Secondary Structure in Protein Analysis

George D. Rose
The Johns Hopkins University, Baltimore, Maryland, USA

Proteins are linear, unbranched polymers of the 20 naturally occurring amino acid residues. Under physiological conditions, most proteins self-assemble into a unique, biologically relevant structure: the native fold. This structure can be dissected into chemically recognizable, topologically simple elements of secondary structure: α-helix, 3_{10}-helix, β-strand, polyproline II helix, turns, and Ω-loops. Together, these six familiar motifs account for $\sim$95% of the total protein structure, and they are utilized repeatedly in mix-and-match patterns, giving rise to the repertoire of known folds. In principle, a protein's three-dimensional structure is predictable from its amino acid sequence, but this problem remains unsolved. A related, but ostensibly simpler, problem is to predict a protein's secondary structure elements from its sequence.

Protein Architecture

A protein is a polymerized chain of amino acid residues, each joined to the next via a peptide bond. The backbone of this polymer describes a complex path through three-dimensional space called the "native fold" or "protein fold."

COVALENT STRUCTURE

Amino acids have both backbone and side chain atoms. Backbone atoms are common to all amino acids, while side chain atoms differ among the 20 types. Chemically, an amino acid consists of a central, tetrahedral carbon atom, $(-\overset{|}{\underset{|}{C}}-)$, linked covalently to (1) an amino group ($-NH_2$), (2) a carboxyl group ($-COOH$), (3) a hydrogen atom ($-H$) and (4) the side chain ($-R$). Upon polymerization, the amino group loses an $-H$ and the carboxy group loses an $-OH$; the remaining chemical moiety is called an "amino acid residue" or, simply, a "residue." Residues in this polymer are linked via peptide bonds, as shown in Figure 1.

DEGREES OF FREEDOM IN THE BACKBONE

The six backbone atoms in the peptide unit $[C\alpha(i)-CO-NH-C\alpha(i+1)]$ are approximately coplanar, leaving only two primary degrees of freedom for each residue. By convention, these two dihedral angles are called ϕ and ψ (Figure 2). The protein's backbone conformation is described by the ϕ,ψ-specification for each residue.

CLASSIFICATION OF STRUCTURE

Protein structure is usually classified into primary, secondary, and tertiary structure. "Primary structure" corresponds to the covalently connected sequence of amino acid residues. "Secondary structure" corresponds to the backbone structure, with particular emphasis on hydrogen bonds. And "tertiary structure" corresponds to the complete atomic positions for the protein.

Secondary Structure

Protein secondary structure can be subdivided into repetitive and nonrepetitive, depending upon whether the backbone dihedral angles assume repeating values. There are three major elements (α-helix, β-strand, and polyproline II helix) and one minor element (3_{10}-helix) of repetitive secondary structure (Figure 3). There are two major elements of nonrepetitive secondary structure (turns and Ω-loops).

REPETITIVE SECONDARY STRUCTURE: THE α-HELIX

When backbone dihedral angles are assigned repeating ϕ,ψ-values near $(-60°, -40°)$, the chain twists into a right-handed helix, with 3.6 residues per helical turn. First proposed as a model by Pauling, Corey, and Branson in 1951, the existence of this famous structure was experimentally confirmed almost immediately by
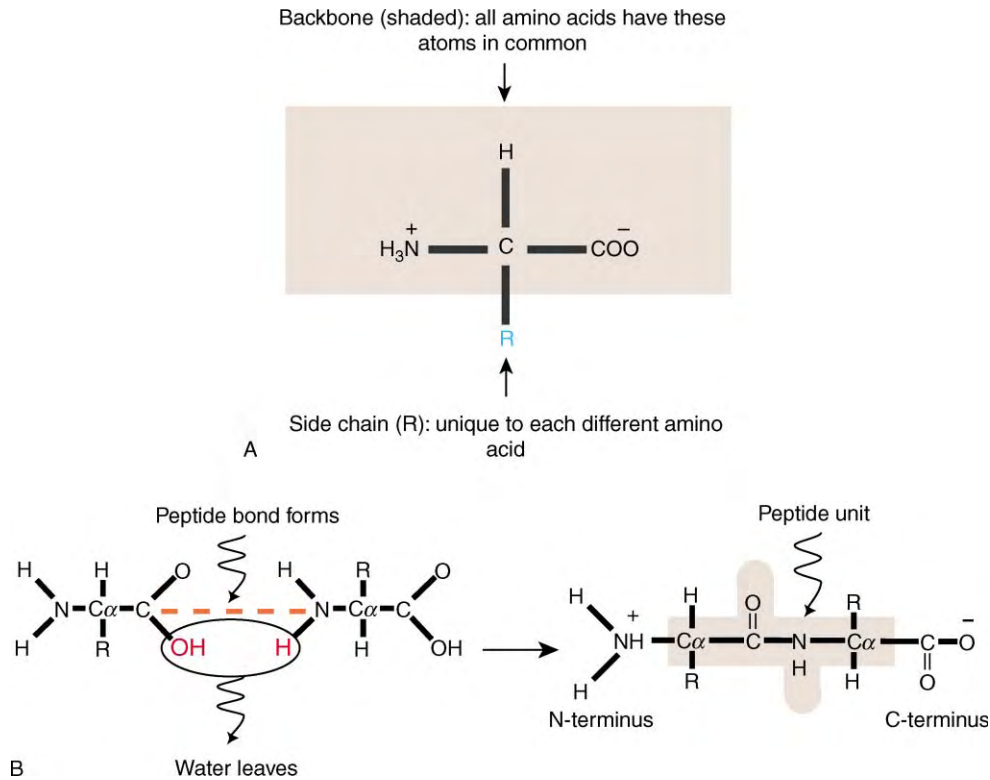
FIGURE 1   (A) A generic amino acid. Each of the 20 naturally occurring amino acids has both backbone atoms (within the shaded rectangle) and side chain atoms (designated R). Backbone atoms are common to all amino acids, while side chain atoms differ among the 20 types. Chemically, an amino acid consists of a tetrahedral carbon atom (−C−), linked covalently to (1) an amino group (−NH$_2$), (2) a carboxyl group (−COOH), (3) a hydrogen atom (−H), and (4) the side chain (−R). (B) Amino acid polymerization. The α-amino group of one amino acid condenses with the α-carboxylate of another, releasing a water molecule. The newly formed amide bond is called a *peptide bond* and the repeating unit is a *residue*. The two chain ends have a free α-amino group and a free α-carboxylate group and are designated the amino-terminal (or N-terminal) and the carboxy-terminal (or C-terminal) ends, respectively. The *peptide unit* consists of the six shaded atoms (Cα−CO−NH−Cα), three on either side of the peptide bond.

Perutz in ongoing crystallographic studies, well before elucidation of the first protein structure.

In an α-helix, each backbone N−H forms a hydrogen bond with the backbone carbonyl oxygen situated four residues away in the linear sequence chain (toward the N-terminus): N−H$(i)\cdots$O=C$(i-4)$. The two sequentially distant hydrogen-bonded groups are brought into spatial proximity by conferring a helical twist upon the chain. This results in a rod-like structure, with the hydrogen bonds oriented approximately parallel to the long axis of the helix.

In globular proteins, the average length of an α-helix is 12 residues. Typically, helices are found on the outside of the protein, with a hydrophilic face oriented toward the surrounding aqueous solvent and a hydrophobic face oriented toward the protein interior.

Inescapably, end effects deprive the first four amide hydrogens and last four carbonyl oxygens of Pauling-type, intra-helical hydrogen bond partners. The special hydrogen-bonding motifs that can provide partners for these otherwise unsatisfied groups are known as "helix caps."

In globular proteins, helices account for $\sim 25\%$ of the structure on average, but this number varies. Some proteins, like myoglobin, are predominantly helical, while others, like plastocyanin, lack helices altogether.

REPETITIVE SECONDARY STRUCTURE: THE 3$_{10}$-HELIX

When backbone dihedral angles are assigned repeating ϕ,ψ-values near $(-50°, -30°)$, the chain twists into a right-handed helix. By convention, this helix is named using formal nomenclature: 3$_{10}$ designates three residues per helical turn and 10 atoms in the hydrogen bonded ring between each N−H donor and its C=O acceptor. (In this nomenclature, the α-helix would be called a 3.6$_{13}$ helix.)

Single turns of 3$_{10}$ helix are common and closely resemble a type of β-turn (see below). Often, α-helices terminate in a turn of 3$_{10}$ helix. Longer 3$_{10}$ helices are sterically strained and much less common.
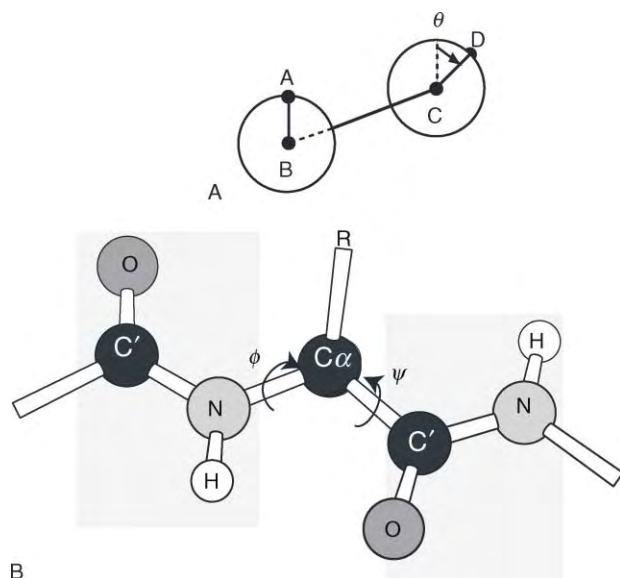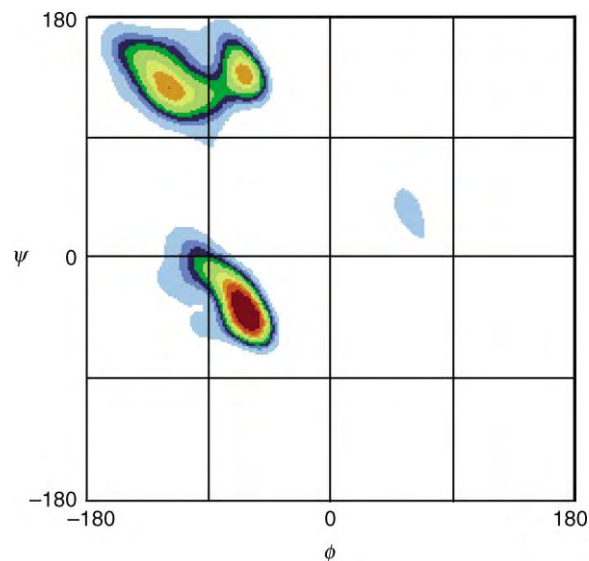
FIGURE 2 (A) Definition of a dihedral angle. In the diagram, the dihedral angle, θ, measures the rotation of line segment CD with respect to line segment AB, where A, B, C, and D correspond to the x,y,z-positions of four atoms. (θ is calculated as the scalar angle between the two normals to planes A–B–C and B–C–D.) By convention, clockwise rotation is positive and $\theta = 0°$ when A and D are eclipsed. (B) Degrees of freedom in the protein backbone. The peptide bond (C'–N) has partial double bond character, so that the six atoms, Ca(i)–CO–Ca($i + 1$), are approximately co-planar. Consequently, only two primary degrees of freedom are available for each residue. By convention, these two dihedral angles are called ϕ and $\psi_0 \phi$ is specified by the four atoms C'(i)–N–Ca–C'($i + 1$) and ψ by the four atoms N(i)–Ca–C'–N($i + 1$). When the chain is fully extended, as depicted here, $\phi = \psi = 180°$.

FIGURE 3 A contoured Ramachandran (ϕ, ψ) plot. Backbone ϕ,ψ-angles were extracted from 1042 protein subunits of known structure. Only nonglycine residues are shown. Contours were drawn in population intervals of 10% and are indicated by the ten colors (in rainbow order). The most densely populated regions are colored red. Three heavily populated regions are apparent, each near one of the major elements of repetitive secondary structure: α-helix ($\sim -60°$, $-40°$), β-strand ($\sim -120°, 120°$), P_{II} helix ($\sim -70°, 140°$). Adapted from Hovmöller, S., Zhou, T., and Ohlson, T. (2002). Conformation of amino acids in proteins. *Acta Cryst.* **D58**, 768–776, with permission of IUCr.

REPETITIVE SECONDARY STRUCTURE: THE β-STRAND

When backbone dihedral angles are assigned repeating ϕ,ψ-values near ($-120°, -120°$), the chain adopts an extended conformation called a β-strand. Two or more β-strands, aligned so as to form inter-strand hydrogen bonds, are called a β-sheet. A β-sheet of just two hydrogen-bonded β-strands interconnected by a tight turn is called a β-hairpin. The average length of a single β-strand is seven residues.

The classical definition of secondary structure found in most textbooks is limited to hydrogen-bonded backbone structure and, strictly speaking, would not include a β-strand, only a β-sheet. However, the β-sheet is tertiary structure, not secondary structure; the intervening chain joining two hydrogen-bonded β-strands can range from a tight turn to a long, structurally complex stretch of polypeptide chain. Further, approximately half the β-strands found in proteins are singletons and do not form inter-strand hydrogen bonds with another β-strand. Textbooks tend to blur this issue.

Typically, β-sheet is found in the interior of the protein, although the outermost parts of edge-strands usually reside at the protein's water-accessible surface.

Two β-strands in a β-sheet are classified as either parallel or anti-parallel, depending upon whether their mutual N- to C-terminal orientation is the same or opposite, respectively.

In globular proteins, β-sheet accounts for about 15% of the structure on an average, but, like helices, this number varies considerably. Some proteins are predominantly sheet while others lack sheet altogether.

REPETITIVE SECONDARY STRUCTURE: THE POLYPROLINE II HELIX (P_{II})

When backbone dihedral angles are assigned repeating ϕ,ψ-values near ($-70°, +140°$), the chain twists into a left-handed helix with 3.0 residues per helical turn. The name of this helix is derived from a poly-proline homopolymer, in which the structure is forced by its stereochemistry. However, a polypeptide chain can adopt a P_{II} helical conformation whether or not it contains proline residues.

Unlike the better known α-helix, a P_{II} helix has no intrasegment hydrogen bonds, and it is not included in the classical definition of secondary structure for this reason. This extension of the definition is also needed in the case of an isolated β-strand. Recent studies have shown that the unfolded state of proteins is rich in P_{II} structure.

NONREPETITIVE SECONDARY STRUCTURE: THE TURN

Turns are sites at which the polypeptide chain changes its overall direction, and their frequent occurrence is the reason why globular proteins are, in fact, globular.

Turns can be subdivided into β-turns, γ-turns, and tight turns. β-turns involve four consecutive residues, with a hydrogen bond between the amide hydrogen of the 4th residue and the carbonyl oxygen of the 1st residue: $N-H(i) \cdots O=C(i-3)$. β-turns are further subdivided into subtypes (e.g., Type I, I′, II, II′, III,...) depending upon their detailed stereochemistry. γ-turns involve only three consecutive, hydrogen-bonded residues, $N-H(i) \cdots O=C(i-2)$, which are further divided into subtypes.

More gradual turns, known as "reverse turns" or "tight turns," are also abundant in protein structures. Reverse turns lack intra-turn hydrogen bonds but nonetheless, are involved in changes in overall chain direction.

Turns are usually, but not invariably, found on the water-accessible surface of proteins. Together, β, γ- and reverse turns account for about one-third of the structure in globular proteins, on an average.

NONREPETITIVE SECONDARY STRUCTURE: THE Ω-LOOP

Ω-loops are sites at which the polypeptide loops back on itself, with a morphology that resembles the Greek letter "Ω" although often with considerable distortion. They range in length from 6–16 residues, and, lacking any specific pattern of backbone-hydrogen bonding, can exhibit significant structural heterogeneity.

Like turns, Ω-loops are typically found on the outside of proteins. On an average, there are about four such structures in a globular protein.

Identification of Secondary Structure from Coordinates

Typically, one becomes familiar with a given protein structure by visualizing a model – usually a computer model – that is generated from experimentally determined coordinates. Some secondary structure types are well defined on visual inspection, but others are not. For example, the central residues of a well-formed helix are visually unambiguous, but the helix termini are subject to interpretation. In general, visual parsing of the protein into its elements of secondary structure can be a highly subjective enterprise. Objective criteria have been developed to resolve such ambiguity. These criteria have been implemented in computer programs that accept a protein's three-dimensional coordinates as input and provide its secondary structure components as output.

INHERENT AMBIGUITY IN STRUCTURAL IDENTIFICATION

It should be realized that objective criteria for structural identification can provide a welcome self-consistency, but there is no single "right" answer. For example, turns have been defined in the literature as chains sites at which the distance between two α-carbon atoms, separated in sequence by four residues, is not more than 7Å, provided the residues are not in an α-helix: distance$[C\alpha(i)-C\alpha(i+3)] \leq 7$Å and $C\alpha(i)-C\alpha(i+3)$ not α-helix. Indeed, turns identified using this definition agree quite well with one's visual intuition. However, the 7Å threshold is somewhat arbitrary. Had 7.1Å been used instead, additional, intuitively plausible turns would have been found.

PROGRAMS TO IDENTIFY STRUCTURE FROM COORDINATES

Many workers have devised algorithms to parse the three-dimensional structure into its secondary structure components. Unavoidably, these procedures include investigator-defined thresholds. Two such programs are mentioned here.

The *D*atabase of *S*econdary *S*tructure Assignments in *P*roteins

This is the most widely used secondary structure identification method available today. Developed by Kabsch and Sander, it is accessible on the internet, both from the original authors and in numerous implementations from other investigators as well.

The *d*atabase of *s*econdary *s*tructure assignments in *p*roteins (DSSP) identifies an extensive set of secondary structure categories, based on a combination of backbone dihedral angles and hydrogen bonds. In turn, hydrogen bonds are identified based on geometric criteria involving both the distance and orientation between a donor–acceptor pair. The program has criteria for recognizing α-helix, 3_{10}-helix, π helix, β-sheet (both parallel and anti-parallel), hydrogen-bonded turns and reverse turns. (Note: the π-helix is rare and has been omitted from the secondary structure categories.)

*Pro*tein *S*econdary *S*tructure Assignments

In contrast to DSSP, protein secondary structure assignments (PROSS) identification is based solely on backbone dihedral angles, without resorting to hydrogen

bonds. Developed by Srinivasan and Rose, it is accessible on the *internet*.

PROSS identifies only α-helix, β-strand, and turns, using standard ϕ,ψ definitions for these categories. Because hydrogen bonds are not among the identification criteria, PROSS does not distinguish between isolated β-strands and those in a β-sheet.

Prediction of Protein Secondary Structure from Amino Acid Sequence

Efforts to predict secondary structure from amino acid sequence dates back to the 1960s to the works of Guzzo, Prothero and, slightly later, Chou and Fasman. The problem is complicated by the fact that protein secondary structure is only marginally stable, at best. Proteins fold cooperatively, with secondary and tertiary structure emerging more or less concomitantly. Typical peptide fragments excised from the host protein, and measured in isolation, exhibit only a weak tendency to adopt their native secondary structure conformation.

PREDICTIONS BASED ON EMPIRICALLY DETERMINED PREFERENCES

Motivated by early work of Chou and Fasman, this approach uses a database of known structures to discover the empirical likelihood, f, of finding each of the twenty amino acids in helix, sheet, turn, *etc*. These likelihoods are equated to the residue's normalized frequency of occurrence in a given secondary structure type, obtained by counting. Using alanine in helices as an example

$$\text{fraction Ala in helix} = \frac{\text{occurrences of Ala in helices}}{\text{occurrences of Ala in database}}$$

This fraction is then normalized against the corresponding fraction of helices in the database:

$$f_{\text{Ala}}^{\text{helix}} = \frac{\text{fraction Ala in helix}}{\text{fraction helices in database}}$$

$$= \frac{\dfrac{\text{occurrences of Ala in helices}}{\text{occurrences of Ala in database}}}{\dfrac{\text{number of residues in helices}}{\text{number of residues in database}}}$$

A normalized frequency of unity indicates no preference – i.e., the frequency of occurrence of the given residue in that particular position is the same as its frequency at large. Normalized frequencies greater than/less than unity indicate selection for/against the given residue in a particular position.

These residue likelihoods are then used in combination to make a prediction. When only a small number of proteins had been solved, these data-dependent f-values fluctuated significantly as new structures were added to the database. At this point there are more than 22 000 structures in the Protein Data Bank (www.rcsb.org), and the f-values have reached a plateau.

DATABASE-INDEPENDENT PREDICTIONS: THE HYDROPHOBICITY PROFILE

Hydrophobicity profiles have been used to predict the location of turns in proteins. A hydrophobicity profile is a plot of the residue number versus residue hydrophobicity, averaged over a running window. The only variables are the size of the window used for averaging and the choice of hydrophobicity scale (of which there are many). No empirical data from the database is required. Peaks in the profile correspond to local maxima in hydrophobicity, and valleys to local minima. Prediction is based on the idea that apolar sites along the chain (i.e., peaks in the profile) will be disposed preferentially to the molecular interior, forming a hydrophobic core, whereas polar sites (i.e., valleys in the profile) will be disposed to the exterior and correspond to chain turns.

NEURAL NETWORKS

More recently, neural network approaches to secondary structure prediction have come to dominate the field. These approaches are based on pattern-recognition methods developed in artificial intelligence. When used in conjunction with the protein database, these are the most successful programs available today.

A neural network is a computer program that associates an input (e.g., a residue sequence) with an output (e.g., secondary structure prediction) through a complex network of interconnected nodes. The path taken from the input through the network to the output depends upon past experience. Thus, the network is said to be "trained" on a dataset.

The method is based on the observation that amino acid substitutions follow a pattern within a family of homologous proteins. Therefore, if the sequence of interest has homologues within the database of known structures, this information can be used to improve predictive success, provided the homologues are recognizable. In fact, a homologue can be recognized quite successfully when the sequence of interest and a putative homologue have an aligned sequence identity of 25% or more.

Neural nets provide an information-rich approach to secondary structure prediction that has become increasingly successful as the protein databank has grown.

PHYSICAL BASIS OF SECONDARY STRUCTURE

An impressive number of secondary structure prediction methods can be found in the literature and on the web. Surprisingly, almost all are based on empirical likelihoods or neural nets; few are based on physicochemical theory.

In one such theory, secondary structure propensities are predominantly a consequence of two competing local effects – one favoring hydrogen bond formation in helices and turns, and the other opposing the attendant reduction in sidechain conformational entropy upon helix and turn formation. These sequence-specific biases are densely dispersed throughout the unfolded polypeptide chain, where they serve to pre-organize the folding process and largely, but imperfectly, anticipate the native secondary structure.

WHY AREN'T SECONDARY STRUCTURE PREDICTIONS BETTER?

Currently, the best methods for predicting helix and sheet are correct about three-quarters of the time. Can greater success be achieved?

Several measures to assess predictive accuracy are in common use, of which the Q3 score is the most widespread. The Q3 score gives the percentage of correctly predicted residues in three categories: helix, strand, and coil (i.e., everything else):

$$Q3 = \frac{\text{number of correctly predicted residues}}{\text{total number of residues}} \times 100$$

where the "correct" answer is given by a program to identify secondary structure from coordinates, e.g., DSSP. At this writing, (Position-Specific PREDiction algorithm) PSIPRED has an overall Q3 score of 78%.

Is greater prediction accuracy possible? It has been argued that prediction methods fail to achieve a higher rate of success because some amino acid sequences are inherently ambiguous. That is, these "conformational chameleons" will adopt a helical conformation in one protein, but the identical sequence will adopt a strand conformation in another protein. Only time will tell whether current efforts have encountered an inherent limit.

SEE ALSO THE FOLLOWING ARTICLES

Amino Acid Metabolism • Multiple Sequence Alignment and Phylogenetic Trees • Protein Data Resources • X-Ray Determination of 3-D Structure in Proteins

GLOSSARY

α-helix   The best-known element of secondary structure in which the polypeptide chain adopts a right-handed helical twist with 3.6 residues per turn and an $i \rightarrow i - 4$ hydrogen bond between successive amide hydrogens and carbonyl oxygens.

β-strand   An element of secondary structure in which the chain adopts an extended conformation. A β-sheet results when two or more aligned β-strands form inter-strand hydrogen bonds.

Chou–Fasman   Among the earliest attempts to predict protein secondary structure from the amino acid sequence. The method, which uses a database of known structures, is based on the empirically observed likelihood of finding the 20 different amino acids in helix, sheet or turns.

