGly (GDG) recognition site and the diphosphate– $Mg^{2+}$ -binding motif consisting of GlyAspGly-X<sub>26</sub>-AsnAsn (GDG-X<sub>26</sub>-NN) sequence of amino acids suggested by the Perham group.<sup>20</sup> Historically, and chemically, the thiazolium ring is central to catalysis as reported in early studies by Breslow,<sup>21</sup> due to its ability to form a key nucleophilic center the C2-carbanion/ylide (YI) or carbene, depending on one's viewpoint regarding the relative importance of the resonance contributions. The 4-aminopyrimidine moiety has gained more favor as an important contributor to catalysis since the appearance of the X-ray structures showing its conserved proximity to the C2 thiazolium atom and the possibility for its participation in acid–base catalysis.<sup>22</sup>

The goal of this chapter is to touch on themes of particular interest to the author in the vast field of thiamin chemistry and enzymology, with an attempt to point out perceived gaps in our understanding, notwithstanding the massive amount of research devoted to this coenzyme for several decades. As the methods have improved, the issues could be studied in ever greater detail. Near to completion of this chapter, there appeared a very comprehensive review with somewhat similar goals in *Chemical Reviews* by Kluger and Tittmann.<sup>23</sup>

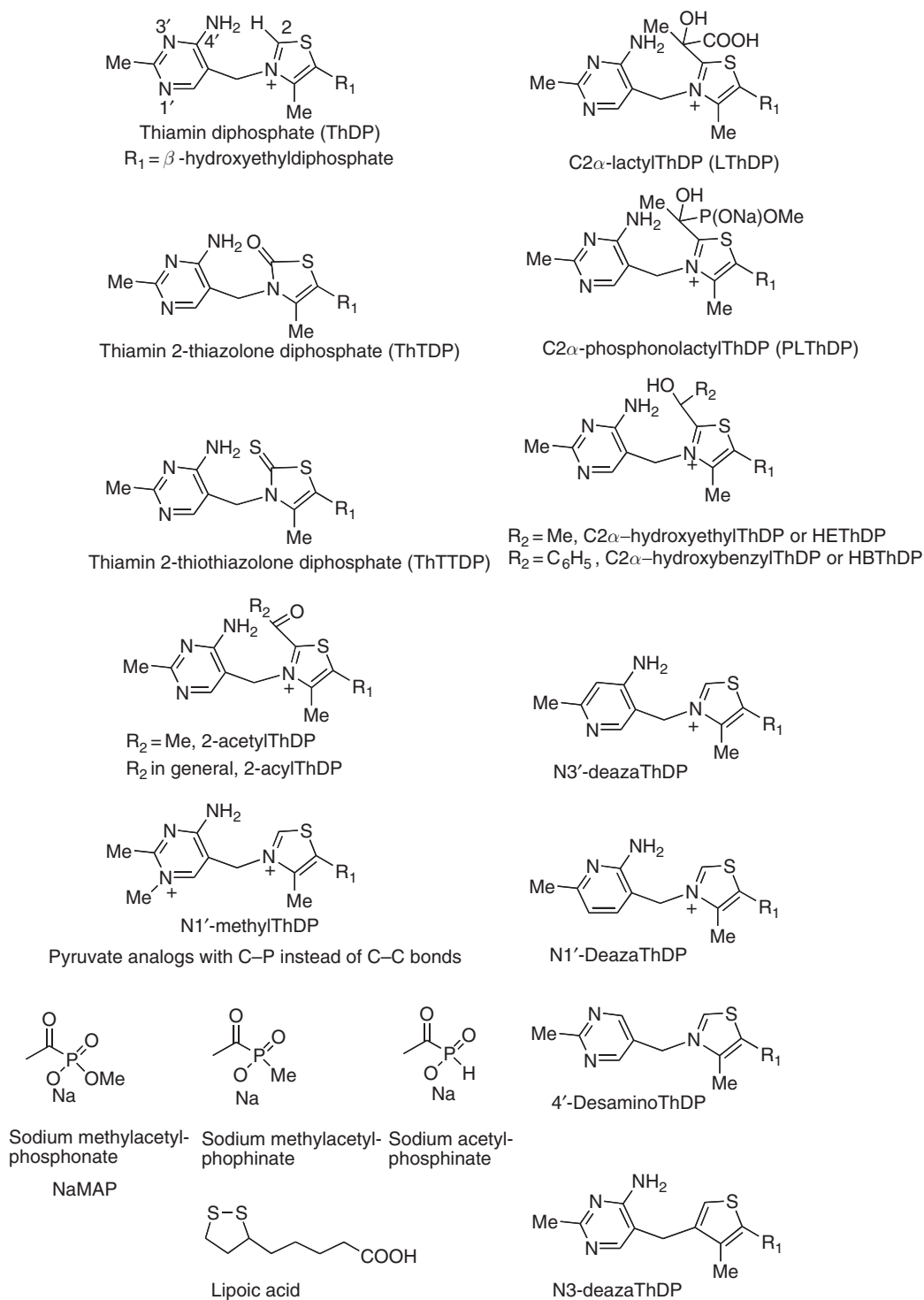
### 7.16.2 Detection of Thiamin Diphosphate-Related Intermediates and Their Kinetic Fates, and the Information Gained from Such Data

The presentation of thiamin-related and thiamin-bound intermediates represent pre-, or postsubstrate (or substrate analog) binding, an important distinction needed since the recent identification of several forms of ThDP on the enzymes.

#### 7.16.2.1 Thiamin Diphosphate-Related Intermediates Prior to Substrate Addition

In the recent past, the author with collaborators established the presence of various tautomeric and ionization states of ThDP (**Schemes 1 and 2**). With the publication of the first high-resolution X-ray structures of ThDP enzymes,<sup>16–19,22</sup> two issues were clarified for the first time: (1) The conformation of the bound coenzyme (defining the dihedral angles formed by the thiazolium and 4'-aminopyrimidine rings with respect to the bridge methylene group was quite different (so-called V conformation) from that found and predicted for both free



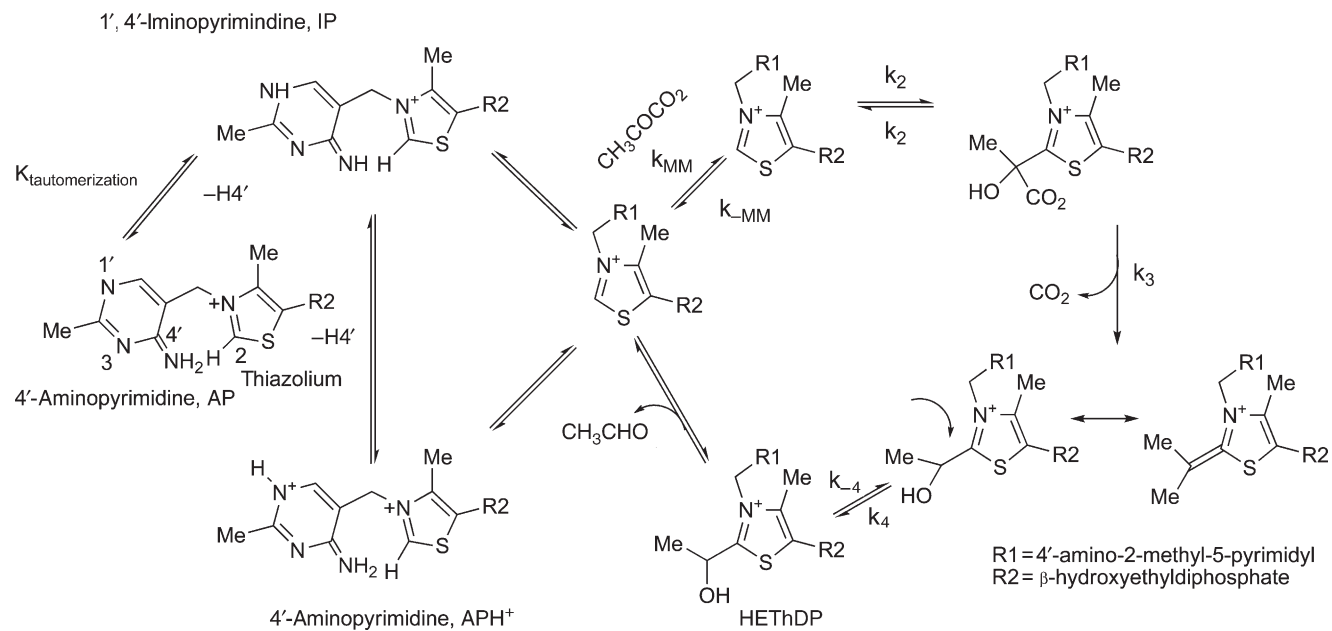


**Figure 1** Compounds discussed in the chapter.

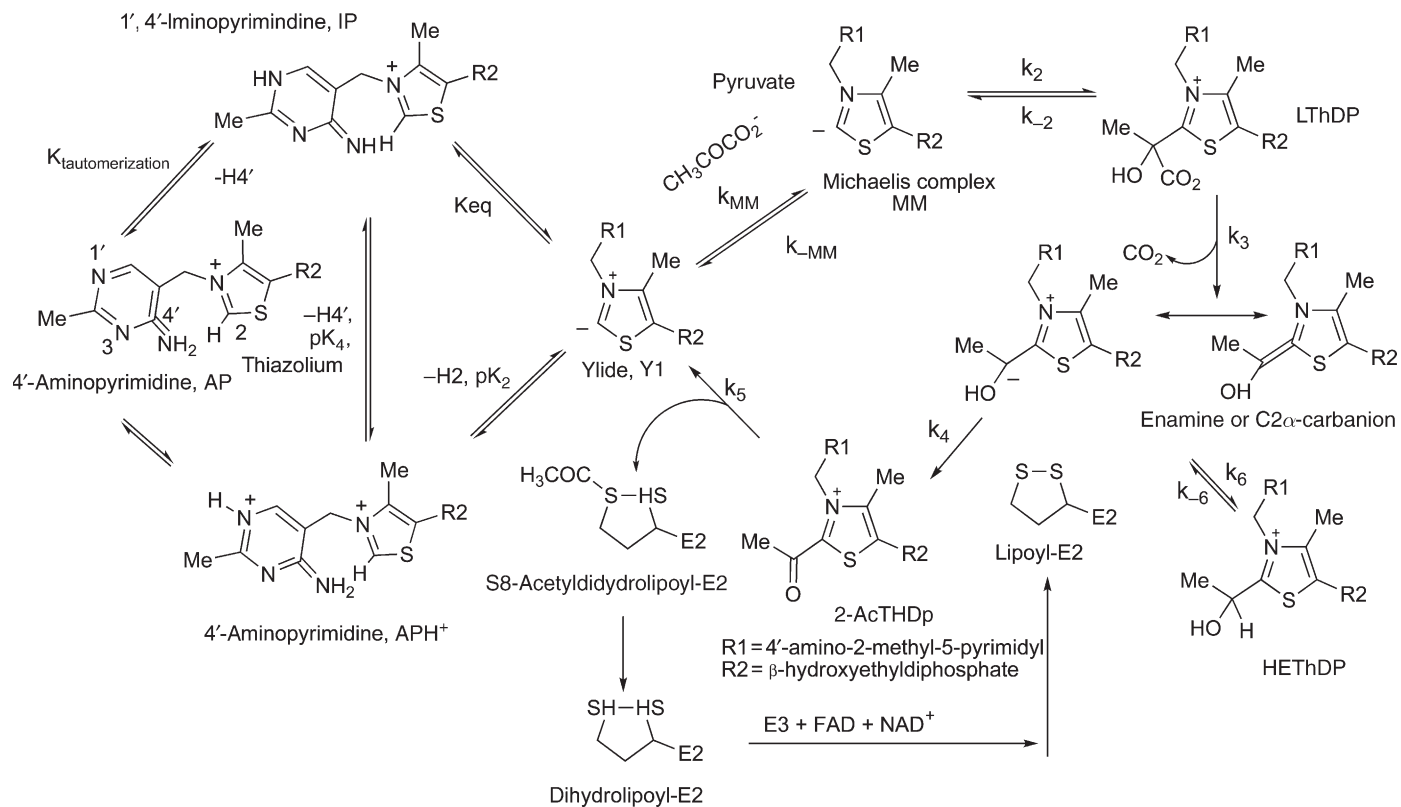
coenzyme (F) and coenzyme with a C2 substituent (S)<sup>24–26</sup>; (2) This V conformation brings the N4' and C2 atoms to within 3.5 Å of each other, certainly consistent with participation of the 4'-aminopyrimidine ring in catalysis, as suggested by two groups over the years,<sup>9,27–29</sup> but receiving limited enthusiasm until the structures

**Table 1** Representative enzymatic reactions of thiamin diphosphate referred to in this chapter

<i>Name of enzyme</i>	<i>Type of reaction catalyzed</i>	<i>Reaction</i>	<i>Cofactors in addition to ThDP and divalent metal</i>
Pyruvate decarboxylase	Decarboxylation of pyruvic acid	$\text{CH}_3\text{COCOOH} \rightarrow \text{CH}_3\text{CHO} + \text{CO}_2$	None
Pyruvate dehydrogenase complex	Oxidative decarboxylation of pyruvic acid	$\text{CH}_3\text{COCOOH} + \text{NAD}^+ + \text{CoASH} \rightarrow \text{CH}_3\text{COSCoA} + \text{NADH} + \text{CO}_2$	$\text{NAD}^+$ , lipoic acid, FAD, coenzyme A
Pyruvate oxidase	Oxidative decarboxylation of pyruvic acid	$\text{CH}_3\text{COCOOH} \rightarrow \text{CH}_3\text{COOH}$ or to $\text{CH}_3\text{COOPO}_3^- + \text{CO}_2$	FAD
Benzoylformate decarboxylase	Decarboxylation of benzoylformic acid	$\text{C}_6\text{H}_5\text{COCOOH} \rightarrow \text{C}_6\text{H}_5\text{CHO} + \text{CO}_2$	None
Benzaldehyde lyase	Reversible benzoin condensation, a carboligase/lyase	$\text{C}_6\text{H}_5\text{COCH(OH)C}_6\text{H}_5 \rightarrow 2\text{C}_6\text{H}_5\text{CHO}$	None
Glyoxylate carboligase	Decarboxylase and carboligase	$2\text{CHOCOOH} \rightarrow \text{HOOCCH}_2\text{CHO}$	Catalytically inactive FAD
Pyruvate ferredoxin oxidoreductase	Oxidative decarboxylation of pyruvic acid	$\text{CH}_3\text{COCOOH} + \text{Fe}_4\text{S}_4 + \text{CoASH} \rightarrow \text{CH}_3\text{CoSCoA} + \text{CO}_2$	$\text{Fe}_4\text{S}_4$ clusters, coenzyme A
Transketolase	Carboligase/lyase	$\text{Xylose-5-P} + \text{Ribose-5-P} \rightarrow \text{Sedoheptulose-7-P} + \text{Glyceraldehyde-3-P}$	None
Acetolactate synthase	Decarboxylase and carboligase	$2\text{CH}_3\text{COCOOH} \rightarrow \text{CH}_3\text{COC(CH}_3\text{)(OH)COOH} + \text{CO}_2$	Catalytically inactive FAD
$\text{N}^2$ -(2-carboxylethyl)arginine synthase	Transferase	$\text{D-Glyceraldehyde-3-P} + \text{Arginine} \rightarrow \text{N}^2$ -(2-Carboxylethyl)-L-arginine + Phosphate	None



**Scheme 1** Mechanism of yeast pyruvate decarboxylase YPDC.



**Scheme 2** Mechanism of *E. coli* pyruvate dehydrogenase complex with role of ThDP on E1ec.

appeared. With the renewed interest in the topic, a need arose to identify spectral signatures for various tautomeric and ionization states of the 4'-aminopyrimidine ring of ThDP.

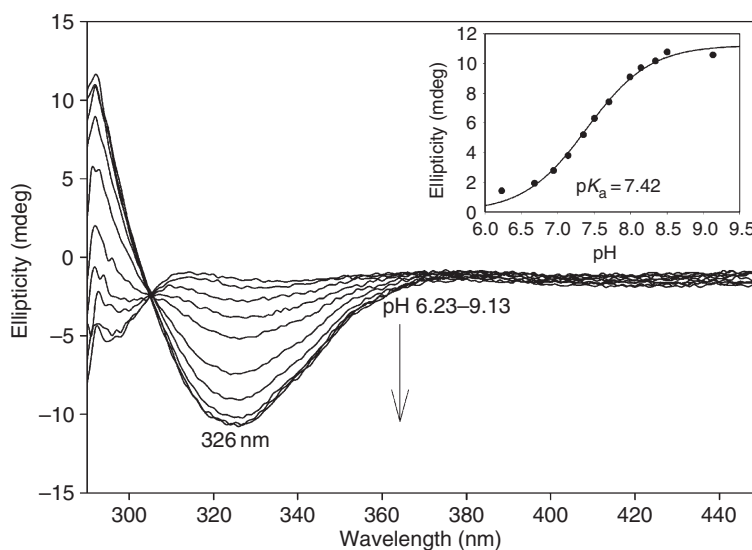
#### 7.16.2.1.1 The 4'-aminopyrimidine form of thiamin diphosphate

The signature for this species is a negative circular dichroism (CD) band centered near 320–330 nm and is well illustrated with the enzyme benzaldehyde lyase (BAL) (see near-UV CD at high pH in **Figure 2**). While this CD band has long been observed on the enzyme TK,<sup>30</sup> it had been suggested to be the result of a charge transfer (CT) transition between ThDP and an amino acid side chain on TK. A number of studies on YPDC and E1ec and their variants, as well as chemical model studies in the author's laboratory strongly suggest that in fact this UV/CD band is due to a CT transition between the 4'-aminopyrimidine ring as donor and the thiazolium ring as acceptor. The band has now been observed on a number of ThDP enzymes and as to whether or not it is observed depends more on the pH of measurement and also to a significant extent on the enzyme environment.

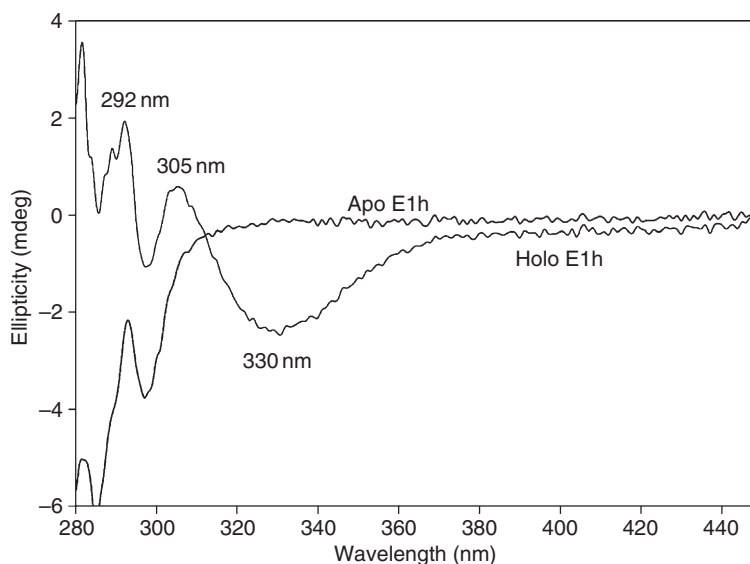
#### 7.16.2.1.2 The 1',4'-iminopyrimidine form of thiamin diphosphate<sup>31–35</sup>

The notion that the 4'-aminopyrimidine could exist in the 1',4'-iminopyrimidine tautomeric form was suggested earlier by models attempting to mimic the reactivity of such a tautomer. The N1'-methyl analog of both the 4-aminopyrimidine ring and of ThC itself was synthesized and gave evidence of two relevant points: (1) in this N1'-methylpyrimidinium the  $pK_a$  of the exocyclic amine was reduced to  $\sim 12$ – $12.5$ ,<sup>27,36</sup> offering early rationalization for the presence of conserved glutamate as a catalyst for the amino=imino tautomerization and (2) with the positive charge on the ring, the amino protons undergo differential exchange rates and the exchange is buffer catalyzed.<sup>29</sup> The possibility that the imino tautomer may indeed have a spectroscopic signature was first indicated by Sergienko's results on the slow E477Q variant of YPDC.<sup>31</sup> Inspired by these results, the old models were dusted off and in a series of chemical model studies, Zhang<sup>31</sup> and later Baykal<sup>36</sup> showed that an appropriate chemical model will give rise to a UV absorption in the 300–310 nm range. It bode well for future studies that the <sup>15</sup>N chemical shifts of the three species on the left-hand side of **Schemes 1 and 2**, the two neutral and one positively charged forms of the 4'-aminopyrimidine are quite distinct (earlier nuclear magnetic resonance (NMR) studies on this issue were published 30 years ago by J. D. Roberts and coworkers<sup>37</sup>).

The recognition that the CD bands corresponding to the AP and IP forms have different phases, allowing us to observe both bands simultaneously notwithstanding their proximity, made the CD method particularly useful for such studies. The signature for this IP species is a positive CD band centered near 300–314 nm and is well illustrated on E1h, where both the IP and AP tautomeric forms can be seen simultaneously (**Figure 3**). To the author's knowledge, no electronic absorption characteristic of the APH<sup>+</sup> form or the YL has yet been proposed.



**Figure 2** pH-dependence of the CD band at 326 nm corresponding to the AP form of ThDP on BAL.



**Figure 3** CD spectra of E1h in the absence (apo E1h) and presence of ThDP (holo E1h). The positive CD band at 305 nm is assigned to IP and the negative at 330 nm to the AP tautomers of ThDP.

### 7.16.2.1.3 Determination of $pK_a$ for the enzyme-bound $APH^+$ form<sup>35</sup>

As the pH is lowered in **Figure 2**, the amplitude of the band for the AP form diminishes and titrates with an apparent  $pK_a = 7.42$  for the  $(AP+IP)/[APH^+]$  equilibrium on BAL. This  $pK_a$  in water for ThDP is 4.85,<sup>37</sup> while on the enzymes it ranges from 5.6 to 7.5 (**Table 2**).<sup>35</sup> It was concluded from data in **Table 2** that the  $pK_a$  for the  $APH^+$  coincides with the pH of optimum activity for each enzyme, indicating that all three forms IP, AP, and  $APH^+$  must be readily accessible during the catalytic cycle. The  $pK_a$  elevation on the enzymes could be rationalized by the presence of the highly conserved glutamate near the N1' position of ThDP (residue E571 on E1ec) that would tend to make the 4'-aminopyrimidine ring more basic. The tautomeric equilibrium constant  $K_{\text{tautomer}}$  in conjunction with the  $pK_a$ s led to novel insight regarding ThDP catalysis, best viewed by the thermodynamic box for enzymes that are not substrate activated (left-hand side of **Schemes 1 and 2**), such as E1h and pyruvate oxidase from *Lactobacillus plantarum* (POX). For these enzymes, both the IP and AP forms could be monitored over a wide pH range, providing both  $pK_a$  and  $K_{\text{tautomer}}$  within reasonable error limits. The equilibria shown in **Schemes 1 and 2** are valid prior to addition of substrate and lead to the following tantalizing conclusions: (1) On POX and E1h,  $pK_1$  and  $pK_4$  have similar magnitudes; the enzymes shifted the  $pK_4$  from 12 in water<sup>27</sup> to 5.6 and 7.0, respectively (see the left triangle in **Schemes 1 and 2**). (2) With a known forward rate constant from  $APH^+$  to the YI of  $\sim 50 \text{ s}^{-1}$  determined for E1h,<sup>38</sup> and assuming a diffusion-controlled reverse protonation rate constant of  $10^{10} \text{ s}^{-1} \text{ mol}^{-1}$  (giving a  $pK_2$  of 8.3 on E1h compared to an estimate in water of 17–19<sup>39</sup>), one could next speculate about the right triangle in **Schemes 1 and 2**. The most important conclusion is that the proton-transfer equilibrium constant for  $[IP]/[YI]$  is  $10^1$ – $10^2$  on E1h. These thermodynamic parameters are the first estimates on any ThDP enzyme and should be generally applicable to ThDP enzymes. The results also suggested conditions under which a significant fraction of the thiazolium ring may be in the conjugate base YI form.

**Table 2** Correlation of  $pK_a$  of enzyme-bound  $APH^+$  and pH optimum of enzyme activity<sup>35</sup>

Enzyme	pH optimum activity	$pK_a$ for the $([AP]+[IP])/([APH^+])$
BAL	6.5–7.5	$7.42 \pm 0.02$
BFDC	6.0–8.5	$7.54 \pm 0.11$
POX	5.6–6.2	$5.56 \pm 0.03$
E1h	7.0–7.5	$7.07 \pm 0.07$



#### 7.16.2.1.4 The C2-carbanion/ylide/carbene

According to the Breslow findings, proton loss at the thiazolium C2 position is required to initiate the catalytic cycle. In 1997 there were two reports with significant implications regarding this issue:

1. According to  $^{13}\text{C}$  NMR measurements, on the YPDC the thiazolium ring C2H of bound ThDP is in its undissociated state both in the absence and in the presence of the substrate activator surrogate pyruvamide (this enzyme has long been known to be substrate activated), that is, no evidence was found for the presence of the conjugate base in the activated or unactivated forms of the enzyme.<sup>40</sup> In the same report, a  $^1\text{H}$ -NMR-based method was developed for measuring the rate of H/D exchange at the C2 position of bound ThDP, providing the rate constant for the dissociation of the C2H to the YI.
2. Coincidentally, in a series of groundbreaking publications in organic chemistry, Arduengo and colleagues<sup>41</sup> showed that the conjugate bases of imidazolium and indeed of thiazolium salts could be generated and their structure evaluated by NMR methods. These 'stable carbenes' have been used to synthesize some of the most important currently used catalysts, for example, for olefin metathesis. Most relevant to this review, the report by Arduengo and coworkers showed that the  $^{13}\text{C}$  chemical shift of the C2 resonance shifted from 157 to 253 ppm on conversion of their model thiazolium compound to its conjugate base, thereby providing the all important guide for future attempts to observe the YI. It is important to appreciate that determination of the state of ionization and tautomerization of enzyme-bound ThDP by solution NMR methods poses several challenges, both in the absence and in the presence of substituents at the C2 atom: (1) the size of ThDP enzymes (>120 kDa) leads to broadened lines, (2) for many ThDP enzymes it is difficult to reversibly remove ThDP to replenish with labeled coenzyme, and of course (3) *de novo* synthesis required for specific labeling of ThDP is time consuming and pricy.

#### 7.16.2.2 Thiamin-Bound Intermediates with Substrate or Substrate Analog Present

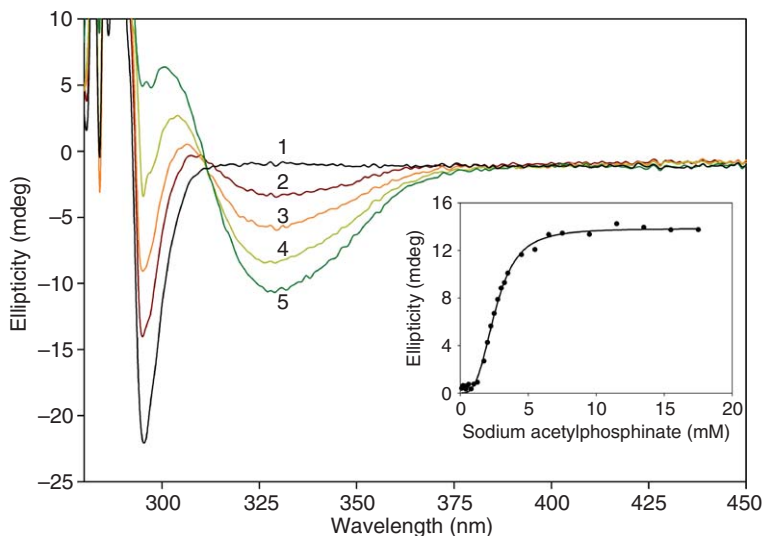
There is now available a very powerful array of methods to monitor the kinetic fate of each intermediate along the catalytic cycle of both YPDC and the 2-oxoacid dehydrogenase complexes represented in **Schemes 1 and 2**. A major breakthrough in looking at the rate constants for individual steps was proposed by Tittmann and Hübner (henceforth the TH method).<sup>42</sup> The method takes advantage of the known stability of the intermediates in **Schemes 1 and 2** under acidic conditions, so that using either rapid quench or manual quench methods one can 'freeze' the intermediates under acidic conditions while also precipitating the enzyme. Furthermore, the chemical shifts of the C6'H resonances of each intermediate are sufficiently distinct from each other and from that of the unsubstituted ThDP at the high magnetic fields currently available. This makes  $^1\text{H}$ -NMR an efficient method for evaluation of the relative concentration of the intermediates under steady-state conditions. Assuming a series of first-order reactions for the forward pathway, and with the known turnover number of the particular enzyme, rate constants could be calculated for individual reaction steps.

The author's group has demonstrated that mutations (substitutions of the protein variant) distant from the active site (such as with some of the loops, see Section 7.16.4.1) could have a dramatic influence not only changing the rate, but even changing the rate-limiting step. Therefore, further developments of this method were needed to address these issues on more complex systems such as with the multienzyme complexes. The ThDP-bound intermediates have been synthesized and their chemistry established by the author and others.<sup>14,15,42–45</sup>

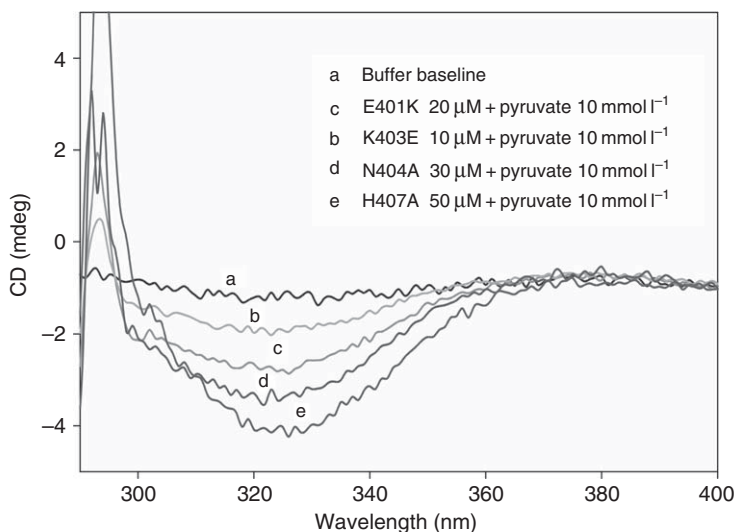
##### 7.16.2.2.1 The Michaelis–Menten complex

The earliest detection of a Michaelis–Menten complex (MC) was on addition of a substrate analog methyl acetylphosphonate (MAP) and acetylphosphinate to several ThDP enzymes. An example is shown with acetylphosphinate added to YPDC (**Figure 4**) leading to a negative CD band at  $\sim 325\text{--}335\text{ nm}$ , very reminiscent of the band observed for the AP form. In this example, the addition of ThDP alone (curve 1) did not display the AP form, the negative CD band only appeared after addition of substrate analog, hence the band must pertain to an MC-type complex.<sup>46</sup> Similar results were also seen when low concentrations of pyruvate were added to E1ec.<sup>33</sup>

Clear formation of the MC complex also resulted when adding pyruvate to the 'inner loop' E1ec variants (**Figure 5**).<sup>46</sup> Especially important support for the claim that the MC was indeed being detected is provided by kinetic measurements: both stopped-flow photodiode array (PDA) spectra in the absorption mode, and



**Figure 4** Near-UV CD spectra of YPDC. Spectrum 1. YPDC with  $0.1 \text{ mmol l}^{-1}$  ThDP,  $2.0 \text{ mmol l}^{-1}$   $\text{MgCl}_2$ . Spectra 2,3,4,5 after addition of 2–17.5  $\text{mmol l}^{-1}$  acetylphosphinate. The spectra show presence of 1',4'-iminophosphinolactyl-ThDP at 302 nm and of the Michaelis complex at 328 nm. Inset: dependence of 1',4'-iminophosphinolactyl-ThDP formation at 302 nm on acetylphosphinate.

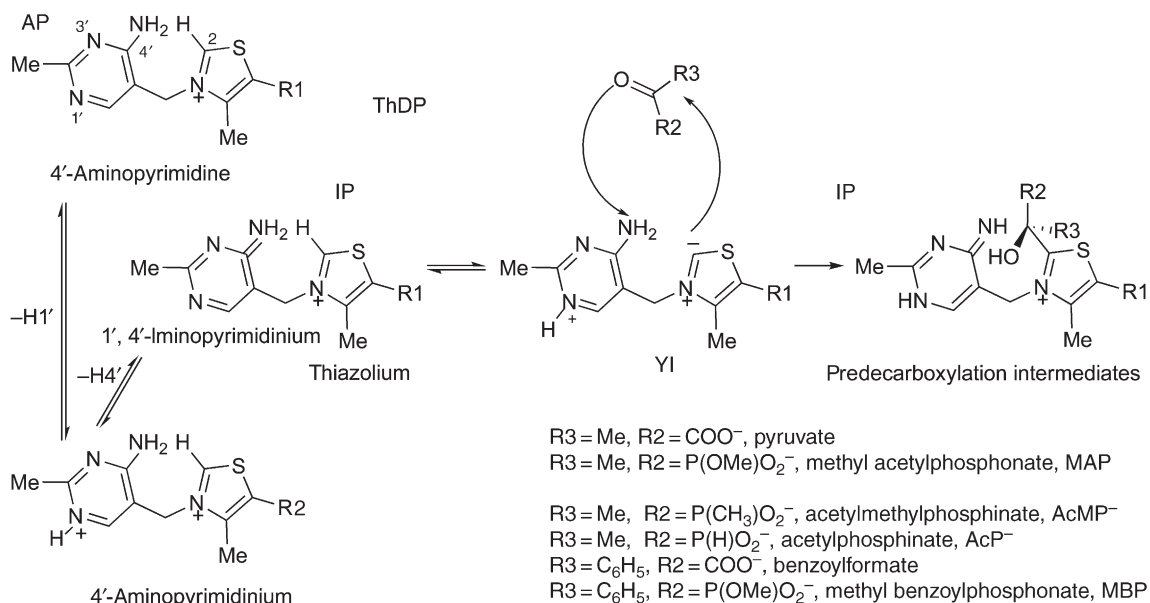


**Figure 5** Michaelis complex observed by CD with pyruvate added to ThDP-inner loop E1ec variants.

stopped-flow CD spectra at the appropriate wavelength, showed the formation of the absorbance/CD band attributed to MC formation, within the dead-time of the stopped-flow instruments ( $<1 \text{ ms}$ ), as expected of a noncovalent MC (see **Figure 16**).

#### 7.16.2.2.2 The covalent substrate–thiamin diphosphate predecarboxylation complex (LThDP and analogs)

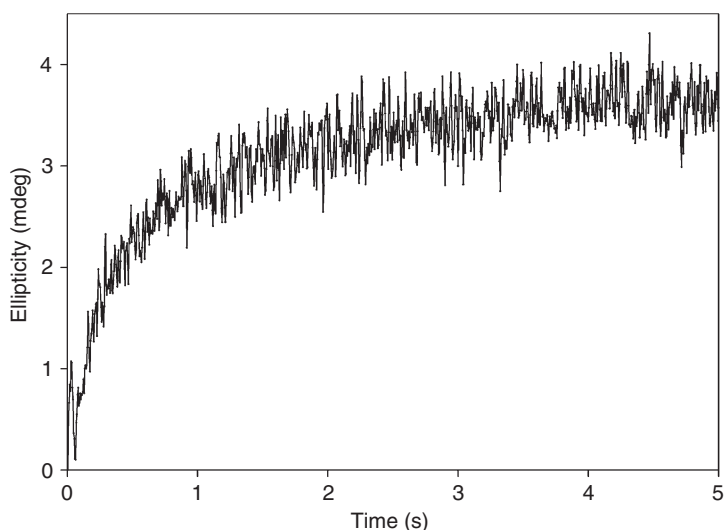
**7.16.2.2.2(i) Observation of the intermediate analogs derived from substrate analog phosphonates and phosphinates** The initial identification of the IP form (positive CD band, 300–314 nm) resulted from the formation of a stable predecarboxylation adduct of ThDP with (1) MAP or acetylphosphinate<sup>32,33</sup>



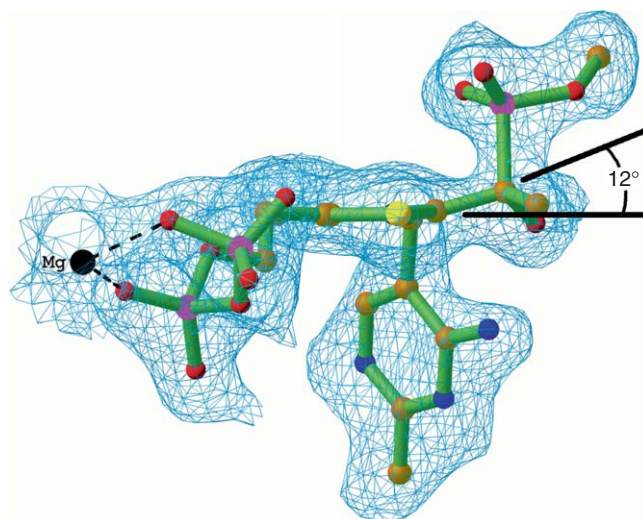
**Scheme 3** Mechanism of formation of LThDP and analog adducts.

(CH<sub>3</sub>C(C=O)P(H)O<sub>2</sub>Na)<sup>35</sup> and (2) the aromatic 2-oxo acid analog methyl benzoylphosphonate (MBP),<sup>47,48</sup> according to **Scheme 3**. With six enzymes tested so far, the IP form appeared on the stopped-flow timescale (either absorption or CD mode): the reaction is efficiently catalyzed by all of the enzymes (**Figure 6**). There is an important additional finding shown in **Figure 4** resulting from mixing YPDC and acetylphosphinate.<sup>35</sup> Since we are seeing evidence for coexistence of the MC and the covalent predecarboxylation intermediate, the results are consistent with ‘alternating active site reactivity’ suggested by the author’s group earlier for YPDC and benzoylformate decarboxylase (BFDC).<sup>49,50</sup>

The product phosphonamandelylThDP (PMThDP) formed on BFDC from MBP and ThDP was also confirmed (Fourier-transform mass spectrometry (FT-MS)) in solution,<sup>48</sup> and of PLThDP (from MAP. ThDP) by X-ray methods on E1ec<sup>51</sup> and on POX.<sup>52</sup> The X-ray structures of the PMThDP adduct on BAL



**Figure 6** Reaction of E1ec + acetylphosphinate by stopped-flow CD. Predecarboxylation intermediate formed with *ks* of  $4.44 \pm 0.34 \text{ s}^{-1}$  and  $0.593 \pm 0.064 \text{ s}^{-1}$ .



**Figure 7** Nonplanar distortion at PLThDP's C2 $\alpha$  atom relative to the thiazolium ring on E1ec.

and BFDC have also been determined to high resolution and are informative,<sup>47,48</sup> as they provide unequivocal proof of configuration at the C2 $\alpha$  position, not accessible with the pyruvate analogs (the methyl and OH groups have similar electron densities at resolutions currently achieved).<sup>51,52</sup>

Additionally, all LThDP-like structures displayed some distortion where the C2–C2 $\alpha$  bond is found noncoplanar with the thiazolium ring (unlike in the structure of PLThDP in the absence of enzyme), although the Dunathan hypothesis (according to which the C2 $\alpha$ –P bond should be essentially perpendicular to the thiazolium plane) is adhered to (Figure 7).<sup>47,48,51,52</sup>

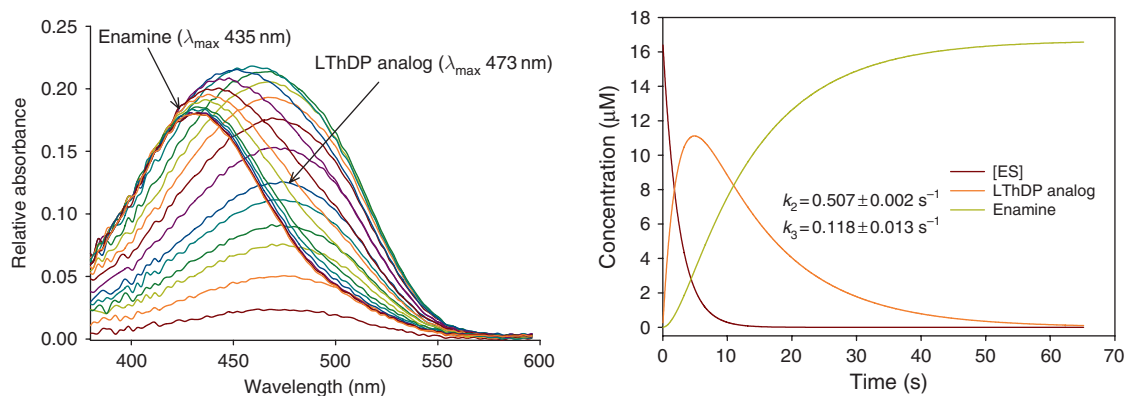
In some favorable cases, such as with BAL, the CD band for the predecarboxylation intermediate (via the IP form) could be observed from the slow substrates benzoylformate or phenylpyruvic acid.<sup>53</sup> This is feasible since BAL, while a carboligase/lyase enzyme, also catalyzes decarboxylation of aromatic 2-oxoacids, albeit very slowly.

The LThDP was also observed by cryocrystallographic X-ray methods on POX.<sup>52</sup>

**7.16.2.2(ii) Observation of LThDP analogs from chromophoric substrate analogs** Recently, the formation of the predecarboxylation adduct formed with ThDP from a chromophoric substrate analog (*E*)-2-oxo-4(pyrid-3-yl)-3-butenic acid (3-PKB) (as well as its ortho and para isomers) was also observed on three enzymes, YPDC, BFDC, and BAL.

With a stopped-flow photodiode array (SF-PDA) instrument, two transients resulted on mixing YPDC with 3-PKB: the first transient formed T<sub>1</sub> was assigned to the predecarboxylation intermediate analogous to LThDP (LThDP\* with  $\lambda_{\max}$  near 470 nm), while the second transient T<sub>2</sub> was assigned to the enamine ( $\lambda_{\max}$  near 430 nm). The second transient T<sub>2</sub> was formed at the same rate as the rate of depletion of T<sub>1</sub> (Figure 8; Scheme 4). In a series of studies on YPDC,<sup>54–56</sup> BFDC,<sup>57</sup> and BAL,<sup>53</sup> this compound provided outstanding information about the rates of formation of these two important intermediates, not available from any other source. With this method, evidence was obtained on YPDC whether an active center residue (1) had a role pre- or postdecarboxylation, (2) was involved in active center communication, and (3) led to the alternating active site behavior.<sup>54,55</sup> The 3-PKB with YPDC mobile loop variants (see Section 7.16.4) showed that substitutions along the 290–304 loop of YPDC could change the relative barriers for the two steps, not noted before.<sup>56</sup> Where applicable, this method is complementary to the TH method, which could not differentiate the enamine from the HETThDP in Scheme 1; these species are at protolytic equilibrium, the acid quench converts the enamine to HETThDP.

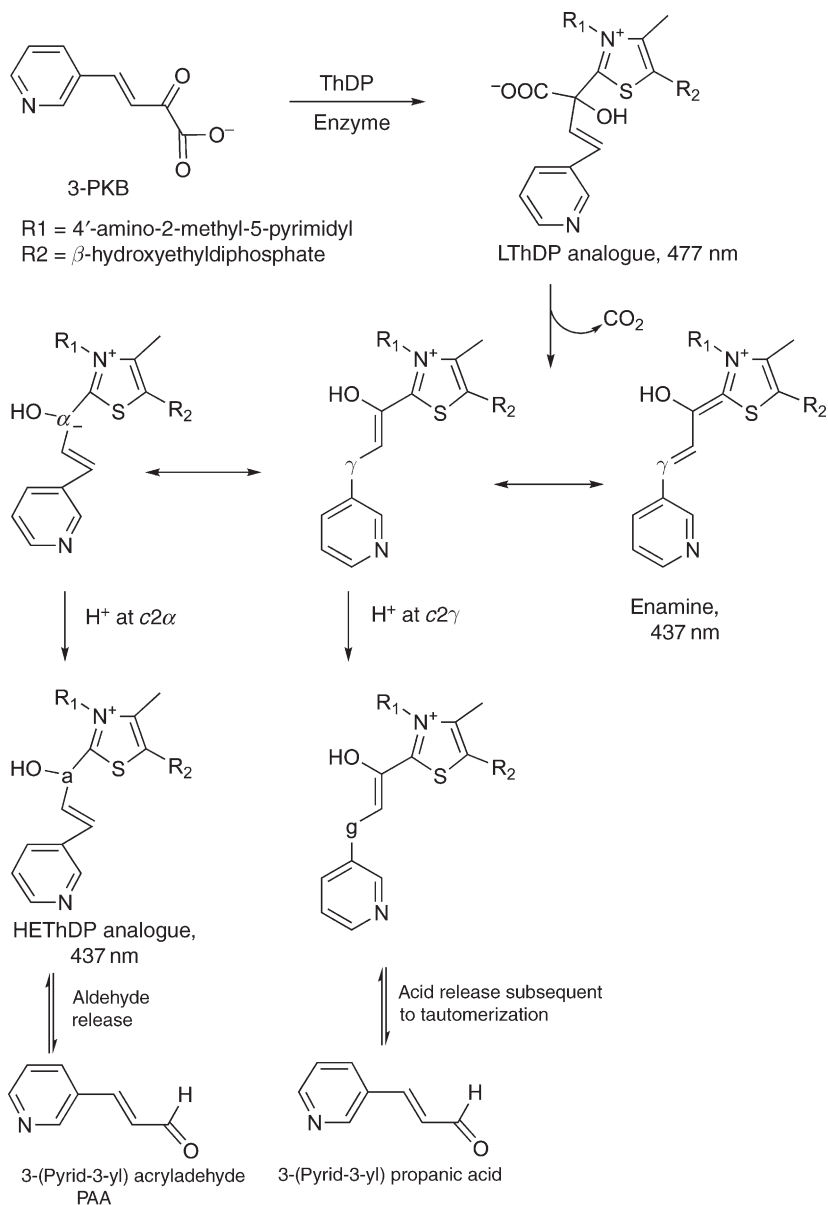
The following evidence was obtained to support the assignments. (1) A thiazolium model compound (without the 4-aminopyrimidine) was synthesized, which on treatment with strong base could generate the enamine, giving an absorbance with  $\lambda_{\max}$  of 435 nm,<sup>58</sup> the same wavelength found with 3-PKB or its isomers on YPDC, or indeed from the phenyl analog of 3-PKB (*E*)-C<sub>6</sub>H<sub>5</sub>CH=CHCOCOOH studied some years ago in



**Figure 8** SF-PDA spectra on mixing YPDC and 3-PKB. Left: YPDC ( $33.2 \mu\text{mol l}^{-1}$  active sites) mixed with  $20 \text{ mmol l}^{-1}$  3-PKB. Spectra recorded every 40 ms. Right: Time course of concentration changes of intermediates from deconvolution;  $[\text{ES}]_0 =$  active sites.

the author's laboratory.<sup>59</sup> Therefore, the assignment of the transient species  $T_2$  observed near 430–440 nm to the enamine has solid support from chemical models. (2) Assignment of the 475–480 nm transient  $T_1$  was based, in part, on the following experimental data: (1) It had been found on the basis of results with six enzymes (BAL, BFDC, YPDC, POX, E1ec, and E1h) that the ThDP is in its IP form whenever there is a tetrahedral  $C2\alpha$  carbon attached to the  $C2$  carbon (i.e., when there are four substituents at the  $C2\alpha$  carbon), while the AP form predominates with only three substituents at the  $C2\alpha$  carbon, as in the enamine. In the case of BAL, simultaneous formation of both the IP form at 310 nm and of the 475 nm band from (*E*)-3-(pyrid-3-yl)-2-propenal (PAA, the product of decarboxylation of 3-PKB) could be observed by CD, affirming that this transition does not correspond to the enamine.<sup>57</sup> (2) Essentially the same absorbance resulted when 3-PKB or its decarboxylated product PAA was added to BAL or BFDC, suggesting a similar type of tetrahedral intermediate either the LThDP-like (predecarboxylation) or the HEThDP-like (postdecarboxylation, resulting from protonation of the enamine at  $C2\alpha$ ). (3) The transient  $T_1$  at 475 nm on the three enzymes studied (BAL, BFDC, and YPDC) always preceded the formation of the transient  $T_2$  corresponding to the enamine at 435 nm. The time course therefore suggests that on the decarboxylation pathway, the intermediate corresponding to the 475 nm transient precedes the enamine, leaving only the LThDP-like intermediate or the MC. As mentioned above, the stopped-flow evidence on the MC showed that it is formed within the dead-time (1–2 ms) of the instrument, ruling out the MC as being responsible for the 475 nm excitation, and leaving the LThDP-like intermediate as the most plausible one corresponding to it. (4) In earlier studies *p*-nitrobenzoylformic acid (NBFA) was used as an alternate substrate for both YPDC and BFDC. With BFDC, rather analogous results were obtained with NBFA to those seen with 3-PKB: in a stopped-flow PDA experiment two consecutive transients were formed,  $T_1'$  with  $\lambda_{\text{max}}$  near 620 nm, the second one  $T_2'$  with  $\lambda_{\text{max}}$  near 400–410 nm.<sup>50</sup> The  $T_2'$  was again assigned to the enamine (the conjugated system is shorter by a  $C=C$  bond than with 3-PKB), and  $T_1'$  to the predecarboxylation intermediate once more. (5) Finally, Brandt at Brandeis succeeded in growing crystals of BFDC with both PAA and 3-PKB and solved the structure to high resolution.<sup>57</sup> The structure with PAA clearly indicated covalent binding to ThDP as the  $C2\alpha$ -hydroxymethyl derivative with the vinylpyridyl substituent attached to the  $C2\alpha$  atom, clearly suggesting a tetrahedral rather than a trigonal environment at that atom. The finding that the thiazolium and vinylpyridine planes were not coplanar with each other also argues against an enamine structure.

These results with the chromophoric substrates raise an interesting issue regarding the nature of the electronic transition responsible for the  $T_1$ , as such observations have now been made on several ThDP enzymes. What is the source of the electronic transition responsible for the transient  $T_1$  at 475 nm derived from either 3-PKB or PAA and seen on three different ThDP enzymes, and of the 620 nm transient  $T_1'$  observed on BFDC derived from NBFA? The author and coworkers have suggested that the broad absorptions correspond to a CT transition between a donor and acceptor. That ThDP can give such CT bands is not unexpected. The



**Scheme 4** Proposed reaction of 3-PKB on decarboxylases.

absorbance/negative CD band near 320–330 nm is a property of the enzyme-bound ThDP,<sup>32,33</sup> and while the intensity may be enhanced by some of the amino acids in the immediate vicinity of the cofactor, extensive studies on TK<sup>60</sup> and E1ec<sup>33</sup> pretty much ruled out any amino acids acting as a donor or acceptor. Rather, the thiazolium ring and the AP form of the ThDP were suggested as acceptor and donor in this CT transition. It is noteworthy that these two rings are not parallel to each other as the coenzyme assumes the V conformation. Nor is there a direct  $\pi$ -type interaction between the two aromatic rings since they are separated by a methylene group. Given that with the LThDP or HEThDP (or their analogs) the IP tautomer predominates, a possible interaction between the pyridine (or pyridinium) rings of 3-PKB or PAA as acceptor and the 4'-iminopyrimidine as donor were suggested. While the pyridinium ionization state would be more appealing as an electron acceptor than pyridine, the  $pK_a$  for the pyridinium salt of 3-PKB in aqueous solution is 4.3, and it is unknown if this  $pK_a$  would be significantly perturbed on the enzyme. The recent publication on the ThDP enzyme



$N^2$ -(2-carboxyethyl)arginine synthase suggested that the key ThDP-bound intermediate 2-acryloylThDP gives rise to a broad absorption with  $\lambda_{\max}$  near 430 nm.<sup>61</sup> With considerable effort, the authors could not reproduce such a long-wavelength absorption via a variety of synthetic models mimicking likely intermediates. A possible explanation of the results on CEAS with the acryloyl side chain at C2, and on YPDC, BFD, BAL with the side chains derived from 3-PKB and PAA, is that the CT bands are a result of cation- $\pi$  interactions: the thiazolium ring is the cation (acceptor) and the double-bonded substituent at C2 $\alpha$  is the donor (it should be noted that in the enamine the positive charge on the thiazolium ring is significantly diminished, and correspondingly there is no CT band observed). While the origin of the CT bands remains an enigma, their very existence is not in doubt, nor is the time course of their formation: formation of the T<sub>1</sub> and T<sub>1'</sub> precede formation of the enamine in each case and are best assigned to predecarboxylation intermediates rather than to the enamine. On BAL, Nemeria observed simultaneous presence of both the negative CD band at 475 nm and the positive one at 312 nm on addition of PAA: the former a direct result of the tetrahedral adduct between PAA and BAL, the band at 312 nm of the IP form characteristic of C2 $\alpha$ -tetrahedrally substituted intermediates.<sup>57</sup>

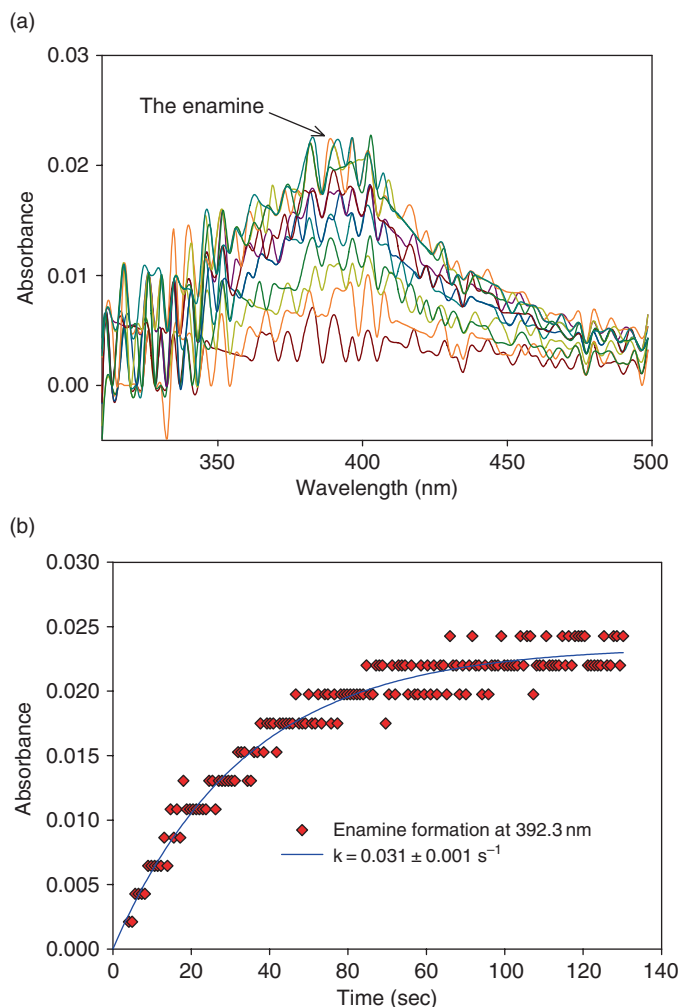
It is also to be noted that semiempirical calculations predict that the two configurations of the enamine (cis and trans) should have virtually identical  $\lambda_{\max}$  values and the enamines derived from NBFA, 3-PKB, or BFA always have their  $\lambda_{\max}$  values in a position predicted by synthetic models.

**7.16.2.2(iii) Reactivity of LThDP on and off the enzymes** The Breslow mechanism has stood the test of time for 50 years,<sup>21</sup> and with the ability to synthesize the putative intermediates LThDP and HETHDP (thanks mostly to Kluger's group<sup>3</sup>), a reexamination of the reactivity of LThDP on and off the enzymes could be carried out, leading to the following results:<sup>42,62</sup> (1) Decarboxylation of model C2 $\alpha$ -lactylthiazolium and C2 $\alpha$ -lactylthiamin salts could achieve first-order rate constant values essentially identical to the enzymic turnover numbers ( $\sim 60 \text{ s}^{-1}$ ) in aprotic solvents.<sup>62</sup> Hence, LThDP can undergo decarboxylation at rates that are catalytically competent with the turnover numbers. In other words, a low polarity active center may be sufficient to ensure that decarboxylation of LThDP is not rate limiting. (2) Both YPDC and E1ec were reconstituted with synthetic LThDP and it was shown that both enzymes catalyze decarboxylation (unpublished observations at Rutgers). Rates at steady state are limited by the slow off-rate for ThDP, so a new molecule of LThDP can be bound. Pre-steady-state experiments were also carried out and demonstrated that the reaction rates with LThDP reconstitution are limited by factors other than catalysis, that is, binding and/or conformational changes. (3) Rapid acid quench of certain YPDC variants reacting with pyruvate enabled direct observation of LThDP under steady-state conditions.<sup>42</sup> It was concluded that LThDP is indeed on the pathway of both YPDC and E1ec.<sup>62</sup>

### 7.16.2.2.3 The first postdecarboxylation intermediate: The enamine/C2 $\alpha$ -carbanion

According to the **Schemes 1 and 2**, the enamine is the only covalent thiamin-bound intermediate capable of being conjugated. Electronic spectral observation of the enzyme-bound enamine derived from aliphatic substrates is difficult due to the expected  $\lambda_{\max}$  near 290–295 nm, according to thiazolium-based models.<sup>63,64</sup>

The enamine intermediate derived from benzoylformate (modeled with  $\lambda_{\max}$  of 380 nm)<sup>65,66</sup> has been observed directly on the enzyme BFDC at 390 nm.<sup>57</sup> While BFDC converts benzoylformate to benzaldehyde, the enzyme also catalyzes a benzoin-type condensation of benzaldehyde in the reverse reaction, similar to the reverse reaction of BAL, and both BFDC and BAL have been used in the chemoenzymatic synthesis of  $\alpha$ -ketols. When reacting the benzaldehyde product with BFDC, there appeared an absorbance (and a CD band) at 390 nm, in the wavelength region predicted by models,<sup>65,66</sup> but no CD band was evident in the 300–310 nm region.<sup>57</sup> Also, when (*R*)-benzoin was added to BAL, the same CD band was formed at 390 nm indicating slow release of the first benzaldehyde, and the stability of the enamine in the forward direction (**Figure 9**).<sup>53</sup> These experiments provided fundamental information: (1) the 'real' enamine could be observed (due to its long  $\lambda_{\max}$  at 390 nm) for the first time derived from benzoin or benzaldehyde; (2) the enamine was found to be in its AP or APH<sup>+</sup>, but not in its IP form; and (3) since it gives rise to a CD signal, the enamine is chiral on the enzyme, even though it is planar and conjugated. Almost certainly, this is imparted by the conserved 'V' conformation of ThDP on all enzymes. This reinforces our findings that all intermediates discussed here are chiral on the enzymes, both pre- and postsubstrate addition. While binding to the chiral protein is expected to induce chirality in all



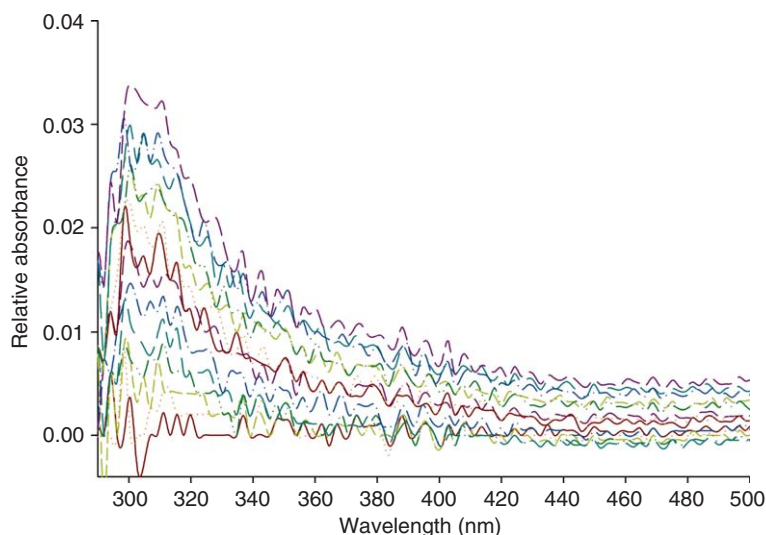
**Figure 9** Stopped-flow PDA experiment on reacting (*R*)-benzoin with BAL. (a) Difference spectra showing enamine formation on mixing BAL ( $33.94 \text{ mmol l}^{-1}$  active site concentration) with an equal volume of  $4 \text{ mmol l}^{-1}$  (*R*)-benzoin at  $20 \text{ }^{\circ}\text{C}$ . The reaction was monitored for 2 min and spectra were recorded every 2.5 ms. (b) Rate of formation of enamine on BAL from (*R*)-benzoin from data in (a). Data were fitted to a single exponential.

ThDP-related intermediates, the CD measurements nicely confirm this expectation for all intermediates, both covalently bound and even noncovalent one such as the MC.

With YPDC, BFDC, and BAL, the enamine could be observed directly at 430 nm with 3-PKB as alternate substrate as seen in [Figure 8](#) on YPDC.<sup>53–57</sup>

#### 7.16.2.2.4 The second postdecarboxylation intermediate, the product-ThDP complex (HEThDP and HBThDP)

**7.16.2.2.4(i) Evidence on YPDC** Evidence for the formation of this intermediate also will be reported ([Table 3](#)) using the TH method. Clear evidence was obtained for HEThDP analog formation from reacting 3(pyrid-3-yl)-acrylaldehyde (PAA) (the product of decarboxylation of 3-PKB), with BAL or BFDC.<sup>57</sup> An absorbance with  $\lambda_{\text{max}}$  470 nm appeared (similar to that observed with 3-PKB in [Figure 8](#)), and was attributed to the HEThDP analog. Recent X-ray results clearly confirmed formation of the tetrahedral postdecarboxylation intermediate on addition of PAA to BFDC.



**Figure 10** SF-PDA spectra from mixing YPDC with the product acetaldehyde.

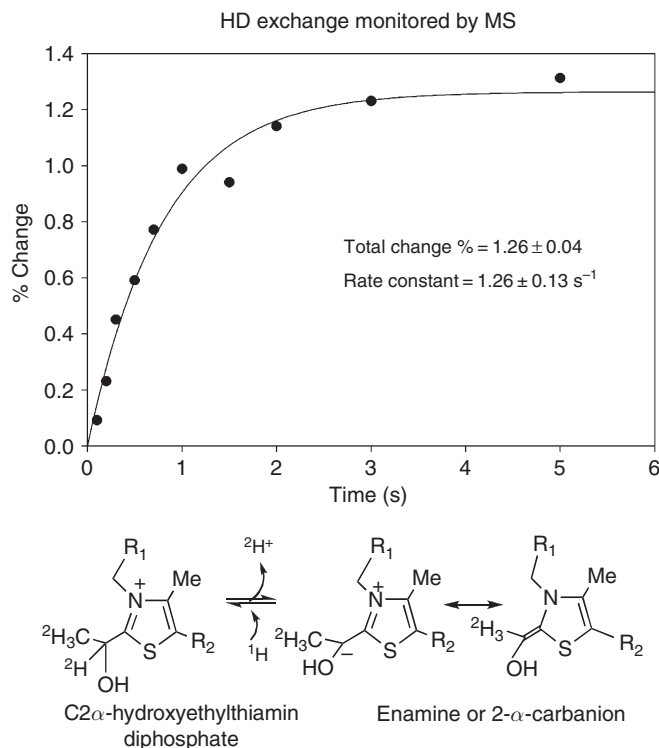
The most striking confirmation of HEThDP formation resulted from mixing acetaldehyde, the product of pyruvate decarboxylation, with YPDC on the SF-PDA instrument (**Figure 10**),<sup>67</sup> giving the characteristic absorption for the IP form,  $\lambda_{\max} = 310$  nm. Here, there is simply no alternative assignment than to the IP form of HEThDP (see **Scheme 1**).

**7.16.2.2.4(ii) Evidence on E1ec**<sup>33,68</sup> While HEThDP is not usually considered to be on the reaction pathway of ThDP-dependent oxidative decarboxylases (**Scheme 2**), groups working with POX<sup>69</sup> and the pyruvate dehydrogenase complexes (PDHcs)<sup>70</sup> have long used HEThDP as an alternate substrate. Chemical model studies from the author's laboratory on the enamine and related intermediates<sup>10,14,15</sup> have unambiguously signaled that HEThDP cannot be oxidized directly by either FAD or lipoic acid as oxidizing partners. Instead, oxidation must be preceded by ionization at the C2 $\alpha$  position to generate the enamine, which is then prone to oxidation by even molecular dioxygen from air. However, the  $pK_a$  at the C2 $\alpha$  position is very high, estimated at  $\sim 17$ – $18$  for HEThDP<sup>71</sup> derived from pyruvic acid. Experiments on YPDC demonstrated substantial lowering of this  $pK_a$  by the enzyme, and a low effective dielectric constant at the active center of YPDC<sup>72</sup> was offered as explanation.

Given that HEThDP is not on the direct pathway of PDHc, reversibility of the reaction from HEThDP to enamine needed to be demonstrated on E1ec. If E1ec catalyzes the exchange in **Figure 8**, this could reflect the intervention of an enzymatic 'solvent effect'. We measured the E1ec-catalyzed pre-steady-state rate constant for the  $^2\text{H} \rightarrow ^1\text{H}$  exchange from the C2 $\alpha$  position of HEThDP- $^2\text{H}_4$  (steps  $k_6/k_{-6}$  in **Scheme 2**), as an indicator of the formation of the enamine (**Figure 11**). E1ec accelerates the rate of ionization of this bond by a factor of  $10^7$ , corresponding to  $10$  kcal mol $^{-1}$  stabilization of the enamine intermediate by the enzyme.<sup>73</sup> This stabilization is a property of the active center per se, and, with our earlier evidence reported on YPDC, suggests that such a 'solvent effect' is likely a general feature of ThDP enzymes.

#### **7.16.2.2.5 2-Acetylthiamin diphosphate, the two-electron oxidation product of the enamine**

This compound is the product of oxidation of the enamine by any one of the following oxidizing agents on enzymes: most commonly by the dithiolane ring of lipoic acid covalently amidated to a lysine side chain in the 2-oxoacid dehydrogenase multienzyme complexes; less frequently by FAD in the pyruvate oxidases – these come in two flavors, forming acetate in *Escherichia coli* and forming acetylphosphate in *L. plantarum*; finally by  $\text{NAD}^+$ .



**Figure 11** Time course of  $^2\text{H} \rightarrow ^1\text{H}$  exchange at the  $\text{C2}\alpha$  of  $\text{HEThDP-}^2\text{H}_4$  monitored by FT-ICR MS.

The reaction in POX has been studied in detail for many years and most recent evidence suggests that the oxidation takes place via single electron transfers with the likely intermediacy of the radical cation species delocalized onto the thiazolium ring (see the next section).

Oxidation of the enamine by lipoic acid is usually viewed as the reductive acetylation of lipoyllysyl-E2. An important mechanistic issue of whether electron and group transfer take place in a single step via a tetrahedral intermediate (essentially ‘cross-linking’ the ThDP-bound enamine on the E1 component with the lipoyl group on E2), or stepwise (where oxidation to the acetylThDP with concomitant reduction of lipoamide-E2 to dihydrolipoamide-E2 precedes acyl transfer to the S8 atom of the dihydrolipoamide-E2) was investigated by Frey and coworkers. Using a number of ingenious ways to generate 2-acetylThDP (e.g., by reversing the reaction by addition acetyl-CoA), they provided evidence that the first option, redox followed by group transfer is the likely scenario.<sup>45</sup> In a model system, Pan in the author’s group demonstrated that the tetrahedral intermediate can indeed be generated from the enamine and lipoic acid analogs, although on critical examination this could be the obligatory intermediate with either scenario.<sup>74</sup> Most importantly, Pan showed that the cleavage of the dithiolane ring by the enamine required electrophilic catalysis. In the model system this could be accomplished by S-methylation, while on the enzymes the reaction presumably uses an acid catalyst such as a histidinium ion, probably H407 on E1ec.<sup>75</sup>

#### 7.16.2.2.6 The $\text{C2}\alpha\text{-hydroxyethylideneThDP}$ radical, the one-electron oxidation product of the enamine

Early and clear evidence for a free-radical mechanism on ThDP enzymes was obtained on pyruvate-ferredoxin oxidoreductase, an enzyme that converts pyruvate to acetyl-CoA in anaerobes.<sup>76</sup> In addition to ThDP, the enzyme has three  $\text{Fe}_4\text{S}_4$  clusters forming a 40–50 Å long electron transfer chain.<sup>77</sup> The stability of the free radical was manifested by the fact that the crystal also displayed an electron paramagnetic resonance signal.<sup>77</sup> A chemical model was generated for the electrochemical oxidation of the enamine leading to dimerization at the  $\text{C2}\alpha$  atom, suggesting significant electron spin density at this atom.<sup>78</sup> Subsequent detailed work on

pyruvate-ferredoxin oxidoreductase clearly showed that the spin density is delocalized into the thiazolium ring but there indeed is a significant fraction at the C2 $\alpha$  atom.<sup>79</sup>

Recent work on pyruvate-ferredoxin oxidoreductase also suggested that the enzyme proceeds via a free-radical mechanism. While it is not formally a redox enzyme, the class of enzymes named acetohydroxyacid synthases and acetolactate synthases also have FAD in addition to the ThDP, although the function of FAD is still somewhat uncertain since the carboligase reactions carried out by these enzymes have no immediately obvious need for an oxidizing coenzyme (there is a detailed discussion of these issues in Kluger and Pike<sup>43</sup>).

A newer member of the nonredox ThDP- and FAD-containing enzymes is glyoxylate carboligase. This enzyme carries out decarboxylative carboligation. Perhaps the most striking feature of glyoxylate carboligase is the replacement of the virtually universally conserved Glu residue opposite the N1' atom of ThDP by a hydrophobic residue on the enzyme.<sup>80</sup> The glyoxylate carboligase and BAL are of particular interest since on neither enzyme is there an acid-base residue within hydrogen bonding distance of the ThDP to assist with proton transfers. There is perhaps no alternative to water carrying out the proton transfers on these two enzymes. Of course, these enzymes also provide strong support for an obligatory catalytic role of the 4'-aminopyrimidine ring of the ThDP.

Adam-Vizi and coworkers have suggested that reactive oxygen species in the mitochondrion target the citric acid cycle enzyme  $\alpha$ -ketoglutarate dehydrogenase.<sup>81</sup> As all other 2-oxoacid dehydrogenase complexes,  $\alpha$ -ketoglutarate dehydrogenase also comprises three principal components. There is evidence provided in the literature that the E1 component can produce the C2 $\alpha$ -hydroxyethylideneThDP radical.<sup>82</sup> Although, as to which of the components of this complex carries the free radical is not yet settled. This tantalizing evidence opens the possibility, although by no means proves, that (1) the 2-oxoacid dehydrogenase complexes carry out the reductive acetylation via a radical mechanism, as yet not clarified, and (2) the reactive oxygen species express their harmful effects on the reductive acetylation of E2 by E1, that is, on intercomponent communication. Much work remains to be done to further support either hypothesis.

### 7.16.3 Determination of Rate-Limiting Steps and Microscopic Rate Constants on ThDP Enzymes

Starting in the 1970s, both solvent and heavy-atom kinetic isotope effects (KIEs) were employed to probe the rate-limiting steps in ThDP enzymes, especially on YPDC. Some of the easier ones to interpret used natural abundance <sup>13</sup>C label at the pyruvate C1 atom to investigate <sup>13</sup>C/<sup>12</sup>C KIEs by measuring the mass ratio of <sup>45</sup>CO<sub>2</sub>/<sup>44</sup>CO<sub>2</sub> using an isotope ratio mass spectrometer.<sup>83–85</sup> Supported by a model study in which the decarboxylation step could be isolated,<sup>86</sup> the results of such studies could inform about the partitioning of the LThDP intermediate reverting to pyruvate and free coenzyme or going forward to decarboxylation to the enamine (Scheme 1). The method suggested that decarboxylation was not rate limiting on YPDC, the forward rate constant being four- to fivefold larger than the reversion to the Michaelis complex.

In a series of joint papers from the groups of Schellenberger and Hübner at Halle and Schowen at Kansas, the rate-limiting steps on  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_{\text{m}}$  were probed on YPDC by a variety of isotope effect measurements. Given that the effects of regulation are indeed an important issue in enzymology, and of course in physiology as well, YPDC with its substrate activation (Hill coefficient nearly 2.0) is an interesting enzyme to study. Especially, the solvent deuterium KIEs suggested that there may be a strong hydrogen bond formed during regulation and, along with heavy atom KIEs enabled the authors to construct a free energy diagram for the reaction pathway including the regulatory mechanism.<sup>87,88</sup> Subsequently, solvent deuterium KIEs were determined on some of the cysteine variants of YPDC (Cys221 is believed to be the site of substrate activation<sup>14,15</sup>), again attempting to delineate the effects of substrate activation.<sup>89,90</sup> While these studies provided valuable insight, they certainly had limitations regarding assignment of specific rate constants to microscopic steps, especially if one wished to apply such methods to the many variants created over the years by site-directed mutagenesis.

A more generally applicable method was reported in 2003 (referred to above as the TH method),<sup>42</sup> capable of trapping ThDP-bound intermediates shown in Schemes 1 and 2 by acid quench after rapid mixing of enzyme and substrate for predetermined time periods on a KINTEK chemical quench instrument.

**Table 3** Microscopic rate constants for the YPDC loop variants from the analysis of intermediate distribution<sup>91</sup>

<i>Variant</i>	$k_{\text{cat}} (\text{s}^{-1})$	$k'_2 (\text{s}^{-1})$ C–C bonding	$k'_3 (\text{s}^{-1})$ CO <sub>2</sub> release	$k'_{4,5} (\text{s}^{-1})$ Acetaldehyde release
Wild type	45	294	105	105
D291A	0.0075	0.0527	0.0450	0.0109
N293A	0.0307	0.842	0.969	0.0330
T294A	0.535	4.86	5.59	0.674
S298A	0.893	10.2	7.49	1.13

Fortuitously, HEThDP, LThDP, and 2-acetylThDP are all (1) stable under these conditions, and (2) have distinct <sup>1</sup>H chemical shifts for their C6'–H resonances. Integration of these resonances provides the relative steady-state distribution of intermediates, and, with the turnover number of the enzyme, the forward rate constants could be calculated for the pathway (rate constants for formation of LThDP ( $k'_2$ ), its decarboxylation ( $k'_3$ ), and acetaldehyde release ( $k_{4,5}$ ).<sup>42</sup> The method has been applied to a number of ThDP enzymes.<sup>23</sup> The power of the experiment is illustrated with the YPDC loop variants (Table 3).

It was concluded that acetaldehyde release is rate limiting for all of the loop variants, hence the loop opening/closing must be impaired.<sup>91</sup> The method has also been applied to active center variants of YPDC,<sup>42</sup> along with several substitutions on the E1ec.<sup>92</sup> Like all experiments, the method has some limitations: it cannot differentiate the level of enamine and HEThDP since under acid quench the former is converted to the latter. Therefore, the HEThDP measured in the quench corresponds to the sum of the relative concentration of enamine and HEThDP.

S. Chakraborty<sup>93</sup> and A. Balakrishnan<sup>94</sup> in the author's laboratory extended the TH method by synthesizing ThDP specifically labeled at both the C2 and C6' positions with <sup>13</sup>C (producing [C2,C6'–<sup>13</sup>C<sub>2</sub>]ThDP. Using 1D-HSQC methods, only protons attached to these two <sup>13</sup>C-labeled atoms were detected. This enabled detection of the aromatic region of the <sup>1</sup>H-NMR spectrum where the C2H and C6'H could be detected even in the presence of other cofactors with aromatic resonances. As an example, the PDHc, even after the protein is acid precipitated, leaves aromatic resonances pertinent to FAD, NAD<sup>+</sup>, and coenzyme A in the supernatant, in addition to the C2H and C6'H of ThDP. An application of this double-labeled compound to a comparison of the rates of formation of HEThDP in the E1ec component and in the entire PDHc-ec is shown in Figure 12.<sup>94</sup> The results demonstrate that the assembly of the complex accelerates the rate of HEThDP formation in this particularly slow E1ec variant by approximately ninefold. Such information is not readily available from any other method. With relatively easy tweaking, the rate of formation of acetyl-CoA could also be monitored, this time in the aliphatic region and using pyruvate with or without <sup>13</sup>C-label at C3.

Finally, it is worth emphasizing once more that the chromophoric-conjugated substrates exploited by the author's group provide complementary information to the TH method, since with the chromophoric substrates the enamine is directly observed and several microscopic rate constants could be directly assessed.

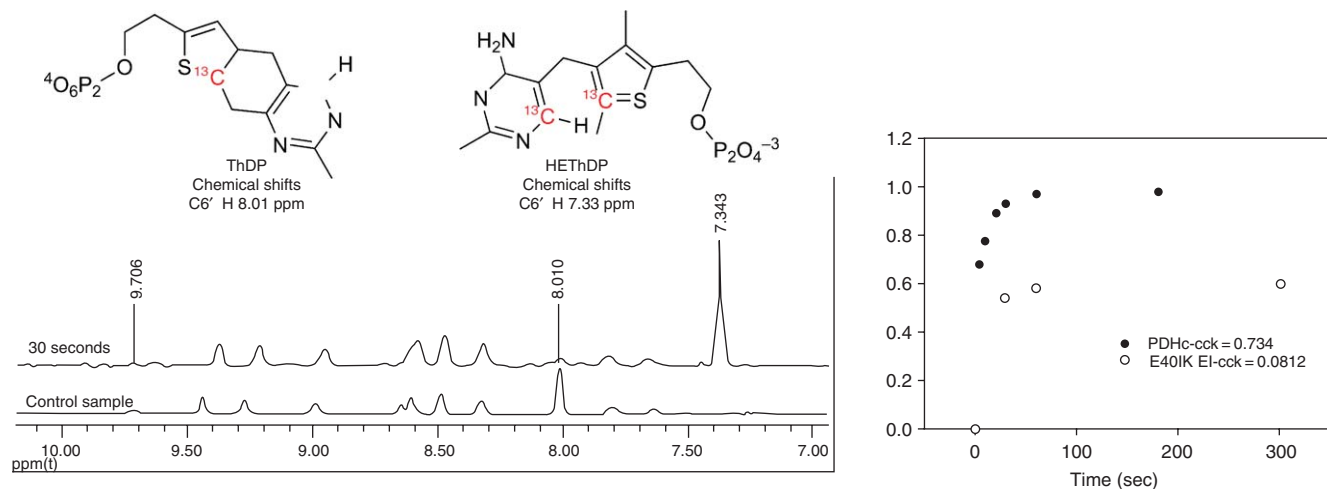
## 7.16.4 Function and Dynamics of Mobile Loops on YPDC and E1ec

An issue of great current interest in mechanistic enzymology is the possibility that loop movement/dynamics may control catalysis (i.e., such movement may be rate limiting) rather than chemical steps. In view of the large superfamily of ThDP enzymes, it was of interest to find evidence pro or con this hypothesis. The author's group explored two enzymes in this regard.

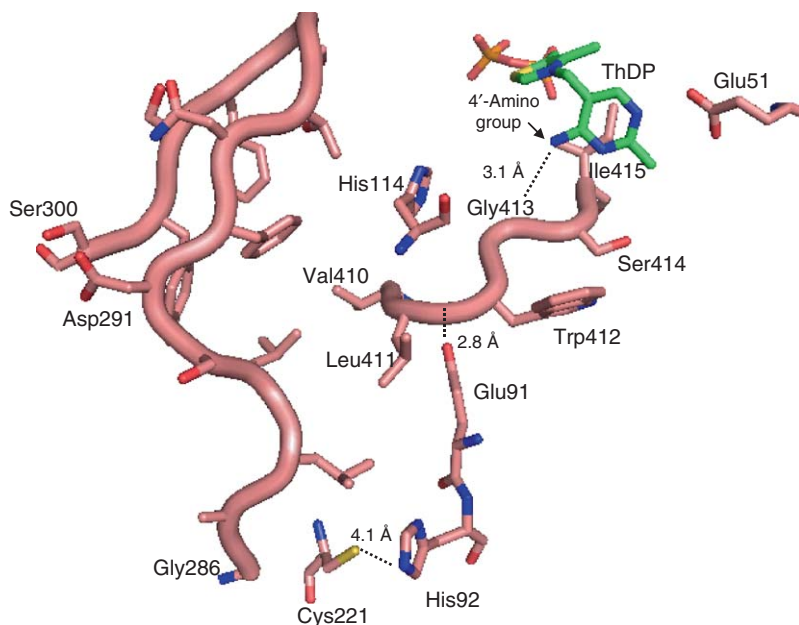
### 7.16.4.1 Evidence on YPDC

A loop encompassing residues 290–304 was among the few residues not seen in the YPDC. ThDP complex (563 amino acids per monomer in a homotetramer).<sup>18,19</sup> The residues were seen, however, in the presence of the substrate activator surrogate pyruvamide.<sup>95</sup> Earlier, some yeast mutants were identified, which resulted in the E291A substitution and the S296–F297 deletion.<sup>96</sup> When YPDC variants with changes reflecting these





**Figure 12** Left: Distribution of ThDP intermediates in the reaction of E401K inner loop variant of *E. coli* PDHc-E1 reconstituted with 1-lipoyl PDHc-E2 and PDHc-E3 and pyruvate analyzed by ID HSQC NMR spectroscopy using the characteristic C6'-H chemical shifts. Recorded at pH 0.75 and 25 °C. Control spectrum, before adding pyruvate. 30 s spectrum after incubation with pyruvate for 30 s at 25 °C. Right: Time course of reactions of E401K E1ec or the same variant reconstituted with the E2ec and E3ec components.



**Figure 13** Location of mobile loop in YPDC with ThDP in green.

substitution/deletion were produced,<sup>96</sup> dramatically reduced (hardly measurable) activity was observed, notwithstanding the distal nature of the loop from the active center (**Figure 13**).

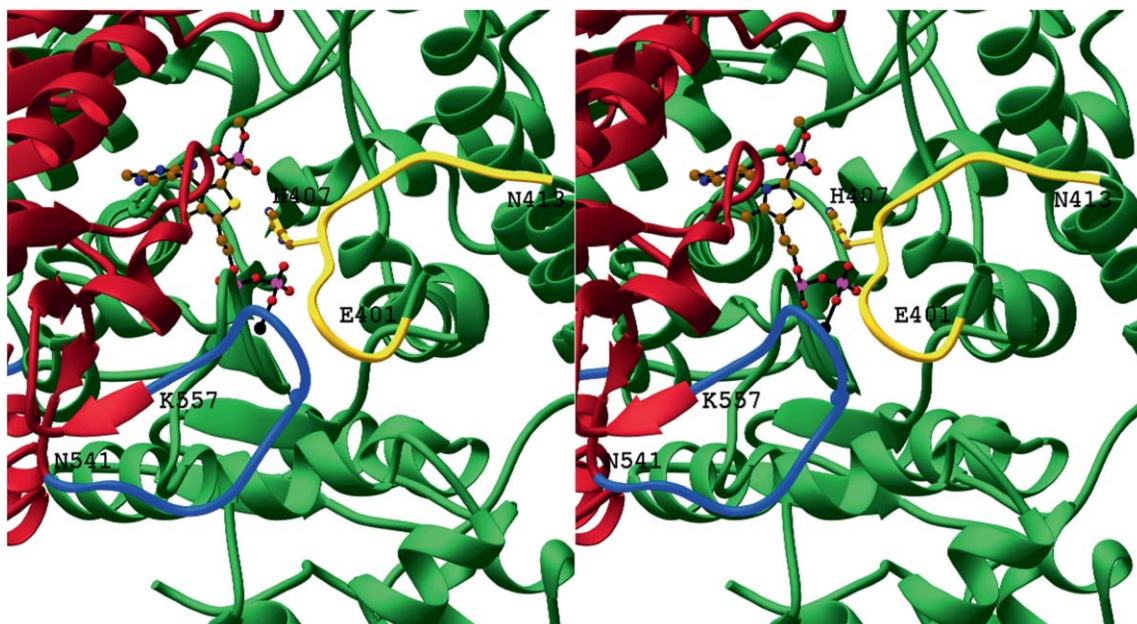
Five variants were constructed along the loop and a battery of experiments were used to show the importance of the loop in the catalytic cycle: While the rate-limiting step was still product release, the relative barriers of the predecarboxylation and the first postdecarboxylation step changed with diminishing activity of the variant, and the subtle effects of substrate activation were no longer detectible with the lowest activity variants (**Table 3**).<sup>56,93</sup>

When the YPDC loop variants were reacted with 3-PKB, an additional factor was noted: ‘misprotonation’ of the enamine in **Scheme 4**, leading to the 3-(3-pyridyl)propanoic acid in place of PAA. This strongly supports the notion of C2 $\alpha$  rather than C2 $\alpha$  protonation with the very weak loop variants, a protonation that is no longer carried out by the usual acid–base triad, but rather by water, which now has freer access to the active center. The fact that the usual acid–base machinery is no longer functioning optimally may indeed account for the slower release of product with these loop variants. Interestingly, while there is a loop including residues 111–113, which according to the pyruvamide-activated YPDC structure could be in van der Waals contact with the 290–304 loop, mutagenesis studies on these residues showed minimal impact on activity, even though these residues are adjacent to His114 and His115, implicated in catalysis by both X-ray and site-directed mutagenesis studies.

## 7.16.4.2 Evidence on E1ec

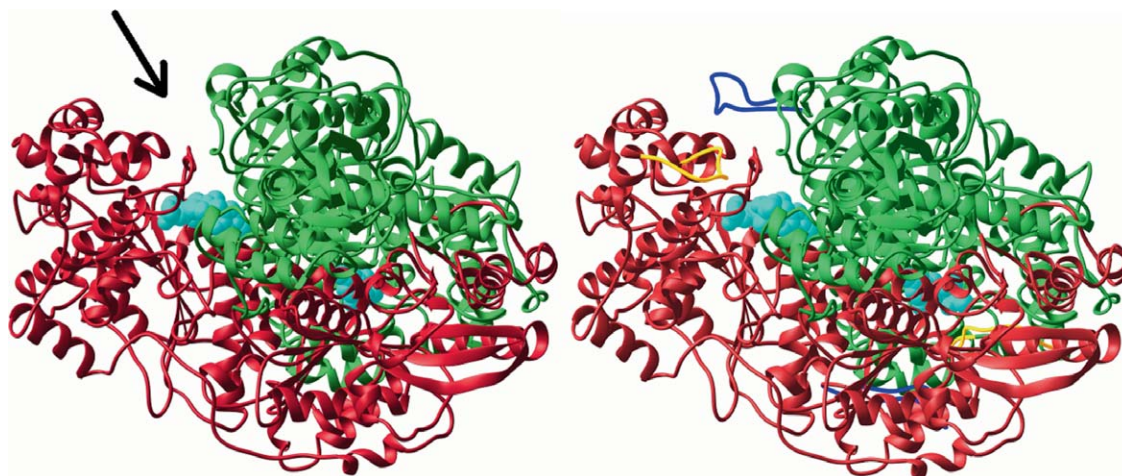
### 7.16.4.2.1 Structural evidence pointing to mobile loops

When the X-ray structure of E1ec was solved in complex with ThDP, there were three regions of poorly defined electron density among the 886 amino acid residues per monomer, corresponding to residues 1–55, 401–413, and 541–557.<sup>97</sup> The N-terminal 1–55 residue region will be discussed (Section 7.16.5.1.1). A remarkable change was seen in the structure of E1ec in complex with a predecarboxylation intermediate phosphonolactyl ThDP (PLThDP – the adduct of MAP with ThDP – see **Scheme 3**), revealing the structure and important interactions of the two loops 401–413 (inner loop) and 541–557 (outer loop) not seen before (**Figure 14**).<sup>98</sup> This signaled a disorder to order transition on forming the predecarboxylation PLThDP intermediate (**Schemes 2 and 3**). On E1ec, the bound PLThDP is stabilized by interactions with inner loop residues E401, H407 (via direct hydrogen bond formation), and Q408 (via a water molecule) (**Figure 14**).<sup>98</sup>



**Figure 14** Stereo view of the newly ordered loops showing their relationship to the active site and enzyme surface. The 'inner' loop (401–413) is colored in yellow while the 'outer' loop (541–557) is colored in blue. Individual subunits in the E1 dimer are shown in red and green, respectively. The PLThDP and side chain for His407 are shown with a ball and stick representation.

The two newly ordered regions interact with each other, and propagate from the active site to the enzyme surface forming a tunnel to accept lipoamide from E2ec for the next reaction (**Figure 15**). This led the authors to propose that the ordering observed in the E1ec–PLThDP complex would also be present when the true intermediate LThDP is formed. Also, the crystal structure of the H407A E1ec in complex with PLThDP provided evidence that in the absence of H407, the two loops are not ordered, hence the H-bond from H407 to LThDP is the key trigger for the disorder to order transformation. The E1ec–PLThDP structure (mimicking



**Figure 15** Left: E1ec–ThDP with inner and outer loops fully disordered (not seen). Arrow indicates active site entrance. Subunits in red and green, ThDP (cyan) space filling. Right: E1ec–PLThDP complex, newly ordered inner (yellow) and outer loop (blue).

LThDP) was among the first examples of a covalently bound predecarboxylation reaction intermediate analog in any ThDP-dependent enzyme.<sup>98</sup>

#### 7.16.4.2.2 Studies on the inner active center mobile loop of E1ec<sup>46,99,100</sup>

The above results prompted the subsequent study on the role and dynamics of the inner loop, starting with mapping of the conserved residues on the loop (Table 4). Kinetic, spectroscopic, and crystallographic studies on some inner loop variants led to the conclusion that charged residues flanking H407 are important for stabilization/ordering of the inner loop thereby facilitating completion of the active site. The results further suggested that a disorder to order transition of the dynamic inner loop is essential for (1) forming LThDP, (2) sequestering active site chemistry from undesirable side carboligase side reactions, as well as (3) communication between the E1ec and E2ec components. The experiments carried out to reach these conclusions are presented briefly as they should be useful for studies of related enzymes.

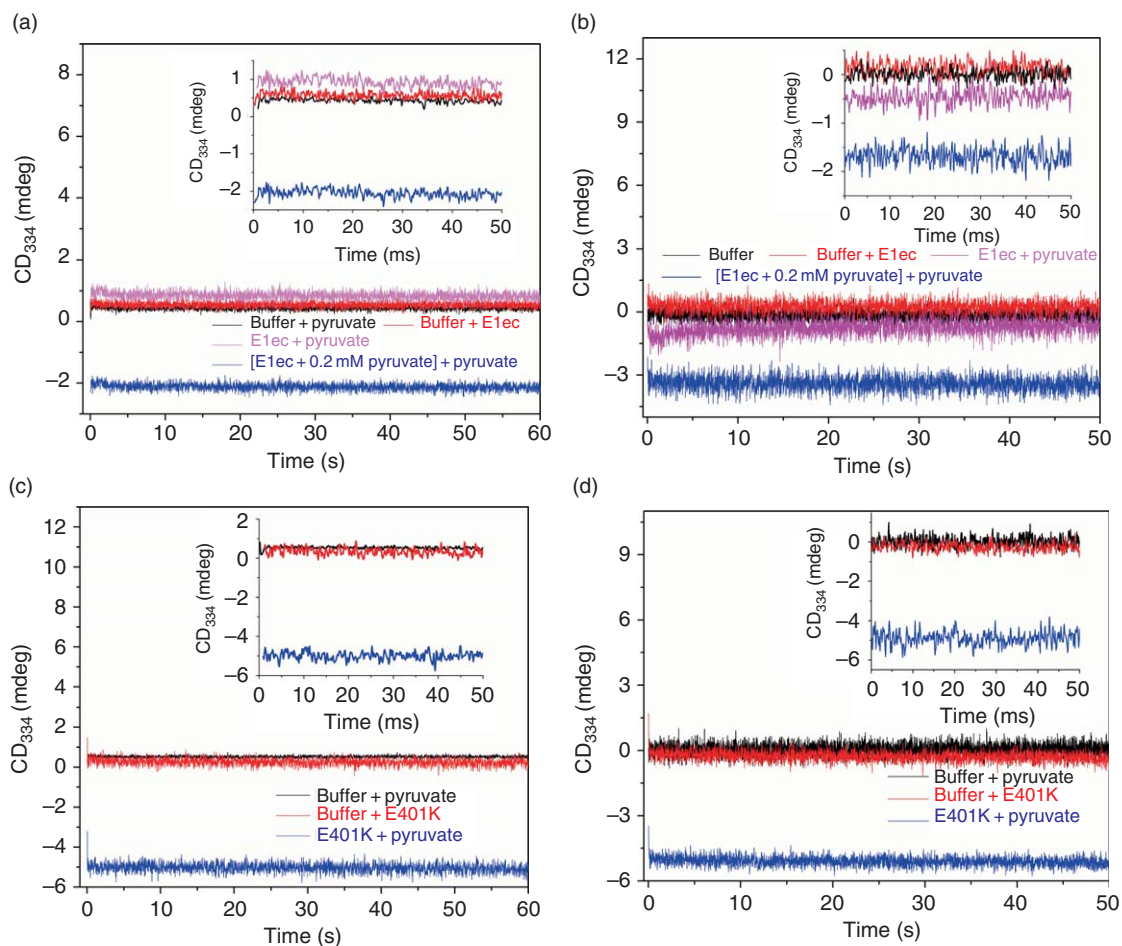
To study the reaction through LThDP formation, the alternate substrate MAP was used. To identify which microscopic step(s) in Scheme 2 respond to loop dynamics, the following experiments were carried out. Interpretation of the results was aided by assignment of a negative CD band centered near 320–330 nm to (1) the 4'-aminopyrimidine tautomer of ThDP (AP) or (2) the MC formed either with substrate or substrate analog MAP (see Figures 4 and 5). On E1ec it was found difficult to observe the AP form. But, addition of pyruvate to the E401K loop variant under a variety of conditions produced the negative CD band corresponding to MC, stabilized as a result of very slow catalysis caused by impaired loop dynamics. It could be shown that the MC is fully formed in E1ec and E401K, as measured with stopped-flow CD, within the dead-time of the instrument (1 ms) (Figure 16). However, formation of PLThDP from MAP (Scheme 3) on E1ec ( $k_1^{\text{obs}} = 3.6 \pm 0.2 \text{ s}^{-1}$  and  $k_2^{\text{obs}} = 0.35 \pm 0.06 \text{ s}^{-1}$ ) was slower than MC formation and significantly slower in E401K ( $k_1^{\text{obs}} = 0.37 \pm 0.05 \text{ s}^{-1}$  and  $k_2^{\text{obs}} = 0.04 \pm 0.01 \text{ s}^{-1}$ ). Clearly, formation of PLThDP, hence formation of C–C covalent bond ( $k_2$  in Scheme 3), and not formation of MC ( $k_{\text{MC}}$ ), is the rate-limiting predecarboxylation step. Since C–C bond formation, but not MC formation, is dramatically slowed down (10-fold compared to E1ec) in E401K, loop dynamics apparently greatly influence covalent addition of substrate to the enzyme-bound ThDP.

To study the sequestration of the active site, the extent and stereochemistry of the carboligase side products was determined. The carboligase reactions on ThDP-dependent decarboxylases arise from the impairment of the reaction subsequent to the decarboxylation step, such that the enamine reacts with either the 2-oxoacid product forming acetolactate or its analogs, or with the acetaldehyde product forming acetoin or its analogs (Scheme 5). While for some ThDP enzymes, such as acetoxyacid synthase, the carboligase reactions comprise the principal products; these products are observed with virtually all ThDP-dependent 2-oxoacid decarboxylases. Importantly, both of these carboligase side reactions form chiral compounds and the enantiomeric excess formed provide stereochemical information about the facial selectivity of the incoming electrophile (2-oxoacid or aldehyde), that is, whether the *re* or *si* face of the electrophile is being presented

**Table 4** Site-directed mutagenesis on loop 401–413 in E1ec

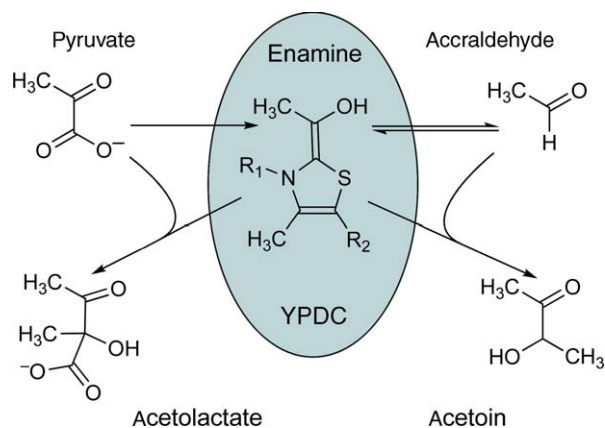
E1ec variant	Overall PDHc-ec activity (units mg <sup>-1</sup> E1) (%)	$k_{\text{cat}}/K_m$ ThDP s <sup>-1</sup> μmol l <sup>-1</sup>	$k_{\text{cat}}/K_m$ pyruvate s <sup>-1</sup> μmol l <sup>-1</sup>	$K_d$ MAP μmol l <sup>-1</sup>
None	11.4 (100)	24.0	0.145	1.0
E401A	1.09 (9.6)	1.16	0.0058	9.0
E401K	0.12 (1.0)	NA	NA	129
K403A	1.32 (11)	0.99	0.0063	12
K403E	0.06 (0.6)	NA	NA	136
H407A	0.025 (0.21)	NA	NA	221
Q408A	5.71 (32)	1.85	0.15	NA
K410A	2.30 (23)	0.92	0.045	NA
K410E	0.46 (3.7)	NA	NA	7.0
K411A	10.7 (68)	7.96	0.150	NA
K411E	7.56 (38)	1.79	0.125	0.92





**Figure 16** Evidence for the formation of the Michaelis complex (MC) between pyruvate and E1ec and variants within the 1 ms dead-time of the stopped-flow CD instrument. (a) E1ec (50 mmol l<sup>-1</sup> active sites in 50 mmol l<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub> (pH = 7.0) containing 0.2 mmol l<sup>-1</sup> ThDP and 1.0 mmol l<sup>-1</sup> MgCl<sub>2</sub>) was mixed with 2.0 mmol l<sup>-1</sup> pyruvate in the same buffer. MC (blue traces) in E1ec could be only seen on preincubation with 0.2 mmol l<sup>-1</sup> pyruvate (E1ec + 0.2 mmol l<sup>-1</sup> pyruvate).<sup>33</sup> Inset refers to the same experiment monitored at shorter time. (b) Experiment in (a) carried out in 50% buffered glycerol ( $\eta = 5.3$ ). Same experiment was performed on E401K at  $h = 1.0$  (buffer) (c) and  $\eta = 5.3$  (d). Insets in (c) and (d) refer to experiments carried out on E401K monitored for shorter timescales. All conditions were same as E1ec except E401K did not require preincubation with pyruvate for MC signal.

to the enamine intermediate. Analogous reaction products are also formed by aromatic 2-oxoacid decarboxylases utilizing ThDP. While the carboligase side reactions are common and virtually ubiquitous at low percentage of total product (1–2% typically), single amino acid substitutions on either YPDC or E1ec were shown to lead to the function being diverted to acetoin synthase or acetolactate synthase activity. As an example, the E477Q substitution converts YPDC to an acetoin synthase, while the D28A (or D28N) substitution to an acetolactate synthase.<sup>50</sup> Similarly, the E636A substitution on E1ec converts that enzyme to an acetolactate synthase<sup>92</sup> with some acetoin synthase activity. Interestingly, the two enzymes lead to preponderance of the opposite enantiomer of both carboligase products, using either pyruvate or  $\beta$ -hydroxypyruvate.<sup>101</sup> When the carboligase products of the inner active center loop variants were evaluated, the results indicated that not only was there a higher yield of carboligase products with the loop variants, but the predominant enantiomer changed as reflected by a change in the phase of the CD signal.<sup>99</sup> Hence, the mobile inner loop not only controls access to the active site, changing from the normal product distribution to more carboligase side product, but the loop even impacts on the facial selectivity toward the electrophile.

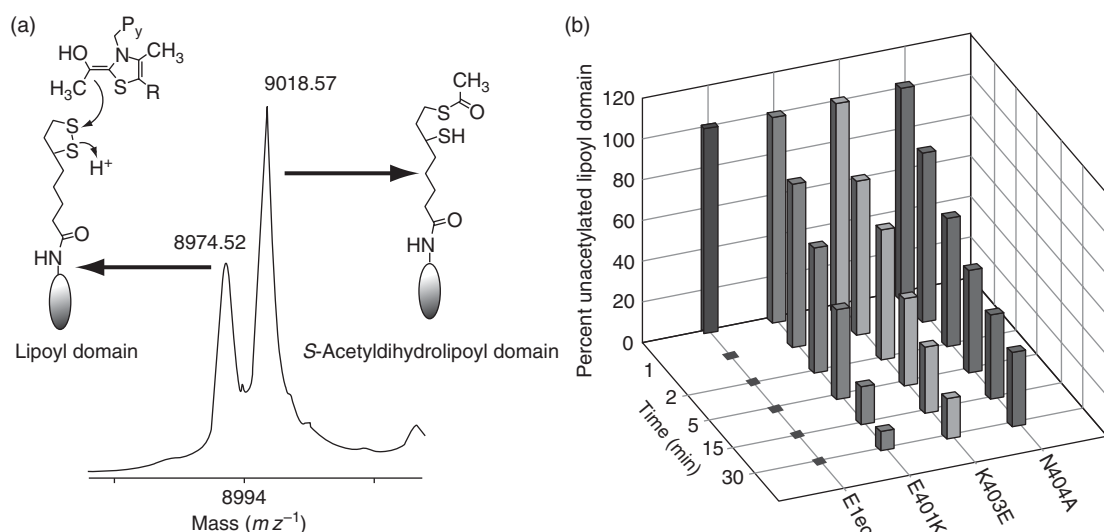


**Scheme 5** Carboligase side reactions on YPDC and other 2-oxoacid decarboxylases.

To study the communication between the E1 and E2 components, matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been very useful, as it can also be utilized to determine the time course of reductive acetylation of the E2 component by the E1 component and pyruvate.<sup>75,102</sup> While the wild-type E1ec can reductively acetylate the E2ec component in less than 30 s (the current limit for manual quench), with the inner loop variants, both unacetylated and acetylated forms are clearly visible even after 30 min, that is, the reductive acetylation was incomplete even after this time period (**Figure 17**).<sup>99</sup> This experiment is made possible by two observations: (1) the reductively acetylated E2 component is stable under acidic conditions and (2) the lipoyl domains of the E2 component are resistant to trypsinolysis.

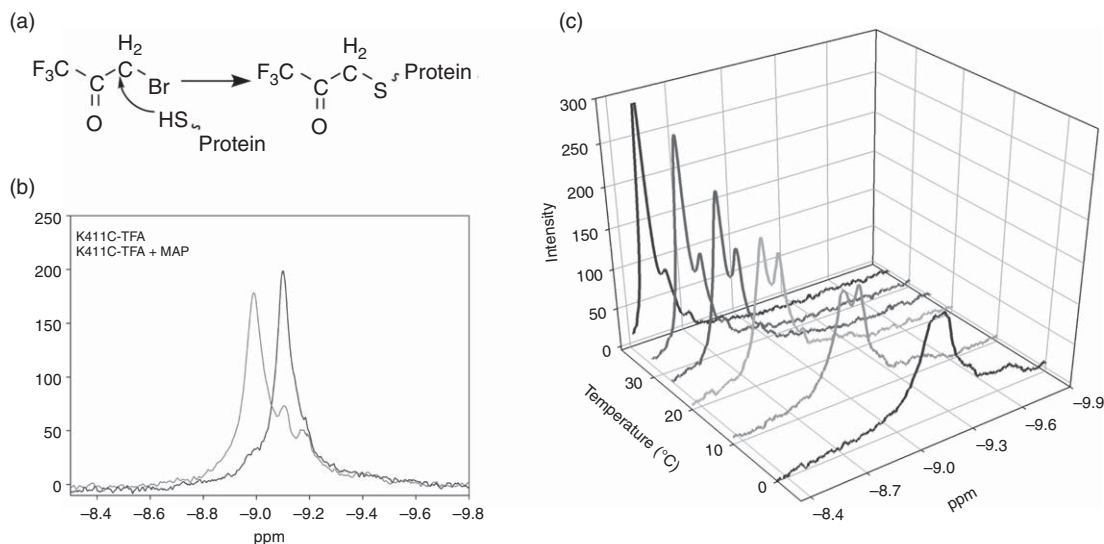
Next, the X-ray crystal structure of a second loop variant E401K, with ThDP and with PLThDP present was solved. Again, the inner and outer loops were unorganized even with PLThDP present (as with the H407A variant).<sup>99</sup> This provided the second structural data point to support the notion that unimpaired inner loop is essential for organization of the two loops.

To investigate the rate of loop movement, a cysteine-less E1ec was constructed where five cysteines were substituted to alanine and the sixth, Cys259 to Asn.<sup>46</sup> This cysteine-less construct still possessed 22% activity for NADH production in the reconstituted PDHc-ec assay. (In this assay all three of the components of the



**Figure 17** Reductive acetylation of lipoyl domain with E1ec inner loop variants. (a) MALDI-TOF/TOF MS method to measure reductive acetylation of lipoyl domain of E2 by E1 and pyruvate. (b) Time course of reductive acetylation of lipoyl domain.





**Figure 18**  $^{19}\text{F}$  NMR on (trifluoroacetylated) K411C-TFA. (a) Covalent modification at K411C with 3-bromo-1,1,1-trifluoroacetone. (b) Effect of MAP on the spectra of K411C-TFA. Grey line is a spectrum of K411C-TFA and shows three resonances at  $-8.993$ ,  $-9.106$ , and  $-9.193$  ppm. The resonance at  $-9.193$  ppm is due to oxidized free label (see (c)). The black trace is a K411C-TFA ( $250\ \mu\text{mol l}^{-1}$ ) with saturating ( $500\ \mu\text{mol l}^{-1}$ ) MAP. The resonances at  $-8.993$  and  $-9.193$  ppm can be seen as small shoulders to a now intense resonance at  $-9.106$  ppm indicating a conformational change in the loop. (c) Effect of temperature on the  $^{19}\text{F}$  NMR spectra of unliganded K411C-TFA ( $250\ \mu\text{mol l}^{-1}$ ).

*E. coli* enzyme are individually expressed, purified, and added in optimal molar ratios leading to reconstituted complex with full activity on the wild-type components.)

Monitoring the kinetics of E1ec and its loop variants at various solution viscosities, it was shown that the rate of a chemical step is modulated by loop dynamics. The cysteine-less E1ec construct was site-specifically labeled on the inner loop with an electron paramagnetic resonance active nitroxide label revealing ligand-induced conformational dynamics of the loop and slow ‘open–close’ conformational equilibrium in the unliganded state. A  $\text{CF}_3\text{C}(=\text{O})\text{CH}_2-$  (trifluoroacetyl)  $^{19}\text{F}$ -NMR label placed at the same residue revealed that the interconversion of the open and close forms of the inner loop takes place with a rate constant of  $0.5\text{--}1.0\ \text{s}^{-1}$  very similar to the turnover number from the E1-specific assay ( $0.38\ \text{s}^{-1}$ ) (Figure 18 – line shape analysis as in the spectral simulations used WINDNMR-Pro<sup>103</sup>). These results suggest a quantitative correlation of E1ec catalysis and loop dynamics for the 200 000-Da protein. Thermochemical studies (isothermal calorimetry) revealed that these motions may promote covalent addition of substrate to the enzyme-bound ThDP by reducing the free energy of activation. The results are consistent with efficient coupling of catalysis and regulation with enzyme dynamics in E1ec, and suggest the mechanism by which it is achieved, and reinforce the hypothesis ascribing catalytic and regulatory roles to enzyme dynamics.<sup>46</sup>

## 7.16.5 Structure–Function Studies in ThDP-Initiated Multienzyme Complex Reactions at the E1 (ThDP-Dependent) Component

### 7.16.5.1 Structural Evidence

To gain better understanding of the assembly of the PDHc-ec, a prominent member of octahedral complexes, studies were initiated to map the interactions in the E1ec–E2ec binary complex. This is very important for a variety of reasons, not the least of which is to build a better model of the complex from rather low-resolution EM reconstruction studies. Also, the communication between the E1 and E2 components has an absolute dependence on ThDP reactivity. Earlier it was shown that residue H407 on the ‘inner loop’ of E1ec participates in the reductive acetylation of E2ec.<sup>75,102</sup>

### 7.16.5.1.1 The amino terminal region of E1ec interacts with E2ec

The loci of interaction between E1ec and E2ec were studied by carrying out both deletion analysis and single-site site-directed mutagenesis experiments of the N-terminal region of the E1ec. The motivation for selecting this region is: (1) In the crystal structure of the E1ec component, the region spanning residues 1–55 did not lead to interpretable electron density, presumably due to mobility.<sup>97</sup> (2) Using FT-MS detected mass distribution, it was demonstrated that E2ec protects the N-terminal region of E1ec (residues 1–80) from trypsinolysis.<sup>104</sup>

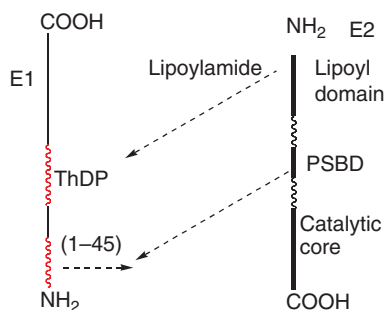
The entire 7–42 N-terminal region of E1ec was scanned (Table 5) leading to the following conclusions regarding this region: (1) decarboxylation of pyruvate in the E1ec-specific reaction is only modestly affected and (2) most significantly, both the overall complex activity (NADH production) and the interaction between E1ec and E2ec are indeed affected with substitutions at Asp7, Asp9, Glu12, Arg14, Asp15, Glu27, and Glu30 (Table 5).<sup>104,105</sup>

Two loci of interaction were suggested between the E1ec and E2ec (Figure 19): (1) the ThDP-bound enamine on E1ec and the lipoylamide of E2ec (as reflected by the ability to reductively acetylate the latter), and (2) the N terminal 1–45 residues of E1ec with some as yet undefined region of E2ec (reflected by overall activity of the PDHc). It could be shown that reductive acetylation is necessary but insufficient for the overall PDHc reaction.<sup>104</sup>

**Table 5** Effect of E1ec N-terminal substitutions on activity and ability to form E1ec–E2ec subcomplex

N-terminal residue of E1ec	Overall PDHc-ec activity (%)	DCPIP E1ec-specific activity (%)	Formation of E1ec–E2ec <sup>a</sup> subcomplex
D7A	0.14	90	–
D9A	0.0	88	–
D9R	3.0	21	±
P10A	107	80	+
I11A	94	62	+
E12D	2.3	10	–
E12Q	3.3	41	
E12R	3.4	100	±
T13A	13	74	+
R14A	0.25	100	–
D15A	0.19	84	–
E21R	69.3	80	+
E26R	87	100	+
E27R	2.7	88	±
E30R	2.9	86	±
D37R	74	94	+
E42R	100	100	+

<sup>a</sup> Gel filtration chromatography. Code for species detected: +, only E1ec–E2ec subcomplex; –, no interaction with E2ec; ±, peaks E1ec and E1ec–E2ec subcomplex detected. Strength of subunit interaction: + > ±, – > –.



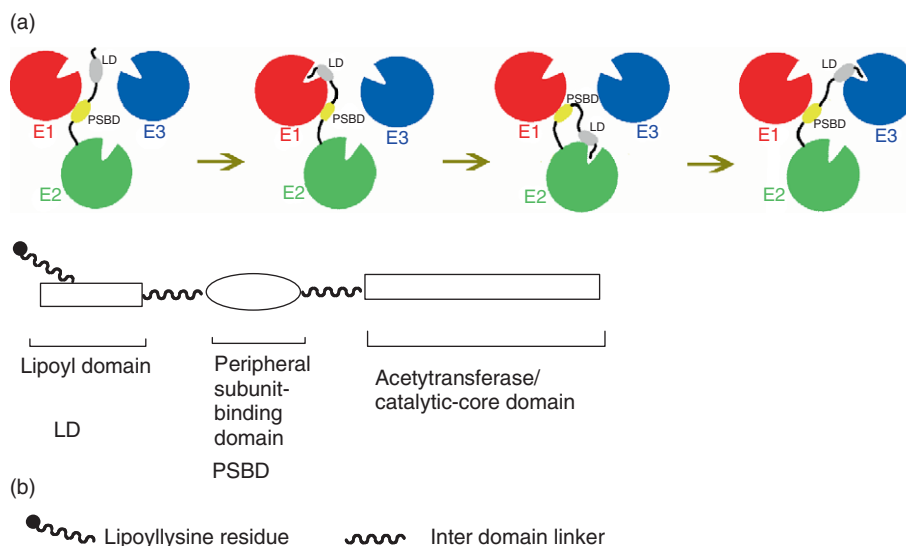
**Figure 19** Two loci of interaction between *E. coli* PDHc-E1 and PDHc-E2.

### 7.16.5.1.2 Region of E2ec interacting with E1ec

It is indeed also relevant which loci of the E2ec component interact with the E1ec component, especially since there is evidence that the same loci of E2ec may interact with the E1ec and the third E3ec component. It was hypothesized that the highly negatively charged N-terminal 1–45 region of E1ec interacts with the positively charged residues of the peripheral subunit-binding domain of E2ec, as was also suggested for the related complex from *Azotobacter vinelandii*.<sup>106–108</sup> The domain structure of the E2ec component is shown in **Scheme 6** and consists of (1) anywhere from 1–3 lipoyl domains (LD), depending on the source of the enzyme, (2) a peripheral subunit-binding domain, and (3) the catalytic or core domain where the acetyl group is transferred to the coenzyme A. The domains are connected by hinges of variable sizes. Therefore, site-directed mutagenesis studies were initiated on the positively charged residues of the peripheral subunit-binding domain and at the N-terminal region of the catalytic core of E2ec.<sup>105</sup> These studies have already identified two amino acids R129 and R150 in the peripheral subunit-binding domain, important not only for E1ec–E2ec, but also for E2ec–E3ec interactions. These results strongly suggest that the binding regions of E2ec for E1ec and E3ec overlap, although they may not be identical. The R129E and R150E substitutions on E2ec were shown to affect overall activity of PDHc by dramatically reducing the rate of reductive acetylation. Residues substituted in the catalytic core K191 and R202 did not affect interactions with the other components.<sup>108</sup> This remarkable result with residues R129 and R150 of E2ec strongly implies that lipoyl domain and the peripheral subunit-binding domain do interact since residues R129 and R150 are located in the peripheral subunit-binding domain while the lipoamide is covalently attached to K41 in the lipoyl domain.

The results so far indicate that the ‘epitopes’ or binary interaction surfaces of E2ec for E1ec and E3ec are certainly overlapping and suggest competition between E1ec and E3ec for E2ec. The precise nature of these interactions still needs further delineation even in the ‘simpler’ bacterial 2-oxoacid dehydrogenase complexes.

In the mammalian enzymes the interactions are even more complicated since while the E1 component interacts with the E2 component, the E2 component does not directly interact with the E3 component, rather the information is relayed from E2 to the E3-binding protein (E3BP), which in turn interacts with the E3 component.



**Scheme 6** (a) Domain structure of the E2 component of the 2-oxoacid dehydrogenase complexes. (b) Schematic layout of domains in the 1-lipoyl E2ec. The lipoyl domain (LD) with the appended lipoamide ‘visits’ the active sites of all three components of PDHc-ec.

## 7.16.6 Probing ThDP Enzyme Mechanisms with Coenzyme Analogs

Over the years, several ThDP analogs have been synthesized and tested as coenzyme surrogates (see [Figure 1](#) for structures). In many cases these analogs are inactive, but can provide structural information since the substrate/product could cocrystallize with the substituted enzyme without being turned over. An important limitation of such methods is that with some enzymes, ThDP is so tightly bound that it can only be removed under harsh conditions and reconstitution even with ThDP to fully active enzyme is nearly impossible.

A very significant early use of coenzyme analogs to study the mechanism of ThDP enzymes reported the *de novo* synthesis of ThDP analogs with substitutions at the pyrimidine ring nitrogen atoms by Schellenberger's group (for a review see Schellenberger<sup>9</sup>). The 1'-deaza, 3'-deaza and 4'-desamino analogs of ThDP were synthesized *de novo* and tested on a number of enzymes for nearly four decades. The information gained from this admirable synthetic achievement clearly showed the absolute requirement for the N1' and N4' atoms, with less requirement for the N3' atom claimed.

When applications of transition state inhibitors to enzymology were introduced by Wolfenden<sup>109</sup> and Lienhard,<sup>110</sup> Lienhard suggested two compounds, thiamin thiazolone diphosphate (ThTTDP) and thiamin 2-thiothiazolone diphosphate (ThTTDP), as potential transition state analogs. The C2=O and C2=S bonds, respectively, in place of C2-H in ThDP, resemble the C2=C2 $\alpha$  double bond of the enamine resonance structure in [Schemes 1 and 2](#). These compounds have been synthesized and used to good effect on a variety of ThDP enzymes. For reasons not yet clear, they appear to be most inhibitory toward the E1 components of the 2-oxoacid dehydrogenase complexes,<sup>111</sup> less so against decarboxylases, both nonoxidative and oxidative and transketolase. An interesting observation was made by N. Nemeria when reacting ThTTDP with the E1 component of the *E. coli* PDHc.<sup>112</sup> On titration of the ThTTDP into a solution of the enzyme, a new positive CD band centered at 325 nm developed and enabled calculation of the  $K_d$  for this inhibitor. This provides a very sensitive probe of the active center of these enzymes since only a 'native' protein allows development of this CD band. There are now several structures of ThDP enzymes with one or another of these compounds bound and they show that the ThDP environment is not significantly perturbed by the substitution. Furey and Sax and coworkers hypothesized on the basis of the structural evidence with the *E. coli* E1 component that the number of hydrogen bonds to the coenzyme/analog increased going from ThDP to the transition state analog,<sup>112</sup> perhaps accounting for the stronger binding to the enzyme of the latter. By extrapolation, the additional hydrogen bonds may explain stronger affinity of the enzyme to the enamine-like transition state than to ThDP itself.

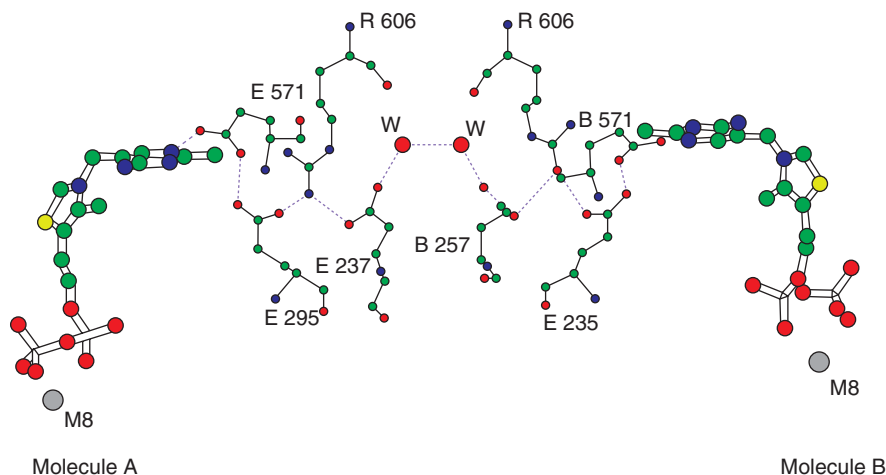
A more recent and potentially very useful analog was synthesized by Leeper's group at Cambridge: 3-deazaThDP, where the thiazolium N is replaced by a carbon atom, that is, replacing the positively charged thiazolium ring by the neutral thiophene ring.<sup>13,113,114</sup> This compound is a superb inhibitor of some of the enzymes tested, being a virtually ideal structural mimic of ThDP and the neutral charge on the thiophene ring also resembles the neutral charge on the enamine intermediate. Substituting ThDP with 3-deazaThDP also allows crystallization with substrate and product, throwing light onto the nature of the Michaelis complex. C2-substituted analogs of 3-deazaThDP have also been synthesized, some resembling the covalent complexes of ThDP in [Schemes 1 and 2](#).

All of these compounds have significant utility in ThDP enzymology. One can assume that similarly imaginative ThDP analogs will be synthesized in the future.

## 7.16.7 Active Center Communication in ThDP Enzymes

### 7.16.7.1 The Proton Wire Mechanism

While there have been occasional reports of active center communication in ThDP enzymes, there was an intriguing formal suggestion for such communication by the Perham group of a 'proton wire' composed of a series of acidic residues connecting ThDPs at two active centers.<sup>115</sup> The basis of the proposal is that the conserved Glu residues located within hydrogen bonding distance of the N1' atom is an integral part of this proton wire. The proposal is applicable to the E1 components of two classes of bacterial complexes, that from *Bacillus stearothermophilus* and E1ec ([Figure 20](#)), and possibly to the human E1 component as well. In an



**Figure 20** Putative information transfer pathway between ThDPs on two monomers of the *E. coli* pyruvate dehydrogenase complex E1 component.

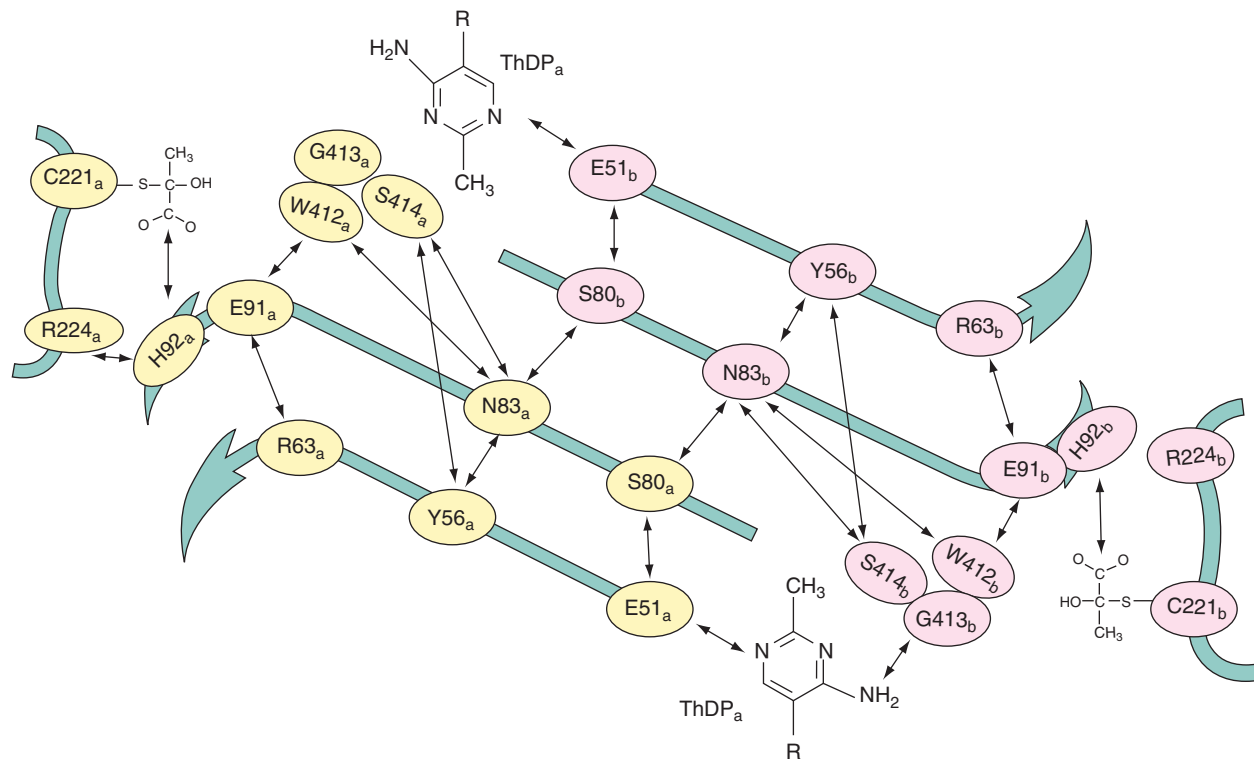
accompanying letter it was pointed out that with many ThDP enzymes, such as YPDC, there is no such proton wire via acidic groups, rather there is a hydrogen bonding network as drawn in [Figure 21](#).<sup>20</sup> Inspired by the proton wire proposal, a potential pathway of communication between the ThDP.Mg<sup>2+</sup>s at the two active centers of E1ec is being investigated<sup>116</sup> ([Figure 20](#)). Extensive biochemical–mechanistic and structural studies were carried out at the highly conserved E571 residue, known to play an important role during intramolecular proton transfer involving the IP tautomer of the ThDP ([Scheme 2](#)). The E571A variant shows greatly reduced catalytic activity ~1%. The crystal structures of E571A with no ThDP, with ThDP, and with PLThDP were solved to further characterize the role of E571 in both catalysis and cofactor binding. The E571A substitution leads to conformational changes effectively preventing proper 4'-aminopyrimidine ring orientation within the active site, leading to a disordered cofactor. In the E571A–PLThDP complex, binding of PLThDP was found unaffected, and there were no significant structural changes observed for this E571A–PLThDP complex relative to the E1ec–PLThDP complex.<sup>117</sup> The analyses indicate that the conserved hydrogen bond between Glu571 and the N1' atom of ThDP is required for proper cofactor binding and associated conformation, as well as for efficient catalytic activity, but is not required for stabilizing reaction intermediates, or for organizing the mobile inner and outer loops. The results also support the idea of nonequivalence of active sites within the E1 dimer, as in some of the dimeric structures different conformations are found for corresponding active site residues.

### 7.16.7.2 Spectroscopic Evidence for Nonequivalence of Active Sites

The CD spectroscopic experiments on E1h and POX signaled the coexistence of the AP and IP forms on the enzymes prior to addition of substrate. While the results could be explained by the existence of dimers of E1h with mixtures of IP<sub>2</sub> and AP<sub>2</sub> active centers, this is less logical than the IP AP combination. It is important to note that at the molecular level and at any instant, there is one and only one tautomeric form present at any active site, in other words we could not be observing an equilibrating mixture of tautomers.<sup>68</sup>

### 7.16.7.3 Kinetic Evidence for Nonequivalence of Active Sites

Using the <sup>1</sup>H-NMR methods, Tittmann and collaborators could demonstrate that the rate of C2H→D exchange displayed half-of-the-sites reactivity in E1h since the rate constant for the first active site equaled or exceeded the turnover number, while for the second site it was hardly faster than the rate constant for the



**Figure 21** Substrate activation pathway and active center communication on YPDC.



same exchange in the absence of enzyme. Similar observations were made by Tittmann on some of the 'proton wire' residues of E1ec (in preparation).

A number of results pointing to active center communication in ThDP enzymes were recently summarized in a review on the subject.<sup>118</sup>

### 7.16.8 Continuing Fascination with Thiamin Enzymes as Paradigm for Enzymatic Solvent Effects

In 1970, there appeared two relevant publications modeling ThDP-catalyzed reactions in solvents of different dielectric solvent. Both Lienhard<sup>119,120</sup> and coworkers and Kemp and O'Brien<sup>121</sup> presented model studies, which suggested that a solvent of low dielectric medium accelerated thiamin-catalyzed reactions. At the same time, Ullrich and Donner presented the first enzymatic evidence suggesting that the enzyme yeast pyruvate decarboxylase has a hydrophobic active center on the basis of studies with a fluorescent label.<sup>122</sup>

The author's group has for the past decade also probed the active center environment of ThDP enzymes using a variety of experimental approaches.

#### 7.16.8.1 Evidence from Observation of the Enamine on YPDC

With the background knowledge that YPDC has a rather forgiving active center, which accepts 2-oxoacids with large substituents in place of the methyl group, and that the wavelength of maximum absorption of the enamine derived from HBThDP is near 380 nm and far from the protein envelope, we tested whether the enzyme would convert the second postdecarboxylation intermediate HBThDP to the enamine. Therefore, racemic HBThDP was synthesized and added to the E91D YPDC variant (this variant has a highly desirable property in its ability to readily exchange its ThDP with its analogs, unlike the wild-type enzyme), and resulted in a slow development of the  $A_{380}$ , suggesting formation of the enamine, that is, reversal of the protonation in **Scheme 1**.<sup>72</sup> Given that the  $pK_a$  for this ionization is 15–16 in the absence of enzyme,<sup>65,66</sup> this result strongly suggested that the YPDC could stabilize the enamine relative to an aqueous environment. In an attempt to determine the nature of the YPDC active center environment, the fluorescence emission of thiochrome diphosphate, a competitive inhibitor (competing with ThDP) of a number of ThDP enzymes, was measured on YPDC and in a series of 1-alkanols ranging from methanol to 1-hexanol.<sup>72</sup> The correlation of solvent dielectric with emission maximum for the 1-alkanols was virtually linear, enabling us to interpolate the value on YPDC as between 1-hexanol and 1-pentanol, that is, 11–13. While such correlations are not quantitative, the message is unmistakable: the active center environment of YPDC resembles that of 1-pentanol or 1-hexanol more than it resembles the aqueous milieu.

#### 7.16.8.2 Evidence from Solvent Effects on Decarboxylation Rate Constants in Model Compounds

The effect of solvent of model decarboxylation reactions carried out earlier by Lienhard<sup>119</sup> and Kluger<sup>5</sup> was reexamined by Zhang *et al.* in both C2 $\alpha$ -lactylthiazolium salts and C2 $\alpha$ -lactylthiamin.<sup>62</sup> This model isolates the decarboxylation step of **Schemes 1 and 2** by monitoring enamine formation via oxidative trapping with DCPIP. The most significant result of the study was that first-order rate constants for decarboxylation could be achieved in low dielectric media that were virtually the same as the turnover number per subunit for a typical decarboxylase such as YPDC. The conclusion is that the enzymatic environment per se could accelerate the decarboxylation reaction so that little additional catalysis is required. The Dunathan hypothesis, according to which the bond between the C2 $\alpha$  atom and the departing carboxyl carbon atom should be perpendicular to the thiazolium plane, so that the incipient p orbital on cleavage of this bond will be in direct conjugation with the aromatic  $\pi$  system, is strictly adhered to. Interestingly, this imperative is apparent even in the C2 $\alpha$ -lactylthiazolium model, as well as in the X-ray structures of all of the predecarboxylation intermediates on several enzymes.

### 7.16.8.3 Evidence from Acceleration of the Reaction Step Leading to the Enamine on E1ec

While HETHDP is not directly on the pathway of the E1 component of the 2-oxoacid dehydrogenase complexes, ionization of its C2 $\alpha$ H could be used to test the reversibility of enamine protonation, and to probe the effect of the active center on this reaction. To test whether the E1ec could accelerate the C2 $\alpha$ H ionization step in **Scheme 2**, HETHDP was synthesized with four deuterons in the C2 $\alpha$ -hydroxyalkyl side chain by condensing ThDP with acetaldehyde- $d_4$ . This deuterated HETHDP was then rapidly mixed with apo-E1ec (devoid of ThDP) on a KINTEK rapid quench instrument in H<sub>2</sub>O and the reaction was stopped by acid quench. The hypothesis used was that the rate of D to H exchange would reflect the rate of the C2 $\alpha$ H ionization step. The exchange was quantified using FT-MS, which displayed linear integration response for as little as 1% conversion to product. Compared to model systems, the rate of this reaction was accelerated on the enzyme by  $\sim 10^7$ . Since there is no amino acid in the active site of E1ec charged with the specific function of ionizing this C2 $\alpha$ H bond, it was concluded that the enzyme environment was largely responsible for the observed rate acceleration, again pointing to as much as 10 kcal mol<sup>-1</sup> rate acceleration due to the enzyme environment.<sup>73</sup>

### 7.16.8.4 Spectroscopic Observations Suggesting an Apolar Environment

The position of the CD<sub>max</sub> for the band associated with the 1',4'-iminopyrimidine tautomer induced on formation of the tetrahedral adduct at C2 on addition of MBP was 299 nm with BFDC and 312 nm with BAL.<sup>48</sup> A model study generating the IP form by addition of base to N1'-methyl-4-aminopyrimidinium salt displayed very dramatic shift in  $\lambda_{\text{max}}$  for the IP form with solvent dielectric constant, the lower dielectric constant shifting the  $\lambda_{\text{max}}$  to longer wavelength.<sup>36</sup> On the basis of these comparisons, a more hydrophobic environment could be attributed to BAL than to BFDC. In fact, inspection of the active centers of ThDP enzymes suggests that BAL and glyoxylate carboligase appear to have the active centers with the least hydrophilic residues.

In a different experiment, when PAA was added to BFDC and the  $\lambda_{\text{max}}$  was monitored for the formation of the adduct at 475 nm, there was an interesting time dependence observed leading to diminution of the signal with time according to an 'approach to equilibrium' kinetic scheme. It is likely that the enamine was being generated on the enzyme at a slow rate.<sup>57</sup> When the external oxidizing agent 2,6-dichlorophenolindophenol was added (this compound has been shown to react with the enamine very fast, unlike with 3-PKB or PAA), only a fraction of the absorption was quenched, presumably the fraction due to the presence of the enamine. When the same experiment was carried out on BAL with PAA, the negative CD band at 475 nm formed again, suggesting formation of the PAA-ThDP adduct. With BAL, however, the entire CD signal at 475 nm could be quenched by 2,6-dichlorophenolindophenol suggesting that there is a protolytic equilibrium established on the enzyme between the bound PAA and the bound enamine. The implication of these studies is that, depending on the enzyme environment, the equilibrium between the C2 $\alpha$ -hydroxyalkyl derivative and the enamine can be established, the stability of the enamine increasing with lower effective dielectric constant of the active center.

The results again suggest a more hydrophobic active center for BAL than BFDC.

These few examples suggest the following: (1) Indeed there is an important 'environmental' contribution to the catalytic rate acceleration on ThDP enzymes. (2) Such factors could account for as much as  $10^7$ – $10^9$  of the total rate acceleration. This constitutes a very significant fraction of the total rate acceleration provided by the protein over and above that afforded by ThDP itself, which on YPDC was estimated to be  $10^{12}$ – $10^{13}$ -fold.<sup>87,88</sup> (3) Such effects are almost certainly present on other ThDP enzymes as well, and researchers should be on the lookout for them.

### 7.16.9 Perspective for Future Studies

In the thiamin model field, even 50 years after the seminal Breslow discovery, to the author's knowledge, there is not yet available a synthetic model juxtaposing the N4' and C2 atoms to within distances found on the enzymes. This poses a significant synthetic challenge, but a successful model resembling the V conformation,

conserved on all ThDP enzymes to date, would carry a significant payoff. On enzymes, it will be important to establish whether the reductive acetylations take place via ionic or radical mechanisms, as recently suggested.

Perhaps the 'holy grail' in the ThDP enzyme field is high-resolution structural studies of the entire 2-oxoacid dehydrogenase complex from any source, although lower-resolution studies from electron microscopy have appeared already. Given the recently published implication of the human PDHc in cancer, and its earlier revealed possible roles in obesity and diabetes, such studies are likely to increase in importance as a guide in drug design.

## Acknowledgments

The author appreciates financial support for his research on thiamin over several decades by the NIH (most recently GM050380) and earlier by the NSF.

### Abbreviations

<b>3-PKB</b>	( <i>E</i> )-4-(pyrid-3-yl)-2-oxo-3-butenic acid
<b>AP</b>	the canonical 4-aminopyrimidine tautomer of ThDP or its C2-substituted derivatives
<b>APH<sup>+</sup></b>	the N1-protonated 4-aminopyridinium form of ThDP or its C2-substituted derivatives
<b>BAL</b>	benzaldehyde lyase
<b>BFDC</b>	benzoylformate decarboxylase
<b>CD</b>	circular dichroism
<b>E1ec</b>	the first component of the <i>E. coli</i> PDHc
<b>E1h</b>	the first component of the human PDHc
<b>HBThDP</b>	C2 $\alpha$ -hydroxybenzylThDP, the adduct of benzaldehyde and ThDP
<b>HETHDP</b>	C2 $\alpha$ -hydroxyethylThDP, the adduct of acetaldehyde and ThDP
<b>IP</b>	the 1',4'-iminopyrimidine tautomer of ThDP or its C2-substituted derivatives
<b>LThDP</b>	C2 $\alpha$ -lactylThDP, the adduct of pyruvic acid and ThDP
<b>MAP</b>	acetylphosphonic acid monomethyl ester
<b>MBP</b>	benzoylphosphonic acid monomethyl ester
<b>MThDP</b>	C2 $\alpha$ -mandelylThDP, the adduct of benzoylformic acid and ThDP
<b>PDA</b>	photodiode array
<b>PDHc</b>	pyruvate dehydrogenase complex
<b>PLThDP</b>	C2 $\alpha$ -phosphonolactylThDP, the adduct of MAP and ThDP
<b>PMTThDP</b>	C2 $\alpha$ -phosphonomandelylThDP, the adduct of MBP and ThDP
<b>POX</b>	pyruvate oxidase from <i>Lactobacillus plantarum</i>
<b>SF</b>	stopped-flow
<b>ThDP</b>	thiamin diphosphate
<b>ThMP</b>	thiamin monophosphate
<b>ThTP</b>	thiamin triphosphate
<b>TK</b>	transketolase
<b>YI</b>	ylide
<b>YI</b>	the C2-carbanion/ylide/carbene form conjugate base of ThDP
<b>YPDC</b>	pyruvate decarboxylase from the yeast <i>Saccharomyces cerevisiae</i>

## References

1. T. B. Begley; A. Chatterjee; J. W. Hanes; A. Hazra; S. E. Ealick, *Curr. Opin. Chem. Biol.* **2008**, *12*, 118–125.
2. L. Bettendorff, *Nat. Chem. Biol.* **2007**, *3*, 454–455.

3. L. O. Krampitz, *Thiamin Diphosphate and Its Catalytic Functions*; Marcel Dekker: New York, 1970.
4. H. Z. Sable, C. J. Gubler, Eds., *Thiamin, 20 Years of Progress*; *Ann. N.Y. Acad. Sci.* **1982**, 378, 7–122.
5. R. Kluger, *Chem. Rev.* **1987**, 87, 863–876.
6. A. Schellenberger, R. L. Schowen, Eds., *Thiamin Pyrophosphate Biochemistry*; CRC Press: Boca Raton, FL, 1988; Vol. 1.
7. H. Bisswanger, J. Ullrich, Eds., *Biochemistry and Physiology of Thiamin Diphosphate Enzymes*; VCH Publishers: Weinheim, Germany, 1991.
8. H. Bisswanger, A. Schellenberger, Eds., *Biochemistry and Physiology of Thiamin Diphosphate Enzymes*; A.u.C. Intemann, Wissenschaftlicher Verlag: Prien, Germany, 1996.
9. A. Schellenberger, *Biochim. Biophys. Acta* **1998**, 1385, 177–186.
10. R. A. Frank; F. J. Leeper; B. F. Luisi, *Cell. Mol. Life Sci.* **2007**, 64, 892–905.
11. M. Pohl; G. A. Sprenger; M. Müller, *Curr. Opin. Biotechnol.* **2004**, 15, 335–342.
12. F. Jordan, M. S. Patel, Eds., *Thiamine. Catalytic Mechanisms in Normal and Disease States*; Marcel Dekker, Inc: New York, 2004.
13. F. Jordan, *FEBS Lett.* **1999**, 457, 298–301.
14. F. Jordan, *Nat. Prod. Rep.* **2003**, 20, 184–201.
15. F. Jordan; N. S. Nemeria, *Bioorg. Chem.* **2005**, 33, 190–215.
16. Y. Lindqvist; G. Schneider; V. Ermler; M. Sundstrom, *EMBO J.* **1992**, 11, 2373–2379.
17. Y. Muller; G. Schulz, *Science* **1993**, 259, 965–967.
18. F. Dyda; W. Furey; S. Swaminathan; M. Sax; B. Farrenkopf; F. Jordan, *Biochemistry* **1993**, 32, 6165–6170.
19. P. Arjunan; T. Umland; F. Dyda; S. Swaminathan; W. Furey; M. Sax; B. Farrenkopf; Y. Gao; D. Zhang; F. Jordan, *J. Mol. Biol.* **1996**, 256, 590–600.
20. C. F. Hawkins; A. Borges; R. N. Perham, *FEBS Lett.* **1989**, 255, 77–82.
21. R. Breslow, *J. Am. Chem. Soc.* **1957**, 79, 1762–1763.
22. Y. A. Muller; Y. Lindqvist; W. Furey; G. E. Schulz; F. Jordan; G. Schneider, *Structure* **1993**, 1, 95–103.
23. R. Kluger; K. Tittmann, *Chem. Rev.* **2008**, 108, 1797–1833.
24. M. Sax; P. Pulsinelli; J. Pletcher, *J. Am. Chem. Soc.* **1974**, 96, 155–165.
25. J. Pletcher; M. Sax; G. Blank, *J. Am. Chem. Soc.* **1977**, 99, 1396–1403.
26. F. Jordan, *J. Am. Chem. Soc.* **1976**, 98, 808–813.
27. F. Jordan; Y. H. Mariam, *J. Am. Chem. Soc.* **1978**, 100, 2534–2541.
28. F. Jordan; G. Chen; S. Nishikawa; B. Sundoro-Wu, *Ann. N. Y. Acad. Sci.* **1982**, 378, 14–21.
29. F. Jordan, *J. Org. Chem.* **1982**, 47, 2748–2753.
30. G. A. Kochetov; R. A. Usmanov, *Biochem. Biophys. Res. Commun.* **1970**, 41, 1134–1140.
31. F. Jordan; Z. Zhang; E. Sergienko, *Bioorg. Chem.* **2002**, 30, 188–198.
32. F. Jordan; N. Nemeria; S. Zhang; Y. Yan; P. Arjunan; W. Furey, *J. Am. Chem. Soc.* **2003**, 125, 12732–12738.
33. N. Nemeria; A. Baykal; E. Joseph; S. Zhang; Y. Yan; W. Furey; F. Jordan, *Biochemistry* **2004**, 43, 6565–6575.
34. N. Nemeria; S. Chakraborty; A. Baykal; L. G. Korotchkina; M. S. Patel; F. Jordan, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 78–82.
35. N. Nemeria; L. G. Korotchkina; M. J. McLeish; G. L. Kenyon; M. S. Patel; F. Jordan, *Biochemistry* **2007**, 46, 10739–10744.
36. A. Baykal; L. Kakalis; F. Jordan, *Biochemistry* **2006**, 45, 7522–7528.
37. A. H. Cain; G. R. Sullivan; J. D. Roberts, *J. Am. Chem. Soc.* **1977**, 99, 6423–6425.
38. F. Seifert; R. Golbik; J. Brauer; H. Lilie; K. Schröder-Tittmann; E. Hinze; L. G. Korotchkina; M. S. Patel; K. Tittmann, *Biochemistry* **2006**, 45, 12775–12785.
39. M. W. Washabaugh; W. P. Jencks, *Biochemistry* **1988**, 27, 5044–5053.
40. D. Kern; G. Kern; H. Neef; K. Tittmann; M. Killenberg-Jabs; C. Wikner; G. Schneider; G. Hübner, *Science* **1997**, 275, 67–70.
41. A. J. Arduengo; J. R. Goerlich; W. J. Marshall, *Liebigs Ann. Recueil* **1997**, 365–374.
42. K. Tittmann; R. Golbik; K. Uhlemann; L. Khailova; G. Schneider; M. Patel; F. Jordan; D. M. Chipman; R. G. Duggleby; G. Hübner, *Biochemistry* **2003**, 42, 7885–7891.
43. R. Kluger; D. C. Pike, *J. Am. Chem. Soc.* **1977**, 99, 4504–4506.
44. R. Kluger; D. C. Pike, *J. Am. Chem. Soc.* **1979**, 101, 6425–6428.
45. P. A. Frey, *Chemical Intermediates in Catalysis by Thiamine Diphosphate*. In *Thiamine. Catalytic Mechanisms in Normal and Disease States*; F. Jordan, M. S. Patel, Eds.; Marcel Dekker, Inc.: New York, 2004.
46. S. Kale; G. Ulas; J. Song; G. W. Brudvig; F. Jordan, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 1158–1163.
47. G. S. Brandt; N. Nemeria; S. Chakraborty; M. J. McLeish; A. Yep; G. L. Kenyon; G. A. Petsko; F. Jordan; D. Ringe, *Biochemistry* **2008**, 47, 7734–7743.
48. G. S. Brandt; M. Kneen; S. Chakraborty; A. Baykal; N. Nemeria; A. Yep; D. Ruby; G. A. Petsko; G. L. Kenyon; M. J. McLeish; F. Jordan; D. Ringe, *Biochemistry* **2009**, 48, 3247–3257.
49. E. A. Sergienko; F. Jordan, *Biochemistry* **2001**, 40, 7382–7403.
50. E. A. Sergienko; J. Wang; L. Polovnikova; M. S. Hasson; M. J. McLeish; G. L. Kenyon; F. Jordan, *Biochemistry* **2000**, 39, 13862–13869.
51. P. Arjunan; M. Sax; A. Brunskill; K. Chandrasekhar; N. Nemeria; S. Zhang; F. Jordan; W. Furey, *J. Biol. Chem.* **2006**, 281, 15296–15303.
52. G. Wille; D. Meyer; A. Steinmetz; E. Hinze; R. Golbik; K. Tittmann, *Nat. Chem. Biol.* **2006**, 2, 324–328.
53. S. Chakraborty; N. Nemeria; A. Yep; M. J. McLeish; G. L. Kenyon; F. Jordan, *Biochemistry* **2008**, 47, 3800–3809.
54. S. Zhang; F. Jordan, submitted.
55. S. Chakraborty; F. Jordan, submitted.
56. E. Joseph; F. Jordan, submitted.
57. S. Chakraborty; N. S. Nemeria; A. Balakrishnan; G. S. Brandt; M. M. Kneen; A. Yep; M. J. McLeish; G. L. Kenyon; G. A. Petsko; D. Ringe; F. Jordan, *Biochemistry* **2009**, 48, 981–994.
58. X. Zeng; A. Chung; M. Haran; F. Jordan, *J. Am. Chem. Soc.* **1991**, 113, 5842.
59. S. Menon-Rudolph; S. Nishikawa; X. Zeng; F. Jordan, *J. Am. Chem. Soc.* **1992**, 114, 10110–10112.

60. L. Meshalkina; U. Nilsson; C. Wikner; T. Kostikowa; G. Schneider, *Eur. J. Biochem.* **1997**, *244*, 646–652.
61. M. Merski; C. A. Townsend, *J. Am. Chem. Soc.* **2007**, *129*, 15750–15751.
62. S. Zhang; M. Liu; Y. Yan; Z. Zhang; F. Jordan, *J. Biol. Chem.* **2004**, *279*, 54312–54318.
63. F. Jordan; Z. H. Kudzin; C. B. Rios, *J. Am. Chem. Soc.* **1987**, *109*, 4415–4417.
64. C. B. Rios, Generation, Structure and Chemical Properties of Enamines Related to the Key Intermediate in Thiamin Diphosphate Dependent Enzymatic Pathways. Ph.D. Thesis, Rutgers University Graduate Faculty at Newark, NJ, 1988.
65. G. L. Barletta; W. P. Huskey; F. Jordan, *J. Am. Chem. Soc.* **1992**, *114*, 7607–7608.
66. G. L. Barletta; W. P. Huskey; F. Jordan, *J. Am. Chem. Soc.* **1997**, *119*, 2356–2362.
67. S. Chakraborty, Detection, Characterization and Determination of Kinetic Fates of Enzyme-bound Intermediates in Thiamin Diphosphate Dependent Enzymes. Ph.D. Thesis, Rutgers University Graduate Faculty at Newark, NJ, 2006.
68. K. J. Gruys; C. J. Halkides; P. A. Frey, *Biochemistry* **1987**, *26*, 7575–7585.
69. B. L. Bertagnolli; L. P. Hager, *J. Biol. Chem.* **1991**, *266*, 10168–10173.
70. L. S. Khailova; L. G. Korochkina; S. E. Severin, *Ann. N. Y. Acad. Sci.* **1989**, *573*, 37–54.
71. Y. Zou, Investigating the Mechanism of Catalysis in Thiamin Dependent Enzymes. Ph.D. Thesis, Rutgers University Graduate Faculty at Newark, NJ, 1999.
72. F. Jordan; H. Li; A. M. Brown, *Biochemistry* **1999**, *38*, 6369–6373.
73. S. Zhang; L. Zhou; N. Nemeria; Y. Yan; Z. Zhang; Y. Zou; F. Jordan, *Biochemistry* **2005**, *44*, 2237–2243.
74. K. Pan; F. Jordan, *Biochemistry* **1998**, *37*, 1357–1364.
75. N. Nemeria; P. Arjunan; A. Brunskill; A. Sheibani; W. Wei; Y. Yan; S. Zhang; F. Jordan; W. Furey, *Biochemistry* **2002**, *41*, 15459–15467.
76. L. Kerscher; D. Oesterhelt, *Eur. J. Biochem.* **1981**, *116*, 595–600.
77. E. Chabriere; M. Charon; A. Volbeda; L. Pieulle; E. Hatchikian; J. Fontecilla-Camps, *Nat. Struct. Biol.* **1999**, *6*, 182–190.
78. G. Barletta; A. C. Chung; C. B. Rios; F. Jordan; J. M. Schlegel, *J. Am. Chem. Soc.* **1990**, *112*, 8144–8149.
79. S. O. Mansoorabadi; J. Seravalli; C. Furdul; V. Krymov; G. J. Gerfen; T. P. Begley; J. Melnick; S. W. Ragsdale; G. H. Reed, *Biochemistry* **2006**, *45*, 7122–7131.
80. A. Kaplun; E. Binshtein; M. Vyazmensky; A. Steinmetz; Z. Barak; D. M. Chipman; K. Tittmann; B. Shaanan, *Nat. Chem. Biol.* **2008**, *4*, 113–118.
81. L. Tretter; V. Adam-Vizi, *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2005**, *360*, 2335–2345.
82. R. A. Frank; C. W. Kay; J. Hirst; B. F. Luisi, *J. Am. Chem. Soc.* **2008**, *130*, 1662–1668.
83. F. Jordan; D. J. Kuo; E. U. Monse, *J. Am. Chem. Soc.* **1978**, *100*, 2872–2878.
84. M. J. DeNiro; S. Epstein, *Science* **1977**, *197*, 261–263.
85. M. H. O’Leary, *Biochem. Biophys. Res. Commun.* **1976**, *73*, 614–618.
86. F. Jordan; D. J. Kuo; E. U. Monse, *J. Org. Chem.* **1978**, *43*, 2828–2830.
87. D. W. Huhta; T. Heckenthaler; F. J. Alvarez; J. Ermer; G. Hübner; A. Schellenberger; R. L. Schowen, *Acta Chem Scand.* **1992**, *46*, 778–788.
88. F. J. Alvarez; J. Ermer; G. Huebner; A. Schellenberger; R. L. Schowen, *J. Am. Chem. Soc.* **1995**, *117*, 1678–1683.
89. J. Wang; R. Golbik; B. Seliger; M. Spinka; K. Tittmann; G. Hübner; F. Jordan, *Biochemistry* **2001**, *40*, 1755–1763.
90. W. Wei; M. Liu; F. Jordan, *Biochemistry* **2002**, *41*, 451–461.
91. E. Joseph; W. Wei; K. Tittmann; F. Jordan, *Biochemistry* **2006**, *45*, 13517–13527.
92. N. Nemeria; K. Tittmann; E. Joseph; L. Zhou; M. B. Vazquez-Coll; P. Arjunan; G. Hübner; W. Furey; F. Jordan, *J. Biol. Chem.* **2005**, *280*, 21473–21482.
93. A. Balakrishnan; S. Chakraborty; N. Nemeria; F. Jordan, in preparation.
94. A. Balakrishnan, unpublished.
95. G. Lu; D. Dobritzsch; S. Baumann; G. Schneider; S. König, *Eur. J. Biochem.* **2000**, *267*, 861–867.
96. I. Eberhardt; H. Cederberg; S. Konig; H. Li; F. Jordan; S. Hohmann, *Eur. J. Biochem.* **1999**, *262*, 191–202.
97. P. Arjunan; N. Nemeria; A. Brunskill; K. Chandrasekhar; M. Sax; Y. Yan; F. Jordan; J. R. Guest; W. Furey, *Biochemistry* **2002**, *41*, 5213–5221.
98. P. Arjunan; M. Sax; A. Brunskill; K. Chandrasekhar; N. Nemeria; S. Zhang; F. Jordan; W. Furey, *J. Biol. Chem.* **2006**, *281*, 15296–15303.
99. S. Kale; P. Arjunan; W. Furey; F. Jordan, *J. Biol. Chem.* **2007**, *282*, 28106–28116.
100. S. Kale, Characterization of the Dynamics in the E1 Component of *Escherichia coli* Pyruvate Dehydrogenase. Ph.D. Dissertation, Rutgers University Graduate Faculty at Newark, Newark, NJ, 2008.
101. A. Baykal; S. Chakraborty; A. Dadoo; F. Jordan, *Bioorg. Chem.* **2006**, *34*, 380–393.
102. W. Wei; H. Li; N. Nemeria; F. Jordan, *Protein Expr. Purif.* **2003**, *28*, 140–150.
103. The spectral simulations used WINDNMR-Pro , H.J. Reich, Dynamic NMR Spectra for Windows. *J. Chem. Educ. Software* **1996** 3D2.
104. Y.-H. Park; W. Wei; L. Zhou; N. Nemeria; F. Jordan, *Biochemistry* **2004**, *43*, 14037–14046.
105. Y.-H. Park, Protein Interactions in the Pyruvate Dehydrogenase Complex from *Escherichia coli*. Ph.D. Thesis, Rutgers University Graduate Faculty at Newark, NJ, 2008.
106. A. F. Hengeveld; A. de Kok, *FEBS Lett.* **2002**, *522*, 173–176.
107. E. Schulze; A. H. Westphal; C. Veeger; A. de Kok, *Eur. J. Biochem.* **1992**, *206*, 427–435.
108. S. S. Mande; S. Sarfaty; M. D. Allen; R. N. Perham; W. G. J. Hol, *Structure* **1996**, *4*, 277–286.
109. R. Wolfenden, *Nature* **1969**, *223*, 704–705.
110. G. E. Lienhard, *Science* **1973**, *180*, 149–154.
111. J. A. Gutowski; G. E. Lienhard, *J. Biol. Chem.* **1976**, *251*, 2863–2866.
112. N. Nemeria; Y. Yan; Z. Zhang; A. M. Brown; P. Arjunan; W. Furey; J. R. Guest; F. Jordan, *J. Biol. Chem.* **2001**, *276*, 45969–45978.
113. W. Versées; S. Spaepen; M. D. Wood; F. J. Leeper; J. Vanderleyden; J. Steyaert, *J. Biol. Chem.* **2007**, *282*, 35269–35278.
114. C. L. Berthold; D. Gocke; M. D. Wood; F. J. Leeper; M. Pohl; G. Schneider, *Acta Cryst. D Biol. Crystallogr.* **2007**, *63*, 1217–1224.
115. R. A. W. Frank; C. M. Titman; J. V. Pratar; B. F. Luisi; R. N. Perham, *Science* **2004**, *306*, 872–876.



116. F. Jordan, *Science* **2004**, *306*, 818–820.  
117. P. Arjunan; N. Nemeria; K. Tittmann; W. Furey; F. Jordan, in preparation.  
118. F. Jordan; N. Nemeria; E. Sergienko, *Acc. Chem. Res.* **2005**, *38*, 755–763.  
119. J. Crosby; G. E. Lienhard, *J. Am. Chem. Soc.* **1970**, *92*, 5707–5716.  
120. J. Crosby; R. Stone; G. E. Lienhard, *J. Am. Chem. Soc.* **1970**, *92*, 2891–2900.  
121. D. S. Kemp; J. T. O'Brien, *J. Am. Chem. Soc.* **1970**, *92*, 2554–2555.  
122. J. Ullrich; I. Donner, *Hoppe-Seylers Z. Physiol. Chem.* **1970**, *351*, 1030–1034.

### Biographical Sketch



Frank Jordan was born in Budapest, Hungary and immigrated to the United States in 1957. He received his undergraduate training at Drexel University in Philadelphia (B.Sc. in 1964 in chemistry) and his Ph.D. at the University of Pennsylvania's Department of Chemistry in 1967 working with Edward R. Thornton. This was followed by a NATO fellowship in the laboratory of the late Bernard Pullman at the Sorbonne in Paris (1967–68) on problems in quantum biochemistry and an NIH fellowship in the laboratory of the late Frank H. Westheimer at the Department of Chemistry at Harvard, carrying out research on a model study for the enzyme acetoacetate decarboxylase (1968–70). In 1970, he joined the Department of Chemistry at Rutgers in Newark as an assistant professor, rising to the rank of full professor in 1979 and to his current title as Rutgers University Board of Governors Professor of Chemistry in 1997. He served as chair of his department from 1985–91, 1994–2000, and 2008–present. He was a recipient of (1) the Rutgers University Board of Trustees Award for Excellence in Research in 1983, (2) a Johnson and Johnson Research Discovery Fellowship in 1988–90, (3) the Honor Scroll of the American Institute of Chemists, New Jersey Section in 1995, and (4) the 1998 Excellence in Education Award of the North Jersey Section of the American Chemical Society. He was elected a Fellow of the American Association for the Advancement of Science (AAAS) in 1995. He has also served on NIH Study Sections and on the editorial board of the *Journal of Biological Chemistry*. In addition to studies described in this chapter, he has a long-standing interest in the structure, mechanism, and folding of serine proteases. He is the author or coauthor of more than 210 publications and reviews and 3 patents, and has graduated 55 Ph.D. students from his laboratories.



## 7.17 The Biosynthesis of Folate and Pterins and Their Enzymology

**Markus Fischer**, University of Hamburg, Hamburg, Germany

**Beat Thöny**, University Children's Hospital, Zürich, Switzerland

**Silke Leimkühler**, University of Potsdam, Potsdam, Germany

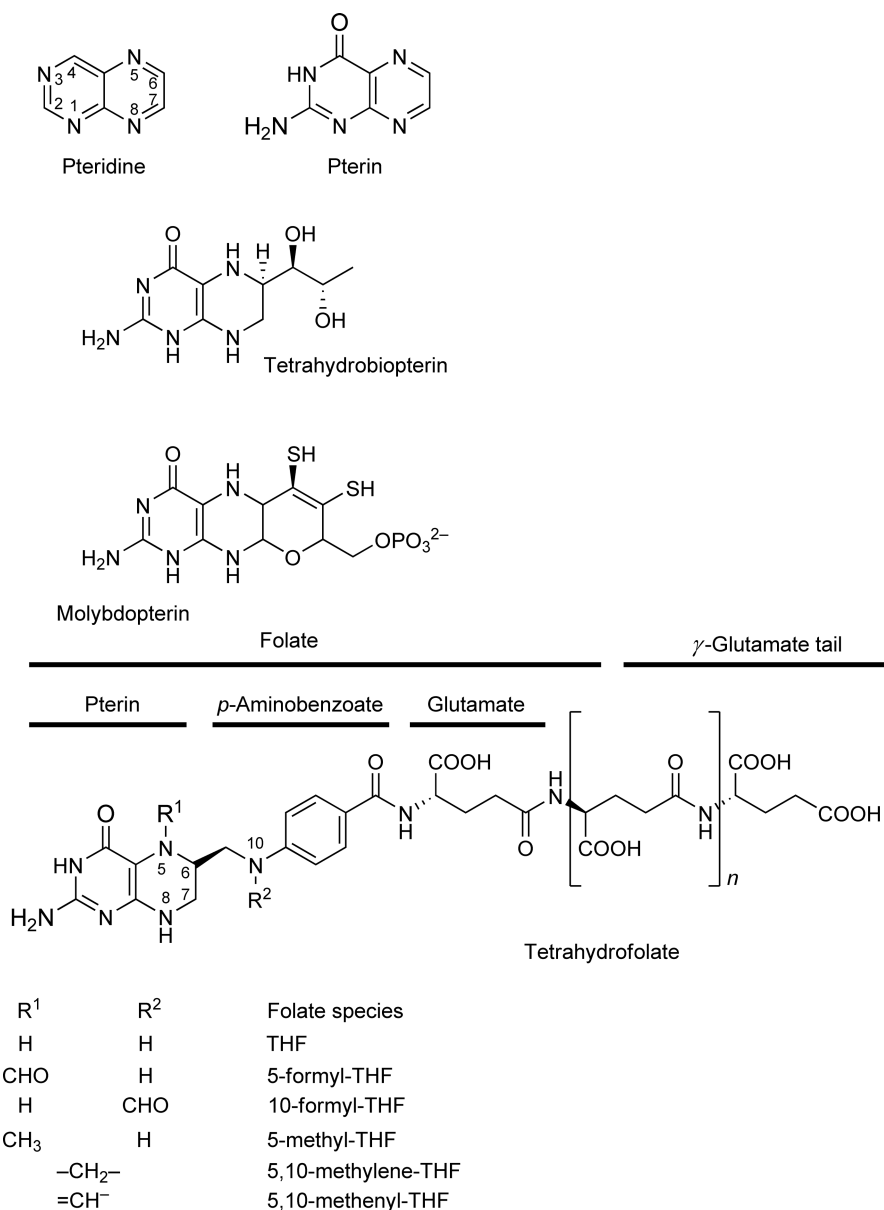
© 2010 Elsevier Ltd. All rights reserved.

<b>7.17.1</b>	<b>Introduction</b>	600
<b>7.17.2</b>	<b>The Biosynthesis of Folate</b>	602
7.17.2.1	An Introduction into the Biosynthesis of Tetrahydrofolate	602
7.17.2.2	Functions of Folate-Type Coenzymes	602
7.17.2.2.1	Tetrahydrofolate derivatives	602
7.17.2.2.2	Metabolic interconversion of tetrahydrofolate derivatives and one-carbon transfer reactions	603
7.17.2.3	Biosynthesis of Tetrahydrofolate	603
7.17.2.3.1	Biosynthesis of 4-aminobenzoate	605
7.17.2.3.2	The pterin branch	606
7.17.2.4	Tetrahydrofolate Enzymes as Drug Targets	613
<b>7.17.3</b>	<b>Tetrahydrobiopterin Metabolism and Function</b>	616
7.17.3.1	An Introduction to Tetrahydrobiopterin Biosynthesis	616
7.17.3.2	Functions of Tetrahydrobiopterin	616
7.17.3.2.1	Cofactor-dependent enzyme systems	616
7.17.3.2.2	Cellular and systemic functions of BH <sub>4</sub>	619
7.17.3.3	Biosynthesis of BH <sub>4</sub>	620
7.17.3.3.1	Reaction mechanism of the <i>de novo</i> pathway	620
7.17.3.3.2	Alternative routes for biosynthesis of BH <sub>4</sub>	623
7.17.3.4	Regeneration of BH <sub>4</sub>	624
7.17.3.4.1	Reaction mechanism of the regenerative pathway	625
7.17.3.5	Tetrahydrobiopterin in Disease	627
7.17.3.6	Animal Models to Study BH <sub>4</sub> Function and Deficiency	628
<b>7.17.4</b>	<b>Molybdopterin and Molybdenum Cofactor Biosynthesis</b>	628
7.17.4.1	An Introduction to Molybdopterin Biosynthesis and Molybdoenzymes	628
7.17.4.2	The Families of Molybdoenzymes	630
7.17.4.2.1	The xanthine oxidase family	630
7.17.4.2.2	The sulfite oxidase family	630
7.17.4.2.3	The DMSO reductase family	630
7.17.4.2.4	The enzymology of <i>Rhodobacter capsulatus</i> xanthine dehydrogenase (XDH)	630
7.17.4.3	Biosynthesis of Molybdopterin and the Molybdenum Cofactor	632
7.17.4.3.1	Conversion of guanosine monophosphate to precursor Z	632
7.17.4.3.2	Insertion of sulfur into precursor Z and formation of molybdopterin	633
7.17.4.3.3	Insertion of molybdenum into molybdopterin	635
7.17.4.4	Additional Modification of Molybdenum Cofactor	636
7.17.4.4.1	Attachment of guanosine monophosphate	636
7.17.4.4.2	Formation of sulfurated molybdenum cofactor	637
7.17.4.5	The Storage of Molybdenum Cofactor and Its Insertion into Molybdoenzymes	637
7.17.4.5.1	Chaperones for the insertion and targeting of bis-molybdopterin guanine dinucleotide and sulfurated molybdenum cofactor into prokaryotic molybdoenzymes	637

7.17.4.5.2	Molybdenum cofactor carrier proteins in eukaryotes	638
7.17.4.6	The Importance of Molybdenum Cofactor in Humans	638
7.17.4.6.1	Molybdenum cofactor deficiency and isolated sulfate oxidase deficiency	638
7.17.4.6.2	Animal models and first therapies for molybdenum cofactor deficiency	639
References		641

### 7.17.1 Introduction

Pterins belong to the pteridine family of heterocycles<sup>1</sup> (Scheme 1). Pteridines were first discovered in the pigments of butterfly wings<sup>2</sup> (from Greek *pteros* = wing) and are composed of fused pyrimidine and pyrazine rings containing a wide variety of substitutions on this structure.<sup>1</sup> In pterins, the pyrimidine ring has a carbonyl



**Scheme 1** Structures of pteridine, pterin, tetrahydrobiopterin, molybdopterin, and tetrahydrofolate.

oxygen and an amino group.<sup>3</sup> Pterins perform many roles in biological processes as cofactors in enzymatic catalysis. Folates (from Latin *folium*=leaf) are 'conjugated' pterins (**Scheme 1**), which contain para-aminobenzoic acid and glutamates in addition to the pterin ring system.<sup>1,3</sup> In the form of a series of tetrahydrofolate (THF) compounds, folate derivatives are substrates in a number of single-carbon-transfer reactions.<sup>4</sup>

This chapter gives an overview of the biosynthesis of THF, (6*R*)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH<sub>4</sub>), and molybdopterin (MPT), the organic compound of the molybdenum cofactor (Moco), and discusses the enzymology of the proteins that bind the three classes of cofactors (structures are shown in **Scheme 1**). The biosynthesis of THF, BH<sub>4</sub>, and MPT start from guanosine triphosphate (GTP), but the reactions for the conversion of GTP into the three cofactors are diverse and involve different enzymatic compounds.

THF (vitamin B<sub>9</sub>) and its folate derivatives are essential cofactors in one-carbon (C1) transfer reactions and absolutely required for the synthesis of a variety of different compounds. Even when microorganisms are supplied with exogenous sources of all metabolites whose biosynthesis requires folate-type cofactors, they require the coenzyme for the initiation of protein biosynthesis. Reduced folate supplies the cellular C1 pool with C1 carrier units, which act as enzyme cofactors in C1 transfer reactions essential for the synthesis of purines, thymidylate, glycine, methionine, pantothenic acid, and *N*-formylmethionyl-tRNA in all organisms. Most plants, microbial eukaryotes, and prokaryotes synthesize folate *de novo*. Mammalian cells are unable to carry out *de novo* synthesis of folate. Therefore, folate is considered a vitamin in these organisms, and the cells use a carrier-mediated active transport system for the uptake of folate. Folate deficiency causes massive incorporation of uracil into the DNA and chromosome breaks during the repair by uracil-DNA glycosylase and apyrimidinic endonuclease.<sup>5</sup> The biosynthesis of THF has been studied in some detail and the first step is catalyzed by GTP cyclohydrolase I which converts GTP into dihydroneopterin triphosphate (DHNTTP). The triphosphate motif is removed and the resulting 7,8-dihydro-D-neopterin is converted into 6-hydroxymethyl-dihydropterin by dihydroneopterin aldolase (DHNA). The consecutive action of different enzymes finally results in 6-(*S*)-tetrahydrofolate through different intermediates. Enzymes of the folate biosynthesis pathway are important drug targets for the treatment of infections with bacteria and certain malignancies. Potent inhibitors of folate biosynthesis were discovered about two decades earlier than the vitamin whose biosynthesis they interfere with. In fact, sulfonamide drugs that interfere with the formation of the THF precursor, dihydropteroate synthase, were discovered and widely used therapeutically more than a decade prior to the structure elucidation of folic acid.

BH<sub>4</sub> is an essential cofactor of a set of enzymes that are of central metabolic importance, that is, the hydroxylases of the three aromatic amino acids phenylalanine, tyrosine, and tryptophan; of ether lipid oxidase; and of the three nitric oxide synthase (NOS) isoenzymes.<sup>6,7</sup> As a consequence, BH<sub>4</sub> plays a key role in a vast number of biological processes and pathological states associated with neurotransmitter formation, vasorelaxation, and immune response, and is present in probably every cell or tissue of higher organisms. In mammals, its biosynthesis is controlled by hormones, cytokines, and certain immune stimuli. With regard to human disease, BH<sub>4</sub> deficiency due to autosomal recessive mutations in all enzymes (except sepiapterin reductase) has been described as a cause of hyperphenylalaninemia. Furthermore, several neurological diseases, including Dopa-responsive dystonia (DSD) and also Alzheimer's disease, Parkinson's disease, autism, and depression, have been suggested to be a consequence of restricted cofactor availability.

BH<sub>4</sub> is formed *de novo* from GTP through a sequence of three enzymatic steps carried out by GTP cyclohydrolase I (EC 3.5.4.16) (GTPCH), 6-pyruvoyl tetrahydropterin synthase (EC 4.6.1.10) (PTPS), and sepiapterin reductase (EC 1.1.1.153) (SR). Cofactor regeneration requires pterin-4a-carbinolamine dehydratase and dihydropteridine reductase. Based on gene cloning, recombinant expression, mutagenesis studies, structural analysis of crystals, and NMR studies, reaction mechanisms for the biosynthetic and recycling enzymes were proposed. With regard to the regulation of cofactor biosynthesis, the major controlling point is GTPCH, the expression of which may be under the control of cytokine induction.

MPT is the chemical backbone for the coordination of transition metals such as molybdenum and tungsten, which are coordinated with the dithiolene group on the 6-alkyl side chain of MPT.<sup>8,9</sup> In its molybdenum-bound state, the so-called Moco is an essential component of a diverse group of redox-active molybdoenzymes, comprising of more than 40 enzymes catalyzing a variety of transformations at carbon, sulfur, and nitrogen atoms.<sup>10</sup> Some of the better known molybdoenzymes include sulfite oxidase, xanthine dehydrogenase/oxidase, nitrate reductase, formate dehydrogenase, and carbon monoxide dehydrogenase. In bacteria, the enzymes generally catalyze nonessential redox reactions and are often used as terminal electron acceptors during

anaerobic respiration. In humans, sulfite oxidase is essential for the last step of the degradation of the sulfur-containing amino acids, cysteine and methionine, and it catalyzes the oxidation of sulfite to sulfate, which is further excreted in the urine.<sup>11</sup> Sulfite oxidase deficiency leads to severe neurological abnormalities such as attenuated growth of the brain, seizures, and dislocated ocular lenses and usually results in death in early childhood.<sup>12</sup> So far, no effective therapy is available to cure the disease.

The biosynthesis of Moco starts with GTP, which is converted in a GTP cyclohydrolase-independent reaction to the first stable intermediate precursor Z, a cyclic pyranopterin monophosphate.<sup>13,14</sup> Precursor Z is converted into MPT by the insertion of two sulfur atoms, forming the dithiolene group which coordinates the molybdenum atom.<sup>15</sup> After insertion of molybdate, Moco is formed that can be further modified and is inserted into various molybdoenzymes.<sup>16</sup>

## **7.17.2 The Biosynthesis of Folate**

### **7.17.2.1 An Introduction into the Biosynthesis of Tetrahydrofolate**

Well before the structural elucidation of THF, the sulfonamides which were later shown to act through the inhibition of THF biosynthesis were introduced as the first synthetic antibacterial agents with a broad action spectrum.<sup>17,18</sup> Later on, enzymes involved in the mammalian metabolism of folates acquired additional medical relevance as targets of cytostatic and immunosuppressive agents. Owing to these medical aspects, certain enzymes of the THF biosynthesis and utilization pathways have been investigated in substantial detail. The coverage of different enzymes of the folate pathway in the literature varies by more than two orders of magnitude with dihydrofolate reductase as the front runner and as one of the most investigated proteins. The volume of published literature on folate biosynthesis and related areas defies comprehensive treatment within the format of the present chapter.

### **7.17.2.2 Functions of Folate-Type Coenzymes**

Folates are essential in all living systems with the exception of methanogenic bacteria, where they are replaced by methanopterin derivatives. The superfamily of folate and methanopterin-type coenzymes serves the transfer of C1 fragments at the levels of methyl, methylene, and formyl groups (whose oxidation levels correspond to those of methanol, formaldehyde, and formate), whereas biotin is the cofactor in charge of CO<sub>2</sub>, the C1 compound with the highest level of oxidation in biological systems. The various C1 moieties carried by THF-type coenzymes serve as building blocks for the biosynthesis of purines, pyrimidines, and methionine. Formyltetrahydrofolate also affords formylmethionine tRNA that serves as the initiator tRNA<sup>Met-i</sup> in eubacterial and mitochondrial protein biosynthesis.<sup>19</sup>

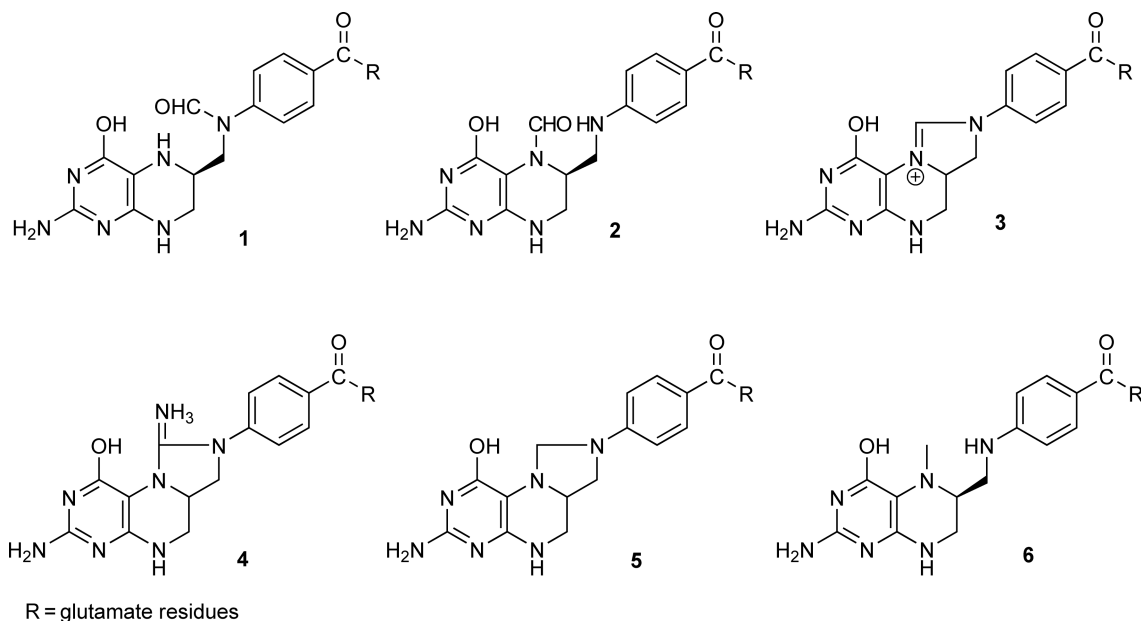
For thymidylate synthase in animals, plants, and many but not all bacteria, methylene tetrahydrofolate serves not only as a C1 transponder but also as a reductive agent;<sup>20,21</sup> the dihydrofolate resulting from that reaction needs to be recycled by the salvage action of dihydrofolate reductase, the most intensely studied enzyme in the family of folate-cofactor processing enzymes.

It should also be noted that methenyltetrahydrofolate serves as an optical transponder in DNA photolyases that are involved in the photochemically driven repair of photodamaged DNA in numerous organisms, albeit not in mammals.<sup>22</sup>

The cytoplasmic, mitochondrial, and chloroplast compartments are all involved in aspects of folate coenzyme metabolism in eukaryotic organisms.

#### **7.17.2.2.1 Tetrahydrofolate derivatives**

The structural complexity of folate-type coenzymes is quite large by comparison with other coenzymes for several reasons. (1) C1 fragments can be present in three different oxidation levels. (2) C1 fragments can be attached at N-5 and/or at N-10 (specifically, C1 fragments at the oxidation levels of formaldehyde and formate can attach to both N-5 and N-10 under formation of five-membered heterocyclic ring systems). (3) Whereas THF carries a single glutamate residue, the coenzyme forms carry polyglutamyl tails of variable lengths (**Scheme 1**). The various cofactor forms are summarized in **Figure 1**.



**Figure 1** THF derivatives. **1**, 10-formyl-THF; **2**, 5-formyl-THF; **3**, 5,10-methenyl-THF; **4**, 5-formimino-THF; **5**, 5,10-methylene-THF; **6**, 5-methyl-THF.

#### 7.17.2.2 Metabolic interconversion of tetrahydrofolate derivatives and one-carbon transfer reactions

Major enzyme reactions involved in the transfer of C1 moieties and the concomitant transformations of the cognate coenzyme forms are summarized in **Table 1**. Whereas animals use folate-type coenzymes predominantly for the biosynthesis of nucleotides, plants and many microorganisms require them also for the biosynthesis of several essential amino acids. A detailed description of the metabolic pathways that depend on the cooperation of folate-type coenzymes is beyond the scope of the present chapter.

#### 7.17.2.3 Biosynthesis of Tetrahydrofolate

Dihydrofolate is assembled by a convergent pathway from a pteridine precursor (6-hydroxymethyldihydropterin pyrophosphate), 4-aminobenzoate, and glutamate (**Figure 2**). The following paragraphs describe the biosynthesis pathways of 4-aminobenzoate and the pteridine precursor, 6-hydroxymethyl-dihydropterine

**Table 1** Major metabolic interconversion reaction of THF derivatives

Enzyme	Folate interconversion
5,10-Methenyl-THF cyclohydrolase	10-Formyl-THF → 5,10-methenyl-THF
5,10-Methylene-THF dehydrogenase	5,10-Methenyl-THF → 5,10-methylene-THF
5,10-Methylene-THF reductase	5,10-Methylene-THF → 5-methyl-THF
10-Formyl-THF dehydrogenase	10-Formyl-THF → THF
10-Formyl-THF-AICAR formyl transferase	10-Formyl-THF → THF
10-Formyl-THF-GAR formyl transferase	10-Formyl-THF → THF
Thymidylate synthase	5,10-Methylene-THF → DHF
Methionine synthase	5-Methyl-THF → THF
7,8-Dihydrofolate reductase	DHF → THF
Serine hydroxymethyl transferase	THF → 5,10-methylene-THF
Glutamate formimino transferase	THF → formimino-THF
Formimino-THF-cyclodesaminase	Formimino-THF → 5,10-methenyl-THF
10-Formyl-THF synthetase	THF → 10-formyl-THF

AICAR, 4-amino-5-imidazolecarboxamidribotide; GAR, glycinamideribotide.

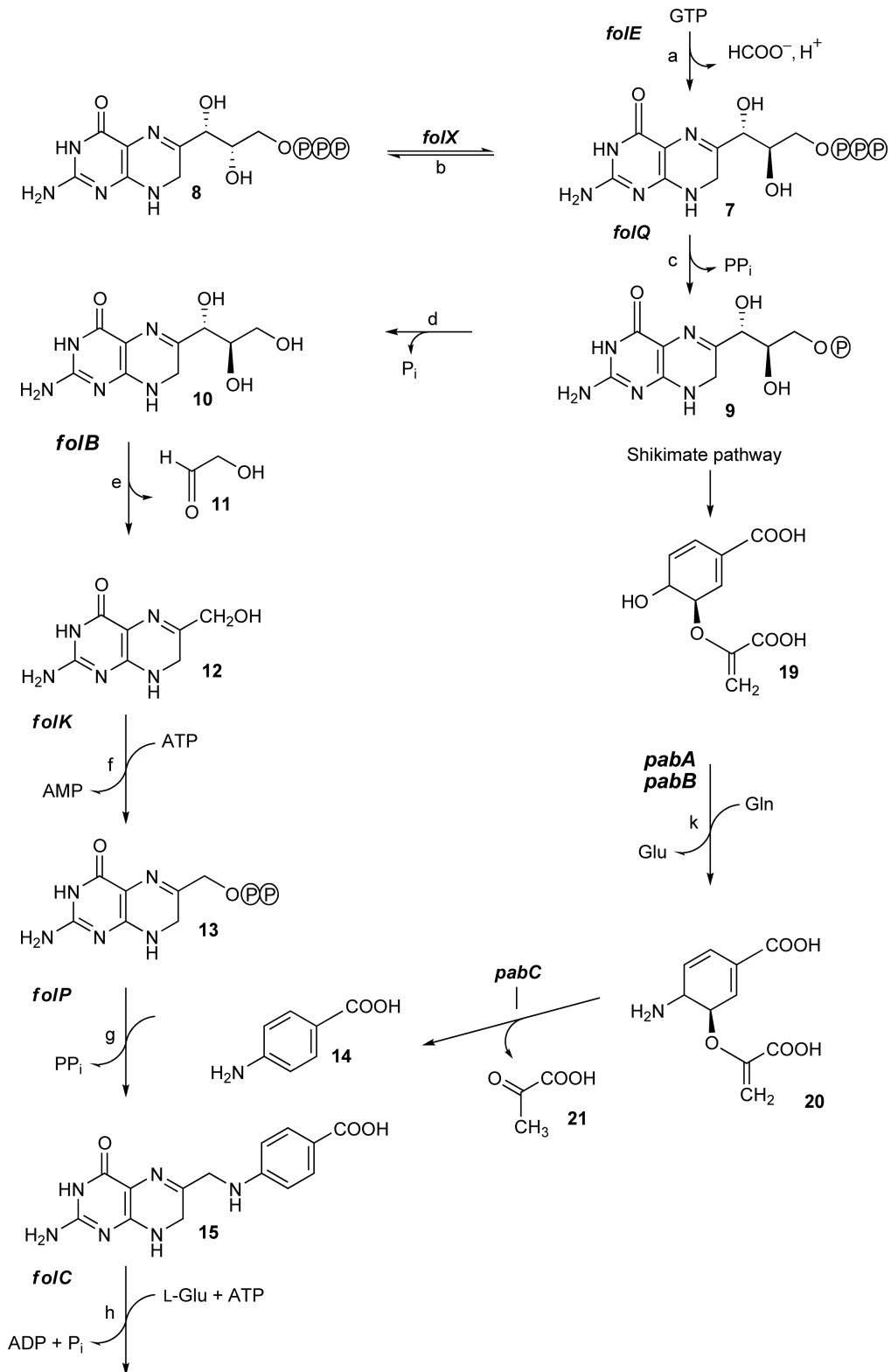
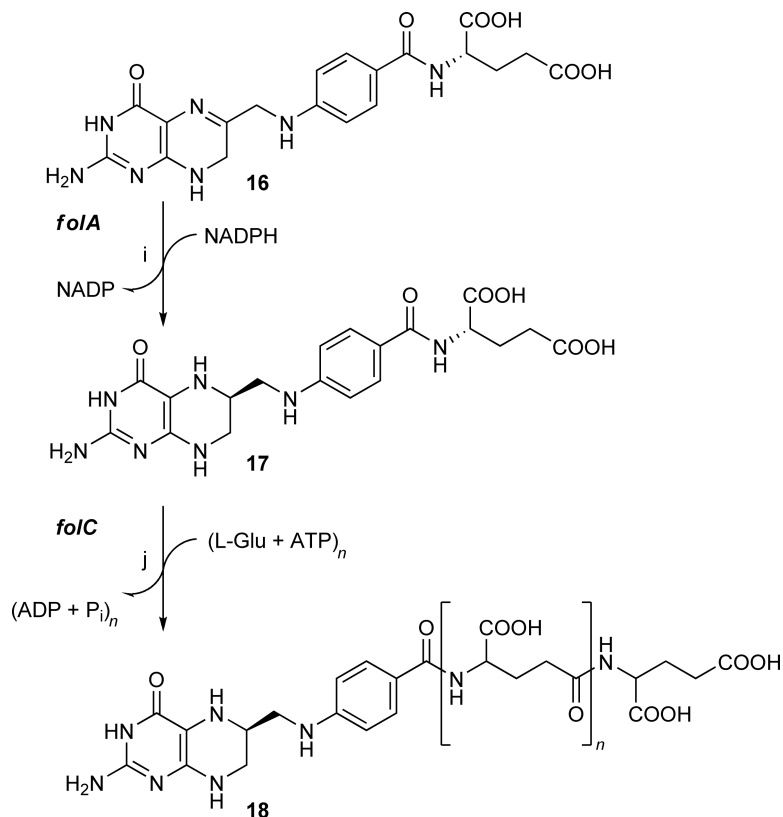


Figure 2 (Continued)





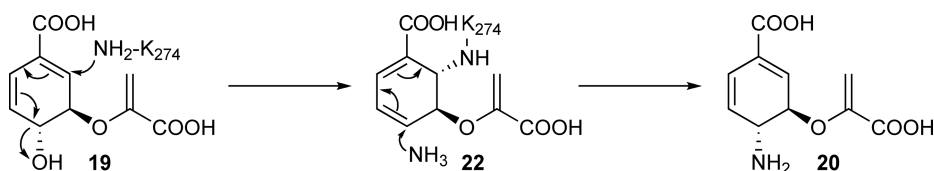
**Figure 2** Biosynthesis of folates. a, GTP cyclohydrolase I; b, dihydroneopterin epimerase; c, nudix hydrolase; d, phosphatase; e, dihydroneopterin aldolase; f, hydroxymethyl-dihydropterin pyrophosphokinase; g, dihydropteroate synthase; h, dihydrofolate synthetase; i, dihydrofolate reductase; j, folylpolyglutamate synthetase; k, aminodeoxychorismate synthase; l, aminodeoxychorismate lyase. **7**, 7,8-dihydro-D-neopterin 3'-triphosphate; **8**, 7,8-dihydro-L-monapterin 3'-triphosphate; **9**, 7,8-dihydro-D-neopterin 3'-monophosphate; **10**, 7,8-dihydro-D-neopterin; **11**, glycolaldehyde; **12**, 6-hydroxymethyl-7,8-dihydropterin; **13**, 6-hydroxymethyl-7,8-dihydropterin 3'-pyrophosphate; **14**, *p*-aminobenzoate; **15**, dihydropteroate; **16**, dihydrofolate; **17**, tetrahydrofolate; **18**, tetrahydrofolate polyglutamates; **19**, chorismate; **20**, 4-amino-4-deoxychorismate; **21**, pyruvate. Gene names based on *E. coli* are given in italics.

pyrophosphate, and the assembly of dihydrofolate and its conversion into THF. Three-dimensional structures have been determined for representatives of all pathway enzymes.

### 7.17.2.3.1 Biosynthesis of 4-aminobenzoate

Whereas it had been known that 4-aminobenzoate (*p*-aminobenzoate, **14**) is derived from the shikimate pathway (**Figure 2**) through chorismate (**19**, **Figure 3**), the details of that transformation have been elucidated relatively recently.

**7.17.2.3.1(i) 4-Amino-4-deoxychorismate synthase (EC 6.3.5.8)** PabAB protein (**Figure 2**, k) catalyzes the substitution of the position 4 hydroxyl group of chorismate (**19**) by an amino group. In closer detail, the first



**Figure 3** Mechanism of 4-amino-4-deoxychorismate synthase.

reaction step consists of a vinylogous addition/elimination reaction. More specifically, the amino group of lysine 274 (in case of the *Escherichia coli* protein) is covalently linked to the substrate under formation of **22** and a water molecule (Figure 3).<sup>23</sup>

An ammonia molecule that is obtained by the hydrolysis of glutamine is then added to the position 4 carbon and the bond between the substrate and lysine 274 is broken by way of a second vinylogous addition/elimination reaction. The overall reaction sequence is similar to the biosynthetic formation of anthranilic acid that serves as a biosynthetic precursor of tryptophan.

Interestingly, PabAB protein can catalyze the formation of 4-aminobenzoate at a slow rate even after mutagenesis of lysine 274; in the mutant, the function of the lysine amino group can be taken over by yet another ammonia molecule.

The structure of the PabA protein of *E. coli* has been determined by X-ray crystallography.<sup>24</sup>

In plants, the functional equivalent of the *E. coli* PabAB protein is a fusion protein with two folding domains that are equivalent to the subunits of the *E. coli* heterodimer.<sup>25</sup>

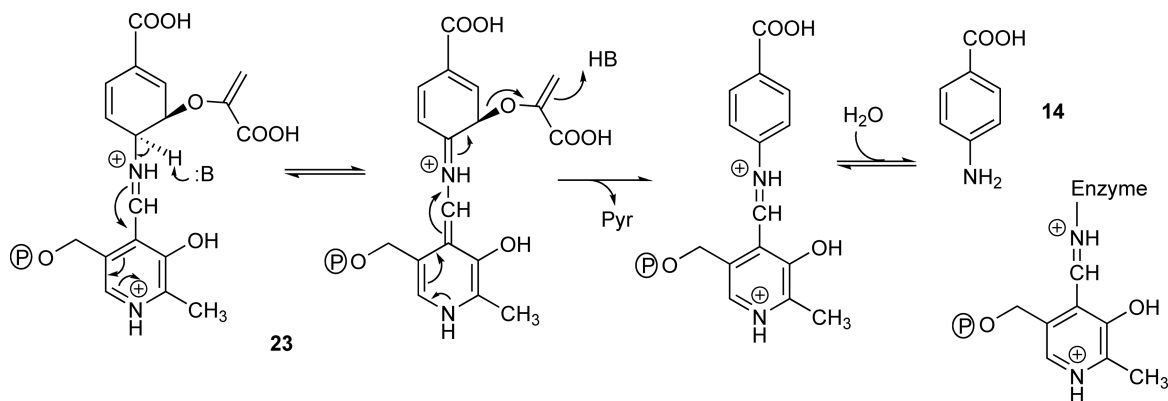
**7.17.2.3.1(ii) 4-Amino-4-deoxychorismate lyase (EC 4.1.3.38)** PabC protein (Figure 2, l) catalyzes the elimination of pyruvate (**21**) from 4-amino-4-deoxychorismate (**20**). The enzyme utilizes pyridoxal phosphate as a cofactor. The reaction mechanism shown in Figure 4 implies that pyridoxal phosphate forms a Schiff base (**23**) with the substrate, 4-amino-4-deoxychorismate. Abstraction of a proton from C-4 of the amino-4-deoxychorismate moiety is believed to result in a shift of the imine bond. Pyruvate can then be eliminated from the Schiff base form **23**, and the pyridoxal phosphate moiety can be transferred back to a lysine residue (specifically, lysine 159 of *E. coli* PabC protein) under liberation of 4-amino benzoate (**14**) as shown by X-ray crystallography (Figure 4).<sup>27</sup>

4-Amino-4-deoxychorismate lyases from fungi are devoid of sequence similarity with the bacterial or plant enzymes.<sup>28</sup>

### 7.17.2.3.2 The pterin branch

**7.17.2.3.2(i) Guanosine triphosphate cyclohydrolase I (EC 3.5.4.16)** The first committed step (Figure 2, a) in the biosynthesis of the heterocyclic moiety of THF involves a ring expansion affording the pteridine derivative, DHNTP, from GTP. That complex reaction is catalyzed by a GTP cyclohydrolase that was discovered and studied in the 1960s by the research groups of Brown and Shiota using bacterial cell extracts (for a review of the early work see Brown and Williamson<sup>29</sup>). Later studies identified enzymes of other pathways that also catalyze the hydrolytic opening of the imidazole ring of GTP but afford different products.<sup>30,31</sup>

The GTP cyclohydrolase involved in folate biosynthesis was subsequently designated GTP cyclohydrolase I, whereas enzymes involved in the biosynthetic pathways of riboflavin and methanopterin were designated GTP cyclohydrolase II and III, respectively.



**Figure 4** Proposed mechanism for the role of pyridoxal phosphate in the 4-amino-4-deoxychorismate lyase reaction (see Section 7.17.3.3.1(ii)). Pyr, pyruvate.<sup>26</sup>

Whereas plants and microorganisms use GTP cyclohydrolase I for the biosynthesis of THF, vertebrates use an orthologue of the microbial GTP cyclohydrolase I for the biosynthesis of the cofactor BH<sub>4</sub>.<sup>32,33</sup>

Type I GTP cyclohydrolases from bacteria, fungi, and animals are homodecamers, those of plants are homodimers.<sup>34,35</sup> The enzymes from plants contain two GTPCHI-like domains in tandem. Neither domain of these proteins has a full set of the residues involved in substrate binding and catalysis in other GTPCHIs. Hence, it is believed that both domains have evolved from a common ancestor by gene duplication. Both domains of the plant enzymes have similarity to the microbial or mammalian enzymes. Neither domain of these proteins has a full set of the residues involved in substrate binding and catalysis in other GTPCHIs. Neither domain expressed separately shows GTPCHI activity.<sup>35</sup>

Formally, the homodecameric enzymes of bacteria can be described as D<sub>5</sub>-symmetric dimers of homopentamers.<sup>32,34,36,37</sup> The 10 topologically active sites are located at the interface regions of two respective subunits in one of the pentameric rings and one respective subunit in the adjacent pentameric ring.

The classical type I GTP cyclohydrolases all carry one essential Zn<sup>2+</sup> ion per subunit (see, however, below for a different class of type I cyclohydrolases). In the decameric enzymes from microorganisms and animals, the metal ion is complexed by two cysteine and one histidine residues. The enzyme from *E. coli* or *H. sapiens* can lose the metal relatively easily.<sup>32</sup>

The hypothetical reaction mechanism of GTP cyclohydrolase I is shown in **Figure 5**. The first reaction step in the trajectory of GTP cyclohydrolase I involves the addition of a water molecule to C-8 in the imidazole ring of GTP (**Figure 5**). The X-ray structure of the *E. coli* enzyme shows a fixed water molecule that is complexed by the Zn<sup>2+</sup> ion and His112 that is believed to act as the nucleophile for the hydrolytic attack of C-8.<sup>38</sup> In the intermediate **26** with a quaternary C-8 atom, the bond between C-8 and N-9 can be broken under formation of the formamide derivative **27**. The Zn<sup>2+</sup> is believed to serve a second time as a Lewis acid for the activation of a water molecule, this time for addition to the formyl carbon with subsequent cleavage of the bond between N-7 and C-8 under the formation of formate. The ribosyl side chain must then undergo an Amadori rearrangement in preparation for ring closure. The Amadori compound is then ready for cyclization with the formation of DHNTP.

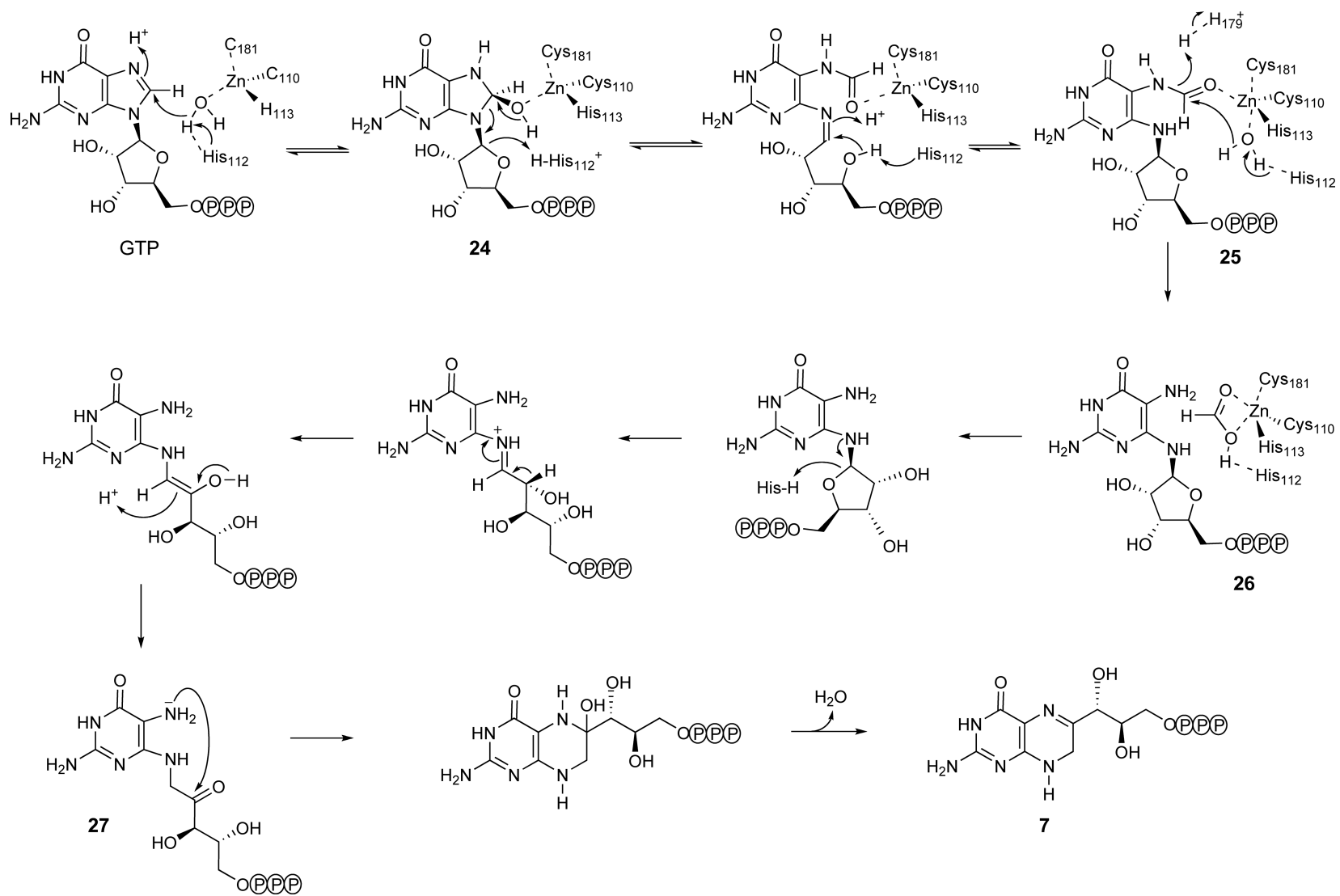
Rather surprisingly, the hydration of the imidazole carbon and the cleavage of the first carbon–nitrogen bond are relatively rapid reactions. The rate-determining step is located late in the reaction trajectory. It is therefore possible to distinguish at least four optical transients under single turnover conditions.<sup>39</sup>

If His179 is replaced in the *E. coli* enzyme, the bond between C-8 and N-9 can be cleaved, but the resulting formamide derivative (**27**) cannot be hydrolyzed. However, the resulting intermediate can be converted into DHNTP (**7**) by wild-type GTP cyclohydrolase I.<sup>40</sup> In fact, **27** can be shown to fulfill the criteria for a kinetically competent intermediate.<sup>39,40</sup>

Comparison of numerous genomes recently revealed the absence of obvious GTP cyclohydrolase orthologues in certain bacteria, which must be assumed to be engaged in the biosynthesis of folic acid. Genomic analysis culminated in the identification of *folE2* genes (as opposed to *folE* genes specifying classical type I cyclohydrolases). The *folE2* gene of *Thermotoga maritima*, *Bacillus subtilis*, and *Acinetobacter baumannii* could complement a *folE* mutant of *E. coli*. The *folE2* genes of *Thermotoga maritima*, *Neisseria gonorrhoeae*, and *B. subtilis* could all be expressed in recombinant *E. coli* strains and their GTP cyclohydrolase I activity was confirmed by *in vitro* assays.<sup>42</sup> A close analysis has revealed a low level of sequence similarity between *folE* and *folE2* sequences. Interestingly, however, the *folE2* genes fail to specify conserved amino acid residues that appear as plausible ligands for Zn<sup>2+</sup> ions. The metal requirement, if any, of *folE2* cyclohydrolases is as yet unknown.

**7.17.2.3.2(ii) Dephosphorylation of dihydroneopterin triphosphate** The conversion of the product of GTP cyclohydrolase I, dihydroneopterin 3'-triphosphate (**7**), into dihydroneopterin (**10**), is believed to proceed in two steps involving (1) the removal of a pyrophosphate unit (**Figure 2**, c) affording dihydroneopterin 3'-phosphate (**9**) and (2) hydrolysis of that monophosphate (**Figure 2**, d) affording dihydroneopterin (**10**).

Early reports on the isolation of a pyrophosphate lyase and a phosphatase activity from *E. coli* had no apparent follow up.<sup>43</sup> Later, it was proposed that the release of pyrophosphate could proceed without enzyme catalysis under the influence of magnesium or calcium ions.<sup>44</sup>



**Figure 5** Hypothetical reaction mechanism for GTP cyclohydrolase I. <sup>34,38-40,41</sup>

More recently, a previously unannotated open reading frame (*ylgG*, renamed *folQ*) that is located in a cluster together with three genes specifying known folate biosynthesis enzymes in *Lactobacillus lactis* was assigned as a hydrolase (Figure 2, c) of the nudix (nucleoside diphosphate linked to x) hydrolase superfamily.<sup>45</sup> Deletion of that gene resulted in accumulation of DHNTP (7) and reduced excretion of folate, but apparently not in outright folate deficiency. Recombinant FolQ protein was found to cleave DHNTP *in vitro*. On this basis, the protein was assigned as a bona fide folate biosynthesis enzyme.<sup>45</sup>

It is unknown how the phosphate residue of dihydroneopterin 3'-monophosphate (9) is removed. This is reminiscent of the fact that the dephosphorylation of 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione 5'-phosphate in the biosynthetic pathway of the benzopteridine derivative, riboflavin (vitamin B<sub>2</sub>), was never assigned to a specific enzyme activity (for review see Fischer and Bacher<sup>46</sup>); surprisingly, very large amounts of the vitamin (at least 15 g l<sup>-1</sup>) can be obtained by fermentation using strains that hyperexpress riboflavin biosynthesis enzymes but not any known phosphatase.<sup>47,48</sup>

**7.17.2.3.2(iii) Dihydroneopterin aldolase (EC 4.1.2.25)** DHNA (Figure 2, e) catalyzes the cleavage of the trihydroxypropyl side chain of its substrate 10 affording hydroxymethyl-dihydropterin (12) and glycol aldehyde (11).<sup>49,50</sup> The enzyme from *E. coli* is a homooctamer consisting of 122 amino acid residues per monomer. The toroid-shaped protein is characterized by D<sub>4</sub> symmetry. The topologically equivalent subunits are all located at interfaces between two subunits. The DHNA fold is topologically similar to 7,8-dihydroneopterin triphosphate epimerase,<sup>51</sup> 6-pyruvoyl tetrahydropterin synthase (PTPS),<sup>52,53</sup> the C-terminal domain of GTP cyclohydrolase I<sup>34</sup> and both domains of urate oxidase (UO1 and UO2).<sup>54</sup> Structures are known from DHNA of *E. coli*, *Mycobacterium tuberculosis*, *Staphylococcus aureus*, *Arabidopsis thaliana* and of a bifunctional DHNA/6-hydroxymethyl-7,8-dihydropterin pyrophospho-kinases from *Streptococcus pneumoniae*.<sup>55–58</sup>

The enzyme requires Mg<sup>2+</sup> but no other cofactors for its activity. Mutation analysis showed that the conserved residue Glu73 in *E. coli* DHNA is important for substrate binding, but its role in catalysis is minor. On the other hand, residues Glu21 and Lys98 of the *E. coli* enzyme are important for both substrate binding and catalysis.<sup>59</sup>

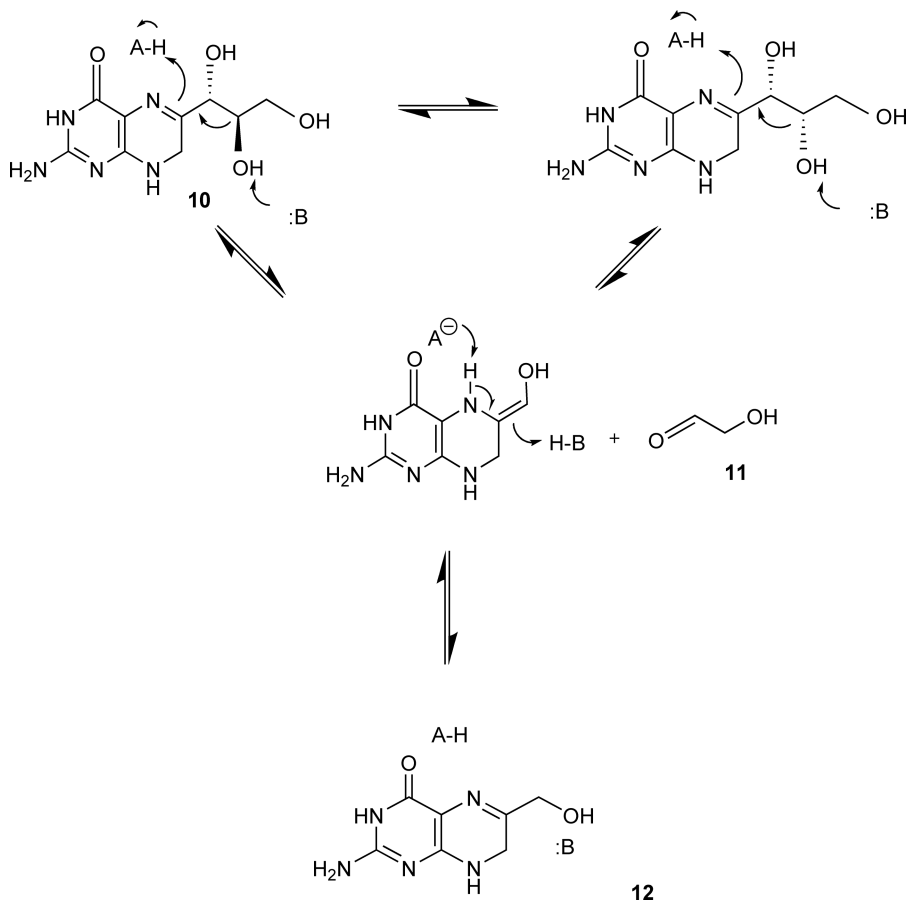
An impressive number of different catalytic mechanisms is employed by aldolases catalyzing the cleavage of variety of substrates. In case of class I aldolases, bond cleavage is preceded by the formation of a Schiff base involving a carbonyl group of the substrate and a conserved lysine residue.<sup>60</sup> Class II aldolases operate without covalent catalysis but use divalent cations, thiamine pyrophosphate, pyridoxal phosphate, or phosphoenolpyruvate for catalysis and are also designated metalloaldolases.<sup>61–63</sup> Dihydroneopterin (10), the substrate of DHNA, can be viewed as an intramolecular Schiff base, and that motif appears to play an equivalent role similar to Schiff base motif involved in classical class I aldolases by facilitating the cleavage of the C-1'/C-2' bond (Figure 6). As a result of the different reaction mechanism and a completely different overall fold as compared to class I aldolases it was proposed that DHNA present a new class of aldolase enzymes, both mechanistically as well as structurally.<sup>55</sup>

The DHNA of *E. coli* is thermally stable up to relatively high temperatures.<sup>49</sup> This property in conjunction with the relatively low-subunit molecular weight and the high degree of symmetry (all subunits are positionally equivalent) made it suitable as a paradigm protein for the exploration of NMR methods aimed at the escalation of the molecular weight range for NMR structure analysis.<sup>64</sup>

Besides the cleavage of C-1'/C-2' bond of the substrate, the DHNA of *E. coli* can also catalyze the epimerization of the 1' position of dihydroneopterin.<sup>65,66</sup> That reaction is believed to proceed through bond cleavage followed by religation, that is, a retroaldol/aldol mechanism.<sup>59</sup>

A paralogous enzyme from *E. coli* catalyzes the epimerization of the C-1' position in DHNTP (Figure 2, b), a biosynthetic precursor of dihydroneopterin in the THF pathway.<sup>49</sup> That enzyme is structurally similar to DHNA and is likewise characterized by D<sub>4</sub> symmetry.<sup>51,65,66</sup> The metabolic role, if any, of its product, dihydromonapterin triphosphate (8), is unknown. However, dihydromonapterin is obviously reduced to the level of tetrahydromonapterin in *E. coli*.<sup>67</sup>

**7.17.2.3.2(iv) 6-Hydroxymethyl-dihydropterin pyrophosphokinase (EC 2.7.6.3)** 6-Hydroxymethyl-dihydropterin pyrophosphokinase (HPPK) (Figure 2, f) catalyzes the transfer of a pyrophosphate moiety to



**Figure 6** Proposed mechanism for the dihydroneopterin aldolase catalyzed reaction.<sup>55,59</sup>

6-hydroxymethyl-dihydropterin (**12**). The enzyme was discovered, like most other enzymes of the folate pathway, in the 1960s.<sup>68</sup> In recent literature, the enzyme is frequently designated HPPK. It has been recently studied in some detail as it is considered as a potential target for novel antifolate-type anti-infective drugs.<sup>69,70</sup> Kinetic analysis indicates a bi-bi mechanism.<sup>71</sup> The three-dimensional structures of several orthologues including *Yersinia pestis*, *S. pneumoniae*, *Saccharomyces cerevisiae*, and *E. coli* have been determined by X-ray structure analysis, in case of the enzyme from *E. coli* to a resolution of 1.25 Å.<sup>58,70,72,73</sup> X-ray data on a variety of complexes with substrates and product analogues indicate conformational changes of the protein that are associated with the reaction trajectory.<sup>74–80</sup>

**7.17.2.3.2(v) Dihydropteroate synthase (EC 2.5.1.15)** Dihydropteroate synthase (**Figure 2, g**) catalyzes the reaction of 6-hydroxymethyl-dihydropterin 1'-diphosphate (**13**) with 4-aminobenzoate (**14**) affording dihydropteroate (**15**) and inorganic pyrophosphate. The enzyme is the target of sulfonamide drugs and, in consequence of its therapeutic relevance, has been studied in considerable detail, although not to the same extent as dihydrofolate reductase (**Figure 2, i**). Databases record more than 700 papers relating to the enzyme.

Crystal structures have been reported for dihydropteroate synthases from several pathogenic organisms (*S. pneumoniae*, *S. aureus*, *M. tuberculosis*, *Bacillus anthracis*).<sup>81–84</sup> The structure represents an eight-stranded  $\alpha\beta$ -barrel. The substrate, 6-hydroxymethyl dihydropterin 6-diphosphate, binds in a deep cleft; the second substrate, 4-aminobenzoate, binds closer to the protein surface.<sup>85</sup>

Sulfonamides act as competitive inhibitors of dihydropteroate synthase.<sup>86</sup> They can also serve as substrates that are converted by the enzyme into sulfonamide analogues of dihydropteroate.<sup>87,88</sup>



These products are unable to serve as substrate for glutamyl transfer and are excreted into the culture medium by *E. coli*. However, products of this type have also been reported more recently to interfere with folate-mediated metabolism. Moreover, the consumption of the dihydropteroate synthase substrate by reaction with the sulfonamide-type substrate analogue may contribute to overall folate-coenzyme depletion.

Mutations in the dihydropteroate synthase gene have been reported in sulfonamide-resistant strains of *Pneumocystis jirovecii*, the causal agent of *Pneumocystis jirovecii* pneumonia (PCP).<sup>89–92</sup> Similar mutations have been observed in *S. pneumoniae* and *Plasmodium falciparum*.<sup>93–95</sup>

**7.17.2.3.2(vi) Dihydrofolate synthetase (EC 6.3.2.12)** Dihydrofolate synthetase (**Figure 2**, h) catalyzes the condensation of dihydropteroate (**15**) with glutamate in a reaction requiring ATP that is converted into ADP and inorganic phosphate yielding dihydrofolate (**16**). By comparison with several other enzymes of the pathway, dihydrofolate synthetase has been less intensively studied.