DSSP   The most widely used method to parse x, y, z-coordinates for a protein structure into elements of secondary structure.

hydrophobicity   A measure of the degree to which solutes, like amino acids, partition spontaneously between a polar environment (like the outside of a protein) and an organic environment (like the inside of a protein).

hydrophobicity profile   A method to predict the location of peptide chain turns from the amino acid sequence by plotting averaged hydrophobicity against residue number. The method does not require a database of known structure.

neural network   A pattern recognition method – adapted from artificial intelligence – that has been highly successful in predicting protein secondary structure when used in conjunction with an extensive database of known structures.

peptide chain turn   A site at which the protein changes its overall direction. The frequent occurrence of turns is responsible for the globular morphology of globular (i.e., sphere-like) proteins.

secondary structure   The backbone structure of the protein, with particular emphasis on hydrogen bonded motifs.

tertiary structure   The three-dimensional structure of the protein.

FURTHER READING

Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002). *Biochemistry*, 5th edition. W.H. Freeman and Company, New York.

Holm, L., and Sander, C. (1996). Mapping the protein universe. *Science* **273**, 595–603.

Hovmöller, S., Zhou, T., and Ohlson, T. (2002). Conformation of amino acids in proteins. *Acta Cryst.* **D58**, 768–776.

Jones, D. T. (1999). Protein secondary structure based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202.

Mathews, C., van Holde, K. E., and Ahern, K. G. (2000). *Biochemistry*, 3rd edition. Pearson Benjamin Cummings, Menlo Park, CA.

Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Adv. Prot. Chem.* **34**, 168–340.

Rose, G. D., Gierasch, L. M., and Smith, J. A. (1985). Turns in peptides and proteins. *Adv. Prot. Chem.* **37**, 1–109.

Voet, D., and Voet, J. G. (1996). *Biochemistry*, 2nd edition. Wiley, New York.

BIOGRAPHY

George Rose is Professor of Biophysics and Director of the Institute for Biophysical Research at Johns Hopkins University. He holds a Ph.D. from Oregon State University. His principal research interest is in protein folding, and he has written many articles on this topic. He serves as the consulting editor of *Proteins: Structure, Function and Genetics* and as a member of the editorial advisory board of *Protein Science*. Recently, he was a Fellow of the John Simon Guggenheim Memorial Foundation.

Secretases

Robert L. Heinrikson
The Pharmacia Corporation, Kalamazoo, Michigan, USA

**Secretases are proteolytic enzymes involved in the processing
of an integral membrane protein known as Amyloid precursor
protein, or APP. β-Amyloid (Aβ) is a neurotoxic and highly
aggregative peptide that is excised from APP by secretase
action, and that accumulates in the neuritic plaque found in the
brains of Alzheimer's disease (AD) patients. The amyloid
hypothesis holds that the neuronal dysfunction and clinical
manifestation of AD is a consequence of the long-term
deposition and accumulation of Aβ, and that this peptide of
40–42 amino acids is a causative agent of AD. Accordingly, the
secretases involved in the liberation, or destruction of Aβ are
of enormous interest as therapeutic intervention points toward
treatment of this dreaded disease.**

Background

Proteolytic enzymes play crucial roles in a wide variety
of normal and pathological processes in which they
display a high order of selectivity for their substrate(s)
and the specific peptide bonds hydrolyzed therein. This
article concerns secretases, membrane-associated pro-
teinases that produce, or prevent formation of, a highly
aggregative and toxic peptide called β-amyloid (Aβ).
This Aβ peptide is removed from a widely distributed
and little understood Type I integral membrane protein
called amyloid precursor protein (APP). The apparent
causal relationship between Aβ and AD has fueled an
intense interest in the secretases responsible for its
production. Herein will be discussed the current under-
standing of three of the most-studied secretases, α-, β-,
and γ-secretases. A schematic representation of the Aβ
region of APP showing the amino acid sequence of Aβ
and the major sites of cleavage for these three secretases
is given in Figure 1. Aβ is produced by the action of
β- and γ-secretases, and there is an intense search
underway for inhibitors of these enzymes that might
serve as drugs in treatment of Alzheimer's disease (AD).
The α-secretase cleaves at a site near the middle of Aβ,
and gives rise to fragments of Aβ that lack the
potential for aggregation; therefore, amplification of
α-secretase activity might be seen as another approach
to AD therapy.

α-Secretase

The activity responsible for cleavage of the Lys^{16}-Leu^{17}
bond within the Aβ region (Figure 1) is ascribed to
α-secretase. This action prevents formation of the 40–42
amino acid residue Aβ and leads to release of soluble
APPα and the membrane-bound C83-terminal fragment.
α-Secretase competes with β-secretase for the APP
substrate, but the α-secretase product, soluble APPα
(pathway A, Figure 1) is generated at a level about 20
times that of the sAPPβ released by β-secretase (pathway
B). Because α-secretase action prevents formation of the
toxic Aβ peptide, augmentation of this activity could
represent a useful strategy in AD treatment, and this has
been done experimentally by activators of protein kinase
C (PKC) such as phorbol esters and by muscarinic
agonists. The specificity of the α-secretase for the
Lys^{16}-$\downarrow$-Leu^{17} cleavage site (Figure 1) appears to be
governed by spatial and structural requirements that
this bond exist in a local α-helical conformation and be
within 12 or 13 amino acids distance from the membrane.
α-Secretase has not been identified as any single
proteinase, but two members of the ADAM (a disintegrin
and metalloprotease) family, ADAM-10 and ADAM-17
are candidate α-secretases. ADAM-17 is known as TACE
(tumor necrosis factor-α-converting enzyme) and TACE
cleaves peptides modeled after the α-secretase site at the
Lys^{16}-$\downarrow$-Leu position. This was also shown to be the case
for ADAM-10; overexpression of this enzyme in a human
cell line led to several-fold increase in both basal and
PKC-inducible α-secretase activity. As of now, it remains
to be proven whether α-secretase activity derives from
either or both of these ADAM family metalloproteinases,
or whether another as yet unidentified proteinase carries
out this processing of APP.

β-Secretase

The enzyme responsible for cleaving at the amino-
terminus of Aβ is β-secretase (Figure 1). In the mid-1980s,
when Aβ was recognized as a principal component of AD
neuritic plaque, an intense search was begun to identify
the β-secretase. Finally, in 1999, several independent
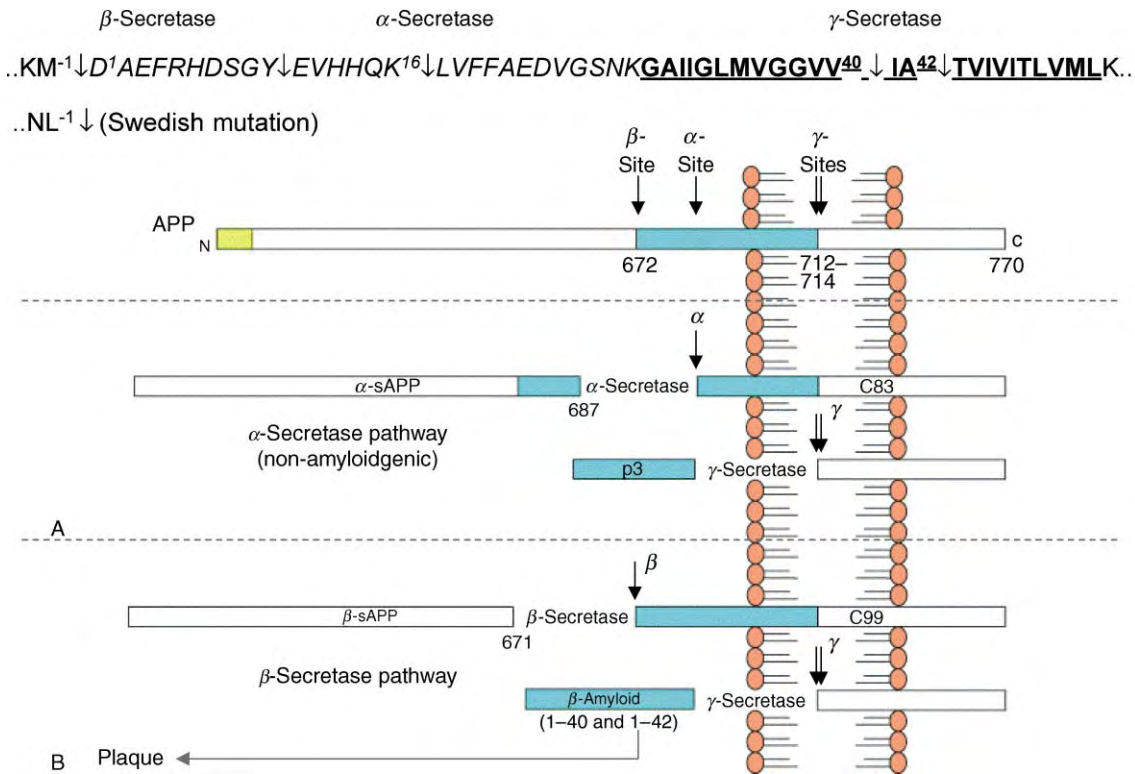
FIGURE 1 A schematic overview of APP processing by the α-, β-, and γ-secretases. The top panel shows the amino acid sequence of APP upstream of the transmembrane segment (underlined, bold), and encompassing the sequences of $A\beta_{1-40}$ and $A\beta_{1-42}$ (D^1-V^{40}, and D^1-A^{42}, respectively). The β-secretase cleaves at D^1 and Y^{10}; the α-secretase at Lys^{16}, and the γ-secretase at Val^{40} and/or Ala^{42}. Below the sequence is a representation of APP emphasizing its membrane localization and the residue numbers of interest in β- and γ-secretase processing. Panel A represents the non-amyloidogenic α-secretase pathway in which sAPPα and C83 are generated. Subsequent hydrolysis by the γ-secretase produces a p3 peptide that does not form amyloid deposits. Panel B represents the amyloidogenic pathway in which cleavage of APP by the β-secretase to liberate sAPPβ and C99 is followed by γ-secretase processing to release β-amyloid peptides ($A\beta_{1-40}$ and $A\beta_{1-42}$) found in plaque deposits.

laboratories published evidence demonstrating that β-secretase is a unique member of the pepsin family of aspartyl proteinases. This structural relationship to a well-characterized and mechanistically defined class of proteases gave enormous impetus to research on β-secretase. The preproenzyme consists of 501 amino acids, with a 21-residue signal peptide, a prosegment of about 39 residues, the catalytic bilobal unit with active site aspartyl residues at positions 93 and 289, a 27-residue transmembrane region, and a 21-residue C-terminal domain. The membrane localization of β-secretase makes it unique among mammalian aspartyl proteases described to date. Another interesting feature of the enzyme is that, unlike pepsin, renin, cathepsin D, and other prototypic members of the aspartyl proteases, it does not appear to require removal of the prosegment as a means of activation. A furin-like activity is responsible for cleavage in the sequence Arg-Leu-Pro-Arg-$\downarrow$-Glu^{25} of the proenzyme, but this does not lead to any remarkable enhancement of activity, at least as is seen in recombinant constructs of pro-β-secretase. β-Secretase has been referred to by a number of designations in the literature, but the term BACE (β-site APP cleaving

enzyme) has become most widely adopted. With the discovery of the β-secretase, it was recognized that there was another human homologue of BACE with a transmembrane segment and this has now come to be called BACE2. This may well be a misnomer, since the function of BACE2 has yet to be established, and it is not clear that APP is a normal substrate of this enzyme. At present, BACE2 is not considered to be a secretase.

There is considerable experimental support for the assertion that BACE is, in fact, the β-secretase involved in APP processing. The enzyme is highly expressed in brain, but is also found in other tissues, thus explaining the fact that many cell types can process $A\beta$. Use of antisense oligonucleotides to block expression of BACE greatly diminishes production of $A\beta$ and, conversely, overexpression of BACE in a number of cell lines leads to enhanced $A\beta$ production. BACE knockout mice show no adverse phenotype, but have dramatically reduced levels of $A\beta$. This not only demonstrates that BACE is the true β-site APP processor, but also that its elimination does not pose serious consequences for the animal, a factor of great importance in targeting BACE for inhibition in AD therapy.

Much of the evidence in support of the amyloid hypothesis comes from the observation of mutations near the β- and γ-cleavage sites in APP that influence production of Aβ and correlate directly with the onset of AD. One such mutation in APP, that invariably leads to AD in later life, occurs at the β-cleavage site where Lys-Met^{-1} is changed to Asn-Leu^{-1} (Figure 1). This so-called Swedish mutation greatly enhances production of Aβ, and as would be expected, β-secretase hydrolyzes the mutated Leu-Asp^1 bond in model peptides ~ 50 times faster than the wild-type Met-Asp^1 bond. It is important to recognize that BACE cleavage is required for subsequent processing by the γ-secretase; in this sense, a BACE inhibitor will also block γ-secretase. Another BACE cleavage point is indicated in Figure 1 by the arrow at Y$^{10} \downarrow$ E^{11}; the A$\beta_{11-40 \text{ or } 42}$ subsequently liberated by γ-secretase action also forms amyloid deposits and is found in neuritic plaque.

In all respects, therefore, BACE fits the picture expected of β-secretase, and because of its detailed level of characterization and its primary role in Aβ production, it has become a major target for development of inhibitors as drugs to treat AD. Great strides in this direction have become possible because of the availability of three-dimensional (3-D) structural information on BACE. The crystal structure of BACE complexed with an inhibitor is represented schematically in Figure 2. Homology with the pepsin-like aspartyl proteases is reflected in the similar folding pattern of BACE, with extensive β-sheet organization, and the proximal location of the two aspartyl residues that comprise the catalytic machine for peptide bond cleavage. The C-terminal lobe of the molecule is larger than is customarily seen in the aspartyl proteases, and contains extra elements of structure with as yet unexplained impact on function. In fact, before the crystal structure was solved, it was thought that this larger C-terminal region might contribute a spacer to distance the catalytic unit from the membrane and to provide mobility. This appears not to be the case. As denoted by the arrow in Figure 2, there is a critical disulfide bridge linking the C-terminal region just upstream of the transmembrane segment to the body of the molecule. Therefore, the globular BACE molecule is proximal to the membrane surface and is not attached via a mobile stalk that would permit much motion. This steric localization would be expected to limit the repertoire of protein substrates that are accessible to BACE as it resides in the Golgi region. Crystal structures of BACE/inhibitor complexes have revealed much about the nature of protein-ligand interactions, and information regarding the nature of binding sites obtained by this approach will be of critical importance in the design and development of inhibitors that will be effective drugs in treatment of AD.

FIGURE 2 Schematic representation of the 3-D structure of the BACE (β-secretase) catalytic unit as determined by x-ray crystallography. Arrows and ribbons designate β-strands and α-helices, respectively. An inhibitor is shown bound in the cleft defined by the amino- (left) and carboxyl- (right) terminal halves of the molecule. The C-terminus of the catalytic unit is marked C to indicate the amino acid residue immediately preceding the transmembrane and cytoplasmic domains of BACE. These latter domains were omitted from the construct that was solved crystallographically. The arrow marks a disulfide bridge, which maintains the C-terminus in close structural association with the body of the catalytic unit. The catalytic entity as depicted sits directly on the membrane surface, thereby restricting its motion relative to protein substrates. (Courtesy of Dr. Lin Hong, Oklahoma Medical Research Foundation, Oklahoma City, OK.)

γ-Secretase

γ-Secretase activity is produced in a complex of proteins and is yet to be understood in terms of the actual catalytic entity and mechanism of proteolysis. This secretase cleaves bonds in the middle of the APP segment that traverses the membrane (underlined and boldface in Figure 1), and its activity is exhibited subsequent to cleavages at the α- or β-sites. In Figure 1, the γ-secretase cleavage sites are indicated by two arrows. Cleavage at the Val^{40}-Ile^{41} bond liberates the more abundant 40-amino acid residue Aβ (Aβ_{1-40}). Cleavage at Ala^{42}-Thr^{43} produces a minor Aβ species, Aβ_{1-42}, but one that appears to be much more hydrophobic and aggregative, and it is the 42-residue Aβ that is believed to be of most significance in AD pathology. As was the case for APP β-site mutations, there are human APP mutants showing alterations in the vicinity of the γ-site, and these changes, powerfully associated with onset of AD, lead to higher ratios of Aβ_{1-42}.

Central to the notion of the γ-secretase is the presence of presenilins, intregral membrane proteins with mass ~ 50 kDa. There are a host of presenilin mutations in familial AD (FAD) that are associated with early onset disease and an increased production of the toxic Aβ_{1-42}. This correlation provides strong

support for the involvement of presenilin in AD, and its presence in γ-secretase preparations implies that it is either a proteolytic enzyme in its own right, or can contribute to that function in the presence of other proteins. In fact, much remains to be learned about the presenilins; it has been difficult to obtain precise molecular and functional characterization because of their close association with membranes and other proteins in a complex. Modeling studies have predicted a variable number of transmembrane segments (6–8), but presenilin function is predicated upon processing by an unknown protease to yield a 30 kDa N-terminal fragment (NTF) and a 20 kDa C-terminal fragment (CTF). These accumulate *in vivo* in a 1:1 stoichiometry within high molecular weight complexes with a variety of ancillary proteins. Some of the cohort proteins identified in the multimeric presenilin complexes displaying γ-secretase activity include catenins, armadillo-repeat proteins that appear not to be essential for γ-secretase function, and nicastrin. Nicastrin is a Type I integral membrane protein with homologues in a variety of organisms, but its function is unknown. It shows intracellular colocalization with presenilin, and is able to bind the NTF and CTF of presenilin as well as the C83 and C99 C-terminal APP substrates of γ-secretase. Interestingly, down-regulation of the nicastrin homologue in *Caenorhabditis elegans* gave a phenotype similar to that seen in worms deficient in presenilin and notch. Evidence that nicastrin is essential for γ-secretase cleavage of APP and notch adds to the belief that nicastrin is an important element in presenilin, and γ-secretase function. Efforts to delineate other protein components of γ-secretase complexes and to understand their individual roles in the enzyme function represent a large current research effort. Recently, two additional proteins associated with the complex have been identified through genetic screening of flies and worms. The *aph-1* gene encodes a protein with 7 transmembrane domains, and the *pen-2* gene codes for a small protein passing twice through the membrane. Both of these putative members of the γ-secretase complex are new proteins whose functions, either with respect to secretase activity or in other potential systems, remain to be elucidated.

At present, it is still unclear as to how γ-secretase exerts its function. What is known, however, is that γ-secretase is able to cleave at other peptide bonds in APP near the γ-site in addition to those indicated in Figure 1, and is involved with processing of intra-membrane peptide bonds in a variety of additional protein substrates, including notch. This lack of specificity is a major concern in developing drugs for AD targeted to γ-secretase that do not show side effects due to inhibition of processing of these additional, functionally diverse protein substrates.

SEE ALSO THE FOLLOWING ARTICLES

Amyloid • Metalloproteinases, Matrix

GLOSSARY

Aβ The peptide produced from APP by the action of β- and γ-secretases. Aβ shows neurotoxic activity and aggregates to form insoluble deposits seen in the brains of Alzheimer's disease patients. The α-secretase hydrolyzes a bond within the Aβ region and releases fragments which do not aggregate.

Alzheimer's disease (AD) A disease first described by Alois Alzheimer in 1906 characterized by progressive loss of memory and cognition. AD afflicts a major proportion of our aging population and is one of the most serious diseases facing our society today, especially in light of increasing human longevity. The secretases represent important potential therapeutic intervention points in AD treatment.

proteinases Enzymes that hydrolyze, or split peptide bonds in protein substrates; also referred to as proteolytic enzymes.

secretase A proteinase identified with respect to its hydrolysis of peptide bonds within a region of a Type I integral membrane protein called APP. These cleavages are responsible for liberation, or destruction of an amyloidogenic peptide of about 40 amino acid residues in length called Aβ.

FURTHER READING

Esler, W. P., and Wolfe, M. S. (2001). A portrait of Alzheimer secretases – New features and familiar faces. *Science* **293**, 1449–1454.

Fortini, M. E. (2002). γ-Secretase-mediated proteolysis in cell-surface-receptor signaling. *Nat. Rev.* **3**, 673–684.

Glenner, G. G., and Wong, C. W. (1984). Alzheimer's disease: Initial report of the purification and characterization of a novel cerbro-vascular amyloid protein. *Biophys. Res. Commun.* **120**, 885–890.

Hendriksen, Z. J. V. R. B., Nottet, H. S. L. M., and Smits, H. A. (2002). Secretases as targets for drug design in Alzheimer's disease. *Eur. J. Clin. Invest.* **32**, 60–68.

Sisodia, S. S., and St. George-Hyslop, P. H. (2002). γ-Secretase, notch, Aβ and Alzheimer's disease: Where do the presenilins fit in? *Nat. Rev.* **3**, 281–290.

BIOGRAPHY

Robert L. Heinrikson is a Distinguished Fellow at the Pharmacia Corporation in Kalamazoo, MI. Prior to his industrial post, Dr. Heinrikson was Full Professor of Biochemistry at the University of Chicago. His principal area of research is protein chemistry, with an emphasis on proteolytic enzymes as drug targets. Dr. Heinrikson is on the editorial board of four journals, including the *Journal of Biological Chemistry*. He is a member of the American Society of Biochemistry and Molecular Biology and Phi Beta Kappa.

Secretory Pathway

Karen J. Colley
University of Illinois at Chicago, Chicago, Illinois, USA

The eukaryotic cell is separated into several functionally distinct, membrane-enclosed compartments (Figure 1). Each compartment contains proteins required to accomplish specific functions. Consequently, each protein must be sorted to its proper location to ensure cell viability. Proteins possess specific signals, either encoded in their amino acid sequences or added as posttranslational modifications, which target them for the various compartments of the cell. The pioneering work of Dr. George Palade provided scientists with their first picture of the functional organization of the mammalian secretory pathway. Later work showed that the secretory pathway acts as a folding, modification, and quality control system for proteins that function in the endoplasmic reticulum (ER) and Golgi apparatus, and for those that are targeted to the lysosome, plasma membrane, and extracellular space. This article will focus on protein targeting to and within the compartments of the secretory pathway, and how proteins within this pathway function to ensure that correctly folded and modified proteins are delivered to the cell surface and secreted from cells.

Targeting of New Proteins to the Secretory Pathway

WHAT KINDS OF PROTEINS ARE TARGETED TO THE SECRETORY PATHWAY?

The proteins that are targeted to the secretory pathway can be separated into two groups – those that function in the ER and Golgi to ensure proper protein folding and modification (i.e., resident proteins), and those that are processed in the ER and Golgi, and are transported to later compartments like the lysosome, plasma membrane, and extracellular space (Figure 1). Each of these proteins not only possesses a signal to enter the secretory pathway, but also may have a secondary signal to localize it to a particular organelle within the pathway.

SIGNALS AND MECHANISMS OF SECRETORY PATHWAY ENTRY

The 1999 Nobel Prize in physiology or medicine was awarded to Dr. Günter Blobel for his contributions to our understanding of the mechanism of secretory pathway entry. Dr. Blobel and his colleagues found that in order to enter the secretory pathway, proteins are synthesized with an amino terminal signal peptide that allows them to cross the membrane of the endoplasmic reticulum (ER). The signal peptide is recognized by a complex of proteins and ribonucleic acid called the signal recognition particle (SRP) (Figure 2). As the signal peptide emerges from the ribosome during translation, SRP binds to it and halts translation, and then targets the new protein–ribosome complex to the cytoplasmic face of the ER membrane where it binds to the SRP receptor. Subsequently, the new protein–ribosome complex is released from SRP and its receptor, and transferred to an aqueous membrane channel known as the "translocon." Translation resumes and the new protein is co-translationally transferred through the translocon into the lumen of the ER, where in many cases the signal peptide is cleaved by a specific signal peptidase (Figure 2).

SOLUBLE AND INTEGRAL MEMBRANE PROTEINS

Soluble proteins are completely translocated across the ER membrane into the lumen (Figure 2). These proteins will either remain in the ER, be targeted to another organelle, or be secreted from the cell. Integral membrane proteins that possess one or more hydrophobic membrane-spanning regions will use these sequences to insert into the membrane of the ER and either stay as ER-resident transmembrane proteins, or be targeted to another cellular membrane. A type-I membrane protein has a cleavable signal peptide and a separate hydrophobic stretch of amino acids that acts as a membrane-spanning region. This type of protein has its amino terminus in the lumen of an organelle or the outside of the cell (which are topologically equivalent), and its carboxy terminus in the cytoplasm (Figure 2). In contrast, a type-II membrane protein has an uncleavable signal peptide, or signal anchor that is not at the protein's amino terminus and serves a dual function as both signal peptide and a membrane-spanning region. Type-II membrane proteins employ a more elaborate insertion mechanism than do type-I membrane proteins.