Certain bacteria including *E. coli* form bifunctional enzymes with dihydrofolate and folylpolyglutamate synthetase (**Figure 2**, j) activity.<sup>96,97</sup> Both the conversion of dihydropteroate into dihydrofolate and the subsequent attachment of additional glutamate moieties by the bifunctional *E. coli* enzyme have been shown to involve the transient formation of acyl phosphate intermediates.<sup>98</sup>

The formation of folate and the condensation with additional glutamate units appear to proceed at a single active site.<sup>99</sup> The structure of the bifunctional protein specified by the *E. coli folC* in complex with ADP has been determined by X-ray crystallography.<sup>100</sup> The enzyme has a single catalytic site and a single ATP-binding site. On the other hand, it is believed to have distinct binding sites for dihydropteroate and THF.<sup>101</sup> The binding of glutamate as well as other details of the reaction mechanism are unknown. A ter-ter mechanism has been determined by kinetic analysis.<sup>102</sup>

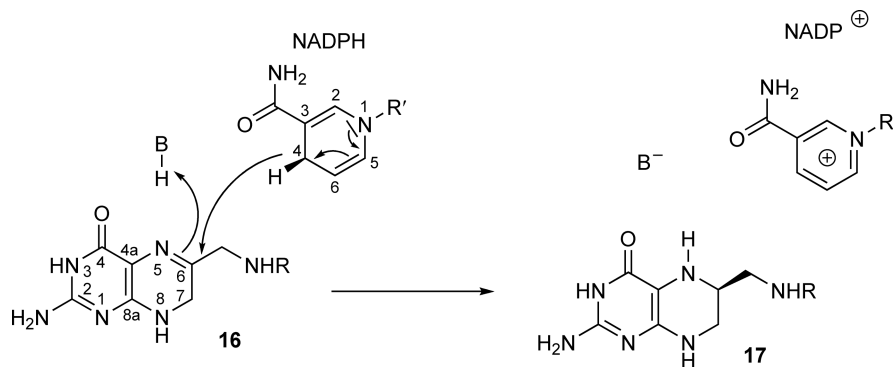
**7.17.2.3.2(vii) Dihydrofolate reductase (EC 1.5.1.3)** Dihydrofolate reductase (**Figure 2**, i) commands a special position in the dihydrofolate pathway since it serves several different roles. (1) In animals, plants, and numerous eubacteria, the enzyme is required for the recycling of dihydrofolate that is generated by the catalytic action of thymidylate synthase, which uses methylenetetrahydrofolate as donor of a C1 unit and as a donor of reduction equivalents.<sup>103</sup> (2) In organisms that biosynthesize THF *de novo*, with the exception of certain bacteria that use flavocoenzyme-dependent thymidylate synthases in conjunction with flavocoenzyme-dependent dihydropterin reductase (see Section 7.17.2.3.2(viii)), dihydrofolate reductase catalyzes the final step of the *de novo* biosynthesis affording THF (**17**) from dihydrofolate (**16**). (3) In folic acid auxotrophs including animals and certain bacteria, the enzyme is used for converting nutritional folates such as dihydrofolate and folate into the coenzyme form, THF.

Surprisingly, however, recent evidence showed that dihydrofolate reductase is not a universal enzyme. About one-third of the sequenced eubacteria use a flavin-dependent thymidylate synthetase (FDTS) (ThyX protein, EC 2.1.1.148) that does not require the supply of reduction equivalents by the THF-type cofactor, they are NADPH oxidases that use flavin adenine nucleotide (FAD) to mediate hydride transfer; the genomes of these respective organisms do not specify orthologues of dihydrofolate reductase (see Section 7.17.2.3.2(viii)).<sup>104–108</sup>

Dihydrofolate is a target for antibacterial and antiprotozoan drugs. Moreover, human dihydrofolate reductase is a target for cytostatic and immunosuppressive agents. This dual drug target aspect was one of the reasons why dihydrofolate reductase became one of the most intensely studied proteins. A database search retrieved over 15 000 articles and over 400 reviews relating to the enzyme, and more than 100 crystal structures are in the public domain. The structure and function of the protein have been addressed by a very wide range of experimental techniques including X-ray and NMR structure analysis, steady-state and presteady-state kinetics, and single molecule spectroscopy as well as computer simulations using classical and quantum mechanics approaches.

The dihydrofolate reductases of animals and of certain eubacteria are monomeric proteins with molecular weights around 20 kDa. Other eubacteria, ciliate, protozoa, plants, and certain viruses specify bifunctional proteins (DHFR-TS) characterized by dihydrofolate reductase and thymidylate synthetase domains.

The reaction catalyzed by dihydrofolate reductase appears superficially simple and involves the transfer of a hydride ion from the 4 position of NADH to the 6 position of the substrate, dihydrofolate (**Figure 7**).<sup>110,109</sup>



**Figure 7** The reaction catalyzed by DHFR.<sup>109</sup>

Mammalian dihydrofolate reductases can also catalyze the transfer of a hydride ion to the C-7 position of folate in a reaction affording dihydrofolate, thus enabling the utilization of the fully oxidized vitamin from nutritional sources.<sup>111</sup>

Despite its apparent simplicity, on the basis of an extraordinarily large amount of experimental data, the reaction catalyzed by dihydrofolate reductase has been dissected into a complex series of reaction steps. A comparison between the *E. coli* enzyme and an R-plasmid-coded dihydrofolate reductase involved in trimethoprim resistance indicates that the chromosomal enzyme has evolved to a virtually perfect catalytic process.<sup>112,113</sup> The experimentally observed velocities of the hydride transfer reaction has been explained by tunneling;<sup>114</sup> in fact, dihydrofolate reductase is a paradigm enzyme for the study of tunneling processes in biological systems.

A systematic search for dihydrofolate analogues in complete genomes has failed to locate plausible orthologues of the classical dihydrofolate reductase family comprising the enzyme specified by the *folA* gene of *E. coli*. A different type of dihydrofolate reductase is specified in *E. coli* by the *folM* gene; this protein is a member of the short-chain dehydrogenase reductase family.<sup>115</sup> An *E. coli* plasmid has been shown to specify a third dihydrofolate reductase type that has been interpreted as a primitive enzyme that has not undergone extensive evolution.<sup>116</sup>

**7.17.2.3.2(viii) Dihydropteroylglutamate reductase** Thymidylate synthetase (ThyA protein, EC 2.1.1.45) catalyzes the conversion of uridine 5'-phosphate into thymidine 5'-phosphate. In animals, plants, and many bacteria, methylenetetrahydrofolate serves as a donor of a formaldehyde moiety as well as of the required reduction equivalents. The cofactor is oxidized to the level of dihydrofolate and needs to be reduced to the level of THF by the action of dihydrofolate reductase. This causes a requirement for a relatively high dihydrofolate reductase activity.

Relatively recently, it was found that certain bacteria including *Cblamydia trachomatis* use thymidylate synthases requiring flavocoenzymes (ThyX protein, EC 2.1.1.148) as donors of reduction equivalents.<sup>108</sup> Although both enzymes, ThyA and ThyX, catalyze the formation of thymidylate *in vitro*, their reductive mechanisms are dramatically different. ThyX catalysis results in the formation of THF, and not dihydrofolate, as the product of the methylation reaction.<sup>104</sup>

These organisms are typically devoid of dihydrofolate reductase.<sup>107</sup> However, a recent study showed that the *folP* gene of *Helicobacter pylori* specifies a bifunctional dihydropteroylglutamate synthase/dihydropteroylglutamate reductase that can be supplied with reduction equivalents by soluble FMNH<sub>2</sub> or FADH<sub>2</sub>, in the absence of a dihydrofolate reductase.<sup>117</sup> Whereas classical dihydrofolate reductases serve the *de novo* biosynthesis and/or the recycling of THF, dihydrofolate recycling function appears to be unnecessary in organisms using a flavocoenzyme-dependent thymidylate synthase. Thus, two folate biosynthesis pathways differing in the mode of conversion of dihydropteroylglutamate to THF must now be considered. Further details of the folate biosynthesis pathway through the reduction of dihydropteroylglutamate by the recently discovered dihydropteroylglutamate synthase/dihydropteroylglutamate reductase remain to be established.

**7.17.2.3.2(ix) Folylpolyglutamate synthase (EC 6.3.2.17)** Folylpolyglutamate synthase (**Figure 2**, j) catalyzes the attachment of 1–7 glutamate residues to THF through pseudopeptide linkages involving the  $\gamma$ -carboxylic group of glutamate;<sup>118,119</sup> the reaction requires ATP. The enzyme appears to be universal in all organisms, with the exception of methanogenic bacteria that use tetrahydromethanopterin instead of THF (**17**). The substrate, THF, can be provided by *de novo* synthesis (in plants and many microorganisms) or from nutritional folate and folate derivatives. The polyglutamate tails of dihydrofolate cofactor serve the interaction with the cognate apoenzymes as well as their retention inside cells.

The reaction mechanism of folylpolyglutamate synthetase (**Figure 2**, j) is virtually identical with that of dihydrofolate synthetase (**Figure 2**, h). Specifically, reaction of the respective substrate with ATP affords an acylphosphate intermediate (**28**, **29**) that can subsequently form a pseudopeptide bond with an additional glutamate residue.<sup>98</sup> As described above, certain microorganisms including *E. coli* express enzymes that can catalyze the conversion of dihydropteroate (**15**) into dihydrofolate (**16**) (**Figure 8(a)**) as well as the conversion of THF (**17**) into polyglutamyl derivatives (**18**) (**Figure 8(b)**).<sup>120</sup> The sequences of folylpolyglutamate synthases of folate-auxotrophs requiring exogenous folates (i.e., animals and folate-auxotrophic microorganisms) are similar to those of the bifunctional FolC protein of *E. coli*.

The three-dimensional structure of folylpolyglutamate synthase of *Lactobacillus casei* and *E. coli* have been determined by X-ray crystallography.<sup>100,121</sup> The catalytically active conformation of the protein is obtained by bonding of a folate-type substrate; binding of ATP is not, *per se*, sufficient to generate that conformational state.

Folylpolyglutamate synthetases, including the bifunctional FolC proteins of bacterial origin, are members of a superfamily including the Mur proteins involved in bacterial cell wall biosynthesis.<sup>122</sup>

**7.17.2.3.2(x) Cellular topology of the pathway enzymes** In plants, three different cellular compartments participate in the biosynthesis of THF.<sup>123,124</sup> The biosynthesis of the pteridine precursor proceeds in the cytoplasm. The biosynthesis of 4-aminobenzoate (**14**) from the common shikimate pathway intermediate, chorismate (**19**), proceeds in chloroplasts. Both structural components are then imported into the mitochondria, where they are assembled to THF (**17**). THF can be exported to the cytoplasmic compartment or into chloroplasts where it can also be converted to the polyglutamyl forms by cytoplasmic or plastid enzymes. Polyglutamyl forms can be imported into the vacuole for storage.

Certain protozoa including the *Plasmodium* species causing malaria and *Toxoplasma gondii* causing toxoplasmosis synthesize THF *de novo*.<sup>125</sup> These apicomplexan pathogens are characterized by an organelle designated apicoplast that is believed to be a homologue of plant chloroplasts. The apicoplast is involved in the biosynthesis of isoprenoids and of THF. *Plasmodium* appears to operate the shikimate pathway for the singular purpose to provide 4-aminobenzoate (**14**) as a precursor for the biosynthesis of folate-type coenzymes, whereas aromatic amino acids are acquired from the host cell.<sup>126</sup>

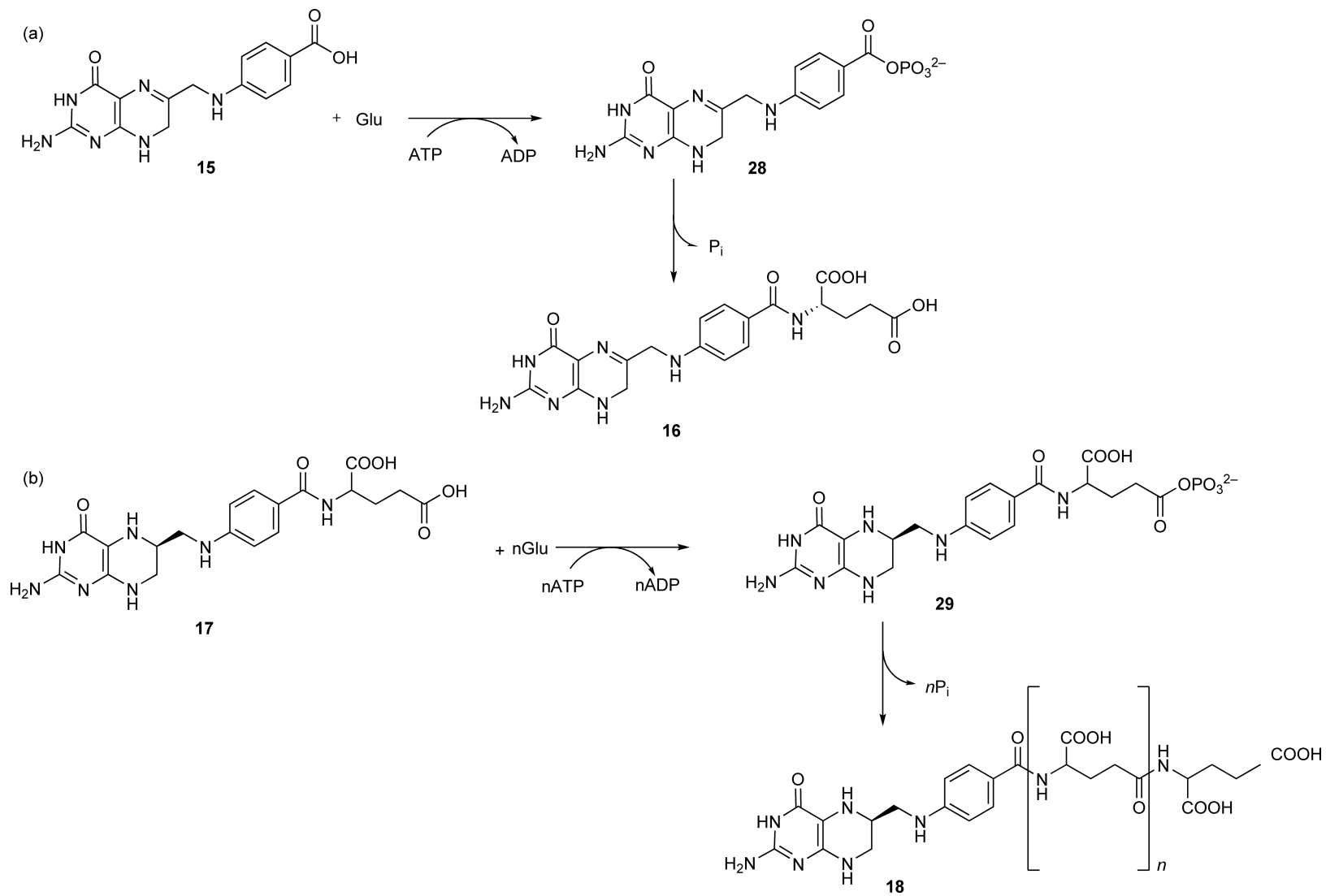
Notably, a DHNA has not been identified in *Plasmodium* despite considerable efforts. The most likely explanation is that the cognate gene cannot be identified on the basis of similarity to known DHNA genes and may have originated independently.

#### 7.17.2.4 Tetrahydrofolate Enzymes as Drug Targets

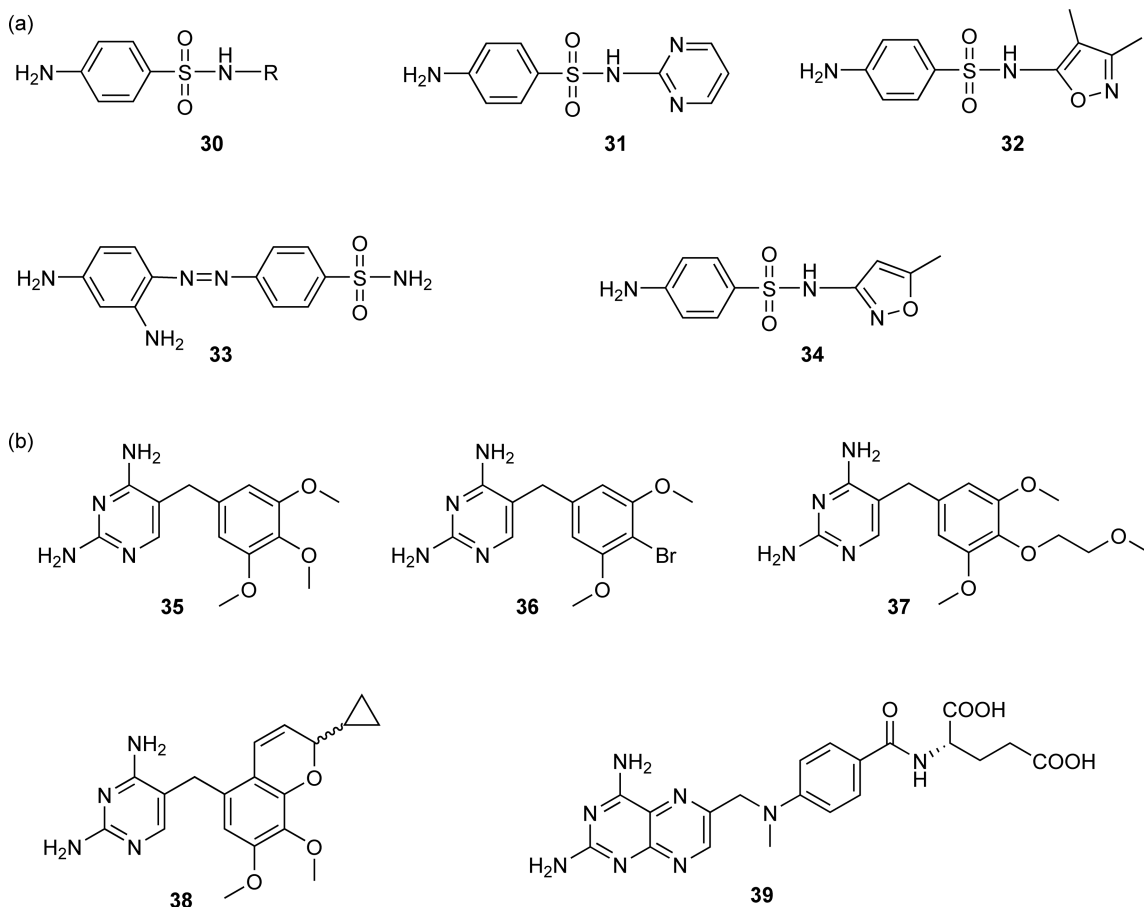
THFs or structural analogues thereof (methanopterin) are universally required for the maintenance of C1 metabolism. C1 synthones play an indispensable role for the biosynthesis of nucleic acid components and of methionine. Even when eubacteria are supplied with exogenous sources of all metabolites whose biosynthesis requires folate-type cofactors, they require the coenzyme for the initiation of protein biosynthesis.

Enzymes of the folate biosynthesis pathway are important drug targets for the treatment of infections with bacteria and of certain malignancies. In fact, sulfonamide (**Figure 9**) drugs that interfere with the formation of the THF precursor, dihydropteroate, were discovered and widely used therapeutically more than a decade prior to the structure elucidation of folic acid.<sup>87,88,128–131</sup>

Early in the twentieth century, Paul Ehrlich introduced the arseno dye salvarsan (Arsphenamine) for the treatment of syphilis. Salvarsan was the first synthetic anti-infective drug and thus represented a major breakthrough despite its toxicity.



**Figure 8** Reaction mechanism of (a) dihydrofolate synthetase (DHFS)<sup>98</sup> and (b) folylpolyglutamate synthetase (FPGS).<sup>98</sup>



**Figure 9** (a) Dihydropteroate synthase inhibitors as antibacterial agents. **30**, Sulfonamide core structure; **31**, sulfadiazine; **32**, sulfisoxazole; **33**, prontosil; **34**, sulfamethoxazole (SMX)<sup>127</sup> and (b) inhibitors of dihydrofolate reductase. **35**, Trimethoprim (TMP); **36**, brodimoprim; **37**, tetroxoprim; **38**, iclaprim; **39**, methotrexate.<sup>127</sup>

In the 1930s, by extension of Ehrlich's concept, Domagk and coworkers studied the influence of various dyestuffs on bacteria. The azo dye Prontosil (**33**) showed antibacterial activity *in vitro* and was rapidly introduced for the treatment of patients. Subsequent work showed that it was the sulfanilamide (more generally designated sulfonamide) motif of Prontosil and not the azo motif that was crucial for antibacterial activity, and a wide variety of other sulfonamides were developed. These could be used for a variety of bacteria and of certain protozoa including the *Plasmodium* spp. that cause malaria. When the structure of folic acid was elucidated in the 1940s, it was found that the antibacterial activity of sulfonamides could be antagonized by 4-aminobenzoate (**14**) that was a structural component of folic acid connected to a glutamate residue through a carboxamide motif.

Subsequent work established that the sulfonamides inhibit the biosynthesis of folate at the level of dihydropteroate synthase (Figure 2, g). More specifically, the compounds can be used as substrates of the enzyme, but the resulting conjugates cannot serve as substrates for the subsequent enzymatic steps resulting in the addition of an oligoglutamate tail; the products generated by dihydropteroate synthase with sulfa drugs as substrates can therefore not serve the THF requirement of the parasite.

In order to counteract resistance development and to improve antibacterial activity, sulfonamides are typically combined with an inhibitor of dihydrofolate reductase (Figure 2, i). Dihydrofolate reductase is a core enzyme of human central intermediary metabolism. Although the bacterial and mammalian dihydrofolate reductases are orthologues with considerable sequence similarity, trimethoprim (**35**) has sufficient selectivity to enable the selective inhibition of the parasite enzyme.

Methotrexate binds human dihydrofolate reductase with very high affinity. Despite the obvious similarity between dihydrofolate (16) and methotrexate (39), they are bound with different topology by the enzyme (Figure 9). Inhibition of dihydrofolate is conducive to a shortage of purine derivatives for DNA synthesis that results in cytotoxicity. The cytotoxicity of methotrexate is the basis for its continued use for the chemotherapy of hematologic malignancies as well as for its use as an immunosuppressive agent.

### **7.17.3 Tetrahydrobiopterin Metabolism and Function**

#### **7.17.3.1 An Introduction to Tetrahydrobiopterin Biosynthesis**

BH<sub>4</sub> is essential for diverse processes and is ubiquitously present in all tissues of higher organisms. It is well established as an essential cofactor for various enzyme activities to activate dioxygen, and for less-defined functions at the cellular level. The latter function of BH<sub>4</sub> seems to have a dual nature, depending on the concentration of the cofactor and the cell type, between proliferative activity and trigger for apoptosis. Since BH<sub>4</sub> (and related tetrahydropterins) can form radicals, it can act as a generator or scavenger of reactive oxygen species in cells. More characteristics have arisen based on recent observations, including chaperon function reported at least for the hepatic phenylalanine hydroxylase. The BH<sub>4</sub> cofactor is endogenously synthesized and a (genetic) deficiency in the biosynthesis or regeneration leads to neurological abnormalities, including DSD or severe monoamine neurotransmitter depletion. However, also under normal BH<sub>4</sub> concentrations its availability becomes limiting in various pathological situations involving endothelial dysfunction, for example, in diabetes or coronary heart diseases.

The following chapter of BH<sub>4</sub> metabolism is an update and extension of a previous review article,<sup>6</sup> and covers mainly structure and function of the ‘mammalian’ enzymes responsible for the synthesis and regeneration of BH<sub>4</sub> in order to understand the pathophysiology and to eventually help patients with different defects involving directly or indirectly the BH<sub>4</sub> cofactor. Comprehensive reviews covering BH<sub>4</sub> biology also in nonmammalian species, as well as regulation of biosynthesis and pharmacological effects can be found in the two articles published by Werner-Felmayer *et al.*<sup>7</sup> and Shi *et al.*<sup>132</sup>

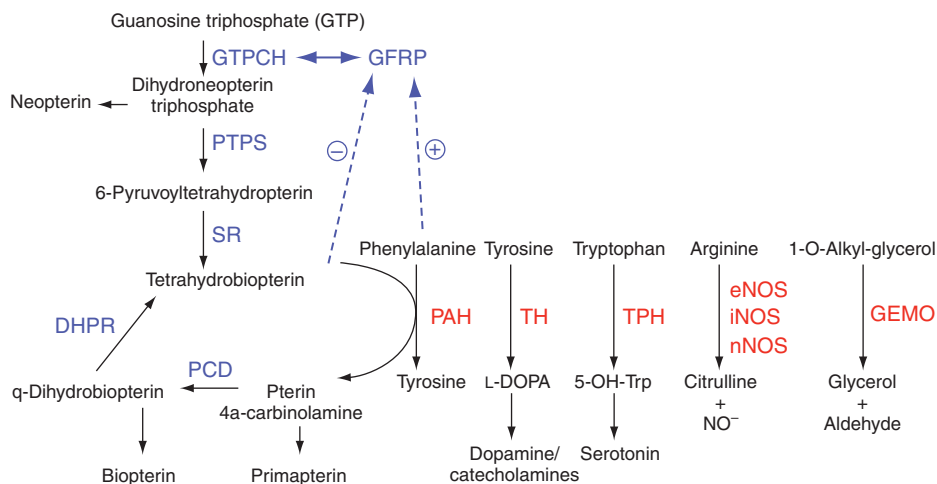
#### **7.17.3.2 Functions of Tetrahydrobiopterin**

##### **7.17.3.2.1 Cofactor-dependent enzyme systems**

The best investigated functions of BH<sub>4</sub>, which is endogenously synthesized by the *de novo* pathway (see Figure 11) is that of its action as a natural cofactor of the aromatic amino acid hydroxylases, phenylalanine-4-hydroxylase (EC 1.14.16.2; PAH), tyrosine-3-hydroxylase (EC 1.14.16.3; TH), and tryptophan-5-hydroxylase type 1 and type 2 (EC 1.14.16.4; TPH1, TPH2),<sup>133,134</sup> as well as of all three forms of nitric oxide synthase, the endothelial, inducible, and neuronal types of NOS (EC 1.14.13.39; e, I, nNOS) (see Figure 10; for a review see Kappock and Caradonna<sup>135</sup>). In addition, BH<sub>4</sub> is required by the enzyme glyceryl-ether monooxygenase (EC 1.14.16.5; GEMO) for hydroxylation of the  $\alpha$ -carbon atom of the lipid carbon chain of glyceryl ether to form  $\alpha$ -hydroxyalkyl glycerol.<sup>136,137</sup> The significance of GEMO in humans and rodents has been well documented; however, there is no record about the consequences of BH<sub>4</sub> deficiency on the alkyl ethers metabolism.

Enzymatic reactions of aromatic amino acid hydroxylases have been intensively studied which revealed that they have many features in common.<sup>138–141</sup> They all have a strict requirement for dioxygen, iron, and a tetrahydropterin cofactor such as BH<sub>4</sub>. The oxidation product of BH<sub>4</sub> is released after each catalytic turnover and regenerated by the enzymes pterin-4 $\alpha$ -carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor-1 $\alpha$  (PCD/DCoH) and dihydropteridine reductase (DHPR) (see Figure 11). The oxidation of BH<sub>4</sub> involves the formation of the 4 $\alpha$ -carbinolamine intermediate (4 $\alpha$ -OH-BH<sub>4</sub> in Figure 12), and this has been shown to be formed in reactions of both PAH and TH.<sup>143–146</sup> Studies on PAH found that a stoichiometric amount of BH<sub>4</sub> can be oxidized in the presence of oxygen, and this yields the reduced enzyme. It has been proposed that this reductive activation of PAH occurs at the redox site and that the enzyme's iron is a part of this redox site. Its reduction from Fe(III) to Fe(II) has been linked to the formation of active PAH. Two electrons from BH<sub>4</sub> are required to reduce the enzyme; one is transferred to Fe(III) and the second apparently to oxygen.<sup>147</sup> A mechanism of the catalytic cycle of nonheme iron-dependent PAH, that hydroxylases L-Phe to L-Tyr using BH<sub>4</sub> and molecular O<sub>2</sub> based on structural data was proposed by several authors and is shown in





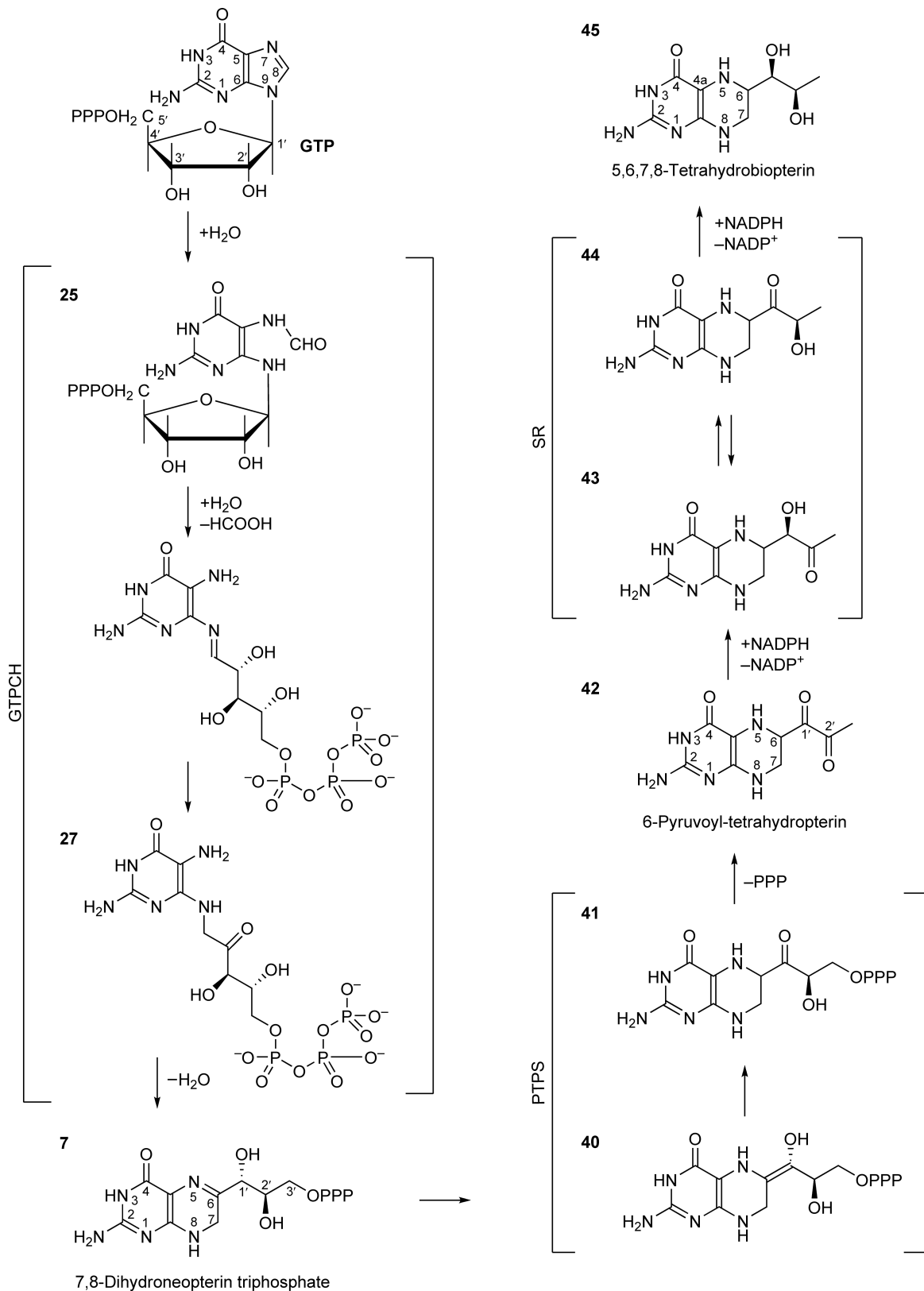
**Figure 10** Overview of BH<sub>4</sub> biosynthesis and regeneration, including cofactor-dependent metabolic pathways. BH<sub>4</sub>-metabolizing enzymes are in blue, cofactor-dependent enzymes are in red (see also text for details and abbreviations of enzymes).

**Figure 12.**<sup>142,148,149</sup> The function of BH<sub>4</sub> for NOS appears to be different, as the cofactor is a one-electron donor and seems not to dissociate from NOS during turnover (see below).

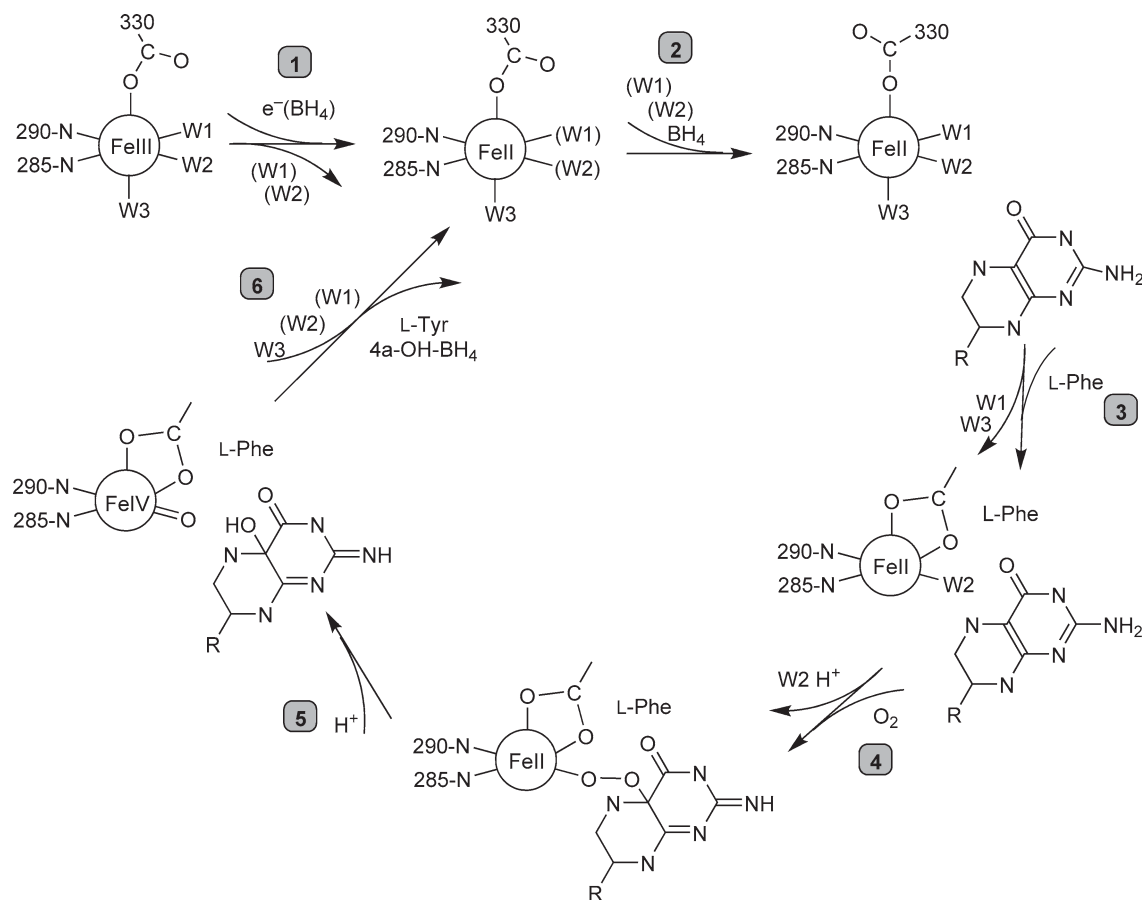
Phenylalanine and BH<sub>4</sub> are the major regulators of PAH.<sup>147</sup> Although phenylalanine is a positive allosteric effector (activator) that converts inactive enzyme to catalytically competent (activated) enzyme, BH<sub>4</sub> is a negative effector that competes with phenylalanine activation to form a dead end complex PAH·BH<sub>4</sub>.<sup>150</sup> Another role of BH<sub>4</sub> on the PAH is the chemical chaperon effect preventing protein misfolding and inactivation during dimer/tetramer expression, and protection from proteolytic cleavage.<sup>151,152</sup> Thus BH<sub>4</sub> plays a central regulatory role in the phenylalanine hydroxylating system. The only other known BH<sub>4</sub>-requiring enzymes in liver, GEMO and NOS, are present in relatively low amounts, and PAH (subunit) and BH<sub>4</sub> concentrations in liver are approximately equal (PAH ~9 μmol l<sup>-1</sup> and BH<sub>4</sub> 5–10 μmol l<sup>-1</sup>).<sup>1135,150–154</sup> Nevertheless, considering the *K<sub>m</sub>* for the cofactor in the PAH (and GEMO) reaction, which was estimated to be 25–30 μmol l<sup>-1</sup>,<sup>153,155</sup> BH<sub>4</sub> is limiting and subsaturating at least in the liver<sup>152,156</sup> (with consequences for BH<sub>4</sub>-responsive PKU; see below). There is no evidence that TH and TPH are regulated by substrate-activated mechanisms similar to those that regulate PAH. All aromatic amino acid hydroxylases are inhibited by catecholamines, but only the inhibition of human TH is competitive with respect to the BH<sub>4</sub> cofactor, and it has been shown that the cofactor can directly displace dopamine from the enzyme active site.<sup>157</sup>

BH<sub>4</sub> is also utilized as a redox cofactor by all NOS isoforms, inducible NOS (iNOS or NOSII), neuronal NOS (nNOS or NOSI), and endothelial NOS (eNOS or NOSIII). All these NOSs share 50–60% sequence homology, and crystal structure determination for several NOS proteins revealed that they have similar homodimeric structures. Each monomer has an N-terminal oxygenase domain with binding sites for arginine, heme, and BH<sub>4</sub>, and a C-terminal reductase domain with binding sites for the cofactors NADPH, FAD, FMN, and Ca<sup>2+</sup>/calmodulin (for recent reviews see Wei *et al.*<sup>158</sup> Stuehr *et al.*<sup>159</sup>). Two BH<sub>4</sub> molecules bind at the dimer interface and are in the vicinity of the arginine substrate binding sites and the heme moieties. NOSs catalyze two sequential monooxygenase reactions, involving NADPH- and O<sub>2</sub>-dependent hydroxylation of L-arginine through an *N*-hydroxy-arginine intermediate to nitric oxide (NO) and L-citrulline. BH<sub>4</sub> is required for catalysis but also has structural functions, including dimer stabilization or promoting its formation, protection against proteolysis, and increased arginine binding. As mentioned before, for NOSs BH<sub>4</sub> is a one-electron donor, and the BH<sub>4</sub> radical (BH<sub>4</sub><sup>·</sup>) remains bound in NOS and is subsequently reduced back to BH<sub>4</sub> by an electron provided by the NOS reductase domain.

In contrast to PAH, requirement for the BH<sub>4</sub> cofactor is much lower for the NOS enzyme. The *K<sub>m</sub>* values for BH<sub>4</sub> for PAH and NOS are 25–30 μmol l<sup>-1</sup> and 0.2–0.3 μmol l<sup>-1</sup>, respectively.<sup>160</sup> Pastor *et al.*<sup>161</sup> questioned the importance of BH<sub>4</sub> competition between these two hepatic enzymes. They showed that basal BH<sub>4</sub> synthesis appears to be adequate to support iNOS activity, whereas BH<sub>4</sub> is increased to support PAH activity. Phenylalanine markedly increased BH<sub>4</sub> biosynthesis (through GTP cyclohydrolase feedback regulatory protein, GFRP), whereas arginine had no effect. The



**Figure 11** Reaction mechanism for the *de novo* pathway for BH<sub>4</sub> biosynthesis (see text for details).



**Figure 12** Reaction pathway of PAH. The catalytic cycle consists of a reduction of the iron center [1], which converts the six-coordinate, high-spin Fe(III) into a 'four-coordinate,' high-spin Fe(II); reversible binding of  $\text{BH}_4$ , which results in a six-coordinate Fe(II) and a change in coordination of Glu330 [2]; reversible binding of L-Phe to give a highly distorted square pyramidal five-coordinate Fe(II), with a bidentate coordination of Glu330 [3]; reversible binding of dioxygen to give a five-coordinate Fe(II)- $\text{O}_2$  intermediate followed by the formation of the putative Fe(II)-O-O- $\text{BH}_4$  intermediate [4]; heterolytic cleavage of the oxygen-oxygen bond and the formation of a 4a-OH- $\text{BH}_4$  and an oxyferryl species [5], and products are finally released [6]. Reproduced with permission from O. A. Andersen; T. Flatmark; E. Hough, *J. Mol. Biol.* **2002**, 320, 1095–1108.

$K_m$ -( $\text{BH}_4$ )-values for the human brain TH range from 0.05 to 0.1  $\text{mmol l}^{-1}$ ,<sup>135</sup> and for the rat brain TPH 30–60  $\mu\text{mol l}^{-1}$ .<sup>162</sup> Nevertheless,  $\text{BH}_4$  (bio) availability has a critical role when the cofactor is also limiting for NOS activity, as the NOS no longer produces NO but instead becomes uncoupled and generates the superoxide. By the uptake of an additional electron, superoxide can form  $\text{H}_2\text{O}_2$  or, by reaction with NO, form the cell or neurotoxic peroxynitrate. Such a mechanism is best documented for the eNOS and endothelial dysfunction in insulin-resistant diabetes, but it may also play an important role in patients with SR or DHPR deficiencies or induce other pathophysiological pathways<sup>163–168</sup> (for reviews see Shi *et al.*,<sup>132</sup> Alp and Channon,<sup>169</sup> and Blau *et al.*<sup>170</sup>).

#### 7.17.3.2.2 Cellular and systemic functions of $\text{BH}_4$

In general,  $\text{BH}_4$  is labile in solution at physiological pH and can readily react with  $\text{O}_2$  to produce free radicals thus generating oxidative species in cells, including superoxide,  $\text{H}_2\text{O}_2$ , and peroxynitrite.<sup>171</sup> On the other hand,  $\text{BH}_4$  can also protect cells against oxidative damage by scavenging radicals.<sup>158</sup> Spontaneous oxygen reactivity, chaperon-like effect on protein synthesis, superoxide formation by NOS-uncoupling, and other functions of  $\text{BH}_4$  all hint toward the importance of the concentration of  $\text{BH}_4$  as a critical factor that determines the beneficial or adverse effects *in vivo*. In this context, toxicology of  $\text{BH}_4$  is also an issue to consider when experimental treatment studies with cells or animals are conducted (for an overview see Blau and Erlandsen<sup>172</sup>). One of the earliest discovered cellular functions of  $\text{BH}_4$  was that as a growth factor of *Cribidia fasciculata* and this was

initially used to measure bipterins in different body fluids and tissues. More recent observations suggested the proliferative activity of BH<sub>4</sub> in hemopoietic cells.<sup>173,174</sup> Exogenous BH<sub>4</sub> was found to stimulate DNA synthesis and induce proliferation of some mouse erythroleukemia clonal cell lines. BH<sub>4</sub> and sepiapterin also enhanced the proliferation of SV40-transformed human fibroblasts, rat C6 glioma cells, and PC12 cells. Subsequently, it has been shown that epidermal factor (EGF), nerve growth factor (NGF), and insulin-like growth factor-1 (IGF-1) increased the proliferation of rat pheochromocytoma-12 (PC12) cells through obligatory elevation of intracellular BH<sub>4</sub>.<sup>175,176</sup> Besides its proliferative activity, BH<sub>4</sub> was also suggested to act as a self-protecting factor for NO toxicity with generation of superoxide in NO-producing neurons.<sup>177</sup> Indeed, strong scavenging activity of BH<sub>4</sub> against superoxide anion radicals was shown in both xanthine/xanthine oxidase (XO) and rat macrophages/phorbol myristate acetate radical-generating systems, and the authors suggested that BH<sub>4</sub> might be useful in the treatment of various diseases whose pathogenesis is actively oxygen related.<sup>178</sup> In another series of experiments, Shimizu *et al.*<sup>179</sup> demonstrated that *S*-nitroso-*N*-acetyl-DL-penicillamine (NO donor)-induced endothelial cell death can be prevented by increasing cellular levels of BH<sub>4</sub>. This finding was an early observation that suggested the cytotoxicity of NO involving H<sub>2</sub>O<sub>2</sub> production, and that scavenging of H<sub>2</sub>O<sub>2</sub> by BH<sub>4</sub> may be at least one of the mechanisms by which BH<sub>4</sub> reduces NO-induced endothelial cell death. In contrast, Cho *et al.*<sup>180</sup> hypothesized that ischemia increases the intracellular BH<sub>4</sub> levels and that the increased BH<sub>4</sub> level plays a critical role in selective neuronal injury through NOS activation. Using a selective inhibitor of GTPCH in animals exposed to transient forebrain ischemia they demonstrated a marked reduction of BH<sub>4</sub> levels, NADPH-diaphorase activity, and caspase-3 gene expression in the CA1 hippocampus. Moreover, delayed neuronal injury in the CA1 hippocampal region was significantly attenuated by the GTPCH inhibitor. These data, in contrast to those by Shimizu *et al.*, suggested that a blockade of BH<sub>4</sub> biosynthesis may provide novel strategies for neuroprotection.

Modulation of BH<sub>4</sub>-dependent gene expression was put forward based on several observations. In the GTPCH-deficient (*bph-1*) mouse, upregulation of liver PAH and brain TH gene expression upon BH<sub>4</sub> treatment was reported; however, no such change was found in the wild-type control mice.<sup>181</sup> In smooth muscle cells, posttranscriptional stabilization of iNOS mRNA was observed indirectly upon varying endogenous BH<sub>4</sub> levels.<sup>182</sup> An effect of BH<sub>4</sub> supplementation on PAH gene expression was also discussed in BH<sub>4</sub>-responsive PKU (see below);<sup>183</sup> however, gene expression or mRNA stability for PAH was not changed, at least not in transgenic mice with different cofactor concentrations present in the liver.<sup>152,184</sup>

In addition to its other functions BH<sub>4</sub> enhances the release of dopamine and serotonin in the rat striatum when administered locally through the dialysis membrane.<sup>185,186</sup> The enhancement of dopamine release persisted even when dopamine biosynthesis or dopamine reuptake was completely blocked, but it was abolished when blockers of voltage-dependent Na<sup>+</sup> or Ca<sup>2+</sup> channels were administered with BH<sub>4</sub>. Further experiments using selective inhibitors of tyrosine, TH, and NOS demonstrated that BH<sub>4</sub> stimulates dopamine release directly, independent of its cofactor action on TH and NOS, by acting from the outside of neurons.<sup>187</sup> The exact mechanism is not entirely clear but it has been shown that arginine also induces a concentration-dependent increase of dopamine release in the superfusate of rat striatum slices, and that it is dependent on the presence of BH<sub>4</sub>.<sup>188</sup>

### 7.17.3.3 Biosynthesis of BH<sub>4</sub>

The *de novo* biosynthesis of the BH<sub>4</sub> cofactor is carried out by three classical enzymes as described below. In addition, a hypothetical alternative way to circumvent the last enzymatic step by involving other reductases is presented (from 41 to 45 in Figure 11). Part of this alternative pathway is also called 'salvage' pathway, which feeds BH<sub>4</sub> from dihydrobiopterin through the NADPH-dependent dihydrofolate reductase. As will be discussed later, this alternative pathway seems to compensate for a genetic defect in the last biosynthetic step at least in peripheral tissues.

#### 7.17.3.3.1 Reaction mechanism of the *de novo* pathway

BH<sub>4</sub> biosynthesis proceeds in the *de novo* pathway in a Mg<sup>2+</sup>-, Zn<sup>2+</sup>-, and NADPH-dependent reaction from GTP through the two intermediates, 7,8-dihydroneopterin triphosphate (H<sub>2</sub>NTP, 7) and 6-pyruvoyl-5,6,7,8-tetrahydropterin (PTP, 42), which have been isolated although they are rather unstable. The three enzymes

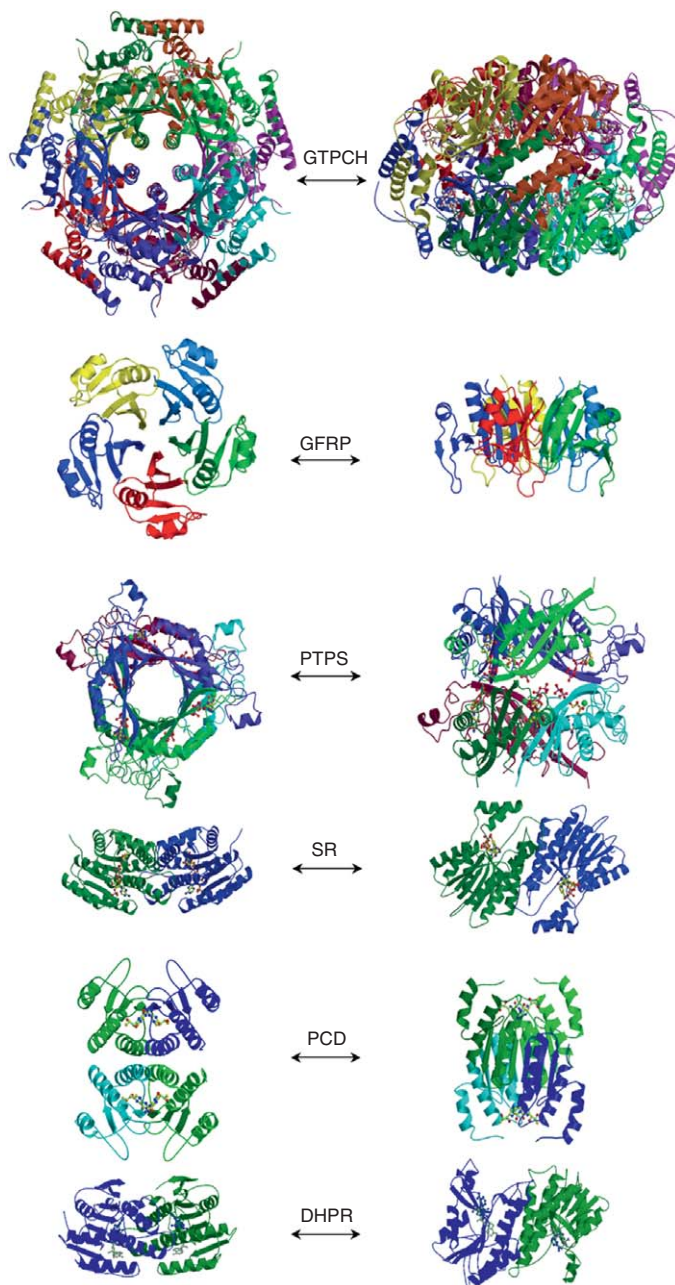
GTPCH, PTPS, and SR are required and sufficient to carry out the proper stereospecific reaction to BH<sub>4</sub> (45). With the crystallographic structures (Figure 13), including the characteristics of the active centers of all three enzymes, the essential information for the interpretation of the reaction mechanism is available. Moreover, NMR studies on the reaction mechanisms of all three enzymes revealed the details of the hydrogen transfer process and the stereochemical course of the reactions.<sup>40</sup>

**7.17.3.3.1(i) Guanosine triphosphate cyclohydrolase I and guanosine triphosphate cyclohydrolase feedback regulatory protein** A catalytic mechanism was proposed based primarily on structural analysis obtained from *E. coli* GTPCH cocrystals with the dGTP analogue and several active site mutants, as described in the chapter about THF biosynthesis in Figure 5.

Structural crystallographic data are available for the recombinant GTPCH from the human and the *E. coli* enzymes,<sup>32,34,36</sup> for the rat GFRP,<sup>189</sup> and for the GTPCH–GFRP co-complex from the rat.<sup>190</sup> The homodecameric GTPCH is composed of two dimers of pentamers (Figure 13). Regarding the *E. coli* enzyme, each subunit contains 221 amino acids and folds into an  $\alpha + \beta$  structure with a predominantly helical N-terminus. The decamer has a toroidal shape with an approximate height of 65 Å and a diameter of 100 Å. It encloses a cavity of 30 Å × 30 Å × 15 Å that is accessible through the pores formed by the five helix bundles in the center of the pentamers, but has no opening at the decamer equator. The active site is located at the interface of three subunits, two from one pentamer and one from the other. There are thus 10 equivalent active sites per functional unit, each containing a zinc ion bound to the active site. In the 360 kDa GTPCH–GFRP complex, the decameric GTPCH is sandwiched by two GFRP homopentamers. The model for the rat complex contains 10 GTPCH units and 10 GFRP subunits, 10 phenylalanine molecules, and 10 zinc ions (Figure 13). Each monomer is folded into an  $\alpha/\beta$ -structure with six-stranded antiparallel  $\beta$ -sheet and two  $\alpha$ -helices, arranged onto a  $\beta\beta\alpha\beta\beta\alpha\beta\beta$  topology. The monomers form a propeller-like pentameric-disk with an overall dimension of 50 Å × 50 Å × 20 Å for this complex. The contacts between GFRP and GTPCH are, besides salt bridging, primarily due to van der Waals interactions, as the concave side of GFRP pentamer exhibits neutral to negative surface potential, whereas the contact surface of the GTPCH is mainly positively charged. The phenylalanine binding sites are located at the interface between GFRP and GTPCH, and phenylalanine binding enhances the association of these proteins. The complex structure suggests that phenylalanine-induced GTPCH–GFRP complex formation enhances GTPCH activity by locking the enzyme in the active site.

**7.17.3.3.1(ii) 6-Pyruvoyl tetrahydropterin synthase (PTPS)** The reaction from H<sub>2</sub>NTP (7) to PTP (42) is catalyzed by PTPS in a Zn<sup>2+</sup>- and Mg<sup>2+</sup>-dependent reaction without consuming an external reducing agent (Figure 10). This conversion involves a stereospecific reduction by an internal redox transfer between atoms N-5, C-6, and C-1', oxidation of both side-chain hydroxyl groups, and an unusual triphosphate elimination at the C-2'–C-3' bond in the side chain. Crystallographic analysis revealed that PTPS is composed of a pair of trimers arranged in a head-to-head fashion to form the functional hexamer.<sup>52</sup> The homohexamer contains six active sites that are located on the interface of three monomers, two subunits from one trimer and one subunit from the other trimer. The catalytic center and the reaction mechanism were studied by crystallographic and kinetic analysis of wild-type and mutant PTPS from the rat.<sup>191,192</sup> In addition, the crystal structure of the inactive mutant Cys42Ala PTPS in complex with its natural substrate H<sub>2</sub>NTP was determined.<sup>192</sup> Each catalytic center harbors a Zn<sup>2+</sup>-metal binding site in a 12-Å deep cavity. The active site pocket with the specific pterin-anchoring Glu residue for salt-bridging plus two hydrogen-bonding amino acids appears to be similar to the equivalent sites in GTPCH, SR, dihydroneopterin epimerase, and neopterin aldolase. The active site pocket contains in addition two catalytic motifs: a Zn<sup>2+</sup>-binding site and an intersubunit catalytic triad formed by a Cys, an Asp, and a His residue. The tetravalent coordination of the transition metal is accomplished through the N $\epsilon$ -atoms of three His residues and a fourth ligand provided by the pyruvoyl moiety of the H<sub>2</sub>NTP substrate. Unfortunately, neither the triphosphate nor the putative Mg<sup>2+</sup> ion moieties could be defined in the electron density map. Zn<sup>2+</sup> plays a crucial role in catalysis as it activates the protons of the substrate and stabilizes the intermediates (40 and 41). The proposed reaction mechanism is the following: Protonation of N-5 and abstraction of a proton from the C-1' side-chain carbonyl atom lead to the N-5-C-6 double-bond reduction (40). Stereospecific protonation of C-6 and oxidation of C-1'–OH to C-1'=O leads to 41. The last step is the abstraction of a proton from the C-2' carbon of the carbonyl side chain, followed by triphosphate elimination and tautomerization to yield PTP (42).





**Figure 13** Three-dimensional structures of  $\text{BH}_4$  metabolizing enzymes. Ribbon-type representations of the main-chain foldings of the proteins and enzymes involved in *de novo*  $\text{BH}_4$  biosynthesis and regeneration are shown: GTPCH, GTP cyclohydrolase I (homodecamer); GFRP, GTP cyclohydrolase feedback regulatory protein (homopentamer); PTPS, 6-pyruvoyl tetrahydropterin synthase (homohexamer); SR, sepiapterin reductase (homodimer); PCD/DCoH, pterin-4a-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor-1 $\alpha$  (homotetramer); DHPR, dihydropteridine reductase (homodimer). Substrates are shown in ball-and-stick representation with atoms in standard colors. On the right side the enzymes are shown rotated by 90° around the *x*-axis. The crystal structure coordinates used are GTPCH from *E. coli* (Protein Data Bank entry code 1GTP), GFRP from rat (1JG5), PTPS from rat liver (1GTQ, 1B66), SR from mouse (1SEP), PCD/DCoH from rat (1DCH, 1DCP), and DHPR from rat liver (1DHR). The figure was made using the programs MOLSCRIPT (Kraulis, 1991, *J. Appl. Crystallogr.* 24, 945), RENDER (Merrit and Murphy, 1994, *Acta Crystallogr. Sct. D Biol. Crystallogr.* D50, 869), and/or PyMOL (<http://pymol.sourceforge.net/>).



The Cys residue (Cys42 in the rat enzyme) of the catalytic triad appears to be the general base for stereospecific protein abstraction in both reaction steps.

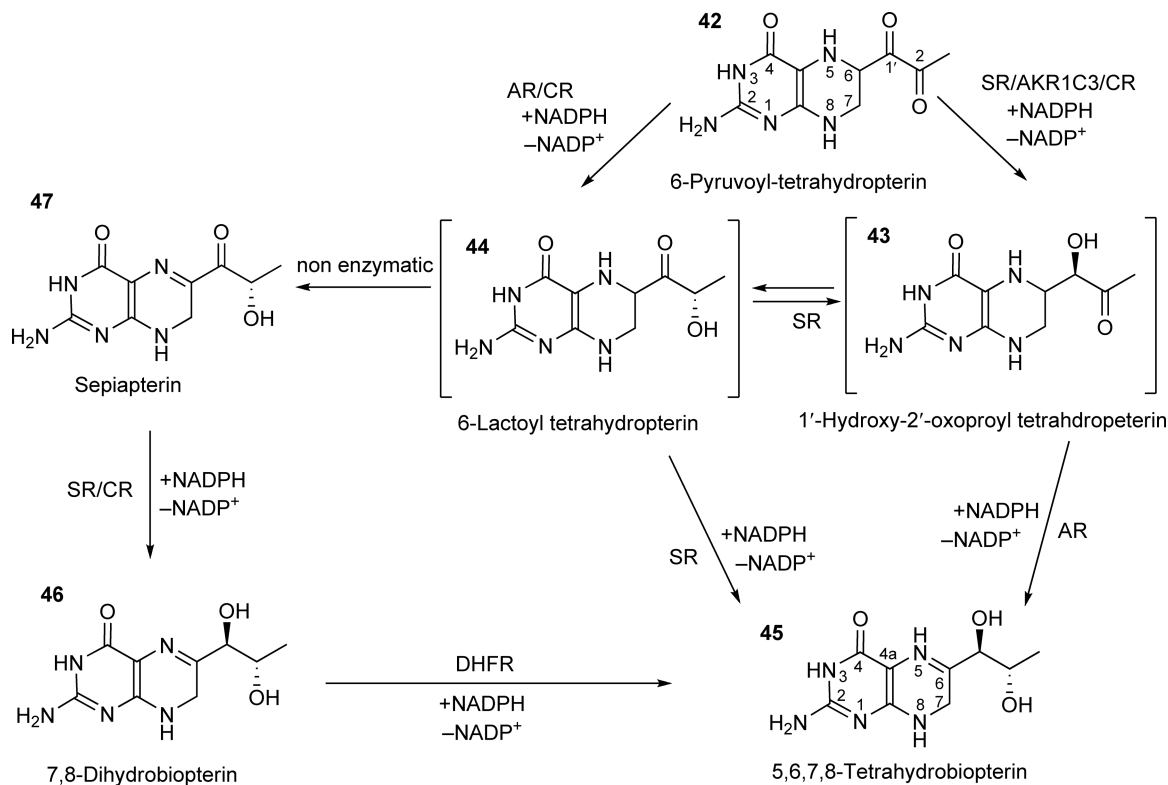
Only the recombinant rat liver enzyme could be crystallized, which yielded interpretable diffraction data.<sup>191</sup> Each subunit folds into a compact, single-domain  $\alpha + \beta$  structure (Figure 13). The monomer consists of a sequential, 4-stranded, antiparallel  $\beta$ -sheet, and 3 monomers assemble into a trimer, forming a 12-stranded antiparallel  $\beta$ -barrel structure surrounded by a ring of  $\alpha$ -helices. Crystallographic and experimental data revealed that the mammalian PTPS is homohexameric and dissociates into trimers. Two trimers arrange in a head-to-head fashion to form a barrel, the functional PTPS hexamer with an overall shape of  $60 \text{ \AA} \times 60 \text{ \AA} \times 60 \text{ \AA}$ . Due to the relatively tilted order of the  $\beta$ -sheet, the pore in the trimer is conically shaped with a 6–12  $\text{ \AA}$  diameter. In the hexameric enzyme, the pore has a smaller opening toward the trimer interface, and also opens up equatorially to the hexamer surroundings. The inside of the barrel accumulates a cluster of basic and aromatic residues stretching radially into the pore. Each subunit of the homohexamer contains one putative active site that is located at the interface of three monomers, two subunits from one trimer and one subunit from the other trimer. Each active site harbors three histidine residues, with the  $N_{\epsilon}$ -atoms coordinating the binding of the  $Zn^{2+}$  metal. In the unliganded state, a water molecule is bound as the fourth ligand.

**7.17.3.3.1(iii) Sepiapterin reductase** The final step is the NADPH-dependent reduction of the two side-chain keto groups of PTP (42) by sepiapterin reductase (SR). The overall structure of SR is a homodimer stabilized by a common four-helix bundle.<sup>193</sup> Each monomer contributes to two  $\alpha$ -helices to the central dimerization domain and forms a separate complex composed of seven parallel  $\beta$ -sheets surrounded by  $\alpha$ -helices. The C-terminal end of the  $\beta$ -sheets contains in close vicinity NADPH and the pterin binding site, the latter comprising a 15- $\text{ \AA}$  deep pocket. The pterin substrate is anchored by the guanidino moiety of a specific Asp. The C-1' carbon of the pterin side chain is in direct proximity to NADPH and a Tyr OH-group. This Tyr is the central active site residue for optimal proton transfer.<sup>194</sup> Based on kinetic, crystallographic, and NMR data, the initial step is the NADPH-dependent reduction at the side-chain C-1'-keto function leading to the formation of 1'-OH-2'-oxopropyl tetrahydropterin (43).<sup>40</sup> An internal rearrangement of the keto group through side-chain isomerization leads to the formation of 1'-keto compound 6-lactoyl tetrahydropterin (44). Intermediate 44 is then reduced to  $BH_4$  (45) in a second NADPH-dependent reduction step. While the pterin substrate remains bound to the active site, the redox cofactor has to be renewed after the first reduction. It is thus assumed that NADP is exchanged at the opening located at the opposite side of the pterin-binding and entry pocket.

The 1.25  $\text{ \AA}$ -crystal structure of the mouse SR in complex with NADP has been solved.<sup>193</sup> The 261 amino acids of the monomer fold into a single domain  $\alpha/\beta$ -structure. A seven-stranded parallel  $\beta$ -sheet in the center of the molecule is sandwiched by two arrays of three  $\alpha$ -helices. The association of two monomers to the active homodimeric SR leads to the formation of a four-helix bundle (Figure 13). Owing to the two-folded crystallographic symmetry of the homodimeric molecule, the parallel  $\beta$ -sheets in monomer A is in an antiparallel orientation relative to the  $\beta$ -sheet of monomer B enclosing an angle of  $90^\circ$ . The overall dimensions of the SR dimer are  $40 \text{ \AA} \times 50 \text{ \AA} \times 80 \text{ \AA}$ . The two substrate pockets bind sepiapterin (or 6-pyruvoyl-tetrahydropterin; 42), and the cofactor NADP/NADPH from opposite sides to the enzyme.

#### 7.17.3.3.2 Alternative routes for biosynthesis of $BH_4$

Besides the involvement in the *de novo* biosynthesis of  $BH_4$ , SR may also participate in the pterin *salvage* pathway by catalyzing the conversion of sepiapterin (Figure 14, 47) into 7,8-dihydrobiopterin (46) that is then transformed to  $BH_4$  by dihydrofolate reductase (DHFR; EC 1.5.1.3).<sup>3</sup> Both reactions consume NADPH. Although SR is sufficient to complete the  $BH_4$  biosynthesis, a family of alternative NADPH-dependent aldo-keto reductases, including carbonyl reductases (CR), aldose reductases (AR), and the  $3\alpha$ -hydroxysteroid dehydrogenase type 2 (AKR1C3) may participate in the diketo reduction of the carbonyl side chain *in vivo*.<sup>195–197</sup> Moreover, based on the discovery of the autosomal recessive deficiency for SR, which presents with neurotransmitter deficiency but without hyperphenylalaninemia (see also below), different routes for the final two-step reaction of  $BH_4$  biosynthesis involving these alternative reductases were proposed<sup>197,198</sup> (Figure 14). CR converts 6-pyruvoyl tetrahydropterin (42) into both, the 2'-oxo-tetrahydropterin (43) and the lactoyl tetrahydropterin (or 1'-oxo-tetrahydropterin; 44); however, the rate of production is much more favorable for the 1'-oxo-tetrahydropterin intermediate. On the contrary, the  $3\alpha$ -hydroxysteroid



**Figure 14** Reaction mechanism for alternative routes for BH<sub>4</sub> biosynthesis (see text for details).

dehydrogenase type 2 efficiently converts compounds 42 into 43. AR can convert 6-pyruvoyl tetrahydropterin into 6-lactoyl tetrahydropterin (44), or to 2'-oxo-tetrahydropterin (43) to BH<sub>4</sub>. In summary, alternative routes for BH<sub>4</sub> in the absence of SR involve in one pathway option through the 2'-oxo-tetrahydropterin the concerted action of 3 $\alpha$ -hydroxysteroid dehydrogenase type 2 and AR, thereby consuming 2 NADPH. In the other pathway option through the 1'-oxo-tetrahydropterin intermediate, AR, CR, and dihydrofolate reductase are required. This second route involves consecutively the compounds 42, 44, 45, 46, and 47, and consumes 3 NADPH due to a nonenzymatic reduction step to sepiapterin (from 44 to 47). It was proposed that due to low expression or activity of dihydrofolate reductase and 3 $\alpha$ -hydroxysteroid dehydrogenase type 2 in the human brain, the biosynthesis from 6-pyruvoyl tetrahydropterin to BH<sub>4</sub> in the absence of SR cannot be completed. This leads to central BH<sub>4</sub> deficiency with accumulation of the unstable 1'-oxo-tetrahydropterin (44) and its degradation products, detected as 'abnormal' pterin metabolites (that is, 46).

Besides that sepiapterin (47) and BH<sub>2</sub> (46) can both be taken up and metabolized by the salvage pathway to replenish the BH<sub>4</sub> pool in the body, uptake of these two compounds in mice turns out to be significantly more efficient than natural BH<sub>4</sub> itself. Furthermore, these *in vivo* studies with mice showed that application of exogenous BH<sub>4</sub>, irrespective whether it was the natural 6R or the unnatural 6S-diastereomer of BH<sub>4</sub>, results in tissue accumulation of 6R-BH<sub>4</sub> (45). Apparently, the 6S-form is oxidized to 7,8-BH<sub>2</sub> and subsequent taken up into tissues, followed by DHFR-dependent reduction to the natural 6R-BH<sub>4</sub>.<sup>199</sup>

#### 7.17.3.4 Regeneration of BH<sub>4</sub>

Regeneration of BH<sub>4</sub> is an essential part of the phenylalanine hydroxylating system (see also 'Cofactor functions'). During the catalytic event of aromatic amino acid hydroxylases, molecular oxygen is transferred to the corresponding amino acid and BH<sub>4</sub> is oxidized to BH<sub>4</sub>-4 $\alpha$ -carbinolamine (Figure 15).<sup>143,200</sup> Two enzymes are

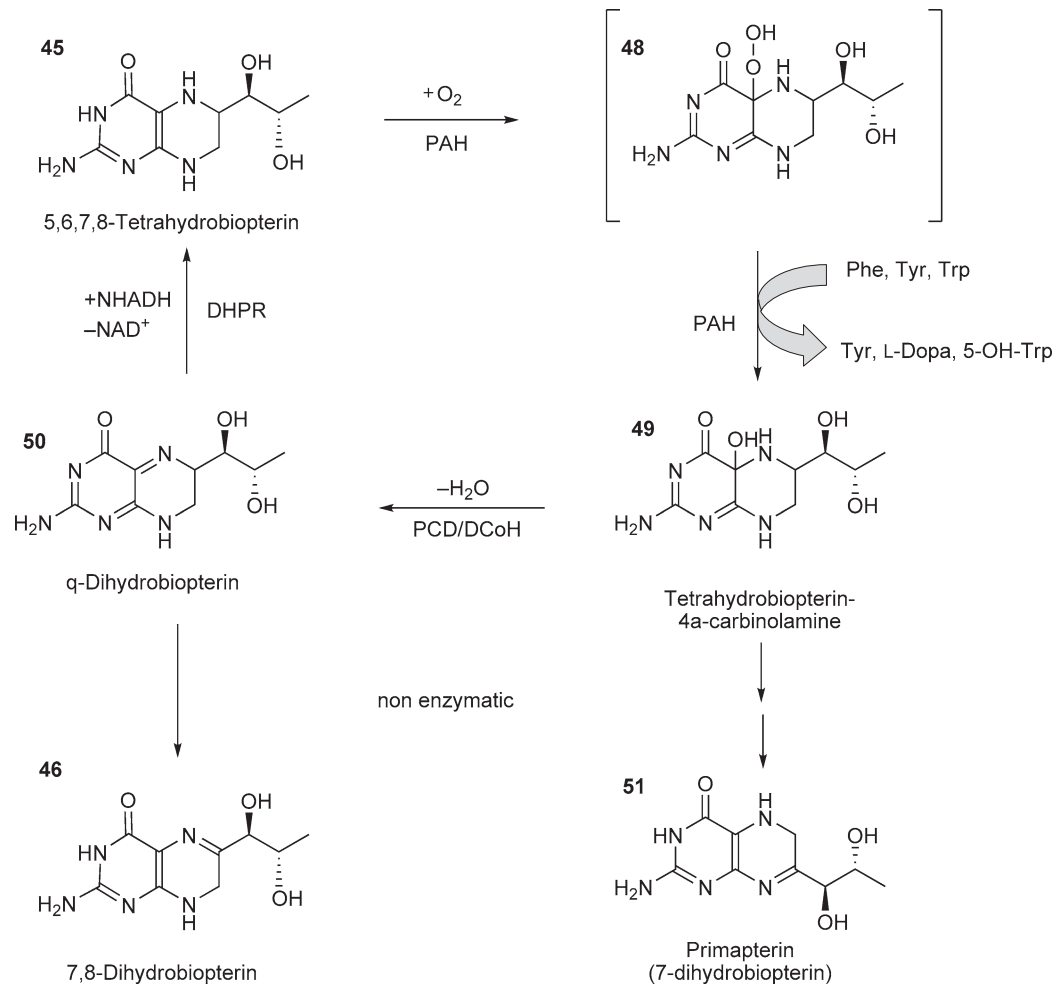
involved in its subsequent dehydration and reduction to BH<sub>4</sub>: pterin-4a-carbinolamine dehydratase (EC 4.2.1.96; PCD) and DHPR (EC 1.6.99.7). Enzymatic recycling of BH<sub>4</sub> is essential for phenylalanine metabolism (1) to ensure a continuous supply of reduced cofactor, and (2) to prevent accumulation of harmful metabolites produced by the rearrangement of BH<sub>4</sub>-4a-carbinolamine. The primary structure of PCD is identical with a protein of the cell nucleus that has transcription function, named dimerization cofactor (DCoH) of hepatocyte nuclear factor 1 $\alpha$  (HNF-1 $\alpha$ ), recently reported to have general transcriptional function.<sup>201–203</sup> In the following PCD will be designated as the dual-function protein PCD/DCoH. Furthermore, a functional homologue, DCoH2 or DCoH $\alpha$ , was identified, which partially complements PCD/DCoH mutants in mice and man.<sup>167,204</sup>

#### 7.17.3.4.1 Reaction mechanism of the regenerative pathway

**7.17.3.4.1(i) Pterin-4a-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor-1 $\alpha$**  The dehydration of BH<sub>4</sub>-4a-carbinolamine (**49**), the first product of the reaction of aromatic amino acid hydroxylases (**Figure 15**), is catalyzed by the enzyme PCD/DCoH. The human cytoplasmic PCD/DCoH, whose sequence is identical to that of the rat protein, is a homotetramer with a molecular mass of 11.9 kDa per subunit.<sup>202,203</sup> Using chemically synthesized pterin-4a-carbinolamine it has been demonstrated that the enzyme shows little sensitivity to the structure or configuration of the 6-substituent of its substrate, and to the 4a(R)- and 4a(S)-hydroxy stereoisomers.<sup>205</sup> Obviously, the binding pocket has a relatively high degree of flexibility and might not be designed to recognize only BH<sub>4</sub>-4a-carbinolamine. X-ray crystal structures of the tetrameric enzyme complexed with the product analogue 7,8-dihydrobiopterin (**46**) revealed four active sites harboring three essential and conserved histidines.<sup>206</sup> Detailed enzymatic studies on the stereospecificity and catalytic function revealed a dehydration mechanism in which the three histidines in PCD/DCoH are crucial for activity.<sup>207–209</sup> The quinonoid dihydrobiopterin (**50**) product is a strong inhibitor of PCD/DCoH with a  $K_i$ -value of about one half of its respective  $K_m$ -value, and no inhibition was observed with 7,8-dihydrobiopterin (**46**).<sup>205</sup> Furthermore, PAH is not inhibited by its cofactor product, BH<sub>4</sub>-4a-carbinolamine, but by primapterin (**51**). In the absence of PCD/DCoH, dehydration of BH<sub>4</sub>-4a-carbinolamine occurs also nonenzymatically, but at a rate that is, at least in the liver, insufficient to maintain BH<sub>4</sub> in the reduced state.<sup>210</sup> As a consequence, liver PCD/DCoH deficiency in man causes BH<sub>4</sub>-4a-carbinolamine to be rearranged through a spiro structure intermediate to dihydroprimapterin (7-substituted dihydrobiopterin; **51**) that is excreted in the urine.<sup>211,212</sup>

The crystal structure of cytoplasmic PCD/DCoH from human and rat liver has been solved.<sup>213,214</sup> The single domain monomer of 103 amino acids comprises of three  $\alpha$ -helices packed against one side of a four-stranded, antiparallel  $\beta$ -sheet. The functional enzyme is a homotetramer where each of the monomers contributes one helix (helix  $\alpha_2$ ) to a central four-helix bundle (**Figure 13**). In the tetramer two monomers form an eight-stranded, antiparallel  $\beta$ -sheet with six helices packing against it from one side. The concave, eight-stranded  $\beta$ -sheet with its two protruding loops at either end is reminiscent of the saddle-like shape seen in the TATA box binding protein. The overall dimensions of the tetramer are 60 Å  $\times$  60 Å  $\times$  60 Å.

To probe the relationship between dehydratase activity and transcriptional coactivator functions, the X-ray crystal structures of the free enzyme and its complex with the product analogue 7,8-dihydrobiopterin were solved.<sup>206</sup> The ligand binds at four sites per tetrameric enzyme, with little apparent conformational change in the protein. The pterin binds within an arch of aromatic residues that extends across one dimer interface. The bound ligand makes contacts with the three conserved histidines, and this arrangement restricts proposals for the enzymatic mechanism of dehydration. PCD/DCoH binds as a dimer to the helical dimerization domain of HNF1 $\alpha$ . A mutant of PCD/DCoH with reduced dehydratase activity was not affected in protein–protein interaction and bound to HNF1 $\alpha$ , showing that enzymatic activity is not essential for HNF1 binding.<sup>205,215</sup> On the contrary, it was reported that PCD/DCoH retained its enzymatic activity while complexed with HNF1 as a  $\alpha_2\beta_2$  heterotetramer.<sup>216</sup> The functional homologue, DCoH2,<sup>167</sup> partially complements PCD/DCoH and, based on crystal structure analysis, adopts as dimer of identical folds with structural differences confined largely to the protein surfaces and the tetramer interface.<sup>204</sup> The two paralogues, PCD/DCoH and DCoH2 can even form homotetramers or mixed heterotetramers in solution. A functional PCD/DCoH homologue *phbB* with dehydratase activity as a dimeric enzyme was described from *Pseudomonas aeruginosa phbB*.<sup>217</sup> The amino acid sequences of the mature human and rat liver proteins are identical, and the mouse varies by only one amino acid.<sup>201,203</sup> Similar proteins for PCD/DCoH and DCoH2 were found in many species throughout the animal



**Figure 15** Reaction mechanism for the BH<sub>4</sub> regenerative pathway (see text for details).

kingdom and in various bacteria (e.g., the Cyanobacterium species *Synechocystis*, *Aquifex aeolicus*, and *P. aeruginosa*; see also Citron *et al.*<sup>202</sup>).

**7.17.3.4.1(ii) Dihydropteridine reductase** The final conversion of quinonoid dihydrobiopterin into BH<sub>4</sub> is carried out by the dimeric dihydropteridine reductase (DHPR) (Figure 15). Although the crystallographic structure of the DHPR–NADH binary complex was solved, the location of the active sites is not known from these studies. Nevertheless, an active site pocket involving the Tyr-X-X-X-Lys motif, typical for short-chain dehydrogenases, was proposed to participate in proton donation. Following the classical mechanisms of dehydration, one molecule of water is released and the product quinonoid dihydrobiopterin is reduced back to BH<sub>4</sub> in a NADH-dependent reaction. This final reaction of the regeneration pathway involves direct hydride transfer from the reduced nicotinamide ring to the quinonoid dihydrobiopterin by DHPR. This reaction is supported by the proposed enzyme mechanism of NAD(P)H-dependent reductases and by the lack of detectable prosthetic groups such as flavin or metal ions.<sup>218</sup> The hydride transfer occurs from the B-face of NADH with transfer of the pro-S hydrogen.

The structure of a binary complex of the rat<sup>219</sup> and the human DHPR<sup>220</sup> has been determined by X-ray crystallography (Figure 13). DHPR is an  $\alpha/\beta$  protein with a central twisted  $\beta$ -sheet flanked on each side by a layer of  $\alpha$ -helices. The  $\beta$ -sheet has seven parallel strands and a single antiparallel strand at one edge leading to the carboxyl terminus of the protein. Connections between individual  $\beta$ -strands involve  $\alpha$ -helices. Exceptionally, two  $\beta$ -strands are joined by a short stretch of polypeptide in random coil conformation. The overall enzyme dimensions are 34 Å × 50 Å × 73 Å. The topology of the backbone folding of DHPR is quite distinct from that of dihydrofolate reductase (DHFR) although the first six strands of the central  $\beta$ -sheet in DHPR have the same overall topological connectivity as that found for the coenzyme-binding domains of several other NAD(P)-dependent dehydrogenases. In contrast to the rat enzyme, human DHPR contains two bound NADH molecules per dimer and despite the sequential amino acid changes there are only small differences between the two structures.<sup>220</sup>

### 7.17.3.5 Tetrahydrobiopterin in Disease

BH<sub>4</sub> deficiency is associated with a rare variant of hyperphenylalaninemia that was originally termed ‘atypical’ or ‘malignant’ phenylketonuria (PKU), as it was unresponsive to low-phenylalanine diet. Today, one differentiates among the BH<sub>4</sub> deficiencies between monoamine neurotransmitter deficiency concomitant with hyperphenylalaninemia, that is, the classical forms of BH<sub>4</sub> deficiencies, and monoamine neurotransmitter deficiency without hyperphenylalaninemia, that is, DRD and SR deficiency.<sup>170</sup> The classical forms present phenotypically with deficit of the neurotransmitters dopamine and serotonin, and progressive neurological symptoms.<sup>221</sup> It is to be stressed that it is a highly heterogeneous group of diseases affecting either all organs, including the central nervous system, or only the peripheral hepatic phenylalanine hydroxylating system.<sup>222–224</sup> Classical BH<sub>4</sub> deficiency can be caused by mutations in genes encoding enzymes involved in its biosynthesis (GTPCH and PTPS)<sup>225</sup> or regeneration (PCD/DCoH and DHPR).<sup>223,224,226</sup> The mutations are all inherited autosomal recessively. Biochemical, clinical, and DNA data of patients with BH<sub>4</sub> deficiencies are tabulated in the BIoDEF and BIoMDB databases and are available on the Internet (<http://www.biopku.org>).<sup>227</sup> DRD and SR deficiency are due to a defective BH<sub>4</sub> biosynthesis that occurs without hyperphenylalaninemia. DRD, initially described as Segawa disease may be caused by autosomal dominant mutations in the *GCH* gene, but about 40% of the patients present without identifiable DNA mutations. Furthermore, most of these inherited dominant alleles have no penetrance in the parent of the patient.<sup>228–232</sup> In contrast, SR deficiency is inherited as an autosomal recessive trait, and exhibits biochemical abnormalities that are related to disease exclusively in neurotransmitter metabolites and pterins in the cerebrospinal fluid (CSF).<sup>170,198</sup> As mentioned before, SR deficiency suggested the presence of an alternative biosynthetic pathway operative in peripheral tissues and involving alternative aldose–ketose reductases and the dihydrofolate reductase.

Decreased levels of BH<sub>4</sub> in the CSF have also been documented in other neurological diseases presenting phenotypically without hyperphenylalaninemia, such as Parkinson’s disease,<sup>233</sup> autism,<sup>234</sup> depression,<sup>235</sup> and Alzheimer’s disease.<sup>236</sup> In some of these, administration of BH<sub>4</sub> has been reported to improve the clinical symptoms.<sup>233,237,238</sup> Unfortunately, others were not able to confirm these results,<sup>185</sup> and thus a benefit of BH<sub>4</sub> therapy at least for this group of diseases is questionable.

Another group of diseases with perturbed BH<sub>4</sub> metabolism in human epidermis are skin disorders, including vitiligo and the Hermansky–Pudlak syndrome. Although the etiology for these disorders is not yet known, both involve lowered PCD/DCoH activities concomitant with 6- and 7-biopterin and H<sub>2</sub>O<sub>2</sub> accumulation in skin, tyrosinase inhibition, and abnormal melanin biosynthesis.<sup>239–242</sup>

Finally, a group of mild phenylketonuric patients with autosomal recessive mutations in the gene encoding PAH but with normal BH<sub>4</sub> metabolism respond to treatment with oral BH<sub>4</sub> as an alternative to a low-phenylalanine diet.<sup>172,243,244</sup> Recent investigations on the mechanism of this BH<sub>4</sub>-responsive PKU revealed a multifactorial basis, including a chemical chaperon effect of BH<sub>4</sub> that prevents degradation of mutant PAH proteins and protection from rapid inactivation. Furthermore, some mutants show a kinetic defect with increased  $K_m$  for BH<sub>4</sub>.<sup>151,245,246</sup> In any case, a PAH allele with residual activity of its expressed PAH enzyme is a prerequisite for response.

### 7.17.3.6 Animal Models to Study BH<sub>4</sub> Function and Deficiency

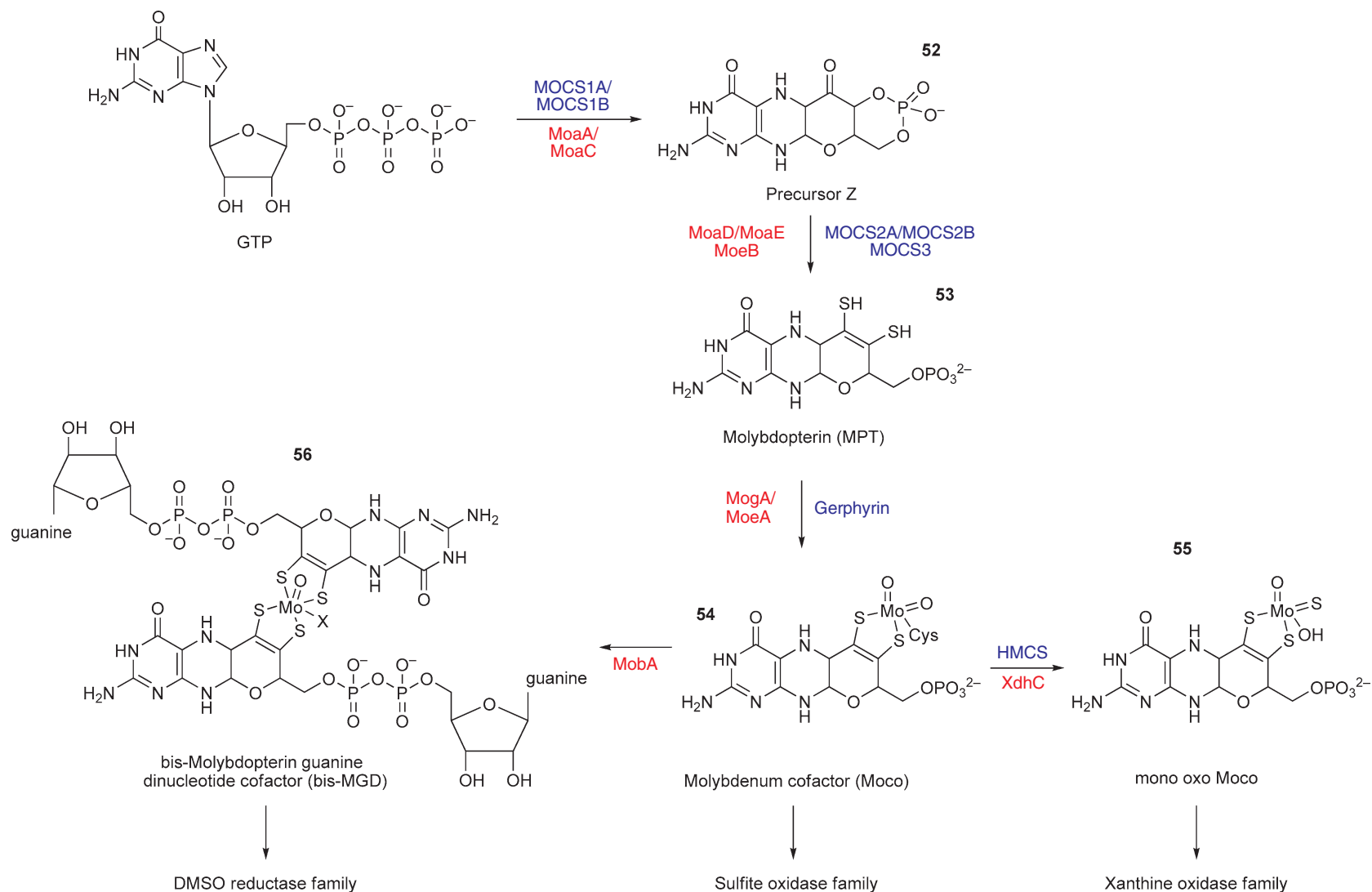
Today, three different animal models are available for BH<sub>4</sub> deficiencies; these are the *bpb-1* mouse and two different knockout strains for PTPS. The *bpb-1* mouse was produced by screening ethylnitrosourea (ENU)-treated mice for the presence of hyperphenylalaninemia.<sup>247</sup> This chemically induced mutant strain has a mild, neonatal, and transient form of hyperphenylalaninemia with lowered GTPCH activity and BH<sub>4</sub> cofactor levels in the liver. In the brain, the mouse showed low levels of BH<sub>4</sub>, catecholamines, serotonin, and TH activity. Thus, biochemical analysis revealed that this strain provides an animal model for DRD.<sup>248</sup> Unfortunately, no DNA mutation could be identified thus far but genomic mapping narrowed the site of mutation to an interval containing the mouse *GCHI* gene.<sup>249</sup> Two PTPS-deficient strains were generated by targeted *Pts*-gene disruption. The mouse presented with perinatal lethality owing to monoamine neurotransmitter depletion, but rescue experiments with neurotransmitter precursors and/or BH<sub>4</sub> revealed that liver PAH and brain TH and TPH might be regulated differentially.<sup>152,250,251</sup> This *Pts*-knockout mouse is thus an important tool to investigate BH<sub>4</sub> deficiency and their consequences for all BH<sub>4</sub>-dependent enzyme systems. Besides animal models for BH<sub>4</sub> deficiency, an endothelium-specific GTPCH-overexpressing mouse was generated with elevated total biopterin content in lung, heart, and aorta, but normal levels in liver and plasma.<sup>163</sup> This transgenic mouse model turned out to be a valuable tool to study eNOS function and vascular disease states.<sup>169</sup>

## 7.17.4 Molybdopterin and Molybdenum Cofactor Biosynthesis

### 7.17.4.1 An Introduction to Molybdopterin Biosynthesis and Molybdoenzymes

The Moco is the essential component of a group of redox enzymes, which are diverse in terms of their phylogenetic distribution and molecular architectures, both at the overall level and in their catalytic geometry.<sup>10</sup> In these enzymes, molybdenum is coordinated to a dithiolene group on the 6-alkyl side chain of a pterin called molybdopterin (MPT)<sup>8</sup> (Scheme 1). Molybdenum is the only second-row transition metal that is required by most living organisms, and the few species that do not require molybdenum use tungsten, which is also coordinated by the dithiolene group of MPT and lies immediately below molybdenum in the periodic table.<sup>252</sup> A wide variety of transformations are catalyzed by these enzymes at carbon, sulfur, and nitrogen atoms, which include the transfer of an oxo group or two electrons to or from the substrate. Molybdenum and tungsten-containing enzymes often catalyze the same kind of chemistry – such as the oxidation of aldehydes to carboxylic acids. More than 40 molybdenum and tungsten enzymes were identified in bacteria, archaea plants, and animals to date. While molybdenum-containing enzymes are found in all aerobic organisms, tungsten-containing enzymes are found only in obligate, typically thermophilic anaerobes.<sup>252</sup> Some of the better-known Moco-containing enzymes include sulfite oxidase, xanthine dehydrogenase, and aldehyde oxidase in humans; assimilatory nitrate reductase in plants; and dissimilatory nitrate reductase, dimethylsulfoxide (DMSO) reductase, and formate dehydrogenase in bacteria and archaea. The mononuclear molybdenum enzymes are categorized on the basis of the structures of their molybdenum centers, dividing them into three families, each with a distinct active site structure and a distinct type of reaction catalyzed (Figure 16): the xanthine oxidase family, the sulfite oxidase family, and the DMSO reductase family.<sup>10</sup> Tungsten-containing enzymes comprise a separate family and are also divided into three different subgroups: the aldehyde ferredoxin oxidoreductase family, the formate dehydrogenase family, and the acetylene hydratase family.<sup>252,253</sup> This chapter focuses on the different classes





**Figure 16** The biosynthesis of Moco and bis-MGD. Shown is a scheme of the biosynthetic pathway for Moco biosynthesis in bacteria and eukaryotes. The proteins involved in the reactions are colored in red for bacterial proteins and blue for human proteins. In bacteria, Moco (54) can be further modified by the attachment of, for example, GMP, forming MGD, and two equivalents of MGD are bound to molybdenum, forming the so-called bis-MGD cofactor (56). Further, Moco can be modified by the replacement of one oxo ligand by a sulfido ligand, forming the monooxo Moco (55). The three molybdenum containing enzyme families are divided into the xanthine oxidase, sulfite oxidase, and DMSO reductase families according to their active site structures.

of molybdoenzymes, the biosynthesis of Moco in bacteria and humans, and its incorporation into specific target proteins. The group of tungsten-containing enzymes is not discussed here as they are the focus of other reviews.<sup>252,254</sup>

## 7.17.4.2 The Families of Molybdoenzymes

### 7.17.4.2.1 The xanthine oxidase family

The XO family is characterized by an MPT-Mo<sup>VI</sup>OS(OH) core in the oxidized state, with one MPT equivalent coordinated to the metal (55, Figure 16). The additional sulfido group is cyanide labile.<sup>255</sup> Removal of the sulfido group results in the formation of an inactive desulfo-enzyme, with an oxygen ligand replacing the sulfur at the Mo active site.<sup>256</sup> Among the members of the xanthine oxidase family are XO, xanthine dehydrogenase (XDH), aldehyde oxidase (AO), aldehyde oxidoreductase, and carbon monoxide (CO) dehydrogenase.<sup>10</sup> These enzymes typically catalyze the hydroxylation of carbon centers. CO dehydrogenase, however, differs in two respects: first, it oxidizes CO to CO<sub>2</sub> and second it was found to harbor a binuclear Mo—Cu center attached to a molybdopterin–cytosine dinucleotide cofactor (MCD),<sup>257</sup> while all other molybdoenzymes of this family contain a mononuclear Mo center at the MPT moiety.

### 7.17.4.2.2 The sulfite oxidase family

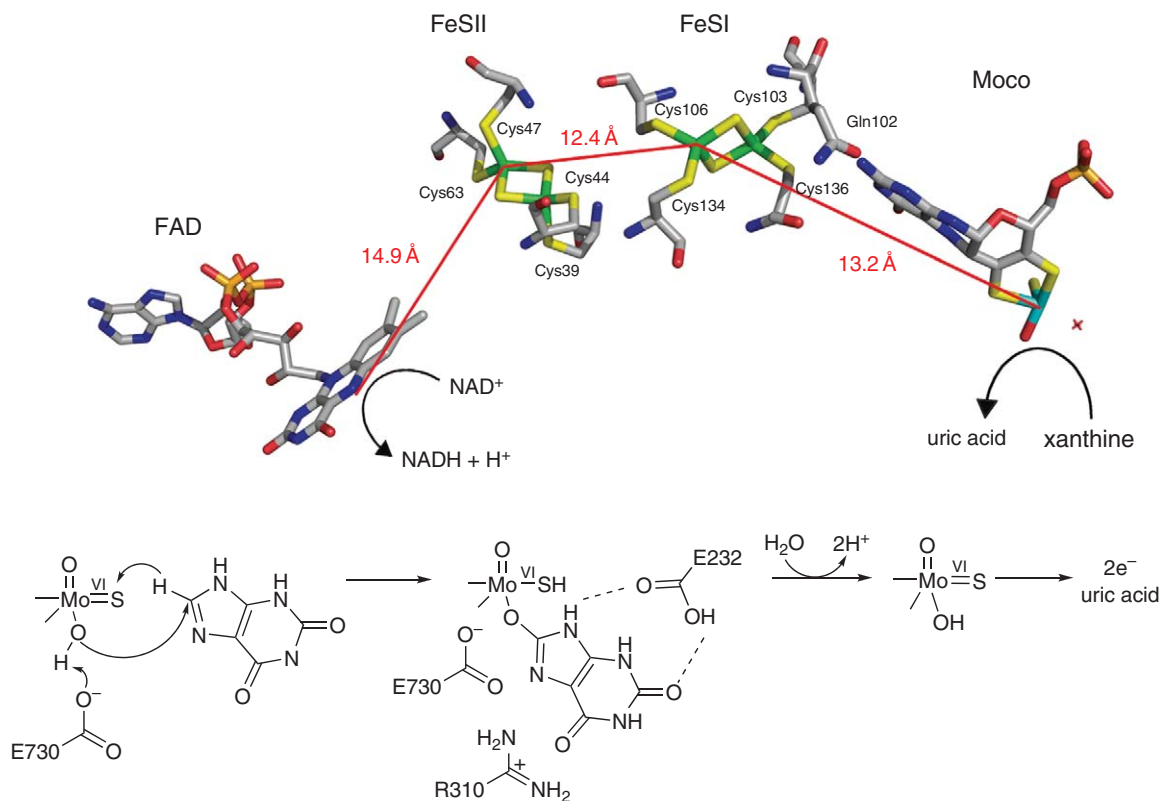
Enzymes of the sulfite oxidase family coordinate a single equivalent of the pterin cofactor with an MPT-Mo<sup>VI</sup>O<sub>2</sub> core in its oxidized state (54, Figure 16), and usually an additional cysteine ligand, which is provided by the polypeptide.<sup>10</sup> Members of this family catalyze the transfer of an oxygen atom either to or from the substrate. Among the members of this family are sulfite oxidase, sulfite dehydrogenase, assimilatory nitrate reductases, and the YedY protein, the catalytic subunit of a sulfite oxidase homologue in *E. coli*.<sup>258</sup> So far, all members of this family contain the MPT-form of Moco without an additional dinucleotide.

### 7.17.4.2.3 The DMSO reductase family

The DMSO reductase family is diverse in both structure and function, but all members have two equivalents of the pterin cofactor bound to the metal.<sup>10</sup> The molybdenum coordination sphere is usually completed by a single Mo=O group with a sixth ligand in the MPT<sub>2</sub>-Mo<sup>VI</sup>O(X) core (56, Figure 16); however, a sulfido ligand has been recently described to replace the oxo group in nitrate reductase from *Desulfovibrio desulfuricans* ATCC 27774.<sup>259</sup> The sixth ligand, X, can be a serine, a cysteine, a selenocysteine, or a hydroxide, and/or water molecule. The reactions catalyzed by members of this family frequently involve oxygen-atom transfer, but dehydrogenation reactions also occur. Members of the DMSO reductase family are exclusively found in eubacteria and include, among other enzymes, the dissimilatory nitrate reductases, formate dehydrogenases, trimethylamine-*N*-oxide reductases, and biotin sulfoxide reductases. All these enzymes have two equivalents of molybdopterin guanine dinucleotide cofactor (MGD), a variant of MPT containing an additional guanosine monophosphate GMP group attached to the phosphate group of MPT (Figure 16).<sup>10</sup> In the absence of oxygen and in the presence of their respective substrates, these enzymes generally serve as terminal reductases. DMSO reductases from phototrophic bacteria are monomeric and lack other redox centers,<sup>260</sup> whereas the membrane-bound enzyme from *E. coli* has two additional subunits, one of which has four different [4Fe–4S] iron–sulfur centers. Most other enzymes of the DMSO reductase family also contain additional subunits and, in many cases an iron–sulfur center in the Moco-containing subunit.<sup>261</sup>

### 7.17.4.2.4 The enzymology of *Rhodobacter capsulatus* xanthine dehydrogenase (XDH)

Xanthine oxidoreductases are the best-studied enzymes of the XO family containing two nonidentical [2Fe–2S] clusters, FAD and the sulfurated form of Moco as catalytically acting units.<sup>262</sup> Mammalian XORs catalyze the hydroxylation of hypoxanthine and xanthine, the last two steps in the formation of urate, and exist originally as the dehydrogenase form (XDH, EC 7.17.1.4) but can be converted to the oxidase form (XO, EC 1.1.3.22) either reversibly by oxidation of sulfhydryl residues of the protein molecule or irreversibly by proteolysis.<sup>263,264</sup> XDH shows a preference for NAD<sup>+</sup> reduction at the FAD reaction site,<sup>265</sup> while XO exclusively uses dioxygen as terminal electron acceptor, leading to the formation of superoxide and hydrogen peroxide.<sup>266</sup> The enzyme has been implicated in diseases characterized by oxygen radical-induced tissue damage such as postischemic reperfusion injury.<sup>267</sup> The oxidation of xanthine takes place at the molybdenum center and the two electrons



**Figure 17** The enzymatic mechanism of *R. capsulatus* XDH. Proposed base catalyzed mechanism for *R. capsulatus* XDH with xanthine and orientation of xanthine at the active site. Shown are residues Glu232, Arg310, and Glu730 interacting with xanthine. Also shown is the coordination of Moco and FeSI, FeSII and FAD and their distances at the active site of *R. capsulatus* XDH using the coordinates from the Protein Data Bank (accession number 1JRO).

thus introduced are rapidly distributed in one-electron steps to the FeS center and FAD according to their relative redox potentials (Figure 17). The reoxidation of the reduced enzyme by the oxidant substrate, either  $\text{NAD}^+$  or molecular oxygen, occurs through FAD in a two-electron transfer step. The two  $[\text{Fe}-2\text{S}]$  clusters (FeSI and FeSII) are indistinguishable in terms of their absorption spectra, but can readily be distinguished by EPR spectroscopy.<sup>268,269</sup> The reaction mechanism of XDH from *R. capsulatus* has been studied in detail and is explained here as a representative mechanism of an enzyme for the XO family (Figure 17).<sup>270-272</sup> *R. capsulatus* XDH is a cytoplasmic enzyme that is highly identical to eukaryotic XORs.<sup>273</sup> Despite differences in subunit composition, the folds of bovine XDH and *R. capsulatus* XDH are very similar. The bacterial enzyme can be described as a butterfly-shaped  $(\alpha\beta)_2$  heterotetramer. Each  $(\alpha\beta)$  dimer represents one half of the active molecule and is encoded by two separate gene products, termed XdhA and XdhB, unlike the  $(\alpha)_2$  dimeric eukaryotic protein, which is derived from a single polypeptide chain.<sup>274</sup> Each subunit of the  $(\alpha\beta)$  heterodimer carries a specific set of cofactors, which are crucial for catalysis and electron transfer. The 50 kDa XdhA subunit harbors two  $[\text{Fe}-2\text{S}]$  clusters as well as a flavin adenine dinucleotide (FAD) cofactor; the 85 kDa XdhB subunit contains the molybdenum cofactor harboring a catalytically essential terminal sulfido ligand.<sup>273</sup> This cofactor is part of the active site binding pocket and catalyzes the oxidative hydroxylation of hypoxanthine to xanthine and further to uric acid.<sup>274</sup> The catalytic sequence of *R. capsulatus* XDH is initiated by abstraction of a proton from the Mo-OH group by the highly conserved active site residue Glu730, followed by nucleophilic attack of the resulting Mo-O<sup>-</sup> on the carbon center of the substrate (C2 in hypoxanthine, C8 in xanthine) and concomitant hydride transfer to the Mo=O of the molybdenum center (Figure 17).<sup>270</sup> Residue Glu232, on the other hand, is involved in both substrate binding and transition state stabilization.<sup>270</sup> It was shown that interaction of Arg310 with the C6 carbonyl group of the substrate xanthine stabilizes negative charge

accumulation on the heterocycle that accompanies nucleophilic attack at C8, thus, stabilizing the transition state and accelerating the reaction of substrate oxidation (Figure 17).<sup>275</sup> The role of R310 was confirmed by cocrystallization of inactive forms of the enzyme with both substrates xanthine and hypoxanthine, showing that the oxygen atom at the C6 position of both substrates is oriented toward R310 in the active site.<sup>276</sup>

### 7.17.4.3 Biosynthesis of Molybdopterin and the Molybdenum Cofactor

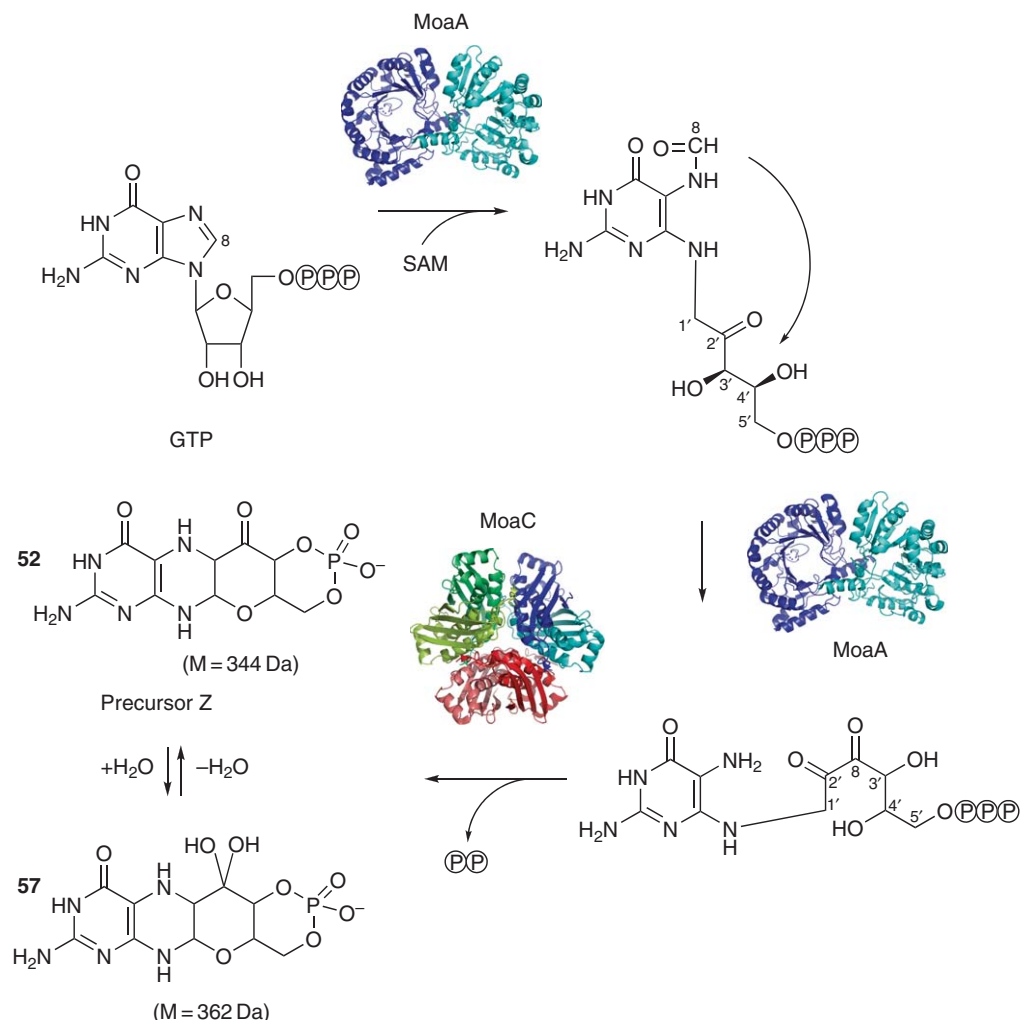
Moco biosynthesis is an ancient, ubiquitous, and highly conserved pathway leading to the biochemical activation of molybdenum.<sup>13</sup> Mutations in the human Moco biosynthetic genes lead to death in early childhood. The affected patients show severe neurological abnormalities such as attenuated growth of the brain, seizures, and dislocated ocular lenses.<sup>277</sup> So far, no effective therapy is available to cure the disease.

Moco biosynthesis has been extensively studied in *E. coli* by using a combination of biochemical, genetic, and structural approaches. Additional insights were provided by studies in eukaryotes such as plants and humans.<sup>16</sup> The biosynthesis of Moco can be divided into four steps in *E. coli* (Figure 16): (1) formation of precursor Z (51), (2) formation of MPT (53) from precursor Z, (3) insertion of molybdenum to form Moco (54), and (4) additional modification of Moco with the attachment of GMP, forming the MGD cofactor (56).<sup>13</sup> In total, more than 10 genes are involved in the biosynthesis of Moco and the corresponding proteins were found to be highly conserved in other organisms (Figure 16). Each of the steps of Moco biosynthesis will be described in detail below and compared to the corresponding reaction by homologous proteins in humans (Figure 16).

#### 7.17.4.3.1 Conversion of guanosine monophosphate to precursor Z

The biosynthesis of Moco starts from 5'-GTP (Figures 16–18), which results in the formation of precursor Z (52), the first stable intermediate of Moco biosynthesis.<sup>14,278–280</sup> Precursor Z is an oxygen-sensitive 6-alkyl pterin with a cyclic phosphate group at the C2' and C4' atoms.<sup>14</sup> The pathways for the synthesis of BH<sub>4</sub>, THF, and riboflavin also start with the conversion of GTP, and for these cofactors the reaction is carried out by the enzymes GTPCH I and II.<sup>6,281</sup> In contrast, formation of precursor Z was shown to be GTPCH I and II independent. It was shown that the C8 atom of 5'-GTP is not released as formate but is retained and incorporated in a rearrangement reaction and inserted between the 2' and 3' ribose carbon atoms thus forming the four carbon atoms of the pyrano ring that is typical for MPT (57, Figure 18).<sup>14</sup> Precursor Z was initially isolated and characterized by Wuebbens and Rajagopalan,<sup>14</sup> and it was reported that precursor Z is relatively stable against oxygen at low pH, with a half-life of several hours. In contrast, BH<sub>4</sub> is more stable at physiological pH.<sup>282</sup> Another structure of precursor Z was proposed by Santamaria-Araujo *et al.*<sup>283</sup> Here, precursor Z was identified in its pyranopterin form containing a geminal diol at the C1' position of the side chain (Figure 18). It was proposed that the geminal diol might have a protective function for the molecule and could drive the subsequent sulfur transfer reaction in a particular direction. However, a recent report about the cocrystal structure of MPT synthase with precursor Z at the active site showed, that most likely, in the protein-bound state precursor Z is bound in the keto form.<sup>284</sup> In *E. coli*, the gene products MoaA and MoaC are essential for the biosynthesis of precursor Z (Figure 18). MoaA belongs to the *S*-adenosylmethionine (SAM)-dependent radical enzyme superfamily, members of which catalyze the formation of protein and/or substrate radicals by reductive cleavage of SAM by a [4Fe–4S] cluster.<sup>285</sup> MoaA, in fact, assembles two oxygen-sensitive [4Fe–4S] clusters, one typical for SAM-dependent radical enzymes at the N-terminus and an additional C-terminal cluster unique to MoaA proteins, as shown in the crystal structure of *S. aureus* MoaA.<sup>280</sup> The core of the protein is characterized by an incomplete ( $\alpha\beta$ )<sub>6</sub> TIM barrel type formed by the N-terminal part of the protein.<sup>280</sup> The structure of *E. coli* MoaC revealed that it is present as a hexamer composed of two dimers with a putative active site located at the dimer interface.<sup>286</sup> Since some radical SAM-dependent enzymes require another protein onto which the radical is transferred, it is believed that MoaC might act in a similar function. However, it has also been speculated that MoaC is involved in the cleavage of the pyrophosphate group, thus forming the cyclic phosphate group of precursor Z (52).<sup>287</sup>

The pathway of Moco biosynthesis is highly conserved among all organisms. Recent studies have identified the human genes involved in the biosynthesis of Moco (Figure 16). They code for proteins with high homologies to the bacterial congeners, but display differences at the gene level.<sup>288</sup> The *MOCS1* locus encodes two proteins, MOCS1A and MOCS1B, which are homologous to MoaA and MoaC, respectively. An unusual bicistronic transcript of the *MOCS1* locus was identified, with open reading frames for both *MOCS1A* and *MOCS1B* in a single transcript, separated by a stop codon.<sup>289</sup> Splice variants of the *MOCS1* locus were identified that bypass the termination codon

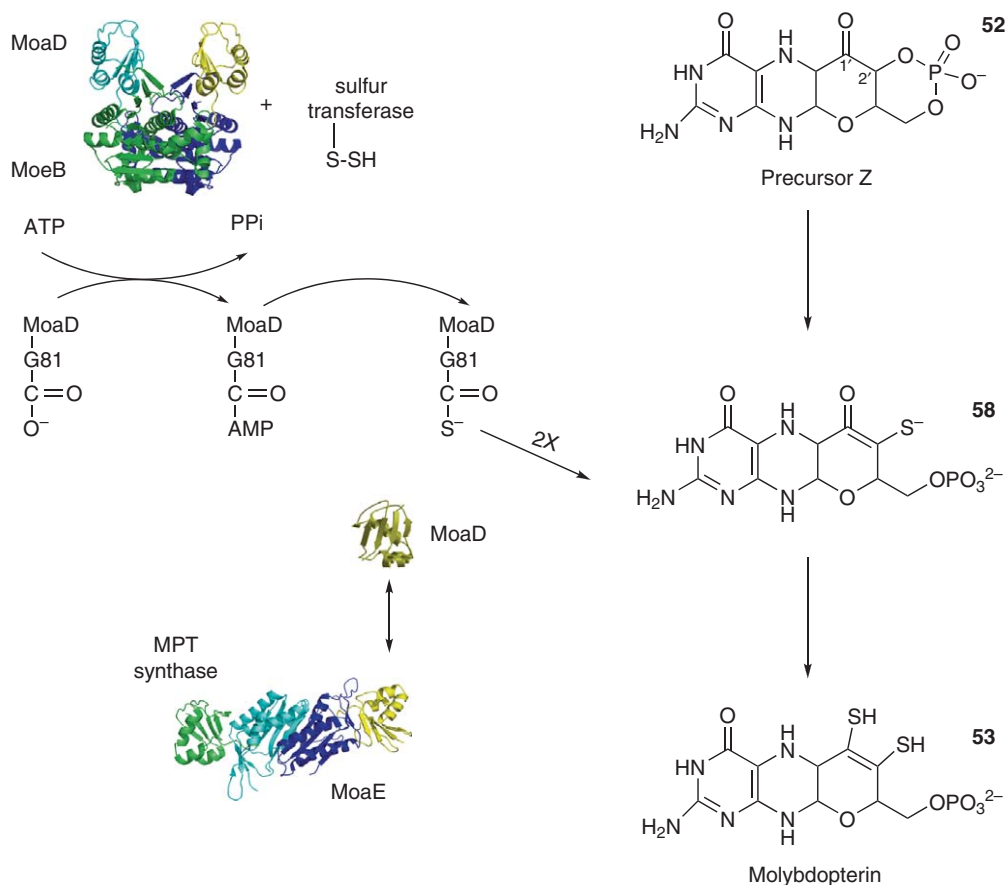


**Figure 18** Synthesis of precursor Z from GTP. All carbon atoms of the GTP are found within precursor Z (**52**). The C8 atom transferred as formyl is inserted between the C2' and C3' atoms of the ribose. This reaction is catalyzed by the MoaA protein, an S-adenosylmethionine (SAM)-dependent enzyme. MoaC is believed to cleave the pyrophosphate group of the intermediate. Precursor Z is shown in the tetrahydropyrano form and in a hydrated product with a geminal diol at the C1' position.

of *MOCS1A*, resulting in a new two-domain protein fusing *MOCS1A* and *MOCS1B*.<sup>290</sup> No evidence was found for the expression of *MOCS1B* from the bicistronic *MOCS1A–MOCS1B* splice-type I cDNA, indicating that *MOCS1B* is only expressed as a fusion with *MOCS1A*, while *MOCS1A* is also expressed as a separate protein.<sup>291</sup> However, the reaction catalyzed by *MOCS1A* and the *MOCS1A–MOCS1B* fusion protein occurs in the same manner as described for the *E. coli* proteins, converting GTP into precursor Z.<sup>287</sup>

#### 7.17.4.3.2 Insertion of sulfur into precursor Z and formation of molybdopterin

For the formation of MPT (**53**) from precursor Z (**52**), two sulfur atoms are incorporated in the C1' and C2' positions of precursor Z (**52**), a reaction catalyzed by MPT synthase<sup>15</sup> (**Figure 19**). The purification of MPT synthase identified a heterotetrameric enzyme, consisting of two small (~8750 Da) and two large subunits (~16850 Da), encoded by *moaD* and *moaE*, respectively.<sup>292,293</sup> Pitterle *et al.*<sup>15</sup> were able to establish an *in vitro* system for MPT synthesis, consisting solely of MPT synthase and precursor Z. It was shown that the small subunit of MPT synthase carries the sulfur in form of a thiocarboxylate at the C-terminal glycine of *MoaD* (**Figure 19**).<sup>294</sup> The high-resolution crystal structure of MPT synthase revealed that the two *MoaE* subunits form a central dimer, and the *MoaD* subunits are located at opposite ends of each dimer<sup>295</sup> (**Figure 19**). A pocket with highly conserved



**Figure 19** The MPT synthase reaction. Precursor Z (**52**) is converted into MPT (**53**) by the transfer of two sulfur groups from the C-terminal thiocarboxylate of the MoaD subunit of MPT synthase. For the regeneration of the sulfur on MoaD, a complex is formed with MoeB. ATP consumption yields adenylated MoaD. MoaD-AMP is susceptible to sulfuration by a protein-bound persulfide group from a sulfur transferase. After the formation of the thiocarboxylate group, MoaD dissociates from the MoeB-dimer and associates with MoaE. Initial attack by the first MoaD thiocarboxylate could occur at either the C1' or C2' position of precursor Z to produce a hemisulfurated precursor Z intermediate (**58**).

amino acids in MoaE can be seen in proximity to the C-terminus of MoaD, building the precursor Z binding site, as seen in the structure of *S. aureus* MPT synthase with bound precursor Z.<sup>284</sup>

The structure of the MoaD protein showed that it has a  $\beta$ -grasp fold characteristic for ubiquitin-like proteins.<sup>295</sup> From the structure of MPT synthase, the question arose whether the addition of the dithiolene sulfurs occurs independently at both active sites, or whether a hemisulfurated intermediate is transferred from one active site to the other for the addition of the second sulfur atom.<sup>294,296</sup> A model proposed by Wuebbens and Rajagopalan<sup>296</sup> favored the release of the first small subunit upon sulfur transfer and, subsequently, binding of a new thiocarboxylated MoaD molecule and transfer of the second sulfur. Purification of a hemisulfurated intermediate that is tightly associated with MPT synthase supported this model.<sup>296</sup> Each precursor Z molecule remains bound at a single active site until conversion of MPT is completed and an exchange of carboxylated and thiocarboxylated MoaD ensures the transfer of two sulfur atoms (Figure 19). In addition, the *S. aureus* MPT synthase–precursor Z structure suggested that the first sulfur is added at the C2' position of precursor Z, resulting in a hemisulfurated precursor Z intermediate (**58**), before the sulfur at the C1' position is added by another sulfurated MoaD subunit.<sup>284</sup>

For the synthase to act catalytically, it is necessary to regenerate its transferable sulfur. Mutant studies of *E. coli* showed the MoeB protein, the MPT synthase sulfurylase, is involved in this activation. It was shown that MoeB activates the C-terminus of MoaD by addition of an acyl-adenylate (Figure 19), a reaction similar to the first step of ubiquitin-targeted protein degradation.<sup>297</sup> In addition, MoeB shares significant sequence similarities to the



ubiquitin-activating enzyme E1.<sup>298,299</sup> In the second reaction, the activated MoaD acyl-adenylate is converted into a thiocarboxylate by action of a persulfide-containing protein. However, the interaction of MoeB with MoaD resembles only the first step of the ubiquitin-targeted protein degradation and a thioester intermediate was not identified between MoaD and MoeB.<sup>297,300</sup> In total, the phylogenetic distribution of the enzymes involved in Moco biosynthesis strongly suggests that the two-component ubiquitin activation systems present in eukaryotes are derived from the simpler and universally distributed MoaD–MoeB pair.<sup>301</sup> In *E. coli*, the sulfur donor for the dithiolene group of MPT is L-cysteine,<sup>302</sup> and the proteins involved in sulfur transfer to MPT synthase is the L-cysteine desulfurase IscS<sup>302</sup> in conjunction with a rhodanese-like protein (S. Leimkühler, unpublished results).

In humans, the *MOCS2* locus encodes the two subunits of MPT synthase, with *MOCS2A* being the MoaD homologue and *MOCS2B* being the MoaE homologue. In analogy to the *MOCS1* locus, *MOCS2A* and *MOCS2B* are transcribed in a bicistronic transcript, however, with overlapping reading frames.<sup>303</sup> *In vitro* translation and mutagenesis experiments showed that *MOCS2A* and *MOCS2B* are translated independently, leading to the synthesis of the 9.8 kDa *MOCS2A* protein and the 20.8 kDa *MOCS2B* protein.<sup>303</sup> Recent studies showed that the expression of the overlapping *MOCS2A* and *MOCS2B* genes is realized by two alternative *MOCS2* splice forms I and III, resulting in different first exons thus leading to alternative transcripts.<sup>304</sup> The human MPT synthase reaction was studied after purification of the separately expressed *MOCS2A* and *MOCS2B* subunits in *E. coli*.<sup>305</sup> Both subunits readily assembled into a functional MPT synthase, which was able to convert precursor Z into MPT *in vitro*. The fact that active hybrid proteins with the *E. coli* large or small subunits of MPT synthase also could be formed shows that the reaction is highly conserved in prokaryotes and eukaryotes.

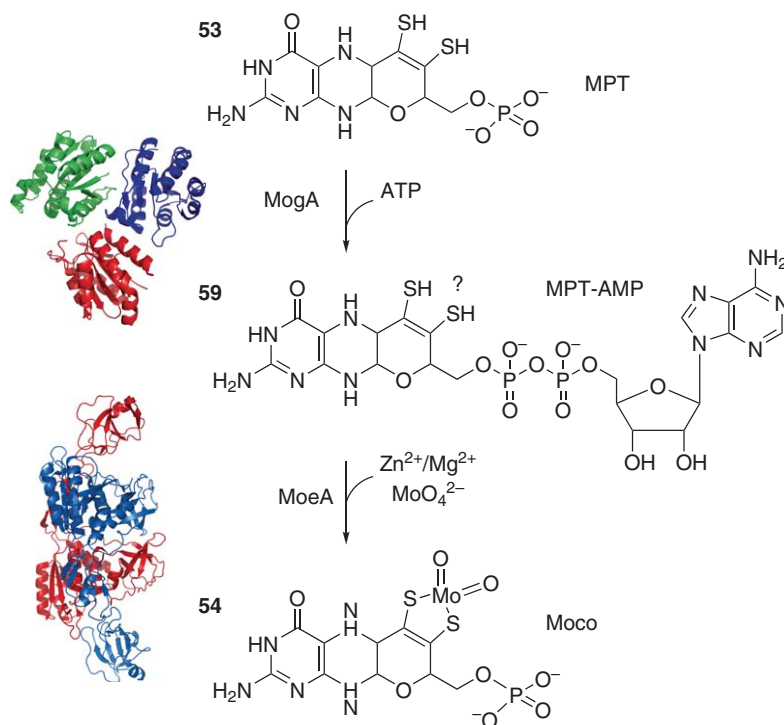
*MOCS3* is the homologue of MoeB in humans. Sequence alignments of *MOCS3* revealed that it is a two-domain protein with an N-terminal MoeB-like domain and an additional C-terminal domain showing homologies to rhodanases.<sup>306</sup> Rhodanases (thiosulfate:cyanide sulfur transferases) are widespread enzymes that catalyze the transfer of a sulfane sulfur from thiosulfate to cyanide *in vitro*.<sup>307</sup> The first physiological role of a rhodanese-like protein was identified for the *MOCS3* protein, showing that it acts as direct sulfur donor for the formation of the thiocarboxylate group on *MOCS2A*.<sup>306</sup> A persulfide group was identified on the highly conserved cysteine residue Cys412 of the six amino acid active site loop of the *MOCS3* C-terminal rhodanese domain.<sup>308</sup> In total, *MOCS3* can be regarded as a multifunctional and multidomain protein combining the adenylation of *MOCS2A* with the subsequent sulfur transfer reaction. However, it was shown *in vitro* that the sulfur for *MOCS3* originates from the L-cysteine desulfurase Nfs1 in humans.<sup>309,310</sup>

#### 7.17.4.3.3 Insertion of molybdenum into molybdopterin

After synthesis of the dithiolene moiety in MPT (53), the chemical backbone is built for binding and coordination of the molybdenum atom. Molybdenum enters the cell as the soluble oxyanion molybdate for which high-affinity transporters have been described in bacteria<sup>311</sup> which also exist in higher eukaryotes such as *Chlamydomonas reinhardtii*<sup>312</sup> and *A. thaliana*.<sup>313</sup>

In the step of molybdenum insertion in *E. coli*, the gene products of *moeA* and *mogA* are involved. MogA was the first protein of Moco biosynthesis for which the crystal structure was solved.<sup>314</sup> The crystal structure of MoeA showed a two-domain structure, with one domain being structurally related to MogA.<sup>315,316</sup> It was suggested that *E. coli* MoeA and MogA proteins are both essential for the incorporation of molybdenum in the cofactor, thus indicating a multistep reaction of molybdenum chelation.<sup>317</sup> However, the proteins were found to have different functions: It was observed that MoeA mediates molybdenum ligation to newly synthesized MPT at low concentrations of  $\text{MoO}_4^{2-}$ , while MogA helps to facilitate this step *in vivo* in an ATP-dependent manner, through an MPT-adenylate intermediate.<sup>318</sup> These results were supported by the identification of MPT and AMP bound to the crystal structure of the *A. thaliana* Cnx1-G domain.<sup>319</sup> In *A. thaliana*, the Cnx1 protein contains a C-terminal domain with homology to *E. coli* MogA and an N-terminal domain with homology to *E. coli* MoeA.<sup>320</sup> Besides defining the MPT-binding site, a novel adenylated form of MPT (MPT-AMP) was identified (59, Figure 20).

It was shown that the Cnx1 G-domain catalyzed this adenylation of MPT in a reaction dependent on  $\text{Mg}^{2+}$  and ATP.<sup>321,322</sup> The Cnx1 E-domain is required for the subsequent  $\text{Mg}^{2+}$ -dependent molybdenum insertion. In addition, the crystal structure of Cnx1-G-MPT-AMP complex showed that copper was bound to the MPT dithiolene sulfurs.<sup>319</sup> Copper was proposed to protect the dithiolene group prior to molybdenum insertion. However, since several metal ions are known to bind tightly to sulfur-rich sites, a more recent report by Morrison *et al.*<sup>323</sup> investigated the effect of copper-limiting reaction conditions on molybdoenzymes in *E. coli* and *R. sphaeroides*. Their results



**Figure 20** Insertion of molybdate into MPT. The *E. coli* MoeA<sup>315,316</sup> and MogA<sup>314</sup> proteins catalyze the specific incorporation of molybdenum into MPT in a multistep reaction with an adenylylated MPT intermediate (MPT-AMP, **59**), which was first identified in the *A. thaliana* Cnx1-G protein.<sup>319</sup> While MogA forms the MPT-adenylate intermediate,<sup>318</sup> MoeA mediates molybdenum ligation to MPT at low concentrations of  $\text{MoO}_4^{2-}$  in a  $\text{Zn}^{2+}$ - and  $\text{Mg}^{2+}$ -dependent manner.<sup>321,322</sup> The involvement of copper to protect the dithiolene sulfurs is unclear.<sup>319,323</sup>

demonstrated that the activities of DMSO reductase and nitrate reductase were not repressed under copper starvation, showing that copper is not strictly required for the biosynthesis of Moco in bacteria.<sup>323</sup> In another report it was shown that copper ( $\text{Cu}^{2+}$ ) and other bivalent ions can be inserted unspecifically into MPT present in heterologously expressed human sulfite oxidase in *E. coli*.<sup>324</sup> The *in vivo* insertion of these metals into the dithiolene group of MPT was shown to be independent of the function of MogA and MoeA in *E. coli*.<sup>324</sup>

In addition, it was shown that active Moco (**54**) can be synthesized *in vitro* by ligation of molybdate to MPT without the requirement of any additional factors.<sup>325</sup> Under the molybdate concentrations used in the assay ( $>5 \text{ mmol l}^{-1}$ ), the involvement of MoeA and MogA was not required for generating the active form of Moco.<sup>325</sup> This shows that the ATP-dependent activation of MPT and molybdenum is only required under physiological concentrations and that *in vitro* copper is not required to protect the dithiolene group of MPT. Apart from the possibility that copper is not physiologically relevant for Moco biosynthesis, it remains possible that there is a difference between prokaryotes and eukaryotes in the requirement of copper during Moco biosynthesis.<sup>323</sup>

In humans, the gephyrin gene encodes a protein with homologies of its N-terminal domain to *E. coli* MogA and its C-terminal domain to MoeA.<sup>326</sup> Apart from the molybdenum cofactor biosynthesis, gephyrin is also involved in synaptic anchoring of inhibitory ligand-gated ion channels. Alternative splicing has been proposed to contribute to gephyrin's functional diversity within the cell.<sup>327</sup>

#### 7.17.4.4 Additional Modification of Molybdenum Cofactor

##### 7.17.4.4.1 Attachment of guanosine monophosphate

Enzymes of the sulfite oxidase family, such as human sulfite oxidase, plant nitrate reductase, or the *E. coli* YedY protein (sulfite oxidase homologue) bind the MPT form of Moco (**54**) without further modification.<sup>258</sup> So far, YedY represents the only molybdoenzyme isolated from *E. coli* that is characterized by the presence of the MPT

form of Moco, however, its function remains unclear.<sup>258,328</sup> In general, Moco formed in *E. coli* is further modified by covalent addition of GMP to the C4' atom of MPT through a pyrophosphate bond.<sup>329</sup> In *E. coli*, the GMP attachment to Moco is catalyzed by the MobA and MobB proteins.<sup>330</sup> Although MobA was shown to be essential for this reaction,<sup>331</sup> the role of MobB remains uncertain. From the crystal structure, it was postulated that MobB is an adapter protein, acting in concert with MobA to achieve the efficient biosynthesis and utilization of MGD.<sup>332</sup> *In vitro*, MobA, GTP, MgCl<sub>2</sub>, and Mo-MPT are sufficient for the formation and insertion of bis-MGD into *R. sphaeroides* DMSO reductase.<sup>333</sup> The crystal structure of monomeric MobA showed a nucleotide-binding Rossmann fold formed by the N-terminal half of the protein.<sup>334,335</sup> It was shown that addition of the dinucleotide to the cofactor occurs after insertion of molybdenum into MPT.<sup>336</sup> In other prokaryotes, variants of the cofactor containing CMP, AMP, or IMP linked to MPT were identified.<sup>329,337</sup> It was shown that in all MGD-containing molybdoenzymes of *E. coli*, the molybdenum atom is coordinated by the dithiolene groups of two MGD molecules, forming the bis-MGD cofactor (56).<sup>338</sup> The formation of bis-MGD is one of the most enigmatic steps in Moco biosynthesis. It is not known whether the two MGD molecules assemble on MobA or instead after the insertion into the target proteins such as DMSO reductase or nitrate reductase A. However, it was shown that molybdenum ligation to MPT has to precede MGD formation.<sup>336</sup> In contrast, eukaryotes solely incorporate the MPT form of the cofactor into their molybdoenzymes without attached nucleotides.

#### 7.17.4.4.2 Formation of sulfurated molybdenum cofactor

Apart from the attachment of additional nucleotides in prokaryotes, Moco can be further modified by an exchange of the equatorial oxygen ligand by sulfur, forming the sulfurated or mono-oxo form of Moco (55).<sup>10</sup> Sulfurated Moco (55) is a form of the cofactor, which is present in all members of the XO family. In eukaryotes, it was shown that the sulfido ligand is added by a specific Moco sulfurase.<sup>339</sup> In general, eukaryotic Moco sulfurases are two-domain proteins with an N-terminal domain showing homologies to bacterial L-cysteine desulfurases (EC 2.1.8.7) of the NifS family and a C-terminal domain with no homologues in prokaryotes.<sup>340</sup> The currently best-characterized Moco sulfurase is ABA3 from *A. thaliana*.<sup>341</sup> L-Cysteine desulfurase activity and sulfur transferase activity of ABA3 were shown to be essential for the activity of aldehyde oxidase (AO), being involved in abscisic acid (ABA) biosynthesis in plants. Bittner *et al.*<sup>342</sup> proposed that after incorporation of Moco, AO and XDH remain inactive until one of the two oxygen ligands is replaced by a sulfur atom delivered by ABA3. By this mechanism, it was suggested that the plant is able to rapidly increase the activities of AO and XDH without *de novo* synthesis of the apoproteins.<sup>342</sup> Studies on recombinant ABA3 showed, that the C-terminal domain binds the sulfurated form of Moco.<sup>343</sup> Moco is most likely sulfurated by the cysteine desulfurase activity of the N-terminal domain.<sup>343</sup> It has been suggested that the role of ABA3 is rather the posttranslational sulfuration of Moco-containing holoenzymes of the XO family rather than the initial insertion of sulfurated Moco.<sup>343</sup> So far, no prokaryotic protein with amino acid sequence identities to both the C-terminal and N-terminal domain of Moco sulfurases has been identified. However, a specific protein involved in the insertion of the Mo=S ligand of Moco was described from *R. capsulatus*, named XdhC.<sup>344</sup> The *xdbc* gene is cotranscribed with *xdbA* and *xdbB*, the structural genes for xanthine dehydrogenase (XDH), and XdhC was found to be essential for XDH activity.<sup>344</sup> XdhC binds stoichiometric amounts of Moco, and is further able to transfer its bound Moco to Moco-free apo-XDH by a specific interaction with the XdhB subunit.<sup>345</sup> XdhC particularly stabilizes the sulfurated form of Moco before the insertion into XDH.<sup>345</sup> The L-cysteine desulfurase NifS4 of *R. capsulatus* was identified to specifically interact with XdhC and sulfurates Moco although it is bound to the latter.<sup>346</sup> Thus, the XdhC–NifS4 pair can be considered as the prokaryotic counterpart to specific Moco sulfurases identified in eukaryotes. Because it was clearly shown that sulfuration of Moco occurs before its insertion into XDH,<sup>346</sup> a regulation of enzyme activity through its sulfuration level does not occur in prokaryotes, another difference to the eukaryotic system.

#### 7.17.4.5 The Storage of Molybdenum Cofactor and Its Insertion into Molybdoenzymes

##### 7.17.4.5.1 Chaperones for the insertion and targeting of bis-molybdopterin guanine dinucleotide and sulfurated molybdenum cofactor into prokaryotic molybdoenzymes

The crystal structures of several molybdoenzymes revealed that Moco is deeply buried inside the proteins, at the end of a funnel-shaped passage giving access only to the substrate.<sup>253</sup> This implied the requirement of specific chaperones for each molybdoenzyme, to facilitate the insertion of Moco. For *R. capsulatus* XDH, the

XdhC protein has been identified to be not only involved in Moco binding and formation of the terminal sulfido ligand, but was also shown to be involved in the maturation of apo-XDH after Moco insertion.<sup>344</sup> In prokaryotes, a number of specific chaperones were identified, such as TorD for *E. coli* trimethylamine oxide (TMAO) reductase (TorA) or NarJ for *E. coli* nitrate reductase A (NarGHI).<sup>347–349</sup> It was shown that TorD or NarJ stabilize a certain conformation of their target proteins. TorD was shown to interact with TorA to generate a form of the protein competent to accept a bis-MGD and to prevent the export of immature TorA to the periplasm.<sup>350–352</sup> NarJ is known to bind on two distinct sites of the aponitrate reductase: It facilitates the correct membrane anchoring by avoiding the anchoring of an immature NarGH complex,<sup>353</sup> and NarJ allows Moco insertion through binding to NarG and the Moco biosynthesis proteins.<sup>354</sup> Thus the chaperones NarJ, TorD, and XdhC are involved in the insertion of Moco and/or in facilitating the correct folding of the corresponding Mo enzyme, but so far, a binding of Moco has been demonstrated only for XdhC.<sup>345</sup> Studies on XdhC showed that the protein is also involved in the distribution and targeting of Moco in the cell.<sup>336</sup> For bacteria, it was shown that the pathways of bis-MGD and Moco biosynthesis diverge after the insertion of the molybdenum atom into MPT. One part of the Moco pool is inserted into molybdoenzymes binding the MPT form of Moco, the other part is further modified by MobA,<sup>331</sup> adding an additional nucleotide and forming the bis-MGD cofactor. For enzymes such as XDH, Moco has to be further modified by exchanging one oxo ligand of the Mo center by sulfur, building the equatorial Mo=S ligand. To prevent the available Moco in the cell being targeted during MGD biosynthesis, XdhC interacts with MobA and thereby prevents the binding of Moco to MobA, simultaneously inhibiting MGD formation.<sup>336</sup> It is likely that other proteins share a similar function with XdhC, regulating and directing Moco targeting by binding different variants of Moco.

The presence of system-specific chaperones involved in Moco insertion in molybdoenzymes seems to be restricted to prokaryotes, since no chaperone so far has been identified for eukaryotic molybdoenzymes. However, oxygen seems to be a major factor of free Moco inactivation. Most likely, Moco stays protein-bound during its biosynthesis until it is inserted into the specific target protein.

#### **7.17.4.5.2 Molybdenum cofactor carrier proteins in eukaryotes**

In *C. reinhardtii* a Moco carrier protein (MocoCP) has been identified.<sup>355</sup> The homotetrameric MocoCP was able to bind and transfer Moco directly to aponitrate reductase and protected Moco against inactivation under aerobic conditions. The direct interaction of MocoCP with nitrate reductase was demonstrated.<sup>356</sup> The crystal structure of the protein showed a homotetrameric organization<sup>357</sup> and revealed that the monomers are arranged in a Rossmann-like fold. Biochemical evidence showed that MCP discriminates between selective binding of Moco/Wco and MPT. MocoCP-bound Moco and Wco were significantly stabilized against degradation.<sup>357</sup> A protein with similar characteristics was also identified in *Vicia faba*.<sup>358</sup> Among higher plants, proteins with homologies to *C. reinhardtii* MCP were found that form a multigene family in *A. thaliana*.<sup>359</sup> The availability of sufficient amounts of Moco is essential for eukaryotes to meet the changing demand of molybdenum enzymes. In contrast to the regulated plant nitrate reductase, the enzymes of the Moco biosynthesis are constitutively expressed at basal levels.<sup>360</sup> Thus, the existence of MCPs would provide a way to ensure storage and basal supply of Moco in eukaryotes. So far, there is no indication that such proteins exist in prokaryotes.

#### **7.17.4.6 The Importance of Molybdenum Cofactor in Humans**

##### **7.17.4.6.1 Molybdenum cofactor deficiency and isolated sulfite oxidase deficiency**

Moco is essential for the activity of sulfite oxidase, XDH, and AO, the three molybdoenzymes present in humans.<sup>288</sup> Human Moco deficiency leads to the pleiotropic loss of all three of these molybdoenzymes and usually progresses to death at an early age.<sup>361</sup> In humans, a combined deficiency of sulfite oxidase and XDH was first described by Duran *et al.*<sup>362</sup> To date, more than 100 cases of Moco deficiency or isolated sulfite oxidase deficiency are known worldwide.<sup>16</sup> Isolated sulfite oxidase deficiency is a related disease in which molybdenum cofactor biosynthesis is normal, but sulfite oxidase activity is altered.<sup>277</sup> The clinical symptoms of Moco deficiency are indistinguishable from those of isolated sulfite oxidase deficiency with a notable exception that xanthinuria is absent in the latter.<sup>277</sup> In both cases, affected neonates show feeding difficulties, neurological abnormalities such as attenuated brain growth, untreatable seizures, dislocated ocular lenses in most cases, and death in early childhood. Although milder symptoms are occasionally observed, none of the treatments tested

to date has produced consistent improvement of the clinical status.<sup>363</sup> Affected individuals normally come to clinical attention shortly after birth because of a failure to thrive, and seizures that are often unresponsive to treatment. Later, if they survive the neonatal period, progressive loss of white matter in the brain, feeding difficulties, and dislocated ocular lenses are observed.<sup>363</sup> The clinical symptoms may result from sulfite toxicity, tissue-specific sulfate deficiency, or a combination of both. Elevated urinary sulfite can be detected using dipsticks; however, a more reliable diagnostic test measures urinary *S*-sulfo-cysteine by mass spectrometry.<sup>364</sup> Based on the presence or absence of precursor Z (52) in clinical samples, Moco deficiency was originally divided into two complementation groups.<sup>365</sup> Group A patients have a defect in the *MOCS1* locus and, therefore, do not produce precursor Z, whereas group B patients have a defect in the *MOCS2* locus and are characterized by the accumulation of precursor Z.<sup>365</sup> Recently, a third group has been identified with genetic defects in gephyrin.<sup>366</sup>

After pinpointing the genes essential for human Moco biosynthesis, the identification of mutations became possible. Moreover, mutations in the *MOCS1* gene were already identified using prenatal diagnosis in cases with a family history in Moco deficiency.<sup>367</sup>

#### 7.17.4.6.2 Animal models and first therapies for molybdenum cofactor deficiency

A mouse model was generated with a defect in the first step of the biosynthesis.<sup>368</sup> The mouse homologue of *MOCS1A* was disrupted by homologous recombination with a targeting vector. Similar to humans, heterozygous mice displayed no symptoms, but homozygous animals died between days 1 and 11 after birth.<sup>368</sup> In contrast to other disorders of pterin metabolism, such as the deficiency of BH<sub>4</sub>, it is difficult to treat Moco deficiency by supplementation with the cofactor, as Moco is extremely unstable in aqueous solutions with neutral pH.<sup>369</sup> Therefore, a method was established to use precursor Z purified from *E. coli* for a substitution therapy of the group A deficient mice.<sup>370</sup> Repeated injections of precursor Z into these mice resulted in a dose-dependent extension of life span,<sup>370</sup> however, this substitution therapy required frequent and lifelong injections of precursor Z because of the short half-life of both precursor Z and the respective molybdoenzymes. Thus, a somatic gene therapy for Moco deficiency was developed. An expression cassette suitable for long-term expression of both *MOCS1A* and *MOCS1B* was constructed in an adenovirus-associated virus (AAV) vector.<sup>371</sup> The vector was used to transduce *MOCS1*-deficient mice after both 1 and 4 days of birth or, after pretreatment with purified precursor Z, after 40 days of birth. The AAV-*MOCS1*-transduced animals showed significantly increased longevity, thus, a single intrahepatic injection of AAV-*MOCS1* into *MOCS1*-deficient mice resulted in fertile adult animals without any pathological phenotypes.<sup>371</sup> As only two-thirds of the known Moco-deficient patients belong to deficiency group A,<sup>372</sup> a therapy for the treatment of the other complementation groups has to be developed. This should be straightforward using the gene therapy approach. However, it remains to be seen whether this therapy will allow reversal of the neurological damage that is observed in many human patients, sometimes already at birth.

## Acknowledgments

The authors thank Drs. Nenad Blau and Aurora Martinez for valuable discussions. This work was supported by the Deutsche Forschungsgemeinschaft (Silke Leimkühler and Markus Fischer) the Cluster of Excellence “Unifying Concepts in Catalysis” coordinated by the Technische Universität Berlin (Silke Leimkühler) and the Swiss National Science Foundation (Beat Thöny).

### Abbreviations

<b>AICAR</b>	4-amino-5-imidazolecarboxamidribotide
<b>AKR1C3</b>	3 $\alpha$ -hydroxysteroid dehydrogenase type 2
<b>AO</b>	aldehyde oxidase
<b>AR</b>	aldose reductase
<b>ATP</b>	adenosine triphosphate
<b>BH<sub>4</sub></b>	(6 <i>R</i> )-L-erythro-5,6,7,8-tetrahydrobiopterin



<b>CO</b>	carbon monoxide
<b>CR</b>	carbonyl reductases
<b>CSF</b>	cerebrospinal fluid
<b>DCoH</b>	dimerization cofactor
<b>DHFR</b>	dihydrofolate reductase
<b>DHFS</b>	dihydrofolate synthetase
<b>DHNA</b>	dihydroneopterin aldolase
<b>DHNTP</b>	dihydroneopterin triphosphate
<b>DHPR</b>	dihydropteridine reductase
<b>DHPS</b>	dihydropteroate synthase
<b>DMSO</b>	dimethylsulfoxide
<b>DRD</b>	DOPA-responsive dystonia
<b>EGF</b>	epidermal factor
<b>ENU</b>	ethylnitrosourea
<b>FDTS</b>	flavin-dependent thymidylate synthetase
<b>GAR</b>	glycinamideribotide
<b>GEMO</b>	glyceryl-ether monooxygenase
<b>GFRP</b>	GTP cyclohydrolase feedback regulatory protein
<b>GMP</b>	guanosine monophosphate
<b>GTP</b>	guanosine triphosphate
<b>GTPCH</b>	GTP cyclohydrolase I
<b>H<sub>2</sub>NTP</b>	7,8-dihydroneopterin triphosphate
<b>HNF-1<math>\alpha</math></b>	hepatocyte nuclear factor 1 $\alpha$
<b>HPPK</b>	6-hydroxymethyl-dihydropterin pyrophosphokinase
<b>IGF-1</b>	insulin-like growth factor-1
<b>MCD</b>	molybdopterin-cytosine dinucleotide cofactor
<b>MGD</b>	molybdopterin guanine dinucleotide cofactor
<b>Moco</b>	molybdenum cofactor
<b>MPT</b>	molybdopterin
<b>MPT-AMP</b>	adenylylated MPT
<b>NGF</b>	nerve growth factor
<b>NOS</b>	nitric oxide synthase
<b>pABA</b>	<i>p</i> -aminobenzoate
<b>PAH</b>	phenylalanine-4-hydroxylase
<b>PC12</b>	pheochromocytoma-12
<b>PCD</b>	pterin-4a-carbinolamine dehydratase
<b>PCD/DCoH</b>	pterin-4a-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor-1 $\alpha$
<b>PKU</b>	phenylketonuria
<b>PTP</b>	6-pyruvoyl-5,6,7,8-tetrahydropterin
<b>PTPS</b>	6-pyruvoyl tetrahydropterin synthase
<b>SAM</b>	S-adenosylmethionine
<b>SR</b>	sepiapterin reductase
<b>TH</b>	tyrosine-3-hydroxylase
<b>THF</b>	tetrahydrofolate
<b>TPH</b>	tryptophan-5-hydroxylase
<b>TS</b>	thymidylate synthetase
<b>XDH</b>	xanthine dehydrogenase
<b>XO</b>	xanthine oxidase



## References

1. S. Kaufman, *Annu. Rev. Biochem.* **1967**, *36*, 171–184.
2. F. G. Hopkins, *Nature* **1889**, *40*, 335.
3. C. A. Nichol; G. K. Smith; D. S. Duch, *Annu. Rev. Biochem.* **1985**, *54*, 729–764.
4. D. R. Appling, *FASEB J.* **1991**, *5* (12), 2645–2651.
5. B. C. Blount; M. M. Mack; C. M. Wehr; J. T. MacGregor; R. A. Hiatt; G. Wang; S. N. Wickramasinghe; R. B. Everson; B. N. Ames, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94* (7), 3290–3295.
6. B. Thöny; G. Auerbach; N. Blau, *Biochem. J.* **2000**, *347* (Pt. 1), 1–16.
7. G. Werner-Felmayer; G. Golderer; E. R. Werner, *Curr. Drug. Metab.* **2002**, *3* (2), 159–173.
8. K. V. Rajagopalan; J. L. Johnson, *J. Biol. Chem.* **1992**, *267*, 10199–10202.
9. J. L. Johnson; B. E. Hainline; K. V. Rajagopalan; B. H. Arison, *J. Biol. Chem.* **1984**, *259*, 5414–5422.
10. R. Hille, *Chem. Rev.* **1996**, *96*, 2757–2816.
11. J. L. Johnson; K. V. Rajagopalan, The Oxidation of Sulphite in Animal Systems. In *Sulphur in Biology*, Ciba Foundation Symposium 72; Excerpta Medica: Amsterdam, 1980; pp 119–133.
12. J. L. Johnson; K. V. Rajagopalan, *J. Clin. Invest.* **1976**, *58*, 551–556.
13. K. V. Rajagopalan, Biosynthesis of the Molybdenum Cofactor. In *Escherichia coli and Salmonella. Cellular and Molecular Biology*; F. C. Neidhardt, Ed.; ASM Press: Washington, DC, 1996; Vol. I, pp 674–679.
14. M. M. Wuebbens; K. V. Rajagopalan, *J. Biol. Chem.* **1993**, *268*, 13493–13498.
15. D. M. Pitterle; J. L. Johnson; K. V. Rajagopalan, *J. Biol. Chem.* **1993**, *268*, 13506–13509.
16. G. Schwarz, *Cell. Mol. Life Sci.* **2005**, *62* (23), 2792–2810.
17. G. H. Hitchings; J. J. Burchall, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1965**, *27*, 417–468.
18. B. I. Schweitzer; A. P. Dicker; J. R. Bertino, *FASEB J.* **1990**, *4* (8), 2441–2452.
19. N. C. Martin, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (3), 1110–1112.
20. N. Kanaan; S. Marti; V. Moliner; A. Kohen, *Biochemistry* **2007**, *46* (12), 3704–3713.
21. T. A. Fritz; L. Liu; J. S. Finer-Moore; R. M. Stroud, *Biochemistry* **2002**, *41* (22), 7021–7029.
22. A. Sancar, *Biochemistry* **1994**, *33* (1), 2–9.
23. Z. He; K. D. Stigers Lavoie; P. A. Bartlett; M. D. Toney, *J. Am. Chem. Soc.* **2004**, *126* (8), 2378–2385.
24. J. F. Parsons; P. Y. Jensen; A. S. Pachikara; A. J. Howard; E. Eisenstein; J. E. Ladner, *Biochemistry* **2002**, *41* (7), 2198–2208.
25. G. J. Basset; E. P. Quinlivan; S. Ravanel; F. Rebeille; B. P. Nichols; K. Shinozaki; M. Seki; L. C. Adams-Phillips; J. J. Giovannoni; J. F. Gregory III; A. D. Hanson, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (6), 1496–1501.
26. J. M. Green; W. K. Merkel; B. P. Nichols, *J. Bacteriol.* **1992**, *174* (16), 5317–5323.
27. T. Nakai; H. Mizutani; I. Miyahara; K. Hirotsu; S. Takeda; K. H. Jhee; T. Yoshimura; N. Esaki, *J. Biochem.* **2000**, *128* (1), 29–38.
28. J. Botet; L. Mateos; J. L. Revuelta; M. A. Santos, *Eukaryot. Cell.* **2007**, *6* (11), 2102–2111.
29. G. M. Brown; J. M. Williamson, Biosynthesis of Folic Acid, Riboflavin, Thiamine, and Pantothenic Acid. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*; F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, H. E. Umbarger, Eds.; Washington, DC: American Society for Microbiology 1987; Vol. 1, pp 521–538.
30. F. Foor; G. M. Brown, *J. Biol. Chem.* **1975**, *250* (9), 3545–3551.
31. D. E. Graham; H. Xu; R. H. White, *Biochemistry* **2002**, *41* (50), 15074–15084.
32. G. Auerbach; A. Herrmann; A. Bracher; G. Bader; M. Gutlich; M. Fischer; M. Neukamm; M. Garrido-Franco; J. Richardson; H. Nar; R. Huber; A. Bacher, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (25), 13567–13572.
33. N. Maita; K. Hatakeyama; K. Okada; T. Hakoshima, *J. Biol. Chem.* **2004**, *279* (49), 51534–51540.
34. H. Nar; R. Huber; W. Meining; C. Schmid; S. Weinkauff; A. Bacher, *Structure* **1995**, *3* (5), 459–466.
35. G. Basset; E. P. Quinlivan; M. J. Ziemak; R. Diaz De La Garza; M. Fischer; S. Schiffmann; A. Bacher; J. F. Gregory III; A. D. Hanson, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (19), 12489–12494.
36. H. Nar; R. Huber; G. Auerbach; M. Fischer; C. Hosl; H. Ritz; A. Bracher; W. Meining; S. Eberhardt; A. Bacher, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92* (26), 12120–12125.
37. Y. Tanaka; N. Nakagawa; S. Kuramitsu; S. Yokoyama; R. Masui, *J. Biochem.* **2005**, *138* (3), 263–275.
38. J. Rebelo; G. Auerbach; G. Bader; A. Bracher; H. Nar; C. Hosl; N. Schramek; J. Kaiser; A. Bacher; R. Huber; M. Fischer, *J. Mol. Biol.* **2003**, *326* (2) 503–516.
39. N. Schramek; A. Bracher; M. Fischer; G. Auerbach; H. Nar; R. Huber; A. Bacher, *J. Mol. Biol.* **2002**, *316* (3), 829–837.
40. A. Bracher; M. Fischer; W. Eisenreich; H. Ritz; N. Schramek; P. Boyle; P. Gentili; R. Huber; H. Nar; G. Auerbach; A. Bacher, *J. Biol. Chem.* **1999**, *274* (24), 16727–16735.
41. A. Bracher; W. Eisenreich; N. Schramek; H. Ritz; E. Gotze; A. Herrmann; M. Gutlich; A. Bacher, *J. Biol. Chem.* **1998**, *273* (43), 28132–28141.
42. B. El Yacoubi; S. Bonnett; J. N. Anderson; M. A. Swairjo; D. Iwata-Reuyl; V. de Crecy-Lagard, *J. Biol. Chem.* **2006**, *281* (49), 37586–37593.
43. Y. Suzuki; G. M. Brown, *J. Biol. Chem.* **1974**, *249* (8), 2405–2410.
44. A. De Saizieu; P. Vankan; A. P. van Loon, *Biochem. J.* **1995**, *306* (Pt. 2), 371–377.
45. S. M. Klaus; A. Wegkamp; W. Sybesma; J. Hugenholtz; J. F. Gregory III; A. D. Hanson, *J. Biol. Chem.* **2005**, *280* (7), 5274–5280.
46. M. Fischer; A. Bacher, *Nat. Prod. Rep.* **2005**, *22* (3), 324–350.
47. J. B. Perkins; J. G. Pero; A. Sloma, Riboflavin Overproducing Bacteria Expressing the Rib Operon of *Bacillus*. *Eur. Pat. Appl.*, EP 405370 A1 910102, 1991.
48. K. P. Stahmann; J. L. Revuelta; H. Seulberger, *Appl. Microbiol. Biotechnol.* **2000**, *53* (5), 509–516.
49. J. B. Mathis; G. M. Brown, *J. Biol. Chem.* **1970**, *245* (11), 3015–3025.
50. J. B. Mathis; G. M. Brown, *Meth. Enzymol.* **1980**, *66*, 556–560.
51. T. Ploom; C. Haussmann; P. Hof; S. Steinbacher; A. Bacher; J. Richardson; R. Huber, *Structure* **1999**, *7* (5), 509–516.
52. H. Nar; R. Huber; C. W. Heizmann; B. Thöny; D. Burgisser, *EMBO J.* **1994**, *13* (6), 1255–1262.

53. T. Ploom; B. Thöny; J. Yim; S. Lee; H. Nar; W. Leimbacher; J. Richardson; R. Huber; G. Auerbach, *J. Mol. Biol.* **1999**, 286 (3), 851–860.
54. N. Colloc'h; M. el Hajji; B. Bachet; G. L'Hermite; M. Schiltz; T. Prange; B. Castro; J. P. Mornon, *Nat. Struct. Biol.* **1997**, 4 (11), 947–952.
55. S. Bauer; A. K. Schott; V. Illarionova; A. Bacher; R. Huber; M. Fischer, *J. Mol. Biol.* **2004**, 339 (4), 967–979.
56. C. W. Goulding; M. I. Apostol; M. R. Sawaya; M. Phillips; A. Parseghian; D. Eisenberg, *J. Mol. Biol.* **2005**, 349 (1), 61–72.
57. M. Hennig; A. D'Arcy; I. C. Hampele; M. G. Page; C. Oefner; G. E. Dale, *Nat. Struct. Biol.* **1998**, 5 (5), 357–362.
58. A. Garcon; C. Levy; J. P. Derrick, *J. Mol. Biol.* **2006**, 360 (3), 644–653.
59. Y. Wang; Y. Li; H. Yan, *Biochemistry* **2006**, 45 (51), 15232–15239.
60. T. Gefflaut; C. Blonski; J. Perie; M. Willson, *Prog. Biophys. Mol. Biol.* **1995**, 63 (3), 301–340.
61. W. Gross; M. G. Bayer; C. Schnarrenberger; U. B. Gebhart; T. L. Maier; H. Schenk, *Plant Physiol.* **1994**, 105 (4), 1393–1398.
62. G. M. Smith; A. S. Mildvan; E. T. Harper, *Biochemistry* **1980**, 19 (6), 1248–1255.
63. B. S. Szwegold; K. Ugurbil; T. R. Brown, *Arch. Biochem. Biophys.* **1995**, 317 (1), 244–252.
64. R. Riek; K. Pervushin; K. Wuthrich, *Trends Biochem. Sci.* **2000**, 25 (10), 462–468.
65. C. Haussmann; F. Rohdich; F. Lottspeich; S. Eberhardt; J. Scheuring; S. Mackamul; A. Bacher, *J. Bacteriol.* **1997**, 179 (3), 949–951.
66. C. Haussmann; F. Rohdich; E. Schmidt; A. Bacher; G. Richter, *J. Biol. Chem.* **1998**, 273 (28), 17418–17424.
67. K. Ikemoto; T. Sugimoto; S. Murata; M. Tazawa; T. Nomura; H. Ichinose; T. Nagatsu, *Biol. Chem.* **2002**, 383 (2), 325–330.
68. D. P. Richey; G. M. Brown, *J. Biol. Chem.* **1969**, 244 (6), 1582–1592.
69. B. Xiao; G. Shi; X. Chen; H. Yan; X. Ji, *Structure* **1999**, 7 (5), 489–496.
70. J. Blaszczyk; Y. Li; S. Cherry; J. Alexandratos; Y. Wu; G. Shaw; J. E. Tropea; D. S. Waugh; H. Yan; X. Ji, *Acta Crystallogr. D Biol. Crystallogr.* **2007**, 63 (Pt. 11), 1169–1177.
71. A. Bermingham; J. R. Bottomley; W. U. Primrose; J. P. Derrick, *J. Biol. Chem.* **2000**, 275 (24), 17962–17967.
72. J. Blaszczyk; G. Shi; H. Yan; X. Ji, *Structure* **2000**, 8 (10), 1049–1058.
73. M. C. Lawrence; P. Iliades; R. T. Fernley; J. Berglez; P. A. Pilling; I. G. Macreadie, *J. Mol. Biol.* **2005**, 348 (3), 655–670.
74. G. Shi; J. Blaszczyk; X. Ji; H. Yan, *J. Med. Chem.* **2001**, 44 (9), 1364–1371.
75. G. Shi; Y. Gong; A. Savchenko; J. G. Zeikus; B. Xiao; X. Ji; H. Yan, *Biochim. Biophys. Acta* **2000**, 1478 (2), 289–299.
76. A. Garcon; A. Bermingham; L. Y. Lian; J. P. Derrick, *Biochem. J.* **2004**, 380 (Pt. 3), 867–873.
77. J. Blaszczyk; Y. Li; Y. Wu; G. Shi; X. Ji; H. Yan, *Biochemistry* **2004**, 43 (6), 1469–1477.
78. O. Keskin; X. Ji; J. Blaszczyk; D. G. Covell, *Proteins* **2002**, 49 (2), 191–205.
79. B. Xiao; G. Shi; J. Gao; J. Blaszczyk; Q. Liu; X. Ji; H. Yan, *J. Biol. Chem.* **2001**, 276 (43), 40274–40281.
80. H. Yan; J. Blaszczyk; B. Xiao; G. Shi; X. Ji, *J. Mol. Graph. Model* **2001**, 19 (1), 70–77.
81. C. Levy; D. Minnis; J. P. Derrick, *Biochem. J.* **2008**, 412, 379–388.
82. K. Babaoglu; J. Qi; R. E. Lee; S. W. White, *Structure* **2004**, 12 (9), 1705–1717.
83. A. M. Baca; R. Sirawaraporn; S. Turley; W. Sirawaraporn; W. G. Hol, *J. Mol. Biol.* **2000**, 302 (5), 1193–1212.
84. I. C. Hampele; A. D'Arcy; G. E. Dale; D. Kostrewa; J. Nielsen; C. Oefner; M. G. Page; H. J. Schonfeld; D. Stuber; R. L. Then, *J. Mol. Biol.* **1997**, 268 (1), 21–30.
85. A. Achari; D. O. Somers; J. N. Champness; P. K. Bryant; J. Rosemond; D. K. Stammers, *Nat. Struct. Biol.* **1997**, 4 (6), 490–497.
86. J. L. McCullough; T. H. Maren, *Antimicrob. Agents Chemother.* **1973**, 3 (6), 665–669.
87. L. Bock; G. H. Miller; K. J. Schaper; J. K. Seydel, *J. Med. Chem.* **1974**, 17 (1), 23–28.
88. S. Roland; R. Ferone; R. J. Harvey; V. L. Styles; R. W. Morrison, *J. Biol. Chem.* **1979**, 254 (20), 10337–10345.
89. P. Iliades; S. R. Meshnick; I. G. Macreadie, *Antimicrob. Agents Chemother.* **2005**, 49 (2), 741–748.
90. C. R. Stein; C. Poole; P. Kazanjian; S. R. Meshnick, *Emerg. Infect. Dis.* **2004**, 10 (10), 1760–1765.
91. L. Huang; K. Crothers; C. Atzori; T. Benfield; R. Miller; M. Rabodonirina; J. Helweg-Larsen, *Emerg. Infect. Dis.* **2004**, 10 (10), 1721–1728.
92. P. Iliades; S. R. Meshnick; I. G. Macreadie, *Antimicrob. Agents Chemother.* **2004**, 48 (7), 2617–2623.
93. G. Swedberg; C. Fermer; O. Skold, *Adv. Exp. Med. Biol.* **1993**, 338, 555–558.
94. P. Lopez; M. Espinosa; B. Greenberg; S. A. Lacks, *J. Bacteriol.* **1987**, 169 (9), 4320–4326.
95. T. Triglia; J. G. Menting; C. Wilson; A. F. Cowman, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, 94 (25), 13944–13949.
96. A. L. Bogнар; C. Osborne; B. Shane; S. C. Singer; R. Ferone, *J. Biol. Chem.* **1985**, 260 (9), 5625–5630.
97. R. Ferone; A. Warskow, *Adv. Exp. Med. Biol.* **1983**, 163, 167–181.
98. R. V. Banerjee; B. Shane; J. J. McGuire; J. K. Coward, *Biochemistry* **1988**, 27 (25), 9062–9070.
99. L. J. Kimlova; C. Pyne; K. Keshavjee; J. Huy; G. Beebakhee; A. L. Bogнар, *Arch. Biochem. Biophys.* **1991**, 284 (1), 9–16.
100. M. Mathieu; G. Debousker; S. Vincent; F. Viviani; N. Bamas-Jacques; V. Mikol, *J. Biol. Chem.* **2005**, 280 (19), 18916–18922.
101. X. Sun; A. L. Bogнар; E. N. Baker; C. A. Smith, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, 95 (12), 6647–6652.
102. D. J. Cichowicz; B. Shane, *Biochemistry* **1987**, 26 (2), 513–521.
103. H. T. Spencer; J. E. Villafranca; J. R. Appleman, *Biochemistry* **1997**, 36 (14), 4212–4222.
104. D. Leduc; F. Escartin; H. F. Nijhout; M. C. Reed; U. Liebl; S. Skouloubris; H. Myllykallio, *J. Bacteriol.* **2007**, 189 (23), 8537–8545.
105. S. Graziani; J. Bernauer; S. Skouloubris; M. Graille; C. Z. Zhou; C. Marchand; P. Decottignies; H. van Tilbeurgh; H. Myllykallio; U. Liebl, *J. Biol. Chem.* **2006**, 281 (33), 24048–24057.
106. S. Graziani; Y. Xia; J. R. Gurnon; J. L. Van Etten; D. Leduc; S. Skouloubris; H. Myllykallio; U. Liebl, *J. Biol. Chem.* **2004**, 279 (52), 54340–54347.
107. H. Myllykallio; D. Leduc; J. Filee; U. Liebl, *Trends Microbiol.* **2003**, 11 (5), 220–223.
108. H. Myllykallio; G. Lipowski; D. Leduc; J. Filee; P. Forterre; U. Liebl, *Science* **2002**, 297 (5578), 105–107.
109. S. J. Benkovic; S. Hammes-Schiffer, *Science* **2003**, 301 (5637), 1196–1202.
110. S. J. Benkovic, *Annu. Rev. Biochem.* **1980**, 49, 227–251.
111. J. Thillet; J. A. Adams; S. J. Benkovic, *Biochemistry* **1990**, 29 (21), 5195–5202.
112. S. G. Amyes; J. T. Smith, *Biochem. Biophys. Res. Commun.* **1974**, 58 (2), 412–418.
113. S. L. Smith; D. Stone; P. Novak; D. P. Baccanari; J. J. Burchall, *J. Biol. Chem.* **1979**, 254 (14), 6222–6225.

114. M. Garcia-Viloca; D. G. Truhlar; J. Gao, *Biochemistry* **2003**, *42* (46), 13558–13575.
115. M. Giladi; N. Altman-Price; I. Levin; L. Levy; M. Mevarech, *J. Bacteriol.* **2003**, *185* (23), 7015–7018.
116. E. E. Howell, *ChemBiochem* **2005**, *6* (4), 590–600.
117. I. Levin; M. Mevarech; B. A. Palfey, *J. Bacteriol.* **2007**, *189* (11), 4062–4069.
118. M. Masurekar; G. M. Brown, *Biochemistry* **1975**, *14* (11), 2424–2430.
119. B. Shane, *J. Biol. Chem.* **1980**, *255* (12), 5655–5662.
120. Y. Sheng; N. Khanam; Y. Tsaksis; X. M. Shi; Q. S. Lu; A. L. Bogнар, *Biochemistry* **2008**, *47* (8), 2388–2396.
121. X. Sun; J. A. Cross; A. L. Bogнар; E. N. Baker; C. A. Smith, *J. Mol. Biol.* **2001**, *310* (5), 1067–1078.
122. J. A. Bertrand; E. Fanchon; L. Martin; L. Chantalat; G. Auger; D. Blanot; J. van Heijenoort; O. Dideberg, *J. Mol. Biol.* **2000**, *301* (5), 1257–1266.
123. F. Rebeille; S. Ravelana; S. Jabrina; R. Doucea; S. Storozhenkob; D. Van Der Straeten, *Physiol. Plant.* **2006**, *126*, 330–342.
124. V. de Crecy-Lagard; B. El Yacoubi; R. D. de la Garza; A. Noiriel; A. D. Hanson, *BMC Genomics* **2007**, *8*, 245.
125. J. E. Hyde, *Acta Trop.* **2005**, *94* (3), 191–206.
126. G. A. McConkey; J. W. Pinney; D. R. Westhead; K. Plueckhahn; T. B. Fitzpatrick; P. Macheroux; B. Kappes, *Trends Parasitol.* **2004**, *20* (2), 60–65.
127. S. Hawser; S. Lociuoro; K. Islam, *Biochem. Pharma.* **2006**, *71*, 941–948.
128. G. M. Brown, *J. Biol. Chem.* **1962**, *237*, 536–540.
129. T. Shiota; M. N. Disraely; M. P. McCann, *J. Biol. Chem.* **1964**, *239*, 2259–2266.
130. G. H. Hitchings, *S. Afr. Med. J.* **1970**, *44* (32 Suppl.), 1–2.
131. R. Then; P. Angehrn, *J. Gen. Microbiol.* **1973**, *76* (2), 255–263.
132. W. Shi; C. J. Meininger; T. E. Haynes; K. Hatakeyama; G. Wu, *Cell Biochem. Biophys.* **2004**, *41* (3), 415–434.
133. D. J. Walther; M. Bader, *Biochem. Pharmacol.* **2003**, *66* (9), 1673–1680.
134. D. J. Walther; J. U. Peter; S. Bashammakh; H. Hortaagl; M. Voits; H. Fink; M. Bader, *Science* **2003**, *299* (5603), 76.
135. T. J. Kappock; J. P. Caradonna, *Chem. Rev.* **1996**, *96* (7), 2659–2756.
136. H. Taguchi; W. L. Armarego, *Med. Res. Rev.* **1998**, *18* (1), 43–89.
137. E. R. Werner; A. Hermetter; H. Prast; G. Golderer; G. Werner-Felmayer, *J. Lipid Res.* **2007**, *48* (6), 1422–1427.
138. A. Tietz; M. Lindberg; E. P. Kennedy, *J. Biol. Chem.* **1964**, *239*, 4081–4090.
139. S. Kaufman, *J. Biol. Chem.* **1957**, *226* (1), 511–524.
140. S. Kaufman, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1971**, *35*, 245–319.
141. R. Shiman; S. H. Jones; D. W. Gray, *J. Biol. Chem.* **1990**, *265* (20), 11633–11642.
142. O. A. Andersen; T. Flatmark; E. Hough, *J. Mol. Biol.* **2002**, *320*, 1095–1108.
143. R. A. Lazarus; S. J. Benkovic; S. Kaufman, *J. Biol. Chem.* **1983**, *258*, 10960–10962.
144. R. A. Lazarus; D. E. Wallick; R. F. Dietrich; D. W. Gottschall; S. J. Benkovic; B. J. Gaffney; R. Shiman, *Fed Proc.* **1982**, *41* (9), 2605–2607.
145. T. A. Dix; D. M. Kuhn; S. J. Benkovic, *Biochemistry* **1987**, *26* (12), 3354–3361.
146. J. Haavik; T. Flatmark, *Eur. J. Biochem.* **1987**, *168* (1), 21–26.
147. S. E. Hufton; I. G. Jennings; R. G. Cotton, *Biochem. J.* **1995**, *311* (Pt. 2), 353–366.
148. K. Teigen; N. A. Froystein; A. Martinez, *J. Mol. Biol.* **1999**, *294*, 807–823.
149. B. E. Eser; E. W. Barr; P. A. Frantom; L. Saleh; J. M. Bollinger; C. Krebs; P. F. Fitzpatrick, *J. Am. Chem. Soc.* **2007**, *129*, 11334–11335.
150. L. J. Mitnaul; R. Shiman, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92* (3), 885–889.
151. A. L. Pey; B. Perez; L. R. Desviat; M. A. Martinez; C. Aguado; H. Erlandsen; A. Gamez; R. C. Stevens; M. Thorolfsson; M. Ugarte; A. Martinez, *Hum. Mutat.* **2004**, *24* (5), 388–399.
152. B. Thöny; Z. Ding; A. Martinez, *FEBS Lett.* **2004**, *577* (3), 507–511.
153. S. Milstien; S. Kaufman, *J. Biol. Chem.* **1975**, *250* (12), 4777–4781.
154. T. Fukushima; J. C. Nixon, *Anal. Biochem.* **1980**, *102* (1), 176–188.
155. K. Teigen; A. Martínez, *J. Biomol. Struct. Dyn.* **2003**, *20* (6), 733–740.
156. S. Kure; K. Sato; K. Fujii; Y. Aoki; Y. Suzuki; S. Kato; Y. Matsubara, *Mol. Genet. Metab.* **2004**, *83* (1–2), 150–156.
157. B. Almas; B. Le Bourdelles; T. Flatmark; J. Mallet; J. Haavik, *Eur. J. Biochem.* **1992**, *209* (1), 249–255.
158. C. C. Wei; B. R. Crane; D. J. Stuehr, *Chem. Rev.* **2003**, *103* (6), 2365–2383.
159. D. J. Stuehr; J. Santolini; Z. Q. Wang; C. C. Wei; S. Adak, *J. Biol. Chem.* **2004**, *279* (35), 36167–36170.
160. J. Giovanelli; K. L. Campos; S. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88* (16), 7091–7095.
161. C. M. Pastor; D. Williams; T. Yoneyama; K. Hatakeyama; S. Singleton; E. Naylor; T. R. Billiar, *J. Biol. Chem.* **1996**, *271* (40), 24534–24538.
162. M. C. Boadle-Biber, *Prog. Biophys. Mol. Biol.* **1993**, *60* (1), 1–15.
163. N. J. Alp; S. Mussa; J. Khoo; S. Cai; T. Guzik; A. Jefferson; N. Goh; K. A. Rockett; K. M. Channon, *J. Clin. Invest.* **2003**, *112* (5), 725–735.
164. M. Pannirselvam; V. Simon; S. Verma; T. Anderson; C. R. Triggle, *Br. J. Pharmacol.* **2003**, *140* (4), 701–706.
165. N. Ihlemann; C. Rask-Madsen; A. Perner; H. Dominguez; T. Hermann; L. Kober; C. Torp-Pedersen, *Am. J. Physiol. Heart Circ. Physiol.* **2003**, *285* (2), H875–H882.
166. C. J. Meininger; S. Cai; J. L. Parker; K. M. Channon; K. A. Kelly; E. J. Becker; M. K. Wood; L. A. Wade; G. Wu, *FASEB J.* **2004**, *18*, 1900–1902.
167. J. H. Bayle; F. Randazzo; G. Johnen; S. Kaufman; A. Nagy; J. Rossant; G. R. Crabtree, *J. Biol. Chem.* **2002**, *277* (32), 28884–28891.
168. M. Delgado-Esteban; A. Almeida; J. M. Medina, *J. Neurochem.* **2002**, *82* (5), 1148–1159.
169. N. J. Alp; K. M. Channon, *Arterioscler. Thromb. Vasc. Biol.* **2004**, *24* (3), 413–420.
170. N. Blau; L. Bonafe; B. Thöny, *Mol. Genet. Metab.* **2001**, *74* (1–2), 172–185.
171. M. Kirsch; H. G. Korth; V. Stenert; R. Sustmann; H. de Groot, *J. Biol. Chem.* **2003**, *278* (27), 24481–24490.
172. N. Blau; H. Erlandsen, *Mol. Genet. Metab.* **2004**, *82* (2), 101–111.

173. K. Tanaka; S. Kaufman; S. Milstien, *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86* (15), 5864–5867.
174. F. Kerler; I. Ziegler; B. Schwarzkopf; A. Bacher, *FEBS Lett.* **1989**, *250* (2), 622–624.
175. P. Z. Anastasiadis; L. Bezin; B. A. Imerman; D. M. Kuhn; M. C. Louie; R. A. Levine, *Eur. J. Neurosci.* **1997**, *9* (9), 1831–1837.
176. J. Tanaka; K. Koshimura; Y. Murakami; Y. Kato, *Neurosci. Lett.* **2002**, *328* (2), 201–203.
177. K. Koshimura; Y. Murakami; J. Tanaka; Y. Kato, *J. Neurosci. Res.* **1998**, *54* (5), 664–672.
178. S. Kojima; S. Ona; I. Iizuka; T. Arai; H. Mori; K. Kubota, *Free Radic. Res.* **1995**, *23* (5), 419–430.
179. S. Shimizu; M. Ishii; Y. Kawakami; K. Momose; T. Yamamoto, *Life Sci.* **1998**, *63* (18), 1585–1592.
180. S. Cho; B. T. Volpe; Y. Bae; O. Hwang; H. J. Choi; J. Gal; L. C. Park; C. K. Chu; J. Du; T. H. Joh, *J. Neurosci.* **1999**, *19* (3), 878–889.
181. K. Hyland; T. L. Munk-Martin, *J. Inherit. Metab. Dis.* **2001**, *24* (Suppl.), 30.
182. P. Linscheid; A. Schaffner; G. Schoedon, *Biochem. Biophys. Res. Commun.* **1998**, *243* (1), 137–141.
183. L. R. Desviat; B. Perez; A. Bèlanger-Quintana; M. Castro; C. Aguado; A. Sánchez; M. J. Garcia, M. Martinez-Pardo, M. Ugarte, *Mol. Genet. Metab.* **2004**, *83*, 157–162.
184. R. Scavelli; Z. Ding; N. Blau; J. Haavik; A. Martinez; B. Thöny, *Mol. Genet. Metab.* **2005**, *86* (Suppl. 1), S153–S155.
185. K. Koshimura; S. Miwa; K. Lee; M. Fujiwara; Y. Watanabe, *J. Neurochem.* **1990**, *54* (4), 1391–1397.
186. N. Mataga; K. Imamura; Y. Watanabe, *Brain Res.* **1991**, *551* (1–2), 64–71.
187. K. Koshimura; S. Miwa; Y. Watanabe, *J. Neurochem.* **1994**, *63* (2), 649–654.
188. L. P. Liang; S. Kaufman, *Brain Res.* **1998**, *800* (2), 181–186.
189. G. Bader; S. Schifffmann; A. Herrmann; M. Fischer; M. Gutlich; G. Auerbach; T. Ploom; A. Bacher; R. Huber; T. Lemm, *J. Mol. Biol.* **2001**, *312* (5), 1051–1057.
190. N. Maita; K. Okada; K. Hatakeyama; T. Hakoshima, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (3), 1212–1217.
191. D. Bürgisser; B. Thöny; U. Redweik; D. Hess; C. W. Heizmann; R. Huber; H. Nar, *J. Mol. Biol.* **1995**, *253*, 358–369.
192. T. Ploom; B. Thöny; S. Lee; H. Nar; W. Leimbacher; R. Huber; G. Auerbach, *J. Mol. Biol.* **1999**, *286*, 851–860.
193. G. Auerbach; A. Herrmann; M. Gütlich; M. Fischer; U. Jacob; A. Bacher; R. Huber, *EMBO J.* **1997**, *16*, 7219–7230.
194. K. Fujimoto; S. Y. Takahashi; S. Katoh, *Biochim. Biophys. Acta* **2002**, *1594* (1), 191–198.
195. S. Milstien; S. Kaufman, *J. Biol. Chem.* **1989**, *264* (14), 8066–8073.
196. Y. S. Park; C. W. Heizmann; B. Wermuth; R. A. Levine; P. Steinerstauch; J. Guzman; N. Blau, *Biochem. Biophys. Res. Commun.* **1991**, *175*, 738–744.
197. T. Iino; M. Tabata; S. Takikawa; H. Sawada; H. Shintaku; S. Ishikura; A. Hara, *Arch. Biochem. Biophys.* **2003**, *416* (2), 180–187.
198. L. Bonafe; B. Thöny; J. M. Penzien; B. Czarnecki; N. Blau, *Am. J. Hum. Genet.* **2001**, *69* (2), 269–277.
199. K. Sawabe; K. O. Wakasugi; H. Hasegawa, *J. Pharmacol. Sci.* **2004**, *96* (2), 124–133.
200. S. Kaufman, *J. Biol. Chem.* **1970**, *254*, 4751–4759.
201. D. Mendel; P. A. Kahvari; P. B. Conley; M. K. Graves; L. P. Hansen; A. Admon; G. R. Crabtree, *Science* **1991**, *254*, 1762–1767.
202. B. A. Citron; M. D. Davis; S. Milstien; J. Gutierrez; D. B. Mendel; G. R. Crabtree; S. Kaufman, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 11891–11894.
203. C. R. Hauer; I. Rebrin; B. Thöny; F. Neuheiser; H. C. Curtius; P. Hunziker; N. Blau; S. Ghisla; C. W. Heizmann, *J. Biol. Chem.* **1993**, *268* (7), 4828–4831.
204. R. B. Rose; K. E. Pullen; J. H. Bayle; G. R. Crabtree; T. Alber, *Biochemistry* **2004**, *43* (23), 7345–7355.
205. I. Rebrin; S. W. Bailey; S. R. Boerth; M. D. Ardell; J. E. Ayling, *Biochemistry* **1995**, *34*, 5801–5810.
206. J. D. Cronk; J. A. Endrizzi; T. Alber, *Protein Sci.* **1996**, *5*, 1963–1972.
207. S. Köster; B. Thöny; P. Macheroux; H.-C. Curtius; C. W. Heizmann; W. Pfeleiderer; S. Ghisla, *Eur. J. Biochem.* **1995**, *231*, 414–423.
208. I. Rebrin; B. Thöny; S. W. Bailey; J. E. Ayling, *Biochemistry* **1998**, *37* (32), 11246–11254.
209. S. Köster; G. Stier; R. Ficner; M. Holzer; H. C. Curtius; D. Suck; S. Ghisla, *Eur. J. Biochem.* **1996**, *241* (3), 858–864.
210. S. W. Bailey; S. R. Boerth; S. B. Dillard; J. E. Ayling, *Adv. Exp. Med. Biol.* **1993**, *338*, 47–54.
211. H.-C. Curtius; C. Adler; I. Rebrin; C. W. Heizmann; S. Ghisla, *Biochem. Biophys. Res. Commun.* **1990**, *172*, 1060–1066.
212. M. D. Davis; S. Kaufman; S. Milstien, *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 385–389.
213. R. Ficner; U. H. Sauer; G. Stier; D. Suck, *EMBO J.* **1995**, *14* (9), 2034–2042.
214. J. A. Endrizzi; J. D. Cronk; W. Wang; G. R. Crabtree; T. Alber, *Science* **1995**, *268* (5210), 556–559.
215. D. J. Sourdive; C. Transy; S. Garbay; M. Yaniv, *Nucleic Acids Res.* **1997**, *25* (8), 1476–1484.
216. K. H. Rhee; G. Stier; P. B. Becker; D. Suck; R. Sandaltzopoulos, *J. Mol. Biol.* **1997**, *265* (1), 20–29.
217. G. Zhao; T. Xia; J. Song; R. A. Jensen, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91* (4), 1366–1370.
218. W. L. Armarego; D. Randles; P. Waring, *Med. Res. Rev.* **1984**, *4* (3), 267–321.
219. K. I. Varughese; M. M. Skinner; J. M. Whiteley; D. A. Matthews; N. H. Xuong, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89* (13), 6080–6084.
220. Y. Su; K. I. Varughese; N. H. Xuong; T. L. Bray; D. J. Roche; J. M. Whiteley, *J. Biol. Chem.* **1993**, *268* (36), 26836–26841.
221. I. Smith; O. H. Wolff, *Lancet* **1974**, *2* (7880), 540–544.
222. N. Blau; B. Thöny; R. G. H. Cotton; K. Hyland, Disorders of Tetrahydrobiopterin and Related Biogenic Amines. In *The Metabolic and Molecular Bases of Inherited Disease*, 8th ed.; C. R. Scriver; A. L. Beaudet; W. S. Sly; D. Valle; B. Vogelstein, Eds.; McGraw-Hill: New York, 2001; pp 1725–1776.
223. B. Thöny; F. Neuheiser; L. Kierat; M. O. Rolland; P. Guibaud; T. Schluter; R. Germann; R. A. Heidenreich; M. Duran; J. B. de Klerk; J. E. Ayling; N. Blau, *Hum. Genet.* **1998**, *103* (2), 162–167.
224. B. Thöny; F. Neuheiser; L. Kierat; M. Blaskovics; P. H. Arn; P. Ferreira; I. Rebrin; J. Ayling; N. Blau, *Am. J. Hum. Genet.* **1998**, *62* (6), 1302–1311.
225. B. Thöny; N. Blau, *Hum. Mutat.* **1997**, *10* (1), 11–20.
226. I. Dianzani; L. de Sanctis; P. M. Smooker; T. J. Gough; C. Alliaudi; A. Brusco; M. Spada; N. Blau; M. Dobos; H. P. Zhang; N. Yang; A. Ponzzone; W. L. Armarego; R. G. Cotton, *Hum. Mutat.* **1998**, *12* (4), 267–273.
227. B. Thöny; N. Blau, *Hum. Mutat.* **2006**, *27* (9), 870–878.
228. M. Segawa; A. Hosaka; F. Miyagawa; Y. Nomura; H. Imai, *Adv. Neurol.* **1976**, *14*, 215–233.



229. T. G. Nygaard; J. M. Trugman; J. G. de Yebenes; S. Fahn, *Neurology* **1990**, *40* (1), 66–69.
230. H. Ichinose; T. Ohye; E.-I. Takahashi; N. Seki; T.-A. Tada-aki Hori; M. Segawa; Y. Nomura; K. Endo; H. Tanaka; S. Tsuji, K. Fujita; T. Nagatsu, *Nat. Genet.* **1994**, *8* (3), 236–242.
231. Y. Furukawa; S. J. Kish, *Mov. Disord.* **1999**, *14* (5), 709–715.
232. H. Ichinose; T. Suzuki; H. Inagaki; T. Ohye; T. Nagatsu, *Biol. Chem.* **1999**, *380* (12), 1355–1364.
233. H. C. Curtius; A. Niederwieser; R. Levine; H. Muldner, *Adv. Neurol.* **1984**, *40*, 463–466.
234. Y. Tani; E. Fernell; Y. Watanabe; T. Kanai; B. Langstrom, *Neurosci. Lett.* **1994**, *181* (1–2), 169–172.
235. T. Bottiglieri; K. Hyland; M. Laundry; P. Godfrey; M. W. Carney; B. K. Toone; E. H. Reynolds, *Psychol. Med.* **1992**, *22* (4), 871–876.
236. P. A. Barford; J. A. Blair; C. Eggar; C. Hamon; C. Morar; S. B. Whitburn, *J. Neurol. Neurosurg. Psychiatry* **1984**, *47* (7), 736–738.
237. H. C. Curtius; A. Niederwieser; R. A. Levine; W. Lovenberg; B. Woggon; J. Angst, *Lancet* **1983**, *1* (8325), 657–658.
238. E. Fernell; Y. Watanabe; I. Adolfsson; Y. Tani; M. Bergstrom; P. Hartvig; A. Lilja; A. L. von Knorring; C. Gillberg; B. Langstrom, *Dev. Med. Child Neurol.* **1997**, *39* (5), 313–318.
239. K. U. Schallreuter; J. M. Wood; M. R. Pittelkow; M. Gutlich; K. R. Lemke; W. Rodl; N. N. Swanson; K. Hitzemann; I. Ziegler, *Science* **1994**, *263* (5152), 1444–1446.
240. K. U. Schallreuter; W. D. Beazley; N. A. Hibberts; N. N. Swanson; M. R. Pittelkow, *J. Invest. Dermatol.* **1998**, *111* (3), 511–516.
241. K. U. Schallreuter; J. Moore; J. M. Wood; W. D. Beazley; E. M. Peters; L. K. Marles; S. C. Behrens-Williams; R. Dummer; N. Blau; B. Thöny, *J. Invest. Dermatol.* **2001**, *116* (1), 167–174.
242. K. U. Schallreuter, *Cell. Mol. Biol. (Noisy-le-grand)* **1999**, *45* (7), 943–949.
243. S. Kure; D. C. Hou; T. Ohura; H. Iwamoto; S. Suzuki; N. Sugiyama; O. Sakamoto; K. Fujii; Y. Matsubara; K. Narisawa, *J. Pediatr.* **1999**, *135* (3), 375–378.
244. M. Zurfluh; J. Zschocke; M. Lindner; F. Feillet; C. Chery; A. Burlina; R. C. Stevens; B. Thöny; N. Blau, *Hum. Mutat.* **2007**, *29*, 167–175.
245. H. Erlandsen; A. L. Pey; A. Gamez; B. Perez; L. R. Desviat; C. Aguado; R. Koch; S. Surendran; S. Tyring; R. Matalon; C. R. Scriver; M. Ugarte; A. Martinez; R. C. Stevens, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (48), 16903–16908.
246. A. L. Pey; F. Stricher; L. Serrano; A. Martinez, *Am. J. Hum. Genet.* **2007**, *81* (5), 1006–1024.
247. V. C. Bode; J. D. McDonald; J. L. Guenet; D. Simon, *Genetics* **1988**, *118* (2), 299–305.
248. K. Hyland; R. S. Gunasekara; T. L. Munk-Martin; L. A. Arnold; T. Engle, *Ann. Neurol.* **2003**, *54* (Suppl. 6), S46–S48.
249. J. P. Khoo; T. Nicoli; N. J. Alp; J. Fullerton; J. Flint; K. M. Channon, *Mol. Genet. Metab.* **2004**, *82* (3), 251–254.
250. C. Sumi-Ichinose; F. Urano; R. Kuroda; T. Ohye; M. Kojima; M. Tazawa; H. Shiraishi; Y. Hagino; T. Nagatsu; T. Nomura; H. Ichinose, *J. Biol. Chem.* **2001**, *276* (44), 41150–41160.
251. L. Elzaouk; W. Leimbacher; M. Turri; B. Ledermann; K. Bürki; N. Blau; B. Thöny, *J. Biol. Chem.* **2003**, *278* (30), 28303–28311.
252. R. Hille, *Trends Biochem. Sci.* **2002**, *27*, 360–367.
253. C. Kisker; H. Schindelin; D. C. Rees, *Ann. Rev. Biochem.* **1997**, *66*, 233–267.
254. A. Kletzin; M. W. Adams, *FEMS Microbiol. Rev.* **1996**, *18*, 5–63.
255. S. Gutteridge; S. J. Tanner; R. C. Bray, *Biochem. J.* **1978**, *175* (3), 887–897.
256. R. C. Wahl; K. V. Rajagopalan, *J. Biol. Chem.* **1982**, *257*, 1354–1359.
257. H. Dobbek; L. Gremer; R. Kiefersauer; R. Huber; O. Meyer, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (25), 15971–15976.
258. S. J. Broxk; R. A. Rothery; G. Zhang; D. P. Ng; J. H. Weiner, *Biochemistry* **2005**, *44* (30), 10339–10348.
259. S. Najmudin; P. J. Gonzalez; C. Coelho; A. Mukhopadhyay; N. M. Cerqueira; C. C. Romao; I. Moura; J. J. Moura; C. D. Brondino; M. J. Romao, *J. Biol. Inorg. Chem.* **2008**, *13*, 737–753.
260. H. Schindelin; C. Kisker; J. Hilton; K. V. Rajagopalan; D. C. Rees, *Science* **1996**, *272*, 1615–1621.
261. C. A. McDevitt; P. Hugenholtz; G. R. Hanson; A. G. McEwan, *Mol. Microbiol.* **2002**, *44* (6), 1575–1587.
262. E. Garattini; R. Mendel; M. J. Romao; R. Wright; M. Terao, *Biochem. J.* **2003**, *372* (Pt. 1), 15–32.
263. F. Stirpe; E. Della Corte, *J. Biol. Chem.* **1969**, *244*, 3855–3863.
264. W. R. Waud; K. V. Rajagopalan, *Arch. Biochem. Biophys.* **1976**, *172*, 365–379.
265. L. M. Schopfer; V. Massey; T. Nishino, *J. Biol. Chem.* **1988**, *263* (27), 13528–13538.
266. T. Nishino, *J. Biochem. (Tokyo)* **1994**, *116* (1), 1–6.
267. J. M. McCord, *New Engl. J. Med.* **1985**, *312*, 159–163.
268. J. Hunt; V. Massey; W. R. Dunham; R. H. Sands, *J. Biol. Chem.* **1993**, *268* (25), 18685–18691.
269. G. Palmer; V. Massey, *J. Biol. Chem.* **1969**, *244*, 2614–2620.
270. S. Leimkühler; A. L. Stockert; K. Igarashi; T. Nishino; R. Hille, *J. Biol. Chem.* **2004**, *279*, 40437–40444.
271. E.-Y. Choi; A. L. Stockert; S. Leimkühler; R. Hille, *J. Inorg. Biochem.* **2004**, *98*, 841–848.
272. R. Hille, *Arch. Biochem. Biophys.* **2005**, *433* (1), 107–116.
273. S. Leimkühler; M. Kern; P. S. Solomon; A. G. McEwan; G. Schwarz; R. R. Mendel; W. Klipp, *Mol. Microbiol.* **1998**, *27* (4), 853–869.
274. J. J. Truglio; K. Theis; S. Leimkühler; R. Rappa; K. V. Rajagopalan; C. Kisker, *Structure (Camb)* **2002**, *10* (1), 115–125.
275. J. M. Pauff; C. F. Hemann; N. Jünemann; S. Leimkühler; R. Hille, *J. Biol. Chem.* **2007**, *282* (17), 12785–12790.
276. U. Dietzel; J. Kuper; J. A. Doebbler; A. Schulte; J. J. Truglio; S. Leimkühler; C. Kisker, *J. Biol. Chem.* **2008**, *284*, 8768–8776.
277. J. L. Johnson; M. Duran, Molybdenum Cofactor Deficiency and Isolated Sulfite Oxidase Deficiency. In *The Metabolic and Molecular Bases of Inherited Disease*, 8th ed.; C. R. Scriver; A. L. Beaudet; W. S. Sly; D. Valle; B. Childs; B. Vogelstein, Eds.; McGraw-Hill: New York, 2001; pp 3163–3177.
278. J. L. Johnson; M. M. Wuebbens; K. V. Rajagopalan, *J. Biol. Chem.* **1989**, *264*, 13440–13447.
279. M. M. Wuebbens; K. V. Rajagopalan, *J. Biol. Chem.* **1995**, *270*, 1082–1087.
280. P. Hanzelmann; H. Schindelin, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (35), 12870–12875.
281. A. Bacher; S. Eberhardt; W. Eisenreich; M. Fischer; S. Herz; B. Illarionov; K. Kis; G. Richter, *Vitam. Horm.* **2001**, *61*, 1–49.
282. S. Kaufman, *J. Biol. Chem.* **1967**, *242* (17), 3934–3943.
283. J. A. Santamaria-Araujo; B. Fischer; T. Otte; M. Nimtz; R. R. Mendel; V. Wray; G. Schwarz, *J. Biol. Chem.* **2004**, *279*, 15994–15999.

284. J. N. Daniels; M. M. Wuebbens; K. V. Rajagopalan; H. Schindelin, *Biochemistry* **2007**, *47*, 615–626.
285. H. J. Sofia; G. Chen; B. G. Hetzler; J. F. Reyes-Spindola; N. E. Miller, *Nucleic Acids Res.* **2001**, *29*, 1097–1106.
286. M. M. Wuebbens; M. T. Liu; K. Rajagopalan; H. Schindelin, *Structure* **2000**, *8* (7), 709–718.
287. P. Hanzelmann; H. L. Hernandez; C. Menzel; R. Garcia-Serres; B. H. Huynh; M. K. Johnson; R. R. Mendel; H. Schindelin, *J. Biol. Chem.* **2004**, *279* (33), 34721–34732.
288. J. Reiss, *Hum. Genet.* **2000**, *106*, 157–163.
289. S. Gross-Hard; J. Reiss, *Mol. Genet. Metab.* **2002**, *76*, 340–343.
290. T. A. Gray; R. D. Nicholls, *RNA* **2000**, *6*, 928–936.
291. P. Hanzelmann; G. Schwarz; R. R. Mendel, *J. Biol. Chem.* **2002**, *277*, 18303–18312.
292. D. M. Pitterle; K. V. Rajagopalan, *J. Bacteriol.* **1989**, *171*, 3373–3378.
293. D. M. Pitterle; K. V. Rajagopalan, *J. Biol. Chem.* **1993**, *268*, 13499–13505.
294. G. Gutzke; B. Fischer; R. R. Mendel; G. Schwarz, *J. Biol. Chem.* **2001**, *276*, 36268–36274.
295. M. J. Rudolph; M. M. Wuebbens; K. V. Rajagopalan; H. Schindelin, *Nat. Struct. Biol.* **2001**, *8* (1), 42–46.
296. M. M. Wuebbens; K. V. Rajagopalan, *J. Biol. Chem.* **2003**, *278*, 14523–14532.
297. S. Leimkühler; M. M. Wuebbens; K. V. Rajagopalan, *J. Biol. Chem.* **2001**, *276*, 34695–34701.
298. J. P. McGrath; S. Jentsch; A. Varshavsky, *EMBO J.* **1991**, *10*, 227–236.
299. K. V. Rajagopalan, *Biochem. Soc. Trans.* **1997**, *25*, 757–761.
300. M. W. Lake; M. M. Wuebbens; K. V. Rajagopalan; H. Schindelin, *Nature* **2001**, *414*, 325–329.
301. H. Schindelin, Evolutionary Origin of the Activation Step During Ubiquitin-Dependent Protein Degradation. In *Protein Degradation 1. Ubiquitin and the Chemistry of Life*; J. R. Mayer, A. Ciechanover, M. Rechsteiner, Eds.; Wiley-VCH Verlag: Weinheim, 2005; Vol. 1, pp 21–43.
302. S. Leimkühler; K. V. Rajagopalan, *J. Biol. Chem.* **2001**, *276*, 22024–22031.
303. B. Stallmeyer; G. Drugeon; J. Reiss; A. L. Haenni; R. R. Mendel, *Am. J. Hum. Genet.* **1999**, *64*, 698–705.
304. R. Hahnwald; S. Leimkühler; A. Vilaseca; C. Acquaviva-Bourdain; U. Lenz; J. Reiss, *Mol. Genet. Metab.* **2006**, *89*, 210–213.
305. S. Leimkühler; A. Freuer; J. A. Araujo; K. V. Rajagopalan; R. R. Mendel, *J. Biol. Chem.* **2003**, *278*, 26127–26134.
306. A. Matthies; K. V. Rajagopalan; R. R. Mendel; S. Leimkühler, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5946–5951.
307. D. Bordo; P. Bork, *EMBO Rep.* **2002**, *3*, 741–746.
308. A. Matthies; M. Nimtz; S. Leimkühler, *Biochemistry* **2005**, *44* (21), 7912–7920.
309. K. Krepinsky; S. Leimkühler, *FEBS J.* **2007**, *274*, 2778–2787.
310. Z. Marelja; W. Stöcklein; M. Nimtz; S. Leimkühler, *J. Biol. Chem.* **2008**, *283*, 25178–25185.
311. R. J. Dawson; K. P. Locher, *Nature* **2006**, *443* (7108), 180–185.
312. M. Tejada-Jimenez; A. Llamas; E. Sanz-Luque; A. Galvan; E. Fernandez, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (50), 20126–20130.
313. H. Tomatsu; J. Takano; H. Takahashi; A. Watanabe-Takahashi; N. Shibagaki; T. Fujiwara, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (47), 18807–18812.
314. M. T. Liu; M. M. Wuebbens; K. V. Rajagopalan; H. Schindelin, *J. Biol. Chem.* **2000**, *275*, 1814–1822.
315. S. Xiang; J. Nichols; K. V. Rajagopalan; H. Schindelin, *Structure* **2001**, *9*, 299–310.
316. J. D. Schrag; W. Huang; J. Sivaraman; C. Smith; J. Plamondon; R. Larocque; A. Matte; M. Cygler, *J. Mol. Biol.* **2001**, *310*, 419–431.
317. J. Nichols; K. V. Rajagopalan, *J. Biol. Chem.* **2002**, *277*, 24995–25000.
318. J. D. Nichols; K. V. Rajagopalan, *J. Biol. Chem.* **2005**, *280* (9), 7817–7822.
319. J. Kuper; A. Llamas; H. J. Hecht; R. R. Mendel; G. Schwarz, *Nature* **2004**, *430* (7001), 803–806.
320. B. Stallmeyer; A. Nerlich; J. Schiemann; H. Brinkmann; R. R. Mendel, *Plant J.* **1995**, *8*, 751–762.
321. A. Llamas; R. R. Mendel; G. Schwarz, *J. Biol. Chem.* **2004**, *279* (53), 55241–55246.
322. A. Llamas; T. Otte; G. Multhaup; R. R. Mendel; G. Schwarz, *J. Biol. Chem.* **2006**, *281* (27), 18343–18350.
323. M. S. Morrison; P. A. Cobine; E. L. Hegg, *J. Biol. Inorg. Chem.* **2007**, *12* (8), 1129–1139.
324. M. Neumann; S. Leimkühler, *FEBS J.* **2008**, *275* (22), 5678–5689.
325. S. Leimkühler; K. V. Rajagopalan, *J. Biol. Chem.* **2001**, *276* (3), 1837–1844.
326. B. Stallmeyer; G. Schwarz; J. Schulze; A. Nerlich; J. Reiss; J. Kirsch; R. R. Mendel, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1333–1338.
327. I. Paarmann; T. Saiyed; B. Schmitt; H. Betz, *Biochem. Soc. Trans.* **2006**, *34* (Pt. 1), 45–47.
328. L. Loschi; S. J. Broxk; T. L. Hills; G. Zhang; M. G. Bertero; A. L. Lovering; J. H. Weiner; N. C. Strynadka, *J. Biol. Chem.* **2004**, *279* (48), 50391–50400.
329. J. L. Johnson; N. R. Bastian; K. V. Rajagopalan, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3190–3194.
330. T. Palmer; I. P. Goodfellow; R. E. Sockett; A. G. McEwan; D. H. Boxer, *Biochem. Biophys. Acta* **1998**, *1395*, 135–140.
331. T. Palmer; C.-L. Santini; C. Iobbi-Nivolo; D. J. Eaves; D. H. Boxer; G. Giordano, *Mol. Microbiol.* **1996**, *20*, 875–884.
332. K. McLuskey; J. A. Harrison; A. W. Schuttelkopf; D. H. Boxer; W. N. Hunter, *J. Biol. Chem.* **2003**, *278* (26), 23706–23713.
333. C. A. Temple; K. V. Rajagopalan, *J. Biol. Chem.* **2000**, *275* (51), 40202–40210.
334. M. W. Lake; C. A. Temple; K. V. Rajagopalan; H. Schindelin, *J. Biol. Chem.* **2000**, *275* (51), 40211–40217.
335. C. E. Stevenson; F. Sargent; G. Buchanan; T. Palmer; D. M. Lawson, *Structure* **2000**, *8*, 1115–1125.
336. M. Neumann; W. Stöcklein; S. Leimkühler, *J. Biol. Chem.* **2007**, *282* (39), 28493–28500.
337. G. Börner; M. Karrasch; R. K. Thauer, *FEBS Lett.* **1991**, *290*, 31–34.
338. J. C. Hilton; K. V. Rajagopalan, *Arch. Biochem. Biophys.* **1996**, *325*, 139–143.
339. R. C. Wahl; C. K. Warner; V. Finnerty; K. V. Rajagopalan, *J. Biol. Chem.* **1982**, *257*, 3958–3962.
340. L. Amrani; J. Primus; A. Glatigny; L. Arcangeli; C. Scazzocchio; V. Finnerty, *Mol. Microbiol.* **2000**, *38* (1), 114–125.
341. T. Heidenreich; S. Wollers; R. R. Mendel; F. Bittner, *J. Biol. Chem.* **2005**, *280* (6), 4213–4218.
342. F. Bittner; M. Oreb; R. R. Mendel, *J. Biol. Chem.* **2001**, *276* (44), 40381–40384.
343. S. Wollers; T. Heidenreich; M. Zarepour; D. Zachmann; C. Kraft; Y. Zhao; R. R. Mendel; F. Bittner, *J. Biol. Chem.* **2008**, *283*, 9642–9650.



344. S. Leimkühler; W. Klipp, *J. Bacteriol.* **1999**, *181* (9), 2745–2751.
345. M. Neumann; M. Schulte; N. Jünemann; W. Stöcklein; S. Leimkühler, *J. Biol. Chem.* **2006**, *281* (23), 15701–15708.
346. M. Neumann; W. Stöcklein; A. Walburger; A. Magalon; S. Leimkühler, *Biochemistry* **2007**, *46* (33), 9586–9595.
347. F. Blasco; J. P. Dos Santos; A. Magalon; C. Frixon; B. Guigliarelli; C. L. Santini; G. Giordano, *Mol. Microbiol.* **1998**, *28*, 435–447.
348. M. Ilbert; V. Mejean; M. T. Giudici-Orticoni; J. P. Samama; C. Iobbi-Nivol, *J. Biol. Chem.* **2003**, *278*, 28787–28792.
349. S. Tranier; C. Iobbi-Nivol; C. Birck; M. Ilbert; I. Mortier-Barriere; V. Mejean; J. P. Samama, *Structure* **2003**, *11*, 165–174.
350. O. Genest; M. Ilbert; V. Mejean; C. Iobbi-Nivol, *J. Biol. Chem.* **2005**, *280* (16), 15644–15648.
351. I. J. Oresnik; C. L. Ladner; R. J. Turner, *Mol. Microbiol.* **2001**, *40* (2), 323–331.
352. R. L. Jack; G. Buchanan; A. Dubini; K. Hatzixanthis; T. Palmer; F. Sargent, *EMBO J.* **2004**, *23* (20), 3962–3972.
353. A. Vergnes; J. Pommier; R. Toci; F. Blasco; G. Giordano; A. Magalon, *J. Biol. Chem.* **2006**, *281* (4), 2170–2176.
354. A. Vergnes; K. Gouffi-Belhabich; F. Blasco; G. Giordano; A. Magalon, *J. Biol. Chem.* **2004**, *279* (40), 41398–41403.
355. C. P. Witte; M. I. Igeno; R. Mendel; G. Schwarz; E. Fernandez, *FEBS Lett.* **1998**, *431* (2), 205–209.
356. F. S. Ataya; C. P. Witte; A. Galvan; M. I. Igeno; E. Fernandez, *J. Biol. Chem.* **2003**, *278* (13), 10885–10890.
357. K. Fischer; A. Llamas; M. Tejada-Jimenez; N. Schrader; J. Kuper; F. S. Ataya; A. Galvan; R. R. Mendel; E. Fernandez; G. Schwarz, *J. Biol. Chem.* **2006**, *281* (40), 30186–30194.
358. M. Aguilar; K. Kalakoutskii; J. Cardenas; E. Fernandez, *FEBS Lett.* **1992**, *307* (2), 162–163.
359. R. R. Mendel, *J. Exp. Bot.* **2007**, *58* (9), 2289–2296.
360. R. R. Mendel; G. Schwarz, *Met. Ions. Biol. Syst.* **2002**, *39*, 317–368.
361. J. L. Johnson; S. K. Wadman, Molybdenum Cofactor Deficiency and Isolated Sulfite Oxidase Deficiency. In *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed.; C. R. Scriver; A. L. Beaudet; W. S. Sly; D. Valle, Eds.; McGraw-Hill: New York, 1995; pp 2271–2283.
362. M. Duran; F. A. Beemer; C. van der Heiden; J. Korteland; P. K. de Bree; M. Brink; S. K. Wadman; I. Lombeck, *J. Inher. Metab. Dis.* **1978**, *1*, 175–178.
363. J. L. Johnson, K. E. Coyne, K. V. Rajagopalan, J. L. K. Van Hove, M. Mackay, J. Pitt; A. Boneh, *Am. J. Med. Genet.* **2001**, *104*, 169–173.
364. J. L. Johnson; K. V. Rajagopalan, *J. Inher. Metab. Dis.* **1995**, *18*, 40–47.
365. J. L. Johnson; S. K. Wadman, Molybdenum Cofactor Deficiency. In *The Metabolic Basis of Inherited Disease*, 6th ed.; C. R. Scriver; A. L. Beaudet; W. S. Sly; D. Valle, Eds.; McGraw-Hill: New York, 1989; pp 1463–1475.
366. J. Reiss; S. Gross-Hardt; E. Christensen; P. Schmidt; R. R. Mendel; G. Schwarz, *Am. J. Hum. Genet.* **2001**, *68* (1), 208–213.
367. J. Reiss; E. Christensen; C. Dorche, *Prenat. Diagn.* **1999**, *19*, 386–388.
368. H. J. Lee; I. M. Adham; G. Schwarz; M. Kneussel; J. O. Sass; W. Engel; J. Reiss, *Hum. Mol. Genet.* **2002**, *11* (26), 3309–3317.
369. S. Kramer; R. V. Hageman; K. V. Rajagopalan, *Arch. Biochem. Biophys.* **1984**, *233* (2), 821–829.
370. G. Schwarz; J. A. Santamaria-Araujo; S. Wolf; H. J. Lee; I. M. Adham; H. J. Grone; H. Schwegler; J. O. Sass; T. Otte; P. Hanzelmann; R. R. Mendel; W. Engel; J. Reiss, *Hum. Mol. Genet.* **2004**, *13* (12), 1249–1255.
371. S. Kugler; R. Hahnewald; M. Garrido; J. Reiss, *Am. J. Hum. Genet.* **2007**, *80* (2), 291–297.
372. J. Reiss; J. L. Johnson, *Hum. Mutat.* **2003**, *21*, 569–576.

### Biographical Sketches



Markus Fischer obtained his Ph.D. degree from the Technical University of Munich. In 2003, he completed his habilitation in food chemistry and biochemistry. In 2004, he was honored with the Kurt-Taeufel award from the Gesellschaft Deutscher Chemiker. During the years 2002–06, he worked as a guest lecturer at the German Institute of Science and Technology, Singapore. Since 2006, he is a Full Professor and the Director of the Institute of Food Chemistry at the University of Hamburg. His research interests are in the study of the biosynthesis of riboflavin and folic acid. His research deals with the characterization of biochemical reactions at the level of genes, proteins, metabolites, and mechanisms. The development of drugs and evaluation of drug targets are major objects of the group.



Beat Thöny was born in 1958 in Grüşch, Switzerland. He obtained his academic education at the Swiss Federal Institute of Technology (ETH) Zürich, and his Ph.D. in 1989 from the laboratory of Hauke Hennecke at the ETH Zürich in the field of molecular biology and bacterial genetics. He later took up postdoctoral training and moved to the Department of Biochemistry at Stanford University Medical School to join the laboratory of Arthur Kornberg, studying protein biochemistry and enzymology of DNA replication. Since 1992, he is a research group leader in the Division of Clinical Chemistry and Biochemistry at the University Children's Hospital in Zürich, Switzerland, where he specialized in Clinical-Chemical Analytical Medicine (FAMH) in 2004. In 1998, he obtained his habilitation and in 2005 his titular professorship from the Medical Faculty of the University of Zürich. His responsibilities include molecular diagnosis for genetic and metabolic disorders of childhood, and his research interests focus on metabolic diseases, including gene therapy.



Silke Leimkühler was born in Munich, Germany and obtained her Diploma in Biology at the University of Bielefeld. She received her Ph.D. in 1998 working under the direction of Werner Klipp at the Ruhr-University of Bochum and performed postdoctoral studies at the Duke University in the laboratory of K. V. Rajagopalan from 1999 to 2001. During her postdoctoral training, she was inspired by 'Raj,' the founder of the field of molybdoenzymes, to study the biosynthesis of the molybdenum cofactor and the enzymology of selected molybdoenzymes. She returned to Germany to the TU Braunschweig to continue research on molybdoenzymes, where she obtained her habilitation in biochemistry and microbiology in 2004. The Emmy-Noether program of the Deutsche Forschungsgemeinschaft (DFG) funded this study. In 2004, she received a junior professorship at the University of Potsdam and also a Heisenberg fellowship from the DFG in 2005. Since 2009 she is a Full Professor at the University of Potsdam.

## 7.18 Cofactor Catabolism

Tathagata Mukherjee, Kathryn M. McCulloch, Steven E. Ealick, and Tadhg P. Begley,

Cornell University, Ithaca, NY, USA

© 2010 Elsevier Ltd. All rights reserved.

---

7.18.1	Introduction	649
7.18.2	Vitamin B <sub>6</sub> Catabolism	649
7.18.3	Heme Catabolism	652
7.18.4	Vitamin B <sub>3</sub> Catabolism	657
7.18.5	Vitamin B <sub>2</sub> Catabolism	659
7.18.6	Vitamin B <sub>1</sub> Catabolism	659
7.18.7	Vitamin B <sub>9</sub> , Folate Catabolism	659
7.18.8	Vitamin B <sub>7</sub> , Biotin Catabolism	662
7.18.9	Lipoate Catabolism	663
7.18.10	Other Cofactors	663
7.18.11	Conclusion	663
References		670

---

### 7.18.1 Introduction

In contrast to our understanding of the biosynthesis of cofactors, relatively little is known about cofactor degradation. Some previous research has been carried out to identify intermediates on these catabolic pathways, but very little information is available on the genes involved and on the enzymology. In this chapter we summarize our current understanding of the pyridoxal phosphate, riboflavin, heme, thiamin, biotin, nicotinamide adenine dinucleotide (NAD), folate, lipoate, and coenzyme A catabolic pathways in all life-forms and discuss mechanistic aspects of the most interesting catabolic reactions.

### 7.18.2 Vitamin B<sub>6</sub> Catabolism

The catabolic pathway of vitamin B<sub>6</sub> (**1**) is probably the best studied. In animals (including humans) 4-pyridoxic acid (**4**) is the primary catabolic product of vitamin B<sub>6</sub>, found in urine.<sup>1</sup> It is formed by the oxidation of pyridoxal (**2**) by a nonspecific flavin adenine dinucleotide (FAD)-dependent aldehyde oxidase. The catalytically active form of vitamin B<sub>6</sub>, pyridoxal-5'-phosphate, PLP (**17**), does not undergo similar oxidation to form **4**.<sup>2,3</sup> The various forms of vitamin B<sub>6</sub> (**1**, **2**, **15**) and their respective phosphate esters (**16**, **17**) are readily enzymatically interconvertible.<sup>4</sup> Further degradation of **4** is unlikely in humans as the subjects administered with doses of **4** were found to excrete it quantitatively.<sup>5</sup> Estimation of **4** in urine serves as a nutritional marker, as lower than normal amounts of **4** is indicative of vitamin B<sub>6</sub> deficiency.<sup>6</sup>

The catabolic pathway in microorganisms does not stop at **4**. Selective culture techniques were used by Snell and coworkers<sup>7</sup> to identify 10 different bacterial strains and one yeast strain (**Table 1**), which were able to grow on one or more forms of vitamin B<sub>6</sub> as the sole source of carbon and nitrogen. The catabolic intermediates have been isolated and characterized from the growth medium,<sup>7-9</sup> and two catabolic pathways have been proposed as shown in **Figure 1**. Both these pathways are inducible and the enzymes catalyzing these steps have also been characterized.