Compartments of the eukaryotic cell
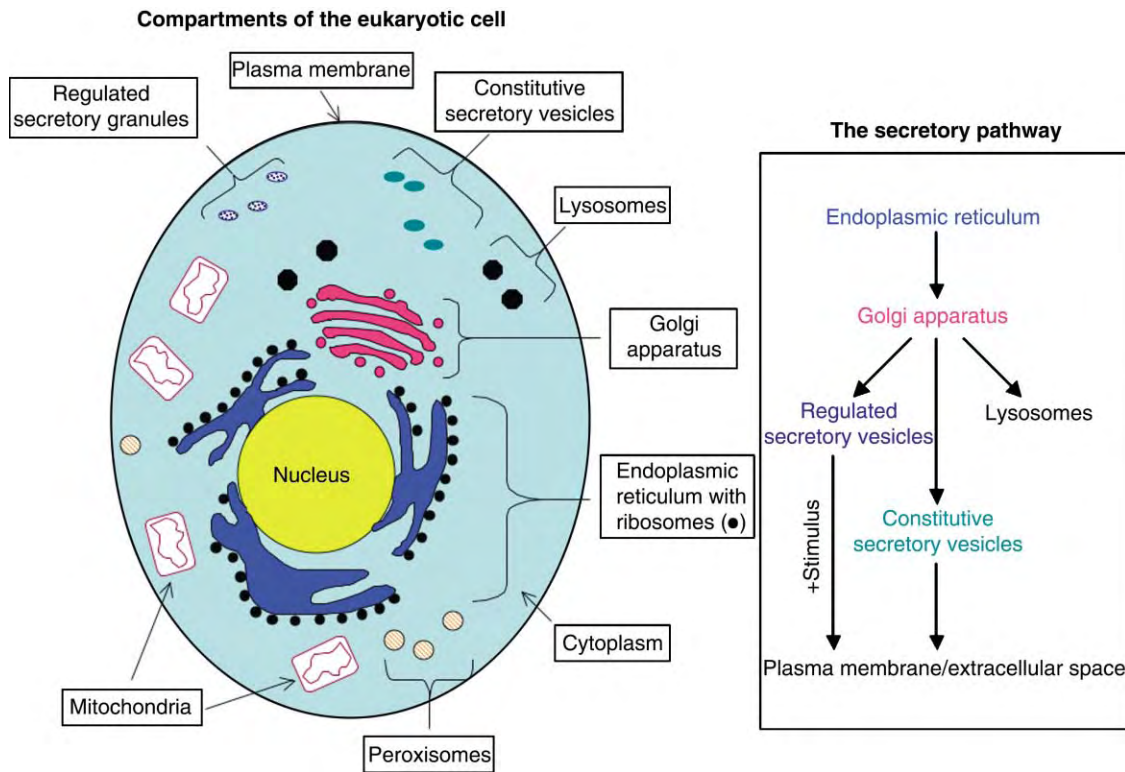
FIGURE 1   Compartments of eukaryotic cells and the organization of the secretory pathway. Diagrammatic representation of the compartments of the eukaryotic cell is shown. The anterograde flow of membrane and protein traffic in the secretory pathway is shown in the box. Anterograde flow is indicated by arrows. Retrograde flow between the ER and Golgi, endosome/lysosome system and Golgi, and plasma membrane and Golgi does occur, but is not shown.

For this reason, a type-II membrane protein will have its carboxy terminus in the lumen of an organelle or outside the cell, and its amino terminus in the cytoplasm (Figure 2). Other proteins span the membrane several times and are called type-III membrane proteins. They can start with either cleavable signal peptides or uncleavable signal anchors and possess variable numbers of hydrophobic membrane-spanning segments.

Protein Folding and Modification in the ER

THE INITIATION OF PROTEIN N-LINKED GLYCOSYLATION IN THE ER

As proteins enter the ER lumen, they fold and assemble with the help of chaperone proteins. Many proteins are also co-translationally modified by the addition of carbohydrates to asparagine residues in the process of N-linked glycosylation (Figure 2). A preformed oligosaccharide, consisting of three glucoses, nine mannoses, and two N-acetylglucosamine residues ($Glc_3Man_9GlcNAc_2$) is transferred to accessible asparagine residues in the tripeptide sequence asparagine-X-serine or threonine (X cannot be proline) by the oligosaccharide protein

transferase complex. Subsequent modification by glycosidases (enzymes that remove monosaccharides) and glycosyltransferases (enzymes that add monosaccharides) in the ER and Golgi lead to the remodeling of the N-linked oligosaccharides. These N-linked carbohydrates help proteins fold, protect them from proteolytic degradation and, in some cases, are critical for modulating and mediating protein and cell interactions at the cell surface and in the extracellular space.

CHAPERONES AND THE ER QUALITY CONTROL SYSTEM

An important function of the ER is to serve as a site of protein folding and quality control. Protein folding in the ER includes the formation of intra-molecular disulfide bonds, prolyl isomerization, and the sequestration of hydrophobic amino acids into the interior of the protein. Protein disulfide bonds are formed as the protein exits the translocon and may at first form incorrectly between cysteine residues close together in the protein's linear amino acid sequence. Thiol-oxidoreductases, such as protein disulfide isomerase (PDI), help to form and reorganize proteins' disulfide bonds into the most energetically favorable configuration. Different types of chaperones monitor a protein's
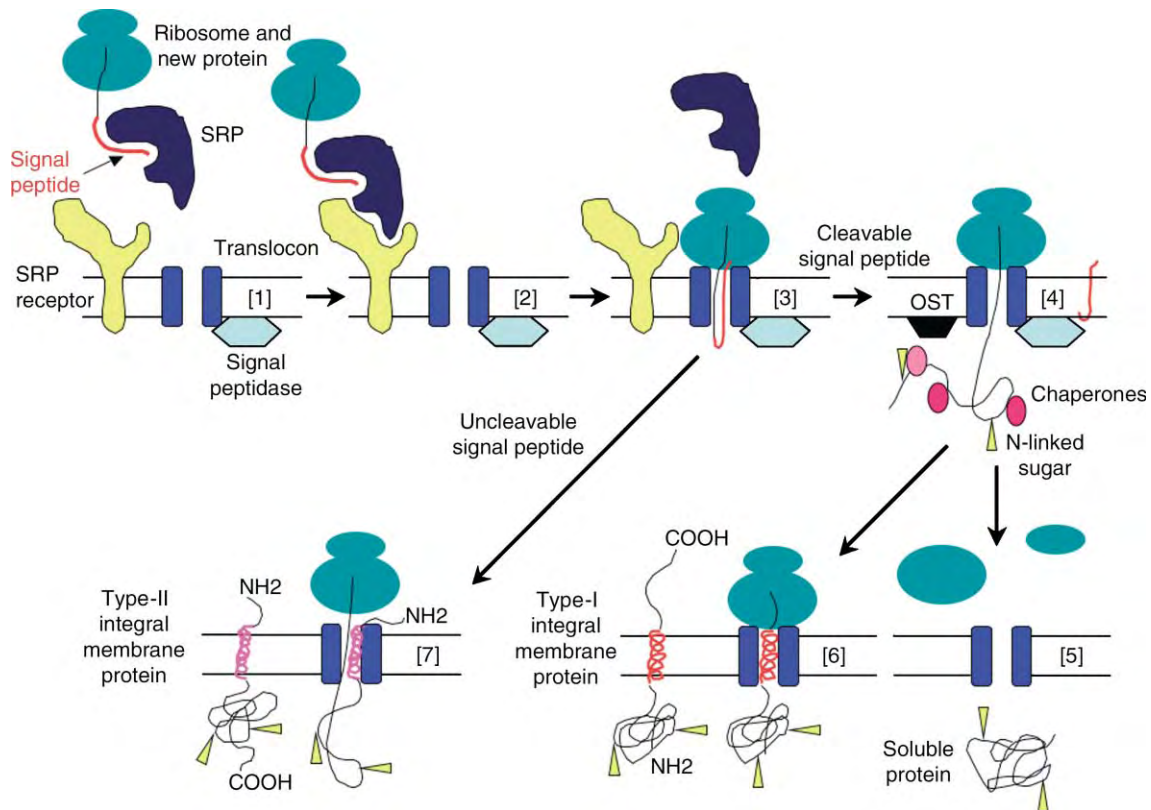
FIGURE 2   Entry into the secretory pathway. Many proteins are targeted for the secretory pathway by an amino terminal hydrophobic signal peptide that allows their co-translational translocation across the ER membrane. [1] Signal recognition particle (SRP) recognizes the new protein's signal peptide. [2] The ribosome–new protein–SRP complex interacts with the SRP receptor on the cytoplasmic face of the ER membrane. [3] The ribosome–new protein complex is transferred to the translocon channel, protein synthesis continues and the protein moves through the aqueous channel. [4] As the new protein enters the lumen of the ER, its signal peptide is cleaved by the signal peptidase, chaperone proteins bind to aid in folding and oligosaccharide protein transferase complex (OST) transfers oligosaccharides (arrowheads) to asparagine residues in the process of N-linked glycosylation. [5] Soluble proteins lack additional hydrophobic sequences and are translocated through the translocon to complete their folding and modification in the ER lumen. [6] Type-I integral membrane proteins have a second hydrophobic sequence that partitions into the lipid bilayer and acts as a membrane-spanning segment. These proteins have their amino termini in the lumen of an organelle or outstide the cell and their carboxy-termini in the cell cytoplasm. [7] Unlike proteins with cleavable amino-terminal signal peptides, type-II integral membrane proteins have an uncleavable signal anchor that target the protein to the secretory pathway and then partitions into the lipid bilayer to act as a membrane-spanning segment. These proteins have their amino termini in the cell cytoplasm and their carboxy-termini in the lumen of an organelle or outside the cell. Soluble and integral membrane proteins that enter at the level of the ER need not stay there, and can be transported out of the ER to other locations in the pathway (see Figure 1, Secretory Pathway box).

folding and prevent exit of unfolded and unassembled proteins from the ER. The chaperone BiP, originally identified as an immunoglobulin heavy-chain-binding protein, interacts with the exposed hydrophobic sequences of folding intermediates of many proteins and prevents their aggregation. Two chaperones called calnexin and calreticulin recognize a monoglucosylated carbohydrate structure ($Glc_1Man_9GlcNAc_2$) that is formed by a special glucosyltransferase that recognizes unfolded or misfolded proteins and adds a single glucose to the $Man_9GlcNAc_2$ structure. Proteins that are not folded properly or are not assembled into oligomers with partner subunits, are prevented from exiting the ER by chaperone interactions, and can be targeted back across the ER membrane through the translocon into the cytoplasm where they are degraded by the proteosome complex in a process called ER associated degradation (ERAD).

Protein Transport through and Localization in the Secretory Pathway

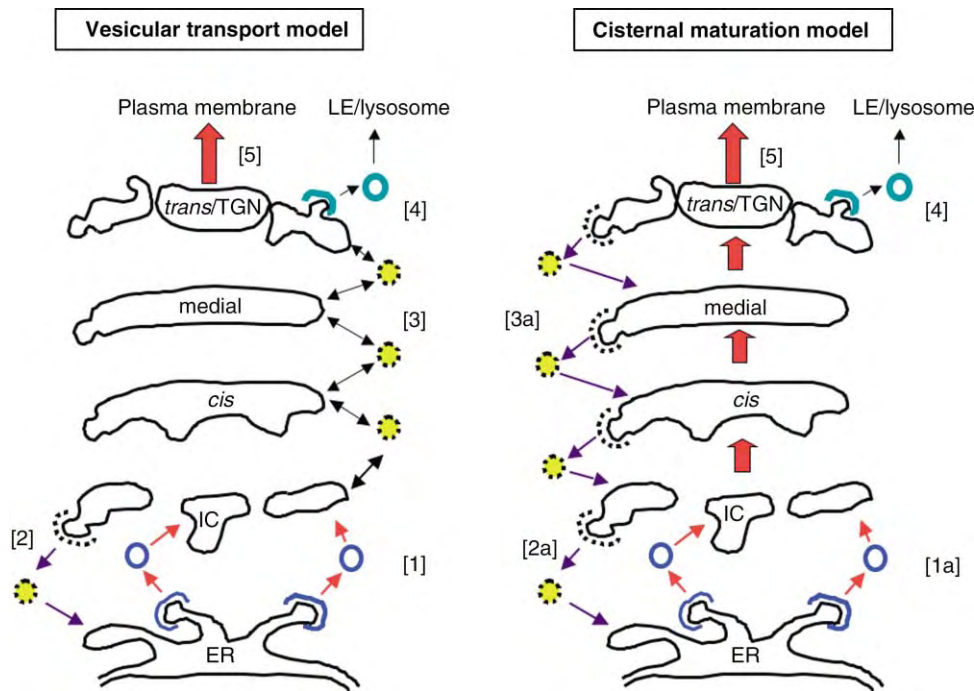## VESICULAR TRANSPORT BETWEEN THE ER AND GOLGI

Proteins move between the ER and Golgi in vesicular carriers. These vesicles are coated with specific sets of cytoplasmic proteins that form the COP-I and COP-II coats. COP-II-coated vesicles move from the ER to the

intermediate compartment (IC)/*cis* Golgi, while COP-I-coated vesicles move from the Golgi back to the ER and may also mediate transport between Golgi cisternae in both the anterograde (toward the plasma membrane) and retrograde (toward the ER) directions (Figure 3). The process of vesicular transport can be separated into three stages—cargo selection and budding, targeting, and fusion. In the first stage, the COP coats serve to select cargo for exit from a compartment and help to deform the membrane for vesicle budding. They assemble on the membrane with the help of small GTPases called ARF (specific for COP I) and Sar1p (specific for COP II). After vesicle budding, the hydrolysis of GTP by ARF and Sar1p leads to the uncoating of the vesicle. This uncoating reveals other vesicle proteins that are essential for vesicle targeting and fusion. In the second and third stages, tethering proteins on the transport vesicle and target membrane interact weakly bringing the membranes together. This allows vesicle-associated SNARE proteins and target membrane-associated SNARE proteins to form complexes. Subsequent conformational changes in the SNARE protein complex bring the membranes together for fusion. Another group of small GTPases (Rabs) control the process of vesicular transport at several levels by recruiting and activating various proteins in the pathway.

PROTEIN LOCALIZATION IN THE ER

Proteins involved in protein folding, modification, and quality control must remain in the ER, while proteins destined for the Golgi, lysosome, cell surface or those that are secreted from the cell must exit. Exit from the ER is a selective process that involves cargo receptors that interact with COP-II coat components (Figure 3). It is likely that most resident ER proteins are not selected to exit the ER. It is clear, however, that some resident proteins escape from the ER and are retrieved from the Golgi and intermediate compartment by COP-I vesicles (Figure 3). These proteins have specific amino acid signals that allow their incorporation into COP-I vesicles either by direct interaction with COP-I components or indirectly by interaction with cargo

FIGURE 3   Comparison of two models of protein transport through the Golgi apparatus. In the vesicular transport model cargo proteins move between the cisternae in vesicles, while Golgi enzymes are retained in their resident cisternae. [1] COP-II-coated vesicles transport new proteins from the ER to the intermediate compartment (IC). [2] Resident ER proteins that escape the ER can be retrieved from the IC in COP-I-coated vesicles. [3] COP-I-coated vesicles also transport anterograde cargo proteins between the Golgi cisternae in both a retrograde and anterograde fashion ("percolating vesicles"). In the cisternal maturation model, cargo proteins enter a new cisterna at the cis face of the stack, and are modified (matured) by "resident" Golgi enzymes that are continuously transported in a retrograde fashion into the sequentially maturing cisternae. [1a] COP-II-coated vesicles transport new proteins from the ER to the IC where a new *cis* cisterna forms. [2a] Resident ER proteins that escape the ER can be retrieved from the IC in COP-I-coated vesicles. [3a] Golgi enzymes are transported in a retrograde fashion in COP-I-coated vesicles to modify the cargo proteins in earlier cisternae. Mechanisms of protein exit from the TGN are common to both models: [4] clathrin-coated vesicles mediate late endosome (LE)-lysosome transport, while [5] other proteins are secreted in either a regulated or constitutive fashion to the plasma membrane or extracellular space.

receptors. For example, mammalian BiP is a soluble ER protein that has the carboxy-terminal four amino acid sequence lysine–aspartate–glutamate–leucine (KDEL). This KDEL sequence is recognized by the KDEL receptor that mediates their incorporation into COP I vesicles moving from the intermediate compartment (IC) back to the ER.

Protein Modification in the Golgi

The Golgi apparatus consists stacks of flattened cisternae that contain enzymes and other proteins involved in the further modification and processing of newly made proteins. It is separated into cis, medial, and trans cisternae, followed by a meshwork of tubules and vesicles called the trans Golgi network (TGN). The process of N-linked glycosylation is completed through the action of glycosidases and glycosyltransferases localized in specific cisternae. Likewise, the glycosylation of serine and threonine residues (O-linked glycosylation) is accomplished by other glycosyltransferases. Additional modifications also occur in the Golgi. For example, proteins and carbohydrate are sulfated by sulfotransferases and some proteins are phosphorylated on serine and threonine residues by Golgi kinases. In addition, proteins like digestive enzymes (trypsin, carboxypeptidase) and hormones (insulin) are made as inactive precursors that must be proteolytically processed to their active forms in the late Golgi or post-Golgi compartments.

Transport of Proteins through the Golgi

Currently there are two different models to explain protein transport through the Golgi (Figure 3). The vesicular transport model proposes that proteins move sequentially between the Golgi cisterna in COP-I-coated vesicles, while the cisternae themselves are stationary. Proteins not retained in the *cis* Golgi, for example, would be incorporated into coated vesicles and be transported to the *medial* Golgi, and then to the *trans* Golgi. Proteins destined for post-Golgi compartments move through successive Golgi cisternae in this fashion, being modified by the resident enzymes in each compartment (Figure 3). In the cisternal maturation model a new cisterna is formed on *cis* face of the Golgi stack from ER-derived membrane. This requires both the anterograde transport of newly synthesized proteins from the ER in COP-II-coated vesicles and the retrograde transport of *cis* Golgi enzymes from the pre-existing *cis* cisterna in COP I-coated vesicles. The new *cis* cisterna and its contents progressively mature through the stack as resident Golgi enzymes are successively introduced by COP-I coated vesicles (Figure 3). In the vesicular transport model, the resident Golgi enzymes are retained in the cisternae while the cargo moves in vesicles between the different cisternae. In contrast, in the cisternal maturation model, the "resident" enzymes are continuously moving in a retrograde fashion, while the anterograde cargo remains in the cisternae. Evidence for both mechanisms is compelling, suggesting that both mechanisms may work in parallel.

Localization of Resident Golgi Enzymes

In the context of the vesicular transport model, Golgi enzymes are retained in specific cisternae. Two mechanisms have been suggested for Golgi protein retention. The "bilayer thickness" mechanism suggests that the relatively short transmembrane regions of Golgi proteins prevent their incorporation into the wider, cholesterol-rich lipid bilayers of the transport vesicles destined for post-Golgi compartments (such as the plasma membrane), and as a result, these proteins are retained in the relatively cholesterol-poor Golgi. The "oligomerization" mechanism predicts that once an enzyme has reached its resident cisterna it forms homo- or hetero-oligomers that prevent its incorporation into transport vesicles moving to the next compartment. In the context of the cisternal maturation model, resident Golgi enzymes are actively incorporated into COP-I vesicles for retrograde transport to a new cisterna, and one might predict that the cytoplasmic tails of these proteins would interact with COP-I-coat components to allow vesicle incorporation. Interestingly, there are only a few examples where the cytoplasmic tail of a Golgi enzyme plays a primary role in its localization, whereas the membrane-spanning regions of these proteins seem to be more critical. Again, it is possible that some or all of these mechanisms work together to maintain the steady-state localization of the resident Golgi proteins.

Protein Exit from the Golgi and Targeting to Post-Golgi Locations

Protein Exit from the Golgi

Once proteins reach the TGN they are sorted to post-Golgi compartments that include the lysosome, the plasma membrane, and the extracellular space (Figure 3). Trafficking to the lysosome–endosome system involves clathrin-coated vesicles similar to those that function in the uptake of proteins in endocytosis. In contrast, transport to the plasma membrane, or exocytosis, can occur either constitutively in secretory vesicles/tubules or in a regulated fashion from secretory granules found in specific cell types.

PROTEIN TARGETING TO THE LYSOSOME

The lysosome is a degradative compartment that contains numerous acid hydrolases that function to digest proteins, lipids, and carbohydrates. The trafficking of the majority of lysosomal enzymes to the lysosome requires mannose 6-phosphate residues on these enzymes' N-linked sugars. The mannose 6-phosphate residues are recognized by receptors in the TGN that mediate the incorporation of the new lysosomal enzymes into clathrin-coated vesicles destined for the late endosome compartment (Figure 3). These clathrin-coated vesicles move from the TGN and fuse with the late endosome, where a decrease in lumenal pH causes the lysosomal enzymes to dissociate from the mannose 6-phosphate receptors. The enzymes are then transported to the lysosome, while the receptors recycle to the TGN. Some lysosomal membrane proteins are also trafficked in clathrin-coated vesicles to the lysosome like the soluble enzymes but without the use of a mannose 6-phosphate marker, while others are transported to the cell surface, incorporated into a different set of clathrin-coated vesicles used in the process of endocytosis, and then trafficked to the lysosome via the late endosome.

CONSTITUTIVE AND REGULATED SECRETION

In many cell types, membrane-associated and soluble proteins move to the plasma membrane constitutively without a requirement for specific signals. Constitutively secreted proteins include receptors, channel proteins, cell adhesion molecules, and soluble extracellular matrix and serum proteins. Other proteins like hormones and neurotransmitters are targeted to secretory granules that are involved in regulated secretion from endocrine and exocrine cells, some types of immune cells, and neurons. These granules remain in a secretion-ready state until extracellular signals that lead to an increase in intracellular calcium levels trigger the exocytosis of their contents.

SEE ALSO THE FOLLOWING ARTICLES

Chaperones, Molecular • Endoplasmic Reticulum-Associated Protein Degradation • Glycoproteins, N-linked • Golgi Complex • Protein Folding and Assembly • Protein Glycosylation, Overview

GLOSSARY

chaperone A protein that aids in the folding and assembly of other proteins, frequently by preventing the aggregation of folding intermediates.

cisternal maturation/progression One model of protein transport through the Golgi apparatus that suggests that secretory cargo enters a new cisternae that forms at the *cis* face of the Golgi stack, and that this cisternae and its cargo progresses or matures through the stack by the sequential introduction of Golgi modification enzymes.

glycosylation The modification of lipids and proteins with carbohydrates in the endoplasmic reticulum and Golgi apparatus of the secretory pathway.

secretory pathway An intracellular pathway consisting of the endoplasmic reticulum, Golgi apparatus, and associated vesicles that is responsible for the folding, modification, and transport of proteins to the lysosome, plasma membrane, and extracellular space.

vesicular transport One model of protein transport through the Golgi apparatus, which suggests that secretory cargo moves sequentially between stationary Golgi cisternae in transport vesicles and is modified by resident Golgi enzymes in the process.

FURTHER READING

Ellgaard, L., Molinari, M., and Helenius, A. (1999). Setting the standards: Quality control in the secretory pathway. *Science* **286**, 1882–1888.

Farquhar, M. G., and Palade, G. E. 1998. The Golgi apparatus: 100 years of progress and controversy. *Trends Cell Biol.* **8**, 2–10.

Intracellular compartments and protein sorting (chapter 12) and intracellular vesicular traffic (chapter 13). In *The Molecular Biology of the Cell* (B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, eds.), 4th edition, pp. 659–766. Garland Science, New York.

Kornfeld, S., and Mellman, I. (1989). The biogenesis of lysosomes. *Annu. Rev. Cell Biol.* **5**, 483–525.

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. *Science* **189**, 347–358.

Rapoport, T. A., Jungnickel, B., and Kutay, U. (1996). Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu. Rev. Biochem.* **65**, 271–303.

Rockefeller University web site describing Dr. Günter Blobel's Nobel Prize research (http://www.rockefeller.edu).

BIOGRAPHY

Karen J. Colley is a Professor in the Department of Biochemistry and Molecular Genetics at the University of Illinois College of Medicine in Chicago. She holds a Ph.D. from Washington University in St. Louis, and received her postdoctoral training at the University of California, Los Angeles. Her principal research interests are in protein trafficking and glycosylation. Her recent studies focus on the elucidation of the signals and mechanisms of Golgi glycosyltransferase localization.

Selenoprotein Synthesis

August Böck
University of Munich, Munich, Germany

Selenoproteins contain one or more residues of the nonstandard amino acid selenocysteine, which is an analogue of cysteine in which a selenol group replaces a thiol. The majority of these proteins catalyze some oxidation/reduction function in which the selenol of the selenocysteine that is present in the active site of the respective enzyme takes part in the reaction. The advantage of having a selenol instead of a thiol lies in the fact that it confers to these enzymes a higher kinetic efficiency. In some biological systems, selenoproteins may also fulfill a structural role because of their capacity to oligomerize proteins via the formation of diselenide or mixed disulfide–selenide bridges. The biosynthesis of selenoproteins is unique since the incorporation of selenocysteine occurs co-translationally by the ribosome and not posttranslationally. Selenocysteine insertion is DNA encoded, requires the function of a cognate tRNA and of a specific translation elongation factor different from elongation factor Tu. Selenocysteine, therefore, has been designated as the 21st amino acid.