In pathway A (**Figure 1**), observed in *Pseudomonas* MA-1, pyridoxal (**2**) is produced either from pyridoxamine (**15**) by a transamination reaction with pyruvate catalyzed by pyridoxamine pyruvate transaminase,<sup>10-13</sup> or from pyridoxine (**1**) by an oxidation reaction catalyzed by the FAD-dependent

**Table 1** List of organisms that were isolated by selective culture techniques, which could use vitamin B<sub>6</sub> as the sole source of carbon and nitrogen<sup>7</sup>

Sl. no.	Organism	Growth after 4 days		
		Pyridoxine	Pyridoxamine	Pyridoxal
1	<i>Pseudomonas</i> sp. IA	Good	Moderate	No
2	<i>Pseudomonas</i> sp. IB	Good	No	No
3	<i>Pseudomonas</i> sp. IC	Good	Moderate	No
4	<i>Pseudomonas</i> sp. ID	Good	No	No
5	<i>Pseudomonas</i> sp. MA	No	Good	No
6	<i>Pseudomonas</i> sp. MB	No	Good	No
7	<i>Pseudomonas</i> sp. MC	Moderate	Good	No
8	<i>Torulopsis niger</i> L	Moderate	No	Moderate
9	<i>Pseudomonas</i> sp. S	No	Poor	Moderate
10	<i>Pseudomonas</i> sp. YA	Poor	Poor	Poor
11	<i>Pseudomonas</i> sp. YB	Poor	Poor	Poor

pyridoxine 4-oxidase.<sup>14</sup> Pyridoxal (2) is then oxidized to form the lactone (3) by the NAD-dependent pyridoxal dehydrogenase,<sup>15</sup> which is then hydrolyzed to form 4 by 4-pyridoxolactonase.<sup>15,16</sup> FAD-dependent 4-pyridoxic acid dehydrogenase catalyzes the two-electron oxidation of 4 to form the aldehyde (5),<sup>17,18</sup> which then undergoes a two-electron oxidation by the NAD-dependent 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylic acid dehydrogenase to form the diacid (6).<sup>17,19</sup> The diacid (6) undergoes a decarboxylation at C4 by an Mn<sup>2+</sup>-dependent 2-methyl-3-hydroxypyridine-4,5-dicarboxylic acid decarboxylase to form 7,<sup>20</sup> which undergoes a unique ring-opening oxidation reaction catalyzed by the FAD-dependent 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (MHPCO) to produce the first acyclic compound (8) of the pathway.<sup>21–23</sup> Compound 8 is hydrolyzed to form succinic semialdehyde (9), acetate, ammonia, and carbon dioxide by 2-(*N*-acetamidomethylene) succinic acid hydrolase.<sup>24,25</sup>

In pathway B (Figure 1), observed in *Pseudomonas* IA and *Arthrobacter* Cr-7, pyridoxine (1) is oxidized to the aldehyde (10) at C5 by an FAD-dependent pyridoxine-5-dehydrogenase,<sup>14,26</sup> which further undergoes oxidation by the NAD-dependent isopyridoxal dehydrogenase to produce the lactone (11),<sup>19</sup> which then undergoes hydrolysis catalyzed by 5-pyridoxolactonase to produce 5-pyridoxic acid (12).<sup>16</sup> FAD-dependent 5-pyridoxic acid oxygenase then catalyzes the oxidative ring opening reaction of 12 to form 13,<sup>23</sup> which is hydrolyzed by 2-hydroxymethyl-(*N*-acetamidomethylene) succinic acid hydrolase to form 2-(hydroxymethyl)-4-oxobutanoate (14), acetate, ammonia, and carbon dioxide.<sup>24,25</sup>

The most interesting enzyme in the pathway is the FAD-dependent oxygenase, MHPCO, that catalyzes the ring cleavage of 7. This enzyme is a member of the aromatic flavoprotein hydroxylase family.<sup>27,28</sup> The enzymatic reaction can be divided into two parts: hydroxylation and subsequent ring-cleavage reaction.<sup>28</sup> MHPCO from *Pseudomonas* sp. MA-1 has been characterized in detail. The binding studies<sup>29</sup> and effect of the substrate and FAD analogues<sup>22</sup> have been reported; the mechanism of hydroxylation<sup>30</sup> and oxygen atom transfer and product rearrangement<sup>28</sup> have been investigated; and the kinetics<sup>21</sup> and thermodynamics of the reaction have been characterized.<sup>31</sup> Based on these experimental results, two mechanisms in which flavin hydroperoxide (18) functions either as an electrophile (Figure 2(c)) or as a nucleophile have been proposed.<sup>29</sup> An alternative mechanistic hypothesis for the formation of 8 through a ketene intermediate (23) is presented in Figure 2(d). A mechanistic proposal for the subsequent hydrolysis reaction, leading to the production of 9 from 8, is shown in Figure 3.

Even though the participating catabolic enzymes were known and the various intermediates were identified, the genes involved have only recently been identified. Yagi and coworkers identified five genes in *Mesorhizobium loti* MAFF303099 involved in pyridoxine (1) catabolism. These genes encoded pyridoxamine pyruvate transaminase (mlr6806),<sup>32</sup> pyridoxine-4-oxidase (mlr6785),<sup>33</sup>

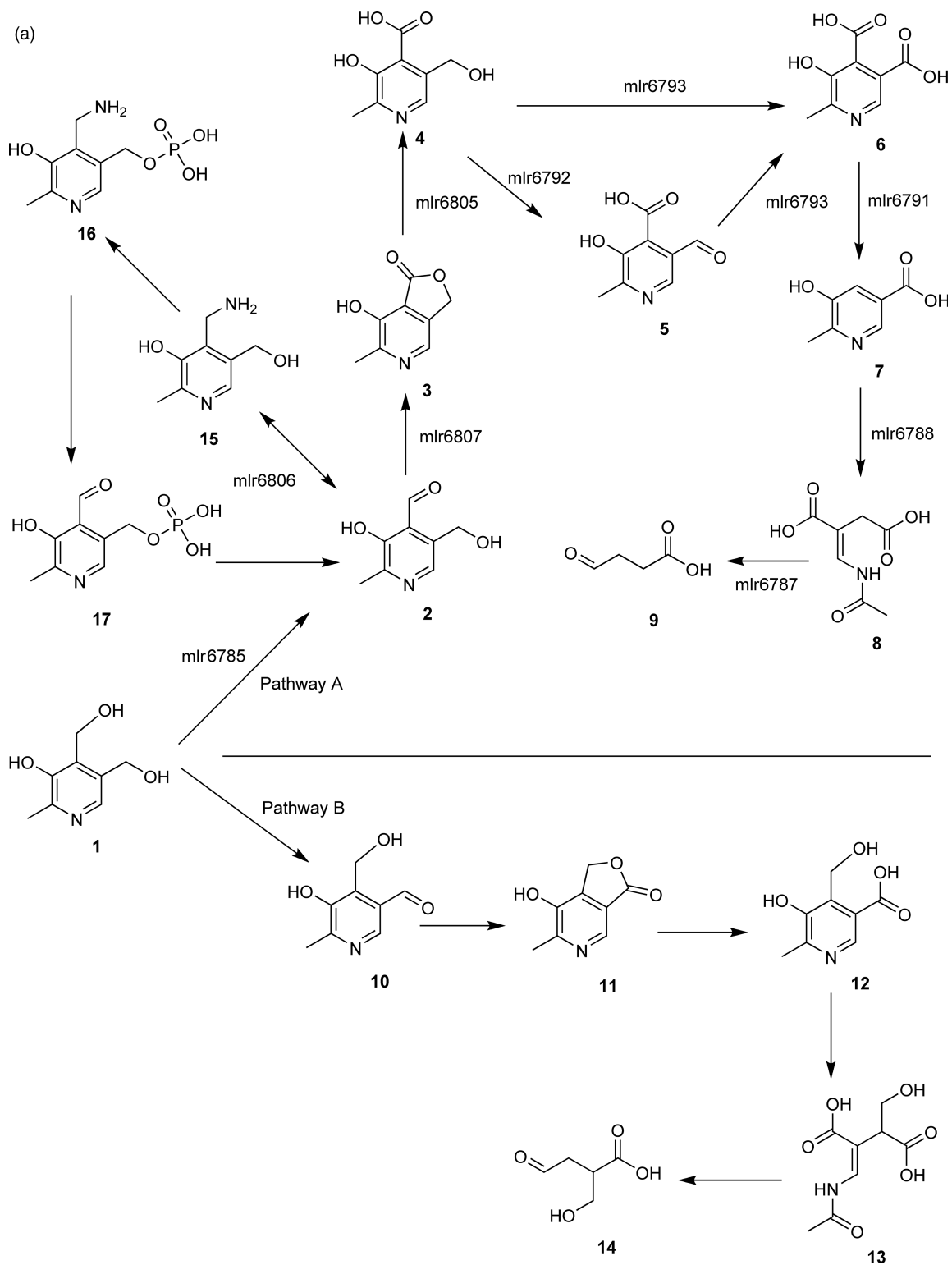
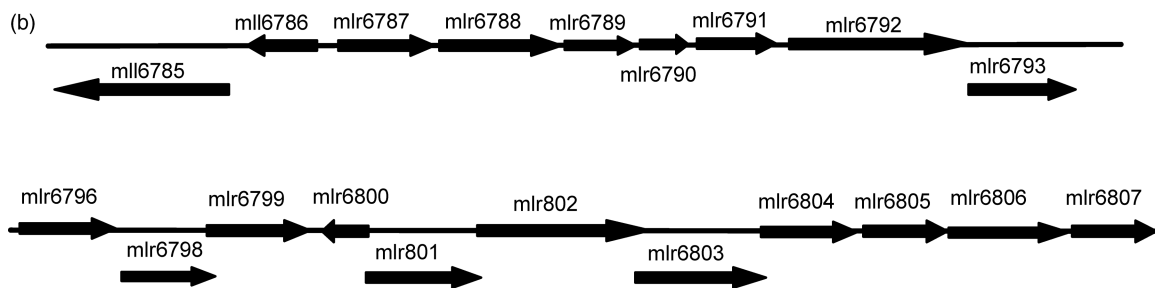


Figure 1 (Continued)



**Figure 1** (a) The vitamin B<sub>6</sub> catabolic pathways. Pathway A is observed in *Pseudomonas* sp. MA-1, while pathway B is observed in *Pseudomonas* IA and in *Arthrobacter* Cr-7. A minor variation of pathway A is seen in *Mesorhizobium loti* MAFF303099. (b) The gene organization in *M. loti* MAFF303099. The genes that have been identified to participate in PLP degradation are shown in catabolic pathway A.

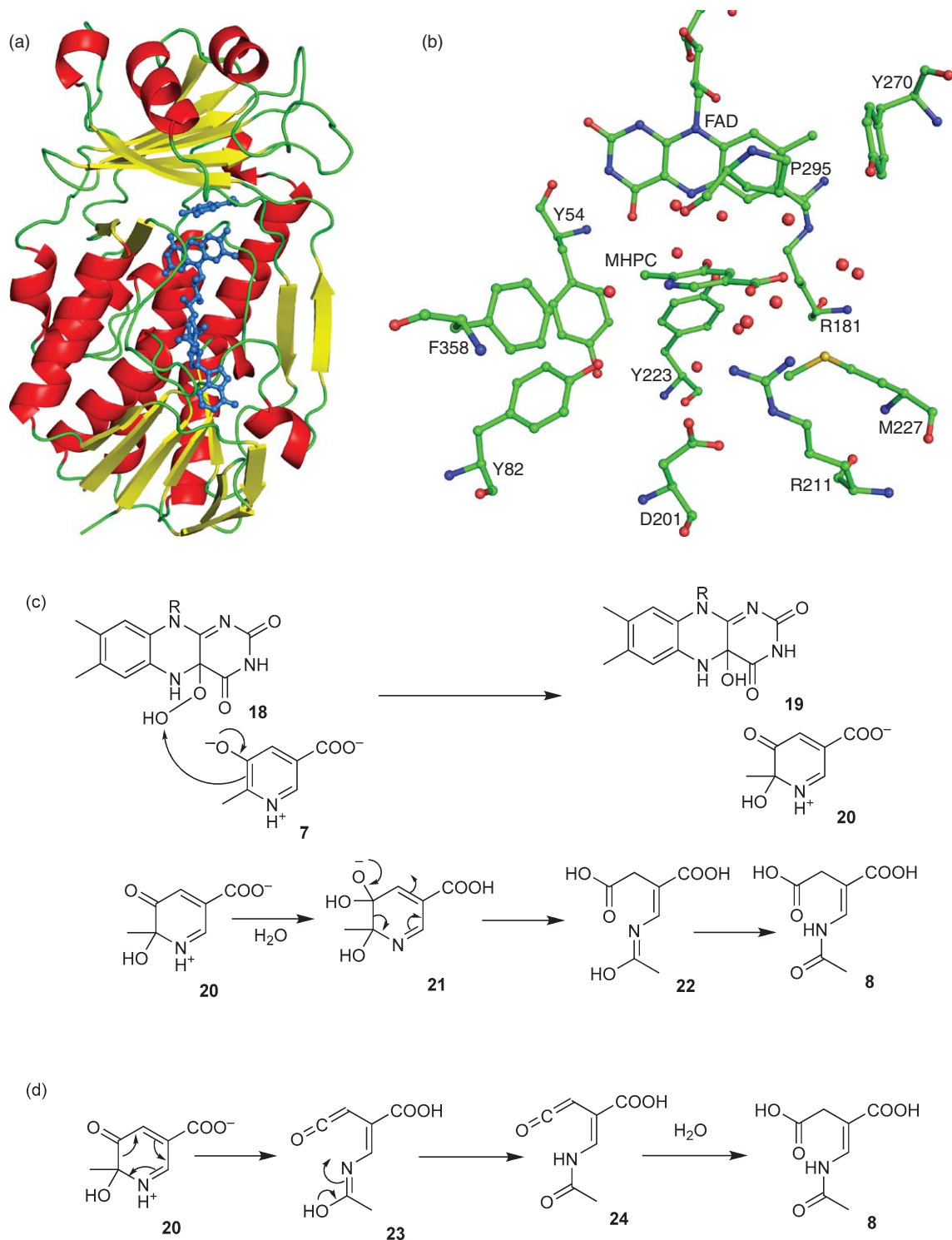
4-pyridoxolactonase (mlr6805),<sup>34</sup> pyridoxal-4-dehydrogenase (mlr6807),<sup>35</sup> and MHPCO (mlr6788).<sup>36</sup> Work done in our lab identified the remaining genes of the pathway in the same organism (4-pyridoxic acid dehydrogenase (mlr6793),<sup>37</sup> 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase (mlr6791),<sup>38</sup> and 2-(*N*-acetamidomethylene) succinic acid hydrolase (mlr6787)<sup>39</sup>) (Figure 1(b)). The crystal structure for 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase (mlr6791)<sup>38</sup> has been solved at 1.9 Å resolution (Figure 4(a)), and it revealed an active site conserved glutamate (Figure 4(b)), which is thought to participate in a retro-aldol type of reaction to decarboxylate (6) (Figure 4(c)). The crystal structure of the oxygenase, MHPCO, has also been solved at 2.1 Å resolution, with flavin and the substrate (7) bound at the active site (Figures 2(a) and 2(b)).<sup>40</sup> We had reported that a four-electron oxidation of 4-pyridoxic acid (4) to the diacid (6) was carried out by 4-pyridoxic acid dehydrogenase (mlr6793).<sup>37</sup> We had not observed any accumulation of the aldehyde (5). Later work by Yagi and coworkers showed that the mlr6792 actually coded for 4-pyridoxic acid dehydrogenase which catalyzed the two-electron oxidation of 4-pyridoxic acid (4) to the aldehyde (5), which in turn undergoes a further two-electron oxidation to form the diacid (6), catalyzed by 5-formyl-3-hydroxy-2-methylpyridine-4-carboxylic acid dehydrogenase (mlr6793).<sup>41</sup>

All the genes involved in the vitamin B<sub>6</sub> degradation have been identified in *M. loti* MAFF303099. Apart from gene identification in *M. loti*, genes encoding pyridoxal 4-dehydrogenase<sup>42</sup> and pyridoxine 4-oxidase,<sup>43,44</sup> have been identified and characterized in *Microbacterium luteolum* YK-1. The gene coding for pyridoxal 4-dehydrogenase has also been identified in *Aureobacterium luteolum*.<sup>45</sup> The remaining PLP catabolic genes have not yet been identified.

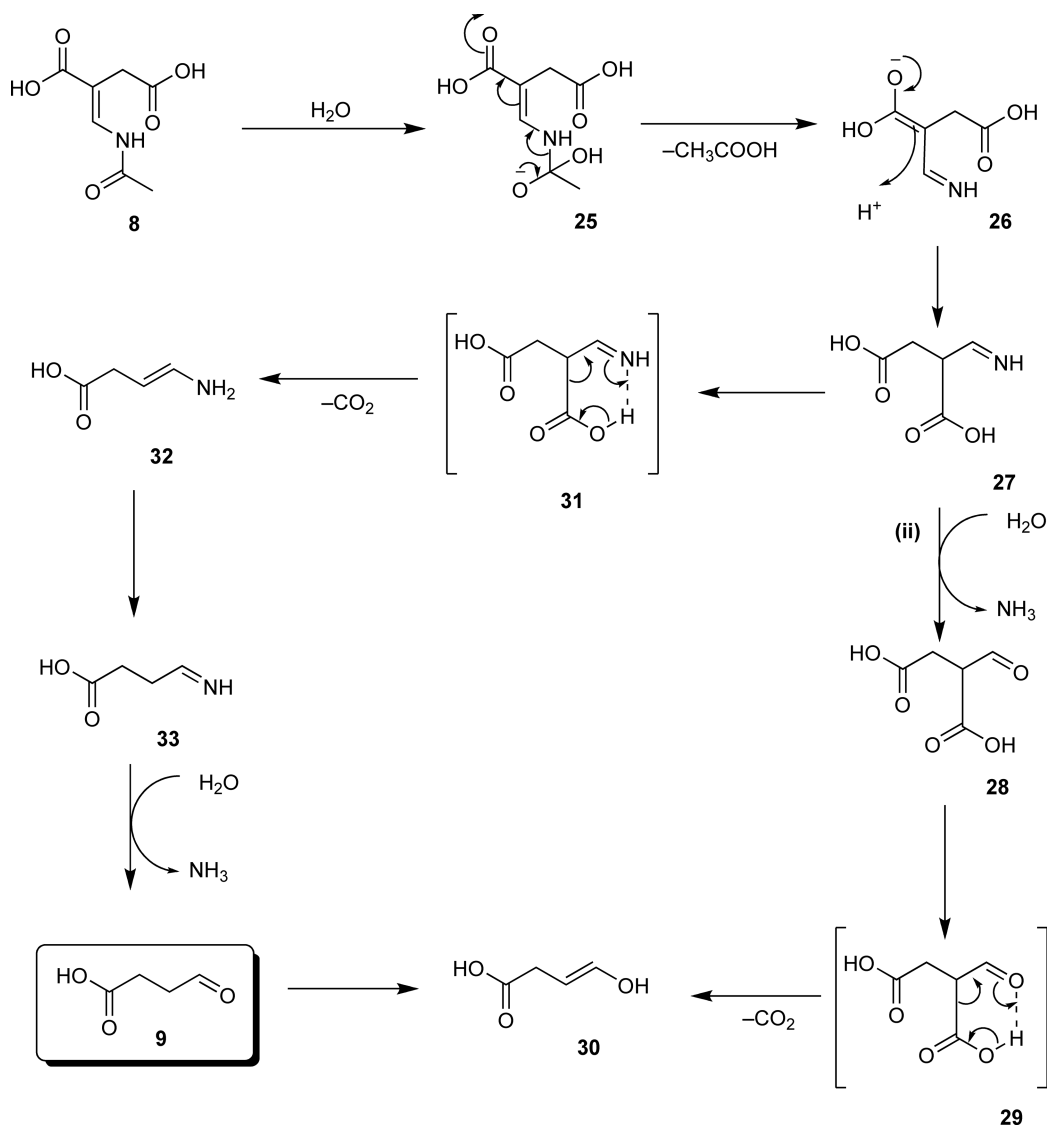
### 7.18.3 Heme Catabolism

The first step of heme (35) catabolism is the same in prokaryotes, eukaryotes, plants, and mammals.<sup>46–49</sup> Heme (35) undergoes oxidative cleavage by the heme cofactor-dependent heme oxygenase to form biliverdin IX $\alpha$  (38) and carbon monoxide and releases iron in the +2 oxidation state<sup>50,51</sup> (Figure 5). This enzyme is used by pathogenic bacteria such as *Corynebacterium* sp.,<sup>52</sup> *Yersinia* sp.,<sup>53</sup> *Pseudomonas aeruginosa*,<sup>54</sup> *Escherichia coli*,<sup>55</sup> *Bacillus anthracis*,<sup>56</sup> and *Staphylococcus aureus*.<sup>57,58</sup> for the uptake of iron from the host. In mammals, biliverdin is further reduced to bilirubin (39) by nicotinamide adenine dinucleotide phosphate (NADPH)-dependent biliverdin reductase,<sup>59</sup> which in turn gets glucuronylated to form (40).<sup>60</sup> Bilirubin (39) is produced in plants and cyanobacteria by flavodoxin-dependent bilin reductases or NADPH-dependent biliverdin reductase.<sup>48</sup> Biliverdin IX $\alpha$  (38) serves as a precursor to various





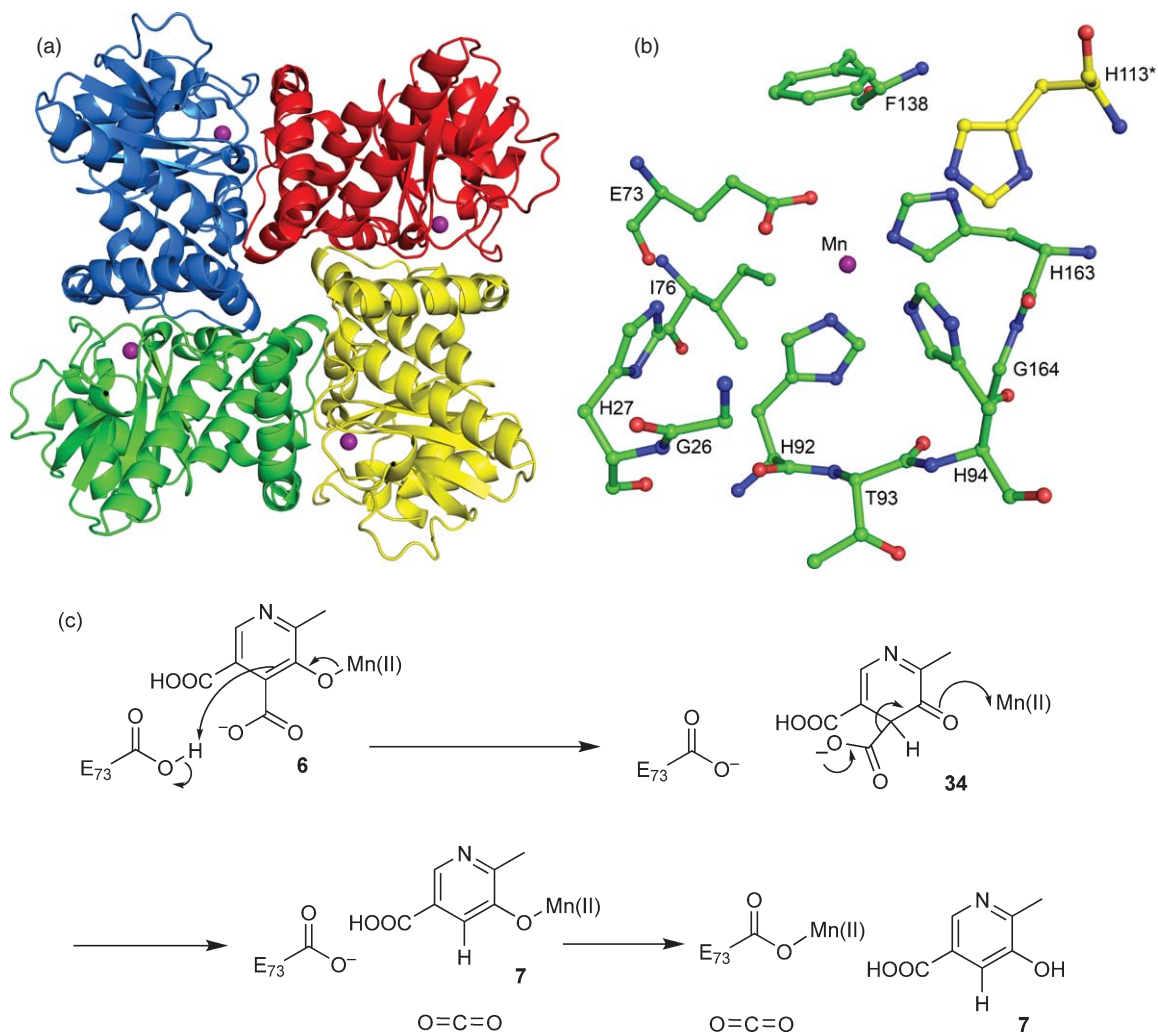
**Figure 2** (a) X-ray crystal structure of MHPCO (2.1 Å, PDB: 3gmc). The monomer is shown with FAD and MHPC bound at the active site. (b) The active site of MHPCO. (c) Proposed mechanism for the oxidative ring-opening reaction catalyzed by MHPCO with flavin hydroperoxide (**18**) acting as an electrophile. (d) An alternative mechanism for the catalytic transformation through a ketene intermediate (**23**).



**Figure 3** Proposed mechanism for the hydrolysis reaction catalyzed by 2-(*N*-acetamidomethylene) succinic acid hydrolase.

phytybilins such as **42–45**, which are produced by phytychromobilin synthase or 3-*Z*-phytychromobilin:ferredoxin oxidoreductase, 3-*Z*-phytycyanobilin:ferredoxin oxidoreductase and 15,16-dihydrobiliverdin:ferredoxin oxidoreductase, and 3-*Z*-phytycoerythrobilin:ferredoxin oxidoreductase, respectively.<sup>48,61</sup>

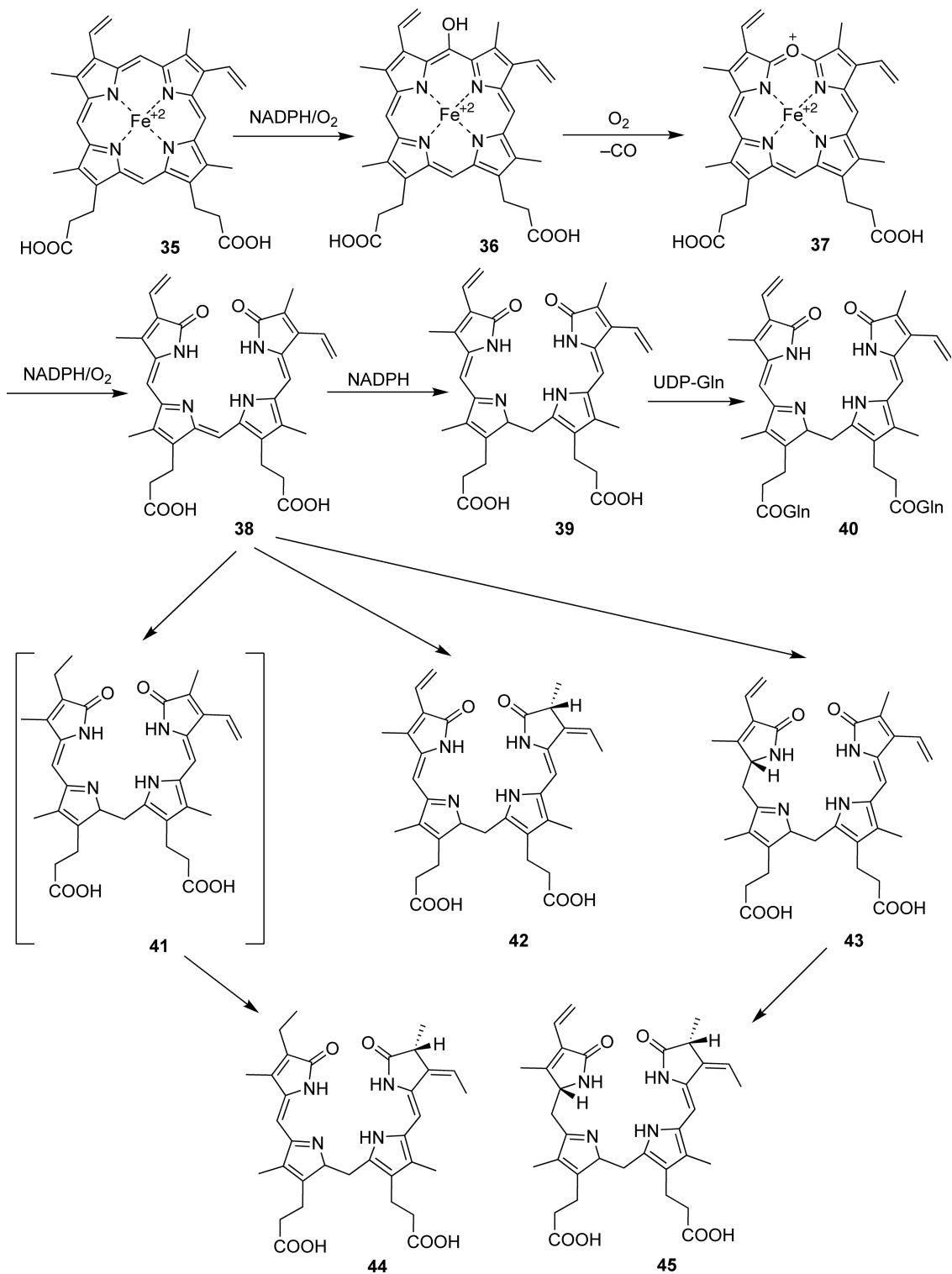
In the majority of cases, the heme oxygenase catalyzed oxidative cleavage of heme (**35**) uses heme as a cofactor and requires three molecules of oxygen<sup>62</sup> and seven electrons<sup>50</sup> to regiospecifically cleave heme (**35**) at the  $\alpha$ -meso-carbon to yield biliverdin IX $\alpha$  (**38**)<sup>63,64</sup> through  $\alpha$ -meso-hydroxyheme (**36**) and verdoheme (**37**)<sup>65,66</sup> as shown in **Figure 5**. The two added oxygen atoms are derived from separate molecules of molecular oxygen.<sup>48</sup> Various structures of heme oxygenase from human, rat, and bacteria are available with cofactor heme bound at the active site, both in the ferrous and ferric state, complexed with carbon monoxide and nitric oxide as well as the product-bound form<sup>67</sup> (**Figure 6**). A possible mechanism for the oxidative cleavage of heme, modified from the literature, is shown in **Figure 7**.<sup>49,68</sup>



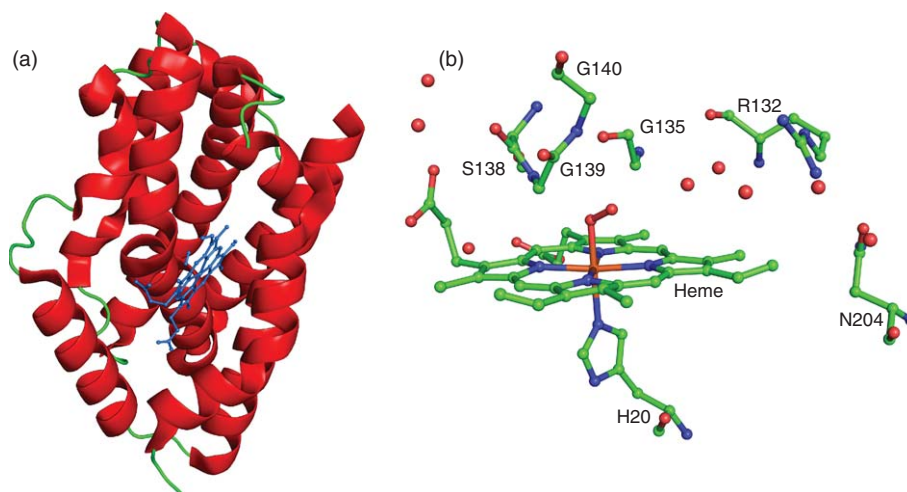
**Figure 4** (a) Crystal structure of 2-methyl-3-hydroxypyridine-4,5-dicarboxylic acid decarboxylase showing the tetrameric structure (1.9 Å, PDB: 2z7b).<sup>38</sup> The tetramer has been color coded by the subunit. The manganese ion is shown in magenta. (b) The active site of the decarboxylase, showing the Mn<sup>2+</sup> ion bound and the conserved glutamate. (c) The proposed mechanism for the decarboxylation of the diacid (**6**).

This mechanism utilizes a simple mechanistic motif consisting of iron(IV) hydroperoxide formation followed by peroxide-mediated chemistry on the porphyrin ring.

The heme oxygenase from *Pseudomonas aeruginosa*, PtgA, catalyzes a different reaction in which  $\beta$ - and  $\delta$ -meso carbon-specific cleavage of the heme (**35**) is seen, producing biliverdin IX $\beta$  and biliverdin IX $\delta$ , respectively.<sup>69,70</sup> Another variation of the heme (**35**) degradation pathway has been reported in the insect *Rhodnius prolixus*, where heme (**35**) is chemically modified before the oxidative cleavage of the tetrapyrrole-producing dicysteinybiliverdin IX $\gamma$ .<sup>71</sup> The IsdG family of heme oxygenases<sup>72</sup> is found exclusively in Gram-positive bacteria degrade heme (**35**) in a novel way to yield a product which is neither biliverdin IX nor the dicysteinybiliverdin IX $\gamma$ . Characterization of this novel product is underway in *Staphylococcus aureus*.<sup>46,73</sup>



**Figure 5** Catabolic pathway of heme.



**Figure 6** (a) Crystal structure of human heme oxygenase-1, showing the heme bound at the active site of the enzyme (1.85 Å, PDB: 1v8x).<sup>137</sup> (b) The active site of the enzyme.

#### 7.18.4 Vitamin B<sub>3</sub> Catabolism

The NAD (62) catabolic pathway, found in bacteria, plants, and animals, is shown in **Figure 8**. NAD (62) or its reduced form (NADH) is deadenylated to produce nicotinamide mononucleotide (NMN) (63) and adenosine 5'-monophosphate (AMP).<sup>74</sup> A nonspecific 5'-ribonucleotide phosphohydrolase then dephosphorylates NMN (63) to form *N*-ribosylnicotinamide (64),<sup>75</sup> which then undergoes hydrolysis to form nicotinamide (66) and *D*-ribose (65) in a reaction catalyzed by *N*-ribosylnicotinamide ribohydrolase.<sup>76</sup> Nicotinamide (66) is also formed by NAD glycohydrolase by the hydrolysis of the *N*-glycoside bond of NAD (62) forming nicotinamide (66) and ADP ribose (67).<sup>77,78</sup> Nicotinamide amidohydrolase forms nicotinate (68) from nicotinamide (66).<sup>79,80</sup> Nicotinic acid (68) can also be formed in a slightly different manner. Hydrolysis of the amide group of NMN (63) by NMN deamidase forms 69,<sup>81–83</sup> which then either undergoes a dephosphorylation by nicotinate *D*-ribonucleotide phosphohydrolase to form 70 followed by hydrolysis of the *N*-glycoside bond by nucleotide phosphatase to form nicotinate (68) and ribose 1-phosphate (71)<sup>76</sup> or  $\beta$ -nicotinate *D*-ribonucleotide (69) could undergo a direct hydrolysis of the *N*-glycoside bond by nicotinate phosphoribosyltransferase to form nicotinate (68) and (72).<sup>84</sup>

Under aerobic conditions nicotinate (68) or niacin (66) (also known as vitamin B<sub>3</sub>) undergoes hydroxylation by a two-component hydroxylase, NicAB, to form 6-hydroxynicotinate (73) (**Figure 9**).<sup>85</sup> 6-Hydroxynicotinate (73) undergoes oxidative decarboxylation at C3 catalyzed by 6-hydroxynicotinate-3-monooxygenase, NicC to form 2,5-dihydroxypyridine (74). 2,5-Dihydroxypyridine dioxygenase, NicX, which is an Fe<sup>2+</sup>-dependent extradiol ring cleavage dioxygenase, converts 74 into *N*-formyl maleamate (75), which undergoes deformylation by NicD to form maleamate (76) and formate.<sup>86–88</sup> Maleamate (76) is hydrolyzed to maleate (77) by NicF, which then isomerizes to fumarate (78) by NicE.<sup>89</sup> The Nic gene cluster was recently identified and characterized in *Pseudomonas putida* KT2440 when the organism was found to grow on nicotinic acid as the sole source of carbon.<sup>90</sup> 6-Hydroxynicotinate (73) has a different fate under anaerobic conditions. In *Eubacterium barkeri*, (79) undergoes reduction by a ferredoxin-dependent oxidoreductase to form 80, followed by hydrolysis to form 2-formylglutarate (81).<sup>91</sup> 2-Formylglutarate (81), in six enzymatic steps, can then form pyruvate (82) and propionate (83).<sup>91</sup>

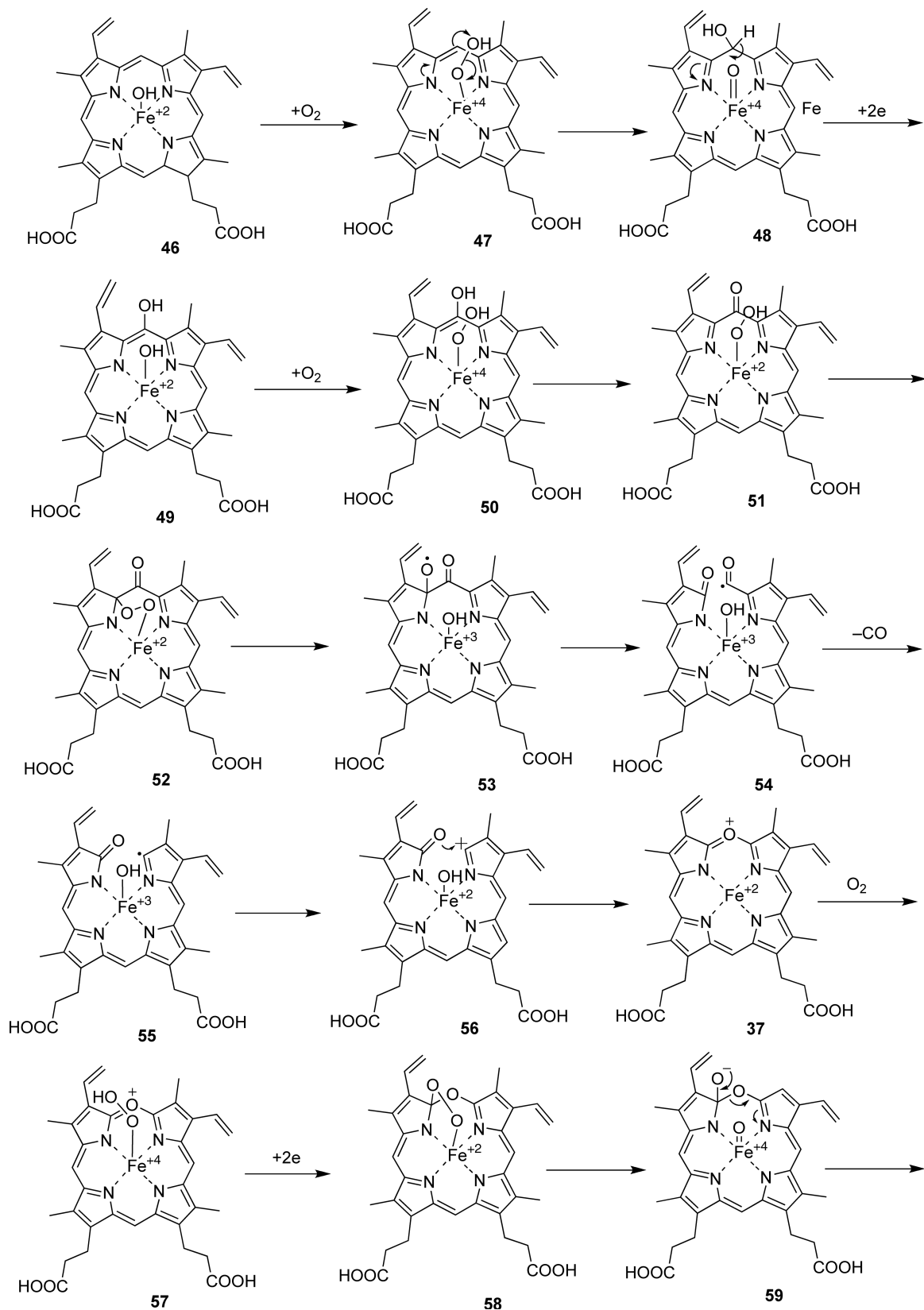
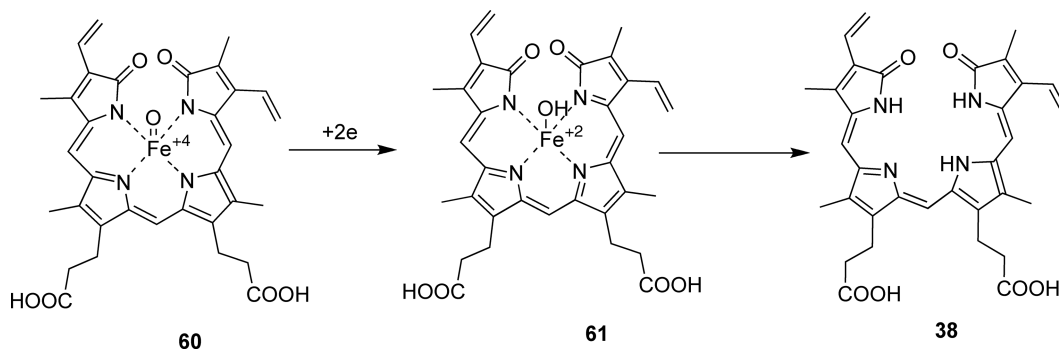


Figure 7 (Continued)





**Figure 7** The mechanism of heme oxygenase: conversion of heme (46) into-biliverdin (38).

### 7.18.5 Vitamin B<sub>2</sub> Catabolism

Much less is known about the catabolic pathway for riboflavin, vitamin B<sub>2</sub> (84). Riboflavinase that catalyzes the hydrolysis of riboflavin (84) to lumichrome (85) and ribitol (86) has been identified in *Devosia riboflavina* (Figure 10).<sup>92–94</sup> However, no further degradation of the isoalloxazine ring was reported.<sup>93</sup> Some intestinal bacteria under strictly anaerobic conditions produced (87) from riboflavin (84).<sup>95</sup> Under aerobic conditions in *Pseudomonas RF*, various intermediates (89–94) of riboflavin (84) degradation have been identified but the enzymes associated with their formation have not yet been discovered.<sup>96,97</sup> A possible mechanism for the conversion of the compound (91) to the lactone (93) has been postulated in Figure 11. In *Shinorhizobium meliloti*, the enzyme BluB is reported to catalyze a remarkable oxidation of riboflavin (84) to 5,6-dimethylbenzimidazole (88), a precursor to vitamin B<sub>12</sub>.<sup>98</sup> The crystal structure of this enzyme has been reported<sup>98</sup> but the mechanism for this oxidation reaction has not yet been established, one possibility is shown in Figure 12.<sup>98–100</sup>

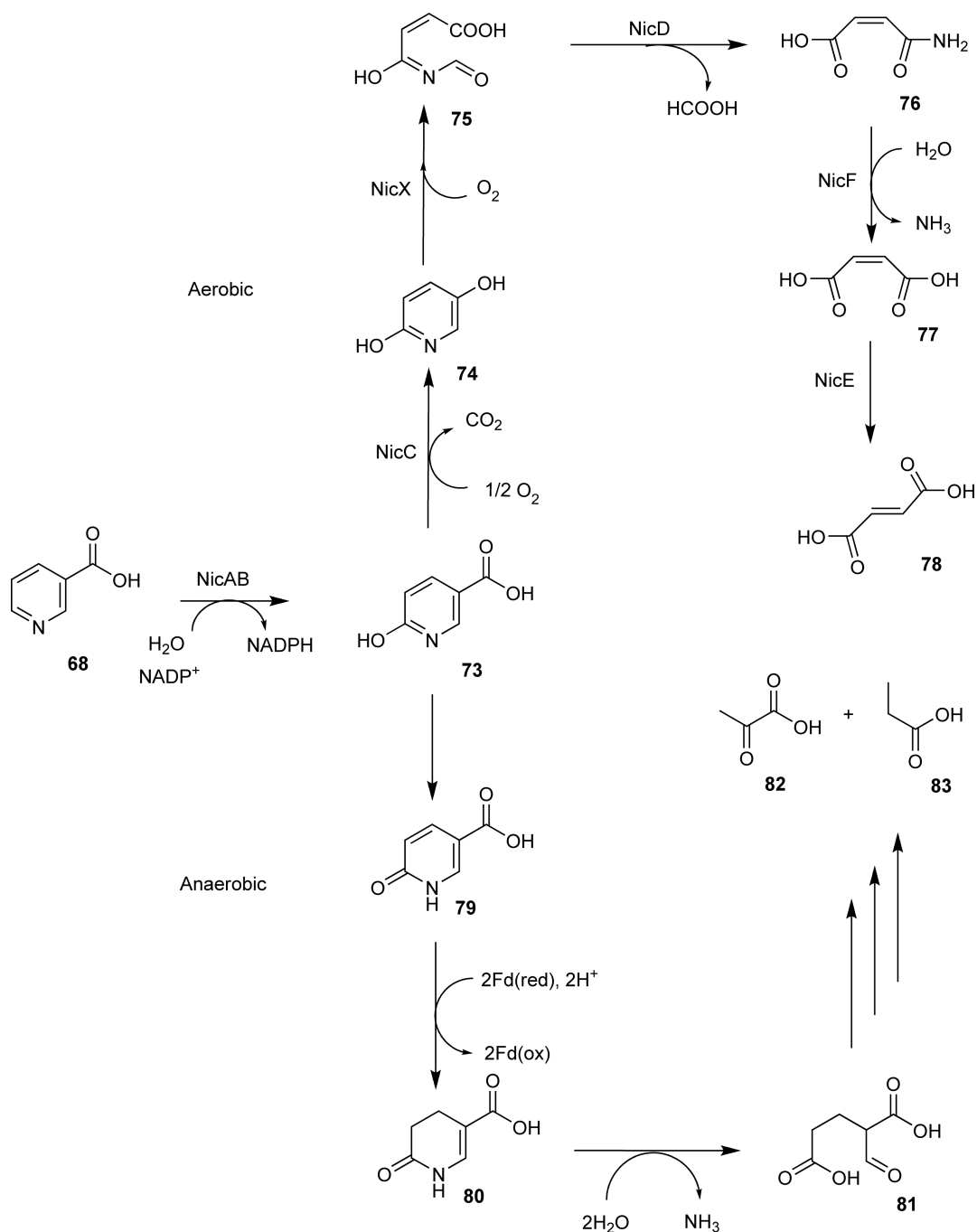
### 7.18.6 Vitamin B<sub>1</sub> Catabolism

Thiaminase I can degrade thiamin (111), thiamin monophosphate (112), and thiamin pyrophosphate (113) to the pyrimidine (114) and the corresponding thiazole (115–117) moiety (Figure 13). The enzyme from *Bacillus thiaminolyticus* has been cloned and characterized.<sup>101,102</sup> Thiaminase I uses nucleophiles such as quinoline, aniline, and pyridine but not water. Thiaminase II on the contrary uses only water as the nucleophile and accepts only thiamin (111) as the substrate.<sup>103–110</sup> Thiamin (111) forms 119 under alkaline conditions. This can be hydrolyzed to desthiothiamin (121) or oxidized to thiamin disulfide (122).<sup>111,112</sup> Compound 119 further degrades to form 124, which is hydrolyzed to 125 by YlmB. Compound 125 is hydrolyzed by TenA to form the pyrimidine 126, an intermediate in the thiamin (111) biosynthetic pathway.<sup>113</sup> The crystal structure of this enzyme has been solved at 2.9 Å resolution and the mechanism of this transformation is presented in Figure 14.<sup>114</sup> Thiamin (111) also undergoes oxidation by a flavin-dependent thiamin oxidase to thiamin acetic acid (128) through 127.<sup>115</sup>

### 7.18.7 Vitamin B<sub>9</sub>, Folate Catabolism

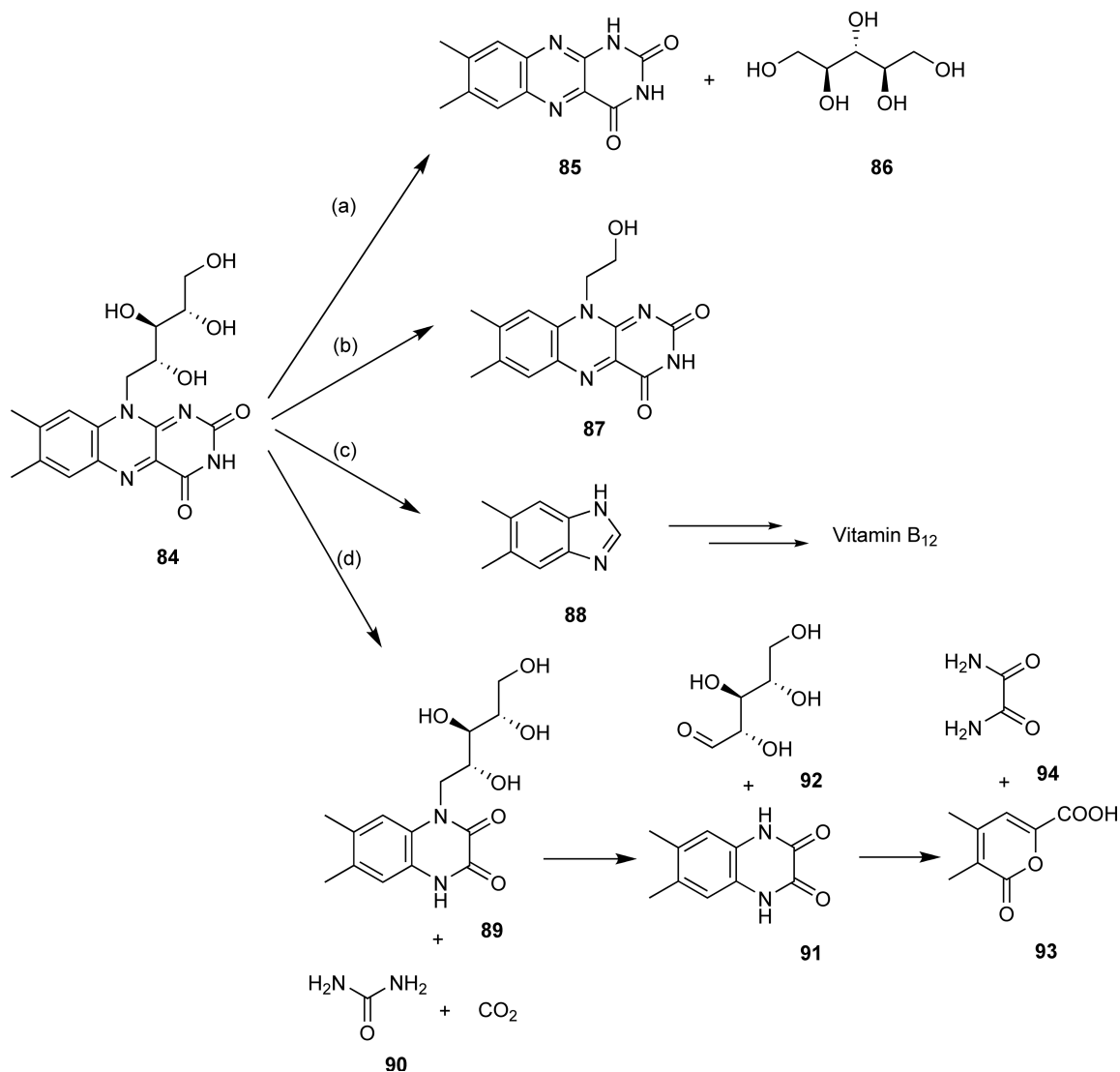
Folypoly- $\gamma$ -glutamate hydrolase catalyzes the hydrolysis of the polyglutamate peptide from intracellular folate-generating folate monoglutamate (135), which can be transported out of the cell.<sup>116</sup> The pteridine ring of the tetrahydrofolate (135) undergoes sequential oxidation to dihydrofolate (137) and





**Figure 9** Catabolic pathway of nicotinic acid (**68**) (niacin, vitamin B<sub>3</sub>).

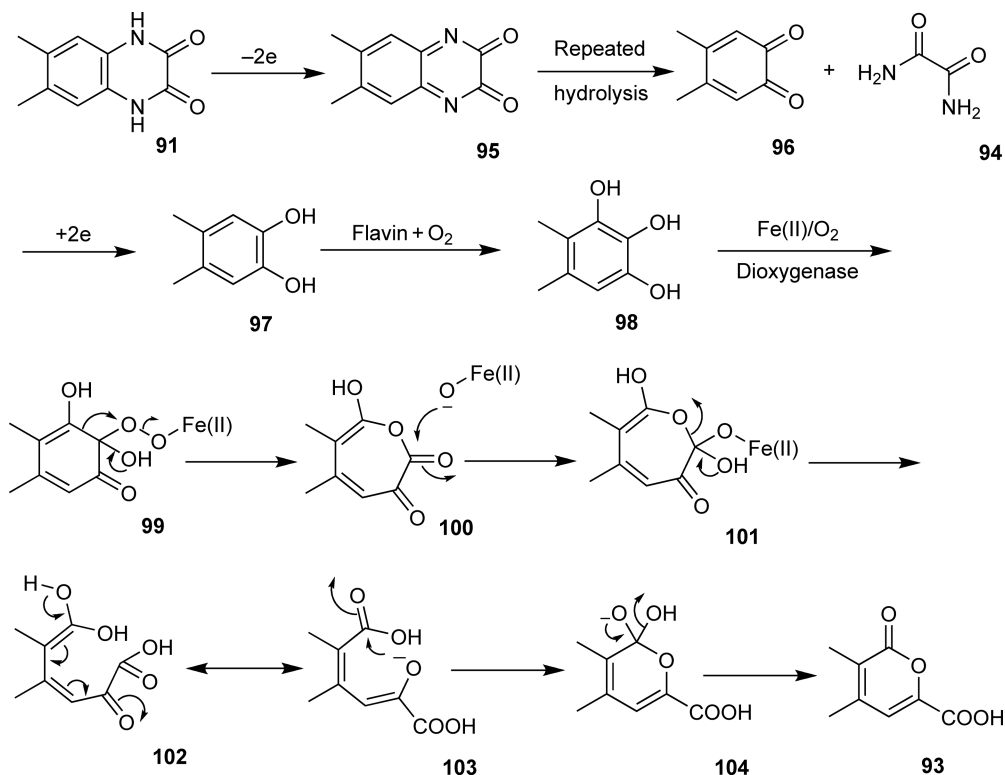
then to folic acid (**138**) through a quinoid dihydrofolate (**136**) intermediate. Alternatively, oxidation of tetrahydrofolate (**135**) can result in the formation of a C9–N10 Schiff base (**139**), which undergoes hydrolysis to yield 6-formyltetrahydropterin (**140**) and *p*-aminobenzoylglutamate (**141**).<sup>116</sup> In rat, 5-formyltetrahydrofolate, the most stable derivative of tetrahydrofolate, is oxidatively cleaved by ferritin to form *p*-aminobenzoylglutamate (**141**) and a pterin derivative.<sup>117</sup> The gene encoding for this protein has not been identified (**Figure 15**).



**Figure 10** Catabolic pathway of riboflavin (**84**) (vitamin B<sub>2</sub>). Riboflavin degradation (a) catalyzed by riboflavinase, (b) under anaerobic conditions, (c) catalyzed by BluB, and (d) under aerobic conditions.

### 7.18.8 Vitamin B<sub>7</sub>, Biotin Catabolism

Biotin (**142**) is covalently attached to biotin-dependent enzymes through an amide linkage. Recycling of biotin (**142**) is achieved by biotinidase, which hydrolyzes the amide bond releasing biotin (**142**).<sup>118</sup> Several products of vitamin B<sub>7</sub> (**142**) catabolism have been identified in *pseudomonads* grown on D-biotin (**142**) as the sole source of carbon, nitrogen, and sulfur.<sup>119–126</sup> No enzymes associated with these transformations have yet been characterized, though a plausible pathway for D-biotin (**142**) degradation to carbon dioxide and ammonia has been proposed<sup>123,126</sup> (Figure 16(a)). The identification of intermediates (**152–156**)<sup>119</sup> suggests that D-biotin (**142**) can be degraded by a  $\beta$ -oxidation pathway similar to that used for fatty acids. Degradation of D-biotin (**142**) has also been observed in fungi,<sup>127–129</sup> rats,<sup>130,131</sup> and in humans.<sup>132</sup>



**Figure 11** Mechanistic hypothesis for the conversion of **91** into the lactone (**93**).

### 7.18.9 Lipoate Catabolism

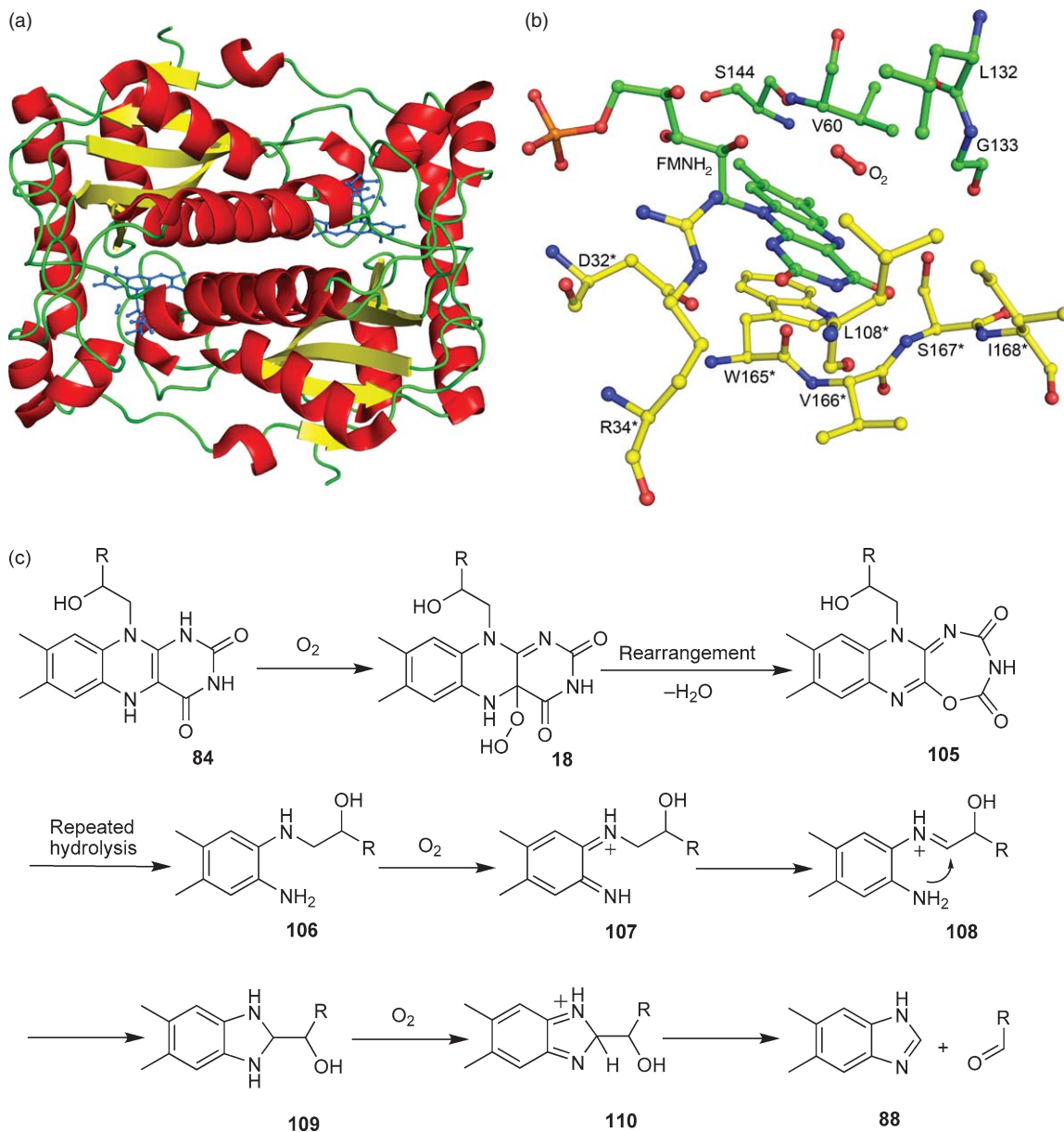
The various intermediates in lipoate (**157**) catabolism have been identified in pseudomonads that were grown on lipoate (**157**) as the sole source of carbon and sulfur<sup>133,134</sup> (Figure 17). These degraded products have also been detected in the rat.<sup>135,136</sup> Bisorlipoate (**158**), tetranorlipoate (**160**), and  $\beta$ -hydroxybisorlipoate (**159**) are the result of  $\beta$ -oxidation of the valeric side chain of lipoate (**157**).<sup>134</sup> In addition, the dithiolane ring undergoes oxidation to thiolsulfonates (**161–163**).<sup>133</sup> No enzymes or genes associated with these transformations have yet been identified.

### 7.18.10 Other Cofactors

Scant information is available on the degradation of pantothenate (**164**), vitamin B<sub>5</sub>, which is a precursor to coenzyme A. In *Pseudomonas fluorescens*, pantothenate (**164**) is hydrolyzed by pantothenase to form pantoate (**165**) and  $\beta$ -alanine (**166**) (Figure 18). Nothing is known about the degradation of menaquinone or ubiquinone.

### 7.18.11 Conclusion

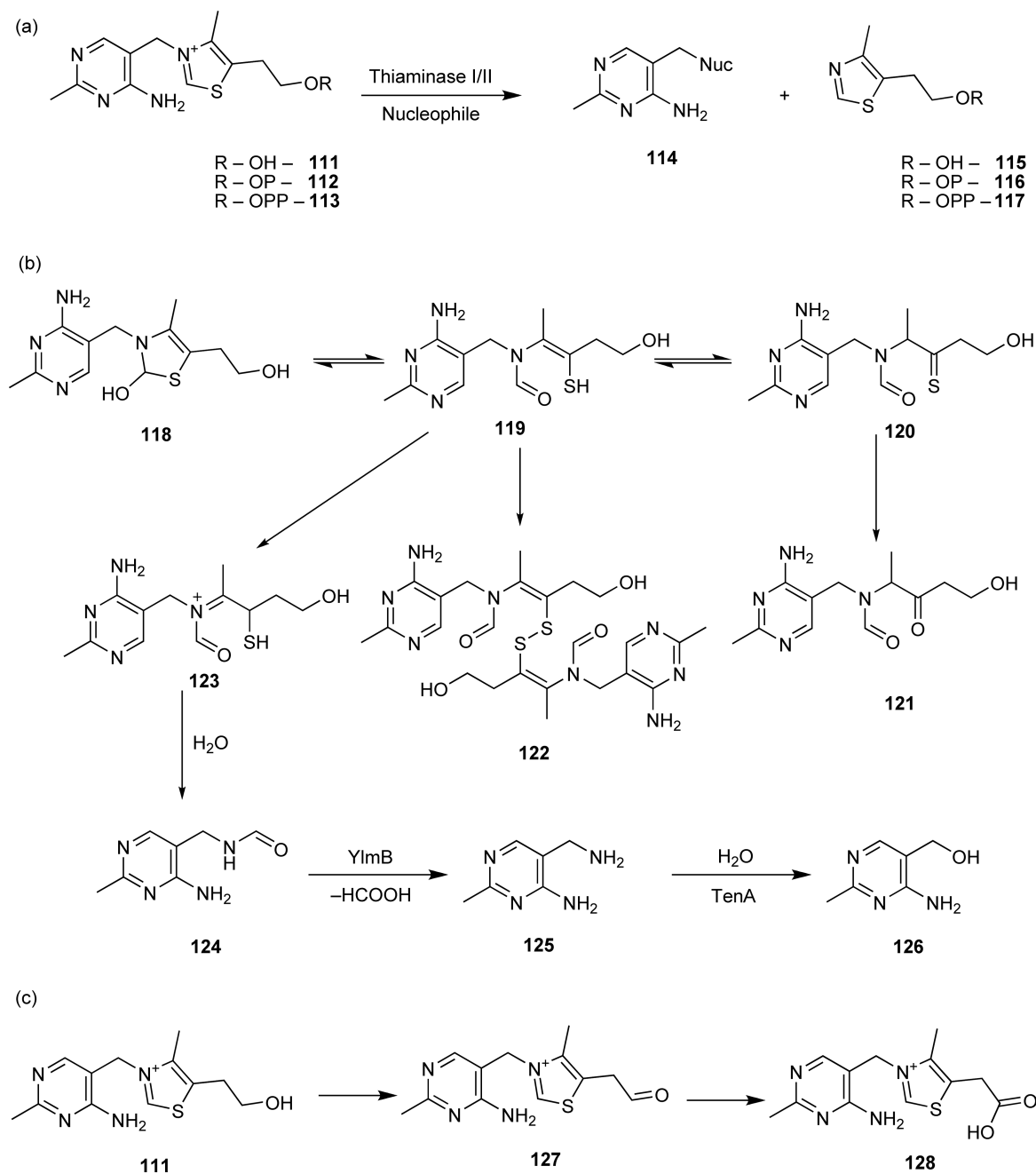
Cofactor catabolism is still an understudied area in cofactor chemistry and much research needs to be done to identify catabolic intermediates, the enzymes that catalyze the formation of these intermediates and the corresponding genes that encode these enzymes. A better knowledge of the *in vivo* stability of the cofactors



**Figure 12** (a) Crystal structure of BluB showing the flavin bound at the active site (2.9 Å, PDB: 2isi).<sup>98</sup> (b) The active site of BluB. (c) Proposed mechanism of BluB, the enzyme that generates 5,6-dimethylbenzimidazole (**88**), a precursor to vitamin B<sub>12</sub>, from riboflavin (**84**).<sup>100</sup>

is important for understanding cellular homeostasis and has potential applications in the biotechnology of vitamin production by fermentation as well as determining the sensitivity of the cell to antibiotics targeted toward vitamin biosynthetic enzymes. Cofactor catabolism is also likely to reveal new and interesting enzymology, as already found, for example, in heme and PLP catabolism (e.g., heme oxygenase and MHPCO).

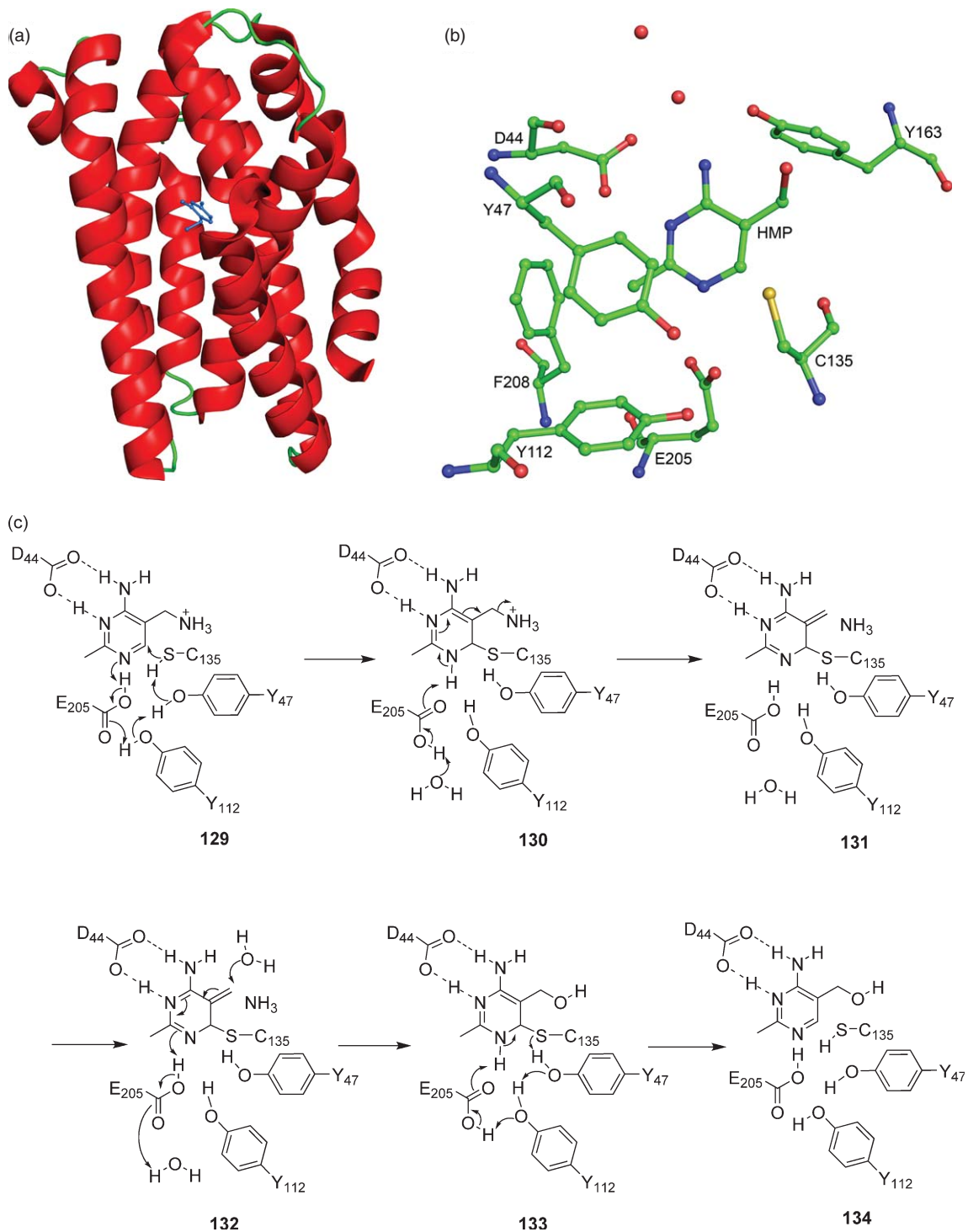




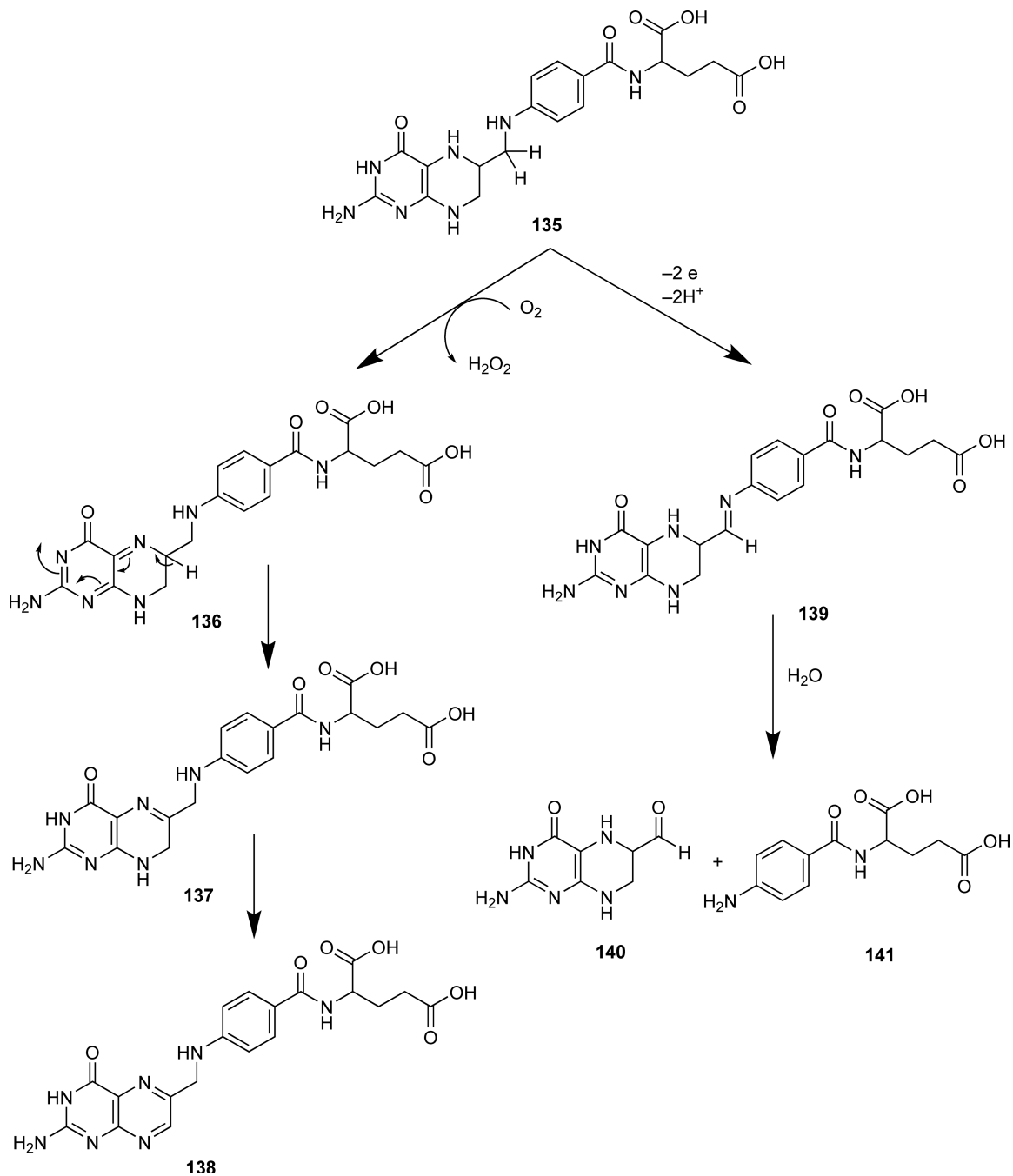
**Figure 13** Catabolic pathway of thiamin (**111**) (vitamin B<sub>1</sub>). (a) Thiaminase catalyzed degradation of thiamin (**111**); (b) salvage of base-degraded thiamin (**118**); and (c) thiamin degradation by oxidation of thiazole side chain of thiamin (**111**).

## Acknowledgment

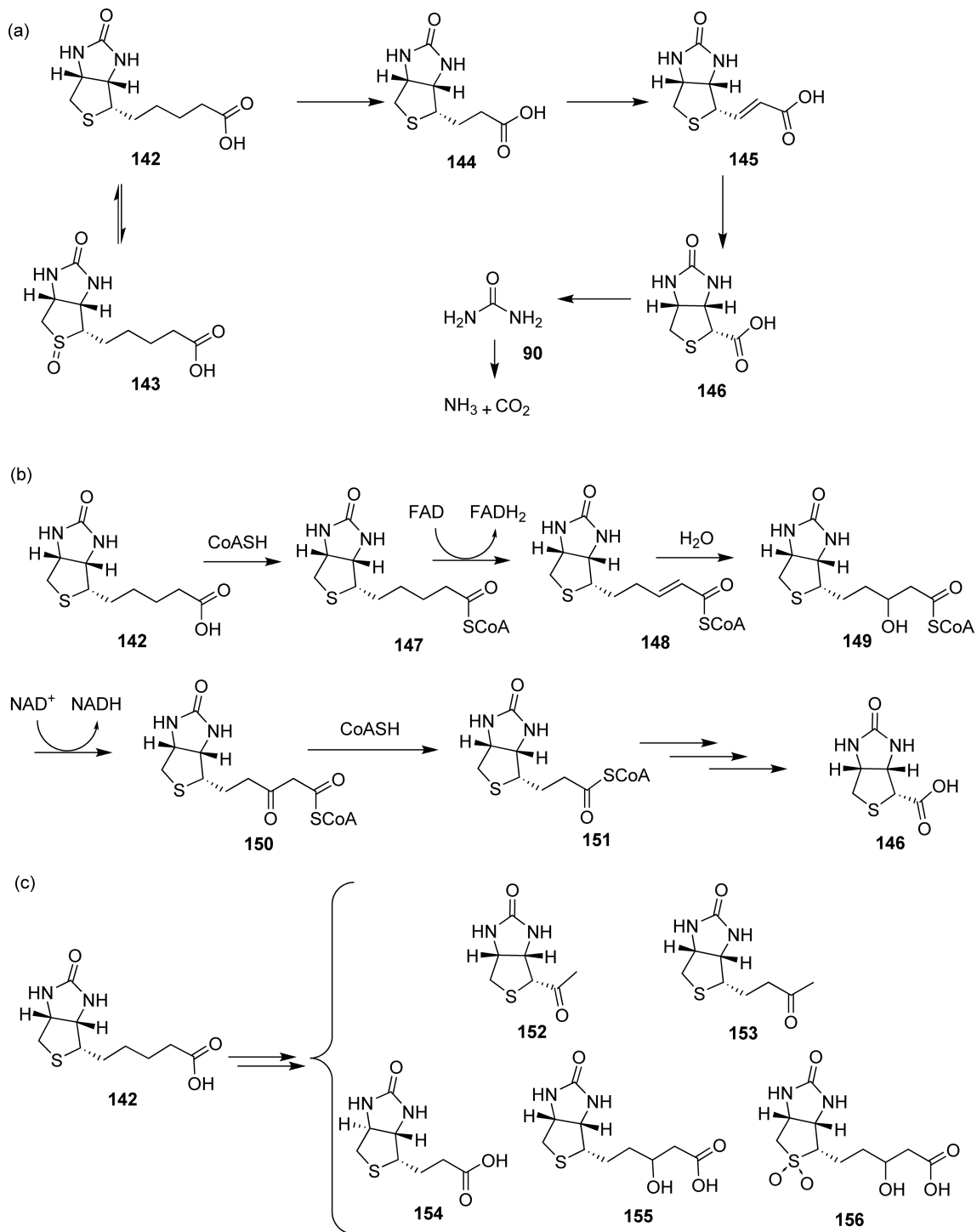
This research was supported by a grant from the National Institute of Health (GM069618).



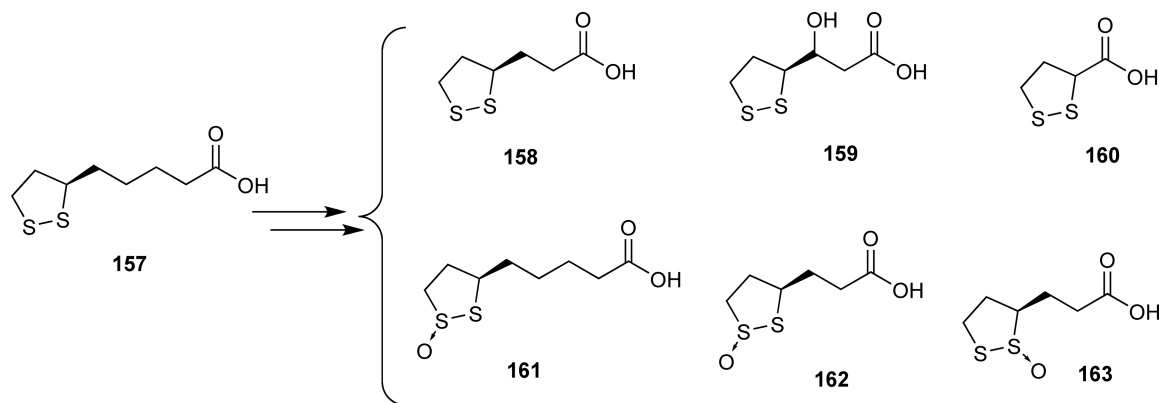
**Figure 14** (a) Crystal structure of TenA showing the 4-amino-2-methyl-5-hydroxymethylpyrimidine (**126**) bound at the active site (2.5 Å, PDB: 1yak).<sup>138</sup> (b) The active site of TenA. (c) Proposed mechanism of the reaction catalyzed by TenA.<sup>113</sup>



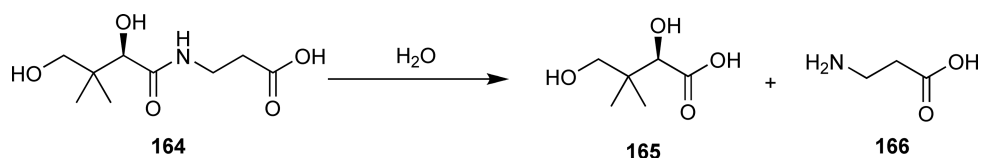
**Figure 15** Catabolic pathway of tetrahydrofolate (135) (vitamin B<sub>9</sub>).



**Figure 16** (a) Proposed biotin (**142**) catabolic pathway in pseudomonads; (b) proposed mechanism of formation of **146**; and (c) various other catabolic intermediates that have been isolated.



**Figure 17** Intermediates of liponic acid (157) catabolism in *pseudomonads*.



**Figure 18** Reaction catalyzed by pantothenase.

## Glossary

**catabolism** The phenomenon of degradation of a metabolite in the living system. The degradation could lead to molecules, which are excreted or it could lead to intermediates that are channelized into other biosynthetic pathways.

**cofactors** These are small molecules (inorganic as well as organic) that are required by various enzymes for their biological activity. All organic cofactors are vitamins. Vitamin B<sub>1</sub> – thiamin, vitamin B<sub>2</sub> – riboflavin, vitamin B<sub>3</sub> – niacin, vitamin B<sub>5</sub> – pantothenate, vitamin B<sub>6</sub> – pyridoxine, vitamin B<sub>7</sub> – biotin, vitamin B<sub>9</sub> – folate, and so on.

**PDB** PDB is the abbreviation for the Protein Data Bank, which has all structural information of a protein that has been crystallized and characterized. The information can be accessed through the web at <http://www.rcsb.org/pdb/home/home.do>.

**vitamins** These are organic compounds that are not synthesized in humans and are absolutely essential for human survival. Vitamins must be acquired from food that we consume (pills, plant, and animal products). Imbalance of the vitamins causes various disorders in our system and eventually leads to death.

**X-ray crystallography** A technique that uses X-rays to elucidate the three-dimensional structure of a crystal. From the X-ray scattering of a crystal (protein as well as small molecules), the position of the atoms within the crystal is revealed.

## Abbreviations

<b>ADP</b>	Adenosine 5' diphosphate
<b>AMP</b>	adenosine 5'-monophosphate
<b>FAD</b>	flavin adenine dinucleotide
<b>MHPC</b>	2-methyl-3-hydroxypyridine-5-carboxylic acid

<b>MHPCO</b>	2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase
<b>NAD</b>	nicotinamide adenine dinucleotide
<b>NADH</b>	nicotinamide adenine dinucleotide, reduced form
<b>NADP</b>	nicotinamide adenine dinucleotide phosphate
<b>NMN</b>	nicotinamide mononucleotide
<b>PDB</b>	Protein Data Bank
<b>PLP</b>	pyridoxal-5'-phosphate

## References

1. J. W. Huff; W. A. Perlzweig, *Science* **1944**, *100*, 15.
2. R. Schwartz; N. O. Kjeldgaard, *Biochem. J.* **1951**, *48*, 333–337.
3. J. C. Rabinowitz; E. E. Snell, *Proc. Soc. Exp. Biol. Med.* **1949**, *70*, 235–240.
4. E. E. Snell; B. E. Haskell, *The Metabolism of Vitamin B6*. In *Comprehensive Biochemistry*; M. Florkin, E. H. Stotz, Eds.; Elsevier/ North-olland New York 1971; Vol. 21, pp 47–67.
5. S. K. Reddy; M. S. Reynolds; J. M. Price, *J. Biol. Chem.* **1958**, *233*, 691–696.
6. A. Baysal; B. A. Johnson; H. Linkswiler, *J. Nutr.* **1966**, *89*, 19–23.
7. V. W. Rodwell; B. E. Volcani; M. Ikawa; E. E. Snell, *J. Biol. Chem.* **1958**, *233*, 1548–1554.
8. R. W. Burg; V. W. Rodwell; E. E. Snell, *J. Biol. Chem.* **1960**, *235*, 1164–1169.
9. M. Ikawa; V. W. Rodwell; E. E. Snell, *J. Biol. Chem.* **1958**, *233*, 1555–1559.
10. J. E. Ayling; H. C. Dunathan; E. E. Snell, *Biochemistry* **1968**, *7*, 4537–4542.
11. J. E. Ayling; E. E. Snell, *Biochemistry* **1968**, *7*, 1626–1636.
12. J. E. Ayling; E. E. Snell, *Biochemistry* **1968**, *7*, 1616–1625.
13. H. Kolb; R. D. Cole; E. E. Snell, *Biochemistry* **1968**, *7*, 2946–2954.
14. T. K. Sundaram; E. E. Snell, *J. Biol. Chem.* **1969**, *244*, 2577–2584.
15. R. W. Burg; E. E. Snell, *J. Biol. Chem.* **1969**, *244*, 2585–2589.
16. Y. J. Jong; E. E. Snell, *J. Biol. Chem.* **1986**, *261*, 15112–15114.
17. R. W. Burg; V. W. Rodwell; E. Snell, *J. Biol. Chem.* **1960**, *235*, 1164–1169.
18. T. Yagi; G. M. Kishore; E. E. Snell, *J. Biol. Chem.* **1983**, *258*, 9419–9425.
19. Y. C. Lee; M. J. Nelson; E. E. Snell, *J. Biol. Chem.* **1986**, *261*, 15106–15111.
20. E. E. Snell; A. A. Smucker; E. Ringelmann; F. Lynen, *Biochem. Z.* **1964**, *341*, 109–119.
21. G. M. Kishore; E. E. Snell, *J. Biol. Chem.* **1981**, *256*, 4228–4233.
22. G. M. Kishore; E. E. Snell, *J. Biol. Chem.* **1981**, *256*, 4234–4240.
23. L. G. Sparrow; P. P. Ho; T. K. Sundaram; D. Zach; E. J. Nyns; E. E. Snell, *J. Biol. Chem.* **1969**, *244*, 2590–2600.
24. E. J. Nyns; D. Zach; E. E. Snell, *J. Biol. Chem.* **1969**, *244*, 2601–2605.
25. M. S. Huynh; E. E. Snell, *J. Biol. Chem.* **1985**, *260*, 2379–2383.
26. Y. J. Jong; M. J. Nelson; E. E. Snell, *J. Biol. Chem.* **1986**, *261*, 15102–15105.
27. M. J. Nelson; E. E. Snell, *J. Biol. Chem.* **1986**, *261*, 15115–15120.
28. P. Chaiyen; P. Brissette; D. P. Ballou; V. Massey, *Biochemistry* **1997**, *36*, 8060–8070.
29. P. Chaiyen; P. Brissette; D. P. Ballou; V. Massey, *Biochemistry* **1997**, *36*, 13856–13864.
30. P. Chaiyen; J. Sucharitakul; J. Svasti; B. Entsch; V. Massey; D. P. Ballou, *Biochemistry* **2004**, *43*, 3933–3943.
31. P. Chaiyen; P. Brissette; D. P. Ballou; V. Massey, *Biochemistry* **1997**, *36*, 2612–2621.
32. Y. Yoshikane; N. Yokochi; K. Ohnishi; H. Hayashi; T. Yagi, *Biochem. J.* **2006**, *396*, 499–507.
33. B. Yuan; Y. Yoshikane; N. Yokochi; K. Ohnishi; T. Yagi, *FEMS Microbiol. Lett.* **2004**, *234*, 225–230.
34. J. Funami; Y. Yoshikane; H. Kobayashi; N. Yokochi; B. Yuan; K. Iwasaki; K. Ohnishi; T. Yagi, *Biochim. Biophys. Acta* **2005**, *1753*, 234–239.
35. N. Yokochi; S. Nishimura; Y. Yoshikane; K. Ohnishi; T. Yagi, *Arch. Biochem. Biophys.* **2006**, *452*, 1–8.
36. B. Yuan; N. Yokochi; Y. Yoshikane; K. Ohnishi; T. Yagi, *J. Biosci. Bioeng.* **2006**, *102*, 504–510.
37. T. Mukherjee; C. Kinsland; T. P. Begley, *Bioorg. Chem.* **2007**, *35*, 458–464.
38. T. Mukherjee; K. M. McCulloch; S. E. Ealick; T. P. Begley, *Biochemistry* **2007**, *46*, 13606–13615.
39. T. Mukherjee; D. G. Hilmey; T. P. Begley, *Biochemistry* **2008**, *47*, 6233–6241.
40. K. M. McCulloch; T. Mukherjee; S. E. Ealick; T. P. Begley, *Biochemistry* **2009**, *48*, 4139–4149.
41. Y. N. F. Ge; Y. Yoshikane; K. Ohnishi; T. Yagi, *J. Biochem.* **2008**, *143*, 603–609.
42. N. Yokochi; Y. Yoshikane; Y. Trongpanich; K. Ohnishi; T. Yagi, *J. Biol. Chem.* **2004**, *279*, 37377–37384.
43. Y. Yoshikane; N. Yokochi; K. Ohnishi; T. Yagi, *Protein Expr. Purif.* **2004**, *34*, 243–248.
44. Y. Kaneda; K. Ohnishi; T. Yagi, *Biosci. Biotechnol. Biochem.* **2002**, *66*, 1022–1031.
45. Y. Trongpanich; K. Abe; Y. Kaneda; T. Morita; T. Yagi, *Biosci. Biotechnol. Biochem.* **2002**, *66*, 543–548.
46. M. L. Reniere; V. J. Torres; E. P. Skaar, *Biometals* **2007**, *20*, 333–345.
47. M. D. Maines, *Antioxid. Redox Signal.* **2005**, *7*, 1761–1766.
48. N. Frankenberg-Dinkel, *Antioxid. Redox Signal.* **2004**, *6*, 825–834.
49. A. Wilks, *Antioxid. Redox Signal.* **2002**, *4*, 603–614.



50. P. R. Montellano, *Curr. Opin. Chem. Biol.* **2000**, *4*, 221–227.
51. Y. Liu; P. R. Ortiz de Montellano, *J. Biol. Chem.* **2000**, *275*, 5297–5307.
52. M. P. Schmitt, *J. Bacteriol.* **1997**, *179*, 838–845.
53. I. Stojiljkovic; K. Hantke, *Mol. Microbiol.* **1994**, *13*, 719–732.
54. W. Zhu; A. Wilks; I. Stojiljkovic, *J. Bacteriol.* **2000**, *182*, 6783–6790.
55. M. D. Suits; G. P. Pal; K. Nakatsu; A. Matte; M. Cygler; Z. Jia, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16955–16960.
56. E. P. Skaar; A. H. Gaspar; O. Schneewind, *J. Bacteriol.* **2006**, *188*, 1071–1080.
57. R. Wu; E. P. Skaar; R. Zhang; G. Joachimiak; P. Gornicki; O. Schneewind; A. Joachimiak, *J. Biol. Chem.* **2005**, *280*, 2840–2846.
58. E. P. Skaar; A. H. Gaspar; O. Schneewind, *J. Biol. Chem.* **2004**, *279*, 436–443.
59. M. D. Maines; G. M. Trakshel, *Arch. Biochem. Biophys.* **1993**, *300*, 320–326.
60. M. Beaulieu; E. Levesque; D. W. Hum; A. Belanger, *Biochem. Biophys. Res. Commun.* **1998**, *248*, 44–50.
61. N. Frankenberger; K. Mukougawa; T. Kohchi; J. C. Lagarias, *Plant Cell* **2001**, *13*, 965–978.
62. R. Tenhunen; H. Marver; N. R. Pimstone; W. F. Trager; D. Y. Cooper; R. Schmid, *Biochemistry* **1972**, *11*, 1716–1720.
63. R. Tenhunen; H. S. Marver; R. Schmid, *Trans. Assoc. Am. Phys.* **1969**, *82*, 363–371.
64. R. Tenhunen; H. S. Marver; R. Schmid, *J. Biol. Chem.* **1969**, *244*, 6388–6394.
65. J. C. Docherty; B. A. Schacter; G. D. Firneisz; S. B. Brown, *J. Biol. Chem.* **1984**, *259*, 13066–13069.
66. J. C. Docherty; G. D. Firneisz; B. A. Schacter, *Arch. Biochem. Biophys.* **1984**, *235*, 657–664.
67. M. Unno; T. Matsui; M. Ikeda-Saito, *Nat. Prod. Rep.* **2007**, *24*, 553–570.
68. T. Yoshida; C. T. Migita, *J. Inorg. Biochem.* **2000**, *82*, 33–41.
69. G. A. Caignan; R. Deshmukh; A. Wilks; Y. Zeng; H. W. Huang; P. Moenne-Loccoz; R. A. Bunce; M. A. Eastman; M. Rivera, *J. Am. Chem. Soc.* **2002**, *124*, 14879–14892.
70. M. Ratliff; W. Zhu; R. Deshmukh; A. Wilks; I. Stojiljkovic, *J. Bacteriol.* **2001**, *183*, 6394–6403.
71. G. O. Paiva-Silva; C. Cruz-Oliveira; E. S. Nakayasu; C. M. Maya-Monteiro; B. C. Dunkov; H. Masuda; I. C. Almeida; P. L. Oliveira, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8030–8035.
72. E. P. Skaar; O. Schneewind, *Microbes Infect.* **2004**, *6*, 390–397.
73. E. P. Skaar; M. Humayun; T. Bae; K. L. DeBord; O. Schneewind, *Science* **2004**, *305*, 1626–1628.
74. D. N. Frick; M. J. Bessman, *J. Biol. Chem.* **1995**, *270*, 1529–1534.
75. M. Proudfoot; E. Kuznetsova; G. Brown; N. N. Rao; M. Kitagawa; H. Mori; A. Savchenko; A. F. Yakunin, *J. Biol. Chem.* **2004**, *279*, 54687–54694.
76. Y. Takagi; B. L. Horecker, *J. Biol. Chem.* **1957**, *225*, 77–86.
77. G. Orsomando; V. Polzonetti; P. Natalini, *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* **2000**, *126*, 89–98.
78. J. A. Jorge; H. F. Terenzi, *J. Gen. Microbiol.* **1984**, *130*, 1563–1568.
79. C. Yan; D. L. Sloan, *J. Biol. Chem.* **1987**, *262*, 9082–9087.
80. M. Wintzerith; A. Dierich; P. Mandel, *Biochim. Biophys. Acta* **1980**, *613*, 191–202.
81. W. Cheng; J. Roth, *J. Bacteriol.* **1995**, *177*, 6711–6717.
82. D. Hillyard; M. Rechsteiner; P. Manlapaz-Ramos; J. S. Imperial; L. J. Cruz; B. M. Olivera, *J. Biol. Chem.* **1981**, *256*, 8491–8497.
83. T. Imai, *J. Biochem.* **1973**, *73*, 139–153.
84. C. T. Grubmeyer; J. W. Gross; M. Rajavel, *Methods Enzymol.* **1999**, *308*, 28–48.
85. V. N. Gladyshev; S. V. Khangulov; T. C. Stadtman, *Biochemistry* **1996**, *35*, 212–223.
86. J. Gauthier; S. C. Rittenberg, *J. Biol. Chem.* **1971**, *246*, 3737–3742.
87. E. J. Behrman; R. Y. Stanier, *J. Biol. Chem.* **1957**, *228*, 923–945.
88. R. B. Cain; C. Houghton; K. A. Wright, *Biochem. J.* **1974**, *140*, 293–300.
89. K. Hatakeyama; M. Goto; M. Kobayashi; M. Terasawa; H. Yukawa, *Biosci. Biotechnol. Biochem.* **2000**, *64*, 1477–1485.
90. J. I. Jimenez; A. Canales; J. Jimenez-Barbero; K. Ginalski; L. Rychlewski; J. L. Garcia; E. Diaz, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11329–11334.
91. A. Alhapel; D. J. Darley; N. Wagener; E. Eckel; N. Elsner; A. J. Pierik, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12341–12346.
92. J. W. Foster; T. Yanagita, *J. Biol. Chem.* **1956**, *221*, 593–607.
93. J. W. Foster, *J. Bacteriol.* **1944**, *47*, 27–41.
94. J. W. Foster, *J. Bacteriol.* **1944**, *48*, 97–111.
95. D. W. West; E. C. Owen, *Br. J. Nutr.* **1973**, *29*, 33–41.
96. D. R. Harkness; E. R. Stadtman, *J. Biol. Chem.* **1965**, *240*, 4089–4096.
97. D. R. Harkness; L. Tsai; E. R. Stadtman, *Arch. Biochem. Biophys.* **1964**, *108*, 323–333.
98. M. E. Taga; N. A. Larsen; A. R. Howard-Jones; C. T. Walsh; G. C. Walker, *Nature* **2007**, *446*, 449–453.
99. L. A. Maggio-Hall; P. C. Dorrestein; J. C. Escalante-Semerena; T. P. Begley, *Org. Lett.* **2003**, *5*, 2211–2213.
100. S. E. Ealick; T. P. Begley, *Nature* **2007**, *446*, 387–388.
101. N. Campobasso; C. A. Costello; C. Kinsland; T. P. Begley; S. E. Ealick, *Biochemistry* **1998**, *37*, 15981–15989.
102. C. A. Costello; N. L. Kelleher; M. Abe; F. W. McLafferty; T. P. Begley, *J. Biol. Chem.* **1996**, *271*, 3445–3452.
103. A. Fujita, *J. Vitaminol. (Kyoto)* **1972**, *18*, 67–72.
104. A. Fujita; K. Ueda, *J. Vitaminol. (Kyoto)* **1958**, *4*, 163–171.
105. A. Fujita, *Adv. Enzymol. Relat. Subj. Biochem.* **1954**, *15*, 389–421.
106. A. Fujita; Y. Nose; K. Kuratani, *J. Vitaminol. (Kyoto)* **1954**, *1*, 1–7.
107. A. Fujita; Y. Nose; S. Kozuka; T. Tashiro; K. Ueda; S. Sakamoto, *J. Biol. Chem.* **1952**, *196*, 289–295.
108. A. Fujita; Y. Nose; K. Ueda; E. Hasegawa, *J. Biol. Chem.* **1952**, *196*, 297–303.
109. A. Fujita; T. Tashiro, *J. Biol. Chem.* **1952**, *196*, 305–311.
110. A. Fujita; Y. Nose; S. Uyeo; J. Koizumi, *J. Biol. Chem.* **1952**, *196*, 313–320.
111. G. D. Maier; D. E. Metzler, *J. Am. Chem. Soc.* **1957**, *79*, 4386–4391.
112. J. M. El Hage Chahine; J. E. Dubois, *J. Am. Chem. Soc.* **1983**, *105*, 2335–2340.
113. A. H. Jenkins; G. Schyns; S. Potot; G. Sun; T. P. Begley, *Nat. Chem. Biol.* **2007**, *3*, 492–497.
114. A. L. Jenkins; Y. Zhang; S. E. Ealick; T. P. Begley, *Bioorg. Chem.* **2008**, *36*, 29–32.

115. C. Gomez-Moreno; D. E. Edmondson, *Arch. Biochem. Biophys.* **1985**, *239*, 46–52.
116. J. R. Suh; A. K. Herbig; P. J. Stover, *Annu. Rev. Nutr.* **2001**, *21*, 255–282.
117. J. R. Suh; E. W. Oppenheim; S. Girgis; P. J. Stover, *J. Biol. Chem.* **2000**, *275*, 35646–35655.
118. B. Wolf, *J. Nutr. Biochem.* **2005**, *16*, 441–445.
119. W. B. Im; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1973**, *248*, 7798–7805.
120. M. N. Kazarinoff; W. B. Im; J. A. Roth; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1972**, *247*, 75–83.
121. J. A. Roth; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1970**, *245*, 6264–6268.
122. W. B. Im; J. A. Roth; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1970**, *245*, 6269–6273.
123. S. Iwahara; D. B. McCormick; L. D. Wright; H. C. Li, *J. Biol. Chem.* **1969**, *244*, 1393–1398.
124. H. Ruis; R. N. Brady; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1968**, *243*, 547–551.
125. R. N. Brady; H. Ruis; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1966**, *241*, 4717–4721.
126. R. N. Brady; L. F. Li; D. B. McCormick; L. D. Wright, *Biochem. Biophys. Res. Commun.* **1965**, *19*, 777–782.
127. H. C. Li; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1968**, *243*, 4391–4395.
128. H. C. Li; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1968**, *243*, 6442–6445.
129. J. P. Tepper; D. B. McCormick; L. D. Wright, *J. Biol. Chem.* **1966**, *241*, 5734–5735.
130. H. M. Lee; L. D. Wright; D. B. McCormick, *Proc. Soc. Exp. Biol. Med.* **1973**, *142*, 439–442.
131. H. M. Lee; N. E. McCall; L. D. Wright; D. B. McCormick, *Proc. Soc. Exp. Biol. Med.* **1973**, *142*, 642–644.
132. J. Zemleni; D. B. McCormick; D. M. Mock, *Am. J. Clin. Nutr.* **1997**, *65*, 508–511.
133. H. C. Furr; H. H. Chang; D. B. McCormick, *Arch. Biochem. Biophys.* **1978**, *185*, 576–583.
134. H. H. Chang; M. L. Roza; D. B. McCormick, *Arch. Biochem. Biophys.* **1975**, *169*, 244–251.
135. J. T. Spence; D. B. McCormick, *Arch. Biochem. Biophys.* **1976**, *174*, 13–19.
136. E. H. Harrison; D. B. McCormick, *Arch. Biochem. Biophys.* **1974**, *160*, 514–522.
137. M. Unno; T. Matsui; G. C. Chu; M. Couture; T. Yoshida; D. L. Rousseau; J. S. Olson; M. Ikeda-Saito, *J. Biol. Chem.* **2004**, *279*, 21055–21061.
138. A. V. Toms; A. L. Haas; J. H. Park; T. P. Begley; S. E. Ealick, *Biochemistry* **2005**, *44*, 2319–2329.

### Biographical Sketches



Tathagata Mukherjee received his Bachelor's in chemistry from St. Stephen's College, Delhi University, India in 2002 and then went on to complete his Master's degree in chemistry from the Indian Institute of Technology (IIT), Delhi, India in 2004. He enrolled at the Cornell University, where he completed his Ph.D. studies in 2009 under Professor Tadhg Begley focusing on the mechanistic enzymology and assignment of gene function. Shortly he will be joining Dr. Clifton Barry III at the National Institute of Health (NIH) for postdoctoral research in the field of tuberculosis.



Kathryn M. McCulloch received her Bachelor's degree in chemistry from Ball State University, Muncie, IN, in 2004. She is now at Cornell University pursuing her doctoral degree with Professor Steven Ealick and studying enzyme structure and function through X-ray crystallography.



Steven E. Ealick is the William T. Miller Professor of Chemistry and Chemical Biology at Cornell University. He received his BS degree in chemistry from Oklahoma State University in 1972 and his Ph.D. degree in physical chemistry from the University of Oklahoma in 1976. He is also Director of the Northeast Collaborative Access Team at the Advanced Photon Source, Argonne National Laboratory. His research focuses on protein structure and function using X-ray crystallography, including enzyme mechanism, structure-based drug design, and protein evolution. He has published widely in the areas of thiamin biosynthesis, purine biosynthesis, polyamine biosynthesis, and cytokine structure and function. He is also interested in the development of instrumentation for synchrotron beamlines used for protein crystallography.



Tadhg P. Begley obtained his B.Sc. from The National University of Ireland in Cork in 1977 and his Ph.D. from the California Institute of Technology (P. Dervan) in 1982. He carried out postdoctoral studies at the University of Geneva (W. Oppolzer) and at MIT (C. Walsh). After 23 years in the Cornell Chemistry Department, he recently moved to Texas A&M University where he is the Derek H. R. Barton Professor of Chemistry. Begley's research is focused on the mechanistic enzymology of complex organic transformations, particularly those found on the vitamin biosynthetic pathways.

## 7.19 Protein-Derived Cofactors

Victor L. Davidson, University of Mississippi Medical Center, Jackson, MS, USA

© 2010 Elsevier Ltd. All rights reserved.

---

<b>7.19.1</b>	<b>Introduction</b>	675
<b>7.19.2</b>	<b>Pyruvoyl Cofactor – A Product of Peptide Cleavage</b>	677
7.19.2.1	Pyruvoyl Cofactor Biosynthesis	677
7.19.2.2	Pyruvoyl Cofactor-Dependent Catalysis	678
<b>7.19.3</b>	<b>4-Methylideneimidazole-5-One – A Product of Peptide Cyclization and Dehydration</b>	680
7.19.3.1	MIO Biosynthesis	680
7.19.3.2	MIO-Dependent Catalysis	681
<b>7.19.4</b>	<b>Protein-Derived Quinone Cofactors</b>	682
7.19.4.1	Topaquinone	682
7.19.4.1.1	TPQ biosynthesis	683
7.19.4.1.2	TPQ-dependent catalysis	685
7.19.4.2	Lysine Tyrosylquinone	686
7.19.4.2.1	LTQ biosynthesis	686
7.19.4.2.2	LTQ-dependent catalysis	686
7.19.4.3	Tryptophan Tryptophylquinone	688
7.19.4.3.1	TTQ biosynthesis	688
7.19.4.3.2	TTQ-dependent catalysis	689
7.19.4.4	Cysteine Tryptophylquinone	690
7.19.4.4.1	CTQ biosynthesis	693
7.19.4.4.2	CTQ-dependent catalysis	693
<b>7.19.5</b>	<b>Functional Covalently Cross-Linked Amino Acid Residues in Metalloproteins</b>	695
7.19.5.1	Cysteine-Tyrosine, the Galactose Oxidase Cofactor	695
7.19.5.1.1	Galactose oxidase cofactor biosynthesis	695
7.19.5.1.2	Galactose oxidase cofactor-dependent catalysis	696
7.19.5.2	Cross-Linked Amino Acid Residues in Heme Enzymes	697
7.19.5.2.1	KatG cofactor biosynthesis	697
7.19.5.2.2	KatG-dependent catalysis	698
7.19.5.3	Cross-Linked Amino Acid Residues in Complex Copper Enzymes	698
<b>7.19.6</b>	<b>Green Fluorescent Protein and Protein-Derived Fluorophores</b>	701
7.19.6.1	GFP-Fluorophore Biosynthesis	701
7.19.6.2	Related Fluorescent Protein Fluorophores	701
<b>7.19.7</b>	<b>Pyrrroloquinoline Quinone, a Protein-Derived Exogenous Cofactor</b>	703
7.19.7.1	PQQ Biosynthesis	703
7.19.7.2	PQQ-Dependent Catalysis	704
<b>7.19.8</b>	<b>Lantibiotics, Protein-Derived Antibiotic Peptides</b>	706
<b>7.19.9</b>	<b>Conclusions and Perspectives</b>	707
<b>References</b>		708

---

### 7.19.1 Introduction

Many enzymes are initially synthesized as inactive forms, which require exogenous cofactors for their activity. Only after the inactive apoprotein combines with the cofactor it becomes the active holoenzyme. Cofactors may be dissociable or tightly bound with the latter often referred to as a prosthetic group. A cofactor may be a metal (e.g., iron or copper), an organic compound (e.g., pyridoxal phosphate or flavin), or an organometallic compound (e.g., heme or cobalamin). Cofactors are necessary to assist amino acid residues at the active sites

of enzymes in certain types of catalysis. Of the 20 most common amino acids occurring in proteins, relatively few are chemically well suited to function as catalysts. Most amino acid side chains are essentially inert. Those that are most commonly seen functioning in enzyme active sites, such as histidine, serine, cysteine, and carboxylic residues, are limited in function to acid–base chemistry and electron donation during nucleophilic catalysis. To circumvent the limited chemical versatility of amino acids many enzymes must utilize cofactors. Most organic and organometallic cofactors are natural products, which in humans are typically derived from vitamins.

It has recently been demonstrated that certain enzymes have evolved an alternative method to introduce new catalytic functional groups at their active sites that circumvents the need for exogenous cofactors. These enzymes utilize protein-derived cofactors, which are catalytic or redox-active centers that form by post-translational modification of one or more amino acid residues. Protein-derived cofactors add a new dimension to our view of the field of posttranslational modification of proteins.<sup>1,2</sup> The amino acid sequence of a protein is determined from the sequence of the gene from which it is encoded. It has long been recognized that after the completion of translation on the ribosome many newly synthesized proteins require further posttranslational modifications to enable them to perform their biological functions. Typical posttranslational modifications include phosphorylation, methylation, glycosylation, prenylation, and proteolytic cleavage to remove a signal or localization pre-sequence.<sup>3</sup> These well-characterized posttranslational modifications primarily serve to regulate the biological activity of the protein, usually an enzyme, or to target the protein to its site of action in the cell. However, the modified amino acid residues do not typically acquire any new functional properties because of these posttranslational modifications. This is a distinction from the posttranslational modifications which generate protein-derived cofactors.

Protein-derived cofactors are not typical natural products in the sense that they are not synthesized *de novo* from small metabolites. Rather their synthesis is on a protein template. Protein-derived cofactors have for the most part gone unrecognized until recent years because for most there is no predicative sequence motif from which to identify the site of modification. The biosynthesis of these protein-derived cofactors typically involves oxygenation of aromatic residues, covalent cross-linking of amino acid residues, or cyclization or cleavage of internal amino acid residues. The resulting cofactors typically function at the enzyme active site either by providing an electrophilic site to interact with substrates or by stabilizing free radical intermediates, functions which the unmodified amino acid residues are incapable of performing.

The presence of protein-derived cofactors has typically been revealed when high-resolution X-ray crystal structures of the host enzymes became available and advanced mass spectrometry technology allowed identification and confirmation of the modified peptides and amino acid residues. It is probable that additional protein-derived cofactors can be revealed as more structures of enzymes become available. The current status of this rapidly developing area of biochemical research is summarized in this chapter. The first two protein-derived cofactors to be discussed require not only modification of amino acid side chains, but also modification of the protein backbone during biosynthesis. The pyruvoyl cofactor requires internal cleavage of the gene-encoded polypeptide chain. Methylidene imidazolone (MIO) requires internal cyclization of the peptide chain with subsequent chemical modifications. In each case, the modifications introduce an electrophilic site for catalysis at the enzyme active site. An alternative strategy for introduction of an electrophilic site for catalysis is exhibited by covalent quinoproteins, which possess quinone functional groups derived from tyrosine and tryptophan residues. Examples of functional cross-linked amino acid residues at the active sites of a variety of metalloproteins are described. In these examples, radical mechanisms are described for both the biosynthesis of the protein-derived cofactor and catalysis by the mature cofactor. Although green fluorescent protein (GFP) is not an enzyme, the biogenesis of its protein-derived fluorophore is described, as its mechanism of biosynthesis is similar to that of the protein-derived cofactors, in particular MIO. As with the protein-derived cofactors, the posttranslational modifications that yield the fluorescent protein fluorophores also endow the gene-encoded polypeptide with a new biological function. Pyrroloquinoline quinone (PQQ) is an endogenous cofactor, which is also discussed in this review because, in contrast to the biosynthesis of most endogenous cofactors, PQQ is not synthesized *de novo* from precursor metabolites but instead from a polypeptide precursor. In this sense, it is also a protein-derived cofactor. Finally, lantibiotics are briefly discussed. Although not enzymes, this class of naturally occurring proteins exhibits a wide array of posttranslational modifications, which endow the protein with its antibiotic properties. It will be clear from the examples that the knowledge of



the sequence of a gene does not necessarily tell you everything that you need to know about the gene product. An amazing array of covalent posttranslational modifications are now known to occur, which yield proteins with unexpected structures and functions that as yet cannot be predicted solely from sequence analysis.

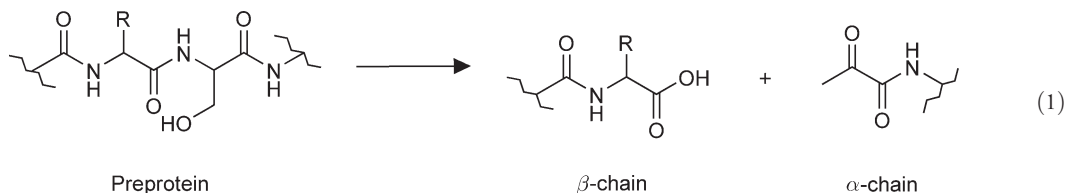
### 7.19.2 Pyruvoyl Cofactor – A Product of Peptide Cleavage

The carbonyl functional group provides an electrophilic site for nucleophilic attack. As none of the naturally occurring amino acids contains a ketone or aldehyde moiety, it is not possible for a protein-bound carbonyl group to be present at an enzyme active site, unless it is provided by a cofactor. The most commonly seen reactive carbonyl group in enzymes is that provided by the pyridoxal phosphate cofactor which is derived from vitamin B<sub>6</sub>.<sup>4</sup> The carbonyl serves as a site for covalent adduct formation with a nucleophilic substrate, often an amine or the amino group of amino acid substrates. An alternative strategy for introduction of a reactive carbonyl at an enzyme active site has been described for a class of enzymes that possess a covalent pyruvoyl cofactor.

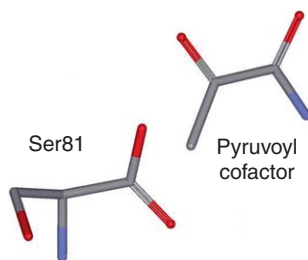
Pyruvoyl cofactor<sup>5</sup> is derived from the posttranslational modification of an internal amino acid residue, and it does not equilibrate with exogenous pyruvate. Enzymes that possess this cofactor play an important role in the metabolism of biologically important amines from bacterial and eukaryotic sources. These enzymes include aspartate decarboxylase,<sup>6</sup> arginine decarboxylase,<sup>7</sup> phosphatidylserine decarboxylase,<sup>8</sup> *S*-adenosylmethionine decarboxylase,<sup>9</sup> histidine decarboxylase,<sup>10</sup> glycine reductase,<sup>11</sup> and proline reductase.<sup>12</sup>

#### 7.19.2.1 Pyruvoyl Cofactor Biosynthesis

A common feature of this class of enzymes is that the catalytic pyruvoyl residue is covalently bound to the N-terminus of the  $\alpha$ -chain, one of two nonidentical subunits that comprise the holoenzyme. The two-subunit active enzyme is initially synthesized as an inactive single subunit pro-enzyme (Equation (1)). The cleavage of the polypeptide and formation of the pyruvoyl cofactor is a self-processing event, which is promoted by the tertiary structure of the pro-enzyme that orients the reactive moieties into position to facilitate the chemical reactions necessary for the biosynthesis of the protein-derived cofactor. During this posttranslational modification a specific internal residue, serine or cysteine, becomes the pyruvoyl prosthetic group concomitant with the cleavage of the pro-enzyme. **Figure 1** shows the pyruvoyl cofactor of histidine decarboxylase from *Lactobacillus* 30A<sup>10</sup> at the C-terminus of the  $\alpha$ -chain and the newly formed N-terminus of the  $\beta$ -chain that result from the internal polypeptide cleavage. For each of the pyruvoyl-dependent decarboxylases, the site of cleavage of the polypeptide, is at a specific serine residue. The chemical mechanism of pyruvoyl group formation from an internal serine is well-established,<sup>5</sup> and has been confirmed and extended as a result of X-ray crystallographic studies of the mature enzyme and mutant forms of the pro-enzyme, which cannot be processed.



The proposed mechanism for biosynthesis of the pyruvoyl cofactor (**Scheme 1**) occurs through nonhydrolytic serinolysis, in which the side-chain hydroxyl group supplies an oxygen that becomes part of the C-terminus of the  $\beta$ -chain. The remainder of the serine residue is converted into ammonia and the pyruvoyl group that blocks the N-terminus of the  $\alpha$ -chain. The serine oxygen attacks the carbonyl carbon to form a cyclic oxyoxazolidine intermediate (**1**), which then converts into a linear ester intermediate<sup>13</sup> (**2**). The ester undergoes  $\beta$ -elimination creating dehydroalanine (**3**) at the N-terminus of the site of cleavage with release of what becomes the  $\beta$ -chain of the mature dimeric enzyme. Tautomerization of dehydroalanine produces an



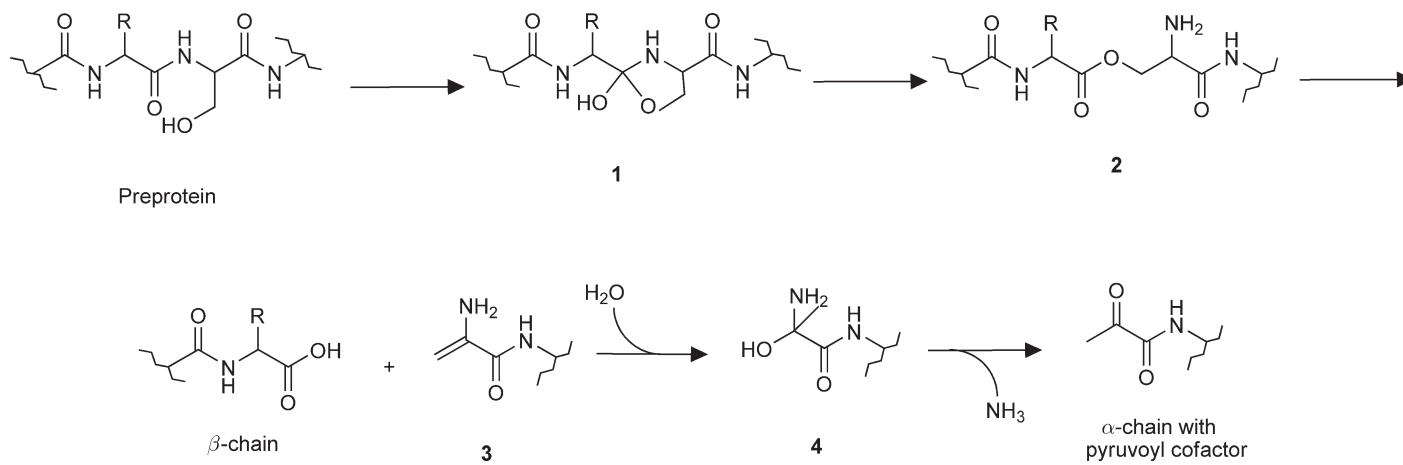
**Figure 1** Structure of the pyruvoyl cofactor in histidine decarboxylase from *Lactobacillus* 30A. In this enzyme, modification of Ser82 and cleavage of the pro-enzyme between Ser81 and Ser82 yields the pyruvoyl cofactor at the C-terminus of the  $\alpha$ -chain, and Ser81 at the N-terminus of the  $\beta$ -chain of the holoprotein. These residues are displayed as sticks colored gray for carbon, red for oxygen, and blue for nitrogen. The coordinates from PDB entry 1jen were used to display this structure.

imine, which converts first into a carbinolamine (4), and then deaminates to form the covalently attached pyruvoyl group, which is present at the end of the  $\alpha$ -chain in the mature enzyme. The importance of the serine residue in this process was demonstrated by site-directed mutagenesis of this serine residue in the *S*-adenosylmethionine decarboxylase pro-enzyme. Changing the serine residue to an alanine completely, prevented the processing and the formation of active enzyme.<sup>14</sup> As a consequence of this serine to alanine mutation which prevented cleavage and pyruvoyl cofactor formation, it was also possible to determine the crystal structure of the unprocessed pro-enzyme.<sup>7</sup> This allowed identification of amino acid residues, which stabilize the cyclic oxyoxazolidine intermediate (1) by hydrogen bond interactions and revealed that a histidine residue provides a proton to the nitrogen atom of the anionic oxyoxazolidine ring during conversion into the ester intermediate (3).

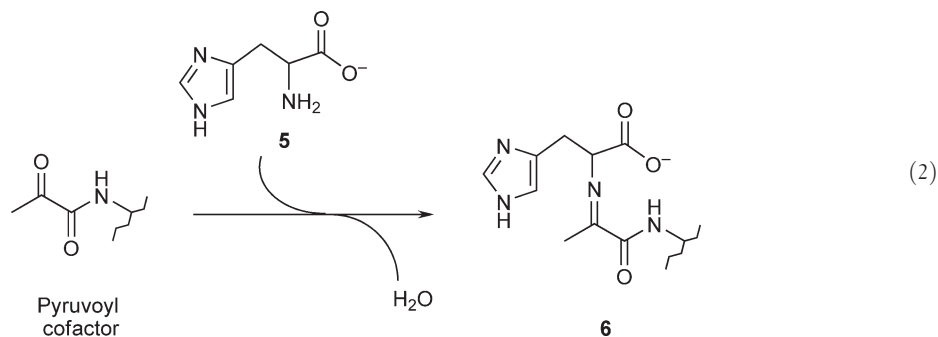
Two enzymes that contain the pyruvoyl cofactor are not decarboxylases; D-proline reductase and glycine reductase. These enzymes were originally reported to contain the pyruvate in an ester linkage, but later studies have demonstrated its presence at the N-terminus of one of the subunits linked by the peptide amide bond. In contrast to the pyruvoyl-dependent decarboxylases, the site of internal cleavage and modification of these reductases is a cysteine rather than a serine.<sup>11,12</sup> The mechanism of post-translational biosynthesis of the pyruvoyl cofactor in these enzymes could conceivably proceed through the same mechanism shown in **Scheme 1** but with the cysteine sulfur performing the role of the serine oxygen.

### 7.19.2.2 Pyruvoyl Cofactor-Dependent Catalysis

The mechanism of action of pyruvoyl cofactor-dependent decarboxylases is believed to involve initial formation of a Schiff base, a covalent adduct between the carbonyl carbon of the pyruvoyl cofactor and the  $\alpha$ -amino group of the substrate. The proposed initial Schiff base-adduct formation for the reaction cycle of histidine decarboxylase is shown in Equation (2). The amino group of the histidine substrate is probably deprotonated by an active-site base to generate the neutral amine (5), which initiates a nucleophilic attack on the carbonyl carbon of the pyruvoyl cofactor displacing the oxygen to form the initial histidine-pyruvoyl Schiff base adduct (6). The decarboxylation reaction then proceeds through a mechanism<sup>5</sup> similar to that catalyzed by pyridoxal phosphate-dependent decarboxylases<sup>4</sup> but utilizing the pyruvoyl cofactor as an electron sink to stabilize reaction intermediates. Loss of the carboxyl group as CO<sub>2</sub> followed by hydrolysis of the Schiff base yields the original form of the cofactor with release of the histamine product. The reactions of the other decarboxylases whose substrates are also amino acids or amino acid derivatives will probably proceed through a similar mechanism. The reaction mechanisms of the glycine and proline reductases, which possess the pyruvoyl cofactor are less clear. A significant distinction between these enzymes and the pyruvoyl cofactor-dependent decarboxylases is that the reductases also possess selenocysteine,<sup>15,16</sup> which also participates in catalysis.



**Scheme 1** Mechanism of pyruvoyl cofactor biosynthesis.



### 7.19.3 4-Methylideneimidazole-5-One – A Product of Peptide Cyclization and Dehydration

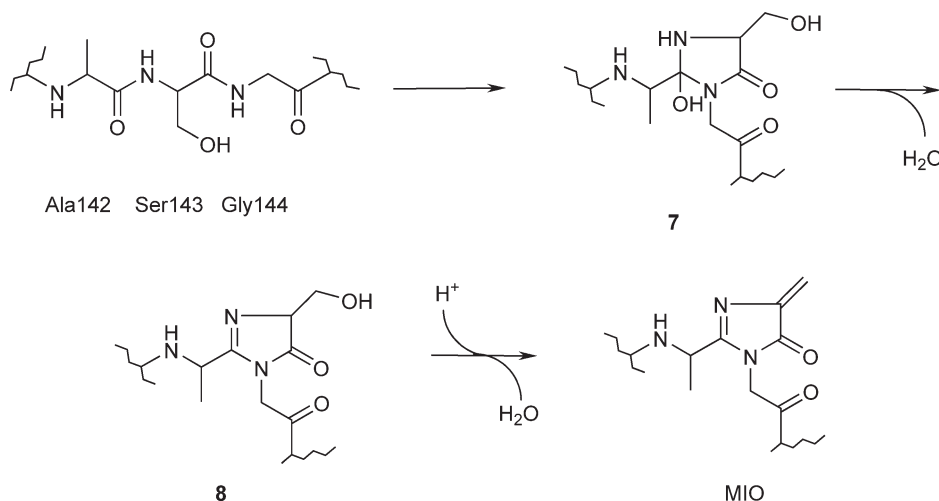
Another strategy for introduction of a reactive electrophile at an enzyme active site is the biosynthesis of the MIO cofactor. MIO has been shown to be the prosthetic group of two enzymes, phenylalanine ammonia-lyase and histidine ammonia-lyase. This cofactor was previously thought to be dehydroalanine. Dehydroalanine is found in food proteins that have been heated or treated with alkali, which can dehydrate serine residues.<sup>17</sup> Naturally occurring dehydroalanine-containing peptides are found in microorganisms and are usually toxic or antibiotic.<sup>18</sup> When the crystal structure of histidine ammonia-lyase from *Pseudomonas putida* was determined it revealed that the covalent cofactor was not dehydroalanine but a modified dehydroalanine, MIO<sup>19</sup> (Figure 2). Although the X-ray structure of phenylalanine ammonia-lyase is not yet solved, MIO could be identified by difference UV spectroscopy,<sup>20</sup> and the presence of MIO in this enzyme was supported by observed sequence homology with histidine ammonia-lyase, and results of mutagenesis studies with phenylalanine ammonia-lyase from *Petroselinum crispum*.<sup>21</sup> MIO has recently been shown to also be the cofactor of another enzyme, the SgcC4 L-tyrosine 2,3-aminomutase, which catalyzes the formation of (*S*)- $\beta$ -tyrosine in the biosynthetic pathway of the enediyne antitumor antibiotic C-1027.<sup>22</sup> A related enzyme phenylalanine aminomutase is responsible for the synthesis of  $\beta$ -amino acids, incorporated into other classes of natural products, and participates in taxol biosynthesis,<sup>23</sup> and the biosynthesis of andrimid, a polyketide nonribosomal peptide.<sup>24</sup>

#### 7.19.3.1 MIO Biosynthesis

On the basis of the results of site-directed mutagenesis and X-ray crystallographic studies, it was proposed that the mechanism of MIO biosynthesis in histidine ammonia-lyase occurs through a self-processing mechanism, with modification of a consecutive sequence of three amino acid residues of the order alanine, serine, and



**Figure 2** Structure of the MIO cofactor in histidine ammonia lyase from *Ps. putida*. In this enzyme modification of Ala142, Ser143, and Gly144 yields the MIO cofactor, which is displayed as sticks colored gray for carbon, red for oxygen, and blue for nitrogen. The coordinates from PDB entry 1b8f were used to display this structure.



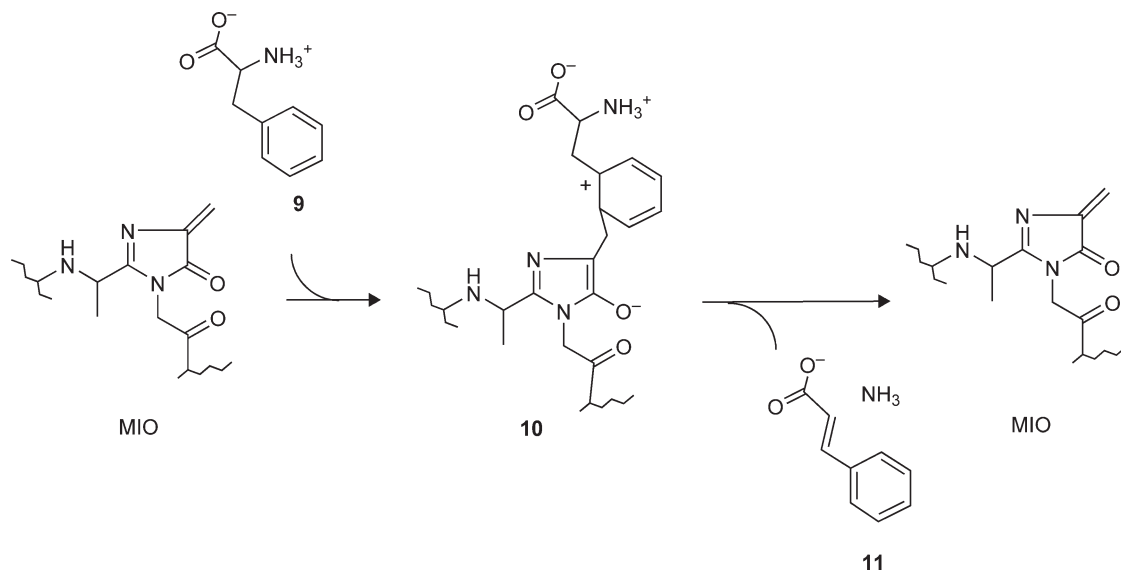
**Scheme 2** Mechanism of MIO biosynthesis.

glycine.<sup>25,26</sup> The proposed mechanism of biosynthesis is shown in **Scheme 2**. The MIO cofactor in histidine ammonia-lyase is formed from the amino acid residues Ala142, Ser143, and Gly144. The initial step in the biosynthetic process is a nucleophilic attack by the amide nitrogen of Gly144 on the carbonyl oxygen of Ala142, which results in an intrachain cyclization to form the first cyclic intermediate (**7**) in the biosynthetic process. This is followed by two consecutive dehydration reactions. In the first, the newly formed OH from the carbon of Ala142 is lost to form the second cyclic intermediate (**8**). Then the OH, which was present originally on the side chain of Ser143, is lost to yield the MIO cofactor. The critical nature of the serine residue in this biosynthetic process was confirmed by site-directed mutagenesis studies, where mutation of the Ser143 in histidine ammonia-lyase,<sup>27</sup> or the corresponding Ser202 in phenylalanine ammonia-lyase<sup>28</sup> prevented the formation of MIO.

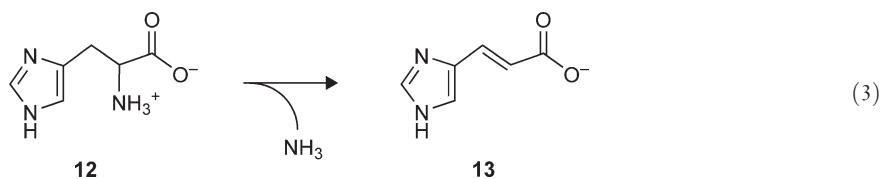
### 7.19.3.2 MIO-Dependent Catalysis

MIO-dependent enzymes catalyze the nonoxidative deamination of their respective substrates. Phenylalanine ammonia-lyase is an important plant enzyme that eliminates ammonia from phenylalanine to form *trans*-cinnamic acid, a precursor of lignins, flavanoids, and coumarins. Histidine ammonia-lyase catalyzes the first step in histidine degradation. Mechanistic studies of these enzymes suggest that MIO forms a covalent adduct, with the side chain of its amino acid substrate during catalysis, which facilitates the elimination of ammonia to yield the product.<sup>29</sup> The dehydroalanine moiety of MIO is particularly electrophilic. In each enzyme it is believed that the side chain of the amino acid substrate, the phenyl ring of phenylalanine or the imidazole ring of histidine, is the target for electrophilic attack to yield the initial adduct. The proposed initial step in the reaction cycle of phenylalanine ammonia-lyase is shown in **Scheme 3**. A reaction between the phenylalanine substrate (**9**) and the dehydroalanine moiety of MIO yields an initial enzyme–substrate adduct (**10**). This adduct formation with MIO increases the acidity of the protons on the  $\beta$ -carbon of phenylalanine facilitating deprotonation, which is then followed by elimination of ammonia and release of the *trans*-cinnamic acid product (**11**).

Histidine ammonia-lyase is the first enzyme in the degradation pathway of  $L$ -histidine and catalyzes the nonoxidative deamination of histidine (**12**) to form *trans*-urocanic acid (**13**) plus ammonia (Equation (3)). Histidine ammonia-lyase is present in several bacteria and in animals. The mechanism for the reaction that is catalyzed by histidine ammonia-lyase is presumed to be similar to that described above for phenylalanine ammonia-lyase<sup>25</sup> (see **Scheme 3**).



**Scheme 3** MIO-dependent conversion of phenylalanine into *trans*-cinnamic acid by phenylalanine ammonia-lyase.



### 7.19.4 Protein-Derived Quinone Cofactors

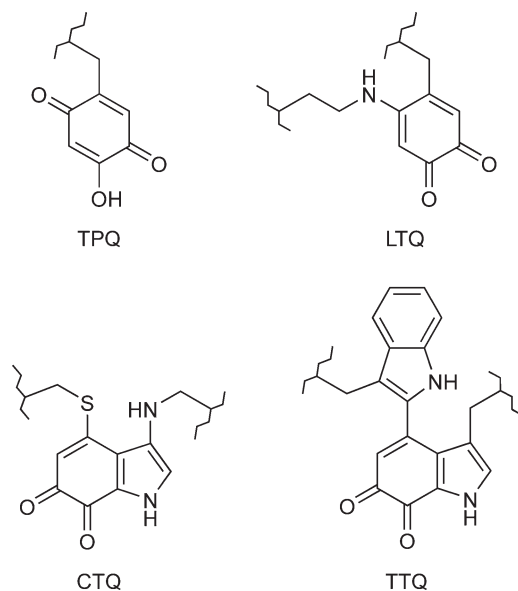
Quinoproteins are a class of enzymes with cofactors possessing a quinone moiety, which participates directly in catalysis. Several dehydrogenases have been characterized which possess the dissociable cofactor 4,5-dihydro-4,5-dioxo-1-H-pyrrolo[2,3-f]quinoline-2,7,9-tricarboxylic acid, which is now commonly called pyrroloquinoline quinone or PQQ.<sup>30</sup> Earlier, several other enzymes were believed to possess covalently attached PQQ as a cofactor. This is now known to be false and an unfortunate consequence of erroneous analytical techniques.<sup>31</sup> Although there are no examples of enzymes possessing covalently bound PQQ, covalent quinoproteins with other protein-derived quinones have been identified.

Covalent quinoproteins possess protein-derived cofactors derived from aromatic amino acid residues. These enzymes contain a posttranslationally modified tyrosine or tryptophan residue into which one or two oxygens has been incorporated (Figure 3). In some cases, the quinolated amino acid residue is also covalently cross-linked to another amino acid residue on the polypeptide. Tyrosine-derived quinone cofactors<sup>32</sup> occur in oxidases from bacterial, mammalian, and plant sources. Tryptophan-derived quinone cofactors<sup>33</sup> have been found thus far in bacterial dehydrogenases.

#### 7.19.4.1 Topaquinone

Topaquinone or 2,4,5-trihydroxyphenylalanine quinone (TPQ)<sup>34</sup> (Figure 3) was the first protein-derived quinone cofactor to be characterized. It is the prosthetic group of the copper-containing amine oxidases. These amine oxidases utilize a wide range of substrates and are involved in a wide range of physiological functions. TPQ-containing amine oxidases can be found throughout nature and have been purified from many sources including plants, mammals, fungi, and bacteria. In microorganisms, these amine oxidases play a nutritional role allowing primary amines to be utilized, as the sole source of carbon and nitrogen for growth. In plants,





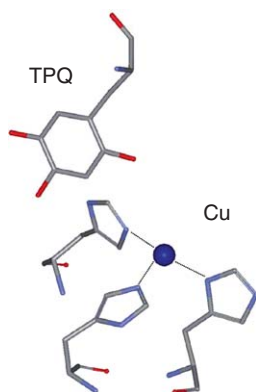
**Figure 3** Quinoprotein cofactors. TPQ, 2,4,5-trihydroxyphenylalanine quinone; LTQ, lysine tyrosylquinone; CTQ, cysteine tryptophylquinone; TTQ, tryptophan tryptophylquinone.

TPQ-containing amine oxidases are thought to play roles in the biosynthesis of hormones, cell walls, and alkaloids.<sup>35</sup> In animals, the physiological roles of these amine oxidases, are not completely understood. Different classes of TPQ-dependent copper amine oxidases have been described from mammalian sources, which include plasma amine oxidase, diamine oxidase, retinal amine oxidase,<sup>36</sup> and semicarbazide-sensitive amine oxidase.<sup>37</sup> This latter class has multiple functions depending on its identity and location, which include glucose homeostasis, lymphocyte adhesion, and adipocyte maturation.<sup>38</sup> In fact, it has been shown that vascular adhesion protein-1 (VAP-1), is also a member of this class of TPQ-containing amine oxidases.<sup>39</sup> VAP-1 is a human endothelial sialoglycoprotein, whose cell surface expression is induced under inflammatory conditions.

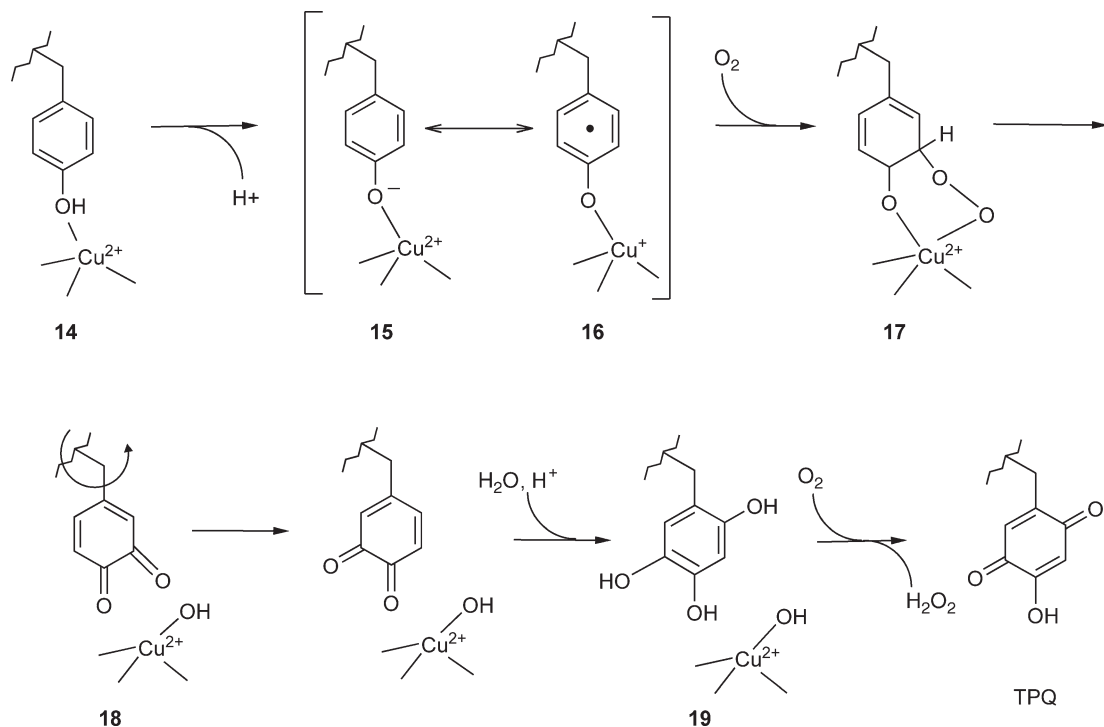
#### 7.19.4.1.1 TPQ biosynthesis

The biogenesis of the TPQ cofactor, which requires the insertion of two oxygens into a specific tyrosine residue, has been shown to be a self-processing event, which requires molecular oxygen as well as copper.<sup>40,41</sup> The tyrosine residue, which is posttranslationally modified, is found within a relatively conserved sequence of Thr-X-X-Asn-Tyr-Asp/Glu, in which the conserved tyrosine residue is modified to form TPQ. This is one of the few examples, thus far of a sequence motif that is predictive for the presence of a protein-derived cofactor. A recombinant amine oxidase has been expressed in copper-depleted *E. coli* cells and isolated as a copper-free apo-form in which the tyrosine residue has not been modified. Generation of the active amine oxidase with correctly formed TPQ was achieved on addition of copper and exposure to oxygen. **Figure 4** shows the mature TPQ cofactor and nearby copper in the crystal structure of the amine oxidase from *Arthrobacter globiformis*.<sup>42</sup>

The mechanism proposed for TPQ biogenesis is shown in **Scheme 4**.<sup>42–45</sup> A copper-binding site is located in close proximity to the tyrosine, which is modified on the polypeptide (**14**). Deprotonation of the tyrosine by a nearby basic amino acid residue yields a charge transfer complex between the tyrosinate and the bound  $\text{Cu}^{2+}$  (**15**), which is in resonance stabilization with a  $\text{Cu}^+$ -tyrosyl radical intermediate (**16**). Electron transfer from either the tyrosinate- $\text{Cu}^{2+}$  or  $\text{Cu}^+$ -tyrosyl radical to molecular oxygen yields superoxide, which then reacts with the tyrosine at the C5 position to generate a copper-coordinated peroxide intermediate (**17**). Fragmentation of the oxygen–oxygen bond in this intermediate results in the formation of  $\text{Cu}^{2+}$ -coordinated hydroxide and dopaquinone (4,5-dihydroxyphenylalanine quinone) (**18**). At this point in the biosynthetic process, it is proposed that the side chain of the dopaquinone intermediate rotates  $180^\circ$  around its  $\text{C}_\beta$ – $\text{C}_\gamma$  bond. Evidence for the rotation of the modified tyrosine during cofactor biogenesis is supported by crystal structures of the unmodified apoprotein, and the zinc-substituted protein in which the posttranslational modification has

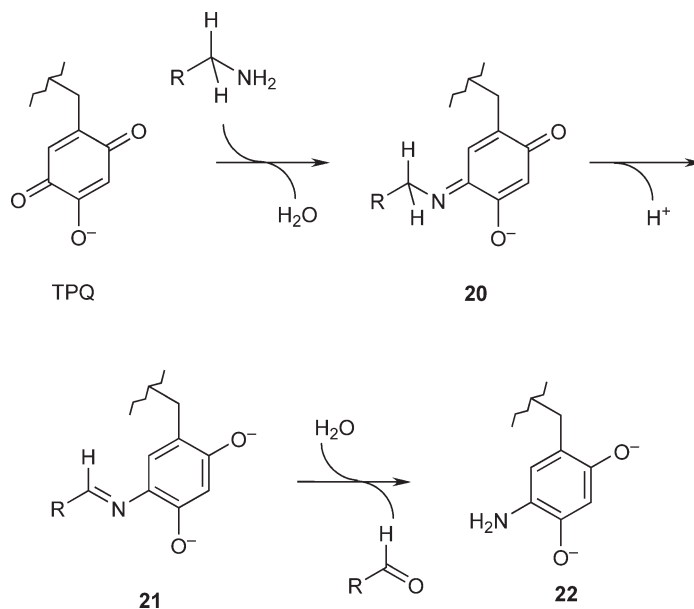


**Figure 4** Structure of the TPQ cofactor in *A. globiformis* amine oxidase. In this enzyme, modification of Tyr382 yields the TPQ cofactor. TPQ and the three histidine residues that coordinate the copper in the active site are displayed as sticks colored gray for carbon, red for oxygen, and blue for nitrogen. The copper is displayed as a dark blue sphere. The coordinates from PDB entry 1ivx were used to display this structure.



**Scheme 4** Mechanism of TPQ biosynthesis.

not occurred.<sup>42,46</sup> The rotation of **18** is followed by addition of hydroxide at the C2 position, which is followed by aromatization to yield 2,4,5-trihydroxyphenylalanine (**19**). This is later oxidized by a second molecule of  $O_2$  to form TPQ and  $H_2O_2$ . Although there is general agreement on this mechanism of TPQ biosynthesis it should be noted that Ni(II) and Co(II) have been shown *in vitro* to substitute for copper and support TPQ biosynthesis in *Hansenula polymorpha*<sup>47</sup> and *A. globiformis*<sup>48</sup> amine oxidases. This would seem to be more consistent with an electrostatic rather than a redox-active function for copper. This may indicate that of the resonance stabilized pair of intermediates **15** and **16**; intermediate **15** is more relevant to the biosynthesis mechanism.

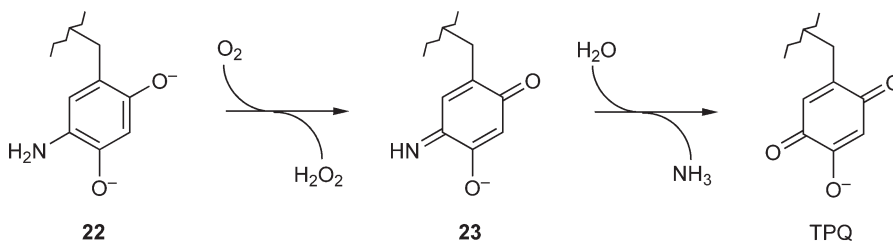


**Scheme 5** The reductive half-reaction of TPQ-dependent amine oxidases.

#### 7.19.4.1.2 TPQ-dependent catalysis

The overall reaction catalyzed by the TPQ-dependent copper amine oxidases is the conversion of a primary amine substrate into its corresponding aldehyde plus ammonia with molecular oxygen and water as co-substrates and  $\text{H}_2\text{O}_2$  as an additional product. The catalytic mechanism consists of two-half reactions.<sup>49</sup> In the initial step in the reductive half-reaction (**Scheme 5**), the nitrogen of the substrate initiates a nucleophilic attack on the quinone carbon at the C5 position of the TPQ cofactor displacing the oxygen to form a substrate-TPQ Schiff base adduct (**20**). Deprotonation of the substrate-derived C1 carbon of the covalently bound substrate in **20** results in reduction of the cofactor and yields an intermediate in which the Schiff base is now between the nitrogen and substrate-derived C1 carbon (**21**). This deprotonation is performed by an aspartic acid residue, which is strictly conserved among the amine oxidases.<sup>50</sup> Hydrolysis of **21** releases the aldehyde product and yields the aminoquinol form of the cofactor with the substrate-derived amino group still covalently bound (**22**). In the oxidative half-reaction (**Scheme 6**) the aminoquinol is re-oxidized by  $\text{O}_2$  forming  $\text{H}_2\text{O}_2$  and the oxidized iminoquinone cofactor (**23**), which still retains the substrate-derived amino group. The ammonia product is released by hydrolysis of **23** to regenerate TPQ.

An interesting aspect of the elucidation of the structure and mechanisms of TPQ biogenesis and function is that for these amine oxidases certain amino acid residues have multiple roles during cofactor biogenesis and catalysis. This is clearly true for the tyrosine residue, which is converted into TPQ. It has also been demonstrated by site-directed mutagenesis that a second strictly conserved tyrosine residue, which is present in the active site is also required for TPQ biogenesis and influences catalysis.<sup>51</sup> Furthermore, the bound copper serves dual functions in cofactor biosynthesis, as shown in **Scheme 4** as well as during catalysis. The latter function appears to be in stabilization of transient intermediates in the oxidative half reaction.<sup>52</sup>



**Scheme 6** The oxidative half-reaction of TPQ-dependent amine oxidases.

### 7.19.4.2 Lysine Tyrosylquinone

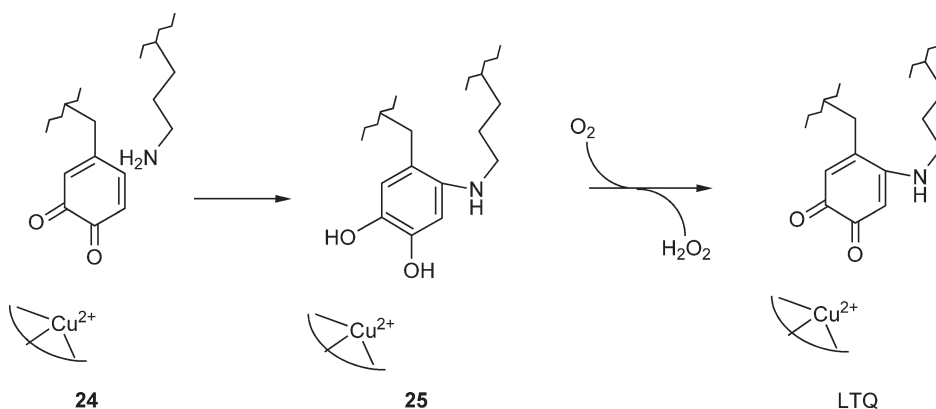
Lysine tyrosylquinone (LTQ) (**Figure 3**) is the protein-derived cofactor of mammalian lysyl oxidase, an important enzyme in the metabolism of connective tissue. Lysyl oxidase catalyzes the posttranslational modification of elastin and collagen.<sup>53</sup> It oxidizes selected peptidyl lysine residues to peptidyl  $\alpha$ -aminoadipic  $\delta$ -semialdehyde residues. This initiates formation of the covalent cross-linkages that insolubilize these extracellular proteins. This enzyme also contains copper as a second prosthetic group.

#### 7.19.4.2.1 LTQ biosynthesis

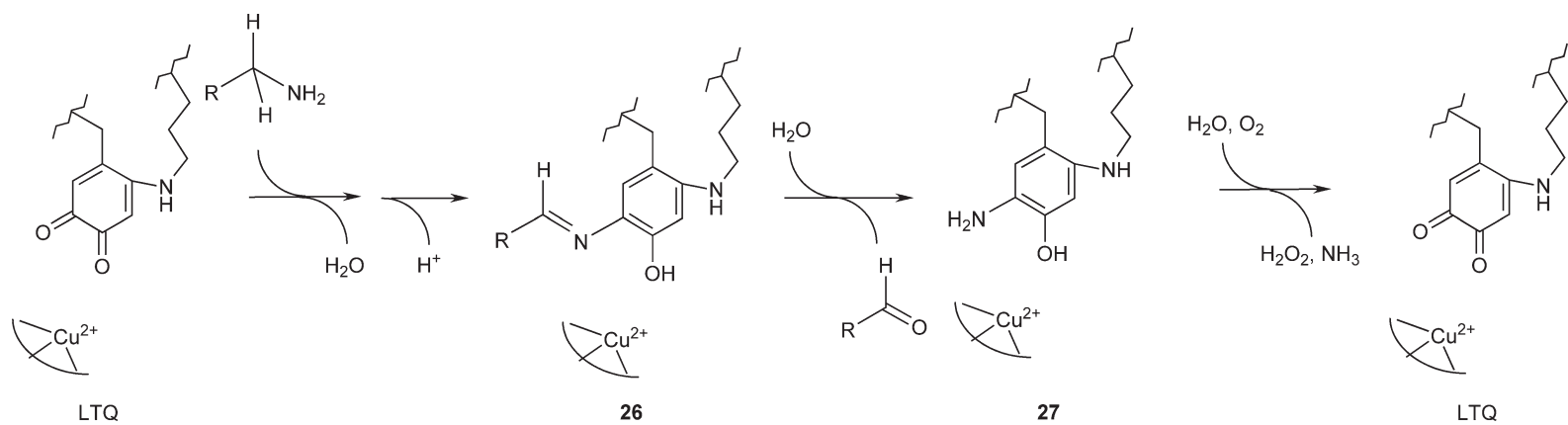
The LTQ cofactor is formed by posttranslational modification in which one atom of oxygen is incorporated into the tyrosine ring, and a covalent bond is formed between the C5 carbon of the tyrosine ring and the side-chain nitrogen of a lysine residue.<sup>54</sup> This bond is formed at the same position at which the last oxygen atom is inserted into the mature TPQ cofactor in the amine oxidases (see **Scheme 4**). The mechanism of LTQ biogenesis has not yet been characterized. LTQ enzymes do not exhibit the conserved sequence motif that is characteristic of TPQ enzymes. However, given the presence of copper at the active site of both TPQ and LTQ enzymes it seems reasonable to assume that the initial steps in the mechanism of biosynthesis of LTQ is similar to that described earlier for TPQ in **Scheme 4**, but with the lysine nitrogen adding to the orthoquinone intermediate (**Scheme 7**). The modified tyrosine in lysyl oxidase is probably converted into a dopaquinone intermediate (**24**) as was seen during TPQ biosynthesis. This intermediate is then the target of nucleophilic attack at the C2 position by the amino nitrogen of the side chain of a nearby lysine residue followed by aromatization to yield a reduced lysylquinol intermediate (**25**). This is then oxidized by a second molecule of O<sub>2</sub> to form LTQ and H<sub>2</sub>O<sub>2</sub>.

#### 7.19.4.2.2 LTQ-dependent catalysis

The catalytic mechanism of lysyl oxidase also consists of reductive and oxidative half-reactions. The proposed mechanism for the overall reaction catalyzed by lysyl oxidase, which was originally proposed by Williamson and Kagan,<sup>55</sup> and is adapted to use the now known LTQ cofactor as shown in **Scheme 8**. In the reductive half-reaction, the nitrogen of the peptidyl lysine substrate initiates a nucleophilic attack on the quinone carbon, at the C5 position of the LTQ cofactor, displacing the oxygen to form a substrate-LTQ Schiff base adduct, which is then deprotonated at the substrate-derived C1 carbon with concomitant reduction of the cofactor to yield an intermediate in which the Schiff base is now between the N and substrate-derived C1 carbon (**26**). Hydrolysis of **26** releases the peptidyl  $\alpha$ -aminoadipic  $\delta$ -semialdehyde product and yields the aminoquinol form of the cofactor with the substrate-derived amino group still covalently bound (**27**). In the oxidative half-reaction **27** is re-oxidized by O<sub>2</sub> forming H<sub>2</sub>O<sub>2</sub> and the oxidized iminoquinone, which is converted back to LTQ with release of the ammonia product by hydrolysis.



**Scheme 7** Mechanism of LTQ biosynthesis.



**Scheme 8** The reaction mechanism of LTQ-dependent lysyl oxidase.

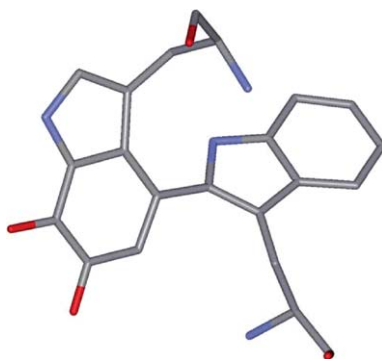
### 7.19.4.3 Tryptophan Tryptophylquinone

Tryptophan-derived quinone cofactors have been identified in amine dehydrogenases from a variety of bacterial sources.<sup>33</sup> TTQ (2',4-bis(tryptophan-6,7-dione)<sup>30</sup> (**Figure 3**) is the protein-derived cofactor of bacterial methylamine dehydrogenase and aromatic amine dehydrogenase.<sup>56</sup> In each of these enzymes two oxygens have been incorporated into a tryptophan side chain to generate a quinone, and the side chain of that residue has also become covalently cross-linked to another tryptophan residue to form TTQ. These TTQ-dependent dehydrogenases are soluble enzymes localized in the periplasmic space of Gram-negative bacteria. Each is an inducible enzyme that allows the host bacterium to utilize particular primary amines as a sole source of carbon and energy. Each enzyme catalyzes the oxidative deamination of the primary amine. These dehydrogenases do not utilize pyridine nucleotides (i.e., NAD<sup>+</sup> or NADP<sup>+</sup>) as electron acceptors. Instead, during the reductive half-reaction the quinone cofactor transfers electrons to a specific redox protein,<sup>57</sup> which mediates the transfer of these electrons to the membrane-bound respiratory chain (**Figure 5**).

#### 7.19.4.3.1 TTQ biosynthesis

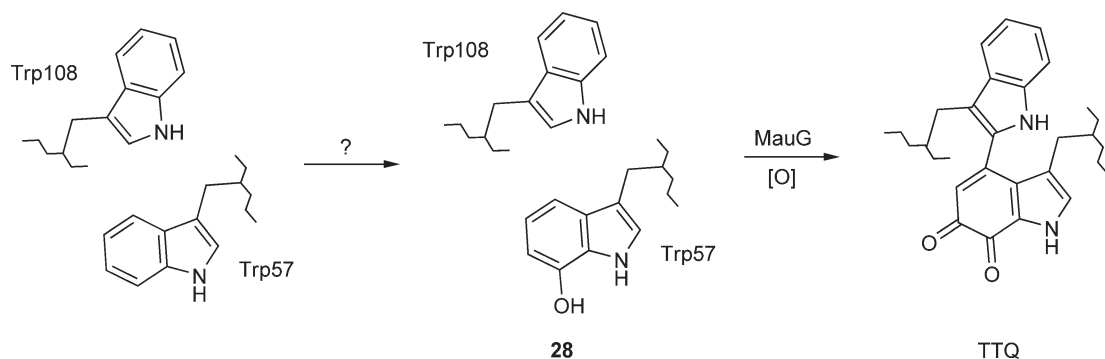
TTQ is formed by a posttranslational modification of two tryptophan residues of the polypeptide chain. Two atoms of oxygen are incorporated into the indole ring of one of the tryptophan residues and a covalent bond between the indole rings of the two tryptophan residues is formed. In contrast to the biogenesis of the tyrosine-derived cofactors TPQ and LTQ, the biogenesis of TTQ is not a self-processing event. The mechanism by which this occurs has not been completely elucidated, but it is known to require the action of other enzymes that are subject to the same genetic regulation as the structural genes for the enzyme.<sup>58,59</sup> The mechanism of biosynthesis of TTQ has been studied most extensively in *Paracoccus denitrificans*. The genes, which encode the two methylamine dehydrogenase subunits, together with nine other genes that relate to methylamine dehydrogenase expression and function, are clustered in the methylamine utilization (*mau*) locus with a gene order of *mauRFBEDACJGMN*.<sup>58,59</sup> The structural genes for the  $\alpha$  and  $\beta$  subunits of methylamine dehydrogenase are *mauB* and *mauA*, respectively.<sup>60</sup> The  $\beta$  subunit contains the tryptophan residues that are modified to TTQ, with  $\beta$ Trp57 incorporating the two oxygens and being cross-linked to  $\beta$ Trp108.<sup>61</sup> Amicyanin,<sup>62</sup> a type 1 copper protein, which is the obligate electron acceptor of methylamine dehydrogenase, is encoded by *mauC*.<sup>60</sup> The first gene, *mauR*, is a transcriptional activator<sup>63</sup> and *mauJ*, *mauM*, and *mauN* appear to encode proteins with unknown function.<sup>59</sup> The four other genes, *mauFEDG*, were each shown to be essential for methylamine dehydrogenase biosynthesis.<sup>58,59</sup> A role specifically in TTQ biosynthesis has been demonstrated for *mauG*. Its gene product, MauG, is a di-heme enzyme which has been purified and characterized.<sup>64</sup> Although MauG possesses two  $c$ -type hemes its physical properties and reactivity are more similar to P450-type cytochromes and oxygen-binding heme proteins than to typical  $c$ -type cytochromes.<sup>64,65</sup>

The biogenesis of TTQ during methylamine dehydrogenase biosynthesis was studied using a recombinant expression system, which contains the structural genes for methylamine dehydrogenase as well as *mauFEDG*.<sup>66</sup> To test the role of MauG in TTQ biogenesis, the *mauG* gene was either deleted or inactivated by site-directed



**Figure 5** Structure of the TTQ cofactor in methylamine dehydrogenase from *P. denitrificans*. In this enzyme oxygenation of  $\beta$ Trp57 and cross-linking with  $\beta$ Trp108 yields the TTQ cofactor, which is displayed as sticks colored gray for carbon, red for oxygen, and blue for nitrogen. The coordinates from PDB entry 2bbk were used to display this structure.





**Scheme 9** Intermediates in TTQ biosynthesis.

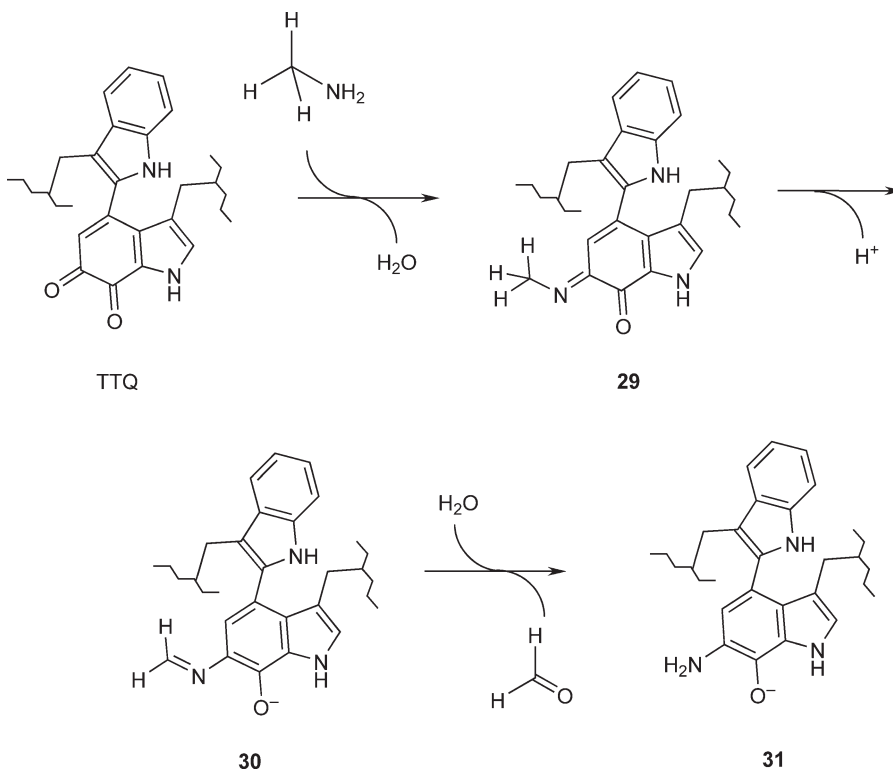
mutagenesis, and wild-type methylamine dehydrogenase was expressed in the background of the missing or inactive MauG. The resultant methylamine dehydrogenase was characterized by mass spectrometry, electrophoretic, and kinetic analyses.<sup>67</sup> The majority species that was isolated was a biosynthetic intermediate of methylamine dehydrogenase with incompletely synthesized TTQ containing  $\beta$ Trp57, which was mono-hydroxylated at the C7 position and with no covalent cross-link to residue  $\beta$ Trp108 (**28**) (**Scheme 9**). The position of insertion of oxygen into  $\beta$ Trp57 was determined by isotope labeling studies of the MauG-dependent biosynthetic reaction *in vitro* in the presence of  $^{18}\text{O}_2$ .<sup>68</sup> Incubation of this intermediate *in vitro* with purified MauG resulted in completion of TTQ biosynthesis and formation of active methylamine dehydrogenase. The MauG-dependent TTQ biosynthesis could be achieved with oxidation equivalents ([O] in **Scheme 9**) provided by molecular oxygen plus a variety of artificial electron donors or by  $\text{H}_2\text{O}_2$ .<sup>65</sup> Although the mechanism by which the first oxygen is inserted into  $\beta$ Trp57 is unknown, it is documented that MauG is able to catalyze the insertion of the second oxygen, and the covalent bond formation between indole rings during the biosynthesis of TTQ in methylamine dehydrogenase.<sup>67,69</sup>

#### 7.19.4.3.2 TTQ-dependent catalysis

The overall reaction catalyzed by the TTQ-dependent amine dehydrogenases is the conversion of a primary amine substrate to its corresponding aldehyde plus ammonia with two substrate-derived electrons transferred through TTQ to an electron acceptor. Although the reductive half-reaction of TTQ-dependent dehydrogenases is similar to that of TPQ-dependent oxidases, the oxidative reactions are not since molecular oxygen is not the electron acceptor. The oxidative half-reactions of TTQ-dependent dehydrogenases also differ from those of most other known dehydrogenases, which use pyridine nucleotides ( $\text{NAD}^+$  or  $\text{NADP}^+$ ) as electron acceptors. The oxidative reactions of TTQ-dependent enzymes involve interprotein electron transfer usually to a type 1 copper protein. For most methylamine dehydrogenases, the electron acceptor is amicyanin,<sup>62</sup> and for aromatic amine dehydrogenase, it is azurin.<sup>70</sup> The fact that these soluble dehydrogenases transfer electrons to other soluble redox proteins has been exploited to develop these systems, particularly the methylamine dehydrogenase–amicyanin complex as models for the study of interprotein electron transfer reactions.<sup>57,71</sup> The electron transfer reactions between each TTQ-dependent dehydrogenase and its respective copper protein have been extensively studied.<sup>57</sup> X-ray crystal structures of protein electron transfer complexes of methylamine dehydrogenase and amicyanin<sup>72,73</sup> from *P. denitrificans*, and aromatic amine dehydrogenase and azurin<sup>74</sup> from *Alcaligenes faecalis* have been determined.

The reductive half-reaction of methylamine dehydrogenase<sup>56</sup> is shown in **Scheme 10**. The methylamine substrate initiates a nucleophilic attack on the quinone carbon at the C6 position of the TTQ cofactor displacing the oxygen to form a substrate–TTQ Schiff base adduct (**29**). The reactivity of the C6 position was demonstrated by covalent adduct formation at this position by hydrazines which are inactivators of methylamine dehydrogenase.<sup>75</sup> Deprotonation of the substrate-derived carbon of **29** by an active-site amino acid residue results in reduction of the cofactor and yields an intermediate in which the Schiff base is now between the nitrogen and substrate-derived carbon (**30**). Hydrolysis of **30** releases the formaldehyde product and yields the aminoquinol form of the cofactor with the substrate-derived amino group still covalently bound (**31**).

TTQ is a two-electron carrier and the electron acceptor for methylamine dehydrogenase, amicyanin, is a one-electron carrier. Thus, the oxidative half-reaction of methylamine dehydrogenase must occur in two

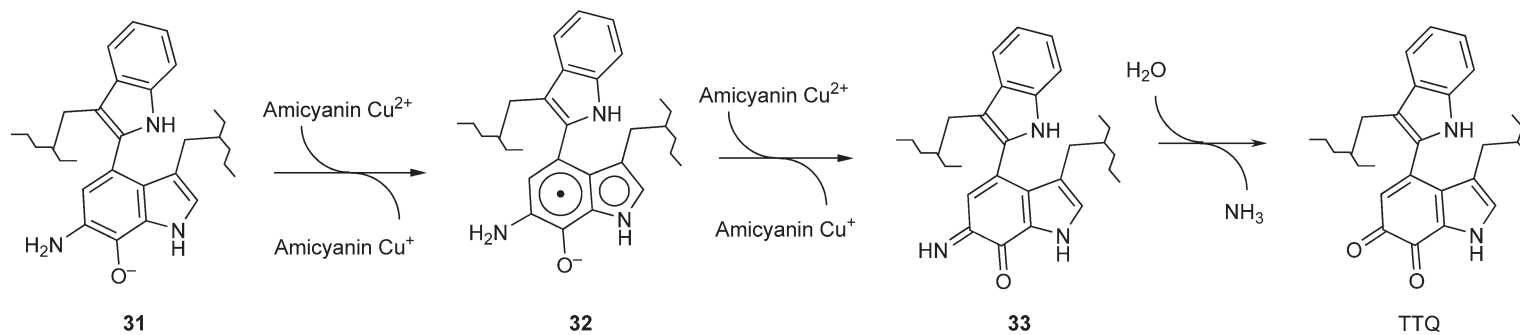


**Scheme 10** The reductive half-reaction of TTQ-dependent methylamine dehydrogenase.

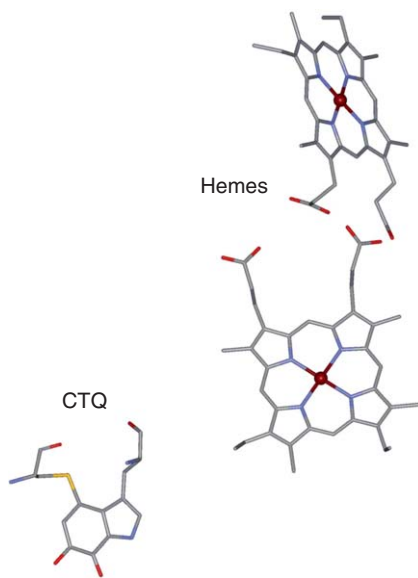
one-electron oxidations as shown in **Scheme 11**. The product of the first one-electron transfer to an oxidized amicyanin is an aminosemiquinone TTQ (**32**).<sup>76,77</sup> The immediate product of the second one-electron transfer to an oxidized amicyanin is an aminoquinone TTQ (**33**), which retains the substrate-derived amino group. The ammonia product is released by hydrolysis of **33** to regenerate TTQ. Alternatively, in the presence of excess methylamine in the steady state, the substrate-derived amino group may be displaced from **33** by another molecule of methylamine rather than water.<sup>78</sup> Investigation of the pH dependence of the redox potentials for the oxidized/reduced couples of methylamine dehydrogenase indicated that the two-electron transfer is linked to the transfer of a single proton.<sup>79</sup> This correlated with structural analysis, which revealed that the quinol oxygen at the C7 position is shielded from the solvent and therefore is not protonated. Thus as indicated in **Scheme 11**, the aminoquinol (**30**) is singly protonated and anionic, and the aminosemiquinone (**32**) is unprotonated and anionic.

#### 7.19.4.4 Cysteine Tryptophylquinone

Cysteine tryptophylquinone (CTQ) has been identified as the protein-derived cofactor of quinohemoprotein amine dehydrogenase from at least two sources, *P. denitrificans*<sup>80</sup> and *Ps. putida*.<sup>81</sup> These enzymes contain CTQ as well as two covalently bound *c*-type hemes (**Figure 6**). CTQ is formed by posttranslational modifications during which two atoms of oxygen are incorporated into the indole ring of a tryptophan residue and a covalent bond is formed between the modified indole ring and the sulfur of a cysteine residue. Quinohemoprotein amine dehydrogenase is also a soluble enzyme localized in the periplasmic space of Gram-negative bacteria. It is an inducible enzyme that allows the host bacterium to utilize certain primary amines, particularly butylamine and substituted phenylethylamines as a source of carbon and energy.<sup>82,83</sup> Each enzyme catalyzes the oxidative deamination of the primary amine. Quinohemoprotein amine dehydrogenase does not utilize NAD<sup>+</sup> or NADP<sup>+</sup> as electron acceptors but rather soluble periplasmic metalloproteins; cytochrome *c*-550 in *P. denitrificans*<sup>84</sup> and azurin in *Ps. putida*.<sup>83</sup>

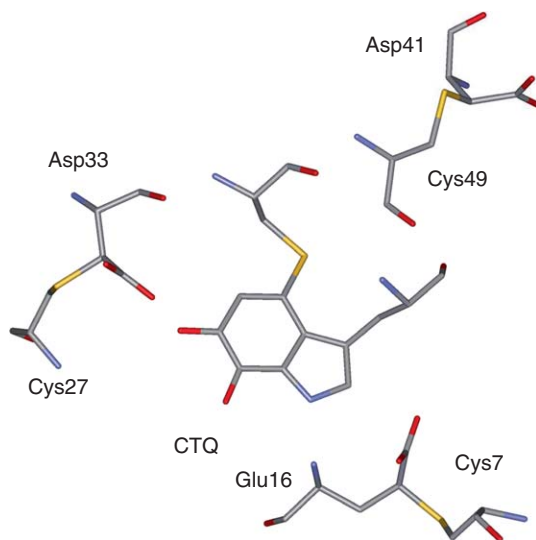


**Scheme 11** The oxidative half-reaction of TTQ-dependent methylamine dehydrogenase.



**Figure 6** Structure and relative orientations of the CTQ and heme cofactors in quinohemoprotein amine dehydrogenase from *P. denitrificans*. In this enzyme, oxygenation of  $\gamma$ Trp43 and cross-linking with  $\gamma$ Cys37 yields the CTQ cofactor. CTQ and the two hemes present on the  $\alpha$  subunit are displayed as sticks colored gray for carbon, red for oxygen, blue for nitrogen, yellow for sulfur, and dark red for iron. The coordinates from PDB entry 1pby were used to display this structure.

Although CTQ- and TTQ-dependent enzymes possess the same tryptophylquinone moiety at their active sites, the overall structures of the CTQ- and TTQ-dependent enzymes are quite different. The overall structures of the TTQ enzymes, methylamine dehydrogenase<sup>61</sup> and aromatic amine dehydrogenase<sup>85</sup> are similar and each is a heterotetramer of two identical larger  $\alpha$  subunits of molecular weight of 40 000–50 000, and two identical smaller  $\beta$  subunits of molecular weight of approximately 15 000 that contain TTQ. In contrast, the CTQ-bearing quinohemoprotein amine dehydrogenase is an  $\alpha\beta\gamma$  heterotrimeric protein.<sup>80,81</sup> The smallest  $\gamma$  subunit contains CTQ. In addition to CTQ, the  $\gamma$  subunit contains three novel thioether cross-links that are formed between the sulfur of a cysteine residue and the  $\beta$ -methylene carbon of an aspartic acid residue, or between the sulfur of a cysteine residue and the  $\gamma$ -methylene carbon of a glutamic acid residue (**Figure 7**).



**Figure 7** Structure and relative orientations of the cross-linked amino acid residues in quinohemoprotein amine dehydrogenase from *P. denitrificans*. The three pairs of cross-linked amino acid residues and CTQ are displayed as sticks colored gray for carbon, red for oxygen, blue for nitrogen, and yellow for sulfur. The coordinates from PDB entry 1pby were used to display this structure.

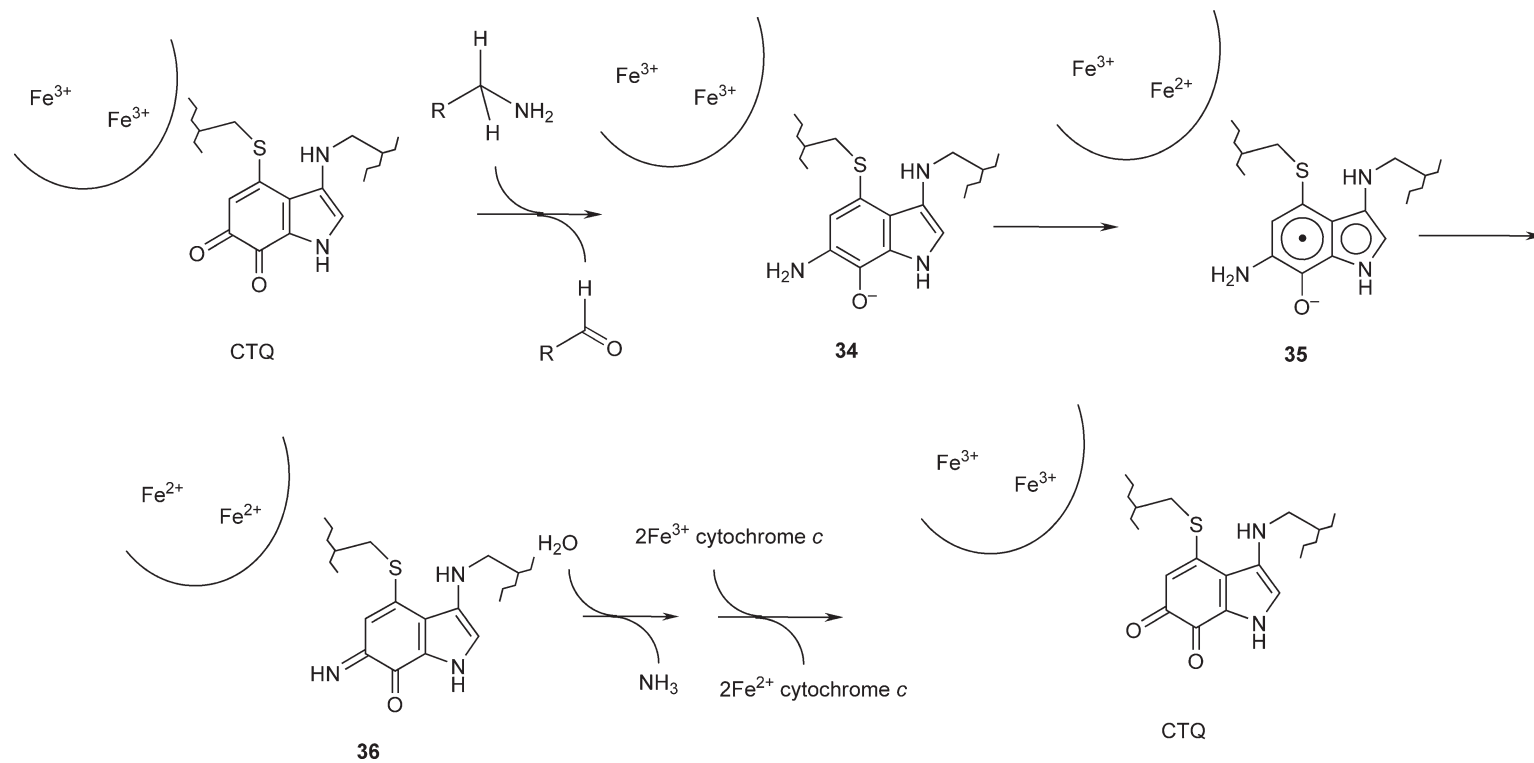
These cross-linked amino acid residues appear to have no direct participation in the catalytic mechanism. They do play a structural role in determining the tertiary structure of the  $\gamma$  subunit. It is interesting to note that the TTQ-dependent methylamine dehydrogenase does not have these thioether cross-linked residues, but does have six intra-subunit disulfide bonds between cysteine residues, which play a structural role in determining the tertiary structure of the TTQ-bearing  $\beta$  subunit of that enzyme.<sup>61</sup> The  $\alpha$  subunit of QHNDH contains two  $c$ -type hemes. One heme  $c$  is solvent-accessible and the other is fully buried within the  $\alpha$  subunit and located approximately 9 Å from the tryptophylquinone moiety of CTQ on the  $\gamma$  subunit. The  $\alpha$  and  $\gamma$  subunits sit on the surface of the  $\beta$  subunit that with the  $\gamma$  subunit forms the enzyme active site.

#### 7.19.4.4.1 CTQ biosynthesis

Relatively little is known about the mechanism of biogenesis of CTQ. The gene cluster, which encodes quinohemoprotein amine dehydrogenase<sup>80</sup> is completely different from that which encodes methylamine dehydrogenase. Present in the quinohemoprotein amine dehydrogenase operon is an open reading frame (ORF2), which encodes a putative radical *S*-adenosylmethione protein,<sup>86</sup> which is not present in the methylamine dehydrogenase operon. Site-specific mutations in the sequence motifs for the putative iron–sulfur cluster or *S*-adenosylmethionine-binding site in the ORF2 protein demonstrated that the ORF2 protein is necessary for the posttranslational processing of the  $\gamma$ -subunit, most probably participating in the formation of the intra-peptidyl thioether cross-links.<sup>87</sup> No gene analogous to *mauG* is present in the quinohemoprotein amine dehydrogenase gene cluster. Although perhaps coincidental, it is interesting that a gene encoding a di-heme protein is also present in the quinohemoprotein amine dehydrogenase gene cluster, but in this case, the gene product is not a free cytochrome but rather a subunit of quinohemoprotein amine dehydrogenase, which associates closely with the CTQ-bearing subunit. It is intriguing to speculate that the di-heme subunit may perform a role in CTQ biogenesis similar to that of the di-heme MauG during TTQ biogenesis, but there is no evidence for or against such a hypothesis to date.

#### 7.19.4.4.2 CTQ-dependent catalysis

The reaction mechanism of quinohemoprotein amine dehydrogenase<sup>88</sup> is essentially the same as that of the TTQ-dependent dehydrogenases. The major difference is that the immediate electron acceptors for the two electrons of reduced CTQ are the two hemes, which are present on the  $\alpha$  subunit, rather than exogenous copper proteins. After reduction by reduced CTQ, the di-heme subunit then transfers electrons to an endogenous cytochrome  $c$ . The proposed mechanism for the overall reaction catalyzed by the CTQ-dependent quinohemoprotein amine dehydrogenase of *P. denitrificans* is shown in [Scheme 12](#).<sup>88</sup> The two irons present in the hemes of the di-heme subunit of the enzyme are shown to the left of the CTQ cofactor in each intermediate. The reductive half-reaction of quinohemoprotein amine dehydrogenase is essentially the same as that for methylamine dehydrogenase, which is shown in [Scheme 10](#). The primary amine substrate initiates a nucleophilic attack on the quinone carbon at the C6 position of the CTQ cofactor, displacing the oxygen to form a substrate-TTQ Schiff base adduct. Deprotonation of the substrate-derived carbon by an active-site amino acid residue results in reduction of the cofactor and yields an intermediate in which the Schiff base is now between the nitrogen and substrate-derived carbon. Hydrolysis then releases the aldehyde product and yields the aminoquinol form of the cofactor with the substrate-derived amino group still covalently bound ([34](#)). CTQ is a two-electron carrier and the immediate electron acceptor, the nearest heme on the  $\alpha$  subunit, is a one-electron carrier. Thus, the oxidative half-reaction of CTQ must occur in two one-electron oxidations. The product of the first one-electron transfer to the first heme is an aminosemiquinone CTQ ([35](#)). Interheme electron transfer and the second one-electron transfer to heme from [35](#) will yield as an immediate product an aminoquinone CTQ ([36](#)), which still retains the substrate-derived amino group. The ammonia product is released by hydrolysis of [36](#) to regenerate CTQ, and electron transfer from the hemes of the  $\alpha$  subunit to two molecules of cytochrome  $c$  regenerate the di-ferric form of the  $\alpha$  subunit to yield the fully oxidized quinohemoprotein amine dehydrogenase.



**Scheme 12** The reaction mechanism of CTQ-dependent quinohemoprotein amine dehydrogenase.



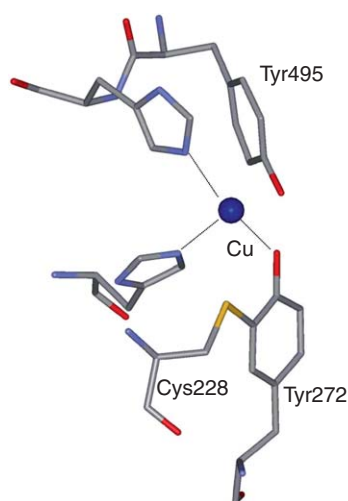
## 7.19.5 Functional Covalently Cross-Linked Amino Acid Residues in Metalloproteins

### 7.19.5.1 Cysteine-Tyrosine, the Galactose Oxidase Cofactor

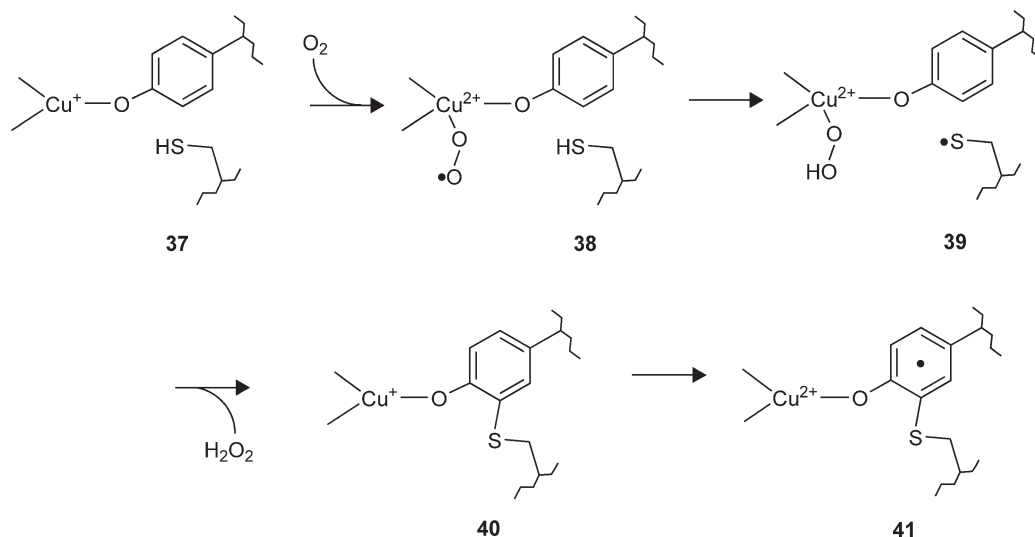
One of the first descriptions of the functional cross-linking of amino acid residues within a protein to generate a reactive prosthetic group is the cysteine–tyrosine cross-linked protein-derived cofactor of the fungal enzyme galactose oxidase. The crystal structure of galactose oxidase from *Dactylium dendroides*<sup>89</sup> revealed the presence of a thioether bond covalently linking Cys228 and Tyr272, with this tyrosine also acting as a ligand to a copper, which is present in the active site (Figure 8). The side chains of these two residues also form an aromatic plane that stacks with the indole ring of a tryptophan residue. This type of cysteine–tyrosine–copper cofactor is also present in glyoxal oxidase<sup>90</sup> from fungi and yeast. This family of metal coordinated cysteine–tyrosine cross-linked cofactors may extend beyond the copper-containing galactose and glyoxal oxidases. A similar cysteine–tyrosine cross-linked species has been identified in the active site of the mammalian enzyme cysteine dioxygenase. However, this is a mono-nuclear iron-dependent enzyme rather than copper-dependent enzyme.<sup>91</sup>

#### 7.19.5.1.1 Galactose oxidase cofactor biosynthesis

Galactose oxidase is initially synthesized as a pro-enzyme with a pre-sequence. Initial studies suggested that this pre-sequence was required for generation of the galactose oxidase cofactor and is cleaved during processing to generate the mature enzyme. When cells expressing galactose oxidase were grown in copper-depleted media the isolated galactose oxidase was produced in a precursor form lacking the thioether bond and also possessing an additional 17-amino acids at the N-terminus. The formation of the cofactor was a self-processing event. The cysteine–tyrosine cross-link formed spontaneously *in vitro* in the presence of O<sub>2</sub> and copper, yielding a free radical coupled-Cu<sup>2+</sup> cofactor in the active site of the mature enzyme along with cleavage of the pre-sequence.<sup>92</sup> The crystal structure of this precursor protein has been determined and compared to that of the mature protein.<sup>93</sup> The pre-sequence did not make direct contact with the active site, but several residues were shown to occupy different positions in the pre-protein from those that they occupy in the mature enzyme. In particular, the tyrosine and cysteine residues to be cross-linked, as well as the tryptophan that  $\pi$ -stacks with the cross-link after it is formed are in much different positions in the precursor than in the mature protein. However, the role of the pre-sequence was subsequently brought into question by studies using a recombinant



**Figure 8** Structure of the cysteine–tyrosine–copper cofactor in galactose oxidase from *Dactylium dendroides*. The cross-linked amino acid residues Cys228 and Tyr272, Tyr495, which is believed to serve as an active site base, and histidine residues which provide ligands for the copper in the active site, are displayed as sticks colored gray for carbon, red for oxygen, blue for nitrogen, and yellow for sulfur. The copper is displayed as a dark blue sphere. The coordinates from PDB entry 1gog were used to display this structure.



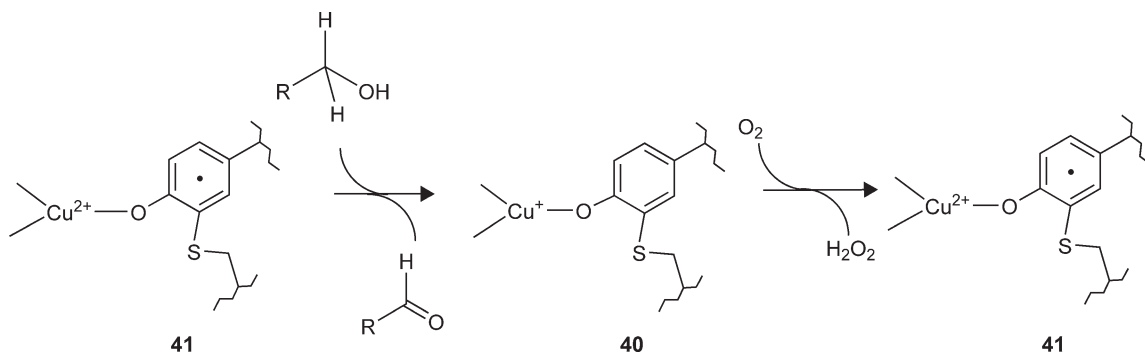
**Scheme 13** Mechanism of galactose oxidase cofactor biosynthesis.

galactose oxidase precursor protein, which was isolated with a processed N-terminus (i.e., no pre-sequence) and no tyrosine–cysteine cross-link. Mature galactose oxidase could be readily formed *in vitro* from that precursor by addition of  $O_2$  and  $Cu^{1+}$ .<sup>94</sup> It was noteworthy in this study that if  $Cu^{2+}$  was used, the maturation process required hours to complete, as opposed to requiring seconds when  $Cu^{1+}$  was used.

A mechanism proposed by Whittaker and Whittaker<sup>94</sup> for biosynthesis of the tyrosine–cysteine cross-link that recognizes the requirement specifically for  $Cu^{1+}$  is shown in **Scheme 13**. It is proposed that  $Cu^{1+}$  initially binds to a pre-organized active site (**37**) in which it is coordinated with Tyr272 and most probably two histidine residues. Dioxygen reacts with this complex to generate an adduct in which electron transfer from copper to oxygen results in an oxygenated complex with superoxide coordinated with  $Cu^{2+}$  (**38**). This oxygenated species then abstracts a hydrogen atom from Cys228, which is in close proximity, to generate a thiyl free radical on that cysteine (**39**). The thiyl free radical then adds to the aromatic ring of Tyr272. This is immediately followed by loss of a proton and reduction of the metal ion restores the aromatic conjugation of the ring.  $H_2O_2$  is lost in this process and the product is the covalently cross-linked cysteine–tyrosine coordinated with  $Cu^{1+}$  (**40**). This species corresponds to the mature cofactor in the fully reduced active site formed during catalytic turnover (discussed later). It then undergoes rapid reaction with a second molecule of  $O_2$  to form the oxidized metallo-radical complex (**41**) seen in the resting state of galactose oxidase. Alternatively it was proposed that the process may proceed through a tyrosyl phenoxyl free radical intermediate capable of undergoing nucleophilic substitution by the neighboring Cys228 thiolate to generate the covalent cross-link.

#### 7.19.5.1.2 Galactose oxidase cofactor-dependent catalysis

The overall reaction catalyzed by galactose oxidase is the oxidation of a primary alcohol to the corresponding aldehyde, coupled to the reduction of dioxygen to hydrogen peroxide.<sup>95</sup> This is shown in **Scheme 14**. The active form of the enzyme consists of the tyrosyl radical at Tyr272 coordinated with  $Cu^{2+}$  (**41**). To initiate the reductive half-reaction, the alcohol substrate will enter the active site and position itself such that the substrate oxygen coordinates with copper. This is followed by deprotonation of the hydroxyl of the substrate by an active site base, which on the basis of the crystal structure is probably Tyr495. Electron transfer occurs from the alkoxide to  $Cu^{2+}$  and is followed by abstraction of a hydrogen atom by the tyrosyl radical from the carbon atom bearing the oxygen. This generates the aldehyde product and the fully reduced active site (**40**). Alternatively, a reversed order of hydrogen atom abstraction and electron transfer has been proposed. In the oxidative half-reaction **40** is re-oxidized by  $O_2$  releasing  $H_2O_2$  and regenerating the oxidized metallo-radical complex (**41**).



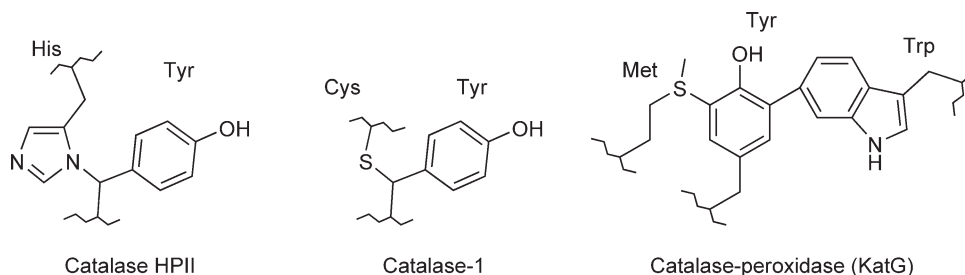
**Scheme 14** Galactose oxidase cofactor-dependent alcohol oxidation.

### 7.19.5.2 Cross-Linked Amino Acid Residues in Heme Enzymes

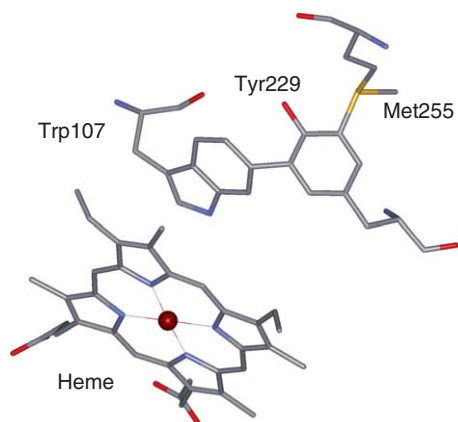
Covalently cross-linked amino acids have been identified in the proximity of the catalytic heme of certain redox enzymes (Figure 9). Cross-linked amino acid residue pairs have been observed in the active sites of two heme enzymes. A histidine-tyrosine cross-link has been identified in catalase HP11,<sup>96</sup> and a cysteine-tyrosine cross-link has been identified in catalase-1.<sup>97</sup> In contrast to the galactose oxidase cofactor, the cross-linked tyrosine residue is not directly coordinated with a metal, but is in close proximity to the heme iron. Also in contrast to the galactose oxidase cofactor, in these enzymes the cross-linking bond is formed with the  $\beta$ -carbon of the tyrosine residue rather than to the aromatic ring. A third heme enzyme, the bifunctional catalase-peroxidase (KatG) possesses a cross-linked triad of amino acid residues at the heme active site. Crystal structures of KatGs from three different sources: *Haloarcula marismortu*,<sup>98</sup> *Burkholderia pseudomonas*,<sup>99</sup> and *Mycobacterium tuberculosis*,<sup>100</sup> each reveal the presence of two covalent bonds between the amino acid side chains of tryptophan, tyrosine, and methionine residues. In this case, there are two covalent cross-links to the tyrosine residue, both to the aromatic ring. Although the three cross-linked amino acid residues are each located in close proximity to the heme in the active site, the amino acids are well separated in the primary sequence. In the *M. tuberculosis* structure these residues are Trp107, Tyr229, and Met255 which are located on the distal side of the active-site heme (Figure 10).

#### 7.19.5.2.1 KatG cofactor biosynthesis

To gain insight into the mechanism of formation of the methionine-tyrosine-tryptophan cross-link of KatG, a protein lacking the cross-links was prepared by expressing KatG in cells deficient in heme by growth of the host cells in iron-depleted media. After reconstitution with heme *in vitro*, formation of the cross-links was observed after addition of peroxyacetic acid.<sup>101</sup> On the basis of the results of those studies, the following mechanism was proposed for the biosynthesis of the cross-linked cofactor *in vivo* (Scheme 15). Initially in the presence of the three amino acid residues at the active site, the heme reacts with  $\text{H}_2\text{O}_2$  (a substrate for mature KatG). Rather than the normal catalytic reaction, this results in the generation a high valent oxoferryl porphyrin cation radical intermediate (also called compound I) (42). The crystal structure of *M. tuberculosis* KatG shows that the indole



**Figure 9** Cross-linked amino acid residues present at the active sites of heme proteins.



**Figure 10** Structure and relative orientations of the methionine–tyrosine–tryptophan and heme cofactors in KatG from *M. tuberculosis*. The cross-linked amino acid residues Met255, Tyr229, and Trp107 and the nearby heme are displayed as sticks colored gray for carbon, red for oxygen, blue for nitrogen, yellow for sulfur, and dark red for iron. The coordinates from PDB entry 1sj2 were used to display this structure.

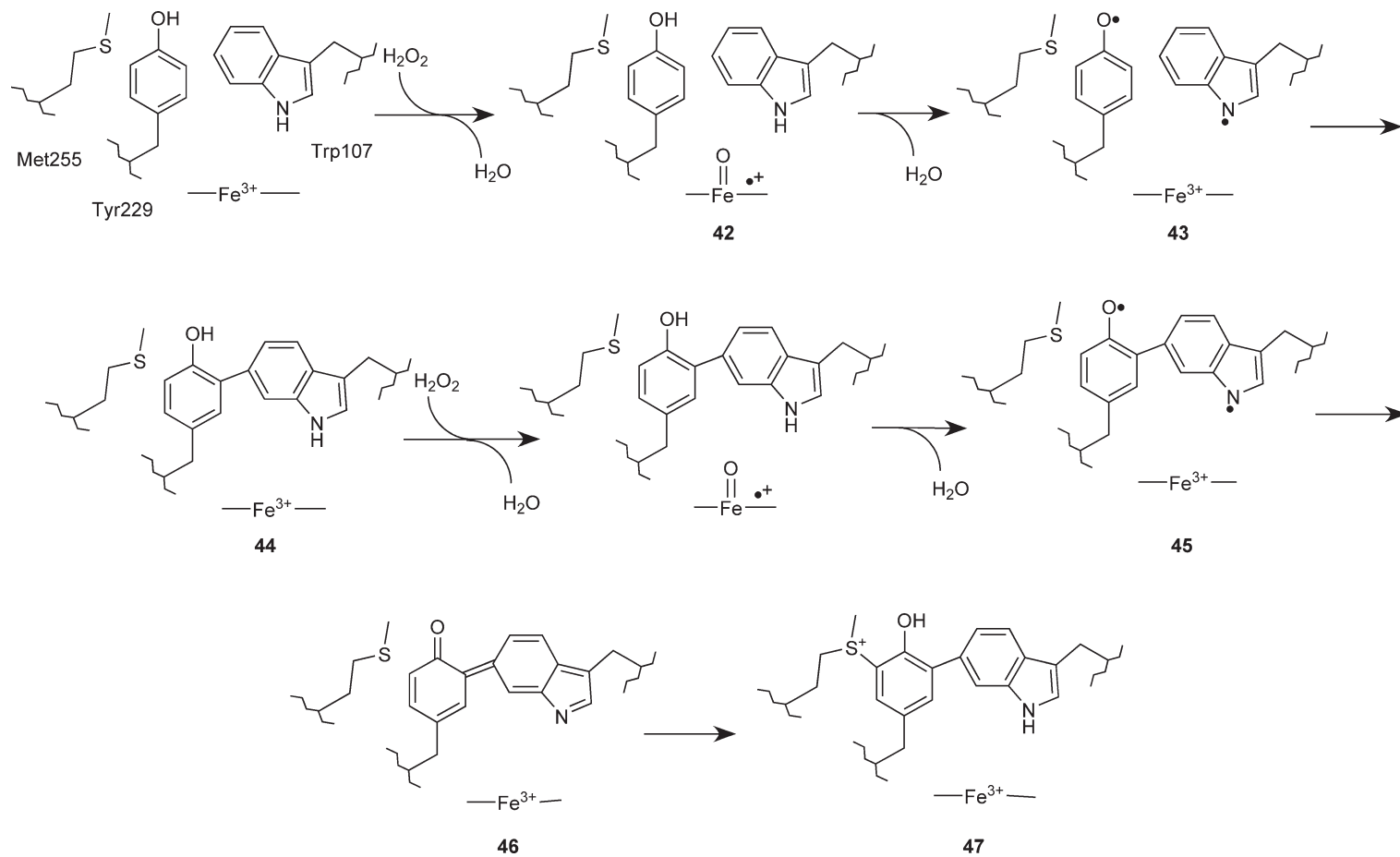
ring of Trp107 and phenyl ring of Tyr229 are within 3.7 and 5.4 Å of the heme, respectively.<sup>100</sup> Compound I is a highly reactive short-lived two-electron oxidant that oxidizes both residues Tyr229 and Trp107 resulting in the formation of an intermediate with free radicals on each amino acid residue with release of water (43). The heme iron returns to its ferric state owing to the oxidation of the amino acid residues. In essence, during this first step in the biosynthetic reaction, the heme functions as a peroxidase by using these two amino acid residues as co-substrates in the reaction. Coupling of the two radicals occurs, resulting in the formation of an intermediate in which the covalent tyrosine–tryptophan bond has been formed (44). At this point in the biosynthetic process, a second molecule of H<sub>2</sub>O<sub>2</sub> enters the active site, and reacts with the ferric heme iron to again generate a high valent oxoferryl porphyrin radical cation intermediate. This results in another oxidation reaction in which the now cross-linked residues Tyr229 and Trp107 are simultaneously oxidized resulting in the formation of an intermediate with free radicals present on each of the cross-linked amino acid residues (45). Rearrangement of this intermediate yields a quinone-like intermediate (46), which is then a target for nucleophilic attack of the sulfur atom of Met255, which results in the formation of the mature active site with the methionine–tyrosine–tryptophan cross-link (47).

#### 7.19.5.2.2 *KatG*-dependent catalysis

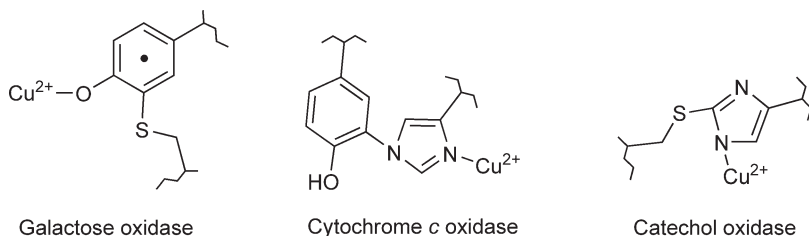
KatGs are present in bacteria, fungi, and plants. They exhibit a high sequence homology with peroxidases from these sources. However, KatGs are bifunctional enzymes which possess both catalase (2H<sub>2</sub>O<sub>2</sub> → H<sub>2</sub>O + O<sub>2</sub>) and peroxidase (2AH + H<sub>2</sub>O<sub>2</sub> → 2A + 2H<sub>2</sub>O) activities. Once the methionine–tyrosine–tryptophan cross-link is complete, the mature KatG is now fully functional, and subsequent reactions with H<sub>2</sub>O<sub>2</sub> will be those of the normal catalytic functions. An interesting conclusion from these findings is that the active-site heme is required to generate the cross-linked amino acids during biosynthesis, and then the cross-linked amino acids are required by the heme to perform its catalytic function in the mature enzyme. This structural feature is not found in the mono-functional peroxidases, suggesting that these cross-linked amino acids may impart certain catalytic activities specifically to the KatGs. Site-directed mutagenesis studies have indeed confirmed that the cross-link is required for catalytic, but not for peroxidatic activity.<sup>102</sup>

#### 7.19.5.3 Cross-Linked Amino Acid Residues in Complex Copper Enzymes

In addition to the tyrosine–cysteine cross-link of galactose oxidase (discussed earlier), other cross-linked amino acids have been identified at the active sites of other copper proteins (Figure 11). In each case these are metalloenzymes possessing multiple metal atoms.



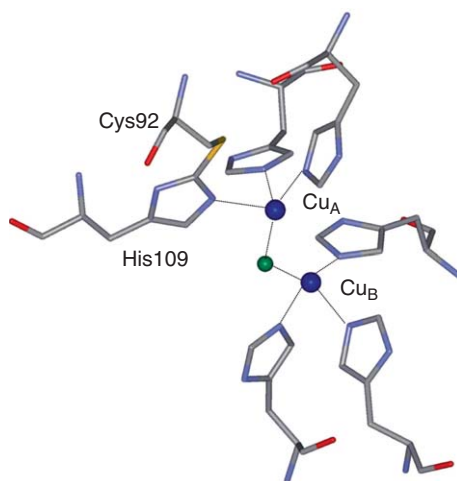
**Scheme 15** Mechanism for biosynthesis of the methionine-tyrosine-tryptophan cross-linked cofactor in KatG.



**Figure 11** Cross-linked amino acid residues present at the active sites of copper proteins.

Catechol oxidase is a ubiquitous enzyme that catalyzes the reaction between catechol and oxygen to yield benzoquinone and water. It is a complex copper-containing protein, which acts also on a variety of substituted catechols. The crystal structure of a plant catechol oxidase reveals that the  $\text{Cu}_A$  of the dinuclear  $\text{Cu}_A\text{--Cu}_B$  metal site is coordinated by three histidine ligands, one of which is covalently cross-linked to a cysteine residue by a thioether bond<sup>103</sup> (Figure 12). A similar cysteine–histidine bond to one of the copper ligands in the dinuclear metal site of Octopus hemocyanin has also been identified in its crystal structure.<sup>104</sup> Hemocyanins play a role in biological oxygen transport similar to that of hemoglobin and carry oxygen in the blood of some mollusks and crustaceans. Rather than using heme iron to bind  $\text{O}_2$ , hemocyanins possess two copper atoms that reversibly bind  $\text{O}_2$ . The mechanisms by which the amino acid cross-links are formed in catechol oxidase and hemocyanin have yet to be determined.

Cytochrome *c* oxidase is a component of the membrane-bound respiratory electron transport chain. It receives electrons from cytochrome *c*, and transfers them with protons to oxygen, converting molecular oxygen into two molecules of water. In the process, it translocates four protons across the membrane helping to establish a chemiosmotic potential that provides the driving force for ATP synthesis through ATP synthase. The crystal structures of bacterial and mammalian cytochrome *c* oxidase<sup>105,106</sup> reveal that a covalent cross-link exists between the side chains of a tyrosine residue, and a histidine residue (Figure 11), which provides one of the copper ligands of the binuclear  $\text{Cu}_B$  component of the  $\text{Cu}_B\text{--heme}_{a_3}$  center. Three categories of proton-pumping heme–copper oxygen reductases are recognized in nature; type A (*aa*<sub>3</sub>-type oxidases), type B (*ba*<sub>3</sub>-type oxidases), and type C (*cbb*<sub>3</sub>-type oxidases).<sup>107</sup> The histidine–tyrosine cross-link has been identified in



**Figure 12** Structure of the cysteine–histidine cross-link and dinuclear copper center of catechol oxidase from sweet potato. The cross-linked amino acid residues Cys92 and His109 and histidines that coordinate the two coppers are displayed as sticks colored gray for carbon, red for oxygen, blue for nitrogen, and yellow for sulfur. The coppers of the dinuclear copper center are displayed as dark blue spheres and a solvent molecule that bridges the two coppers is displayed as a green sphere. The coordinates from PDB entry 1bt3 were used to display this structure.



crystal structures of the type A and type B enzymes and mass spectrometry was used to show that the predicted tyrosine forms a histidine–tyrosine cross-linked cofactor in the active site of the C-type oxygen reductases.<sup>108</sup> Thus, the histidine–tyrosine cross-link is believed to be a common feature of each of the three distinct families of heme–copper oxygen reductases in nature. The mechanism by which the histidine–tyrosine amino acid cross-link is formed has yet to be determined.

### 7.19.6 Green Fluorescent Protein and Protein-Derived Fluorophores

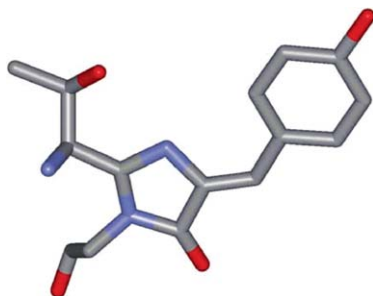
The GFP from the jellyfish *Aequorea victoria*, although not an enzyme, has become widely used as a marker for gene expression and localization. Although the fluorophore of GFP is not technically a protein-derived cofactor, it is a protein-derived fluorophore. This is another example of posttranslational modifications, which endow amino acid residues with a new function. In this case, the new function is not one which assists in catalysis. Instead, the results of these posttranslational modifications create new fluorescent properties, which serve a different biological function. As with most of the protein-derived cofactors discussed earlier, the presence and identity of the fluorophore is not evident from the amino acid sequence of the protein. The structure of the GFP fluorophore and mechanism of its biosynthesis were deduced from structural analyses. The X-ray crystal structure of GFP revealed that the covalently bound fluorescent chromophore is derived from three adjacent amino acids, serine–tyrosine–glycine on the polypeptide chain (**Figure 13**).<sup>109</sup>

#### 7.19.6.1 GFP-Fluorophore Biosynthesis

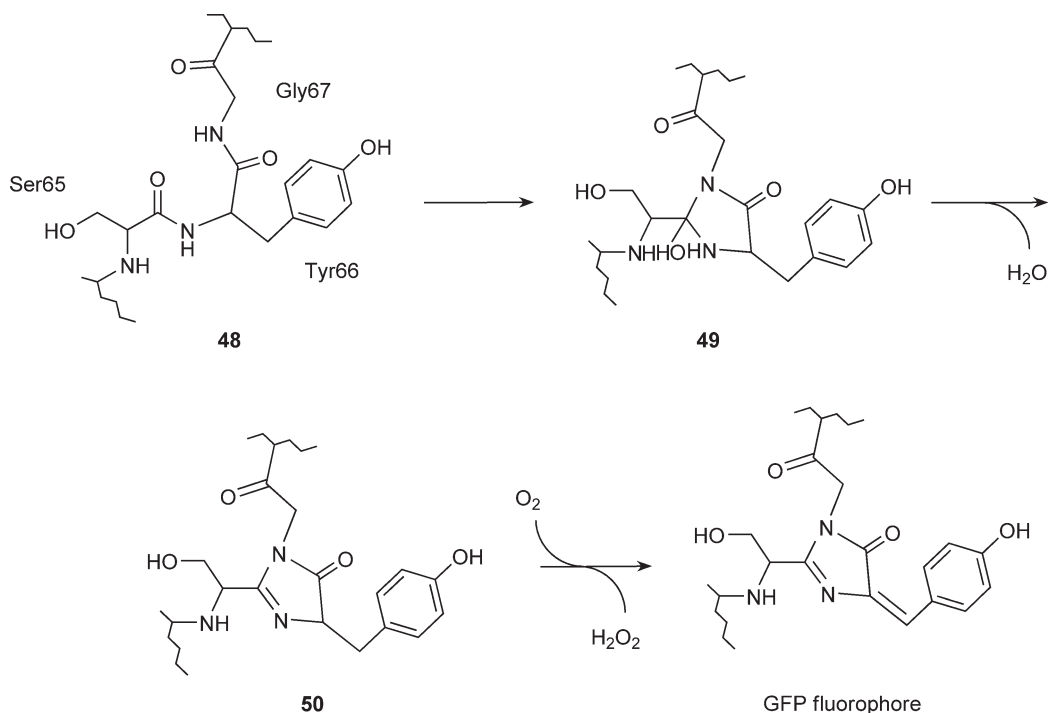
The formation of the GFP fluorophore involves self-processing reactions that bear similarity to the mechanism of formation of MIO discussed earlier (see **Scheme 2**). The proposed mechanism for the biosynthesis of the GFP fluorophore is shown in **Scheme 16**. It is proposed that after translation the protein (**48**) is folded in a way to encourage nucleophilic attack by the amide nitrogen of Gly67 on the amide carbon of Ser65 to form a cyclic intermediate (**49**), which loses water to yield an imidazolidin-5-one intermediate (**50**). This condensation reaction is followed by autoxidation of the  $\alpha$ – $\beta$  bond of Tyr66 to yield the GFP fluorophore.<sup>110</sup> It should be noted that the isolated tripeptide will not spontaneously undergo these reactions. For synthesis of the GFP cofactor to occur, the protein must be folded in such a manner to position the residues to facilitate the biosynthetic reactions.

#### 7.19.6.2 Related Fluorescent Protein Fluorophores

Other naturally occurring fluorescent proteins have been found in aquatic organisms including reef corals and sea anemones.<sup>111–113</sup> The red fluorescent protein DsRed was discovered in the Anthozoan genus *Discosoma*. The DsRed fluorescent protein fluorophore features an imidazoline ring system similar to the GFP fluorophore but

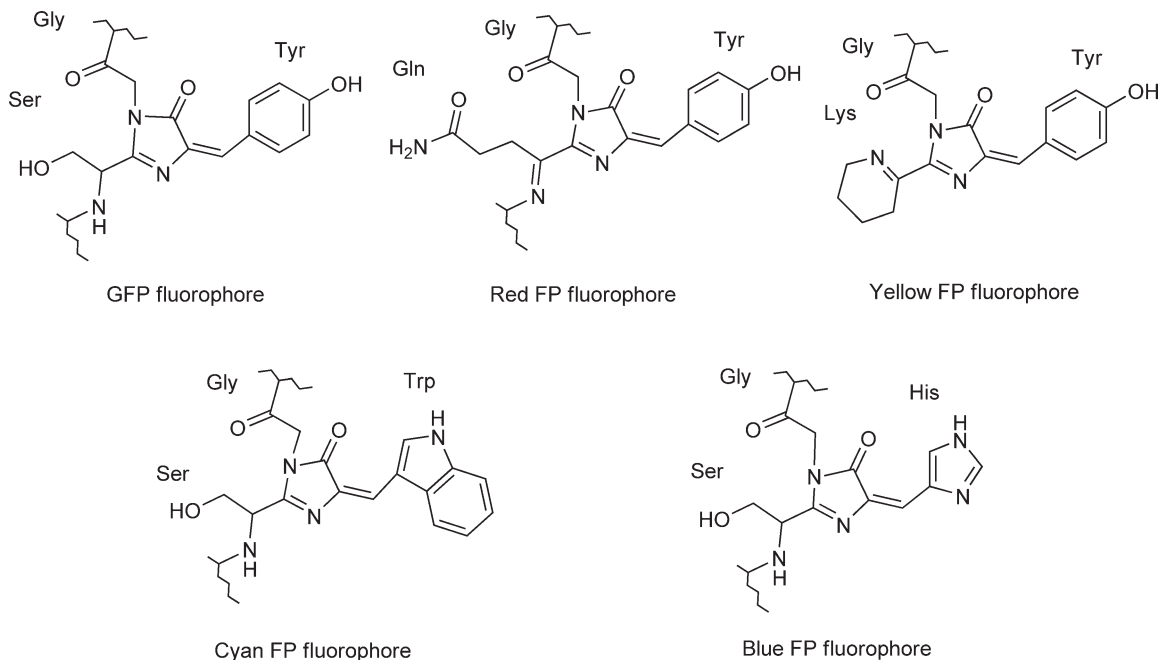


**Figure 13** Structure of the protein-derived fluorophore in green fluorescent protein from *Aequorea victoria*. In this protein modification of Ser65, Tyr66, and Gly67 yields the GFP fluorophore, which is displayed as sticks colored gray for carbon, red for oxygen, and blue for nitrogen. The coordinates from PDB entry 1ema were used to display this structure.