Bacterial Selenoprotein Synthesis

The structure and the function of the components involved in selenocysteine biosynthesis have been characterized to a considerable extent in the case of the bacterial system. The process can be divided into three functional steps, namely the biosynthesis of selenocysteine in the tRNA-bound state, the formation of a complex between elongation factor SelB, GTP, selenocysteyl-tRNA^{Sec} and the mRNA, and the decoding event at the ribosome. As far as it is known, though there are some major differences, similarities also exist between bacterial selenoprotein synthesis and the process characteristic of eukarya and archaea.

SELENOCYSTEINE BIOSYNTHESIS

Figure 1 summarizes the process of selenocysteine biosynthesis as it has been worked out for *Escherichia coli*. It requires the activities of three enzymes, namely seryl-tRNA synthetase (SerS), selenophosphate synthetase (SelD), and selenocysteine synthase (SelA) plus the specific tRNA (tRNA^{Sec}). Seryl-tRNA synthetase charges tRNA^{Sec} with L-serine, selenocysteine synthase converts the seryl-tRNA^{Sec} into selenocysteyl-tRNA^{Sec}

using selenophosphate as a source for activated selenium. Selenophosphate is provided by selenophosphate synthetase from selenide in an ATP-dependent reaction. The genes for these components had been identified with the aid of *E. coli* mutants isolated by Mandrand–Berthelot as being pleiotropically deficient in formate dehydrogenase activities.

tRNA^{Sec}

tRNA^{Sec} (Figure 2) is the key molecule of selenoprotein synthesis since it serves both as the adaptor for selenocysteine biosynthesis and for incorporation of the amino acid at the ribosome. It deviates in size, secondary structure, and in normally conserved sequence positions from canonical elongator tRNA species. Because of the elongated extra arm and the one base-pair-extended aminoacyl acceptor arm, tRNA^{Sec} species are the largest members of the elongator tRNA family. All tRNA^{Sec} species identified thus far possess a UCA anticodon which enables them to pair with UGA stop codons (but only if these are in a special mRNA sequence context). Moreover, tRNA^{Sec} species deviate from canonical elongator tRNA species in sequence positions which are usually invariant and which are involved in the establishment of novel tertiary interactions within the molecule.

On the basis of its serine identity elements, tRNA^{Sec} is charged by the cellular seryl-tRNA synthetase which also aminoacylates serine inserting isoacceptors. However, both the affinity and the rate of aminacylation are diminished in comparison to the charging of tRNA^{Ser}, resulting in an overall 100-fold reduced efficiency.

Selenocysteine Synthase

The overall reaction catalyzed by selenocysteine synthase consists in the exchange of the hydroxylgroup of the serine moiety of seryl-tRNA^{Sec} by a selenol group (Figure 1). The reaction occurs in two steps; first, the amino group of serine forms a Schiff base with the carbonyl of the pyridoxal 5′-phosphate cofactor of selenocysteine synthase leads to the 2,3-elimination of a water molecule and the formation of dehydroalanyl-tRNA^{Sec} and second, nucleophilic addition of reduced
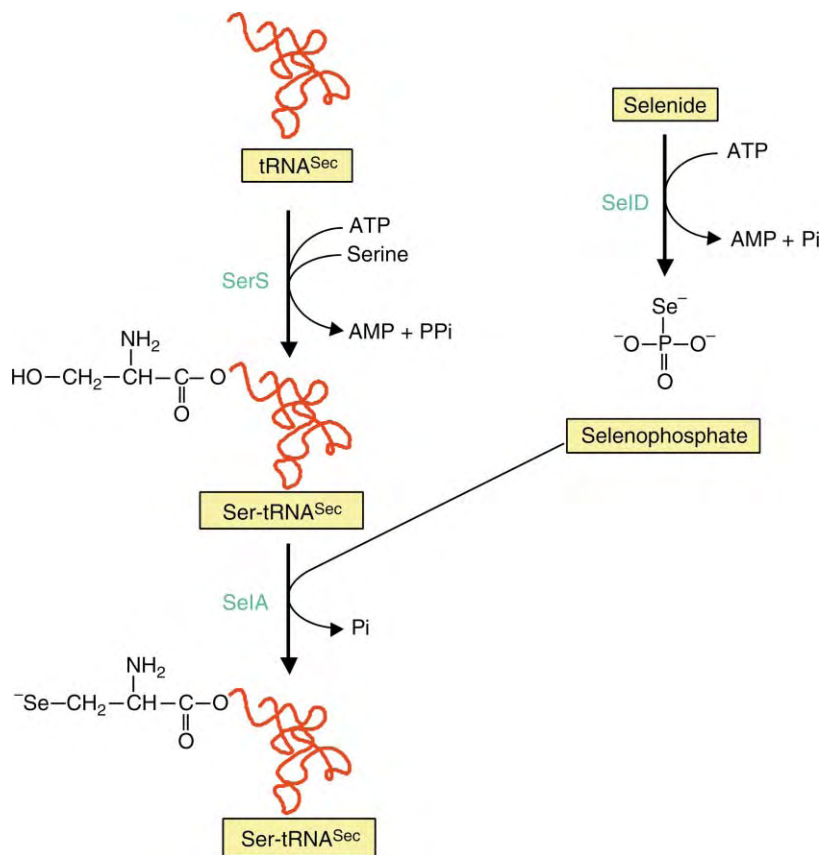
FIGURE 1   Path of selenocysteine biosynthesis. For explanation see text.

selenium to the double bond of dehydroalanyl-tRNA^{Sec} from selenophosphate as a donor yields selenocysteyl-tRNA^{Sec}.

Selenocysteine synthase from *E. coli* is a homodecameric enzyme and low resolution electron microscopy revealed that it is made up of two pentameric rings stacked on top of each other. Two subunits each are able to bind one molecule of seryl-tRNA^{Sec}, so the fully loaded enzyme can complex five charged tRNA molecules. As serine isoacceptor tRNAs are not recognized, the tRNA must have determinants for the specific recognition of seryl-tRNA^{Sec} by selenocysteine synthase and antideterminants for the rejection of seryl-tRNA^{Ser} species. The specificity for discrimination of the selenium donor is not as strict since the purified enzyme accepts thiophosphate instead of selenophosphate as a substrate. This results in the formation of cysteyl-tRNA^{Sec}. So the discrimination between sulfur and selenium must take place at some other step of selenocysteine biosynthesis.

Selenophosphate Synthetase

Purified selenocysteine synthase does not exhibit an absolute requirement for selenophosphate, as a substrate to convert seryl-tRNA^{Sec} into selenocysteyl-tRNA^{Sec}, since the reaction also occurs in the presence of high concentrations of selenide. Even sulfide is accepted although at a very low efficiency. So, the necessity for selenophosphate as a substrate may reside in one or more of the following three aspects, i.e., (1) to discriminate sulfide from selenide, (2) to efficiently use low concentrations of selenium compounds, and (3) to accelerate the reaction rate effected by the activation of the trace element. Indeed, selenophosphate synthetase efficiently discriminates between sulfide and selenide, and thus excludes sulfur from intrusion into the selenium pathway. Selenophosphate synthetase from *E. coli* is a monomeric enzyme with a unique reaction mechanism since formally it transfers the γ-phosphate of ATP to selenide with the intermediate formation of enzyme-bound ADP which is subsequently hydrolysed into AMP and inorganic phosphate.

FORMATION OF THE SelB × GTP × SELENOCYSTEYL-tRNA × SECIS COMPLEX

Elongation factor Tu, which forms a complex with all 20 standard aminoacyl-tRNAs and donates them to the ribosomal A-site, displays only minimal binding affinity
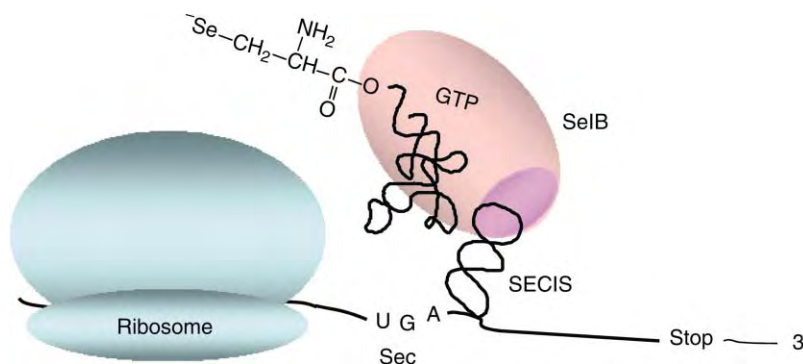
FIGURE 2 Cloverleaf presentation of the structure of tRNA^{Sec} from *E. coli*. Modified bases are shaded. Tertiary interactions via base pairing are indicated by connecting red lines, and those involving intercalation are denoted by arrows. Bases and base pairings deviating from the consensus are indicated in green.

namely the binding of guanosine nucleotides and of charged tRNA. An important difference, however, is that they can discriminate between the serylated and the selenocysteylated forms of tRNA^{Sec}. In this way, the insertion of serine instead of selenocysteine, which would lead to an inactive enzyme, is prevented. The structural basis for this discrimination ability has not yet been resolved. A second difference is that the overall affinity for GTP is about 10-fold higher than that for GDP which obviates the need for the function of a guanosine nucleotide release factor since GDP is chemically replaced by GTP. In accordance, the structure of SelB lacks those subdomains which in elongation factor Tu are responsible for interaction with the GDP release factor EF-Ts. The 25 kDa C-terminal extension of SelB (domain IV) is required for the function in selenoprotein synthesis since its truncation inactivates the molecule. The reason is that subdomain IVb binds to a secondary structure of the mRNA (the SECIS element) coding for selenoprotein synthesis. SelB, thus, is able to form a quaternary complex with GTP and two RNA ligands, namely selenocysteyl-tRNA and the SECIS element (Figure 3). The isolated domains IV or IVb retain the binding capacity for the SECIS element. Formation of the quaternary complex follows random order kinetics. An important feature also is that the stability of the complex is increased when both RNA ligands are bound.

The SECIS element itself is a hairpin structure formed within a section of 39 bases of the selenoprotein mRNA which follows the codon specifying selenocysteine insertion at the 3′-side. Binding of SelB takes place to its apical stem loop minihelix of 17 nucleotides. Genetic and structural analysis showed that bases in the loop region plus a bulged-out U in the helix are required for the interaction with SelB. This apical part of the SECIS element is separated by a short unpaired region from a helix at the base of the hairpin. Pairing within this second helix is not essential but it increases the efficiency of selenocysteine insertion. An absolute requirement,

for selenocysteyl-tRNA^{Sec}. Consequently, insertion of selenocysteine requires the function of an alternate elongation factor which is SelB. SelB from *E. coli* is a 70 kDa protein which contains the sequence elements of elongation factor Tu in the N-terminal two-third of the molecule (designated domains I, II, and III) plus a domain IV of about 25 kDa which can be subdivided into domains IVa and IVb. Domains I, II and III share their functions with those of elongation factor Tu,

FIGURE 3 Translation of prokaryotic selenoprotein mRNA. Note that the SECIS element is within the mRNA reading frame and is complexed to domain IVb of SelB carrying selenocysteyl-tRNA^{Sec} and GTP.

however, is that the codon determining selenocysteine insertion lies within a critical distance relative to the binding site of SelB.

Bacterial SECIS elements lie within the reading frame of their selenoprotein mRNAs; they are thus subject to stringent sequence constraints in order to deliver a functional gene product. However, they do not need to be translated since they also function when placed in the 3′-untranslated region at the correct distance to the selenocysteine codon within an upstream reading frame. The sequence constraint (which depends on the protein to be formed) and the requirement for binding of SelB restricts the number of selenoprotein mRNAs to be expressed in a single organism and explains why the vast majority of selenoprotein genes cannot be heterologously expressed unless the cognate SelB gene is coexpressed. Thus, SelB and their SECIS elements are subject to coevolution.

DECODING EVENT AT THE RIBOSOME

In all biological systems analyzed thus far, selenocysteine insertion is directed by the opal stop codon UGA but only if it is followed by an SECIS element at the correct distance. This violates the dogma that no codon can have more than one meaning within a single cell. The questions to be answered therefore are: (1) what prevents the UGA to be used as a termination signal, and (2) which mechanism interferes with insertion of selenocysteine at ordinary UGA stop codons?

Counteraction of Stop at the UGA?

A convincing answer to the question why the selenocysteine-specific UGA codon does not function as an efficient termination signal must await structural information on the decoding complex. It is, however, clear that termination always competes with selenocysteine insertion, especially under conditions when the capacity for decoding the UGA with selenocysteine is a limiting factor. This can be, for example, a surplus of selenoprotein mRNA in relation to the amount of SelB quaternary complex which forces the ribosome to stall at the UGA. One fact identified to be involved in the suppression of termination is that the base following the UGA at the 3′-side in selenoprotein mRNAs is prodominately an A or C, which renders the UGA a weak termination signal. Also, the two amino acids preceding selenocysteine in the nascent polypeptide chain are predominantly hydrophobic which counteracts the dissociation of the nascent polypeptide from the ribosome, when translation pauses at a "hungry" codon present in the A site. Additional mechanisms, however, must exist which contribute to the suppression of termination.

Selenocysteine Specificity of UGA Codons

From the colinearity between the mRNA nucleotide sequence and the amino acid sequence of the translation product, it is clear that UGA determines the *position* where selenocysteine is to be inserted during translation. The *specificity* of the UGA, however, is determined by the codon context, i.e., by the existence of a SECIS element at the 3′ side. The results of extensive biochemical and biophysical analysis suggest the following scenario for the decoding process: (1) SelB forms the quaternary complex at the mRNA in which the two RNA ligands display cooperativity in their interaction with the protein; (2) in this quaternary complex SelB attains a conformation compatible for interaction with the ribosome which then results in stimulation of GTP hydrolysis by SelB which in turn causes the release of the charged tRNA in the vicinity of the ribosomal A-site; (3) loss of the tRNA ligand causes the SelB protein to return to a conformation with about tenfold lower affinity for the SECIS element. As a consequence, the mRNA is released from the protein and freed for the translation of codons downstream of the UGA. The consequence of the complex cascade of reactions is that the efficiency of the decoding of UGA with selenocysteine is lower than that of any of the standard sense codons. It is also reflected by a considerable pause taking place when the ribosome encounters the quaternary complex at the mRNA. In the absence of selenocysteyl-tRNA, binding of SelB alone to the mRNA does not retard the rate of translation.

Archaeal and Eukaryal Selenoprotein Synthesis

tRNA^Sec species from archaea and eukarya share several structural similarities with the bacterial counterparts but they are more related to each other than either one is to bacterial tRNA^Sec. There is also considerable sequence similarity between selenophosphate synthetases from all three lines of descent rendering their annotation in genome projects easy. On the other hand, homologues for the bacterial selenocysteine synthase have not been identified yet in any of the genomic sequences from organisms known to synthesize selenoproteins.

Whereas UGA directs selenocysteine insertion also in archaea and eukarya, a fundamental difference is that the SECIS element is not positioned within the reading frame but in the 3′-nontranslated region of the mRNA. SECIS elements from organisms of the three lines of descent are different by sequence and by secondary structure. They may be positioned at different distances from the actual termination codon and/or the selenocysteine inserting UGA codon but a critical distance must not be underpassed. It is thought that the selective value for having the SECIS element in the nontranslated
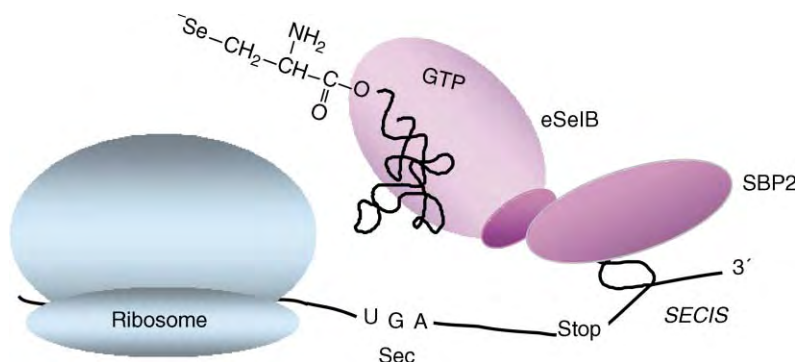
FIGURE 4 Translation of eukaryal selenoprotein mRNA. Note that the SECIS element is in the 3′-nontranslated region and serves as the binding site for SBP2 which in turn interacts with eukaryal SelB protein.

region consists in liberating it from the sequence constraint, and thus, allowing the translation of mRNAs with more than one UGA codon specifying selenocysteine insertion. Indeed, proteins with up to 17 selenocysteine residues are formed in some eukaryotes and in one instance a polypeptide with two such amino acids is synthesized in an archaeon.

Parallel to this deviation in both sequence, structure and position of the SECIS element, there is an alteration of the structure of the archaeal and eukaryal SelB-like translation factors. Domains I, II, and III closely resemble the three homologous domains from the bacterial SelB. However, the C-terminal extension is only short, less than 10 kDa, and accordingly and not unexpectedly, the archaeal and eukaryal SelB homologues do not bind to their cognate SECIS structures. In eukarya a second protein is fulfilling this task, namely SBP2 (SECIS binding protein 2) (Figure 4). There is evidence that SBP2 interacts with the SelB protein by direct contact in the decoding process. This interaction is stabilized in the presence of selenocysteyl-tRNA^{Sec}. However, the precise function of SBP2 has not yet been resolved.

SEE ALSO THE FOLLOWING ARTICLES

EF-G and EF-Tu Structures and Translation Elongation in Bacteria • Ribozyme Structural Elements: Hairpin Ribozyme • Translation Termination and Ribosome Recycling

GLOSSARY

elongation factor Helper protein assisting the ribosome in the polypeptide elongation process.

nonstandard amino acid Amino acid whose insertion is achieved by an expansion of the classical genetic code.

SECIS Selenocysteine insertion sequence of the mRNA which redefines a UGA stop codon, in a sense, codon for the insertion of selenocysteine.

selenoprotein Protein with one or more selenocysteine residues.

stop codon A codon signaling chain termination in protein synthesis in the classical genetic code UGA, UAA or UAG.

tRNA RNA molecule carrying an amino acid at its 3′-end and functioning as an adaptor to incorporate the amino acid into the growing polypeptide chain according to the triplet sequence of the mRNA.

FURTHER READING

Atkins, J. F., Böck, A., Matsufuji, S., and Gesteland, R. F. (1999). Dynamics of the genetic code. In *The RNA World* (R. F. Gesteland, T. R. Cech and J. F. Atkins, eds.) 2nd edition, pp. 637–673. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Copeland, P. R., Fletcher, J. E., Carlson, B. A., Hatfield, D. I., and Driscoll, D. M. (2000). A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J.* **19**, 306–314.

Flohe, L., Andreesen, J. R., Brigelius-Flohe, B., Maiorino, M., and Ursini, F. (2000). Selenium, the element of the moon, in life on earth. *Life* **49**, 411–420.

Hatfield, D. I. (ed.) (2001). *Selenium: Its Molecular Biology and Role in Human Health*. Kluwer, Academic Publishers, New York.

Krol, A. (2002). Evolutionary different RNA motifs and RNA–protein complexes to achieve selenoprotein synthesis. *Biochimie* **84**, 765–774.

Rother, M., Resch, A., Wilting, R., and Böck, A. (2001). Selenoprotein synthesis in archaea. *BioFactors* **14**, 75–83.

Stadtman, T. C. (1996). Selenocysteine. *Annu. Rev. Biochem.* **65**, 83–100.

BIOGRAPHY

August Böck is Professor Emeritus and former holder of the chair in Microbiology at the University of Regensburg from 1971 to 1978 and at the University of Munich from 1978 to 2002. He pursued his education at the University of Munich and his postdoctoral training at Purdue University. His main research interests are in microbial physiology with special emphasis on selenium biochemistry, bacterial metabolism and its regulation, and prokaryotic protein synthesis.

Septins and Cytokinesis

Makoto Kinoshita and Christine M. Field
Harvard Medical School, Boston, Massachusetts, USA

Septins are a family of conserved GTPases that has been identified in most animals from yeast to mammals. Each organism has multiple family members. Biochemical and genetic evidence indicate that multiple septin polypeptides form large, discrete complexes that further multimerize into filaments and higher order assemblies. Septins have been implicated in a variety of cellular processes including cytokinesis, vesicle trafficking, and axon migration. In yeast, they are involved in bud-site selection, cell polarity, and cytokinesis. Their name derives from their requirement during the final separation of the daughter cells in yeast, a process termed septation. While the precise molecular functions of septins are not known, a unifying hypothesis considers septin assemblies as scaffolds that localize, and perhaps regulate, diverse proteins involved in cortical dynamics. The septin scaffold may also have a fence-like function, limiting diffusion of proteins in the plane of the plasma membrane.

Cytokinesis

Cytokinesis is the process that physically separates the two daughter cells at the end of each division cycle. It must be temporally and spatially coupled to chromosome segregation to ensure that each daughter cell receives the correct number of chromosomes. The initiation of cytokinesis is controlled by cell-cycle-regulatory proteins, with the first step, the positioning of the cleavage plane (the site of division) beginning in late anaphase. In metazoa, the division site is determined by microtubules derived from the mitotic spindle. Next, a contractile ring made of actin, myosin-II, and other associated proteins, including septins, assembles at the plasma membrane at the specified position. The cleavage furrow ingresses by a combination of ring contraction driven by myosin-II, and targeted insertion of vesicles near the furrow to supply new plasma membrane. Finally, cytokinesis is completed in a complex process that involves the disassembly of the cleavage furrow and underlying microtubule structures, plasma membrane sealing and abscission (the actual separation) of the daughter cells. Targeted exocytosis and protein degradation are implicated in this completion phase (see Figure 1).

Septins localize to the cleavage furrow in all organisms that have been studied. Deletion, mutation, or inhibition of septins typically results in incomplete or abortive cytokinesis, though the severity of the defect varies between organisms. This suggests a conserved function of septins in cytokinesis that is not required for cleavage plane specification, but is required for normal furrow ingression, and/or completion.

Biochemical and Structural Properties of the Septins

SEPTINS BIND GUANINE NUCLEOTIDE AND FORM COMPLEXES AND FILAMENTS

Sequence analysis shows that all septins have a central globular domain containing conserved motifs found in small GTPases, and most septins have a C-terminal predicted coiled-coil region of variable length. On purification, septins are found in large complexes containing multiple septin polypeptides. The septin complex purified from *Drosophila* embryos is composed of three septin polypeptides with a stoichiometry of 2:2:2. Yeast complexes contain a fourth septin polypeptide, and complexes from mammalian brain are heterogeneous, and may be built from at least six different septin proteins. When viewed by negative-stain electron microscopy (EM) a typical septin preparation appears as filaments 7–9 nm thick and of variable lengths. The shortest filament represents the complex itself and is the building block from which the longer filaments are formed (see Figure 2A and 2B) which show a purified yeast complex.

Purified septin complexes contain tightly bound guanine nucleotide at a level of one molecule per septin polypeptide. The GDP:GTP ratio is ~2:1 for both *Drosophila* embryos and yeast complexes. The role of bound nucleotide in septin biochemistry is still under investigation, and appears to be distinctly different from small GTPases whose function employs
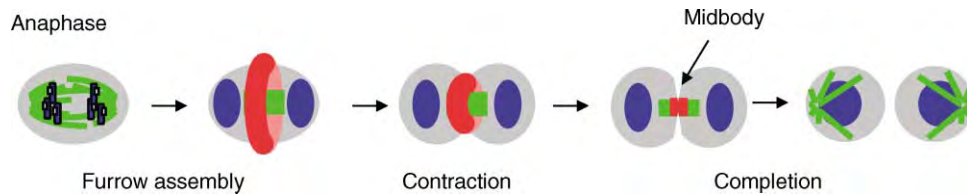
FIGURE 1 Schematic illustration of the different subprocesses of cytokinesis. DNA is shown in blue, microtubules (MT) in green, and the cleavage furrow/contractile ring (CR) in red. When the cleavage furrow assembles and contracts, microtubules become bundled and compacted into the midbody. Cytokinesis is completed by disassembly of the CR and MT structures and fusion of the membrane to create two daughter cells.

rapid exchange and hydrolysis. Isolated septin complexes exchange bound nucleotide very slowly, and in yeast, the majority of bound nucleotide does not turn over. These data suggest that GTP is bound during septin folding or complex assembly, and thereafter is not exchanged, at least on the majority of septins. Thus bound GTP may play a structural role, analogous to GTP bound to α-tubulin, and not a regulatory role, analogous to nucleotide in β-tubulin or small GTPases. However it is possible that GTP exchange and hydrolysis plays a more dynamic regulatory role for a subset of septins.

SEPTIN FILAMENTS CAN FORM HIGHER-ORDER ASSEMBLIES

Unit septin complexes are able to assemble into several different higher-order structures *in vitro*. Septin complexes from all organisms studied tend to polymerize end-to-end to form long filaments of the same thickness as the unit complexes. With yeast septins, these filaments tend to associate side by side in pairs a fixed distance apart, suggesting they may be cross-bridged by one of the septin polypeptides (see Figure 2C).
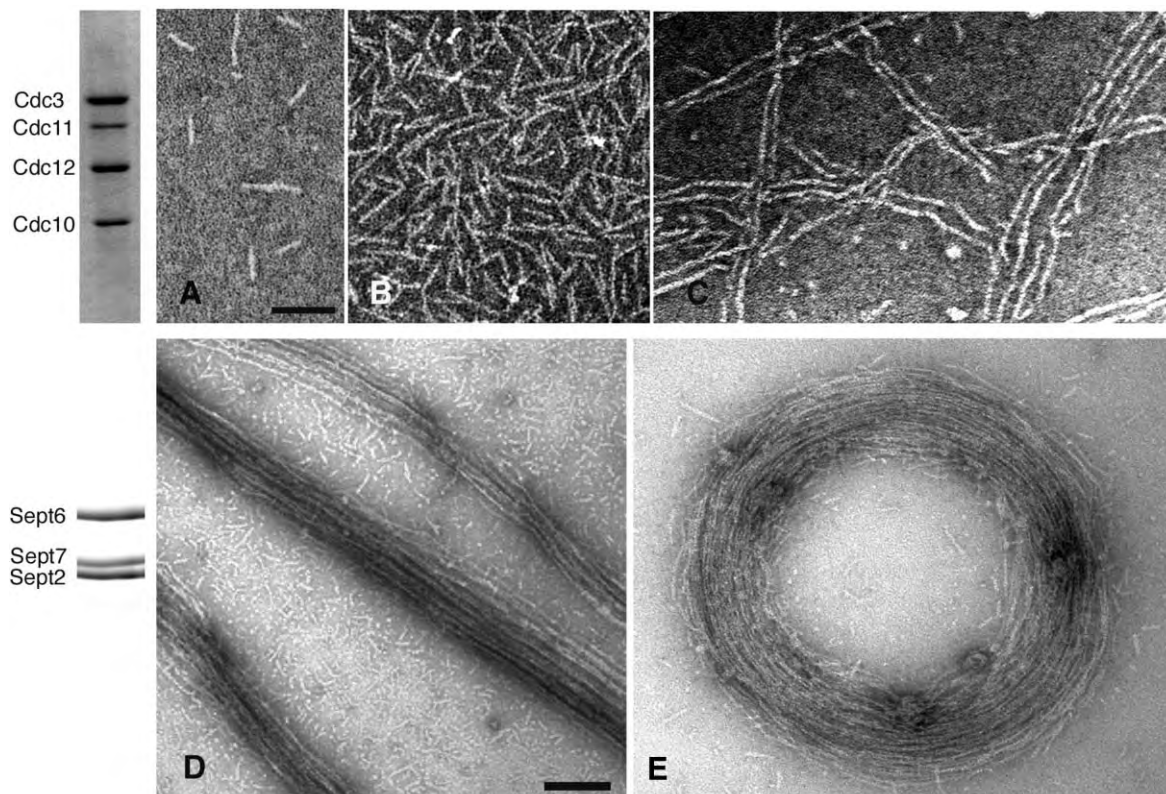
FIGURE 2 Negative stain electron micrographs of septin structures. (A–C) Examples of filamentous structures formed by a four polypeptide septin complex purified from *S. cerevisiae*. (A) Monomers and dimers. (Modified from Byers, B., and Goetsch, L. (1976). A highly ordered ring of membrane-associated filaments in budding yeast. *J. Cell Biol.* **69**, 717–721.) (B) Filaments of variable lengths. (C) Long paired filaments. (Courtesy of J. Frazier.) (D) and (E) are higher order structures formed by a three polypeptide recombinant mammalian septin complex. Complexes polymerize into long filaments that bundle (D) and curl up to form rings (E). Coomassie stained polyacrylamide gel analysis of complexes are on the left. Scale bars are 100 nm. (E is reproduced from Kinoshita, *et al.* (2002). *Develop. Cell.* **3**, 791–802, with permission from Elsevier.)
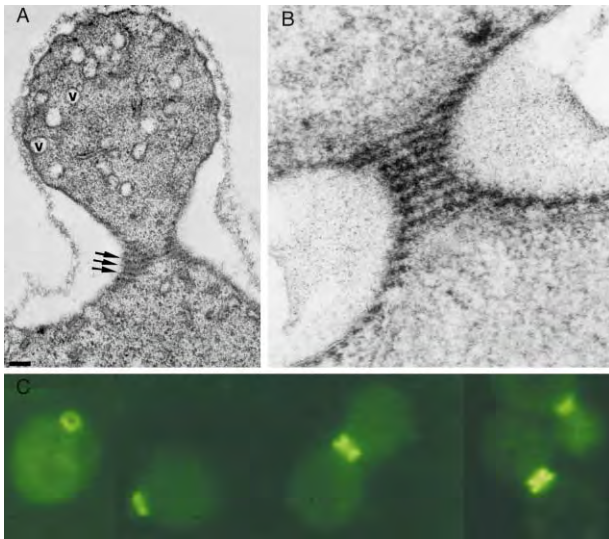
FIGURE 3   Septin structures/localization in *S. cerevisiae*. (A) and (B) are electron micrographs showing grazing sections through an early bud (A) and the neck of a large-budded cell (B) showing the 10 nm neck filaments (arrows in A) Figure 3A reproduced with permission of The Rockefeller University Press from Byers, B., and Goetsch, L. (1976). A highly ordered ring of membrane-associated filaments in budding yeast. *J. Cell Biol.* **69**, 717–721. (B) Reproduced from Strathern, J. N., Jones, E. W. and Broach, J. R. (Eds) (1981) The Molecular Biology of the Yeast Saccharomcyes : Life Cycle and Inheritance, pp. 59–96, with permission of Cold Spring Harbor Laboratory Press. (C) Shows yeast at various stages of the cell cycle stained with an antibody against Cdc3p. (Courtesy of J. Pringle.)

Purified mammalian septin filaments tend to assemble into bundles, that under some conditions curl up into rings and coils of ~ 0.7 μm diameter (see Figures 2D and 2E). Mammalian septins also tend to assemble into rings in cells. This tendency to curve is apparently intrinsic to the septin complex, and may play a role in deforming the plasma membrane in cells. The rings are comparable in size and shape to several septin assemblies in cells, including the yeast bud neck (Figure 3) and septin rings formed in cells under stress (Figure 4C).

Mammalian septin filaments can be recruited to actin bundles by another cytokinesis furrow protein, anillin. Anillin was originally identified as an actin-bundling protein in *Drosophila*. Septins and anillin are abundant in intracellular bridges between daughter cells in conventional cytokinesis and also stable bridges formed as the result of incomplete cytokinesis in *Drosophila* embryos. It is possible that these two proteins have a structural role in supporting a narrow neck in the plasma membrane after the contractile apparatus that formed the neck disassembles at the end of cytokinesis.

SEPTINS INTERACT WITH INOSITOL PHOSPHOLIPIDS

A number of studies have suggested that septins can bind directly to lipid bilayers containing inositol lipids, an activity which may be important in septin targeting or in regulating exocytosis. The question of exactly how septins target to plasma membranes, and how these proteins are involved in vesicular trafficking, are important topics for future study.
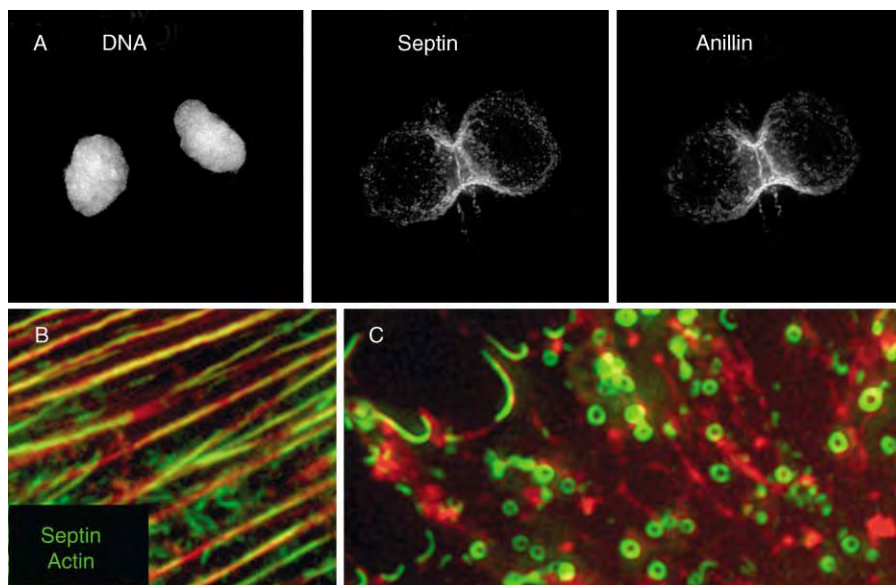
FIGURE 4   Septin structures/localization in mammalian cells. (A) A vertebrate cell in telophase showing sept7 and anillin colocalizing in the cleavage furrow. (Courtesy of Karen Oegema.) (B) and 4C Interphase cells costained for sept2 and actin. (B) Sept2 localizing along actin bundles. (C) A vertebrate cell treated with a drug that depolymerizes actin filaments. Removal of actin causes Sept2 to form rings of similar dimensions to those see by EM *in vitro* (Figure 2E). (4B and 4C are reproduced from Kinoshita, *et al.* (2002). *Develop. Cell* **3**, 791–802, with permission from Elsevier.)

Septin Behavior and Function in Cytokinesis

BUDDING YEAST

Septin proteins were originally identified in budding yeast (*Saccharomyces cerevisiae*) as the protein products of four genes *CDC3*, *CDC10*, *CDC11*, and *CDC12*. Temperature sensitive mutations of these genes exhibited hyperpolarized growth and defects in cell wall deposition and cytokinesis. These septin polypeptides localize to the mother/bud neck late in the G1 phase of the *S. cerevisiae* cell cycle, before the localization of other cleavage furrow components. Septin localization and assembly is controlled at least in part by the GTPase Cdc42, the master regulator of yeast cell polarity, and is independent of other cytoskeleton proteins including actin filaments. By EM, septins appear as 10 nm diameter filaments, termed neck filaments that appear to coil around the bud neck (Figure 3A and 3B). By immunofluorescence they appear as an hourglass-shaped assembly coating the inside of the bud neck. In projection, this hourglass can resemble two rings (Figure 3C). Photobleaching of GFP-tagged septins reveals that septins are quite dynamic before bud emergence, but once they assemble into neck filaments, their turnover rate is slowed considerably, and they can be considered static. This datum correlates well with their GTP exchange properties.

In budding yeast, septin localization at the bud neck is required for the sequential recruitment of all of the cytokinetic machinery including a type II myosin heavy chain (Myo1p), its associated light chain (Mlc1p), a formin homology (FH) protein Bni1p, probably responsible for nucleating contractile ring actin, a PCH protein (Hof1p/Cyk2p), and an IQGAP protein (Iqg1p/Cyk1p). All of these proteins have conserved roles in cytokinesis in other organisms, but it is not clear if their recruitment to the furrow depends on septins in metazoans. Neck filaments are also thought to anchor a chitin synthase complex (Chs3p/4p + Bni4p) responsible for cell-wall deposition during cytokinesis.

In budding yeast, septins are required for localization of many different proteins to the bud neck in addition to the basic cytokinesis machinery. A recent genome wide screen identified 98 proteins that localize to bud necks, and many of these depend on septins for their localization. Well-characterized examples include; the checkpoint kinases, Hsl1p, Gin4p, and Kcc4p, a component of the mitotic exit network (MEN), Dbf2/Mob1, and several proteins involved in bud site selection including Bud4.

Septins also act to restrict diffusion of proteins in the plane of the plasma membrane. The neck filaments form a fence that restricts membrane proteins to the bud, and presumably plays an important role in polarizing the yeast cell. This function may be direct as opposed to being mediated by other proteins dependent on septins for their localization.

Overall, septins play a central role in the cell biology of budding yeast. This role reflects the importance of the bud neck in cell polarity and cell division, and the function of septins as a scaffold for localizing other proteins to this site, and restricting diffusion through it. In organisms that do not grow by polarized budding, septins may be important, but perhaps their role is not as central to the overall biology of the cell.

FISSION YEAST

In fission yeast, *Schizosaccharomyces pombe*, disruption of the septin genes result in a delay in septation (cell–cell separation), but not severe cytokinetic defects as seen in budding yeast. Septins assemble into a single ring structure in late cytokinesis, and are not required to recruit actin, myosin, or other contractile ring components. Stability of the septin ring requires the protein mid2p, that is related to metazoan anillin, suggesting this interaction is conserved. Interestingly, mutations in components of the exocyst, a large complex involved in exocytosis, have a similar septation defect. Thus, in *S. pombe*, the septin scaffold is involved only in the completion stage of cytokinesis, perhaps to target exocytotic vesicles required for membrane fusion or enzymes required for final digestion of the septum.

METAZOA

In animal cells, septin polypeptides are recruited in late anaphase to the equatorial cortex and assemble into the contractile ring at the same time as actin and myosin-II. Perturbation of septins by gene disruption, RNA interference, or microinjection of anti-septin antibodies blocks normal cytokinesis in mammalian cells and fly embryos. However, septins are dispensable in some cases. For instance, nematode eggs can complete cytokinesis without septins at early developmental stages, although, cytokinesis defects manifest in some cell lineages at postembryonic stages. This difference in requirement for septins indicates a divergence in cytokinesis mechanism that we do not understand.

While the concept of septins as a scaffold for recruiting other factors is probably relevant, the exact function of septins during cytokinesis is even less clear in metazoans than it is in yeast. Septins tend to colocalize with actin filaments in interphase cells, and they tightly colocalize with anillin during cytokinesis (Figure 4). Combined with biochemical data reconstituting an actin–septin–anillin interaction *in vitro*, these data suggest a cytokinesis function involving actin filaments. However, disruption of septins does not block positioning or initial contraction of the actomyosin ring, so

septin function is not as central as it is in *S. cerevisiae*. Most likely, septins and anillin function together late in cytokinesis, perhaps during the complex process of completion.

How might septins function in completion? Physical and functional interactions have been shown between mammalian septins and syntaxins (SNARE proteins involved in membrane fusion during secretion) and the exocyst complex (a complex of proteins required for exocytosis). As previously mentioned exocyst mutants in *S. pombe* have a similar phenotype to septin mutants. Thus, it is reasonable to speculate that septins play a role in regulating or targeting vesicle insertion associated with furrow ingression and completion. This role has not yet been explored in budding yeast. Additionally, septins and anillin may assemble into a structure that stabilizes the fully ingressed membrane after the contractile ring has disassembled, but before cytokinesis is completed. These roles might explain why septins are more important during cytokinesis in some cells than others: (1) the amount of new membrane required for cytokinesis may vary according to cell type and (2) the time delay between completing ingression and actually separating the daughter cells is quite variable between species and cell type.

See Also the Following Articles

Cytokinesis • Cytokinin • Mitosis

Glossary

cell cortex A specialized layer of cytoplasm beneath the plasma membrane. In animal cells it contains actin and actin-binding proteins.

cell cycle The sequence of events by which a cell duplicates its contents and divides in two. There is a network of regulatory proteins that govern the progression through the key events such as DNA replication, formation of the mitotic spindle, and segregation of the chromosome. The cell cycle ends with cytokinesis.

exocyst complex Complex of conserved proteins that are an essential part of the exocytotic apparatus.

FH proteins Conserved proteins required to assemble some types of actin structures such as actin cables (in yeast), stress fibers, and the contractile ring.

guanosine di-/tri-phosphate (GDP/GTP) Plays an important role in tubulin stability and microtubule assembly and in signal transduction pathways via small GTPases.

inositol phospholipids Membrane lipids containing inositol and phosphate(s) that are important in various cell-signaling pathways.

midbody The thin intercellular bridge of cytoplasm connecting two daughter cells in late cytokinesis. It contains a tightly packed antiparallel array of microtubules and an electron dense matrix at its center.

photobleach The exposure of a fluorescent probe to light such that it is rendered nonfluorescent or "bleached." Examining the recovery of fluorescence after photobleaching a tagged protein gives an indication of how dynamic it is.

septation In yeast, the formation of a new cell wall or septum to separate two daughter cells. Septation is separable from cytokinesis.

small GTPase GTP-binding and -hydrolyzing switch proteins. They alternate between an active/on state when they are GTP bound and an inactive GDP bound state.

Further Reading

Field, C. M., Li, R., and Oegema, K. G. (1999). Cytokinesis in eukaryotes: A mechanistic comparison. *Curr. Opin. Cell Biol.* **11**, 68–80.

Longtine, M. S., and Bi, E. (2003). Regulation of septin organization and function in yeast. *Trends Cell Biol.* **8**, 403–409.

Longtine, M. S., Demarini, D. J., Valencik, M. L., Al-Awar, O. S., Fares, H., De Virgilio, C., and Pringle, J. R. (1996). The septins, roles in cytokinesis and other processes. *Curr. Opin. Cell Biol.* **8**, 106–119.

Mitchison, T. J., and Field, C. M. (2002). Cytoskeleton: What does GDP do for septins? *Curr. Biol.* **12**, R788–R790.

Moffat, J., and Andrews, B. (2003). Ac'septin' a signal: Kinase regulation by septins. *Dev. Cell* **5**, 528–530.

Rajagopalan, S., Wachtler, V., and Balasubramanian (2003). Cytokinesis in fission yeast: A story of rings, rafts and walls. *Trends Genetics* **19**, 403–408.

Tolliday, N., Bouquin, N., and Li, R. (2001). Assembly and regulation of the cytokinetic apparatus in budding yeast. *Curr. Opin. Microbiol.* **4**, 690–695.

Trimble, W. S. (1999). Septins: A highly conserved family of membrane associated GTPases with functions in cell division and beyond. *J. Membr. Biol.* **169**, 75–81.

Biography

Makoto Kinoshita was a Postdoctoral Research Fellow at Department of Cell Biology, Harvard Medical School and currently is an Assistant Professor at Kyoto University Graduate School of Medicine. His principal research interest is in biochemistry, cell biology, and pathology of mammalian septins.

Christine M. Field is a Research Fellow and a graduate student at Department of Systems Biology, Harvard Medical School. Her principal research interest is in dissecting molecular mechanism of cytokinesis by biochemistry and genetics using *Drosophila* and other organisms.

Serine/Threonine Phosphatases

Thomas S. Ingebritsen
Iowa State University, Ames, Iowa, USA

Phosphorylation of proteins on serine, threonine, and tyrosine residues is a major mechanism for regulating the activity of cell proteins and it plays a central role in virtually all signal transduction pathways in eukaryotes. The steady-state level of phosphorylation of a protein at a particular site depends on the balance of the activities of the protein kinase(s) and protein phosphatase(s) acting on that site. Both protein kinases and protein phosphatases are important targets of cell regulation. This article will focus on the structure, regulation, and function of the two families of protein Ser/Thr phosphatases (PPP and PPM). Members of each family are present in all three domains of life (archae, bacteria, and eukarotes).

Protein Ser/Thr Phosphatase Catalytic Subunit Families

PPP FAMILY

Members of this family possess a common catalytic core (280 residues) and can be further divided into four subfamilies termed PPP1, PPP2A, PPP2B, and PPP5. Three prokaryotic phosphatases: diadenosine tetraphosphatase, $\Phi 80$ phosphatase, λ phosphatase exhibit weaker similarity to the PPP family.

PPM FAMILY

Members include PP2C, *Arabidopsis* ABI1 and ABI2, *Arabidopsis* KAPP, *Bacillus subtilis* SpoIIE phosphatase and pyruvate dehydrogenase phosphatase. The core PPM catalytic domains occur in diverse contexts. For example the catalytic domain of *Arabidopsis* ABI1 is fused to an EF-hand domain, the *Arabidopsis* kinase associated protein phosphatase (KAPP) catalytic domain is fused to a kinase-interaction domain. The kinase interaction domain of KAPP associates with a phosphorylated receptor-like protein. The N terminus of *Bacillus subtilis* SpoIIE phosphatase is fused to a domain with ten membrane-spanning segments.

Biochemical Characterization of Signature Protein Ser/Thr Phosphatases

PP1, PP2A, and PP2B are signature members of the PPP1, PPP2A, and PPP2B subfamilies while PP2C is the signature member of the PPM family. These four enzymes account for the majority of protein Ser/Thr phosphatase activity in cell extracts. The activities of these enzymes can be distinguished in cell extracts based on divalent cation requirements and the effects of physiological and pharmacological inhibitors (Table I). PP1, PP2A, and PP2C have broad and overlapping substrate specificities whereas the substrate specificity of PP2B is more restricted.

Three-Dimensional Structures and Catalytic Mechanism

Three-dimensional structures have been determined for two members of the PPP family (PP1 and PP2B) and for one member of the PPM family. A surprising finding was that the PPP and PPM families have similar three-dimensional architectures even though the primary structures of the two families are unrelated. The three-dimensional structure of the two families is quite distinct from that of the PTP family of protein Tyr phosphatases.

THREE-DIMENSIONAL STRUCTURES

For both the PPM and PPP families, the core structure consists of a pair of mixed β-sheets that form a β-sandwich structure. The catalytic sites possess a binuclear metal ion center which has some similarity to the binuclear metal ion center of purple acid phosphatase (Figure 1). For the PPM family, the two metal ions are Mn^{2+}. In the case of the PPP family, the two metal ions are probably Fe^{2+} and Zn^{2+}, although there is some controversy about whether the second metal ion is Zn^{2+} or Mn^{2+} and also about whether the

TABLE I

Biochemical Characterization of Signature Protein Phosphatases

Type	Protein phosphatase	Substrate specificity	Divalent cation requirement	I_1 and I_2 inhibition	Okadaic acid inhibition (IC_{50})
1	PP1	Broad	None	Yes	20 nM
2	PP2A	Broad	None	No	0.2 nM
2	PP2B	Narrow	Ca^{2+}	No	5 μM
2	PP2C	Broad	Mg^{2+}	No	No effect

PP2B requires μM Ca^{2+} for activity whereas PP2C requires mM Mg^{2+}. Inhibitor-1 (I_1) and inibitor-2 (I_2) are PP1 modulator proteins. Okadaic acid is a pharmacological inhibitor of PPP family members. It is a polyether carboxylic acid that is a diarrhetic shell fish poison and a powerful tumor promoter. There are a number of other pharmacological inhibitors of the PPP family (e.g., microcystin, calyculin A, cantharidin). Okadaic acid, microcystin, and calyculin A inhibit by binding to the phosphatase active site.

Fe is in the 2+ or 3+ oxidation state. In both cases, metal ions are coordinated with water molecules.

CATALYTIC MECHANISM

For both the PPP and PPM family hydrolysis of serine or threonine phosphate esters occurs through a single-step mechanism in which a metal-bound water acts as a nucleophile to attack the phosphorus atom of the substrate phosphate group (Figure 1). The metal ion acts as a Lewis acid to enhance the nucleophilicity of metal-bound water and it also enhances the electrophilicity of the phosphorus atom by coordinating the two oxyanions of the phosphate group. An active site

FIGURE 1 Active site structure and catalytic mechanism of PP1 catalytic subunit. (Reprinted from Barford, D. (1996). Molecular mechanisms of the protein serine/threonine phosphatases. *TIBS* **21**, 407–412.)

His sidechain donates a proton to the leaving oxygen of Ser or Thr. This catalytic mechanism is quite different from that used by the PTP family which involves formation of a phospho-enzyme intermediate.

Subunit Structure of PPP Family Members

Members of the PPP family interact with diverse sets of regulatory subunits, which direct the catalytic subunits to specific subcellular locations, alter substrate specificity and/or confer regulatory properties.

INTERACTION OF PP1 WITH DIVERSE REGULATORY SUBUNITS

The catalytic subunit of PP1 (PP1c) interacts with >50 regulatory subunits, which fall into two classes: targeting subunits and modulator proteins. Targeting subunits direct PP1c to a wide variety of subcellular locations including: glycogen particle, myosin/actin, spliceosomes/RNA, endoplasmic reticulum, proteasomes, nuclear membranes, plasma membranes/cytoskeleton, centrosomes, microtubules, and mitochondria. Targeting subunits also modulate substrate specificities and may regulate phosphatase activity.

Modulators are generally low-molecular-weight, heat-stable proteins that alter PP1 activity or substrate specificity. The activity of some of the modulators (e.g., inhibitor-1, DARPP-32, CPI-17, and G subunit) is regulated through reversible phosphorylation.