**Scheme 16** Mechanism for biosynthesis of the protein-derived fluorophore of GFP.

with a glutamine instead of a serine as the first amino acid residue in the tripeptide sequence (Figure 14). Another distinction is that the DsRed fluorophore is further modified through the oxidation of a second backbone bond to extend the conjugated system.<sup>114</sup> The yellow fluorescent protein, ZsYellow (also referred to as zFP538), was discovered in the Anthozoan button polyp *Zoanthus*. The ZsYellow fluorescent protein



**Figure 14** Protein-derived fluorescent protein fluorophores. The amino acid residues from which each fluorophore is derived are indicated above that portion of the fluorophore.

fluorophore features a novel three-ring system and peptide backbone cleavage which is a consequence of the substitution of lysine for serine as the first amino acid residue in the fluorophore tripeptide sequence (Figure 14).<sup>115</sup> Furthermore, an additional double bond is present in the heterocyclic ring, which results in a greater extent of electron delocalization during excitation when compared to GFP. This accounts for the altered emission wavelengths observed in ZsYellow fluorescent protein relative to GFP.

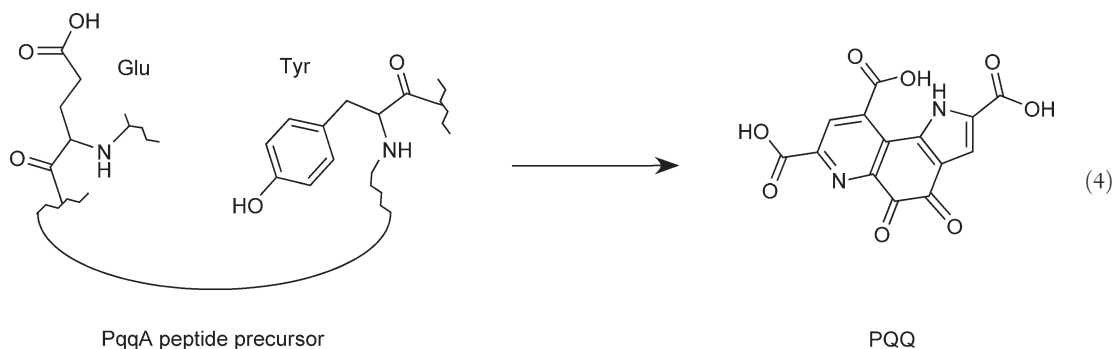
For biotechnological applications it has been possible to generate fluorescent proteins with different properties by site-directed mutagenesis of GFP (Figure 14). Blue fluorescent proteins have been designed, which contain as the critical feature the substitution of histidine for tyrosine as the middle amino acid residue in the chromophore tripeptide sequence.<sup>116</sup> The critical mutation in cyan fluorescent proteins is the substitution of tryptophan for tyrosine as the middle amino acid residue in the tripeptide sequence, which causes the chromophore to form with an indole rather than phenol component.<sup>117</sup>

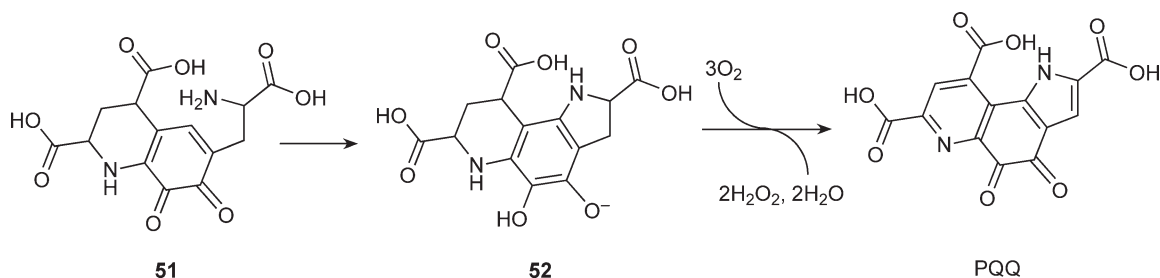
### 7.19.7 Pyrroloquinoline Quinone, a Protein-Derived Exogenous Cofactor

PQQ (4,5-dihydro-4,5-dioxo-1-*H*-pyrrolo[2,3-*f*]quinoline-2,7,9-tricarboxylic acid), which was originally called methoxatin<sup>118</sup> and is now commonly called PQQ<sup>56</sup> is an endogenous cofactor. It is included in this review because in contrast to typical cofactors and vitamins PQQ is not synthesized from free metabolic intermediates but from a gene-encoded precursor peptide. A wide range of bacterial enzymes has been shown to possess tightly but noncovalently bound PQQ as a cofactor. All are pyridine nucleotide-independent dehydrogenases. Glucose and methanol dehydrogenases from several diverse bacterial species were shown to possess PQQ.<sup>119</sup> Several other PQQ-dependent enzymes have been characterized, which use as substrates a variety of alcohols, sugars, and aldehydes.<sup>83</sup> The importance of PQQ to human metabolism and nutrition is a topic of controversy. PQQ has invoked considerable interest because of its presence in foods, its antioxidant properties and its role as a growth-promoting factor.<sup>120</sup> Oral supplementation of PQQ in nanomolar amounts has been shown to increase the responsiveness of B- and T-cells to mitogens and to improve neurologic function and reproductive outcome in rodents. It has been recently reported<sup>121</sup> that PQQ is a vitamin on the basis of analysis and characterization of a mouse homologue of the yeast gene, which encodes  $\alpha$ -amino adipic acid reductase, an enzyme that is crucial for the degradation of the amino acid lysine in mice. The conclusion that  $\alpha$ -amino adipic acid reductase is a PQQ-dependent enzyme has been challenged.<sup>122,123</sup> However, it has been subsequently shown that PQQ nutritional status does influence lysine metabolism and modulates mitochondrial DNA content in the mouse and rat.<sup>122</sup> As such, the role of PQQ in mammals is still not clear.

#### 7.19.7.1 PQQ Biosynthesis

The source of the atoms which comprise the PQQ cofactors was determined from isotopic labeling studies in which cell cultures were fed tyrosine that was labeled at specific positions with <sup>13</sup>C and <sup>15</sup>N. NMR analysis of the PQQ isolated from those cultures indicated that PQQ was formed from a tyrosine and a glutamate, with all carbon and nitrogen atoms of the two amino acids incorporated into PQQ.<sup>124</sup> Remarkably, it was subsequently shown that the amino acid precursors for PQQ are not the free amino acids but tyrosine and glutamate residues of a precursor peptide from which PQQ is synthesized (Equation (4)).





**Scheme 17** Intermediates in PQQ biosynthesis.

Genetic studies have identified an operon designated *pqq*, which contains the genes that are required for PQQ biosynthesis.<sup>125</sup> In the operons from *Klebsiella pneumoniae* and *Methylobacterium extorquens* these have been designated *pqqA–F*. The gene *pqqA* encodes a peptide comprised of 23–39 amino acids depending upon the organism. These polypeptides contain a conserved tyrosine and a glutamate and site-directed mutagenesis studies have confirmed that each of these two residues is required for PQQ biosynthesis.<sup>126</sup>

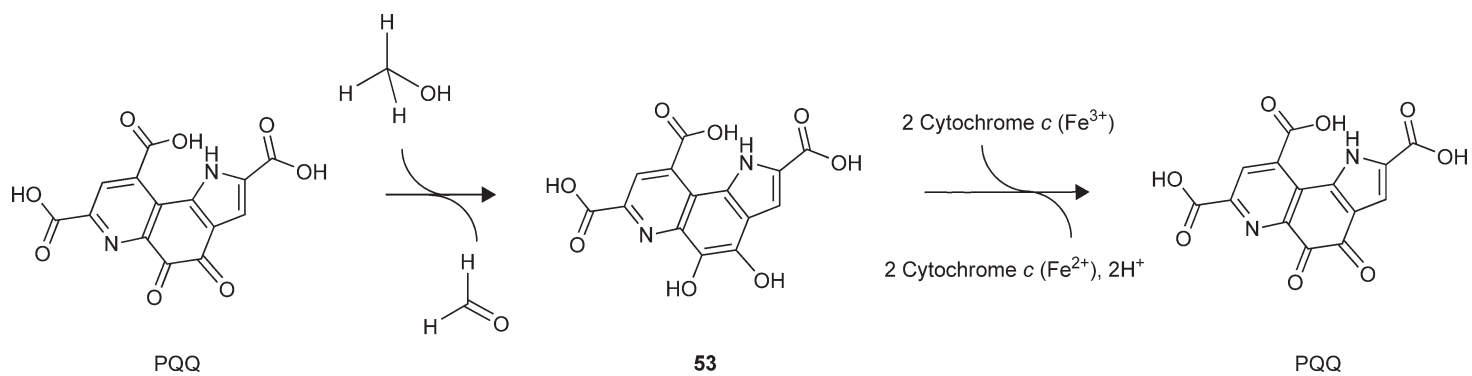
Little is known about the chemical reactions that lead to PQQ biosynthesis. However, it has been possible to study the final steps of PQQ biosynthesis *in vitro* (Scheme 17). This was made possible by the isolation of an intermediate in this process from a *pqqC* mutant strain of *M. extorquens*. That intermediate was characterized by NMR spectroscopy and its structure was determined to be 3a-(2-amino-2-carboxyethyl)-4,5-dioxo-4,5,6,7,8,9-hexahydroquinoline-7,9-dicarboxylic acid (**51**).<sup>127</sup> Incubation of **51** with PqqC, the protein gene product of *pqqC*, results in the formation of PQQ.<sup>128</sup> The crystal structures of PqqC and its complex with PQQ have been described, and the stoichiometry of H<sub>2</sub>O<sub>2</sub> formation and O<sub>2</sub> uptake during the PqqC-dependent PQQ biosynthesis reaction have been determined.<sup>128</sup> A simplified scheme for the PqqC-dependent biosynthesis of PQQ from **51** is shown in Scheme 17. It was proposed<sup>128</sup> that the initial step in the conversion of **51** into PQQ is a ring cyclization to complete formation of the tricyclic ring system to yield intermediate **52**. The cyclization occurs by a 1,4-addition of the primary amine to the conjugated ring structure. Conversion of **52** into PQQ then requires removal of eight electrons and eight protons from the intermediate. Three equivalents of O<sub>2</sub> are consumed and two equivalents of H<sub>2</sub>O<sub>2</sub> are produced per mole of PQQ formed respectively (Scheme 17). The stoichiometry requires that one equivalent of O<sub>2</sub> be reduced by four electrons to form two equivalents of water with the remaining two equivalents of O<sub>2</sub> each undergoing a two-electron reduction to H<sub>2</sub>O<sub>2</sub>.

The process by which PqqA, the initially synthesized precursor peptide, is converted into **51** is not known. Sequencing of the other genes in the *pqq* operon, analysis of these sequences, and genetic studies have suggested possible roles for some of the other genes.<sup>125</sup> The *pqqF* gene is believed to encode a peptidase which would clearly be necessary to cleave an intermediate formed after tyrosine and glutamate condense on the precursor peptide. The *pqqB* gene is thought to encode a transport protein necessary to move PQQ to its site of action in the periplasmic space after synthesis in the cytoplasm. The *pqqE* gene exhibits a sequence motif that suggests that it belongs to the radical *S*-adenosylmethionine superfamily,<sup>86</sup> although no specific function in PQQ biosynthesis has yet been inferred from this.

### 7.19.7.2 PQQ-Dependent Catalysis

PQQ-dependent enzymes are dehydrogenases that catalyze the oxidations of a variety of alcohols and sugars.<sup>125,129</sup> A feature of PQQ-dependent dehydrogenases that distinguishes them is that all are pyridine nucleotide-independent dehydrogenases. They do not use NAD<sup>+</sup> or NADP<sup>+</sup> as electron acceptors. Instead, in their oxidative half-reactions the soluble PQQ-dependent dehydrogenases donate electrons to other soluble redox proteins and the membrane-bound PQQ-dependent dehydrogenases donate electrons to ubiquinone.

The mechanism of the PQQ-dependent oxidation of methanol to formaldehyde which is catalyzed by methanol dehydrogenase is shown in Scheme 18.<sup>56</sup> In the reductive half-reaction methanol is converted into formaldehyde with concomitant reduction of PQQ to the quinol form (**53**). The exact mechanism of the reductive half-reaction has been controversial and two alternative mechanisms have been proposed. In the



**Scheme 18** The reaction mechanism of PQQ-dependent methanol dehydrogenase.

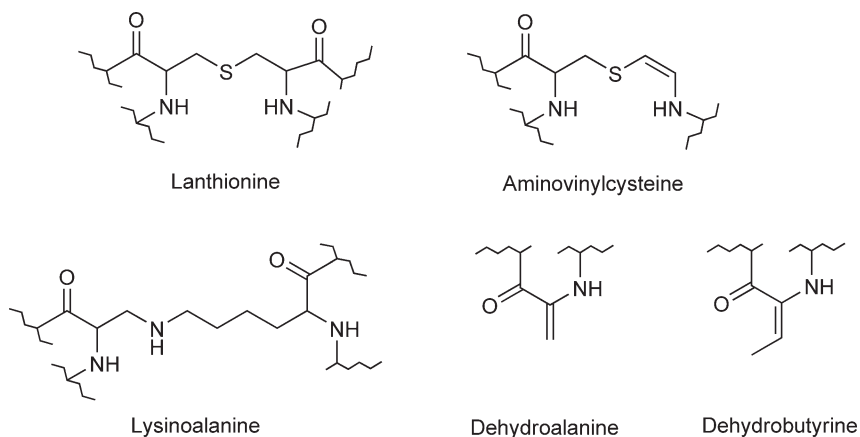
hemiketal mechanism an active-site base abstracts a proton from the methanol substrate to yield a methoxy anion that initiates a nucleophilic attack of the quinone carbon of PQQ to form a covalent adduct. In the alternative hydride transfer mechanism the active-site base again initiates the reaction by abstracting the alcoholic proton from the substrate. However, in this mechanism there is no covalent adduct formed between the substrate and PQQ. Instead, the substrate-derived hydride ion attacks the quinone carbon, resulting in formation of the aldehyde and reduction of PQQ. Structural and computational studies of methanol dehydrogenase support the hydride transfer mechanism.<sup>130</sup> PQQ is a two-electron carrier and the electron acceptor for methanol dehydrogenase is a  $\epsilon$ -type cytochrome,<sup>131</sup> which is a one-electron carrier. Thus, the oxidative half-reaction occurs in two one-electron transfers to two cytochromes through a PQQ semiquinone intermediate.

The mechanism of action of PQQ-dependent sugar dehydrogenases is believed to be similar to that of the alcohol dehydrogenases. One of the most studied of these is glucose dehydrogenase,<sup>132</sup> which converts glucose into D-glucono-1,5-lactone in the reductive half-reaction. On the basis of the structure of the active site of this enzyme it has been proposed that the reaction most probably proceeds through the hydride transfer mechanism, as discussed above for methanol dehydrogenase. This glucose dehydrogenase is membrane associated and in the oxidative half-reaction reduces the membrane-bound ubiquinone to ubiquinol.

### 7.19.8 Lantibiotics, Protein-Derived Antibiotic Peptides

Although lantibiotics are not enzymes and do not contain cofactors they are included here and briefly discussed because they are proteins whose biosynthesis requires extensive and unusual posttranslational modifications. These modifications endow the gene product with a new function, in this case an antimicrobial activity, which would not exist in the absence of these covalent modifications to the protein. They are found in peptides of microbial origin.<sup>133</sup> Their name derives from lanthionine, which is a nonproteinogenic amino acid composed of two alanine residues that are crosslinked on their  $\beta$ -carbon atoms by a thioether linkage (Figure 15). The wide array of fascinating posttranslational modifications which are present in this class of peptide antibiotics<sup>18</sup> is reminiscent of those seen in some protein-derived cofactors.

Lantibiotics are produced by a large number of Gram-positive bacteria such as *Streptococcus* and *Streptomyces* to attack other Gram-positive bacteria. They are initially synthesized as gene-encoded precursor peptides and undergo posttranslational modification to generate the mature peptide. The structural gene for the pre-peptide and the genes involved in biosynthesis, processing, export of the peptide are organized in clusters.<sup>134</sup> Lantibiotics are synthesized on the ribosome as the pre-peptide. The posttranslational modification events associated with lantibiotic biosynthesis include dehydration of specific serines and threonines to form dehydroamino acids, addition of neighboring sulfhydryl groups to form thioethers, formation of lysinoalanine



**Figure 15** Examples of structures of posttranslationally modified amino acids found in lantibiotics.



bridges, formation of novel N-terminal blocking groups, and oxidative decarboxylation of a C-terminal cysteine. Examples of these modifications are shown in **Figure 15**. These types of modifications are similar to those seen in some of the other protein-derived species discussed earlier such as MIO and the GFP fluorophore. However, in contrast to the self-processing mechanisms described earlier for those cofactors, the biosynthesis of the lantibiotics from their precursor peptides is an enzyme-mediated process.<sup>135</sup> In this sense the biosynthesis of lantibiotics is reminiscent of the process of PQQ biosynthesis.

To facilitate the comparison of the gene clusters which are responsible for lantibiotic synthesis from different bacteria they have been assigned a uniform nomenclature in the generic *lan* grouping. LanA is the precursor peptide. LanB, LanC, and LanM are modifying enzymes involved in the dehydration and cyclization reactions. LanP and LanT are processing enzymes that remove the N-terminal leader sequence from the precursor peptide.<sup>136</sup> Several other genes are also present in these clusters which are involved in regulation, transport, and immunity.

Gram-negative bacteria also produce a number of posttranslationally modified peptides with antimicrobial activity, termed microcins, which may be a counterpart to the lantibiotics. Unlike lantibiotics, the post-translational modifications of microcins do not only affect the side chains of the respecting amino acids, but also include the peptide backbone, resulting in the formation of heterocyclic oxazoline and thiazoline rings.<sup>137</sup>

## 7.19.9 Conclusions and Perspectives

The identification of protein-derived cofactors and characterization of the mechanisms of their biogenesis require a re-evaluation of our current ideas about protein evolution and protein structure–function relationships. In some cases, protein-derived cofactor biosynthesis requires large gene clusters with multiple accessory genes and enzyme catalyzed posttranslational modifications. In other cases, the biosynthesis of the protein-derived cofactor is a self-processing event. In some cases, the autocatalytic modification requires the presence of a metal, which remains in the active site and then participates in the catalytic reaction of the mature enzyme. In other cases, no metal is involved but the autocatalytic event is facilitated by the folding of the precursor protein, which orients the reacting groups to allow the biosynthetic reactions to occur. In some cases, a pro-enzyme with a structure different from that of the mature enzyme is required for the biosynthesis of the protein-derived cofactor. If one considers the time-line for the evolution of these proteins it seems that evolution initially directed the protein to allow the formation of the protein-derived cofactor. Only then would the modified protein undergo subsequent evolution to acquire particular functions. It follows that for some proteins such as these; highly conserved elements of structure may not necessarily be related to function but instead to cofactor biogenesis. The characterization of protein-derived cofactors also has implications for protein engineering. A better understanding of the mechanisms of biogenesis of protein-derived cofactors will provide insight into how to design or modify the sequence of existing proteins so that they form protein-derived cofactors that introduce a catalytic or redox center within a protein without the need for addition of exogenous cofactors. The characterization of the mechanisms of biosynthesis and catalysis by protein-derived cofactors is also of fundamental importance. This field has elucidated previously unrecognized natural strategies for catalysis.

### Abbreviations

<b>CTQ</b>	cysteine tryptophylquinone
<b>GFP</b>	green fluorescent protein
<b>KatG</b>	bifunctional catalase–peroxidase
<b>LTQ</b>	lysine tyrosylquinone
<b>MIO</b>	methylidene-imidazolone
<b>ORF2</b>	open reading frame
<b>PQQ</b>	pyrroloquinoline quinone

<b>TPQ</b>	2,4,5-trihydroxyphenylalanine quinone
<b>TTQ</b>	tryptophan tryptophylquinone
<b>VAP-1</b>	vascular adhesion protein-1

## References

1. V. L. Davidson, *Biochemistry* **2007**, *46*, 5283–5292.
2. J. P. Klinman; J. E. Dove, Eds., *Adv. Protein Chem.* **2001**, *58*.
3. C. T. Walsh; S. Garneau-Tsodikova; G. J. Gatto, Jr., *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 7342–7372.
4. A. C. Eliot; J. F. Kirsch, *Annu. Rev. Biochem.* **2004**, *73*, 383–415.
5. P. D. van Poelje; E. E. Snell, *Annu. Rev. Biochem.* **1990**, *59*, 29–59.
6. A. Albert; V. Dhanaraj; U. Genschel; G. Khan; M. K. Ramjee; R. Pulido; B. L. Sibanda; F. von Delft; M. Witty; T. L. Blundell; A. G. Smith; C. Abell, *Nat. Struct. Biol.* **1998**, *5*, 289–293.
7. W. D. Tolbert; Y. Zhang; S. E. Cottet; E. M. Bennett; J. L. Ekstrom; A. E. Pegg; S. E. Ealick, *Biochemistry* **2003**, *42*, 2386–2395.
8. Q. X. Li; W. Dowhan, *J. Biol. Chem.* **1988**, *263*, 11516–11522.
9. B. A. Stanley; A. E. Pegg; I. Holm, *J. Biol. Chem.* **1989**, *264*, 21073–21079.
10. T. Gallagher; D. A. Rozwarski; S. R. Ernst; M. L. Hackert, *J. Mol. Biol.* **1993**, *230*, 516–528.
11. M. Wagner; D. Sonntag; R. Grimm; A. Pich; C. Eckerskorn; B. Sohling; J. R. Andreesen, *Eur. J. Biochem.* **1999**, *260*, 38–49.
12. U. C. Kabisch; A. Grantzdorffer; A. Schierhorn; K. P. Rucknagel; J. R. Andreesen; A. Pich, *J. Biol. Chem.* **1999**, *274*, 8445–8454.
13. J. L. Ekstrom; W. D. Tolbert; H. Xiong; A. E. Pegg; S. E. Ealick, *Biochemistry* **2001**, *40*, 9495–9504.
14. H. Xiong; A. E. Pegg, *J. Biol. Chem.* **1999**, *274*, 35059–35066.
15. U. C. Kabisch; A. Grantzdorffer; A. Schierhorn; K. P. Rucknagel; J. R. Andreesen; A. Pich, *J. Biol. Chem.* **1999**, *274*, 8445–8454.
16. J. R. Andreesen, *Curr. Opin. Chem. Biol.* **2004**, *8*, 454–461.
17. M. Friedman, *J. Agric. Food Chem.* **1999**, *47*, 1295–1319.
18. L. Xie; W. A. van der Donk, *Curr. Opin. Chem. Biol.* **2004**, *8*, 498–507.
19. T. F. Schwede; J. Retey; G. E. Schulz, *Biochemistry* **1999**, *38*, 5355–5361.
20. D. Rother; D. Merkel; J. Retey, *Angew. Chem.* **2000**, *39*, 2462–2464.
21. D. Rother; L. Poppe; G. Morlock; S. Viergutz; J. Retey, *Eur. J. Biochem.* **2002**, *269*, 3065–3075.
22. C. V. Christianson; T. J. Montavon; S. G. VanLanen; B. Shen; S. D. Bruner, *Biochemistry* **2007**, *46*, 7205–7214.
23. K. D. Walker; K. Klettke; T. Akiyama; R. Croteau, *J. Biol. Chem.* **2004**, *279*, 53947–53954.
24. M. Jin; M. A. Fischbach; J. Clardy, *J. Am. Chem. Soc.* **2006**, *128*, 10660–10661.
25. J. Retey, *Biochim. Biophys. Acta* **2003**, *1647*, 179–184.
26. L. Poppe, *Curr. Opin. Chem. Biol.* **2001**, *5*, 512–524.
27. M. Langer; A. Lieber; J. Retey, *Biochemistry* **1994**, *33*, 14034–14038.
28. B. Schuster; J. Retey, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8433–8437.
29. B. Langer; M. Langer; J. Retey, *Adv. Protein Chem.* **2001**, *58*, 175–214.
30. W. S. McIntire; D. E. Wemmer; A. Chistoserdov; M. E. Lidstrom, *Science* **1991**, *252*, 817–824.
31. J. P. Klinman; D. Mu, *Annu. Rev. Biochem.* **1994**, *63*, 299–344.
32. M. Mure, *Acc. Chem. Res.* **2004**, *37*, 131–139.
33. V. L. Davidson, *Bioorg. Chem.* **2005**, *33*, 159–170.
34. S. M. Janes; D. Mu; D. Wemmer; A. J. Smith; S. Kaur; D. Maltby; A. L. Burlingame; J. P. Klinman, *Science* **1990**, *248*, 981–987.
35. A. Cona; G. Rea; R. Angelini; R. Federico; P. Tavladoraki, *Trends Plant Sci.* **2006**, *11*, 80–88.
36. Q. Zhang; Y. Mashima; S. Noda; Y. Imamura; J. Kudoh; N. Shimizu; T. Nishiyama; S. Umeda; Y. Oguchi; Y. Tanaka; T. Iwata, *Gene* **2003**, *318*, 45–53.
37. P. H. Yu; S. Wright; E. H. Fan; Z.-R. Lun; D. Gubisne-Harberle, *Biochim. Biophys. Acta* **2003**, *1647*, 193–199.
38. C. M. Stolen; R. Madanat; L. Marti; S. Kari; G. G. Yegutkin; H. Sariola; A. Zorzano; S. Jalkanen, *FASEB J.* **2004**, *18*, 702–704.
39. T. T. Airene; Y. Nymalm; H. Kidron; D. J. Smith; M. Pihlavisto; M. Salmi; S. Jalkanen; M. S. Johnson; T. A. Salminen, *Protein Sci.* **2005**, *14*, 1964–1974.
40. R. Matsuzaki; T. Fukui; H. Sato; Y. Ozaki; K. Tanizawa, *FEBS Lett.* **1994**, *351*, 360–364.
41. D. Cai; J. P. Klinman, *J. Biol. Chem.* **1994**, *269*, 32039–32042.
42. M. Kim; T. Okajima; S. Kishishita; M. Yoshimura; A. Kawamori; K. Tanizawa; H. Yamaguchi, *Nat. Struct. Biol.* **2002**, *9*, 591–596.
43. B. J. Brazeau; B. J. Johnson; C. M. Wilmot, *Arch. Biochem. Biophys.* **2004**, *428*, 22–31.
44. M. C. Wilce; D. M. Dooley; H. C. Freeman; J. M. Guss; H. Matsunami; W. S. McIntire; C. E. Ruggiero; K. Tanizawa; H. Yamaguchi, *Biochemistry* **1997**, *36*, 16116–16133.
45. J. E. Dove; J. P. Klinman, *Adv. Protein Chem.* **2001**, *58*, 141–174.
46. Z. Chen; B. Schwartz; N. K. Williams; R. Li; J. P. Klinman; F. S. Mathews, *Biochemistry* **2000**, *39*, 9709–9717.
47. N. M. Samuels; J. P. Klinman, *Biochemistry* **2005**, *44*, 14308–14317.
48. T. Okajima; S. Kishishita; Y.-C. Chiu; T. Murakawa; M. Kim; H. Yamaguchi; S. Hirota; S. Kuroda; K. Tanizawa, *Biochemistry* **2005**, *44*, 12041–12048.
49. M. Mure; S. A. Mills; J. P. Klinman, *Biochemistry* **2002**, *41*, 9269–9278.
50. J. M. Murray; C. G. Saysell; C. M. Wilmot; W. S. Tambyrajah; J. Jaeger; P. F. Knowles; S. E. Phillips; M. J. McPherson, *Biochemistry* **1999**, *38*, 8217–8227.
51. J. L. DuBois; J. P. Klinman, *Biochemistry* **2006**, *45*, 3178–3188.

52. D. M. Dooley; M. A. McGuire; D. E. Brown; P. N. Turowski; W. S. McIntire; P. F. Knowles, *Nature* **1991**, *349*, 262–264.
53. H. M. Kagan; W. Li, *J. Cell. Biochem.* **2003**, *88*, 660–672.
54. S. X. Wang; M. Mure; K. F. Medzihradzsky; A. L. Burlingame; D. E. Brown; D. M. Dooley; A. J. Smith; H. M. Kagan; J. P. Klinman, *Science* **1996**, *273*, 1078–1084.
55. P. R. Williamson; H. M. Kagan, *J. Biol. Chem.* **1987**, *262*, 8196–8201.
56. V. L. Davidson, *Adv. Protein Chem.* **2001**, *58*, 95–140.
57. V. L. Davidson, *Arch. Biochem. Biophys.* **2004**, *428*, 32–40.
58. C. J. van der Palen; W. N. Reijnders; S. de Vries; J. A. Duine; R. J. van Spanning, *Antonie Van Leeuwenhoek* **1997**, *72*, 219–228.
59. C. J. van der Palen; D. J. Slotboom; L. Jongejan; W. N. Reijnders; N. Harms; J. A. Duine; R. J. van Spanning, *Eur. J. Biochem.* **1995**, *230*, 860–871.
60. R. J. van Spanning; C. W. Wansell; W. N. Reijnders; L. F. Oltmann; A. H. Stouthamer, *FEBS Lett.* **1990**, *275*, 217–220.
61. L. Chen; M. Doi; R. C. Durley; A. Y. Chistoserdov; M. E. Lidstrom; V. L. Davidson; F. S. Mathews, *J. Mol. Biol.* **1998**, *276*, 131–149.
62. M. Husain; V. L. Davidson, *J. Biol. Chem.* **1985**, *260*, 14626–14629.
63. R. J. Van Spanning; C. J. van der Palen; D. J. Slotboom; W. N. Reijnders; A. H. Stouthamer; J. A. Duine, *Eur. J. Biochem.* **1994**, *226*, 201–210.
64. Y. Wang; M. E. Graichen; A. Liu; A. R. Pearson; C. W. Wilmot; V. L. Davidson, *Biochemistry* **2003**, *42*, 7318–7325.
65. X. Li; L. H. Jones; A. R. Pearson; C. M. Wilmot; V. L. Davidson, *Biochemistry* **2006**, *45*, 13276–13283.
66. M. E. Graichen; L. H. Jones; B. V. Sharma; R. J. van Spanning; J. P. Hosler; V. L. Davidson, *J. Bacteriol.* **1999**, *181*, 4216–4222.
67. A. R. Pearson; T. de la Mora-Rey; M. E. Graichen; Y. Wang; L. H. Jones; S. Marimanikkupam; S. A. Aggar; P. A. Grimsrud; V. L. Davidson; C. W. Wilmot, *Biochemistry* **2004**, *43*, 5494–5502.
68. A. R. Pearson; S. Marimanikkupam; X. Li; V. L. Davidson; C. M. Wilmot, *J. Am. Chem. Soc.* **2006**, *128*, 12416–12417.
69. Y. Wang; X. Li; L. H. Jones; A. R. Pearson; C. M. Wilmot; V. L. Davidson, *J. Am. Chem. Soc.* **2005**, *127*, 8258–8259.
70. Y. L. Hyun; V. L. Davidson, *Biochemistry* **1995**, *34*, 12249–12254.
71. V. L. Davidson, *Acc. Chem. Res.* **2000**, *33*, 87–93.
72. L. Chen; R. Durley; B. J. Poliks; K. Hamada; Z. Chen; F. S. Mathews; V. L. Davidson; Y. Satow; E. Huizinga; F. M. Vellieux, *Biochemistry* **1992**, *31*, 4959–4964.
73. L. Chen; R. C. Durley; F. S. Mathews; V. L. Davidson, *Science* **1994**, *264*, 86–90.
74. N. Sukumar; Z. W. Chen; D. Ferrari; A. Merli; G. L. Rossi; H. D. Bellamy; A. Chistoserdov; V. L. Davidson; F. S. Mathews, *Biochemistry* **2006**, *45*, 13500–13510.
75. E. G. Huizinga; B. A. van Zanten; J. A. Duine; J. A. Jongejan; F. Huitema; K. S. Wilson; W. G. Hol, *Biochemistry* **1992**, *31*, 9789–9795.
76. K. Warncke; H. B. Brooks; G. T. Babcock; V. L. Davidson; J. McCracken, *J. Am. Chem. Soc.* **1993**, *115*, 6464–6465.
77. G. R. Bishop; H. B. Brooks; V. L. Davidson, *Biochemistry* **1996**, *35*, 8948–8954.
78. Z. Zhu; V. L. Davidson, *Biochemistry* **1999**, *38*, 4862–4867.
79. Z. Zhu; V. L. Davidson, *J. Biol. Chem.* **1998**, *273*, 14254–14260.
80. S. Datta; Y. Mori; K. Takagi; K. Kawaguchi; Z. W. Chen; T. Okajima; S. Kuroda; T. Ikeda; K. Kano; K. Tanizawa; F. S. Mathews, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14268–14273.
81. A. Satoh; J. K. Kim; I. Miyahara; B. Devreese; I. Vandenberghe; A. Hacısalihoglu; T. Okajima; S. Kuroda; O. Adachi; J. A. Duine; J. Van Beeumen; K. Tanizawa; K. Hirotsu, *J. Biol. Chem.* **2002**, *277*, 2830–2834.
82. K. Takagi; M. Torimura; K. Kawaguchi; K. Kano; T. Ikeda, *Biochemistry* **1999**, *38*, 6935–6942.
83. O. Adachi; T. Kubota; A. Hacısalihoglu; H. Toyama; E. Shinagawa; J. A. Duine; K. Matsushita, *Biosci. Biotechnol. Biochem.* **1998**, *62*, 469–478.
84. K. Takagi; K. Yamamoto; K. Kano; T. Ikeda, *Eur. J. Biochem.* **2001**, *268*, 470–476.
85. L. Masgrau; A. Roujeinikova; L. O. Johannissen; P. Hothi; J. Basran; K. E. Ranaghan; A. J. Mulholland; M. J. Sutcliffe; N. S. Scrutton; D. Leys, *Science* **2006**, *312*, 237–241.
86. H. J. Sofia; G. Chen; B. G. Hetzler; J. F. Reyes-Spindola; N. E. Miller, *Nucleic Acids Res.* **2001**, *29*, 1097–1106.
87. K. Ono; T. Okajima; M. Tani; S. Kuroda; D. Sun; V. L. Davidson; K. Tanizawa, *J. Biol. Chem.* **2006**, *281*, 13672–13684.
88. D. Sun; K. Ono; T. Okajima; K. Tanizawa; M. Uchida; Y. Yamamoto; F. S. Mathews; V. L. Davidson, *Biochemistry* **2003**, *42*, 10896–10903.
89. N. Ito; S. E. Phillips; C. Stevens; Z. B. Ogel; M. J. McPherson; J. N. Keen; K. D. Yadav; P. F. Knowles, *Nature* **1991**, *350*, 87–90.
90. M. M. Whittaker; P. J. Kersten; D. Cullen; J. W. Whittaker, *J. Biol. Chem.* **1999**, *274*, 36226–36232.
91. C. R. Simmons; Q. Liu; Q. Huang; Q. Hao; T. P. Begley; P. A. Karplus; M. H. Stipanuk, *J. Biol. Chem.* **2006**, *281*, 18723–18733.
92. M. S. Rogers; A. J. Baron; M. J. McPherson; P. F. Knowles; D. M. Dooley, *J. Am. Chem. Soc.* **2000**, *122*, 990–991.
93. S. J. Firbank; M. S. Rogers; C. M. Wilmot; D. M. Dooley; M. A. Halcrow; P. F. Knowles; M. J. McPherson; S. E. Phillips, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12932–12937.
94. M. M. Whittaker; J. W. Whittaker, *J. Biol. Chem.* **2003**, *278*, 22090–22101.
95. J. W. Whittaker, *Arch. Biochem. Biophys.* **2005**, *433*, 227–239.
96. J. Bravo; I. Fita; J. C. Ferrer; W. Ens; A. Hillar; J. Switala; P. C. Loewen, *Protein Sci.* **1997**, *6*, 1016–1023.
97. A. Diaz; E. Horjales; E. Rudino-Pinera; R. Arreola; W. Hansberg, *J. Mol. Biol.* **2004**, *342*, 971–985.
98. Y. Yamada; T. Fujiwara; T. Sato; N. Igarashi; N. Tanaka, *Nat. Struct. Biol.* **2002**, *9*, 691–695.
99. X. Carpena; S. Loprasert; S. Mongkolsuk; J. Switala; P. C. Loewen; I. Fita, *J. Mol. Biol.* **2003**, *327*, 475–489.
100. T. Bertrand; N. A. Eady; J. N. Jones; Jesmin; J. M. Nagy; B. Jamart-Gregoire; E. L. Raven; K. A. Brown, *J. Biol. Chem.* **2004**, *279*, 38991–38999.
101. R. A. Ghiladi; G. M. Knudsen; K. F. Medzihradzsky; P. R. de Montellano, *J. Biol. Chem.* **2005**, *280*, 22651–22663.
102. G. Smulevich; C. Jakopitsch; E. Droghetti; C. Obinger, *J. Inorg. Biochem.* **2006**, *100*, 568–585.
103. T. Klabunde; C. Eicken; J. C. Sacchettini; B. Krebs, *Nat. Struct. Biol.* **1998**, *5*, 1084–1090.
104. M. E. Cuff; K. I. Miller; K. E. van Holde; W. A. Hendrickson, *J. Mol. Biol.* **1998**, *278*, 855–870.
105. C. Ostermeier; A. Harrenga; U. Ermler; H. Michel, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10547–10553.

106. S. Yoshikawa; K. Shinzawa-Itoh; R. Nakashima; R. Yaono; E. Yamashita; N. Inoue; M. Yao; M. J. Fei; C. P. Libeu; T. Mizushima; H. Yamaguchi; T. Tomizaki; T. Tsukihara, *Science* **1998**, *280*, 1723–1729.
107. M. M. Pereira; M. Santana; M. Teixeira, *Biochim. Biophys. Acta* **2001**, *1505*, 185–208.
108. J. Hemp; D. E. Robinson; K. B. Ganesan; T. J. Martinez; N. L. Kelleher; R. B. Gennis, *Biochemistry* **2006**, *45*, 15405–15410.
109. M. Ormo; A. B. Cubitt; K. Kallio; L. A. Gross; R. Y. Tsien; S. J. Remington, *Science* **1996**, *273*, 1392–1395.
110. R. Heim; D. C. Prasher; R. Y. Tsien, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501–12504.
111. Y. A. Labas; N. G. Gurskaya; Y. G. Yanushevich; A. F. Fradkov; K. A. Lukyanov; S. A. Lukyanov; M. V. Matz, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4256–4261.
112. S. G. Dove; O. Hoegh-Guldberg; S. Ranganathan, *Coral Reefs* **2001**, *19*, 197–204.
113. K. A. Lukyanov; A. F. Fradkov; N. G. Gurskaya; M. V. Matz; Y. A. Labas; A. P. Savitsky; M. L. Markelov; A. G. Zaraisky; X. Zhao; Y. Fang; W. Tan; S. A. Lukyanov, *J. Biol. Chem.* **2000**, *275*, 25879–25882.
114. M. A. Wall; M. Socolich; R. Ranganathan, *Nat. Struct. Biol.* **2000**, *7*, 1133–1138.
115. S. J. Remington; R. M. Wachter; D. K. Yarbrough; B. Branchaud; D. C. Anderson; K. Kallio; K. A. Lukyanov, *Biochemistry* **2005**, *44*, 202–212.
116. H. W. Ai; N. C. Shaner; Z. Cheng; R. Y. Tsien; R. E. Campbell, *Biochemistry* **2007**, *46*, 5904–5910.
117. G. D. Malo; L. J. Pouwels; M. Wang; A. Weichsel; W. R. Montfort; M. A. Rizzo; D. W. Piston; R. M. Wachter, *Biochemistry* **2007**, *46*, 9865–9873.
118. H. S. Forrest; S. A. Salisbury; G. Sperl, *Biochim. Biophys. Acta* **1981**, *676*, 226–229.
119. C. Anthony, *Arch. Biochem. Biophys.* **2004**, *428*, 2–9.
120. T. E. Stites; A. E. Mitchell; R. B. Rucker, *J. Nutr.* **2000**, *130*, 719–727.
121. T. Kasahara; T. Kato, *Nature* **2003**, *422*, 832.
122. K. A. Bauerly; D. H. Storms; C. B. Harris; S. Hajizadeh; M. Y. Sun; C. P. Cheung; M. A. Satre; A. J. Fascetti; E. Tchapanian; R. B. Rucker, *Biochim. Biophys. Acta* **2006**, *1760*, 1741–1748.
123. L. M. Felton; C. Anthony, *Nature* **2005**, *433*, E10; discussion E11–E12.
124. C. J. Unkefer; D. R. Houck; B. M. Britt; T. R. Sosnick; J. L. Hanners, *Meth. Enzymol.* **1995**, *258*, 227–235.
125. P. M. Goodwin; C. Anthony, *Adv. Microb. Physiol.* **1998**, *40*, 1–80.
126. N. Goosen; R. G. Huinen; P. van de Putte, *J. Bacteriol.* **1992**, *174*, 1426–1427.
127. O. T. Magnusson; H. Toyama; M. Saeki; R. Schwarzenbacher; J. P. Klinman, *J. Am. Chem. Soc.* **2004**, *126*, 5342–5343.
128. O. T. Magnusson; H. Toyama; M. Saeki; A. Rojas; J. C. Reed; R. C. Liddington; J. P. Klinman; R. Schwarzenbacher, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7913–7918.
129. K. Matsushita; H. Toyama; M. Yamada; O. Adachi, *Appl. Microbiol. Biotechnol.* **2002**, *58*, 13–22.
130. Y. J. Zheng; Z. Xia; Z. Chen; F. S. Mathews; T. C. Bruice, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 432–434.
131. Z.-X. Xia; W.-W. Dai; Y.-N. He; S. A. White; F. S. Mathews; V. L. Davidson, *J. Biol. Inorg. Chem.* **2003**, *8*, 843–854.
132. A. Oubrie; H. J. Rozeboom; K. H. Kalk; A. J. Olsthoorn; J. A. Duine; B. W. Dijkstra, *EMBO J.* **1999**, *18*, 5187–5194.
133. O. McAuliffe; R. P. Ross; C. Hill, *FEMS Microbiol. Rev.* **2001**, *25*, 285–308.
134. W. M. de Vos; O. P. Kuipers; J. R. van der Meer; R. J. Siezen, *Mol. Microbiol.* **1995**, *17*, 427–437.
135. L. Xie; L. M. Miller; C. Chatterjee; O. Averin; N. L. Kelleher; W. A. van der Donk, *Science* **2004**, *303*, 679–681.
136. U. Pag; H. G. Sahl, *Curr. Pharm. Des.* **2002**, *8*, 815–833.
137. S. Duquesne; D. Destoumieux-Garzon; J. Peduzzi; S. Rebuffat, *Nat. Prod. Rep.* **2007**, *24*, 708–734.

### Biographical Sketch



Victor L. Davidson received his B.S. in Biochemistry from the University of Illinois in 1973 and a Ph.D. in Chemistry from the Texas Tech University in 1982. After postdoctoral training at Purdue University from 1982 to 1984 and a research position at the University of California at San Francisco from 1984 to 1988 he joined the faculty at the University of Mississippi Medical Center where he is currently Professor of Biochemistry.

## 7.20 Biosynthesis of the Methanogenic Coenzymes

Laura L. Grochowski and Robert H. White, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

© 2010 Elsevier Ltd. All rights reserved.

7.20.1	Introduction	711
7.20.2	Biosynthesis of Methanofuran	712
7.20.2.1	Introduction	712
7.20.2.2	Tyrosine Decarboxylase	712
7.20.2.3	Biosynthesis of <i>Meso</i> -1,3,4,6-Hexanetetracarboxylic Acid	714
7.20.2.4	Assembly of the Final Structure	716
7.20.3	Biosynthesis of Methanopterin	716
7.20.3.1	Introduction	716
7.20.3.2	Formation of the Pterin	718
7.20.3.3	Biosynthesis of 4-Aminobenzoic Acid	718
7.20.3.4	Formation of 4-( $\beta$ -D-Ribofuranosyl)Aminobenzene-5'-Phosphate	719
7.20.3.5	Biosynthesis of the Methanopterin Side Chain	722
7.20.3.6	Methylation of the Pterin	722
7.20.3.7	Reduction of Dihydromethanopterin	722
7.20.4	Biosynthesis of Coenzyme F <sub>420</sub> and Riboflavin Precursors	724
7.20.4.1	Introduction	724
7.20.4.2	Early Steps in F <sub>420</sub> and Riboflavin Biosynthetic Pathways Leading to 5-Amino-6-Ribitylamino-2,4(1H,3H)-Pyrimidinedione	724
7.20.4.3	Biosynthesis of 7,8-Didemethyl-8-Hydroxy-5-Deazariboflavin	725
7.20.4.4	Assembly of the Side Chains of the Core Structure of the F <sub>420</sub> Coenzyme F <sub>420-0</sub>	726
7.20.4.5	Assembly of the Core Structure of F <sub>420</sub> Coenzymes	728
7.20.4.6	Biosynthesis of the Side Chain	730
7.20.5	Biosynthesis of Riboflavin, FMN, and FAD	730
7.20.6	Biosynthesis of Coenzyme M and Coenzyme B	735
7.20.6.1	Introduction	735
7.20.6.2	Biosynthesis of Coenzyme M	735
7.20.6.3	Biosynthesis of Coenzyme B	737
7.20.6.4	Formation of the Coenzyme M and Coenzyme B Thiols	738
7.20.7	Biosynthesis of Corrinoids and Coenzyme F <sub>430</sub>	740
7.20.7.1	Introduction	740
7.20.7.2	Overview of Archaeal Corrinoid Biosynthesis	740
7.20.7.3	Formation of 1-Amino-2-Propanol	740
7.20.7.4	Attachment of 1-Amino-2-Propanol	742
7.20.7.5	Formation of the Benzimidazole Moiety	742
7.20.7.6	Corrinoid Upper Ligand	743
7.20.7.7	Biosynthesis of Coenzyme F <sub>430</sub>	743
7.20.8	Concluding Remarks	745
References		745

### 7.20.1 Introduction

This review on the biosynthesis of methanogenic coenzymes in the methanoarchaea will constitute an update of published and unpublished work on this topic since it was last reviewed in 2002.<sup>1</sup> For earlier reviews, see White<sup>2</sup> and White and Zhou.<sup>3</sup> Many excellent reviews have covered the involvement of these coenzymes in



methanogenesis.<sup>4,5</sup> However, since their initial discoveries, many of which occurred in the methanogens, each of these coenzymes has now been found to function in different capacities outside of the methanogenic archaea. The structure of each coenzyme as well as their involvement in methanogenesis is outlined in **Figure 1**. We refer to these coenzymes herein as the methanogenic coenzymes, because each functions as a catalyst in the biochemical reduction of CO<sub>2</sub>, acetate, or methylated amines to methane.<sup>6,7</sup> Recent genetic and biochemical data also indicate that many of these coenzymes also function in the anaerobic oxidation of methane,<sup>8–11</sup> where some variations of their structures have been observed, as seen with coenzyme F<sub>430</sub>.<sup>12</sup> The present status of work on the biosynthesis of each of these coenzymes are presented using the figures, which show the proposed biosynthetic pathways, as the focus of the discussion. Although FMN, FAD, and cobalamin analogues are also important coenzymes in methane metabolism,<sup>13,14</sup> they are generally not considered methanogenic coenzymes because they were discovered and their biosynthetic pathways established long before the biochemistry of methane formation was unraveled. It is clear, however, that each plays an essential role in methanogenesis. Their biosynthetic pathways are presented here due to the strong connection between the biosynthesis of riboflavin and F<sub>420</sub> and between the biosynthesis of cobalamin and coenzyme F<sub>430</sub>.

## 7.20.2 Biosynthesis of Methanofuran

### 7.20.2.1 Introduction

Methanofuran is the initial C<sub>1</sub> acceptor molecule in the formation of methane through methanogenesis, the only known biochemical route to methane. Methanofuran is involved in the first two-electron reduction of carbon dioxide to produce the formamide derivative of methanofuran, where the formate is attached to the amino group of methanofuran (step 1, **Figure 1**).<sup>15</sup> Methanofuran is also believed to be involved in the formaldehyde oxidation/detoxification pathway in a number of organisms including *Methylobacterium extorquens* AMI and *Bukbolderia fungorum* LB400.<sup>16,17</sup> Despite the involvement of methanofuran in this first step of methanogenesis, less is known about the enzymes and corresponding genes in methanofuran biosynthesis than any of the other methanogenic coenzymes. Although the primary precursors of methanofuran are known (**Figure 2**), less is known about the individual steps in the biosynthesis of this compound or the corresponding enzymes involved. Indeed, only one enzyme in the methanofuran biosynthetic pathway, tyrosine decarboxylase, has been cloned and characterized.<sup>18</sup>

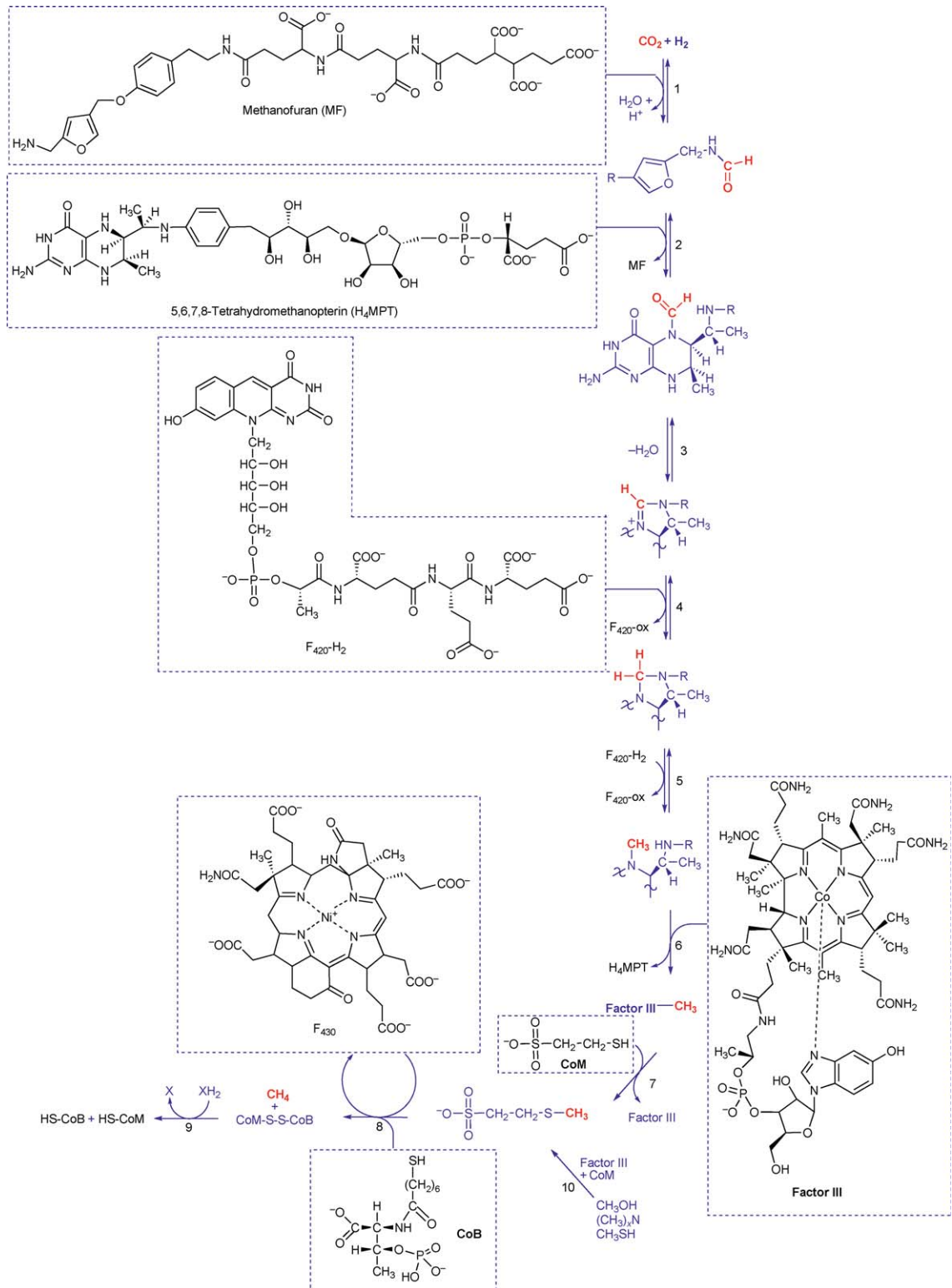
Several structurally different methanofurans are currently known with the nature of the differences residing in modifications of the side chain attached to the basic core structure of 4-[N-(γ-L-glutamyl-γ-L-glutamyl)-p-(β-aminoethyl)phenoxy]methyl-2-(aminomethyl)furan found in all methanofurans.<sup>19</sup> The first methanofuran to be structurally characterized contained *meso*-1,3,4,6-hexanetetracarboxylic acid (HTCA) attached to this core structure (**Figure 2**). *Methanobrevibacter smithii* contained methanofuran-c, which was the same as the originally characterized methanofuran except for the presence of a hydroxyl group at the 2 position of the 1,3,4,6-hexanetetracarboxylic acid moiety of the molecule.<sup>19</sup> It is clear, based on the large numbers of different structures found in the very limited number of methanogens examined, that many more modifications in the structures of methanofurans probably exist in nature.

A series of early biochemical studies have identified the primary precursors of methanofuran as phosphoenolpyruvate, dihydroxyacetone phosphate, tyrosine, glutamate, 2-ketoglutarate, acetate, and CO<sub>2</sub> as illustrated in **Figure 2**. Isotope incorporation studies with <sup>2</sup>H- and <sup>13</sup>C-labeled precursors have shown that the furan moiety of methanofuran is formed from phosphoenolpyruvate and dihydroxyacetone phosphate. Following dehydration, the next step in the biosynthesis of the furan would be a reduction of the carboxylic acid to an aldehyde. The final step in the biosynthesis of the furan moiety of methanofuran is expected to be a transamination reaction to form the 2-aminomethyl subunit.

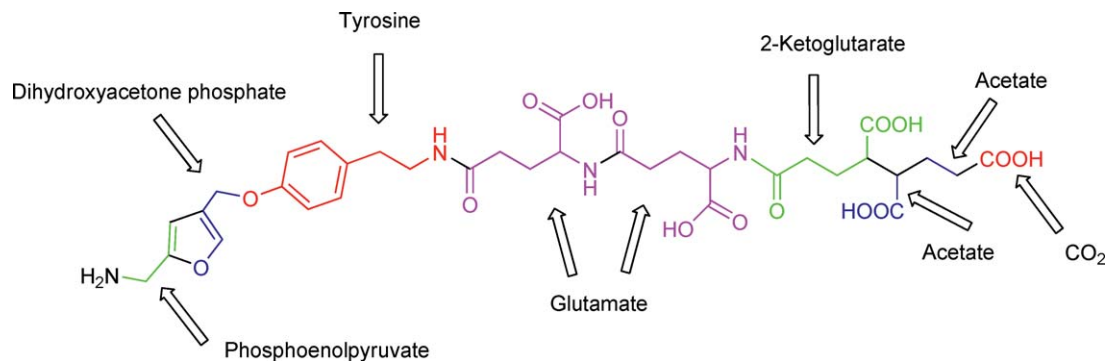
### 7.20.2.2 Tyrosine Decarboxylase

The one enzyme involved in methanofuran biosynthesis that has been characterized is a new tyrosine decarboxylase (MfnA, MJ0050) that is involved in the production of tyramine.<sup>18</sup> MfnA is a PLP-dependent decarboxylase and represents the first archaeal enzyme believed to be specifically involved in the biosynthesis

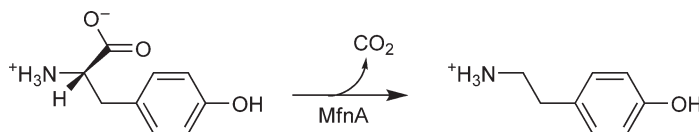




**Figure 1** The methanogenic pathway and the structures of the coenzymes involved.



**Figure 2** Chemical structure of the first characterized methanofuran and its precursors.



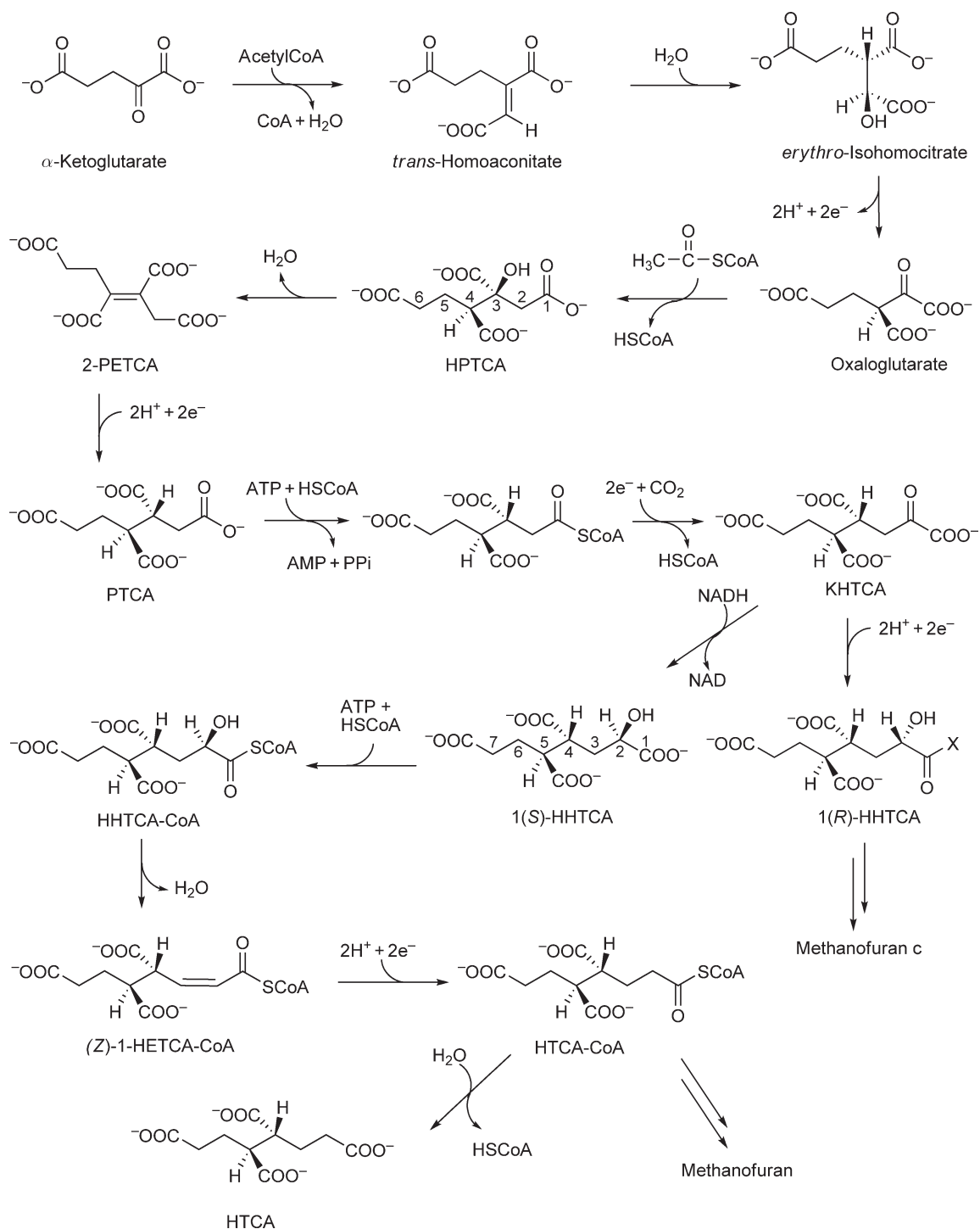
**Figure 3** Reaction catalyzed by tyrosine decarboxylase (MfnA).

of methanofuran. MfnA has been shown to catalyze the formation of tyramine from L-tyrosine as shown in **Figure 3**. The archaeal tyrosine decarboxylase family is in the same protein superfamily as the bacterial and plant tyrosine decarboxylases, though it shares less than 25% amino acid sequence identity to these enzymes.

### 7.20.2.3 Biosynthesis of *Meso*-1,3,4,6-Hexanetetracarboxylic Acid

HTCA (**Figure 4**) is not only a component of the first methanofuran structurally characterized, but also has been found at high concentrations in *Metbanothermobacter thermoautotrophicus* (formerly *Metbanobacterium thermoautotrophicum*  $\Delta H$ ) and has been proposed to serve as an osmolyte.<sup>20</sup> The formation of the HTCA moiety of methanofuran in *Metbanosarcina thermophila* has been studied in some detail by our research team,<sup>21</sup> but much of the information on its biosynthesis has yet to be published. The first step in the biosynthesis of the HTCA moiety of methanofuran is the formation of *trans*-homoaconitate from the condensation of  $\alpha$ -ketoglutarate with acetyl-CoA. This condensation reaction has been shown to be catalyzed by AksA, coded for by the *Metbanococcus jannaschii* MJ0503 gene. This gene product has also been demonstrated to be involved in coenzyme B biosynthesis (**Figure 21**).<sup>22</sup> The *trans*-homoaconitate is hydrated to *erythro*-isohomocitrate, which, after a proposed oxidation, produces oxaloglutaric acid that spontaneously racemizes between the *R* and *S* stereoisomers.

The next step in the formation of HTCA is the elongation of the carbon skeleton of oxaloglutarate by two carbon units. On the basis of the results of the isotope labeling experiments, the added carbons are expected to be derived from acetate, presumably through acetyl-CoA. This reaction is akin to the ketoacid chain elongation process that is important in the generation of fatty acid-like molecules one carbon at a time from acetate.<sup>23</sup> The important difference between this reaction and that normally observed in the ketoacid chain elongation process, is that in the subsequent elimination of water the proton lost does not come from C-2 but from C-4. This is observed by the loss of water from 2-hydroxy-1,2,3,5-pentanetetracarboxylic acid (HPTCA) shown in **Figure 4**. This is made possible because of the presence of the C-4 carboxylic acid and results in the double bond being generated between C-3 and C-4. Both the *erythro*- and *threo*-diastereomers of HPTCA are formed. At this point in the pathway, the production of these two diastereomers generates two parallel pathways, which ultimately lead to the same final product. In order to simplify **Figure 4** only one of these isomers is shown. The *erythro*- and *threo*-diastereomers of HPTCA undergo the elimination of water to form (*Z*)-1,2,3,5-pent-2-enetetracarboxylic acid ((*Z*)-2-PETCA) and (*E*)-1,2,3,5-pent-2-enetetracarboxylic acid ((*E*)-2-PETCA), respectively. The formed 2-PETCA isomers are then reduced to *erythro*- and *threo*-isomers of



**Figure 4** Proposed pathway to *meso*-1,3,4,6-hexanetetra-carboxylic acid (HTCA).

1,2,3,5-pentanetetra-carboxylic acid (PTCA). The PTCA isomers are then presumably activated for addition of the final carbon unit through the formation of a CoA thioester. Reductive carboxylation of the CoA esters of these PTCA isomers then generates the *erythro*- and *threo*-isomers of 1-oxo-1,3,4,6-hexanetetra-carboxylic acid (KHTCA). These KHTCA molecules were found to readily interconvert nonenzymatically between their

(3*R*)- and (3*S*)-isomers. This racemization likely occurred when the molecules contained the enol form of the ketone. NADH-dependent reduction of the KHTCA isomers produced two isomers at C-1 of 1-hydroxy-1,3,4,6-hexanetetracarboxylic acid (HHTCA). Both of these isomers had opposite stereochemistry at C-1 than is found in the 1(*R*)HHTCA isomer present in methanofuran c. This indicates that 1(*R*)HHTCA isomers are not intermediates in HTCA biosynthesis. Dehydration of the coenzyme A derivative of 1(*S*)HHTCA would form the CoA derivative of *erythro*- and *threo*-(*Z*)-1,3,4,6-hex-1-enetetracarboxylic acid ((*Z*)-1-HETCA-CoA). This reaction is analogous to the dehydration of 2-hydroxyglutaryl-CoA by 2-hydroxyglutaryl-CoA dehydratase, which is composed of three subunits,  $\alpha$ (HgdA),  $\beta$ (HgdB), and an activator protein HgdC.<sup>24</sup> A similar enzyme has been implicated in the biosynthesis of  $\beta$ -glutamate in *M. jannaschii*; however, the  $\alpha$  and  $\beta$  subunits are contained within a single protein encoded for by the MJ0007 gene (M. F. Roberts, unpublished data).

The (*Z*)-1-HETCA-CoA isomers would chemically racemize at C-4 and thus reduction of only the *erythro* isomer would produce HTCA or its CoA derivative with the correct stereochemistry at C-4 and C-5. Retaining the CoA portion of the molecule as indicated in **Figure 4** would produce an activated carboxylic acid that could be used for the incorporation of HTCA into the final methanofuran. The biosynthetic intermediates, *erythro*-HPTCA, *threo*-HPTCA, *erythro*-PTCA, and *threo*-PTCA were each identified in cell extracts of *Methanococcus vannielii*, *M. thermophilus* strain TM-1, and *M. thermoautotrophicus*, indicating the widespread occurrence of this pathway in the methanogens.

#### 7.20.2.4 Assembly of the Final Structure

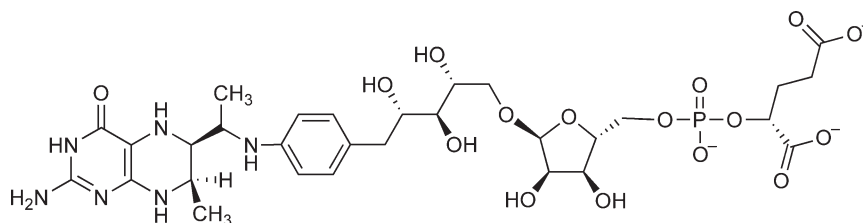
Besides the ether linkage attaching the furan to the tyramine, the rest of the methanofuran structure is held together by amide bonds that would most probably be generated by nonribosomal peptide synthetases. The glutamate dipeptide portion of methanofuran could be formed either prior to or during the final assembly of methanofuran. This reaction would be expected to involve the activation of a glutamate residue (potentially as an adenylate) followed by the formation of the peptide bond. Enzyme classes that could be capable of catalyzing this reaction include the amino acid ligases, peptide synthetases, and the aminoacyl-tRNA synthetases.

### 7.20.3 Biosynthesis of Methanopterin

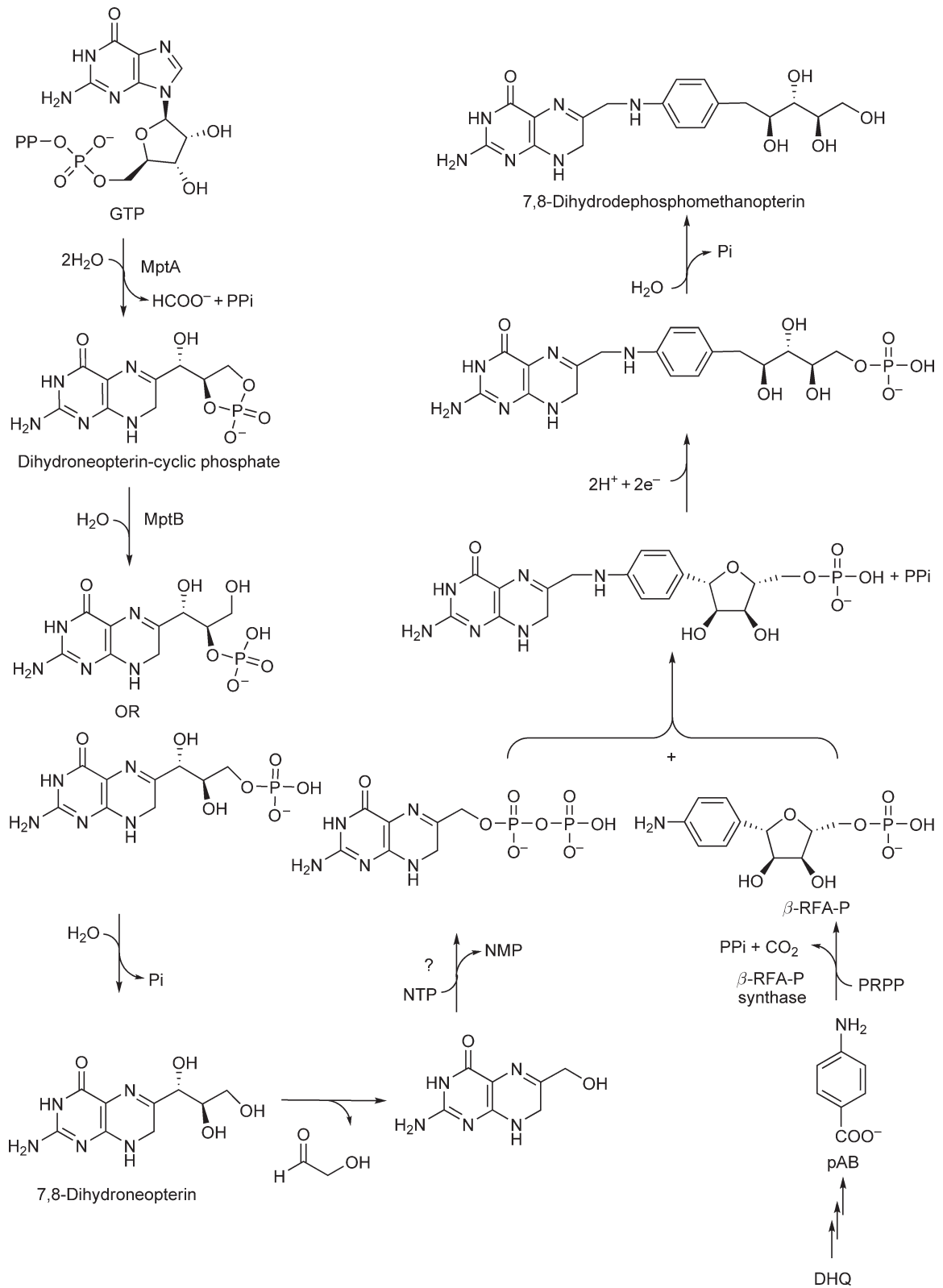
#### 7.20.3.1 Introduction

Tetrahydromethanopterin (H<sub>4</sub>MPT) is involved in the second through sixth steps of methanogenesis where oxidation of the CO<sub>2</sub>-derived formyl group to a methyl group occurs (**Figure 1**). The formyl group is first transferred from formylmethanofuran to N<sup>5</sup> of H<sub>4</sub>MPT. Formylmethanopterin is then cyclized and after two consecutive reductions by F<sub>420</sub>-dependent enzymes produces methyl-H<sub>4</sub>MPT. H<sub>4</sub>MPT (**Figure 5**) and its homologues have always been considered the only C<sub>1</sub>-carrier coenzymes functioning in the methanogens. This dogma has now been broken with the finding that several gene products in the methanogen *Methanosarcina barkeri* are tetrahydrofolate-specific enzymes.<sup>25</sup>

The biosynthesis of H<sub>4</sub>MPT can be viewed as being composed of two converging pathways, one beginning with GTP to generate the pterin portion and the other beginning with 3-dehydroquinate (DHQ) producing 4-aminobenzoic acid (*p*AB), the precursor to the arylamine (5-(4-aminophenyl)-1,2,3,4-tetrahydroxypentane) portion of the final structure (**Figure 6**). After condensation of these two units to generate the functional



**Figure 5** The first structurally characterized tetrahydromethanopterin.



**Figure 6** Proposed pathway for the biosynthesis of H<sub>2</sub>dephosphomethanopterin.





most archaeal genomes.<sup>36,37</sup> Early labeling studies in rumen isolate 10-16B<sup>31</sup> and *M. thermoautotrophicus*<sup>38,39</sup> supported the origin of the *pAB*-derived portion of methanopterin from chorismate or an earlier intermediate in the shikimic acid pathway.<sup>31</sup>

To probe for the possible involvement of intermediates of the shikimic acid pathway in *pAB* biosynthesis, gene knock-outs were constructed in *Metbanococcus maripaludis*.<sup>40</sup> *M. maripaludis* is a strictly anaerobic, methane-producing archaeon and facultative autotroph capable of biosynthesizing all the amino acids and vitamins required for growth. Similar to many other euryarchaeotes, the *M. maripaludis* genome does not encode homologues for the first two steps of the canonical pathway of aromatic amino acid (AroAA) biosynthesis. Instead, it has a different pathway that utilizes two different gene products for the first two steps.<sup>41</sup> On the basis of the information obtained from the knockout studies, *M. maripaludis* probably uses the Mmp0686 (AroA', MJ0400) and Mmp1249 (AroB', MJ1249) gene products for the first two steps in the modified pathway (Figure 8). The Mmp0686 deletion mutant required AroAAs and *pAB* for growth. An *aroD* deletion mutant that blocked the third step of the archaeal pathway, so that it cannot make DHS or chorismate, required AroAAs – but not *pAB* – for growth. These results suggest that the AroAAs are formed from 3-dehydroshikimate (DHS) as expected, but that *pAB* is not. On the basis of the labeling pattern and feeding studies with this mutant, as well as wild-type cells, it was proposed that *pAB* derives from 3-dehydroquinic acid (DHQ). This was confirmed because cell extracts of an *aroD* mutant of *M. maripaludis* produce *pAB* from DHQ but not DHS, indicating that DHQ is the precursor to *pAB*. This result explains the absence of the known genes for *pAB* biosynthesis in the archaea, because they use another pathway.

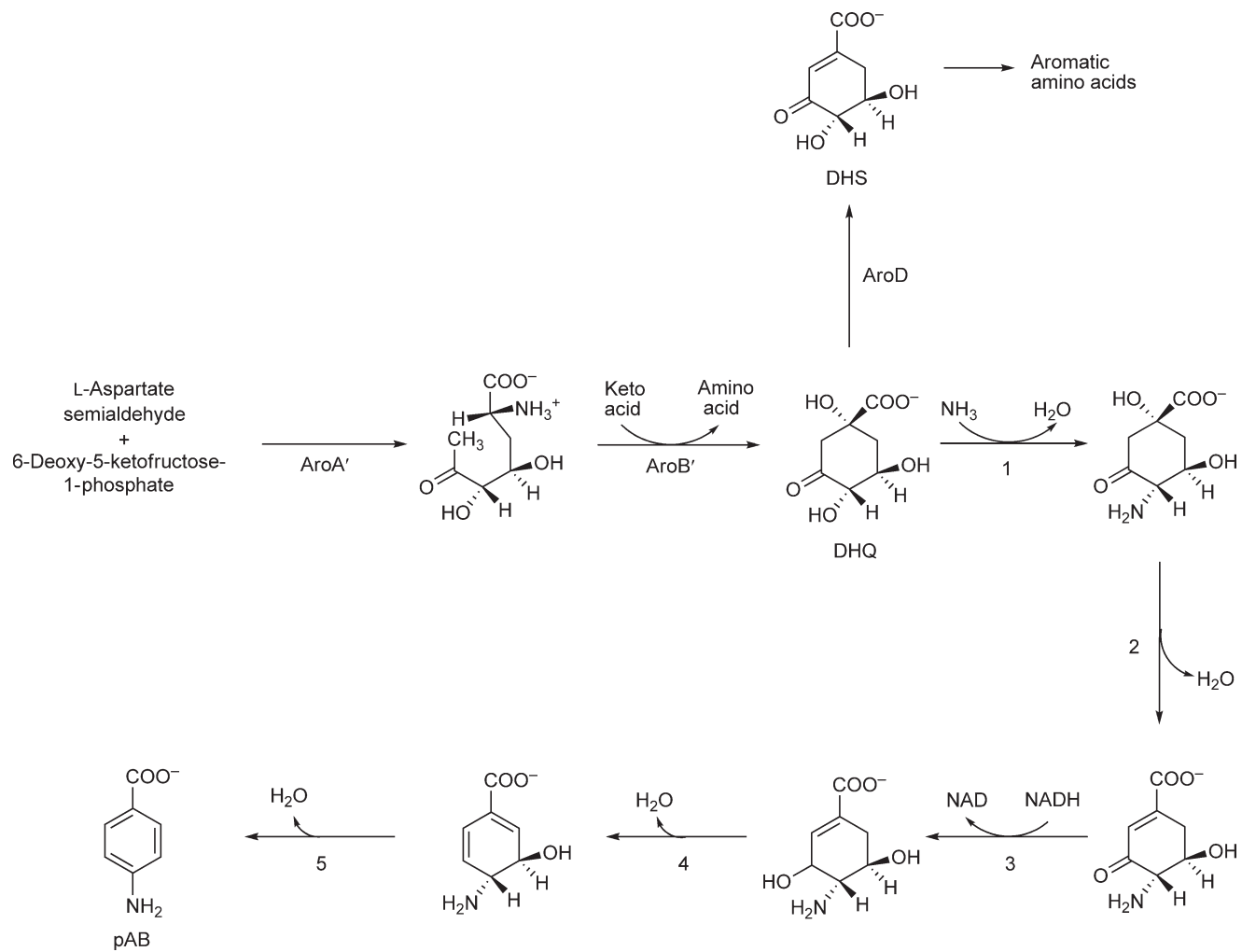
Several different pathways can be envisioned for the conversion of DHQ into *pAB*. One such pathway (Figure 8) has some of the same chemical features as the shikimate pathway but begins with a 4-amino analogue of DHQ. The critical first step in the pathway may proceed in a manner similar to that seen in glucosamine-6-phosphate biosynthesis, in which an  $\alpha$ -hydroxyaldehyde is converted into an  $\alpha$ -aminoaldehyde.<sup>42,43</sup> It is possible that the reduction of the ketone (step 3, Figure 8) could be performed by the normal shikimate enzyme, provided it could function with the 4-amino-containing analogue. Step 4 could occur either by loss of water or by phosphorylation followed by the loss of phosphate as occurs in the shikimate pathway. The last compound in this pathway, *trans*-4-amino-3-hydroxy-1,5-cyclohexadiene-1-carboxylic acid, has been prepared<sup>44</sup> and we have now shown it to be a precursor to *pAB*.

This same pathway could also be used to generate the arylamine present in the modified methanopterin, thermopterin, found in *Metbanococcus thermophilicum*.<sup>45</sup> Thermopterin could be generated from 3,6-dihydroxy-4-aminobenzoic acid (DHAB) instead of *pAB* through a parallel set of reactions. Here the pathway could be similar to *pAB* but starting with dihydroxyacetone or a non-6-deoxy-5-keto sugar, such as 5-ketofructose-1-phosphate, as shown in Figure 9.

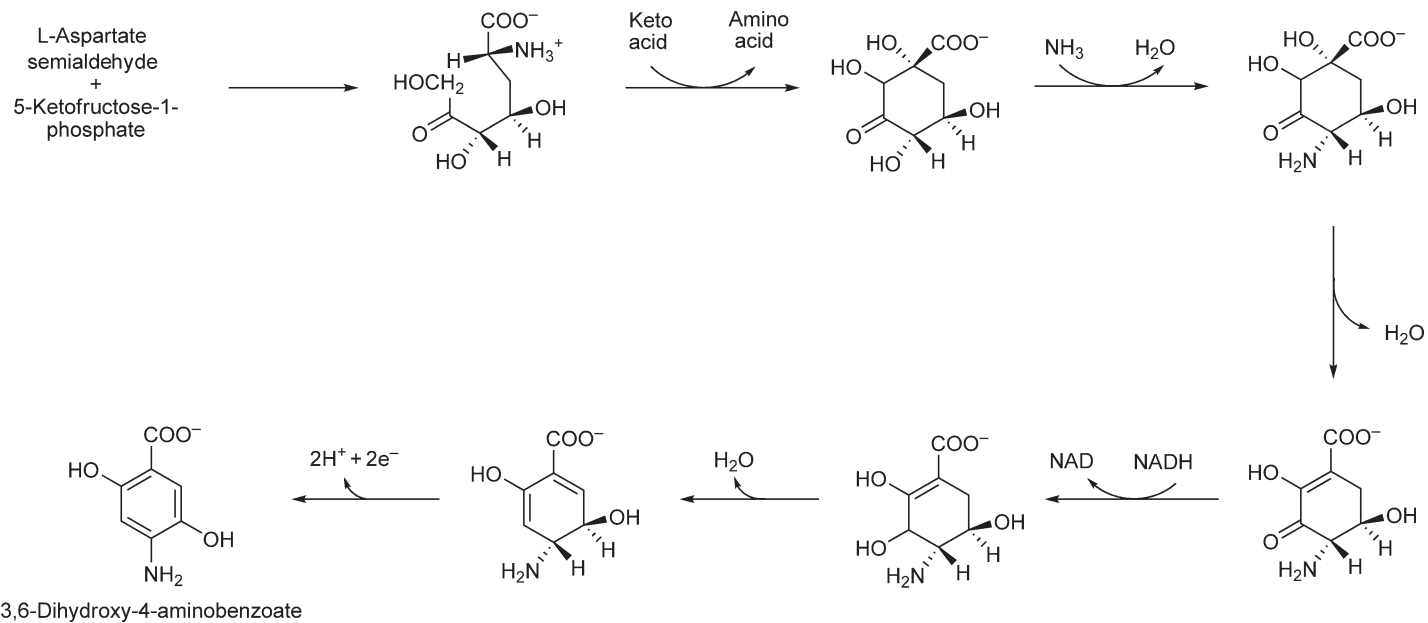
#### 7.20.3.4 Formation of 4-( $\beta$ -D-Ribofuranosyl)Aminobenzene-5'-Phosphate

Early work showed that the precursor to the arylamine of H<sub>4</sub>MPT, 4-( $\beta$ -D-ribofuranosyl)-aminobenzene-5'-phosphate ( $\beta$ -RFA-P), was derived from *pAB* and 5-phospho- $\alpha$ -D-ribose-1-diphosphate (PRPP) (lower right, Figure 6).<sup>32</sup> This reaction is catalyzed by  $\beta$ -RFA-P synthase. The  $\beta$ -RFA-P synthase from *M. thermophila* was purified and its N-terminal sequence was used to identify the corresponding genes from several archaea,<sup>46</sup> including MJ1427 from *M. jannaschii*. The corresponding protein from *Archaeoglobus fulgidus* was expressed in *E. coli* and the enzymatic activity of the recombinant gene product was verified.<sup>36</sup>

Additional work on the mechanism of the recombinant  $\beta$ -RFA-P synthase from *M. jannaschii* (MJ1427) has been reported.<sup>47</sup> Recombinant *M. jannaschii*  $\beta$ -RFA-P synthase was shown to be a homodimer containing two 36.2-kDa subunits. As previously proposed, *M. jannaschii*  $\beta$ -RFA-P synthase<sup>32</sup> was found to lack PLP on the basis of the absence of its expected absorbance. Mass spectral analysis of the enzyme showed that the measured subunit mass of the enzyme was the same as its predicted molecular mass, so the enzyme was not expected to contain a pyruvoyl group. This was confirmed by the absence of pyruvate release upon acid hydrolysis of the enzyme. Steady-state initial velocity and product inhibition kinetic studies indicated an ordered mechanism involving first the binding of PRPP followed by *pAB*. Product release occurred in the order CO<sub>2</sub>,  $\beta$ -RFA-P, and PPI.



**Figure 8** A possible biosynthetic route to 4-aminobenzoic acid in archaea.



**Figure 9** Proposed pathway to 3,6-dihydroxy-4-aminobenzoate (DHAB), a possible precursor to the arylamine in thermopterin.

The demonstration of a growth requirement of *M. barkeri* for folate or pAB has been recently demonstrated.<sup>25</sup> As part of this work, evidence showed that the pAB taken up by the cells was not required for the biosynthesis of tetrahydrosarcinapterin (H<sub>4</sub>SPT), the modified methanopterin present in these methanogens. This result indicated that either another route to  $\beta$ -RFA-P was operating in the cells or that an enzyme-bound form of pAB, not mixing with the pAB from the media, was the source of the pAB used for H<sub>4</sub>SPT biosynthesis.

The bacterium *M. extorquens* has been found to contain a modified H<sub>4</sub>MPT called, dephospho-H<sub>4</sub>MPT, which consists of only the H<sub>4</sub>MPT core structure without the side chain.<sup>48</sup> Dephospho-H<sub>4</sub>MPT would be produced by the reduction of 7,8-dihydrodephosphomethanopterin (Figure 6). This indicates that *M. extorquens* contains the genes/enzymes for the biosynthesis of the core structure but lacks those needed for the biosynthesis of the side chain. This was indicated by the observed absence of dephospho-H<sub>4</sub>MPT in cells with mutation in the *orf4* or *dmrA* (dihydromethanopterin reductase) genes. The *orf4* gene was shown to code for the  $\beta$ -RFA-P synthase. This was the first evidence for the H<sub>4</sub>MPT biosynthetic genes in bacteria. The *orf4* gene has 29% identity to the  $\beta$ -RFA-P synthase from *A. fulgidus*.<sup>46</sup>

### 7.20.3.5 Biosynthesis of the Methanopterin Side Chain

For the most part, less new information has appeared about the enzymes/genes involved in the production of the methanopterin side chain (Figure 10). One of the enzymes that has been characterized is that derived from the MJ0620 gene and is responsible for the addition of the terminal  $\alpha$ -L-linked glutamate present in the modified *M. jannaschii* methanopterin, sarcinapterin.<sup>49</sup> The enzyme was named H<sub>4</sub>MPT:  $\alpha$ -L-glutamate ligase and designated MptN. *In vitro*, MptN was able to utilize either GTP or ATP with similar specific activities. MptN is a member of the ATP-grasp superfamily of amide bond-forming ligases and is a homologue of glutathione synthetase.<sup>49</sup>

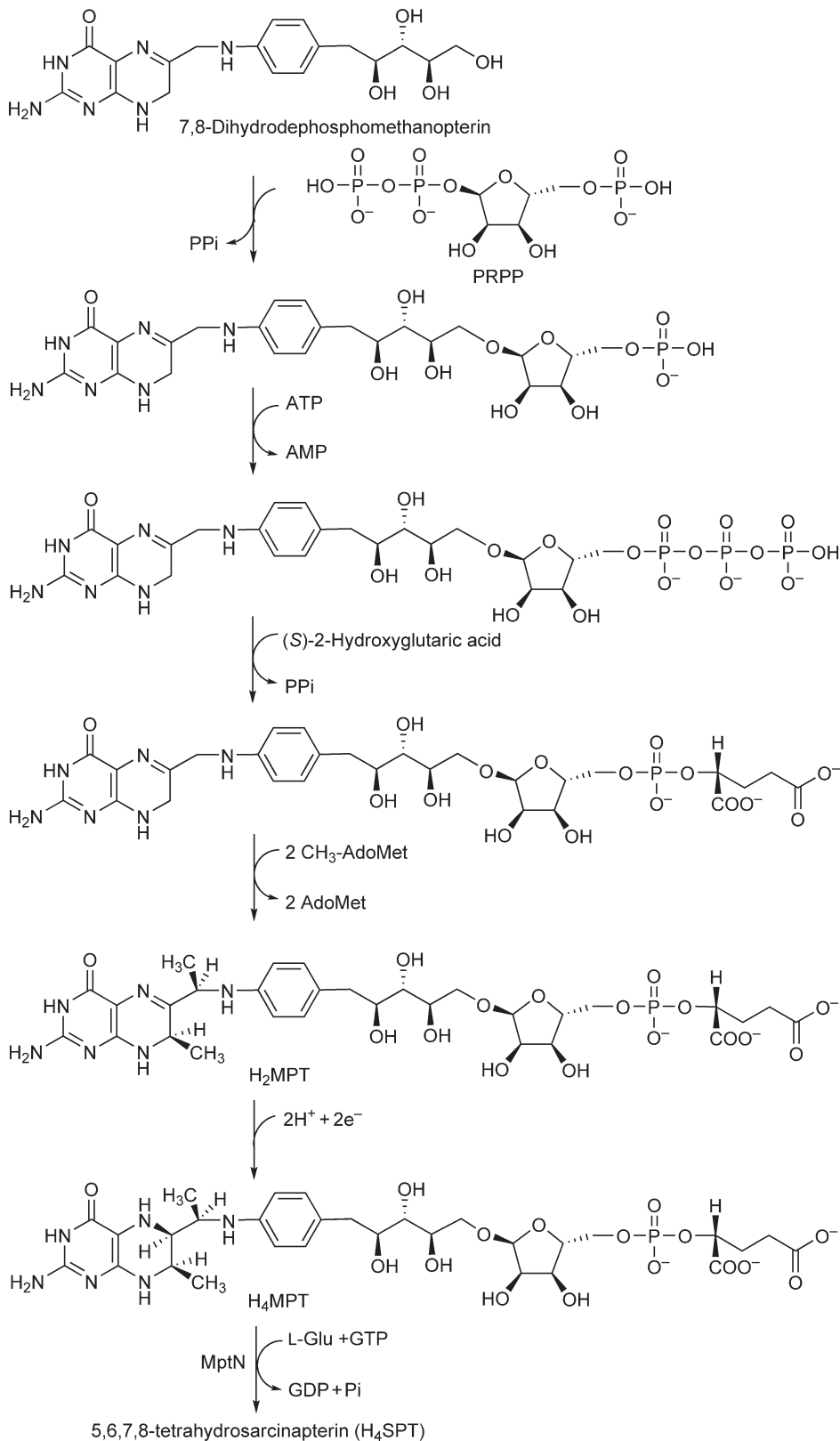
### 7.20.3.6 Methylation of the Pterin

Work is yet to be carried out to identify the methyltransferases involved in the methylation of H<sub>4</sub>MPT. They are expected to be *S*-adenosylmethionine-dependent enzymes that abound in the methanogens, because each methyl group of H<sub>4</sub>MPT has been shown to be derived from the methyl group of methionine.<sup>50</sup> On the basis of the chemistry required, the expectation is that the methylation would occur at the dihydropterin stage, where the imine functional group could generate anionic carbons at C-7 and C-9. The presence of methyl groups in the structure introduces two asymmetric centers but appears to have little effect on the electron density of N<sup>5</sup>. Their presence may only affect the entropy of the reactions involving this coenzyme.<sup>51</sup>

### 7.20.3.7 Reduction of Dihydromethanopterin

The specific enzyme(s) for the reduction of dihydromethanopterins (H<sub>2</sub>MPT) to the functional H<sub>4</sub>MPTs in methanogens has not been established. This is despite the fact that this reduction was implied by experimental observations in Wolfe's lab many years ago.<sup>52</sup> The reaction is analogous to that catalyzed by dihydrofolate reductase (DHFR) in folate biochemistry<sup>53,54</sup> and occurs with the same stereochemistry.<sup>55,56</sup> It has been reported that the *dmrA* gene in the bacterium *M. extorquens* AM1, which functions with a modified H<sub>4</sub>MPT, termed dephosphoH<sub>4</sub>MPT, is a gene encoding the elusive 7,8-dihydrodephosphomethanopterin reductase.<sup>54,57</sup> Despite experimental evidence for the reduction in the archaea, homologues of the *dmrA* gene are not present in the archaeal genomes suggesting that bacteria and archaea produce two evolutionarily distinct forms of H<sub>2</sub>MPT reductase.

As the functional genomics of MPT biosynthesis has evolved it has become increasingly clear that even though folate and MPT share similar C<sub>1</sub> chemistry that is based on the same chemical principles, these coenzymes are not evolutionary related to each other.<sup>51</sup> Thus, the original designation of methanopterin as a modified folate was clearly incorrect.



**Figure 10** A proposed pathway for the biosynthesis of the side chain of H<sub>4</sub>methanopterin starting from H<sub>2</sub> dephosphomethanopterin.

## 7.20.4 Biosynthesis of Coenzyme F<sub>420</sub> and Riboflavin Precursors

### 7.20.4.1 Introduction

The fourth and fifth steps of methanogenesis involve consecutive reductions of the  $N^5N^{10}$ -methyl-H<sub>4</sub>MPT by F<sub>420</sub>-dependent oxidoreductases (**Figure 1**). The product of the first reduction is  $N^5N^{10}$ -methylene-H<sub>4</sub>MPT and that of the second is  $N^5$ -methyl-H<sub>4</sub>MPT. F<sub>420</sub> is also a required cofactor for reverse methanogenesis.<sup>11</sup> In addition, the hydride carrier coenzyme F<sub>420</sub> as F<sub>420</sub>H<sub>2</sub> is involved in many other important biological transformations including the reduction of toxic sulfite to sulfide in *M. jannaschii*<sup>58</sup> and the reduction of oxygen with the F<sub>420</sub>H<sub>2</sub> oxidase (FprA).<sup>59,60</sup> The chromophore and redox active portion of the F<sub>420</sub> molecule is 7,8-didemethyl-8-hydroxy-5-deazariboflavin (Fo). Fo alone serves as a coenzyme of the DNA photolyases used to repair DNA<sup>61</sup> and can function in place of F<sub>420</sub> in some enzymes using F<sub>420</sub>.<sup>62,63</sup>

### 7.20.4.2 Early Steps in F<sub>420</sub> and Riboflavin Biosynthetic Pathways Leading to 5-Amino-6-Ribitylamino-2,4(1H,3H)-Pyrimidinedione

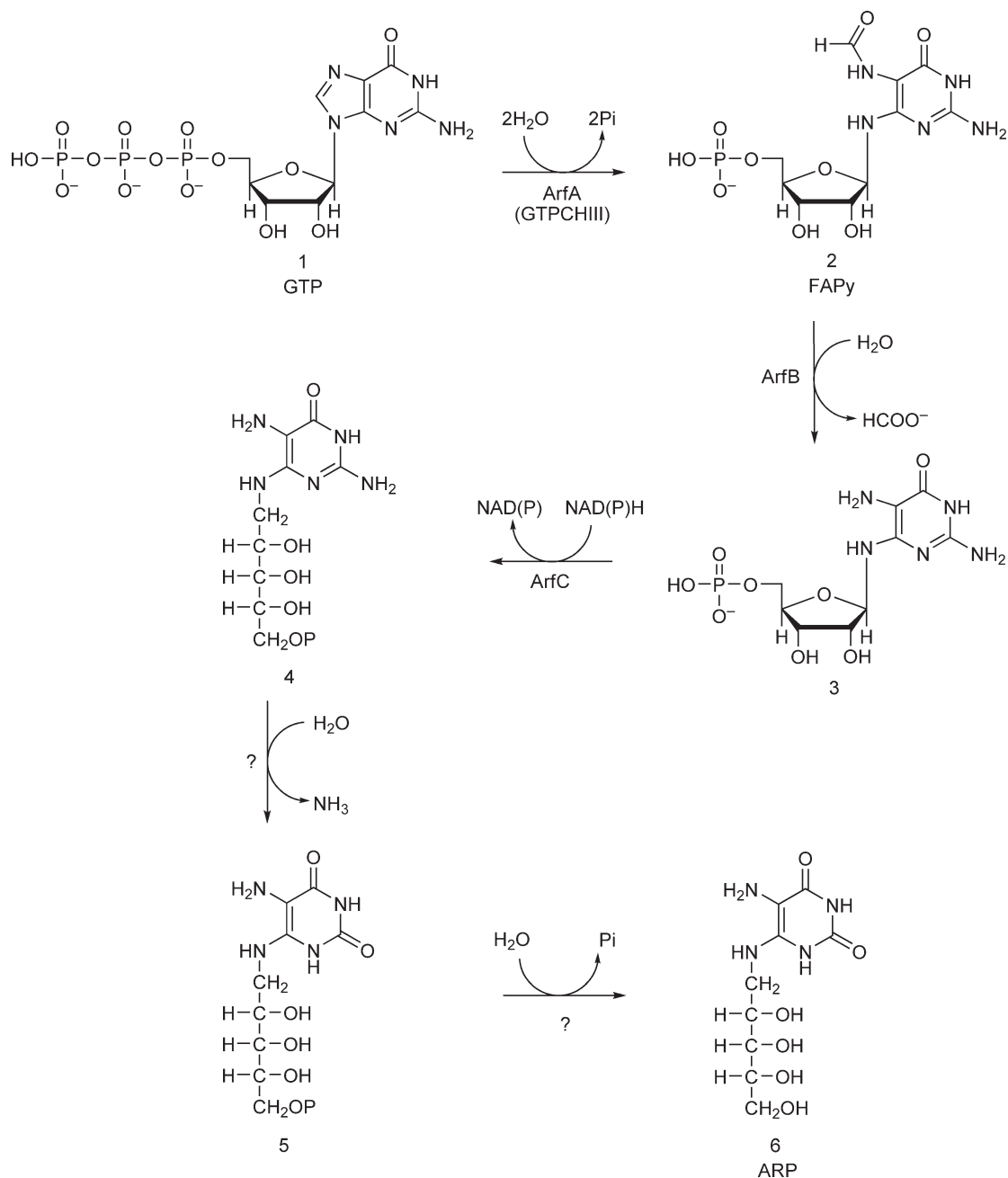
It has been previously shown that the biosynthesis of 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione (ARP), a branch point compound in the biosynthesis of F<sub>420</sub> and riboflavin, occurs by an alternate pathway in the archaea as shown in **Figure 11**.<sup>64</sup> The close structural similarities between the riboflavin and F<sub>420</sub> chromophores has long been recognized, but interestingly each biosynthetic pathway also utilizes similar chemical transformations to get from compound **5** (**Figure 11**) to the final functional coenzymes. These transformations include the removal and subsequent replacement of phosphate and the construction of the three heterocyclic rings of Fo and riboflavin. This is followed by the modification of the C-6 position of the ribitol side chain to finally generate the wide variety of F<sub>420</sub> coenzymes, FMN or FAD. In the case of riboflavin biosynthesis, it has always seemed odd to remove the phosphate and then replace it in the final product, FMN or FAD. This finding may indicate that a F<sub>420</sub> biosynthetic pathway not requiring the phosphate came first and the pathway was later modified by the reintroduction of the phosphate in FAD and FMN.

The first step in the pathway to ARP is the formation of 2-amino-5-formylamino-6-ribosylamino-4(3H)-pyrimidinone 5'-monophosphate from GTP (**Figure 11**). This reaction is catalyzed by an archaeal GTP cyclohydrolase, GTP cyclohydrolase III (GTPCHIII, ArfA), that is coded by the MJ0145 gene in *M. jannaschii*. The product of the GTPCHIII reaction, FAPy, is converted into 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate by a formamide hydrolase, ArfB, that is the MJ0116 gene product in *M. jannaschii*.

The next step in the formation of ARP is the reduction of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (compound **3**, **Figure 11**) by a specific reductase. The X-ray crystallographic structure of the 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate reductase from *M. jannaschii* (ArfC, MJ0671) has recently been published.<sup>65</sup> The structure of the enzyme with the bound NADP coenzyme was also determined. Amazingly, the protomer folding, aside from the addition of a small helix and a  $\beta$ -sheet in the homodimer, has been found to have typical structural features of the oxidoreductase family of dihydrofolate reductase (DHFR). The authors were unable to obtain crystals with bound 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate but this substrate was modeled into the putative active site of the NADP containing structure. This modeling suggested the transfer of the *pro*-R hydrogen of NADPH to the C-1' of the substrate in its imine form.<sup>66</sup> The active-site structure of the enzyme is analogous to that found in DHFR and was most closely aligned with the *Thermotoga martinia* DHFR.<sup>67</sup> Most methanogenic archaea contain no folate coenzymes<sup>25</sup> and thus no DHFR, so it is interesting to speculate that either this enzyme evolved into the DHFRs or was derived from DHFR. If methanogenesis and methanopterin was the first to evolve then perhaps the former is correct.

Two final reactions are required for the transformation of compound **4** in **Figure 11** to ARP. These include a deamination at C-2 of the pyrimidine to produce ARP-P (compound **5**, **Figure 11**). A final dephosphorylation would then give rise to ARP. The enzymes involved in these last two steps are yet to be identified.

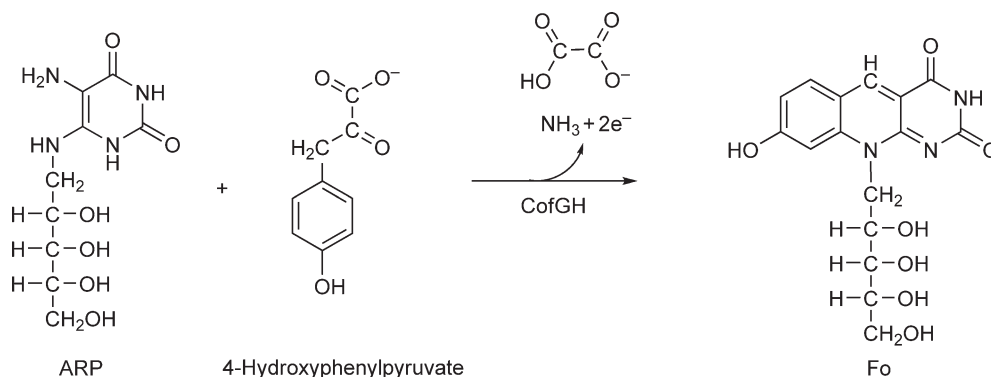




**Figure 11** Early steps in F<sub>420</sub> and riboflavin biosynthesis leading to ARP.

### 7.20.4.3 Biosynthesis of 7,8-Didemethyl-8-Hydroxy-5-Deazariboflavin

Microbes produce Fo by condensing ARP and 4-hydroxyphenylpyruvate, a precursor of tyrosine (Figure 12).<sup>68</sup> The enzyme has been named Fo synthase and in most organisms is composed of two proteins encoded by the *cofG* and *cofH* genes, MJ0446 and MJ1431, respectively, in *M. jannaschii*. Simultaneous expression of CofG and CofH in *E. coli* caused the host cells to produce small amounts of Fo during growth, and activated cell-free extracts containing both of these proteins have been shown to catalyze Fo biosynthesis *in vitro*.<sup>68</sup> Cyanobacterial genomes encode homologues of both *cofG* and *cofH*, necessary to produce the coenzyme for Fo-dependent DNA



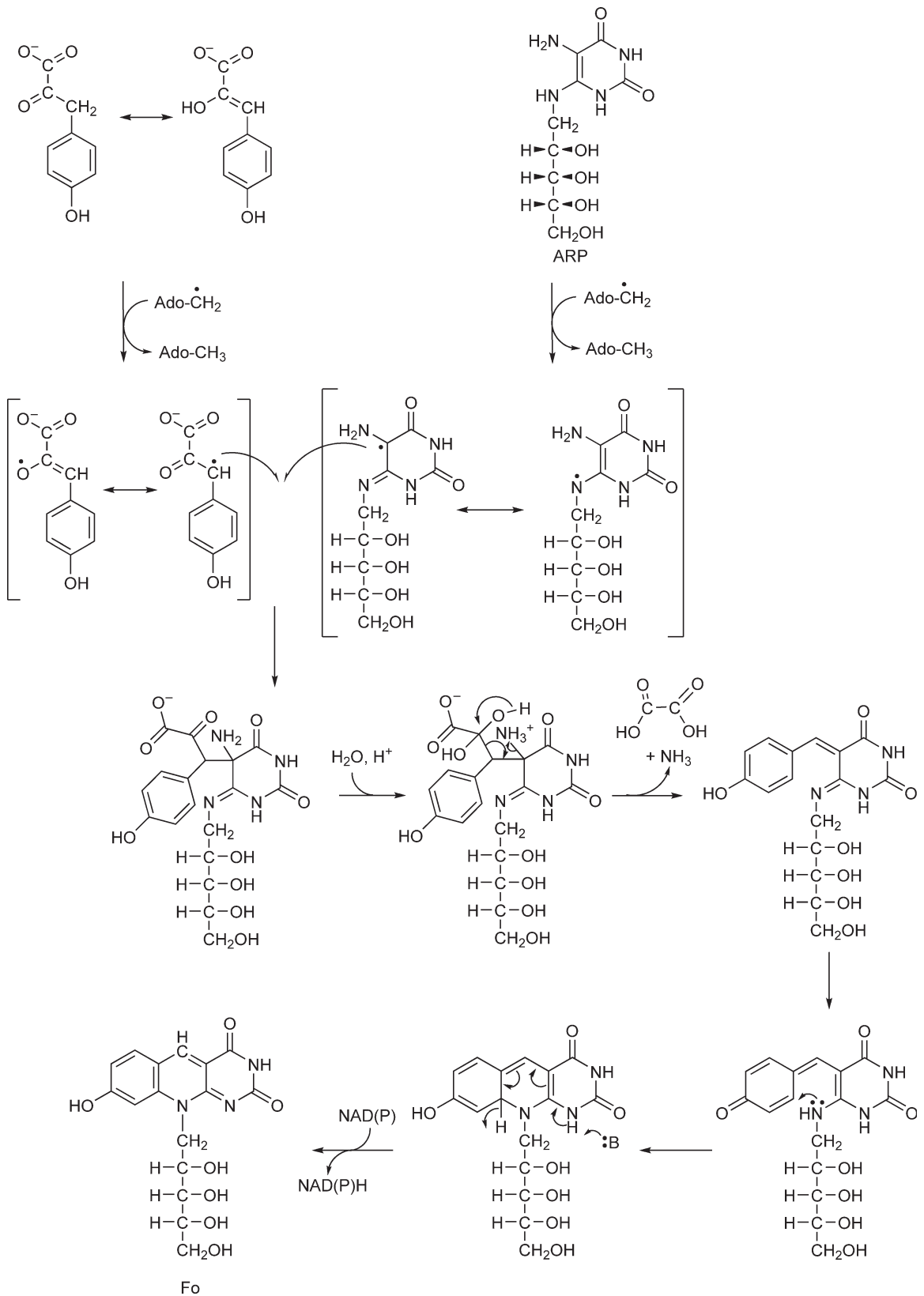
**Figure 12** Fo formation from APR and 4-hydroxyphenylpyruvate.

photolyases.<sup>61,68</sup> A molecular phylogeny of the paralogous *cofG* and *cofH* genes is consistent with the genes being vertically inherited within the euryarchaeal, cyanobacterial, and actinomycetal lineages. The *fbtC* gene, cloned from *Mycobacterium smegmatis*, encodes a Fo synthase within a single polypeptide chain.<sup>69</sup> Ancestors of the cyanobacteria and Actinomycetes must have acquired the two genes, which subsequently fused in Actinomycetes. Both the CofG and CofH proteins are members of the SAM radical family of enzymes,<sup>70</sup> and each has putative radical *S*-adenosylmethionine binding motifs. Preincubation of CofGH with *S*-adenosylmethionine, Fe(II), sulfide, and dithionite stimulates Fo production. Therefore, a radical reaction mechanism is proposed for the biosynthesis of Fo (Figure 13). We currently suspect that the gene product of MJ1488, designated as a dehydrogenase and closely associated genomically with the other F<sub>420</sub> biosynthetic genes, may also be involved in the production of Fo. Exactly where this functions in the biosynthetic pathway is not clear but its possible involvement as a dehydrogenase catalyzing the last step in the pathway, as shown in Figure 13, is proposed. The fact that production of Fo was low in the *E. coli* containing CofG and CofH could be a result of the absence of this gene product. This would indicate either that another enzyme(s) can catalyze the oxidation or the oxidation can proceed with air.