Strong binding of many of the regulatory proteins to PP1c is mediated through a short motif termed RVxF. This motif is found in two-thirds of the targeting subunits and one-half of the modulator proteins. The consensus sequence of the motif is: $(K/R)x_1(V/I)x_2(F/W)$ where x_1 and x_2 may be any residue except a large hydrophobic residue. In some motifs x_1 is absent.
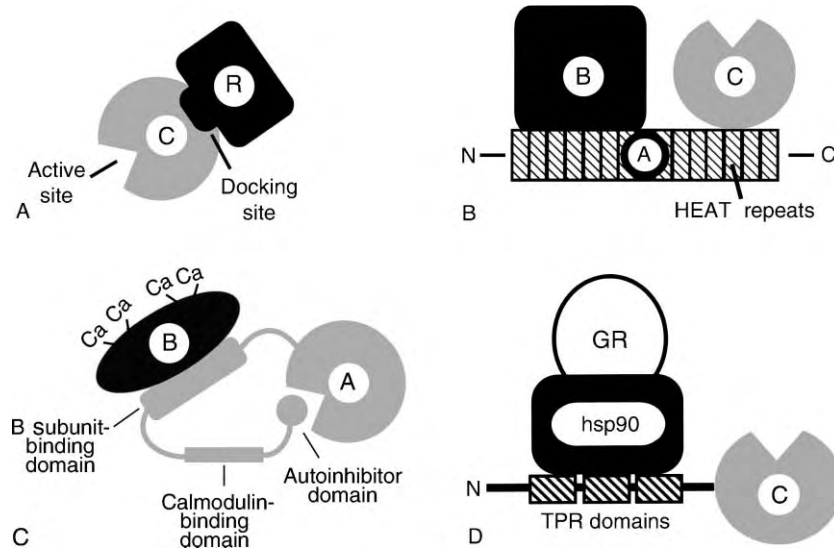
FIGURE 2 Subunit structure of PPP family members. (A) PP1. C and R designate catalytic and regulatory subunits, respectively. The docking site on the C subunit interacts with the RVxF motif of the regulatory subunit. (B) PP2A. The labels C, B, and A designate the catalytic, regulatory, and adaptor subunits, respectively. The three-dimensional structure of the A subunit has been determined. Tandem arrays of HEAT (*h*untingtin-*e*longation factor-*A* subunit-*T*OR) motifs are present in a variety of other proteins. The labels N and C designate the amino and carboxyl termini of the A subunit. (C) PP2B. The labels A and B designate the catalytic and regulatory subunits, respectively. The B subunit has four Ca^{2+}-binding sites. The B subunit-binding, calmodulin-binding and autoinhibitor domains are on a carboxyl-terminal extension of the A subunit. The region from the end of the B subunit-binding domain to the beginning of the autoinhibitor domain is disordered in the absence of calmodulin and thus not visible in the crystal structure. (D) PP5. C and GR designate the catalytic subunit and the glucocorticoid receptor, respectively. TPR domains are characterized by a degenerate 34 amino acid sequence. The label N designates the amino terminus of the C subunit. The three-dimensional structure of the amino terminal extension of C has been determined in the absence of the core catalytic domain.

The RVxF motif binds to a docking site that is remote from the active site (Figure 2A). The docking site consists of a hydrophobic groove which binds the two large hydrophobic residues of the RVxF motif and a cluster of negatively charged residues which interact with basic residues at the N-terminal end of the motif.

Additional interactions between the PP1c and the targeting and modulator proteins are thought to strengthen binding and mediate effects on PP1 activity. For example residues 7–11 of DARPP-32 (KKIQF) bind to the PP1c docking site whereas thr 34 which is phosphorylated by PKA binds to the active site of the phosphatase.

INTERACTION OF PP2A WITH A DIVERSE SET OF REGULATORY SUBUNITS

PP2A diversity is also generated by the interaction of a common catalytic (C) subunit with a diverse set of regulatory (B) subunits. However in this case, the regulatory subunits interact with the C subunit indirectly through an adapter subunit (A), thus forming a heterotrimeric phosphatase complex (Figure 2B).

Over 15 different B subunits are expressed in a tissue- and developmental-specific manner from four families of genes, termed PR55/B, PR61/B′, PR72/B′, and B′. Additional gene products are generated through alternate splicing. Functions of the B subunits include

regulation of PP2A activity, subcellular targeting, and alteration of substrate specificity.

The A subunit is made up of 15 HEAT repeats which form an extended and curved structure reminiscent of a hook or the letter C. The catalytic subunit interacts with the C-terminal HEAT repeats (11–15) whereas the B subunits interact with N-terminal repeats (1–10).

SUBUNIT STRUCTURE OF PP2B

The catalytic (A) subunit of PP2B interacts with two EF-hand-type Ca^{2+}-binding proteins, an integral B subunit which binds in the absence of calcium, and calmodulin which requires calcium for binding. The A subunit has a regulatory C-terminal extension which has binding sites for the two Ca^{2+}-binding proteins as well as an autoinhibitor site (Figure 2C). Binding of Ca^{2+} to the B subunit is absolutely required for phosphatase activity. PP2B activity is further stimulated by Ca^{2+}-calmodulin.

PP5

The catalytic subunit of PP5 has an amino-terminal extension with three TPR domains. These domains are found in a variety of proteins and act as a scaffold that mediates protein–protein interaction. The TPR domains of PP5 mediate interaction of the phosphatase with the heat shock protein, hsp90, and the glucocorticoid

receptor (Figure 2D). TPR domains also suppress the catalytic activity of PP5 25-fold.

Examples of Functions and Regulation of Protein Ser/Thr Phosphatases

PPM FAMILY

PPM family members are involved in stress responses in animals, plants, fungi, and prokaryotes. PP2C antagonizes stress response pathways involving two types of protein kinase cascades: mitogen-activated protein kinase (MAPK) and AMP-activated protein kinase (AMPK). Two other PPM family members (ABI1 and ABI2) play an essential role in the abscisic acid-mediated stress response of plants to water deprivation and in prokaryotes SpoIIE is involved in controlling sporulation.

PPM family members also have other cell functions. For example, the PDH phosphatase is involved in controlling the types of metabolic fuels used in body tissues through a dephosphorylation reaction that activates pyruvate dehydrogenase in mitochondria.

PPP FAMILY

Regulation of PP1 Involving Targeting Subunits

There are four modes of regulation of PP1c involving targeting subunits: inducible expression of targeting subunits, allosteric regulation through targeting subunits, phosphorylation near the RVxF docking motif, and phosphorylation at sites remote from the docking motif.

The expression of two glycogen-targeting subunits in rat liver, G_L and R5, is decreased by diabetes and starvation and is restored by insulin treatment and refeeding, respectively.

G_L-PP1c is subject to allosteric regulation by glycogen phosphorylase *a* (phosphorylated, active form) which binds to a short segment at the C terminus of G_L with nanomolar affinity. Phosphorylase *a* binding inhibits the glycogen synthase phosphatase activity of G_L-PP1c but has no effect on phosphorylase phosphatase activity of the complex. This helps to coordinate the regulation of glycogen breakdown and synthesis in response to glucagon and perhaps to insulin.

G_M (muscle-specific glycogen targeting), NIPP1 (nuclear targeting), and Neurabin I (actin targeting) are phosphorylated by protein kinase A at sites near the RVxF docking motif leading to dissociation of PP1c. In the case of G_M-PP1c this results in decreased activity towards glycogen-bound substrates (glycogen synthase, phosphorylase *a*, and phosphorylase kinase).

The myosin-targeting protein M110 is phosphorylated on sites (Thr 697 and Ser 435) that are distant from the docking motif. The complex of PP1c with M110 is involved in regulating muscle contraction in smooth muscle and nonmuscle cells. Phosphorylation of Ser 435 occurs during mitosis and leads to activation of PP1c and enhanced binding to myosin II.

PP2A

One of the best-documented roles of PP2A is in the regulation of animal growth and development. A role in cell growth was first suggested by the potent inhibition of PP2A by the tumor promoter okadaic acid. Additionally the β-isoform of the A subunit of PP2A has been identified as a candidate tumor-suppressor gene and the myeloid-leukemia-associated protein SET is a potent inhibitor of PP2A.

PP2A dephosphorylates and inactivates protein kinases involved in growth-regulatory signal transduction pathways (e.g., ERK and Mek MAP kinases, protein kinase C, and protein kinase B). Additionally, PP2A is an important cellular target of the SV40 and polyoma DNA tumor viruses. Viral proteins associate with the AC dimer and displace B subunits. This leads to stimulation of growth-related (ERK/Mek) MAP kinase pathways and cell growth.

Genetic approaches in budding and fission yeast as well as in *Drosophila* demonstrate that PP2A has an essential role in regulating the cell cycle. Additionally, PP2A interacts with components of the Wnt signaling cascade, which controls the epithelial–mesenchymal transition during vertebrate development.

Central Role of PP2B in T Cell Activation

Stimulation of the T cell receptor leads to the activation of dual signal transduction pathways involving Ca^{2+} and Ras (Figure 3, left). Elevation of intracellular Ca^{2+} results in activation of PP2B which dephosphorylates NFAT. This leads to activation of NFAT as a transcription factor and translocation of the NFAT from the cytoplasm to the nucleus. The Ras pathway activates a nuclear transcription factor (AP-1) through a protein kinase cascade. NFAT and AP-1 bind cooperatively to DNA regulatory sites resulting in enhanced transcription of cytokine, cell surface receptor, and transcription factor genes.

PP2B is the site of action for the immune-suppressant drugs, Cyclosporin A and FK506, used to prevent rejection in organ transplant procedures. The use of these drugs has revolutionized organ transplant therapy. The two drugs interact with separate intracellular-binding proteins (cyclophin and FK506-binding protein) and the resulting complexes bind to and inhibit the activity of PP2B. This in turn inhibits T cell activation by
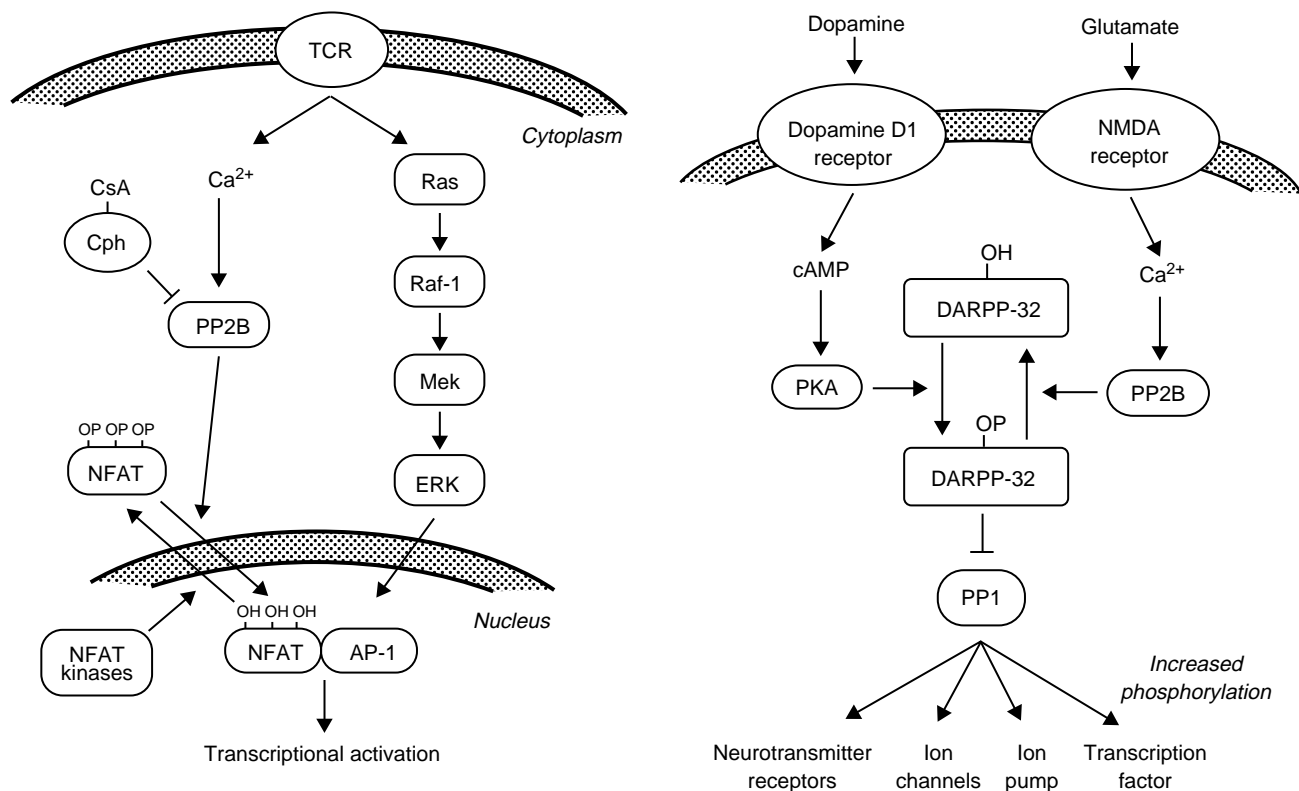
FIGURE 3 (Left) Role of PP2B in T cell activation. Raf-1, Mek, and ERK are protein kinases in the Ras signaling pathway. NFAT and AP-1 are transcription factors. NFAT are a family of transcription factors that exist in an inactive, phosphorylated state in the cytoplasm of resting T cells. These proteins are phosphorylated on multiple serine residues in a regulatory region in the amino-terminal half of the molecule. CsA and Cph designate cyclosporin and cyclophilin, respectively. (Right) Central role of DARPP-32 in the regulation of dopaminoceptive neurons. Inhibition of PP1 is associated with increased activity of NMDA and AMPA glutamate receptors, of L, N, and P type Ca^{2+} ion channels and CREB and with decreased activity of $GABA_A$ receptor, Na^+ channels, and Na^+/K^+ ATPase. There are a variety of other neurotransmitter receptors (opiate, adenosine, VIP) which elevate cAMP in medium spiny neurons. Glutamate acting through AMPA receptors and GABA acting through $GABA_A$ receptors also elevate Ca^{2+} in these neurons.

suppressing transcriptional activation through the NFAT:AP-1 complex.

Similar dual pathways involving PP2B and NFAT family members have been implicated in angiotensin-II-induced cardiac hypertrophy and in the morphogenesis of heart valves.

DARPP-32

DARPP-32 (dopamine and cyclic adenosine 3′,5′-monophosphate-regulated phosphoprotein, 32 kDa) is a specific inhibitor of PP1. It is expressed at high concentrations in medium spiny neurons of the neostriatum where it plays a central role in integrating responses to dopamine (acting via the cAMP) and glutamate (acting via Ca^{2+}) in dopaminoceptive neurons (Figure 3, right). Diseases associated with defects in dopaminergic neurotransmission include Parkinson's disease, Huntington's disease, ADHD, and schizophrenia.

Activation of the cAMP pathway in dopaminoceptive neurons leads to increased phosphorylation of DARPP-32 on Thr 34 and inhibition of PP1. This results in increased phosphorylation of neurotransmitter receptors, voltage-gated ion channels, an electrogenic ion pump (Na^+/K^+ ATPase), and a transcription factor (CREB).

In contrast, glutamate acting through NMDA receptors promotes dephosphorylation of brain proteins through activation of a protein phosphatase cascade involving PP2B and PP1. Activation of NMDA receptors elevates Ca^{2+} which activates PP2B. This leads to DARPP-32 dephosphorylation by PP2B and activation of PP1 through relief of inhibition by DARPP-32.

SEE ALSO THE FOLLOWING ARTICLES

Allosteric Regulation • Angiotensin Receptors • Dopamine Receptors • Protein Kinase B • Protein Kinase C Family • Protein Tyrosine Phosphatases • Pyruvate Dehydrogenase

GLOSSARY

modulator protein   Generally a low-molecular-weight, heat-stable protein that alters protein phosphatase activity or substrate specificity.

protein kinase cascade   A series of protein kinases arranged in a linear fashion in a signal transduction pathway such that an upstream protein kinase phosphorylates and activates the immediate downstream kinase.

protein phosphatase   An enzyme whose physiological function is to remove phosphate groups from serine, threonine, or tyrosine residues of proteins.

targeting subunit   A protein that directs a phosphatase to a specific subcellular location or a specific substrate and may also modulate substrate specificity and regulate phosphatase activity.

transcription factor   A protein that binds to a regulatory site on a gene leading to enhanced transcription of the gene.

FURTHER READING

Barford, D. (1996). Molecular mechanisms of the protein serine/threonine phosphatases. *TIBS* **21**, 407–412.

Ceulemans, H., Stalmans, W., and Bollen, M. (2002). Regulatory-driven functional diversification of protein phosphatase-1 in eukaryotic evolution. *BioEssays* **24**, 371–381.

Cohen, P. (2002). Protein phosphatase 1 – targeted in many directions. *J. Cell Sci.* **115**, 241–256.

Greengard, P., Allen, P., and Nairn, A. (1999). Beyond the dopamine receptor: The DARPP-32/protein phosphatase-1 cascade. *Neuron* **23**, 435–447.

Ingebritsen, T., and Cohen, P. (1983). Protein phosphatases: Properties and role in cellular regulation. *Science* **221**, 331–338.

Janssens, V., and Goris, J. (2001). Protein phosphatase 2A: A highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem. J.* **353**, 417–439.

Rao, A., Luo, C., and Hogan, P. (1997). Transcription factors of the NFAT family: Regulation and function. *Annu. Rev. Immunol.* **15**, 707–747.

Rodriquez, P. (1998). Protein phosphatase 2C (PP2C) function in higher plants. *Plant Mol. Biol.* **38**, 919–927.

BIOGRAPHY

Thomas S. Ingebritsen is an Associate Professor in the Department of Genetics, Development and Cell Biology at Iowa State University. His research interest is the structure, regulation, and function of protein phosphatases. He holds a Ph.D. in Biochemistry from Indiana University and received his postdoctoral training at the University of Dundee, Scotland. Together with Philip Cohen, he developed the scheme for classification of protein serine/threonine phosphatases and he has published extensively in the area of protein phosphorylation and protein phosphatases.

Serotonin Receptor Signaling

Paul J. Gresch and Elaine Sanders-Bush
Vanderbilt University, Nashville, Tennessee, USA

Serotonin (5-hydroxytryptamine; 5-HT) receptors are a family of G-protein-coupled receptors (GPCRs) and one ligand-gated ion channel that transduce an extracellular signal by the neurotransmitter serotonin to an intracellular response. 5-HT receptors are involved in multiple physiological functions such as cognition, sleep, mood, eating, sexual behavior, neuroendocrine function, and gastrointestinal (GI) motility. Since many physiological processes are influenced by 5-HT receptors, it is not surprising that dysfunction and regulation of 5-HT receptors are implicated in numerous disorders and disease states including migraine, depression, anxiety, schizophrenia, obesity, and irritable bowel syndrome. Therefore, understanding 5-HT receptor second messenger systems, their effector linkage, the multiplicity of coupling pathways, and how these pathways are regulated is critical to disease etiology and therapeutic discovery.

Serotonin Synthesis and Metabolism

The neurotransmitter serotonin (5-hydroxytryptamine; 5-HT) is found in the central nervous system, enterochromaffin cells, gastrointestinal tract, and platelets. 5-HT is synthesized from the essential amino acid tryptophan by the rate-limiting enzyme tryptophan hydroxylase and a ubiquitous l-aromatic amino acid decarboxylase. 5-HT is released into the synaptic cleft by exocytosis of vesicles in a TTX-sensitive and Ca^{2+}-dependent manner. Inactivation of 5-HT is mediated by reuptake into the presynaptic terminal through an Na^+-dependent 5-HT transporter. 5-HT is metabolized into the inactive form, 5-hydroxyindole acetic acid, by the enzymes monoamine oxidase and aldehyde dehydrogenase. Levels of synaptic 5-HT can be regulated. For instance, restriction of dietary tryptophan or chemical inhibitors of tryptophan hydroxylase reduce brain levels of 5-HT, while selective 5-HT reuptake inhibitors such as fluoxetine (Prozac) increase the amount of synaptic 5-HT. Released serotonin acts on multiple 5-HT receptors found throughout the body in various tissues including brain, spinal cord, heart, blood, and gut.

Serotonin Receptor Structure and Function

5-HT receptors are classified and characterized by their gene organization, amino acid sequences, pharmacological properties, and second messenger coupling pathways. With the exception of the $5-HT_3$ receptor, the 5-HT receptor family consists of G- protein-coupled receptors (GPCRs). The basic protein structure is predicted to contain seven transmembrane regions, three intracellular loops, and three extracellular loops, with the amino terminus being extracellular and carboxy terminus, intracellular (Figure 1A). These receptors are linked to their signal transduction pathways through guanine nucleotide triphosphate (GTP)-binding proteins (G protein). The sequence of events involves the activation of the cell surface receptor by 5-HT or drugs, binding of receptor and G protein, which promotes the exchange of bound GDP for GTP on the G protein. The G protein is comprised of α-, β-, γ-subunits; the $\beta\gamma$ dimer dissociates when receptor binds and both $G\alpha$ and $G\beta\gamma$ have the ability to interact with effector enzymes. The subunit interactions promote activation or inhibition of adenylate cyclase or activation of phospholipase C. In turn, the effector enzymes generate second messengers that regulate cellular processes such as Ca^{2+} release, and protein kinases and phosphatases. This multistep enzymatic process amplifies receptor signal, and provides the possibility of regulation and crosstalk at multiple levels. The numerous 5-HT receptors are grouped in Table 1 by their traditional (primary) G protein second messenger linkage. The multiple levels of diversity generated by RNA splicing, RNA editing, and promiscuity of receptor G-protein activation are discussed later.

Serotonin Receptors that Inhibit Adenylyl Cyclase

The five members of the $5-HT_1$ receptor family ($5-HT_{1A}$, $5-HT_{1B}$, $5-HT_{1D}$, $5-HT_{1E}$, and $5-HT_{1F}$) and the $5-HT_5$
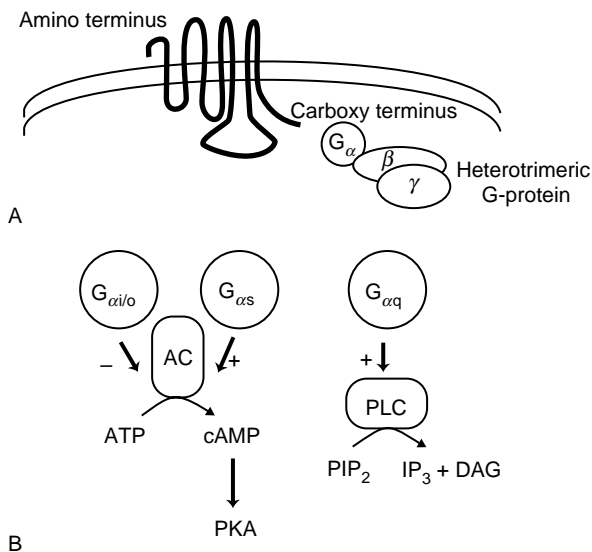
FIGURE 1 A. Schematic drawing of the basic components of a G-protein coupled receptor illustrating the extracellular amino terminus, seven transmembrane domains, three intracellular loops, three extracellular loops, intracellular carboxy terminus, and heterotrimeric G protein. B. The primary signaling pathways of the G-protein coupled receptors. Activation of $G_{\alpha i}$ protein inhibit AC resulting in the decrease of cAMP production. Receptors that stimulate AC through $G_{\alpha s}$ results in increased production of cAMP, with subsequent activation of PKA. Stimulation of PLCβ through $G_{\alpha q}$ results in the cleavage of PIP_2 into IP_3 and DAG.

receptor (5-HT$_{5A}$, 5-HT$_{5B}$ subtype) couple primarily through $G_{i/o}$ proteins to inhibit the membrane-bound enzyme, adenylyl cyclase (AC). This inhibition of AC leads to a decrease of $3'5'$-adenosine monophosphate

TABLE 1

The Primary Signal Transduction Pathway for the Serotonin Receptor Family

G protein	5-HT receptor subtype	Effector linkage
G_i/G_o	1A	Inhibition of adenylate cyclase
	1B	
	1D	
	1E	
	1F	
	5	
G_s	4	Activation of adenylate cyclase
	6	
	7	
G_q/G_{11}	2A	Activation of phospholipase C
	2B	
	2C	
No G protein	3	Ligand-gated ion channel

(cAMP) molecules (Figure 1B). The 5-HT$_1$ receptors are the best characterized of this family. High densities of 5-HT$_{1A}$ receptors are found on the cell bodies of 5-HT neurons in the brainstem nuclei, especially the dorsal raphe. In the dorsal raphe, the 5-HT$_{1A}$ receptor functions as an autoreceptor that reduces cell firing when activated. The receptor elicits neuronal membrane hyperpolarization by activating G protein-linked K$^+$ channels. In addition, 5-HT$_{1A}$ receptors are located on postsynaptic sites in the hippocampus and other limbic brain regions where they also produce hyperpolarization by opening K$^+$ channels. 5-HT$_{1A}$ receptors are targets of a class of antianxiety drugs. Of the other 5-HT$_1$ receptors, much more is known about the 5-HT$_{1B}$ and 5-HT$_{1D}$ subtypes. These receptors are found in basal ganglia and frontal cortex, and function as terminal autoreceptors or heteroreceptors that modulate neurotransmitter release. 5-HT$_{1B/1D}$ heteroreceptors have been proposed to regulate the release of acetylcholine, glutamate, dopamine, norepinephrine, and γ-aminobutyric acid (GABA). Pharmacological and genomic (knock out/deletions) studies suggest that 5-HT$_{1B}$ receptors are involved in aggressive behavior. 5-HT$_{1D}$ receptors have a role in migraine headaches and many antimigraine drugs target this receptor. Less is known about 5-HT$_{1F}$, 5-HT$_{1E}$, and 5-HT$_5$ receptors.