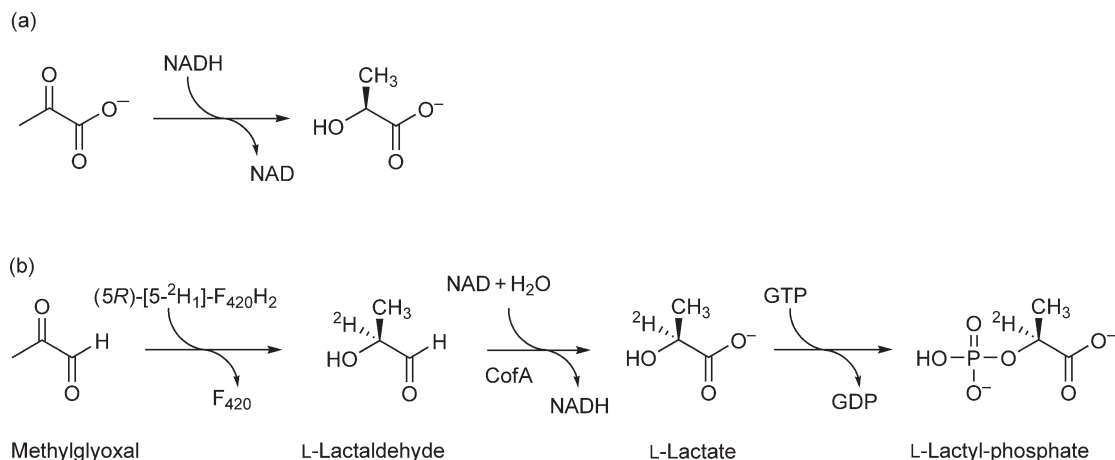
#### 7.20.4.4 Assembly of the Side Chains of the Core Structure of the F<sub>420</sub> Coenzyme F<sub>420</sub>-0

As reported in previous reviews, most of the enzymes involved in the biosynthesis of coenzyme F<sub>420</sub> have already been established.<sup>49,68,71–73</sup> Despite this work, several questions still remain to be answered. One of the early steps in the biosynthesis of coenzyme F<sub>420</sub> requires generation of 2-phospho-L-lactate, which is formed by the phosphorylation of L-lactate. This transformation has been observed in *M. jannaschii* cell free extracts and uses GTP as the phosphoryl donor (unpublished results). Lactate was originally considered to be the first intermediate in the pathway and it would easily arise by the reduction of pyruvate as shown in Figure 14(a). Despite the obviousness of this idea, it was not possible to demonstrate that any enzyme in the methanogens could efficiently reduce pyruvate to lactate.<sup>74,75</sup> Similarly, the direct formation of 2-phospho-L-lactate by reduction of phosphoenolpyruvate could not be demonstrated in *M. jannaschii*. An alternative route for the formation of the required L-lactate is the NAD-dependent oxidation of L-lactaldehyde, Figure 14(b). This route to lactate was confirmed in cell free extracts, and shown to be catalyzed by the MJ1411 gene product, CofA.<sup>73</sup> The lactaldehyde, in turn, was found to be generated either by the F<sub>420</sub>H<sub>2</sub>-dependent reduction of methylglyoxal<sup>76</sup> or by the aldol cleavage of fucose-1-phosphate by fucose-1-phosphate aldolase, the MJ1418 gene product.<sup>73</sup>

At present, the genes encoding the enzyme catalyzing the F<sub>420</sub>H<sub>2</sub>-dependent reduction of methylglyoxal and the kinase catalyzing the phosphorylation of lactate to 2-phospho-L-lactate are unknown. The current data indicate that the enzyme catalyzing the F<sub>420</sub>H<sub>2</sub>-dependent reduction of methylglyoxal uses the *Pro-R* hydrogen of the reduced F<sub>420</sub> coenzyme.<sup>76</sup> If this is confirmed it will be the first example of an F<sub>420</sub>H<sub>2</sub>-dependent enzyme with this stereospecificity, as hydride transfer from F<sub>420</sub> is known to occur from the *Pro-S* face in all of the 11 known F<sub>420</sub>-dependent enzymes.<sup>59,77</sup> The enzyme represents one of an assortment of enzymes in these cells that



**Figure 13** Proposed mechanism for the biosynthesis of F<sub>0</sub> catalyzed by CofGH.



**Figure 14** Two possible routes for the formation of L-lactyl-phosphate. The top of the figure (a) shows the pathway to lactate from pyruvate catalyzed by lactate dehydrogenase and observed in bacteria and eukaryotes. The bottom pathway (b) is that observed in the methanogens.

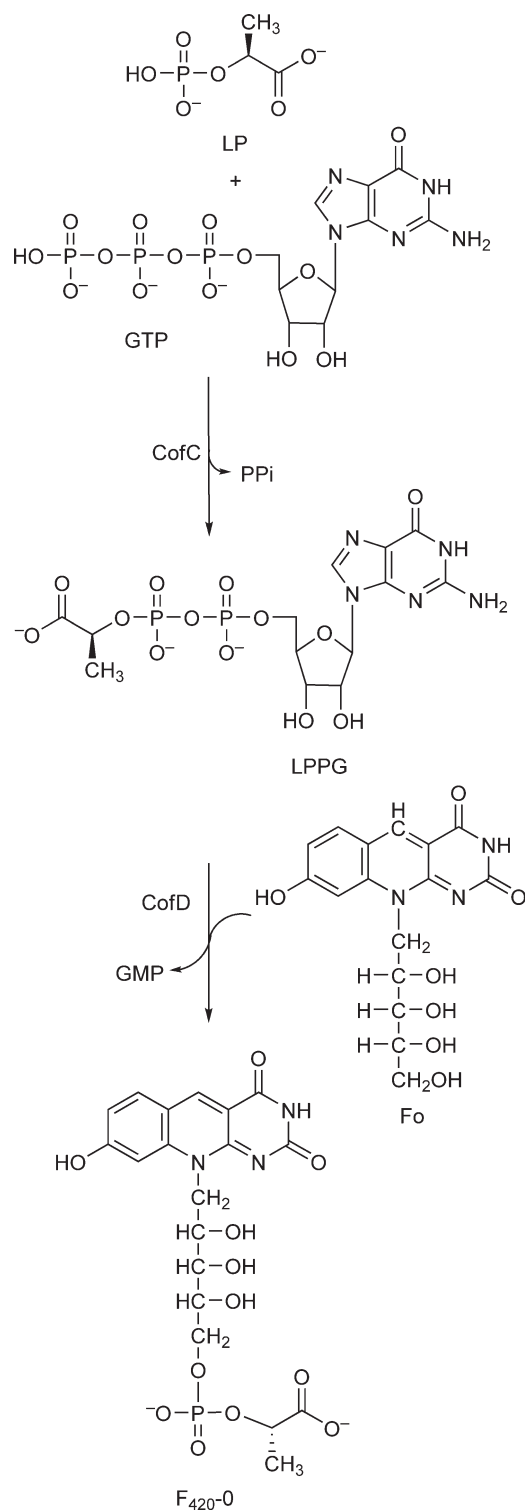
metabolize methylglyoxal. Subsequent work has demonstrated the involvement of methylglyoxal in other aspects of *M. jannaschii* metabolism including AroAA biosynthesis.<sup>78</sup> This was the first documented involvement of methylglyoxal, a known cellular toxin, in metabolism.

#### 7.20.4.5 Assembly of the Core Structure of F<sub>420</sub> Coenzymes

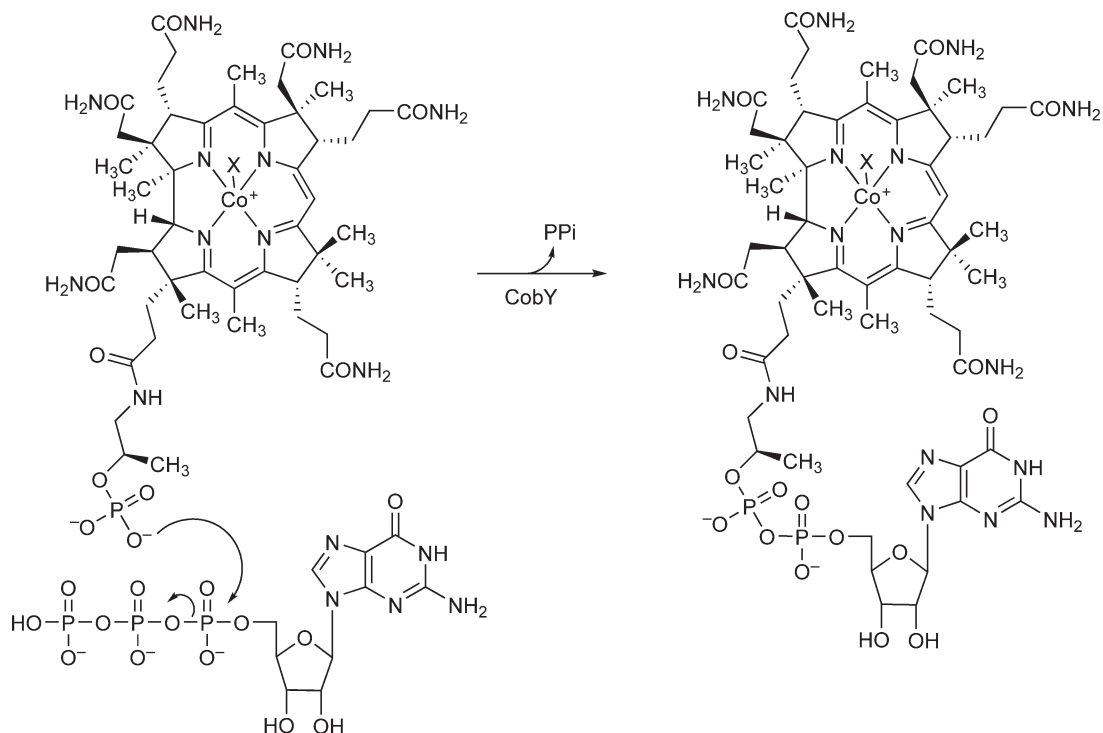
The final assembly of the core F<sub>420</sub> structure, F<sub>420</sub>-0, involves the attachment of 2-phospho-L-lactate to Fo (Figure 15). 2-Phospho-L-lactate is first activated as lactyl(2)diphospho-(5')guanosine (LPPG), to facilitate condensation with the Fo moiety to form F<sub>420</sub>-0. Bioinformatic analysis of archaeal genomes identified a gene, MJ0887, whose gene product has predicted structural similarity to nucleotidyl transferases and could catalyze this reaction. Subsequent cloning and overexpression of this gene and characterization of its gene product confirmed that it is the 2-phospho-L-lactate guanylyltransferase involved in coenzyme F<sub>420</sub> biosynthesis and has been named CofC to reflect this.<sup>79</sup>

This condensation involves a reaction that is analogous to the formation of the adenosylcobinamide guanosine diphosphate intermediate in vitamin B<sub>12</sub> biosynthesis (Figure 16). The corresponding reaction in archaeal B<sub>12</sub> biosynthesis was found to be catalyzed by the MTH1152 gene product in *M. thermoautotrophicus*.<sup>80</sup> Although the *M. thermoautotrophicus* and *Halobacterium* sp. NRC-1 homologues of MJ1117 have been shown to be involved in coenzyme B<sub>12</sub> biosynthesis, we originally proposed that the MJ1117 gene product may be a moonlighting enzyme capable of utilizing both 2-phospho-L-lactate and adenosylcobinamide-phosphate as acceptors for the guanylyltransferase reaction. Our *in vitro* experiments using the heterologous expressed protein clearly demonstrated that the MJ1117 gene product did possess 2-phospho-L-lactate guanylyltransferase activity, albeit with a much lower specific activity than the MJ0887 gene product. In addition, adenosylcobinamide-phosphate guanylyltransferase activity was confirmed by both *in vitro* and gene complementation experiments with the MJ1117 gene product (J. C. Escalante-Semerena, unpublished data).

The activated LP is next transferred to Fo by a 2-phospho-L-lactate transferase (CofD, MJ1256) to form the core F<sub>420</sub>-0 structure as shown in Figure 15. Heterologously expressed CofD from *M. jannaschii* has been characterized and the recombinant enzyme was found to utilize lactyl(2)diphospho-(5')guanosine (LPPG) as the preferred substrate but also had some activity with lactyl(2)diphospho-(5')adenine (LPPA).<sup>81</sup> The crystal structure of CofD from *Methanosarcina mazei* and the structures with either bound Fo and Pi or bound Fo and GDP have been determined.<sup>82</sup> CofD was found to have a Rossmann-fold core structure, and an active-site DxD motif was confirmed by site-directed mutagenesis. Although the binding site was determined from the Fo-bound structure, a consensus Fo/F<sub>420</sub> binding motif for F<sub>420</sub>-utilizing enzymes has still not been identified.



**Figure 15** Biosynthesis of F<sub>420</sub>-0.



**Figure 16** Adenosylcobinamide guanylyltransferase reaction in B<sub>12</sub> biosynthesis is analogous to that in the CofC catalyzed reaction.

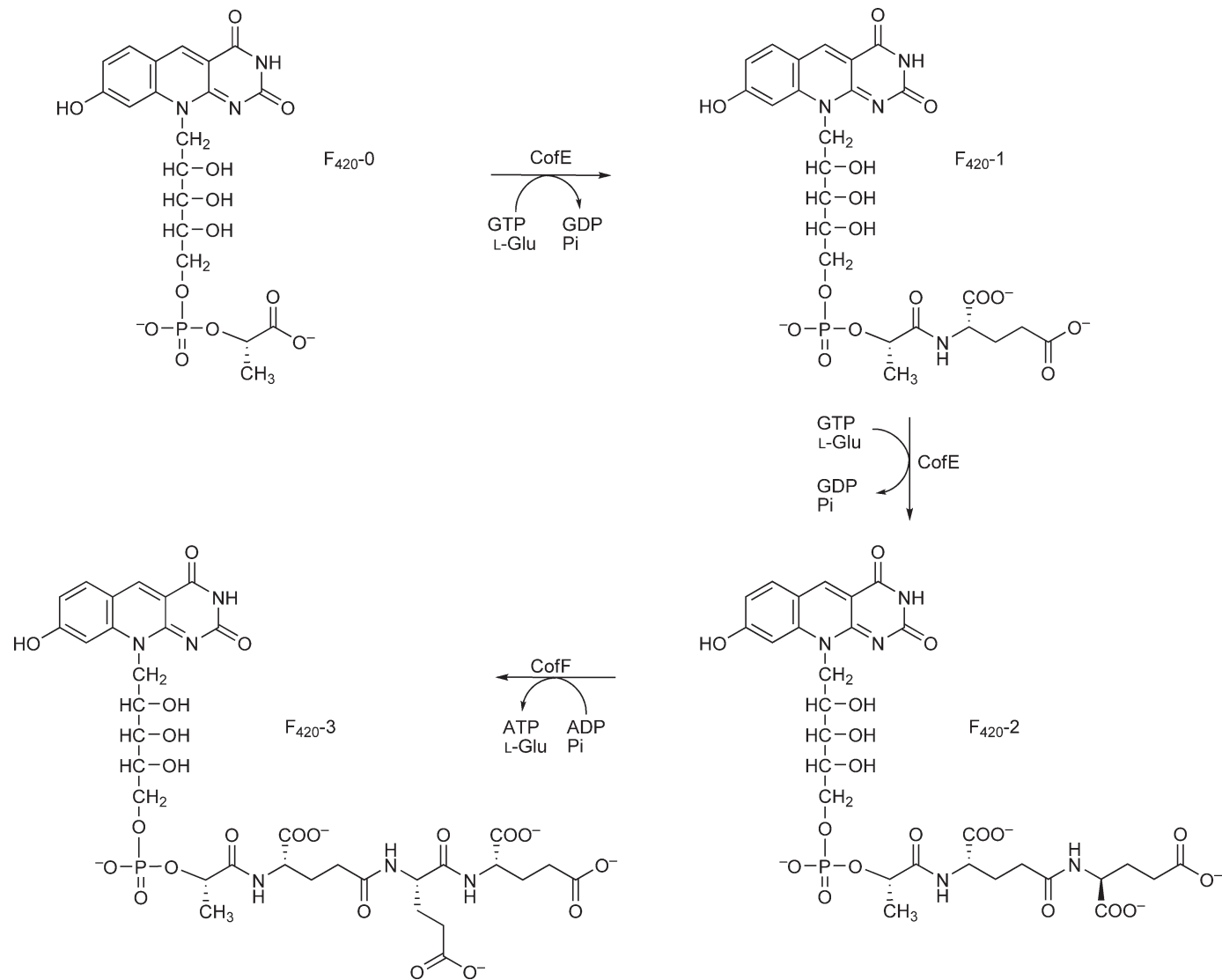
#### 7.20.4.6 Biosynthesis of the Side Chain

The final steps in coenzyme F<sub>420</sub> biosynthesis involve the addition of one or more glutamate residues to the lactate moiety of F<sub>420</sub>-0 (Figure 17). The number and attachment of the glutamate residues varies in F<sub>420</sub> isolated from different organisms. The *M. jannaschii* coenzyme contains three glutamates, with the first attached through the amino group to the carboxylic acid of the lactate. The second glutamate is linked to the  $\gamma$ -carboxylic acid of the first glutamate, and finally, the terminal glutamate is linked to the  $\alpha$ -carboxylic acid of the second glutamate (Figure 17).<sup>72</sup> The first two glutamates are added sequentially by CofE (MJ0768) and the terminal glutamate by CofF (MJ1001).<sup>49</sup> The crystal structure of the *A. fulgidus* CofE (AF2256) has been determined.<sup>83</sup> Similar to the *M. jannaschii* enzyme, *A. fulgidus* CofE catalyzes the GTP-dependent addition of two L-glutamates to F<sub>420</sub>-0 to form F<sub>420</sub>-2. It was shown that the enzyme adds glutamate to both F<sub>420</sub>-0 and F<sub>420</sub>-1 in two distinct steps. The crystal structure of the apo-F<sub>420</sub>-0: $\gamma$ -glutamyl ligase (CofE-AF) from *A. fulgidus* and its complex with GDP at 2.5 and 1.35 Å resolution, respectively, were obtained. The structure of CofE-AF reveals a novel protein fold with intertwined butterfly-like dimer formed by two-domain monomers. GDP and Mn(II) are bound within the putative active site in a large groove at the dimer interface. CofE represents the first member of a new structural family of nonribosomal peptide synthases.<sup>83</sup>

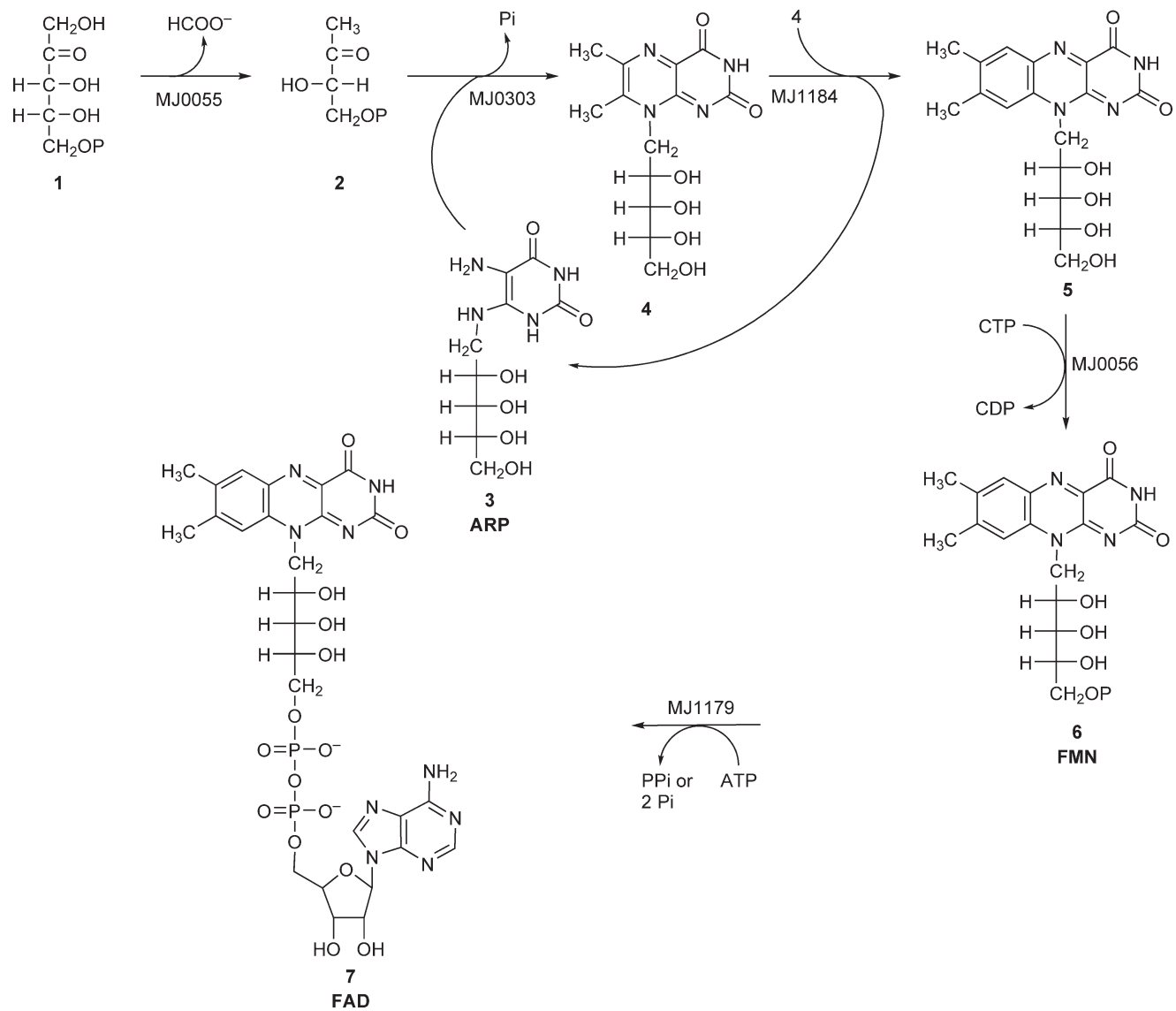
#### 7.20.5 Biosynthesis of Riboflavin, FMN, and FAD

As described above, the steps leading to ARP, the precursor to riboflavin and F<sub>420</sub>, are different in the archaea than all other organisms in that different steps and enzymes are used to generate this centrally important intermediate. The basic chemistry for the subsequent conversion of this compound into riboflavin is universally conserved in all cells and is shown in Figure 18. Key to these reactions are 3,4-dihydroxy-2-





**Figure 17** Pathway for the biosynthesis of F<sub>420</sub>. The F<sub>420-2</sub> was the originally described F<sub>420</sub>. F<sub>420-3</sub> is the structure of the F<sub>420</sub> present in *M. jannaschii*.<sup>72</sup> Other structures with different numbers of glutamates are known.



**Figure 18** Pathway for riboflavin, FMN, and FAD biosynthesis in *Methanococcus jannaschii*.

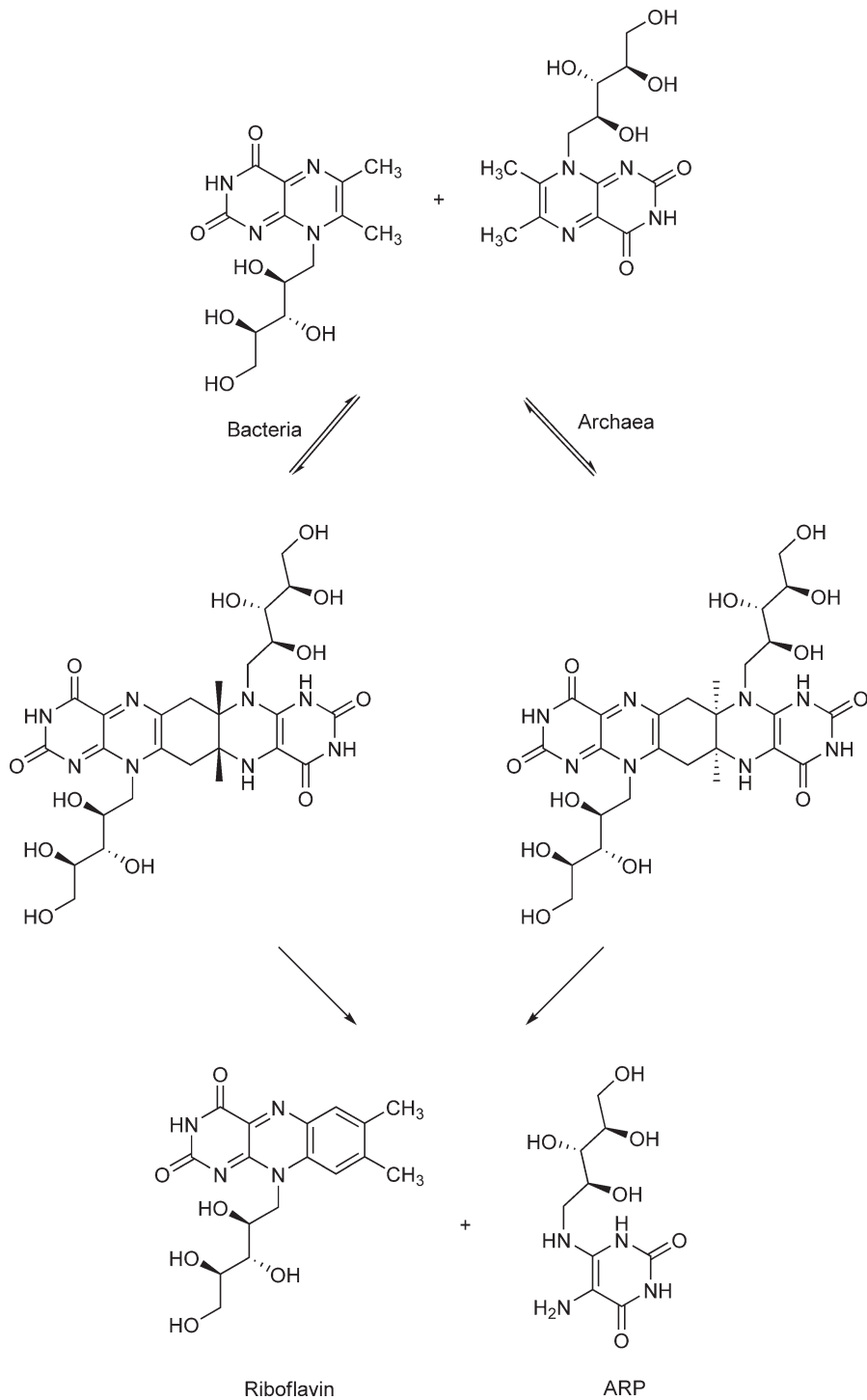
butanone 4-phosphate synthase forming 3,4-dihydroxy-2-butanone 4-phosphate (compound **2**, **Figure 18**) in an amazing enzymatic transformation from ribulose 5-phosphate (compound **1**).<sup>84</sup> We have confirmed that the *M. jannaschii* protein derived from the MJ0055 gene catalyzes this reaction as expected. 6,7-Dimethyl-8-ribityllumazine synthase then catalyzes the condensation of 3,4-dihydroxy-2-butanone 4-phosphate with ARP (compound **3**) to form 6,7-dimethyl-8-ribityllumazine (compound **4**). Riboflavin synthase catalyzes a dismutation of two molecules of 6,7-dimethyl-8-ribityllumazine (compound **4**) to riboflavin (compound **5**) and ARP. It is amazing that these last two steps in riboflavin biosynthesis were demonstrated to occur chemically in work on riboflavin synthesis performed over four decades ago.<sup>85–87</sup>

Two major findings have made it clear that the enzymology behind the formation of riboflavin is quite different in the Archaea compared to all other known organisms. The first was the discovery that the riboflavin synthase functioning in the Archaea is devoid of sequence similarity with those of bacteria and eukaryotes but has significant sequence similarity with the preceding enzyme in the pathway, dimethyl-8-ribityllumazine synthase.<sup>88</sup> In the case of *M. jannaschii*, the MJ0303 gene encodes the dimethyl-8-ribityllumazine synthase<sup>89</sup> and a gene with a similar sequence, MJ1184, encodes the riboflavin synthase.<sup>90</sup> Thus a homologue of lumazine synthase has replaced the canonical riboflavin synthase. The *M. jannaschii* dimethyl-8-ribityllumazine synthase was found to have a mass of 1 MDa and the data indicated that it formed hollow capsids with icosahedral 532 symmetry consisting of 60 subunits.<sup>89</sup> The archaeal riboflavin synthase on the other hand was found to be a homopentamer whereas the riboflavin synthase from plants, fungi, and eubacteria are all trimeric proteins.<sup>90</sup> The X-ray structure of the archaeal riboflavin synthase complexed with the substrate analogue inhibitor 6,7-dioxo-8-ribityllumazine showed five active sites located between the adjacent monomers of the pentamer.<sup>91</sup> Recently, a new group of lumazine synthases (type II lumazine synthases), first identified in *Brucella* spp., have been characterized as homodecamers that are best described as a D5-symmetric dimer of pentamers.<sup>92,93</sup>

The second finding was that the pentacyclic intermediates formed by the trimeric *E. coli* enzyme and the pentameric *M. jannaschii* riboflavin synthase are diastereomers, as shown in **Figure 19**. These pentacyclic diastereomers are formed by the condensation of two molecules of 6,7-dimethyl-8-ribityllumazine with the same regiochemistry in both enzymes.<sup>94</sup> This difference clearly results from the fact that these are different enzymes with different active sites. Each of the diastereomers is a catalytically competent intermediate for its respective enzyme but does not serve as a substrate for the other enzyme.<sup>95</sup>

Phylogenetic analysis of the quaternary arrangements and sequences of lumazine synthases and related pentameric riboflavin synthases derived from archaea, bacteria, plants, and fungi suggests a family of proteins composed of four different clades. These include the archaeal lumazine and riboflavin synthases, type I lumazine synthases, and the eubacterial type II lumazine synthases.<sup>96</sup> We could thus construct a picture of the evolution of the latter steps of riboflavin biosynthesis where initially there were two enzymes catalyzing the riboflavin synthase reaction. The archaeal riboflavin synthase evolved into the different lumazine synthases and the other enzyme evolved into the trimeric enzymes found in present day eubacterial and eukaryotic riboflavin synthases. This would explain how the two different enzymes can use two different diastereomeric intermediates in the generation of riboflavin. That the lumazine synthase evolved from the riboflavin synthases is supported by the finding that it still binds riboflavin.<sup>97</sup> Perhaps the fact that two different enzymes have evolved for this last step in riboflavin reflects on the ease at which riboflavin is formed nonenzymatically.<sup>85–87</sup>

One of the many defining characteristics of flavin biosynthesis in Archaea is the absence of canonical genes encoding riboflavin kinase and FAD synthase that are required for the conversion of riboflavin into the coenzymes FMN and FAD. Genes coding for nonorthogonal replacement of these enzymes must be present in archaea. Archaeal enzymes requiring either or both FMN and FAD exist and both of these coenzymes are present in cell extracts of archaea (R. H. White, unpublished results). After intense research, we have finally identified the riboflavin kinase in *M. jannaschii* as the product of the MJ0056 gene.<sup>98,99</sup> This finding was confirmed by other workers.<sup>99</sup> Recombinant expression of the MJ0056 gene in *E. coli* led to a large increase in the amount of FMN in the *E. coli* cell extract. The unexpected features of the purified recombinant enzyme were its preference for CTP as the source of the phosphoryl donor and the absence of any requirement for added metal ion to catalyze the formation of FMN. The enzyme represents a unique CTP-dependent family of kinases. We have also recently identified the archaeal FAD synthetase that is responsible for transferring



**Figure 19** Structures of the two diastereomeric pentacyclic intermediates produced by riboflavin synthase in bacteria and archaea.

AMP from ATP to FMN. This archaeal FAD synthetase is coded by the MJ1179 gene in *M. jannaschii* (Z. Mashhadi, unpublished). Identification of this riboflavin kinase and FAD synthetase fills another gap in the unique archaeal flavin biosynthetic pathway.

## 7.20.6 Biosynthesis of Coenzyme M and Coenzyme B

### 7.20.6.1 Introduction

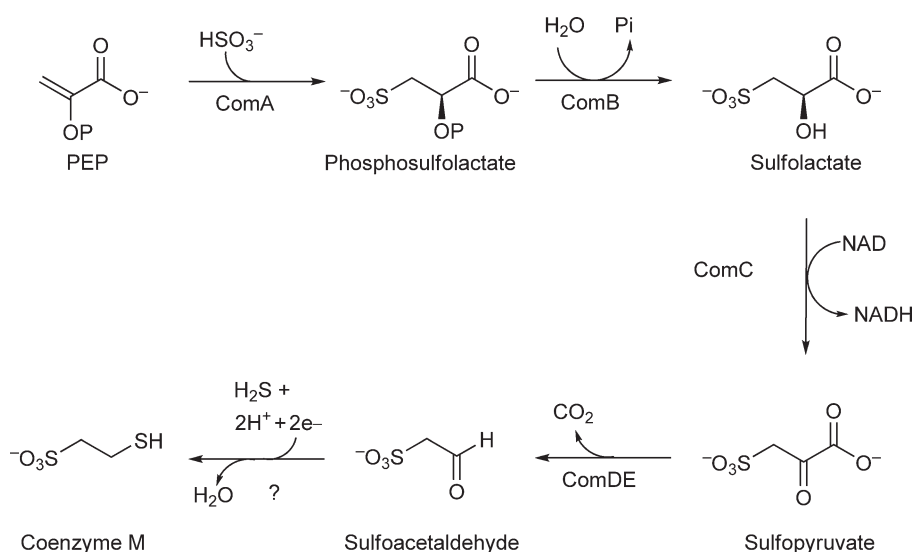
Coenzymes M and B are involved in the final steps of methane formation involving the reduction of methyl-coenzyme M to methane (**Figure 1**). This reaction is catalyzed by methyl-coenzyme M methylreductase, a nickel-dependent oxidoreductase. This enzyme catalyzes the reaction of methyl-coenzyme M with coenzyme B to form methane and a heterodisulfide between coenzymes B and M. The final step in methanogenesis results in the reduction of this heterodisulfide back to coenzymes M and B. This reaction is catalyzed by a heterodisulfide reductase that unexpectedly contains an iron–sulfur cluster.<sup>100</sup>

The pathway for the biosynthesis of coenzymes M and B are shown in **Figures 20 and 21**, respectively. The interesting biochemical connection between these two structurally different coenzymes rests in the biochemistry of the formation of their thiol groups, each of which arises from an aldehyde-containing precursor (**Figure 22**). We discuss the recently published information on the biosynthesis of these two coenzymes and then address a possible common biochemical mechanism for the formation of their thiol groups.

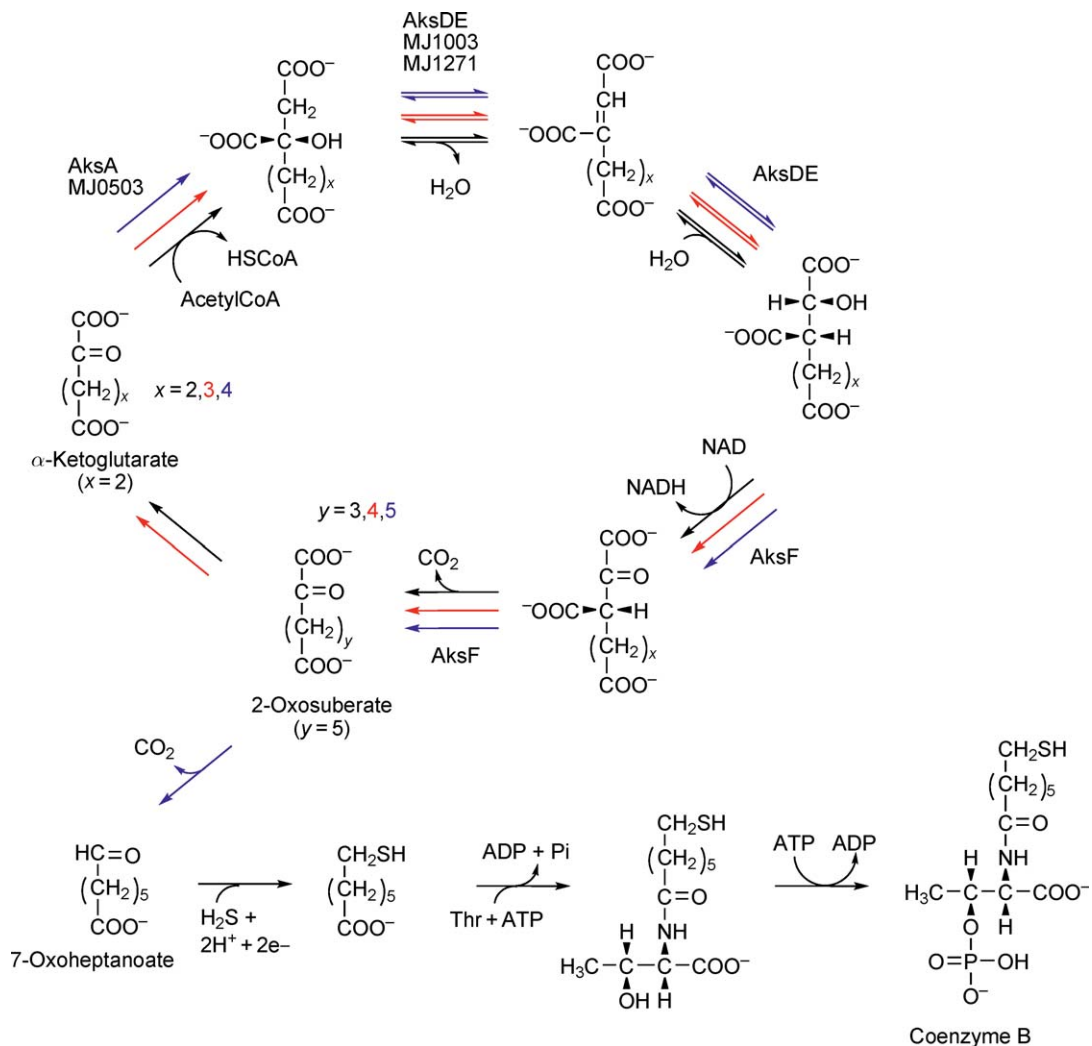
### 7.20.6.2 Biosynthesis of Coenzyme M

Phosphosulfolactate synthase (ComA), catalyzes the first step in coenzyme M biosynthesis (**Figure 20**), which is the stereospecific Michael addition of sulfite to phosphoenolpyruvate to form phosphosulfolactate as shown in **Figure 23**. On the basis of the reaction catalyzed, ComA was predicted to be related to the enolase superfamily of enzymes; however, the overall structure and active-site architecture of ComA are unlike any member of the enolase superfamily. This suggests that ComA is not a member of the enolase superfamily but instead acquired an enolase-type mechanism through convergent evolution.<sup>101</sup>

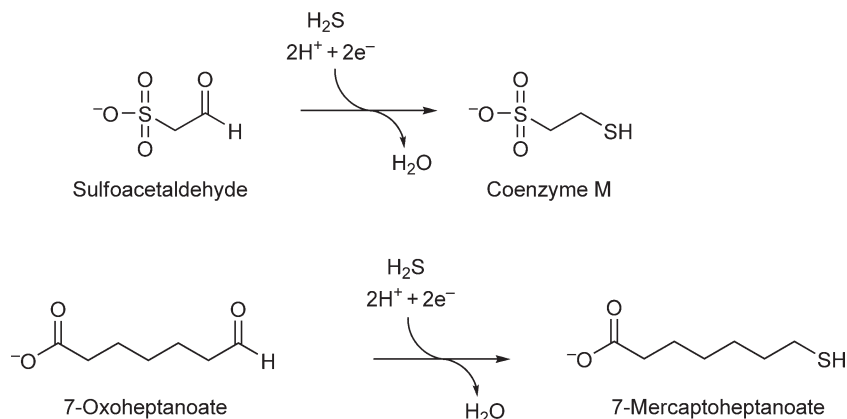
The next step in coenzyme M formation is the dephosphorylation of phosphosulfolactate by a Mg(II)-dependent acid phosphatase, ComB.<sup>102</sup> The third enzyme in the pathway, sulfolactate dehydrogenase (ComC), has also been structurally characterized with the bound reaction product NADH.<sup>103</sup> ComC is present in solution as a dimer, and in the crystal the asymmetric unit contains a tetramer of tight dimers. The dimer is the enzymatically active unit and a portion of each monomer binds NADH at the active site. As a result of this interaction, ComC does not contain the classic ‘Rossmann-Fold’ topology for NADH binding but instead defines a novel fold for NADH binding.



**Figure 20** Pathway for the biosynthesis of coenzyme M.

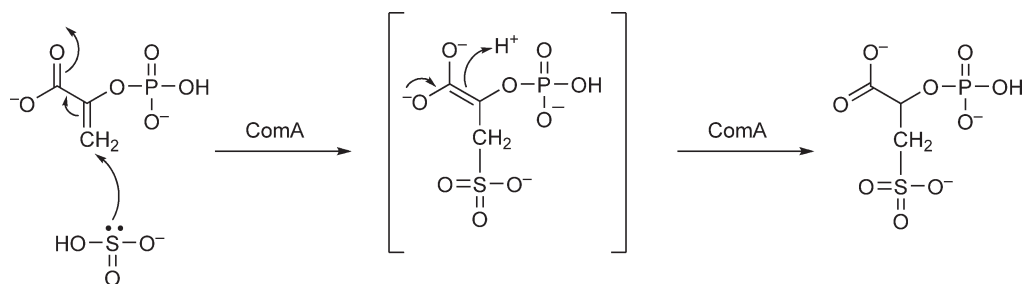


**Figure 21** Pathway for the biosynthesis of coenzyme B. The reactions involved in the biosynthesis of 2-oxosuberate result from three iterations of the 2-oxoacid elongation pathway beginning with  $\alpha$ -ketoglutarate. The black arrows indicate the first cycle, while the red and blue indicate the second and third, respectively.



**Figure 22** Reactions leading to the formation of the thiol groups in coenzymes B and A.





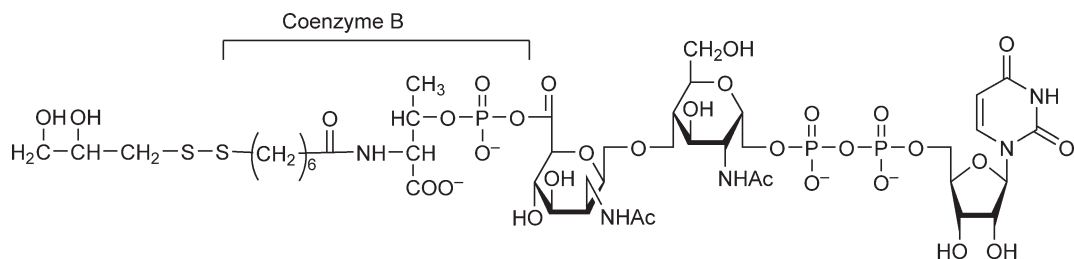
**Figure 23** Mechanism of ComA.

The next step in the biosynthesis of coenzyme M is the decarboxylation of sulfopyruvate by ComDE to form sulfoacetaldehyde. The enzyme(s) catalyzing the last and chemically most interesting step(s) in the biosynthesis of coenzyme M is presently unknown (Figure 20).

### 7.20.6.3 Biosynthesis of Coenzyme B

A major component of coenzyme B is 7-mercaptoheptanoate (Figure 22). This acid is produced by the decarboxylation of 2-oxosuberate to generate 7-oxoheptanoate (Figure 21)<sup>23,104</sup> and the subsequent conversion of the produced aldehyde into the thiol group as shown below. The reactions involved in the biosynthesis of 2-oxosuberate result from three iterations of the 2-oxoacid elongation pathway as occurs in lysine biosynthesis but beginning with  $\alpha$ -ketoglutarate.<sup>23</sup> These reactions require an aconitase-type hydrolyase to catalyze the hydroxyacid isomerization reactions. The genome of *M. jannaschii* contains two homologues of each large (MJ0499 and MJ1003) and small (MJ1271 and MJ1277) subunits of the homoaconitase enzyme that are likely to be involved in the isomerizations required for 2-oxosuberate biosynthesis. These enzymes have been recently heterologously expressed in *E. coli*, purified, reconstituted to generate the Fe-S centers, and tested for activity.<sup>105</sup> Combination of the MJ0499 with the MJ1271 gene products, or the MJ1277 and MJ1003 gene products with the MJ1271 gene product, each copurified as a heterotetramer with two subunits of each of the large and small subunits. Enzymatic analysis showed that only the combination of the MJ0499 and MJ1277 proteins catalyzed the isopropylmalate and citramalate (2-methylmalate) isomerization reactions, which are not related to coenzyme B biosynthesis. On the basis of a phylogenetic analysis of the large subunits it was concluded that the MJ1003 subunit was responsible for homocitrate synthase required for the biosynthesis of coenzyme B.<sup>105</sup> Thus, the current evidence indicates that the enzyme composed of the MJ1003 and the MJ1271 gene products is the desired enzyme for coenzyme B biosynthesis. This awaits experimental confirmation. The remaining reactions for which genes are yet to be identified are the kinase that adds the phosphate to the threonine part of dephosphocoenzyme B and the enzyme condensing the 7-mercaptoheptanoic acid with threonine (Figure 21).

A major concern about the biosynthetic pathway shown in Figure 21 is the possible involvement of the coenzyme B derivative shown in Figure 24, which was identified in *M. thermoautotrophicus*.<sup>106</sup> Although we have not been able to confirm the presence of such a compound in several different methanogens, its presence could indicate that a scaffold may be required for the biosynthesis of coenzyme B. This could explain the low



**Figure 24** Structure of a coenzyme B containing compound identified in *Methanobacterium thermoautotrophicum*.

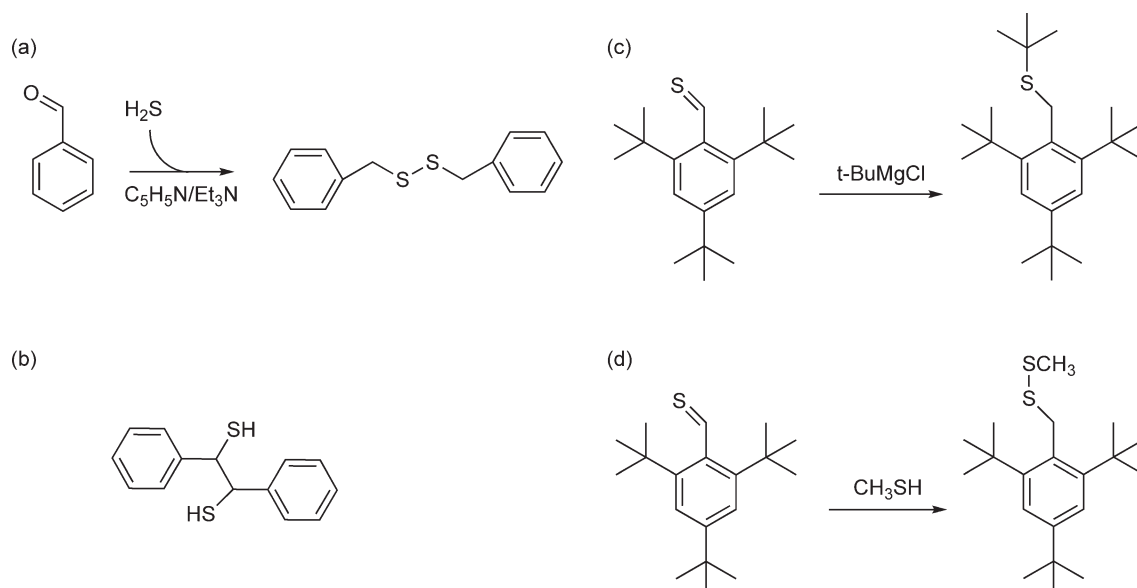
incorporation of labeled threonine into *N*-(7-mercaptoheptanyl)-L-threonine observed in cell extracts of *M. thermophila*.<sup>107</sup> Alternately, this compound could represent an activated intermediate used in the biosynthesis of the N-linked glycan present in methanogens<sup>108</sup> because it contains the same sugar components. The coenzyme B attached to the aminomannuronic acid could be used to activate the coupling of the acid of the aminomannuronic acid to the threonine found in the structure of the N-linked glycan.

#### 7.20.6.4 Formation of the Coenzyme M and Coenzyme B Thiols

A common connection in the biosynthesis of coenzymes M and B is that each involves the conversion of an aldehyde into a thiol.<sup>104,109</sup> In the case of coenzyme M biosynthesis it was originally considered that a thiazolidine adduct with sulfoacetaldehyde was an intermediate in the biosynthesis.<sup>109</sup> Later work has shown that cysteine was readily decomposed to sulfide in *M. jannaschii*,<sup>110</sup> catalyzed by the product of the MJ1025 gene,<sup>111</sup> and that sulfide is the source of sulfur for the synthesis of cysteine and methionine in this methanogen (R. H. White, unpublished). On the basis of these results, identical reactions shown in **Figure 22** most likely account for the last step in the biosynthesis of these two coenzymes. The exact mechanism of this reaction has never been established nor have any of the enzyme(s) or encoding gene(s) involved in carrying out these reactions been identified.

Only a few organic model reactions are known for converting an aldehyde into a thiol. The efficient conversion of aliphatic and aromatic aldehydes into disulfides in the presence of triethylamine and hydrogen sulfide has been described<sup>112</sup> as well as the synthesis of sulfides from thiols and aldehyde or ketones with pyridine-borane in trifluoroacetic acid.<sup>113</sup> Considering that the first example is more biochemically relevant, we have explored the mechanism of how this reaction would proceed with benzaldehyde serving as the aldehyde (**Figure 25(a)**). Benzaldehyde was chosen because it does not contain any hydrogens alpha to the aldehyde, thus preventing the formation of an enol. The desire was that once this mechanism was established it might give us insight into how the biochemical reaction could proceed.

There are three different mechanisms as to how the reaction could proceed: through radicals, hydride transfer, or proton addition. A search for the generation of expected radical coupled products, such as stilbenedithiol (**Figure 25(b)**), proved negative, indicating that radicals were not likely involved. In addition, light had no effect on the reaction. Labeling studies with deuterated precursors showed that the hydrogen incorporated on the aldehyde carbon during reduction came from the sulfide hydrogens. These observations

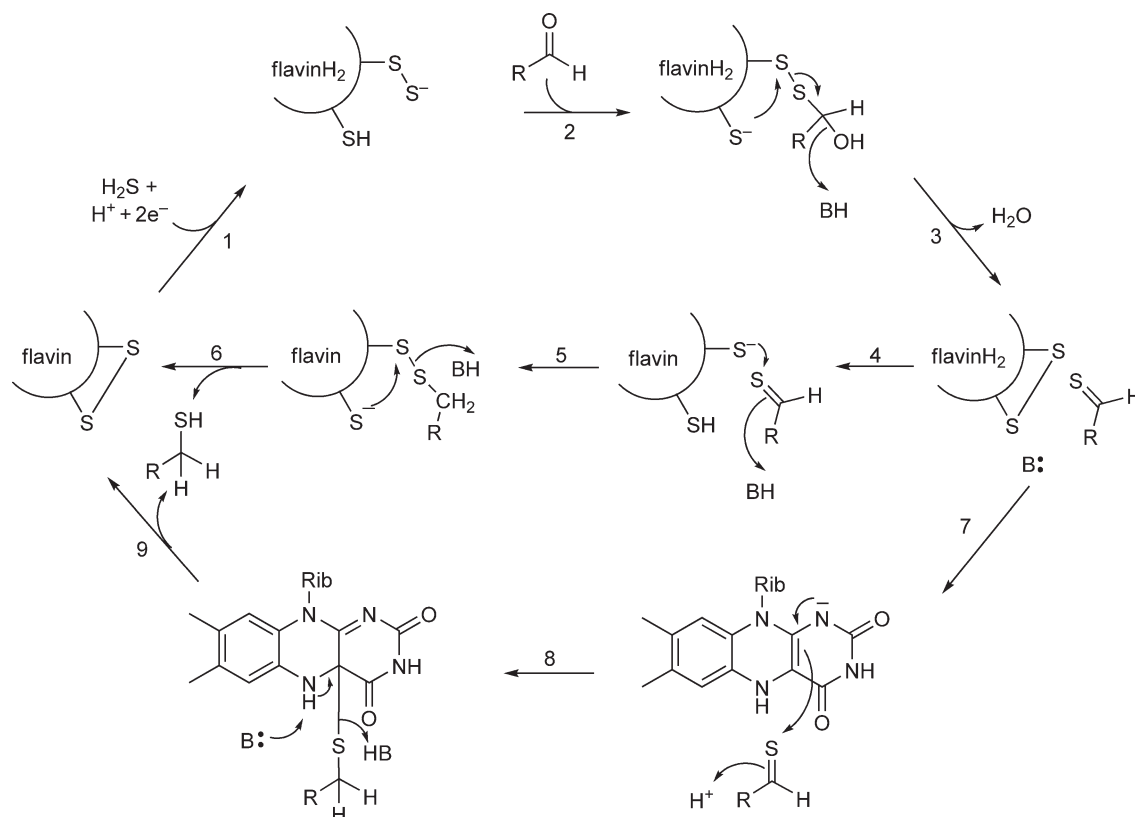


**Figure 25** Model organic reactions for the formation of thiol of the coenzymes A and B.

indicated that hydride reactions probably were not involved in the reaction and that the original aldehyde carbon was likely a carboanion at some point during the reaction. This led to the consideration that the first step in the reaction would be the formation of a thiocarbonyl by the addition of sulfide to the aldehyde followed by elimination of water. The involvement of a thiocarbonyl in the conversion of aldehydes and ketones into thiols has been proposed.<sup>114</sup> Reduction of the thioaldehyde would generate the product. Thiol aldehydes are known not to be stable and readily polymerize. The first stable thioaldehyde, 2,4,6-tri-*t*-butylthiobenzaldehyde (Figures 25(c) and 25(d)), has been prepared and some of its reactions have been studied. It can react at the sulfur with *t*-butylmagnesium chloride to generate the addition product shown in Figure 25(c).<sup>115</sup> We have prepared 2,4,6-tri-*t*-butylthiobenzaldehyde<sup>116</sup> and have demonstrated that it can also undergo the addition of methylthiol to form the mixed disulfide (Figure 25(d)).

Considering all of these observations, a possible mechanism for this interesting biochemical reaction is shown in Figure 26. Central to this proposed mechanism is the involvement of a persulfide (R-S-SH), a thioaldehyde and a protein-bound disulfide in the proposed reaction mechanism. Persulfides are known to be widely involved in the biosynthesis of many important sulfur containing metabolites.<sup>117</sup> Here they function as a protein-bound nucleophilic sulfur that can add to an aldehyde and be reductively released from the protein to form the thioaldehyde (reactions 2–3, Figure 26). The persulfide could be formed by the transfer of hydrogen sulfide to a disulfide (reaction 1, Figure 26).

The known occurrences of thioaldehydes in biochemistry are few. One well-studied example is the involvement of a thioaldehyde in the decarboxylation of cysteine in phosphopantothoenyl-cysteine during coenzyme A biosynthesis.<sup>118,119</sup> In the proposed mechanism for this decarboxylation, a thioaldehyde is generated at the cysteine sulfur by a flavin-dependent oxidation of the thiol. The resulting  $\beta$ -thio keto acid, acting like a  $\beta$ -keto acid, facilitates the decarboxylation of the amide-bound cysteine in phosphopantothoenyl-cysteine substrate. Finally, the flavinH<sub>2</sub> produced in the thiol oxidation is used to reduce the thioaldehyde back to the thiol.



**Figure 26** Proposed mechanism for the enzymatic formation of the thiol of coenzymes M and B.

In the second phase of the enzymatic reaction to form the thiol of coenzymes M and B the thioaldehyde must be reduced to a thiol. This could occur by two separate mechanisms, both of which use the reducing power of flavinH<sub>2</sub> to carry out the reduction. The difference between these two mechanisms is that in the first the flavinH<sub>2</sub> reduces the disulfide to two thiols that in turn reduce the thioaldehyde (reactions 4–6 in **Figure 26**) whereas in the second the flavinH<sub>2</sub> directly reduces the thioaldehyde (reactions 7–9 in **Figure 26**). The mechanism for the direct flavinH<sub>2</sub>-dependent reduction of the thioaldehyde intermediate could be the nucleophilic addition of the reduced flavin (reaction 8) to the thiocarbonyl sulfur to form the 4a adduct to the flavin.<sup>120</sup> Base abstraction of the NH hydrogen and electron delocalization would then generate the thiol (reaction 9). This mechanism would be analogous to the proposed nucleophilic addition of thiol anion to the thioaldehyde sulfur during the reduction (reaction 5). However, not until the actual enzyme(s) involved in this new route to thiols are identified and studied will the true mechanism be revealed.

## 7.20.7 Biosynthesis of Corrinoids and Coenzyme F<sub>430</sub>

### 7.20.7.1 Introduction

Corrinoids are not typically identified as methanogenic cofactors; however, they are important cofactors for several methanogenic enzymes. N<sup>5</sup>-Methyltetrahydromethanopterin: coenzyme M methyltransferase is a corrinoid-dependent enzyme that catalyzes one of the steps in methanogenesis (step 6 + 7, **Figure 1**). Although the most important substrates for methane formation are H<sub>2</sub> + CO<sub>2</sub>, acetate, and methanol, in some marine and brackish environments, methylamines and methylthiol also represent important methanogenic precursors. Archaea also utilize corrinoid-dependent methyltransferases to form methyl-coenzyme M from these methylated compounds (step 10, **Figure 1**). A second nickel-containing tetrapyrrole, coenzyme F<sub>430</sub>, is required by methyl-coenzyme M reductase, which catalyzes the last step in methanogenesis – the reduction of methyl-coenzyme M to methane and coenzyme M (step 8, **Figure 1**). The integral importance of corrinoids and coenzyme F<sub>430</sub> to methanogenesis justifies a discussion of their biosynthesis here.

### 7.20.7.2 Overview of Archaeal Corrinoid Biosynthesis

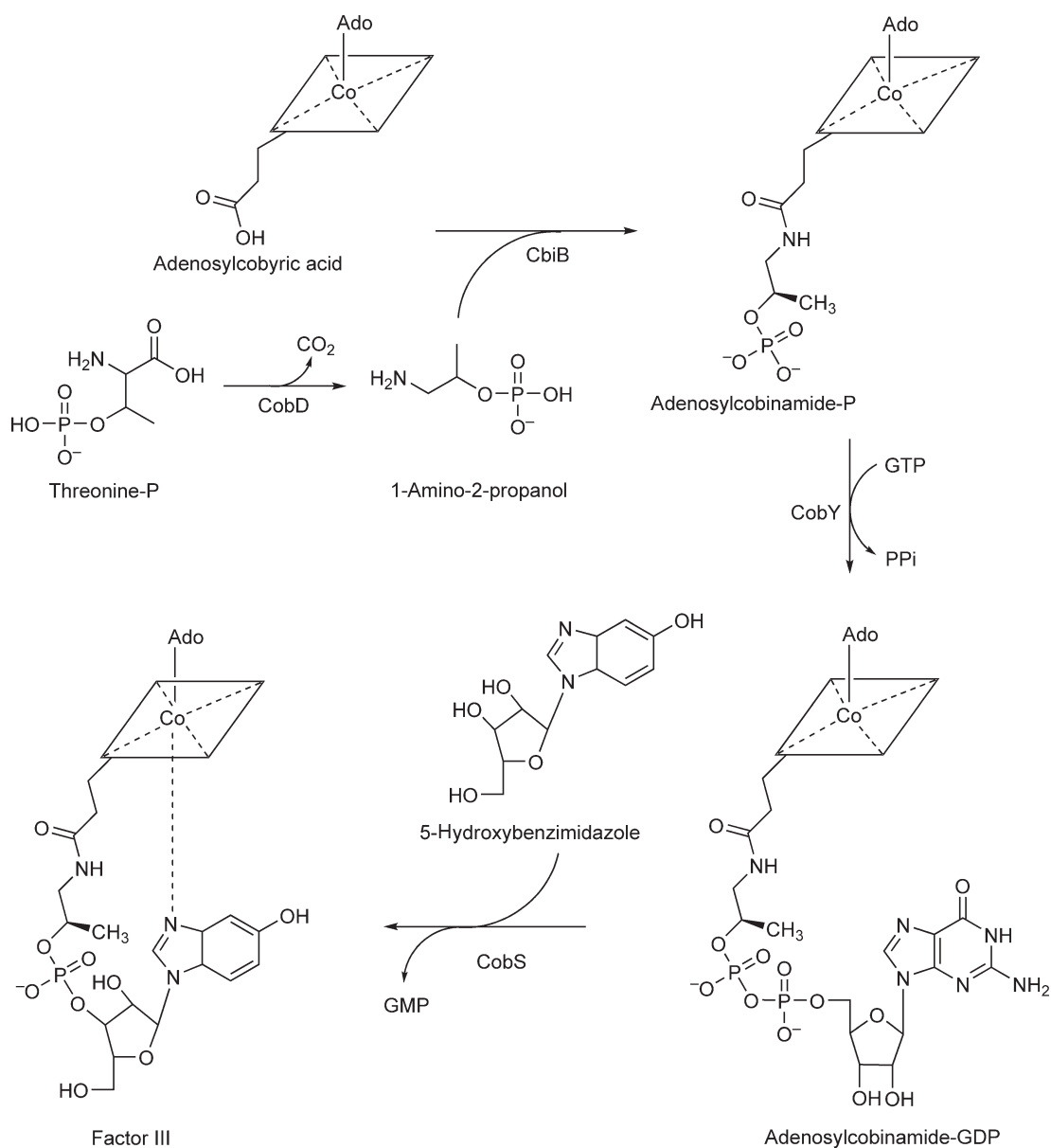
The biosynthesis of corrinoids in archaea is not as well studied as it is in the bacteria. Most of the work done with coenzyme B<sub>12</sub> biosynthesis has been undertaken in *Salmonella typhimurium* LT2, *Propionibacterium freundenreichii* subsp. *sbermanii* and *Bacillus megaterium* as well as the aerobic *Pseudomonas denitrificans*. Work in archaeal corrinoid biosynthesis has been done in *Halobacterium* NRC-1, *M. mazei*, as well as *M. thermoautotrophicus*.<sup>121</sup> Genomic analysis indicates that *M. jannaschii* apparently contains most of the genes necessary for the biosynthesis of corrinoids through the established anaerobic pathway, though several key biosynthetic transformations have yet to be elucidated.<sup>122,123</sup>

Several archaeal corrinoid biosynthetic enzymes have been studied in detail. These include the precorrin 6-X reductase (CbiJ),<sup>124</sup> a C<sub>20</sub> cobalt factor II methyltransferase (CbiL),<sup>125</sup> adenosylcobinamide-phosphate nucleotidyltransferase (CobY),<sup>126,127</sup> adenosylcobyrinic acid synthase (CbiP),<sup>128</sup> and adenosylcobinamide-phosphate synthase (CbiB),<sup>128</sup> a proposed ‘ancestral’ cobaltochelate (CbiX<sup>S</sup>),<sup>129,130</sup> and the last enzyme in the pathway, cobalamin synthase (CobS).<sup>131</sup> Work has also been done on the elucidation of the archaeal cobamide salvage pathway including the characterization of adenosylcobinamide aminohydrolase (CbiZ).<sup>132</sup>

### 7.20.7.3 Formation of 1-Amino-2-Propanol

One aspect of archaeal corrinoid biosynthesis that remains to be verified is the formation of the 1-amino-2-propanol linker in the lower ligand loop of the corrinoid. Possible sources include the reductive amination of lactaldehyde, the reductive amination of pyruvate, or the decarboxylation of threonine. Several bacteria have been shown to directly incorporate <sup>15</sup>N-labeled threonine into aminopropanol and coenzyme B<sub>12</sub>. A threonine decarboxylase (CobD) has been detected in *Streptomyces griseus*<sup>133</sup> and *Salmonella typhimurium*.<sup>134</sup> Studies with CobD from *S. typhimurium* LT2 were able to clearly demonstrate the decarboxylation of threonine-phosphate; however, the timing of the decarboxylation could not be demonstrated and may occur prior to or after addition of the threonine-P to the corrin ring.<sup>134</sup>

Experiments with *M. jannaschii* cell free extracts in our lab have demonstrated that these cells contained measurable levels of free aminopropanol and that while threonine and ATP did not give rise to aminopropanol or aminopropanol-phosphate, incubation with threonine-phosphate did produce aminopropanol-phosphate. The conversion of labeled lactaldehyde into aminopropanol was also observed but at 4% of the activity seen with threonine-phosphate. These data indicate that threonine-phosphate is the precursor to aminopropanol in *M. jannaschii* as shown in **Figure 27**; however an archaeal threonine-phosphate decarboxylase has not yet been identified. *Metbanococcus jannaschii* contains a gene with significant similarity to the *Salmonella* CobD, MJ0955. This gene is annotated as a histidinol-phosphate aminotransferase (hisC) and has 24% identity and 46% positive sequence alignment with the *Salmonella* CobD. The MJ0955 gene product, HisC, was found not to catalyze the formation of aminopropanol from either threonine or threonine-phosphate. *Metbanococcus jannaschii* HisC was able to catalyze the transamination of tyrosine (unpublished results), as was observed with hisC.<sup>135</sup>



**Figure 27** The final steps in the proposed archaeal cobamide biosynthetic pathway.

These results suggest that MJ0955 is a hisC and probably performs the annotated role in histidine biosynthesis. Thus, the identity of the threonine-phosphate decarboxylase remains to be determined. The possible involvement of a pyruvoyl-dependent enzyme in catalyzing this reaction should be considered.

In at least one archaea, *M. thermoautotrophicus*, threonine does not appear to be the precursor to the aminopropanol moiety of cobamide. Labeling patterns of aminopropanol in this organism were found to be similar to pyruvate or lactate and not threonine, leading Eisenreich and Bacher to propose a reductive amination pathway from pyruvate for aminopropanol formation.<sup>136</sup>

#### 7.20.7.4 Attachment of 1-Amino-2-Propanol

Another issue that has not been clearly resolved in the literature is whether aminopropanol or aminopropanol-phosphate is the substrate for attachment to the corrin ring by CbiB. The group of Escalante-Semerena has proposed that aminopropanol-phosphate is the substrate for CbiB in *Halobacterium* NRC-1.<sup>128</sup> This scenario is consistent with the use of threonine-phosphate as a precursor of aminopropanol that we observed in *M. jannaschii* cell-free extracts, as the product of the decarboxylation would be aminopropanol-phosphate and no additional kinase would be required. In a series of knockout and complementation studies, Escalante-Semerena's group has demonstrated that CbiB is also required for cobinamide salvage in *Halobacterium* NRC-1 and that this function can be restored by complementation with the corresponding *Methanococcus mazei* gene. Woodson *et al.*<sup>127</sup> propose that an unidentified hydrolase may be involved in adenosylcobinamide salvage by removal of the nonphosphorylated aminopropanol component of adenosylcobinamide to generate the proposed CbiB substrate, adenosylcobyrinic acid.

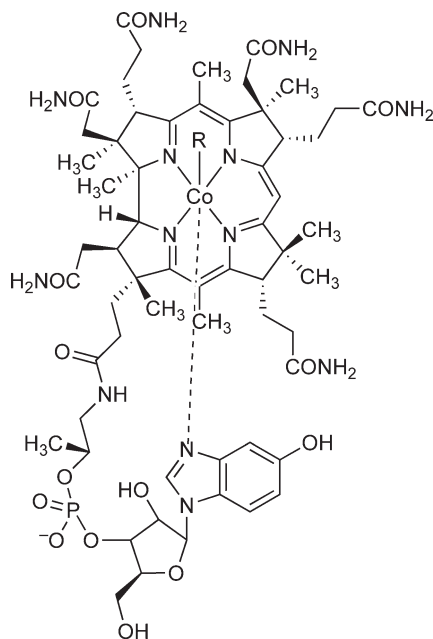
#### 7.20.7.5 Formation of the Benzimidazole Moiety

One difference between the aerobic and anaerobic corrinoid pathways in bacteria is in the formation of the benzimidazole moiety, the 'lower' or ' $\alpha$ -' cobalt ligand. It was known that in anaerobes the 5,6-dimethylbenzimidazole is derived from a flavin, but the nature of the transformation and the enzyme involved was only recently discovered.<sup>137,138</sup>

Archaea have been found to contain a variety of modified cobamides with alternate lower ligands. Analysis of the cobamides isolated from methanogenic archaea has shown that the coenzyme B<sub>12</sub> analogues Co $\alpha$ -[ $\alpha$ -(7-adenyl)]-cobamide (pseudo vitamin B<sub>12</sub>) and Co $\alpha$ -[ $\alpha$ -(5-hydroxybenzimidazolyl)]-cobamide (factor III, **Figure 28**) are the predominant corrinoids. 5-Methylbenzimidazole was found to be the  $\alpha$ -ligand in the cobamide analogue of the acetate metabolizing *Methanotherobacter soehngenii*; however, this strain was later found not to be a pure culture,<sup>139,140</sup> so the source of the 5-methylbenzimidazole is not certain.

The archaeal corrinoid factor III contains 5-hydroxybenzimidazole, that has been shown to be a precursor to the common bacterial ligand, 5,6-dimethylbenzimidazole in the anaerobic corrinoid biosynthetic pathway.<sup>141</sup> Extensive labeling studies in the anaerobe *Eubacterium limosum* has revealed the precursors of 5,6-dimethylbenzimidazole to be erythrose, formate, glutamine, glycine, and methionine.<sup>142</sup> The typical biosynthetic origin of erythrose is the nonoxidative pentose-phosphate pathway; however, many archaea do not possess gene homologues for known pentose-phosphate pathway enzymes and the operation of this pathway in the archaea has been questioned.<sup>143-145</sup> *M. thermoautotrophicus* has been shown to contain corrinoid factor III,<sup>139</sup> but genomic analysis failed to identify a complete set of genes for the nonoxidative pentose-phosphate pathway.<sup>143</sup> Erythrose is also an intermediate in canonical pathways for AroAA and ribose biosynthesis; however, archaea have been found to utilize an alternate pathway that does not require erythrose.<sup>41,144</sup> In addition, an analysis of *M. jannaschii* cell free extract failed to detect the presence of erythrose-4-phosphate. Collectively, these results suggest that erythrose is not an archaeal metabolite and, consequently, an alternate biosynthetic route to the 5-hydroxybenzimidazole must be considered. Evidence for the presence of an alternate archaeal pathway to 5-hydroxybenzimidazole has been obtained through labeling studies in *M. barkeri* and *M. thermoautotrophicus*.<sup>136,142</sup> The hydroxybenzimidazole moiety of factor III from *M. thermoautotrophicus* had the same labeling pattern as the purine imidazole ring.<sup>136,142</sup> This further supports the existence of an alternate pathway for the formation of hydroxybenzimidazole.





**Figure 28** Chemical structure of archaeal cobamide factor III.

It is possible that the existence of pseudo B<sub>12</sub> in many methanogens may be a consequence of their inability to produce erythrose and thus the benzimidazole lower ligand. In these cases, an alternate ligand, such as adenosine, may be utilized. Indeed, studies with *M. thermoautotrophicus* showed that it was not only able to incorporate a variety of benzimidazole and purine bases into its corrinoids, but also that its growth rate with the alternate ligands was not significantly affected.<sup>146</sup> This suggests that the absolute identity of the lower ( $\alpha$ ) ligand is not a strict requirement for activity of cobamide-dependent enzymes in the archaea.

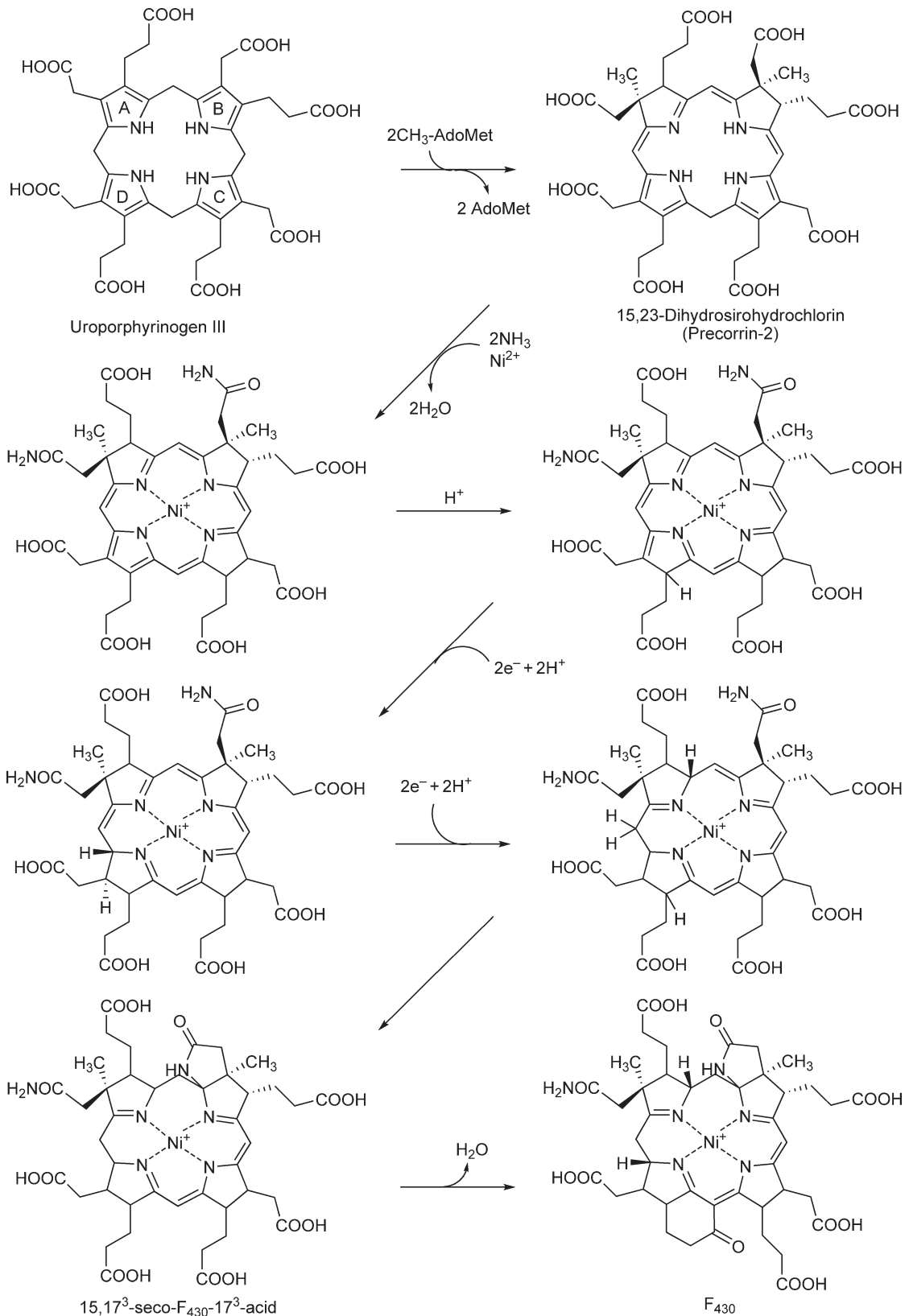
#### 7.20.7.6 Corrinoid Upper Ligand

It is not known if all archaea contain adenylated corrinoids or not, however, because most archaea contain a corrinoid adenosyl transferase (CobA) orthologue it is presumed that the corrin biosynthetic intermediates are adenylated.<sup>128</sup> The archaeal ATP:Co(I)rrinoid adenosyltransferase from *M. mazei* strain Göl has been cloned and characterized.<sup>147</sup> Unlike the bacterial enzyme, the *M. mazei* CobA was found to prefer cobalamin as a substrate over cobinamide, have increased selectivity for ATP over other nucleotides, and was able to utilize 2-deoxynucleotides (dCTP) as well as ribonucleotides.<sup>147</sup> As it is the methylated form of cobalamin, not the adenylated form, that is involved in methanogenesis,<sup>123</sup> a possible role of the adenylated form of the coenzyme has not been determined.

Although the characterization of a corrinoid adenyltransferase provides evidence for the adenylated coenzyme in some archaea, many archaea do not possess homologues of *cobA*. A comparison of the *M. jannaschii* genome with that of *B. megaterium* noted that the *btuR* (equivalent to *cobA* in *Salmonella*) gene is absent, indicating that either *M. jannaschii* contains a nonorthogonal replacement of CobA or the corrinoid in *M. jannaschii* is not adenylated.<sup>123</sup>

#### 7.20.7.7 Biosynthesis of Coenzyme F<sub>430</sub>

Coenzyme F<sub>430</sub> is a nickel-containing tetrapyrrole that is required by methyl-coenzyme M reductase (Figure 29). It is one of a class of nickel chlorins found in nature.<sup>148</sup> Methyl-coenzyme M reductase catalyzes the last step in methanogenesis, the reduction of methyl-coenzyme M to methane and coenzyme M (step 8, Figure 1). The biosynthesis of F<sub>430</sub> follows the common tetrapyrrole biosynthetic pathway from glutamate to



**Figure 29** A proposed pathway for coenzyme F<sub>430</sub> biosynthesis. The order of the reactions is still not known; however, the required transformations are shown.

15,23-dihydrosirohydrochlorin (precorrin-2), the last common intermediate in both corrinoid and F<sub>430</sub> pathways. The intermediacy of both uroporphyrinogen III and 15,23-dihydrosirohydrochlorin is supported by the observed conversion of each to F<sub>430</sub> in *M. thermoautotrophicus*.<sup>149,150</sup> The subsequent conversion of 15,23-dihydrosirohydrochlorin into F<sub>430</sub> has not been studied in detail; however, Thauer has proposed a pathway similar to that shown in **Figure 29**. The individual reactions shown in **Figure 29** were predicted based on the required chemical transformations needed to convert 15,23-dihydrosirohydrochlorin into F<sub>430</sub>; however, the order of the reactions is not known.<sup>151</sup> The formation of 15,17<sup>3</sup>-seco-F<sub>430</sub>-17<sup>3</sup>-acid as well as the subsequent cyclization of the propionic acid group has been demonstrated in several methanogenic archaea.<sup>152</sup> This observation suggests that the formation of the six-membered ring, fused to pyrrole ring D, occurs after the required amidations, acetamide cyclization, and nickel insertion as shown in **Figure 29** and is one of the final steps in F<sub>430</sub> biosynthesis.

## 7.20.8 Concluding Remarks

Methanogens inhabit a wide range of anaerobic habitats from hydrothermal vents to the gastrointestinal tracts of animals, and have thus evolved to grow under extreme environmental conditions. Methanogenesis is a key component of the global carbon cycle and methanogens produce an estimated 75–85% of all methane emissions.<sup>153,154</sup> Consequently, an understanding of the chemistry and biochemistry of methanogenic coenzymes is pivotal for our understanding of global carbon flux. Understanding the details of coenzyme biosynthesis also gives us new insights into the chemical origins of metabolism. Additionally, the encoded enzymes can function as specific drug targets in those few bacteria that produce them. The above studies have shown us that the biosynthesis of methanogenic coenzymes utilizes pathways modified from those producing the widely distributed coenzymes such as FMN, FAD, and B<sub>12</sub> as well as completely new pathways leading to the coenzymes such as F<sub>420</sub>, methanofuran, methanopterin, coenzymes M and B. In the case of F<sub>420</sub> and flavin coenzymes, the biosynthetic pathways branch from a common intermediate. In either case, establishing the enzymes/genes involved in these pathways has and will continue to contribute to the assignment of gene function – an important endeavor in this post genomics age.<sup>155</sup>

## References

1. D. E. Graham; R. H. White, *Nat. Prod. Rep.* **2002**, *19*, 133–147.
2. R. H. White, Biosynthesis of the Methanogenic Cofactors. In *Vitamins and Hormones*; T. Begley, Ed.; Academic Press: New York, 2001; Vol. 61, pp 299–337.
3. R. H. White; D. Zhou, Biosynthesis of the Coenzyme in the Methanogens. In *Methanogenesis Ecology, Physiology, Biochemistry and Genetics*; J. G. Ferry, Ed.; Chapman & Hall: New York, 1993; pp 409–441.
4. J. G. Ferry, *FEMS Microbiol. Rev.* **1999**, *23* (1), 13–38.
5. U. Deppenmeier, *Prog. Nucleic Acid Res. Mol. Biol.* **2002**, *71*, 223–283.
6. J. A. Krzycki, *Curr. Opin. Chem. Biol.* **2004**, *8* (5), 484–491.
7. A. A. DiMarco; T. A. Bobik; R. S. Wolfe, *Annu. Rev. Biochem.* **1990**, *59*, 355–394.
8. L. Chistoserdova; J. A. Vorholt; M. E. Lidstrom, *Genome Biol.* **2005**, *6* (2), 208.
9. S. Shima; R. K. Thauer, *Curr. Opin. Microbiol.* **2005**, *8* (6), 643–648.
10. A. Meyerdierks; M. Kube; T. Lombardot; K. Knittel; M. Bauer; F. O. Glockner; R. Reinhardt; R. Amann, *Environ. Microbiol.* **2005**, *7* (12), 1937–1951.
11. S. J. Hallam; N. Putnam; C. M. Preston; J. C. Detter; D. Rokhsar; P. M. Richardson; E. F. DeLong, *Science* **2004**, *305*, 1457–1462.
12. M. Kruger; A. Meyerdierks; F. O. Glockner; R. Amann; F. Widdel; M. Kube; R. Reinhardt; J. Kahnt; R. Bocher; R. K. Thauer; S. Shima, *Nature* **2003**, *426* (6968), 878–881.
13. R. Hedderich; A. Berkessel; R. K. Thauer, *Eur. J. Biochem.* **1990**, *193* (1), 255–261.
14. E. Muth; E. Morschel; A. Klein, *Eur. J. Biochem.* **1987**, *169* (3), 571–577.
15. J. A. Leigh; K. L. Rinehart, Jr.; R. S. Wolfe, *Biochemistry* **1985**, *24* (4), 995–999.
16. B. K. Pomper; J. A. Vorholt, *Eur. J. Biochem.* **2001**, *268* (17), 4769–4775.
17. C. J. Marx; J. A. Miller; L. Chistoserdova; M. E. Lidstrom, *J. Bacteriol.* **2004**, *186* (7), 2173–2178.
18. N. Kezmarsky; H. Xu; D. Graham; R. White, *Arch. Biochem. Biophys.* **2005**, *435*, 175–182.
19. R. H. White, *J. Bacteriol.* **1988**, *170* (10), 4594–4597.
20. A. Gorkovenko; M. F. Roberts; R. H. White, *Appl. Environ. Microbiol.* **1994**, *60*, 1249–1253.
21. R. H. White, *Biochemistry* **1987**, *26*, 3163–3167.
22. D. M. Howell; K. Harich; H. Xu; R. H. White, *Biochemistry* **1998**, *37* (28), 10108–10117.
23. R. H. White, *Arch. Biochem. Biophys.* **1989**, *270* (2), 691–697.

24. M. Hans; J. Sievers; U. Muller; E. Bill; J. A. Vorholt; D. Linder; W. Buckel, *Eur. J. Biochem.* **1999**, 265 (1), 404–414.
25. B. Buchenau; R. K. Thauer, *Arch. Microbiol.* **2004**, 182 (4), 313–325.
26. H. Xu; R. Aurora; G. D. Rose; R. H. White, *Nat. Struct. Biol.* **1999**, 6 (8), 750–754.
27. R. H. White, *Methods Enzymol.* **1997**, 281, 391–401.
28. D. M. Howell; R. H. White, *J. Bacteriol.* **1997**, 179 (16), 5165–5170.
29. B. El Yacoubi; S. Bonnett; J. N. Anderson; M. A. Swairjo; D. Iwata-Reuyl; V. de Crecy-Lagard, *J. Biol. Chem.* **2006**, 281 (49), 37586–37593.
30. L. L. Grochowski; H. Xu; K. Leung; R. H. White, *Biochemistry* **2007**, 46, 6658–6667.
31. R. H. White, *Arch. Microbiol.* **1985**, 143, 1–5.
32. M. E. Rasche; R. H. White, *Biochemistry* **1998**, 37 (32), 11343–11351.
33. Q. Z. Ye; J. Liu; C. T. Walsh, *Proc. Natl. Acad. Sci. U.S.A.* **1990**, 87 (23), 9391–9395.
34. J. M. Green; W. K. Merkel; B. P. Nichols, *J. Bacteriol.* **1992**, 174 (16), 5317–5323.
35. Z. Chang; Y. Sun; J. He; L. C. Vining, *Microbiology* **2001**, 147 (Pt. 8), 2113–2126.
36. C. von Mering; L. J. Jensen; M. Kuhn; S. Chaffron; T. Doerks; B. Kruger; B. Snel; P. Bork, *Nucleic Acids Res.* **2007**, 35, D358–D362.
37. C. J. Bult; O. White; G. J. Olsen; L. Zhou; R. D. Fleischmann; G. G. Sutton; J. A. Blake; L. M. FitzGerald; R. A. Clayton; J. D. Gocayne; A. R. Kerlavage; B. A. Dougherty; J. F. Tomb; M. D. Adams; C. I. Reich; R. Overbeek; E. F. Kirkness; K. G. Weinstock; J. M. Merrick; A. Glodek; J. L. Scott; N. S. Geoghagen; J. C. Venter, *Science* **1996**, 273 (5278), 1058–1073.
38. P. J. Keller; H. G. Floss; Q. Le Van; B. Schwarzkopf; A. Bacher, *J. Am. Chem. Soc.* **1986**, 108, 344–345.
39. B. Schwarzkopf; B. Reuke; A. Kiener; A. Bacher, *Arch. Microbiol.* **1990**, 153, 259–263.
40. I. Porat; M. Sieprawska-Lupa; Q. Teng; F. J. Bohanon; R. H. White; W. B. Whitman, *Mol. Microbiol.* **2006**, 62 (4), 1117–1131.
41. R. H. White, *Biochemistry* **2004**, 43 (23), 7618–7627.
42. S. Milewski, *Biochim. Biophys. Acta* **2002**, 1597 (2), 173–192.
43. A. Teplyakov; C. Leriche; G. Obmolova; B. Badet; M. A. Badet-Denisot, *Nat. Prod. Rep.* **2002**, 19 (1), 60–69.
44. C.-Y. Teng; B. Ganem; S. Z. Doktor; B. P. Nichols; R. K. Bhatnagar; L. C. Vining, *J. Am. Chem. Soc.* **1985**, 107, 5008–5009.
45. P. C. Raemakers-Franken; R. Bongaerts; R. Fokkens; C. van der Drift; G. D. Vogels, *Eur. J. Biochem.* **1991**, 200 (3), 783–787.
46. J. W. Scott; M. E. Rasche, *J. Bacteriol.* **2002**, 184 (16), 4442–4448.
47. R. V. Dumitru; S. W. Ragsdale, *J. Biol. Chem.* **2004**, 279 (38), 39389–39395.
48. L. Chistoserdova; J. A. Vorholt; R. K. Thauer; M. E. Lidstrom, *Science* **1998**, 281 (5373), 99–102.
49. H. Li; H. Xu; D. E. Graham; R. H. White, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100 (17), 9785–9790.
50. R. H. White, *Biochim. Biophys. Acta* **1998**, 1380 (2), 257–267.
51. B. E. Maden, *Biochem. J.* **2000**, 350 (Pt. 3), 609–629.
52. J. A. Leigh; R. S. Wolfe, *J. Biol. Chem.* **1983**, 258 (12), 7536–7540.
53. J. E. Gready, *Biochemistry* **1985**, 24 (18), 4761–4766.
54. M. A. Caccamo; C. S. Malone; M. E. Rasche, *J. Bacteriol.* **2004**, 186 (7), 2068–2073.
55. J. Schleucher; B. Schworer; C. Zirngibl; U. Koch; W. Weber; E. Egert; R. K. Thauer; C. Griesinger, *FEBS Lett.* **1992**, 314 (3), 440–444.
56. R. White, *Chirality* **1996**, 8, 332–340.
57. C. J. Marx; B. N. O'Brien; J. Breezee; M. E. Lidstrom, *J. Bacteriol.* **2003**, 185 (2), 669–673.
58. E. F. Johnson; B. Mukhopadhyay, *J. Biol. Chem.* **2005**, 280 (46), 38776–38786.
59. H. Seedorf; A. Dreisbach; R. Hedderich; S. Shima; R. K. Thauer, *Arch. Microbiol.* **2004**, 182 (2–3), 126–137.
60. R. K. Seedorf; C. H. Hagemeyer; S. Shima; R. K. Thauer; E. Warkentin; U. Ermler, *FEBS J.* **2007**, 274 (6), 1588–1599.
61. A. P. Eker; P. Kooiman; J. K. Hessels; A. Yasui, *J. Biol. Chem.* **1990**, 265 (14), 8009–8015.
62. R. Michel; C. Massanz; S. Kostka; M. Richter; K. Fiebig, *Eur. J. Biochem.* **1995**, 233 (3), 727–735.
63. J. A. Fox; D. J. Livingston; W. H. Orme-Johnson; C. T. Walsh, *Biochemistry* **1987**, 26 (14), 4219–4227.
64. M. Graupner; H. Xu; R. H. White, *J. Bacteriol.* **2002**, 184 (7), 1952–1957.
65. L. Chatwell; T. Krojer; A. Fidler; W. Romisch; W. Eisenreich; A. Bacher; R. Huber; M. Fischer, *J. Mol. Biol.* **2006**, 359 (5), 1334–1351.
66. P. J. Keller; Q. Le Van; S. U. Kim; D. H. Bown; H. C. Chen; A. Kohnle; A. Bacher; H. G. Floss, *Biochemistry* **1988**, 27 (4), 1117–1120.
67. T. Dams; G. Auerbach; G. Bader; U. Jacob; T. Ploom; R. Huber; R. Jaenicke, *J. Mol. Biol.* **2000**, 297 (3), 659–672.
68. D. E. Graham; H. Xu; R. H. White, *Arch. Microbiol.* **2003**, 180 (6), 455–464.
69. K. P. Choi; N. Kendrick; L. Daniels, *J. Bacteriol.* **2002**, 184 (9), 2420–2428.
70. J. Cheek; J. B. Broderick, *J. Biol. Inorg. Chem.* **2001**, 6 (3), 209–226.
71. H. Li; M. Graupner; H. Xu; R. H. White, *Biochemistry* **2003**, 42 (32), 9771–9778.
72. M. Graupner; R. H. White, *J. Bacteriol.* **2003**, 185 (15), 4662–4665.
73. L. L. Grochowski; H. Xu; R. H. White, *J. Bacteriol.* **2006**, 188, 2836–2844.
74. M. Graupner; H. Xu; R. H. White, *J. Bacteriol.* **2000**, 182 (13), 3688–3692.
75. M. Graupner; R. H. White, *Biochim. Biophys. Acta* **2001**, 1548 (1), 169–173.
76. R. H. White, *Biochemistry* **2008**, 47 (17), 5037–5046.
77. E. Warkentin; B. Mamat; M. Sordel-Klippert; M. Wicke; R. K. Thauer; M. Iwata; S. Iwata; U. Ermler; S. Shima, *EMBO J.* **2001**, 20 (23), 6561–6569.
78. R. H. White; H. Xu, *Biochemistry* **2006**, 45 (40), 12366–12379.
79. L. L. Grochowski; H. Xu; R. H. White, *Biochemistry* **2008**, 47 (9), 3033–3037.
80. M. G. Thomas; J. C. Escalante-Semerena, *J. Bacteriol.* **2000**, 182 (15), 4227–4233.
81. M. Graupner; H. Xu; R. H. White, *Biochemistry* **2002**, 41 (11), 3754–3761.
82. F. Forouhar; M. Abashidze; H. Xu; L. L. Grochowski; J. Seetharaman; M. Hussain; A. Kuzin; Y. Chen; W. Zhou; R. Xiao; T. B. Acton; G. T. Montelione; A. Galinier; R. H. White; L. Tong, *J. Biol. Chem.* **2008**, 283 (17), 11832–11840.
83. B. Nocek; E. Evdokimova; M. Proudfoot; M. Kudritska; L. L. Grochowski; R. H. White; A. Savchenko; A. F. Yakunin; A. Edwards; A. Joachimiak, *J. Mol. Biol.* **2007**, 372 (2), 456–469.
84. G. Richter; M. Kelly; C. Krieger; Y. Yu; W. Bermel; G. Karlsson; A. Bacher; H. Oschkinat, *Eur. J. Biochem.* **1999**, 261 (1), 57–65.

85. T. Rowan; H. C. S. Wood, *Proc. Chem. Soc. (London)* **1963**, 1963, 21–22.
86. K. Kis; K. Kugelbrey; A. Bacher, *J. Org. Chem.* **2001**, *66* (8), 2555–2559.
87. T. Patterson; H. C. S. Wood, *J. Chem. Soc. Commun.* **1969**, 1969, 290–291.
88. S. Eberhardt; S. Korn; F. Lottspeich; A. Bacher, *J. Bacteriol.* **1997**, *179* (9), 2938–2943.
89. I. Haase; S. Mortl; P. Kohler; A. Bacher; M. Fischer, *Eur. J. Biochem.* **2003**, *270* (5), 1025–1032.
90. M. Fischer; A. K. Schott; W. Romisch; A. Ramsperger; M. Augustin; A. Fidler; A. Bacher; G. Richter; R. Huber; W. Eisenreich, *J. Mol. Biol.* **2004**, *343* (1), 267–278.
91. A. Ramsperger; M. Augustin; A. K. Schott; S. Gerhardt; T. Krojer; W. Eisenreich; B. Illarionov; M. Cushman; A. Bacher; R. Huber; M. Fischer, *J. Biol. Chem.* **2006**, *281* (2), 1224–1232.
92. V. Zylberman; P. O. Craig; S. Klinke; B. C. Braden; A. Cauerhff; F. A. Goldbaum, *J. Biol. Chem.* **2004**, *279* (9), 8093–8101.
93. S. Klinke; V. Zylberman; D. R. Vega; B. G. Guimaraes; B. C. Braden; F. A. Goldbaum, *J. Mol. Biol.* **2005**, *353* (1), 124–137.
94. B. Illarionov; W. Eisenreich; N. Schramek; A. Bacher; M. Fischer, *J. Biol. Chem.* **2005**, *280* (31), 28541–28546.
95. M. Fischer; W. Romisch; B. Illarionov; W. Eisenreich; A. Bacher, *Biochem. Soc. Trans.* **2005**, *33*, (Pt. 4), 780–784.
96. V. Zylberman; S. Klinke; I. Haase; A. Bacher; M. Fischer; F. A. Goldbaum, *J. Bacteriol.* **2006**, *188* (17), 6135–6142.
97. M. Fischer; I. Haase; R. Feicht; G. Richter; S. Gerhardt; J. P. Changeux; R. Huber; A. Bacher, *Eur. J. Biochem.* **2002**, *269*, 519–526.
98. Z. Mashhadi; H. Zhang; H. Xu; R. H. White, *J. Bacteriol.* **2008**, *190*, 2615–2618.
99. M. Ammelbur; M. D. Hartmann; S. Djuranovic; V. Alva; K. K. Koretke; J. Martin; G. Sauer; V. Truffault; K. Zeth; A. N. Lupas; M. Coles, *Structure* **2007**, *15* (12), 1577–1590.
100. J. E. Shokes; E. C. Duin; C. Bauer; B. Jaun; R. Hedderich; J. Koch; R. A. Scott, *FEBS Lett.* **2005**, *579* (7), 1741–1744.
101. E. L. Wise; D. E. Graham; R. H. White; I. Rayment, *J. Biol. Chem.* **2003**, *278* (46), 45858–45863.
102. D. E. Graham; M. Graupner; H. Xu; R. H. White, *Eur. J. Biochem.* **2001**, *268* (19), 5176–5188.
103. A. Irimia; D. Madern; G. Zaccai; F. M. Vellieux, *EMBO J.* **2004**, *23* (6), 1234–1244.
104. R. H. White, *Biochemistry* **1989**, *28* (24), 9417–9423.
105. R. M. Drevland; A. Waheed; D. E. Graham, *J. Bacteriol.* **2007**, *189* (12), 4391–4400.
106. F. D. Sauer; B. A. Blackwell; J. K. Kramer; B. J. Marsden, *Biochemistry* **1990**, *29* (33), 7593–7600.
107. B. Solow; R. H. White, *Arch. Biochem. Biophys.* **1997**, *345* (2), 299–304.
108. S. Voisin; R. S. Houlston; J. Kelly; J. R. Brisson; D. Watson; S. L. Bardy; K. F. Jarrell; S. M. Logan, *J. Biol. Chem.* **2005**, *280* (17), 16586–16593.
109. R. H. White, *Biochemistry* **1988**, *27*, 7458–7462.
110. R. H. White, *Biochim. Biophys. Acta* **2003**, *1624* (1–3), 46–53.
111. S.-I. Tchong; H. Xu; R. White, *Biochemistry* **2004**, *44*, 1659–1670.
112. V. Cohen, *Helv. Chem. Acta* **1976**, *59*, 840–844.
113. Y. Kikugawa, *Chem. Lett.* **1981**, 1981, 1157–1158.
114. P. Schneckeburger; P. Adam; P. Albrechi, *Tetrahedron Lett.* **1998**, *39*, 447–450.
115. R. Okazaki, *Chem. Lett.* **1984**, 1984, 101–104.
116. A. Ishii, *Bull. Chem. Soc. Jpn.* **1996**, *69*, 709–717.
117. E. G. Mueller, *Nat. Chem. Biol.* **2006**, *2* (4), 185–194.
118. E. Strauss; T. P. Begley, *J. Am. Chem. Soc.* **2001**, *123* (26), 6449–6450.
119. T. Kupke; W. Schwarz, *J. Biol. Chem.* **2006**, *281* (9), 5435–5444.
120. E. Strauss; T. P. Begley, *Med. Chem. Lett.* **2003**, *13* (3), 339–342.
121. E. Stupperich; I. Steiner; H. J. Eisinger, *J. Bacteriol.* **1987**, *169* (7), 3076–3081.
122. E. Raux; A. Lanois; A. Rambach; M. J. Warren; C. Thermes, *Biochem. J.* **1998**, *335* (Pt. 1), 167–173.
123. E. Raux; A. Lanois; M. J. Warren; A. Rambach; C. Thermes, *Biochem. J.* **1998**, *335* (Pt. 1), 159–166.
124. W. Kim; T. A. Major; W. B. Whitman, *Archaea* **2005**, *1* (6), 375–384.
125. S. Frank; E. Deery; A. A. Brindley; H. K. Leech; A. Lawrence; P. Heathcote; H. L. Schubert; K. Brocklehurst; S. E. Rigby; M. J. Warren; R. W. Pickersgill, *J. Biol. Chem.* **2007**, *282* (33), 23957–23969.
126. M. G. Thomas; J. C. Escalante-Semerena, *J. Bacteriol.* **2000**, *182* (15), 4227–4233.
127. J. D. Woodson; R. F. Peck; M. P. Krebs; J. C. Escalante-Semerena, *J. Bacteriol.* **2003**, *185* (1), 311–316.
128. J. D. Woodson; C. L. Zayas; J. C. Escalante-Semerena, *J. Bacteriol.* **2003**, *185* (24), 7193–7201.
129. A. A. Brindley; E. Raux; H. K. Leech; H. L. Schubert; M. J. Warren, *J. Biol. Chem.* **2003**, *278* (25), 22388–22395.
130. J. Yin; L. X. Xu; M. M. Cherney; E. Raux-Deery; A. A. Brindley; A. Savchenko; J. R. Walker; M. E. Cuff; M. J. Warren; M. N. James, *J. Struct. Funct. Genomics* **2006**, *7* (1), 37–50.
131. L. A. Maggio-Hall; K. R. Claas; J. C. Escalante-Semerena, *Microbiology* **2004**, *150* (Pt. 5), 1385–1395.
132. J. D. Woodson; J. C. Escalante-Semerena, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (10), 3591–3596.
133. A. I. Krasna; C. Rosenblum; D. B. Sprinson, *J. Biol. Chem.* **1957**, *225* (2), 745–750.
134. K. R. Brushaber; G. A. O'Toole; J. C. Escalante-Semerena, *J. Biol. Chem.* **1998**, *273* (5), 2684–2691.
135. F. J. Fernandez; M. C. Vega; F. Lehmann; E. Sandmeier; H. Gehring; P. Christen; M. Wilmanns, *J. Biol. Chem.* **2004**, *279* (20), 21478–21488.
136. W. Eisenreich; A. Bacher, *J. Biol. Chem.* **1991**, *266* (35), 23840–23849.
137. M. J. Gray; J. C. Escalante-Semerena, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (8), 2921–2926.
138. M. E. Taga; N. A. Larsen; A. R. Howard-Jones; C. T. Walsh; G. C. Walker, *Nature* **2007**, *446* (7134), 449–453.
139. H.-P. E. Kohler, *Arch. Microbiol.* **1988**, *150*, 219–223.
140. G. B. Patel, *Int. J. Syst. Bacteriol.* **1992**, *42* (2), 324–326.
141. B. Krautler; J. Moll; R. K. Thauer, *Eur. J. Biochem.* **1987**, *162* (2), 275–278.
142. P. Renz, Biosynthesis of the 5,6-Dimethylbenzimidazole Moiety of Cobalamin and of the Other Based Found in Natural Corrinoids. In *Chemistry and Biochemistry of B<sub>12</sub>*; R. Banerjee, Ed.; John Wiley & Sons, Inc.: New York, 1999.
143. T. Soderberg, *Archaea* **2005**, *1* (5), 347–352.
144. L. L. Grochowski; H. Xu; R. H. White, *J. Bacteriol.* **2005**, *187* (21), 7382–7389.
145. I. Orita; T. Sato; H. Yurimoto; N. Kato; H. Atomi; T. Imanaka; Y. Sakai, *J. Bacteriol.* **2006**, *188* (13), 4698–4704.



146. B. Krautler; H. P. Kohler; E. Stupperich, *Eur. J. Biochem.* **1988**, 176 (2), 461–469.
147. N. R. Buan; K. Rehfeld; J. C. Escalante-Semerena, *J. Bacteriol.* **2006**, 188 (10), 3543–3550.
148. H. L. Sings; K. C. Bible; K. L. Rinehart, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93 (20), 10560–10565.
149. H. Gilles; R. K. Thauer, *Eur. J. Biochem.* **1983**, 135 (1), 109–112.
150. H. Mucha; E. Keller; H. Weber; F. Lingens; W. Trosch, *FEBS Lett.* **1985**, 190 (1), 169–171.
151. R. K. Thauer; L. G. Bonacker, *Ciba Found. Symp.* **1994**, 180, 210–222; discussion 222–227.
152. A. Pfaltz; A. Kobelt; R. Huster; R. K. Thauer, *Eur. J. Biochem.* **1987**, 170 (1–2), 459–467.
153. S. H. Zinder, Physiological Ecology of Methanogens. In *Methanogenesis Ecology, Physiology, Biochemistry & Genetics*; J. G. Ferry, Ed.; Chapman & Hall: New York, 1993; pp 128–206.
154. T. Fenchel; G. M. King; T. H. Blackburn, *Bacterial Biogeochemistry: The Ecophysiology of Mineral Cycling*; Academic Press: San Diego, 1988; p 307.
155. R. H. White, *J. Bacteriol.* **2006**, 188, 3431–3432.

### Biographical Sketches



Laura L. Grochowski was born in Bedford, Pennsylvania, USA, in 1975 and received a BS in Biology from Delaware Valley College in 1997. She studied antibiotic and nonribosomal peptide biosynthesis with Mark Zabriskie at Oregon State University where she obtained a Ph.D. in Medicinal Chemistry in 2004. Laura then joined Bob White's lab in the Department of Biochemistry at Virginia Tech as a postdoctoral associate. Her work focuses on the uncanonical biochemistry and coenzyme biosynthesis of methanogenic archaea.



Robert H. White was born in Oakridge, Tennessee, USA, in 1946 and received his BS in Chemistry from Indiana University, Bloomington, in 1968. He obtained his Ph.D. in Biochemistry from the University of Illinois, Urbana-Champaign, in 1974. After one year postdoctoral positions with Professor Stanley L. Miller and then Professor Trevor McMorris at the University of California, San Diego, he moved to Rice University as a lecturer – spectroscopist. In 1980 he joined the Department of Biochemistry at Virginia Tech, Blacksburg. He has been interested in the biosynthesis of coenzymes and how it relates to the origin of metabolism and life.