Serotonin Receptors Linked to Activation of Adenylate Cyclase

5-HT$_4$, 5-HT$_6$, and 5-HT$_7$ receptors are all coupled to activation of AC. These receptors are linked via the G protein G_s to AC producing an increase cAMP production (Figure 1B). Historically, the first signal-transduction pathway to be linked to a 5-HT receptor was stimulation of AC, characterized in mouse collicular neurons. This receptor now known as the 5-HT$_4$ receptor is also found in hippocampus and peripheral tissues. In the periphery, it releases acetylcholine in the ileum, contracts the esophagus and colon, promotes ion transport in the gut, and elicits cardiac contraction. In the brain, the receptor has been linked to modulation of release of acetylcholine, dopamine, 5-HT, and GABA. Little is known about the 5-HT$_6$ receptor. It is found in the striatum, amygdala, nucleus accumbens, hippocampus, and cortex. 5-HT$_7$ receptors are widely expressed in the brain, with highest expression levels in the thalamus and the hippocampus. The 5-HT$_7$ receptor may have role in circadian rhythms and thermoregulation. Both the 5-HT$_6$ and 5-HT$_7$ receptor have high affinity for many of the atypical antipsychotics leading to speculation of a role for this receptor in schizophrenia.

Serotonin Receptors Coupled to the Activation of Phospholipase C

The 5-HT$_2$ class of receptor (subtype 2A, 2B, 2C) activates the membrane-bound enzyme phospholipase C (PLC) which catalyzes the degradation of the inositol lipid, phosphotidylinositol 4,5 bisphosphate (PIP$_2$) with the production of inositol 1,4,5 triphosphate (IP$_3$) and diacylglycerol (DAG) (Figure 1B). IP$_3$ mobilizes Ca^{2+} from intracellular storage sites; Ca^{2+} then induces multiple responses in the cell including activation of calcium/calmodulin-dependent protein kinase enzymes that phosphorylate protein substrates in the cell. DAG activates another kinase, protein kinase C. The 5-HT$_2$ receptors are coupled via the G protein G$_q$ or G$_{11}$ to activation of PLC.

The 5-HT$_{2A}$ receptor is involved in smooth muscle contraction and platelet aggregation. In the brain, 5-HT$_{2A}$ receptors are found in cerebral cortex, claustrum, and basal ganglia. It is thought that hallucinogenic drugs exert their psychotrophic action by activating 5-HT$_{2A}$ receptors. The 5-HT$_{2B}$ receptor was first described from the stomach fundus and later was identified in the gut, heart, kidney, and lung. Its presence in the brain is less certain. The 5-HT$_{2C}$ receptor is found in choroid plexus, where it regulates cerebral spinal fluid production and ion exchange between the cerebral spinal fluid and brain. The 5-HT$_{2C}$ receptor is also found in various brain regions such as frontal cortex and amygdala. Activation of brain 5-HT$_{2C}$ receptors can lead to hypoactivity and hypophagia. Moreover, atypical antipsychotic drugs block the activation of 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors, indicating that these receptors may be involved in the pathophysiology of schizophrenia.

The Serotonin-3 Receptor is a Ligand-Gated Ion Channel

The 5-HT$_3$ receptor is different from the other 5-HT receptors in that it forms an ion channel that regulates the flux of ions. The structure of receptor is a pentamer, similar to the nicotinic acetylcholine receptor. The receptors are found on neurons in the hippocampus, nucleus tractus solitarius, and area postrema, as well as in the periphery. They are located on pre and postganglionic autonomic neurons and alter GI tract motility and intestinal secretion. When activated by 5-HT, the receptor triggers rapid depolarization due to an inward current by opening a nonselective cation channel (Na$^+$, Ca^{2+} influx and K$^+$ efflux). The receptor is possibly involved in nausea, vomiting, and irritable bowel syndrome.

Alternative Splice Variants of Serotonin Receptors

Functional diversity of proteins can be produced by alternative splicing events. Many of the 5-HT receptor genes contain introns that are subject to alternative splicing with the generation of multiple-receptor mRNAs encoding slightly different proteins, referred to as isoforms. Seven carboxy-terminal splice variants of the 5-HT$_4$ receptor have been described. The most interesting feature of these splice variants is the level of constitutive activity of the receptor, which is markedly increased. Constitutive activity is the ability of a receptor to activate second-messenger pathways spontaneously without the binding of an external ligand. Four carboxy-terminal splice variants of the 5-HT$_7$ receptor have been identified. The functional consequence of these variants is uncertain. An alternatively spliced variant of the 5-HT$_{2C}$ receptor has been described, which encodes a truncated, nonfunctional protein. More work needs to be done to determine the physiological relevance of these RNA splicing events; nonetheless, is it clear that this process leads to additional diversity in 5-HT receptor signaling.

RNA Editing Produces Multiple Functional 5-HT$_{2C}$ Receptor Isoforms

The 5-HT$_{2C}$ receptor undergoes a unique process termed RNA editing to yield multiple-receptor variants. RNA editing is an enzymatic reaction that alters nucleotide sequences of RNA transcripts. For the 5-HT$_{2C}$ receptor, five encoded adenosine residues are converted to inosines by double-stranded RNA adenosine deaminases. In the human 5-HT$_{2C}$ receptor, the adenosines within the predicted intracellular second loop can be converted to inosine at the RNA level resulting in multiple mRNA species with the potential to encode 24 different protein variants (Figure 2). The extensively edited isoforms have different abundances in brain tissue, and the translated proteins exhibit different binding properties, and differential activation of second-messenger systems. For example, 5-HT exhibits a decreased potency when activating the fully edited, VGV, isoform of the human receptor compared with the unedited, INI, form and there is a rightward shift in the dose-response curve for phosphoinositide hydrolysis. In addition, editing can alter the ability of 5-HT$_{2C}$ receptors to couple to multiple G-proteins. For example, the non-edited 5-HT$_{2C}$ receptor functionally couples to G$_q$ and G$_{13}$, whereas the edited 5-HT$_{2C}$ receptors has less coupling to G$_{13}$. RNA editing may have clinical

FIGURE 2 The positions of the editing sites within human 5-HT$_{2C}$ receptors RNA and amino acid sequences are shown for hINI, hVSV, and hVGV edited isoforms of the 5-HT$_{2C}$ receptor. These editing sites are located in the putative second intracellular loop.

significance; recent studies have indicated that alterations in the editing profile of the 5-HT$_{2C}$ receptor are associated with the incidence of suicide and schizophrenia.

Single Nucleotide Polymorphisms Occur in the Serotonin Receptor Family

Normal genetic variations (single-nucleotide polymorphism, SNP) have been identified in almost all 5-HT receptors. Polymorphisms in the coding region of the gene have the potential to alter the receptor's ability to bind ligand, to activate signal-transduction pathways, or to adapt to environmental influences. For example, a polymorphism in the amino terminus of the human 5-HT$_{1A}$ receptor attenuates the down-regulation and desensitization produced by the agonist 8-OH-DPAT. A polymorphic variant in the 5-HT$_{1B}$ receptor in the putative third transmembrane domain alters the binding of the antimigraine drug, sumatriptan. In addition, a polymorphism in the carboxy terminus of 5-HT$_{2A}$ receptor reduces the receptor's ability to mobilize internal Ca^{2+}. Currently, efforts are being made to link the occurrence of 5-HT receptor polymorphisms with various pathological disorders. Future progress in pharmacogenomics (using genetic information to predict drug response) may potentially lead to better design of serotonergic drugs to reduce side effects and target subpopulations of patients with specific therapies dependent on their genetic profile.

Promiscuous Coupling and Crosstalk between 5-HT Receptor Signal-Transduction Pathways

Promiscuous coupling is the ability of a receptor to couple to more than one signal transduction pathway. For example, the 5-HT$_{1A}$ receptor can both inhibit and activate AC. As mentioned above, the primary coupling of this receptor is G$_{\alpha i/o}$ with subsequent inhibition of AC; this has been demonstrated both *in vivo* and in cultured cell systems. However, this receptor has been shown to activate AC, mediated by $\beta\gamma$ subunits released from G$_{\alpha i/o}$, rather than G$_{\alpha s}$ protein. This activation seems to require high receptor occupancy and high expression levels in cell expression systems. In addition, depending on cell type and experimental conditions, the 5-HT$_{1A}$ receptor can activate or inhibit PLC. Many studies suggest that the G-protein $\beta\gamma$-subunits play a role in such crosstalk between signaling pathways. The 5-HT$_{2C}$ receptor (as well as the 5-HT$_{2A}$ receptor) is another promiscuous receptor that may activate multiple signal transduction pathways. The primary coupling of 5-HT$_{2C}$ receptors is to activate PLC activity via G$_q$/G$_{11}$ proteins to produce IP$_3$ and DAG. However, this receptor can activate other phospholipases, for example, phospholipase D through G$_{13}$ proteins and free $\beta\gamma$-subunits. Phospholipase D can influence many neuronal functions including endocytosis, exocytosis, vesicle trafficking, and cytoskeletal dynamics. In addition, 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors can activate phospholipase A-2 activity, which leads to the release of arachidonic acid and potential formation of bioactive eicosanoids such as the prostaglandins. There is evidence that 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors differentially activate PLC compared to PLA2, dependent on the specific agonists employed. Moreover, there may be differences in the desensitization of these two pathways after repeated agonist treatment. The release of arachidonic acid can result in the activation of K$^+$ channels, regulation of neurotransmitter release, and can be a retrograde messenger. It is becoming apparent that specific 5-HT receptors interact with many G proteins, thereby regulating multiple signal-transduction pathways. This diversity may be a target for future therapeutics and interventions.

Summary: Potential Role of Receptor Diversity

There has been much speculation on the origin and significance of the 5-HT receptor diversity. One possible explanation for the numerous subtypes and variants is that the 5-HT neurotransmitter system emerged early

during evolution. Thus, there has been ample time for genetic variation and divergence to occur in the genetic encoding for the receptor proteins. Whatever the process of receptor diversity, the multiple receptor signaling has considerable biological significance. For example, inputs to the dorsal raphe nucleus are integrated into global 5-HT release, via both synaptic and volume transmission, that in turn can modulate multiple and diverse neuronal functions based on receptor subtype. Thus, one transmitter (5-HT) with multiple receptors/multiple pathways can introduce levels of complexity, yet specific physiological responses, that are localized to a given brain region or neural system. These responses can be cell-type specific or even cell-compartment specific. Furthermore, a level of fidelity is introduced based on the receptor's relative affinity for 5-HT. The exact nature of this diversity is at present unclear however the physiological implications are quite apparent, when considering the diverse role of these receptors in many physiological functions and disease states.

SEE ALSO THE FOLLOWING ARTICLES

Adenylyl Cyclases • Neurotransmitter Transporters • Phospholipase C • RNA Editing

GLOSSARY

autoreceptor A receptor on the neuronal cell body or presynaptic terminal can regulate its own cell firing and/or neurotransmitter release and synthesis.

constitutive activity Ability of a receptor to activate second-messenger pathways without the binding of an external ligand.

heteroreceptor A presynaptic receptor that regulates the release of neurotransmitter other than its own natural ligand.

promiscuous coupling The ability of a receptor to couple to more than one signal cascade.

RNA editing A process whereby the nucleotide sequence of RNA transcripts is chemically altered.

FURTHER READING

Aghajanian, G. K., and Sanders-Bush, E. (2002). Serotonin. In *Psychopharmacology: The Fifth Generation of Progress* (K. L. Davis, D. Charney, J. T. Coyle and C. Nemeroff, eds.) Lippincott, Williams, and Wilkins, Philadelphia.

Barnes, N. M., and Sharp, T. (1999). A review of central 5-HT receptors and their function. *Neuropharmacology* **38**, 1083–1152.

Hoyer, D., Clarke, D. E., Fozard, J. R., Hartig, P. R., Martin, G. R., Mylecharane, E. J., Saxena, P. R., and Humphrey, P. P. A. (1994). VII. International Union of Pharmacology classification of receptors for 5-hydroxytrytamine (serotonin). *Pharmacol. Rev.* **46**, 157–203.

Meltzer, H. Y. (1999). The role of serotonin in antipsychotic drug action. *Neuropsychopharmacology* **21**, 106S–115S.

Raymond, J. R., Mukhin, Y. V., Gelasco, A., Turner, J., Collinsworth, G., Gettys, T. W., Grewal, J. S., and Garnovskaya, M. N. (2001). Multiplicity of mechanisms of serotonin receptor signal transduction. *Pharmacol. Ther.* **92**, 179–212.

BIOGRAPHY

Elaine Sanders-Bush is a Professor of Pharmacology and Psychiatry at Vanderbilt University School of Medicine, Nashville, Tennessee. She earned a Ph.D. in pharmacology at Vanderbilt in 1967. Her research focuses on serotonin receptors, applying a multidisciplinary approach to define the role of signal transduction molecules and posttranscriptional and posttranslational modifications that alter receptor function. Dr. Sanders-Bush received a Merit Award from the National Institute of Mental Health in recognition of her research accomplishments.

Paul Gresch is a Postdoctoral Fellow in Dr. Sanders-Bush's laboratory. He earned a Ph.D. in cellular and clinical neurobiology from Wayne State University in Detroit, Michigan in 1999.

Siglecs

Ajit Varki
University of California, San Diego, Calfornia, USA

Sialic acid recognizing Ig-superfamily Lectins (Siglecs) are a major subset of the "I-type lectins." The latter are defined as animal proteins other than antibodies that can mediate carbohydrate (glycan) recognition via immunoglobulin(Ig)-like domains. Siglecs share characteristic amino-terminal structural features that are involved in their sialic acid-binding properties, and can be broadly divided into two groups: an evolutionarily conserved subgroup (Siglecs-1, -2, and -4) and a CD33/Siglec 3-related subgroup (Siglecs -3 and -5 to -11). While the precise functions of Siglecs are unknown, they seem to send inhibitory signals to the cells that express them, in response to recognition events on cell surfaces.

Historical Background and Definition

Sialic acids (Sias) are a family of nine-carbon acidic sugars that typically occupy a terminal position on glycan chains attached to the cell surface of "higher" animals. The immunoglobulin superfamily (IgSf) is an evolutionarily ancient group of proteins whose appearance predated the emergence of the immunoglobulins themselves. Until the 1990s, it was assumed that IgSf members (other than some antibodies) did not mediate carbohydrate recognition. Independent work on CD22 (eventually Siglec-2, a protein on mature resting B cells) and on sialoadhesin (Sn, eventually Siglec-1, a protein on certain macrophage subsets) revealed that their first Ig V-set-like domains could mediate Sia recognition. Homologous features of this V-set Ig-like domain and the adjacent C2-set domain then led to the discovery that two other previously cloned molecules—CD33 (eventually Siglec-3) and Myelin-associated Glycoprotein (MAG, eventually Siglec-4)—also had Sia-binding properties. Following consultation among all researchers working on these proteins, the common name "Siglec" and a numbering system were agreed upon. Criteria for inclusion of IgSf-related proteins as Siglecs are: (1) the ability to recognize sialylated glycans; and (2) significant sequence similarity within the N-terminal V-set and adjoining C2-set domains. Evaluation of the human and mouse genomes eventually defined 11 human and 8 mouse molecules that fulfill these criteria. Since humans have more Siglecs than mice and cloning of the mouse molecules initially lagged behind, the primary numbering system is based on the human molecules.

Two Broad Subgroups of Siglecs

While Siglecs -1, -2, and -4 appear to be evolutionarily rather conserved, the CD33/Siglec-3-related subgroup (Human Siglecs -3 and -5 to -11) appear to be rapidly evolving. Some CD33/Siglec-3-related Siglecs appear to have evolved as hybrids of pre-existing genes and/or by gene conversion. For these reasons, sequence comparisons alone do not allow the conclusive designation of the orthologue status of all mouse genes, and additional features such as gene position and exon structure must be taken into account. Until such issues are resolved, some mouse Siglecs have been assigned a temporary alphabetical designation.

Common Structural Features

All are single-pass Type 1 integral membrane proteins with extra-cellular domains consisting of uniquely similar N-terminal V-set Ig domains, followed by variable numbers of C2-set Ig domains, ranging from 16 in Sn/Siglec-1 to 1 in CD33/Siglec-3. Crystal structures for mouse Siglec-1 and human Siglec-7 indicate that the V-set immunoglobulin-like fold has several unusual features, including an intra-beta sheet disulphide and a splitting of the standard beta strand G into two shorter strands. These features along with certain key amino acid residues appear to be requirements for Sia recognition. In particular, a conserved arginine residue is involved in a salt bridge with the carboxylate of Sia in all instances studied to date.

Cell-Type Specific Expression

With the exception of MAG/Siglec-4 and Siglec-6, expression appears to be confined to the hematopoeitic

and immune systems. Within these systems each Siglec is expressed in a cell-type specific fashion, suggesting that each may be involved in discrete functions. However, systematic studies of Siglec expression outside the hematopoeitic system and during development have not yet been done.

Genomic Organization and Phylogeny

Based on probing for the canonical functional amino acids in the V-set domain of the typical Siglec, there is no evidence for Siglec-like molecules in prokaryotes, fungi or plants, nor in animals of the Protostome lineage, including organisms for which the complete genome is available. In contrast, it is relatively easy to find Siglec-like V-set domains in many vertebrate taxa (Sia recognition by fish and reptile Siglec candidates has not been formally shown as yet). While the relatively conserved Siglecs (-1, -2, and -4) have clear-cut single orthologues that are easy to identify in various species, the remaining "CD33/Siglec-3 related" Siglecs appear to have been evolving rapidly. Most of the latter genes are clustered together in a ~ 500 kb region on human chromosome 19q13.3–13.4.

Siglec Recognition of Sialic Acids and Their Linkages

The first two Siglecs discovered (Sn/Siglec-1 and CD22/Siglec-2) had strikingly different binding properties for sialosides—with Sn preferring alpha2–3 linked targets and CD22 being highly specific for alpha2–6 linkages. In the latter case, the binding affinity was in the low micromolar range. MAG/Siglec-4 also has an extended binding site that is even highly specific for the underlying sugar chain. There is also variable preference for certain types of sialic acids, with Sn and MAG not tolerating the common N-glycolyl modification of Sias. However the CD33/Siglec-3-related Siglecs are more promiscuous in their preferences for different types and linkages of Sias. Of course, many of the less common linkages and types of sialic acids have not been studied for Siglec recognition. The Golgi enzymes that are potential regulators of Siglec functions are primarily the sialyltransferases, and to some extent the enzymes which modify sialic acids. Some Siglecs show preferences for certain macromolecular ligands e.g., CD45 for CD22/Siglec-2, the mucins CD43, and Muc-1 for Sn/Siglec-1, and certain brain glycolipids for MAG/Siglec-4.

Potential Effects of Neu5Gc Loss on Human Siglec Biology

The most common Sias of mammalian cells are N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc). Humans are an exception, because of a mutation in CMP-sialic acid hydroxylase, which occurred after the time ($\sim 5-7$ Ma) when we shared a common ancestor with great apes. The resulting loss of Neu5Gc and increase in Neu5Ac in humans could have potentially altered the biology of the Siglecs. For example, human cells have a higher density of Sn/Siglec-1 ligands than great apes, the distribution of Sn-positive macrophages in humans is different, and a much larger fraction of human macrophages is positive. Other emerging evidence suggests that there are further human-specific changes in Siglec biology that may be related to the loss of Neu5Gc.

Masking and Unmasking of Siglecs Binding Sites on Cell Surfaces

The initial assumption was that Siglecs were involved in intercellular adhesion. However, in most instances, their binding sites appear to be masked by Sias on the same cell surfaces on which they are expressed. Of course, external ligands with very high affinity/avidity may still compete for the endogenous masking ligands. There is also some evidence that unmasking can occur under certain conditions, but it is not known if this is biologically relevant. Overall, the significance of Siglec masking is unclear at this time.

Signaling Motifs in Cytosolic Tails

The CD33-related Siglecs have conserved tyrosine residues in the cytosolic tails, one of which corresponds to a canonical immunoreceptor tyrosine-based inhibition motif (ITIM). Various *in vitro* manipulations of these receptors indicate that these tyrosines are indeed targets for phosphorylation, and that they can modulate signaling events by recruiting certain tyrosine phosphatases. However, the true *in vivo* biological functions of these signaling motifs remain obscure. Another major unresolved question is: what is the connection between extra-cellular sialic acid recognition and signaling via the cytosolic motifs?

Known and Putative Functions of the Siglecs

Various lines of evidence indicate that MAG/Siglec-4 is involved in the maintenance of myelin organization

and in the inhibition of neurite outgrowth during regeneration after injury. It is also reasonably clear that CD22/Siglec-2 functions as an inhibitory component of the antigen receptor complex of B Cells, and is thus involved in regulating the humoral immune response. While Sn/Siglec-1 appears to mediate various macrophage adhesion events *in vitro* and *in vivo*, it is as yet unclear what the functions of these interactions are. Little is known about the functions of CD33-related Siglecs. It has been suggested that these molecules are involved in innate immunity. One hypothesis currently being tested is that Siglecs may be sensors for pathogens that have sialylated cell surfaces and/or express extra cellular sialidases.

SEE ALSO THE FOLLOWING ARTICLES

Immunoglobulin (Fc) Receptors • Lectins • Polysialic Acid in Molecular Medicine

GLOSSARY

immunoglobulin superfamily (IgSf) Proteins that have modules homologous to those of antibodies (immunoglobulins). This is an evolutionarily ancient group of proteins whose appearance actually predated the emergence of the immunoglobulins themselves.

I-type lectins Proteins (other than antibodies) in which immunoglobulin-like modules mediate binding to glycans (sugar chains).

sialic acids These acids are a diverse family of nine-carbon acidic sugars that typically occupy a terminal position on glycan chains attached to the cell surface of "higher" animals of the deuterostome lineage.

siglecs A major subset of the I-type lectins. Name is based on their defining properties, as sialic acid recognizing IgSf lectins.

FURTHER READING

Angata, T., and Brinkman-Van der Linden, E. (2002). I-type lectins. *Biochim. Biophys. Acta* **1572**, 294.

Angata, T., and Varki, A. (2002). Chemical diversity in the sialic acids and related alpha-keto acids: An evolutionary perspective. *Chem. Rev.* **102**, 439–470.

Crocker, P. R., and Varki, A. (2001). Siglecs, sialic acids and innate immunity. *Trends Immunol.* **22**, 337–342.

Crocker, P. R., Mucklow, S., Bouckson, V., McWilliam, A., Willis, A. C., Gordon, S., Milon, G., Kelm, S., and Bradfield, P. (1994). Sialoadhesin, a macrophage sialic acid binding receptor for haemopoietic cells with 17 immunoglobulin-like domains. *EMBO J.* **13**, 4490–4503.

Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le, D. N., Powell, L., Roder, J., Schnaar, R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., Van den Berg, T. K., Van der Merwe, P. A., Watt, M., and Varki, A. (1998). Siglecs: A family of sialic-acid binding lectins [letter]. *Glycobiology* **8**(v2).

Kelm, S., Pelz, A., Schauer, R., Filbin, M. T., Tang, S., De, B. M.-E., Schnaar, R. L., Mahoney, J. A., Hartnell, A., Bradfield, P., and Crocker, P. R. (1994). Sialoadhesin, myelin-associated glycoprotein and CD22 define a new family of sialic acid-dependent adhesion molecules of the immunoglobulin superfamily. *Curr. Biol.* **4**, 965–972.

Powell, L. D., and Varki, A. (1995). I-type lectins. *J. Biol. Chem.* **270**, 14243–14246.

Powell, L. D., Sgroi, D., Sjoberg, E. R., Stamenkovic, I., and Varki, A. (1993). Natural ligands of the B cell adhesion molecule CD22beta carry N-linked oligosaccharides with alpha-2, 6-linked sialic acids that are required for recognition. *J. Biol. Chem.* **268**, 7019–7027.

BIOGRAPHY

Ajit Varki is Professor of Medicine and Cellular and Molecular Medicine, Director of the Glycobiology Research and Training Center, and Coordinator of the project for Explaining the Origin of Humans, at the University of California, San Diego. Dr. Varki's laboratory explores the biology and evolution of sialic acids in health and disease.

Sigma Factors

John D. Helmann
Cornell University, Ithaca, New York, USA

The sigma(σ)-subunit of bacterial RNA polymerase is essential for the process of promoter selection. Interactions between σ and promoter region DNA allow RNAP to recognize and physically open the start sites for transcription initiation. The majority of transcription under most conditions requires a single, essential σ-factor (primary σ). However, in most bacteria there are a number of alternative σ-factors that can activate specific sets of target genes, as defined by their distinct promoter sequences, in response to changing environmental conditions. The activation of groups of coregulated genes (regulons) by alternative σ-factors complements the more familiar gene-regulation mechanisms that rely upon transcription repressors and activators.

Introduction

DNA serves as a storehouse for the genetic information of each organism, but is itself a silent repository. Only upon copying of DNA segments into RNA (transcription), and translation into protein, is the genetic potential of each organism (and each cell type in multicellular organisms) able to be realized. For example, our skin, blood, hair, liver, and lung cells all contain the same DNA, yet they have differentiated into functionally distinct cell types based largely on the selective expression of a small subset of the genes contained within each cell. Similarly, comparative genomics has led to the realization that the evolutionary processes that have led to the divergence of humans and mice, for example, have more to do with changes in "how" genes are regulated than with "which" genes are actually present in each organism. These examples highlight the importance of gene regulation: those mechanisms that selectively activate and repress different genes depending on cell type and local environment.

Gene Regulation in Bacteria

Most of the principles of gene regulation were first developed from studies of bacterial cells and, in particular, *Escherichia coli*. The first step in gene expression is the transcription of the genetic code from DNA into an RNA copy that usually serves to direct protein synthesis (translation). RNA polymerase (RNAP) is the enzyme that copies (transcribes) DNA into RNA.

In bacterial cells, most gene regulation occurs at the level of transcription. Genes that are expressed are actively transcribed and translated into their corresponding protein products. Conversely, genes that are silent are not copied into RNA. Thus, factors that act to control the activity of RNAP are central to many processes of gene regulation. These regulators include activator proteins, repressor proteins, and σ-factors. In order to appreciate how these factors can control gene expression, the biochemical properties of RNAP and the overall features of the transcription cycle are discussed first, followed by σ-factors.

Bacterial RNA Polymerase and the Transcription Cycle

In bacteria, RNAP is a multisubunit protein complex and contains both a catalytic core enzyme and a specificity subunit known as σ. The core enzyme (containing the β, β', two α, and the ω polypeptide chains) is able to catalyze the synthesis of RNA directed by a DNA template. Core enzyme associates reversibly with another polypeptide chain, σ, to form the holoenzyme.

RNAP copies particular DNA segments into RNA chains by first initiating RNA synthesis at a promoter site, elongating the RNA chain, and finally terminating transcription (the transcription cycle). In bacteria, many protein-coding genes are organized into cotranscribed groups called operons. In these cases, each mRNA product can direct the synthesis of multiple protein products.

The first step in the transcription cycle is promoter site localization and requires the presence of the σ-subunit: only holoenzyme can recognize promoter sites. Promoter sites are typically characterized by defined DNA sequences located just upstream of the site where RNA synthesis begins (designated the +1 site). For example, sequences centered ~ 10 bp upstream often have a sequence, designated the "-10 region", similar to 5′TATAAT3′. A different conserved sequence

(TTGACA) is usually located between 16 and 18 bp further upstream and is designated the "−35 region." Together, the −35 and −10 sequences provide the necessary information to allow RNAP to identify a promoter. Specific interactions between the σ-subunit of holoenzyme and the −35 and −10 elements allow RNAP to tightly engage the promoter region. Once bound, RNAP interacts with an extended region of DNA, ultimately contacting the DNA over a region of 60 or more base pairs (from −40 to +20).

Once bound at a promoter site RNAP melts the DNA over approximately one turn of the helix (from about −10 to +2) to expose the template strand of the DNA. The process of converting the double-stranded promoter region DNA into a locally melted structure (called the open complex) requires the action of σ-factor to initiate the melting process and is then driven by the association of the template strand of the DNA with the RNAP active site (comprised largely of the β- and β'-subunits). The open complex is then poised to initiate synthesis of an RNA chain. The region of DNA strand-separation is often referred to as the transcription bubble, to distinguish it from the otherwise double-stranded DNA in the surrounding regions of the genome.

The process of transcription elongation involves the assembly of ribonucleoside triphosphate precursors into an RNA chain containing the sequence complementary to the template DNA. During transcription elongation the DNA template is threaded through the active site of RNAP and, as a result, the transcription bubble is propagated down the DNA with melting of the DNA occurring as the DNA enters the active site and reannealing of the template occurring as the DNA exits the RNAP. The complex containing the DNA template, RNAP, and the growing RNA chain is often referred to as the ternary complex. As the growing RNA chain emerges from RNAP, it is thought to displace the σ-subunit, and the process of transcription elongation is continued by the core RNAP subunits. Elongating RNAP sometimes associates with additional protein factors known as elongation factors which can modulate its properties. In general, elongation is highly processive: the RNAP that begins an RNA chain continues synthesis until the chain is terminated.

The process of transcription termination, like that of initiation, is critical for the controlled expression of genetic information. At the end of many transcription units the DNA sequence encodes a G:C-rich stem loop structure in the RNA chain followed by a short region rich in uridine. This structure functions as a transcription terminator and interacts with the core RNAP leading to dissociation of the ternary complex and release of the terminated RNA chain. In other cases, protein factors such as the rho protein can bind to unstructured regions of RNA and trigger transcription termination. Once released by the termination reaction, the core RNAP must rebind to a σ-subunit to reform holoenzyme and thereby complete the transcription cycle.

Since bacteria lack a nuclear membrane, the processes of transcription and translation can be coupled: ribosomes can bind to the mRNA chain as it emerges from RNAP and immediately initiate protein synthesis. Indeed, along actively transcribed genes several RNAP molecules may be simultaneously engaged in RNA chain elongation with each RNA chain in turn bound by one or more ribosomes engaged in protein synthesis.

Biochemical Properties

As expected for a critical component of the transcription apparatus, σ-factors have been highly conserved through evolution. Indeed, σ-factors from one bacterium can function, at least in some cases, with the core RNAP from distantly related species. In general, all σ-factors share certain defining properties. They all bind to the core RNAP, apparently at a common binding site, to form a holoenzyme. The presence of σ-factor determines the sequence of the promoters that can be bound by the corresponding holoenzyme. The σ-subunit, in those cases that have been studied, also plays a role in the initiation of the DNA-melting step that is an obligatory prelude to transcription initiation. Often, although perhaps not always, σ is released from the ternary complex during the process of RNA chain elongation.

Bacterial σ-factors can be divided, based in part on their protein sequences, into two families: the σ^{70} and the σ^{54} families. All bacteria contain at least one member of the σ^{70} family that is required for the transcription of those genes essential for growth under virtually all conditions (so-called "housekeeping" functions). This is referred to as the primary (or class 1) σ-factor. In addition, most bacteria contain at least one (and as many as 50 or more) alternative σ-factors (Table I). These factors typically control genes needed for specialized functions that may only be expressed under particular growth conditions. Examples include genes needed for stress responses, motility, sporulation, uptake or transport of specific nutrients, or antibiotic production.

Most alternative σ-factors are structurally related to the primary (class 1) σ-factors and are therefore considered to be members of the σ^{70} family. Some alternative σ-factors (class 2) are very similar to the primary σ-factors but are, however, dispensable for growth. One example is the E. coli σ^S (RpoS) protein that becomes active in stationary phase cells. A much larger class of alternative σ-factors are those more distantly related in sequence to the primary σ-factor, and often lacking one or more the conserved regions characteristic of the class 1 and 2 proteins. These alternative σ-factors (classes 3, 4, and 5) control a wide-range of physiological processes and are often only

TABLE I

σ Factors in *E. coli* and *B. subtilis*

Organism	σ	Gene	Function
E. coli	σ^{70}	*rpoD*	Housekeeping genes
	σ^{H}	*rpoH*	Heat shock
	σ^{E}	*rpoE*	Extreme heat shock, periplasmic stress (ECF[a])
	σ^{F}	*fliA*	Flagellar-based motility
	σ^{S}	*rpoS*	Stationary phase adaptations
	σ^{N}	*rpoN,* *glnF*	Nitrogen-regulated genes (σ^{54} family)
	σ^{fecI}	*fecI*	Ferric citrate uptake (ECF)
B. subtilis	σ^{A}	*sigA*	Housekeeping genes
	σ^{B}	*sigB*	General stress response
	σ^{D}	*sigD*	Flagellar-based motility, autolysins
	σ^{E}	*sigE*	Sporulation, early mother cell
	σ^{F}	*sigF*	Sporulation, early forespore
	σ^{G}	*sigG*	Sporulation, late forespore
	σ^{H}	*sigH*	Competence and early sporulation
	σ^{K}	*sigK*	Sporulation, late mother cell
	σ^{L}	*sigL*	Degradative enzymes (σ^{54} family)
	σ^{I}	*ykoZ*	Unknown
	σ^{Xpf}	*xpf*	PBSX regulation
	σ^{X}	*sigX*	Cell envelope modifications (ECF)
	σ^{W}	*sigW*	Antibiotic stress responses (ECF)
	σ^{M}	*sigM*	Cell wall stresses (ECF)
	σ^{V}	*sigV*	Unknown (ECF)
	σ^{Y}	*sigY*	Unknown (ECF)
	σ^{Z}	*sigZ*	Unknown (ECF)
	σ^{ylaC}	*ylaC*	Unknown (ECF)

[a] ECF indicates a factor that is a member of the large extra-cytoplasmic function family of alternative σ factors.

active under very specific growth conditions. Prominent among these alternative σ-factors are the numerically abundant extracytoplasmic function (ECF) σ-factors (class 4 proteins) that appear, in general, to control functions related to the cell surface. Finally, many bacteria have alternative σ-factors with a distinctly different protein sequence and define the σ^{54} family. Unlike members of the σ^{70} family, the σ^{54}-factors form holoenzymes that almost always require ATP-dependent activator proteins to activate their target promoters. The prototype of the σ^{54}-family is the *E. coli* σ^{54} (RpoN) protein that is important for nitrogen regulation of gene expression (Table I). Similar proteins in other bacteria control a variety of other functions.

The division of σ^{70}-family members into different structural classes is based in large part on the comparison of their protein sequences. All primary (class 1) σ-factors (and the closely related class 2 factors) have four conserved regions. Region 1 contains an autoinhibitory domain that prevents the free σ-factor from interacting with DNA. This inhibition is relieved when σ associates with core RNAP. Region 2 contains the key determinants for recognition of the −10 promoter element and for initiation of DNA-melting. Region 3 functions as a spacer domain between regions 2 and 4. Region 4 contains the elements that recognize the −35 promoter element and also serves as a site of contact for some transcription activator proteins. Recent advances in the determination of three-dimensional structures for both σ-factors and holoenzyme have made it clear that these protein sequence regions correspond to folded domains in the σ-structure. The two domains that are required for recognition of promoters, domains 2 and 4, are common to all σ^{70}-family members. In contrast, much of domains 1 and 3 is often deleted in alternative σ-factors.

Roles of σ in Gene Regulation

Just as DNA-binding transcriptional repressors and activators can control the activity of RNAP at promoter sites, alternative σ-factors provide a powerful avenue for the activation of specific sets of genes (the σ-factor regulon) during cell differentiation or in response to various stress conditions. The activity of alternative σ-factors can, in turn, be controlled in numerous ways. In some cases, the σ-factor is only synthesized under very specific conditions. In other cases, the σ-factor is synthesized but is maintained in an inactive state by binding to an inhibitory protein (called an anti-σ-factor) or it is synthesized as an inactive precursor protein that only becomes active when an inhibitory peptide is cleaved by proteolysis. Anti-σ-factors are, in many cases, associated with the cell membrane and thereby able to perceive signals from the external environment. In any event, once the active σ-factor is synthesized or released it can compete with other σ-factors for binding to core RNAP and thereby redirect some of the RNAP to its cognate target promoters.

Roles for alternative σ-factors in bacterial gene regulation are legion. For example, alternative σ-factors play key roles in the expression of heat shock genes, iron uptake systems, the developmental program of sporulation (in gram-positive bacteria), synthesis of flagella and chemotaxis machinery, and various stress responses. A summary of the roles of σ-factors in the two best-studied bacterial systems is presented in Table I.

Phylogenetic Distribution

σ-factors are always found associated with bacterial RNAP but have been functionally replaced by other proteins in the more complex RNAP characteristic of eukaryotes and Archaea. However, σ-like factors

are found in the chloroplasts of photosynthetic eukaryotes which are descended from a cyanobacterium. In contrast to chloroplasts, the transcription apparatus in present day mitochondria is most similar to certain bacteriophage-encoded, single-subunit RNAPs and is unlike the multisubunit RNAP of either bacteria or eukarya.

SEE ALSO THE FOLLOWING ARTICLES

DNA Polymerase I, Bacterial • DNA Polymerase II, Bacterial • DNA Polymerase III, Bacterial • RNA Polymerase Reaction in Bacteria • RNA Polymerase Structure, Bacterial • T7 RNA Polymerase

GLOSSARY

alternative σ A σ-factor that can replace the primary σ-factor in response to a particular stress or signal to activate the expression of a defined target regulon.

promoter DNA sequences necessary for the accurate initiation of transcription by RNA polymerase.

regulon The set of genes and operons that are regulated by a common factor.

RNA polymerase Enzyme that copies DNA sequences into the complementary RNA sequence.

FURTHER READING

Borukhov, S., and Nudler, E. (2003). RNA polymerase holoenzyme: Structure, function and biological implications. *Curr. Opin. Microbiol.* **6**, 93–100.

Borukhov, S., and Severinov, K. (2002). Role of the RNA polymerase sigma subunit in transcription initiation. *Res. Microbiol.* **153**, 557–562.

Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y., and Gralla, J. D. (2000). The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J. Bacteriol.* **182**, 4129–4136.

Gruber, T. M., and Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Ann. Rev. Microbiol.* **57**, 441–466.

Helmann, J. D. (2002). The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.* **46**, 47–110.

Paget, M. S., and Helmann, J. D. (2003). The sigma70 family of sigma factors. *Genome Biol.* **4**, 203.

Wosten, M. M. (1998). Eubacterial sigma-factors. *FEMS Microbiol. Rev.* **22**, 127–150.

BIOGRAPHY

John D. Helmann is a Professor in the Department of Microbiology at Cornell University in Ithaca, New York. His principal research interest is the mechanisms of gene regulation in the gram-positive bacterium *Bacillus subtilis*. Major research directions include the characterization of alternative σ-subunits and regulators that respond to metal ions and oxidative stress. He received a Ph.D. in biochemistry from the University of California, Berkeley and did postdoctoral research at Harvard Medical School.

Sliding Clamps in DNA Replication: *E. coli* β-Clamp and PCNA Structure

Eric R. Goedken and John Kuriyan
University of California, Berkeley, California, USA

In the absence of additional protein factors, the DNA polymerases responsible for replicating entire genomes have a strong tendency to release from the DNA template after synthesizing short segments of DNA. Such enzymes are described as having low "processivity," meaning that polymerase adds only a small number of nucleotides to the growing chain before coming off the DNA template. Rapid replication of the millions of base pairs that constitute the genetic material of cells requires DNA polymerases with extremely high processivity. This is achieved by the utilization of proteins known as "sliding DNA clamps" that convert the poorly processive core polymerase enzymes (capable of synthesizing only tens of nucleotides without dissociation) to highly processive enzymes (capable of synthesizing several thousand nucleotides without dissociation).

Sliding clamps are dimeric or trimeric proteins that form a circular collar through which DNA is threaded. The polymerase enzyme is loosely tethered to the sliding clamp, thereby maintaining proximity to the DNA even when it does let go of the template. The detailed molecular architecture of these clamps is highly conserved from bacteria to eukaryotes, reflecting the fundamentally important role played by these proteins in DNA replication in all organisms.

Clamp Structure and Function

In the well-studied *Escherichia coli* system, the processivity clamp is called the β-subunit or β-clamp, and the polymerase responsible for replicating the genome is DNA polymerase III. The β-clamp is placed at primer/template junctions in an ATP-dependent manner by a complex of protein subunits known as the clamp loader (γ/τ) complex. In eukaryotes, the homologous sliding clamp is called proliferating cell nuclear antigen (PCNA), and the polymerase chiefly responsible for DNA replication is polymerase δ and its clamp loader is the replication factor C(RFC) complex.

Interestingly, clamps from bacteria and eukaryotes differ in the number of subunits that make up the clamp. Prokaryotic clamps such as β, are dimers, whereas eukaryotic PCNA is a trimer. Despite this difference in subunit stoichiometry, the crystal structures of the clamps from both *E. coli* and the yeast PCNA (as well as a number of others) reveal a very similar overall architecture (Figure 1). Each monomer in β consists of three similar domains containing two helices and two four-stranded β sheets, whereas in PCNA there are two of these conserved domains per monomer. Each assembled clamp complex then contains six of these domains despite the difference in subunit number. The conservation of this overall structure of the clamps throughout evolution is demonstrated by the presence of similarly shaped sliding clamps in archaebacteria and in certain bacteriophages such as T4 and RB69.

Both PCNA and β are highly negatively charged proteins. The distribution of charge has a marked asymmetry with positive charge concentrated around the interior hole through which negatively charged DNA can pass (Figure 1). In both proteins the α-helices of DNA, that line this central hole are oriented roughly perpendicular to the direction of the phosphate backbone of DNA as modeled inside the clamp. The orientation of the helices as well as their separation from DNA suggests that the electrostatic interaction between the protein and DNA is non-specific, thereby facilitating the sliding of the protein ring along the nucleic acid strands.

Biochemical results have indicated that the clamp is not broken into separate monomers prior to loading onto DNA and that only one of the two interfaces of the *E. coli* β-clamp dimer is opened before loading. Molecular dynamics simulations of β suggest that each monomer has an energetic preference for a state that is less "curved" than required to close both interfaces in the circular dimeric ring. This may provide a "spring-loaded" tension that makes it easier for the clamp loader to open the clamp.
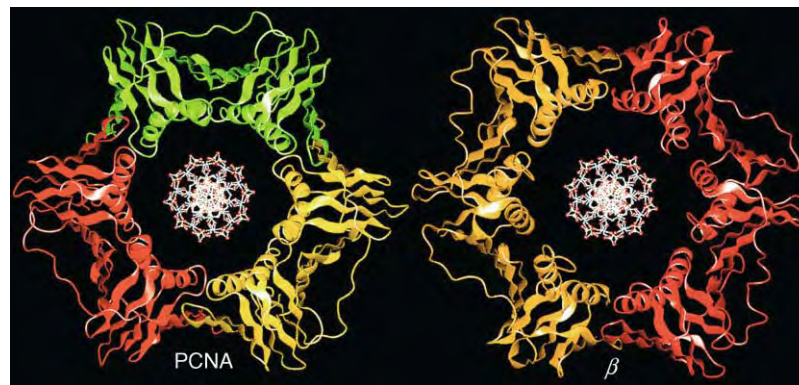
FIGURE 1   Ribbon diagrams of the structures of processivity clamps. DNA is shown as modeled in the interior of each clamp. Left, yeast PCNA; right, *E. coli* β-clamp.

Interaction with Other Proteins

Extended and flexible C-terminal segments of DNA polymerases are responsible for their interaction with the sliding clamps. A crystal structure of a trimeric bacteriophage clamp bounds to a peptide corresponding to the C-terminal region of the cognate DNA polymerase has been determined, showing the region of the clamp that interacts with the polymerase. This binding region turns out to be similar to the one seen earlier for the interaction between an inhibitory peptide (a segment of the cell-cycle protein p21/waf1-cip1) bound to PCNA. The peptide binds in a hydrophobic interface between the two domains that make up the clamp subunit. A recent crystal structure has revealed that the interaction between the clamp loader and the clamp has elements that resemble the polymerase-clamp interaction. A specific helix in the N-terminal domain of the δ-subunit of the *E. coli* clamp loader complex plugs into hydrophobic region of the clamp between the second and third domain of β. This suggests that the site of interaction for both the clamp loader and polymerase is located on the same region of the clamp and is likely conserved throughout the diversity of sliding clamps utilized in nature.

Clamp Loader Structure

The clamp loader in *E. coli* (γ complex) consists of five polypeptide chains: three γ-subunits, a δ-subunit and a δ′-subunit. The δ-subunit mentioned above is the "wrench" that opens the β-clamp. The wrench subunit is controlled by the coordinated actions of three γ "motor" subunits that hydrolyze ATP and an inert, δ′ "stator" subunit. A similar pentameric complex having sequence homology exists in eukaryotes (the RFC complex), and this loads the eukaryotic clamp (PCNA) onto DNA.

Structural studies have shown that each of the subunits of the *E. coli* clamp loader has a similar three-domain fold, with the first two domains being structurally related to the subunits of AAA+ATPases. Similar results have been obtained for the small subunit of the archaebacterial clamp loader complex. A flexible linker region tethers these nucleotide-binding domains to a helical C-terminal domain that clusters the five subunits into a pentameric assembly (Figure 2). ATP binding is required for the activation of the clamp loader, and nucleotide is potentially bound in each of the three γ-subunits. Once ATP is bound, the γ complex can bind to β and open one of its dimeric interfaces.
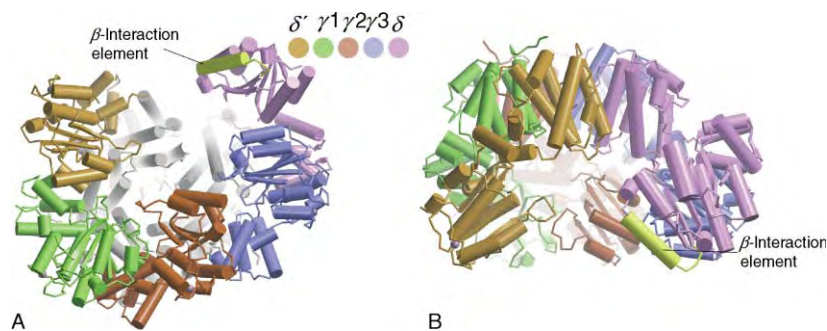
FIGURE 2   Two ribbon diagram of the structure of the *E. coli* clamp loader.

Upon interaction with DNA, the γ complex hydrolyzes these bound nucleotides and releases the β-clamp around the nucleic acid, where it is ready to tether the polymerase.

The mechanism by which ATP binding activates the clamp loader for binding to the clamp remains poorly understood, as is mechanism for coupling ATP hydrolysis to release the clamp around DNA at the proper sites.

SEE ALSO THE FOLLOWING ARTICLES

DNA Polymerase δ, Eukaryotic • Processivity Clamps in DNA Replication: Clamp Loading

GLOSSARY

clamp loader A multisubunit protein assembly that places sliding clamps around DNA at primer-template junctions in an ATP-dependent manner; examples include the γ complex in *E. coli* and the RFC complex in eukaryotes.

processivity The property of an enzyme that allows it to catalyze the addition of repeated rounds of reaction without releasing a substrate.

sliding clamp A circular multisubunit protein (either dimeric or trimeric) that slides along DNA to increase the processivity of DNA polymerase.

FURTHER READING

Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* **87**, 297–306.

Huang, C. C., Hearst, J. E., and Alberts, B. M. (1981). Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. *J. Biol. Chem.* **256**, 4087–4094.

Jeruzalmi, D., O'Donnell, M., and Kuriyan, J. (2001). Crystal structure of the processivity clamp loader γ complex of *E. coli* DNA polymerase III. *Cell* **106**, 429–441.

Jeruzalmi, D., Yurieva, O., Zhao, Y., Young, M., Stewart, J., Hingorani, M., O'Donnell, M., and Kuriyan, J. (2001). Mechanism of processivity clamp opening by the δ subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell* **106**, 417–428.

Kelman, Z., and O'Donnell, M. (1995). DNA polymerase III holoenzyme: Structure and function of a chromosomal replicating machine. *Annu. Rev. Biochem.* **64**, 171–200.

Kong, X. P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992). Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* **69**, 425–437.

Krishna, T. S. R., Kong, X.-P., Gary, S., Burgers, P., and Kuriyan, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **9**, 1233–1243.

Oyama, T., Ishino, Y., Cann, I. K., Ishino, S., and Morikawa, K. (2001). Atomic structure of the clamp loader small subunit from *Pyrococcus furiosus*. *Mol. Cell.* **8**, 8455–8463.

Shamoo, Y., and Steitz, T. A. (1999). Building a replisome from interacting pieces: Sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* **99**, 155–166.

BIOGRAPHY

John Kuriyan is Chancellor's Professor in the Department of Molecular and Cell Biology and Department of Chemistry at the University of California at Berkeley and is an investigator in the Howard Hughes Medical Institute. His principal fields of interest are DNA replication and signal transduction. These areas are under study in his research group through structural techniques like X-ray crystallography, but also through molecular dynamics simulations and biophysics. He holds a Ph.D. from the Massachusetts Institute of Technology and received postdoctoral training at Harvard University.

Eric R. Goedken is an American Cancer Society postdoctoral fellow at the University of California, Berkeley, where he studies the mechanism of DNA polymerase clamp loaders.

Small GTPases

Adam Shutes and Channing J. Der
University of North Carolina, Chapel Hill, North Carolina, USA

Small GTPases (20–25 kDa) are a class of guanine nucleotide-binding proteins involved in the regulation of numerous signaling pathways and cellular processes in eukaryotic cells. The small GTPases are monomeric, distinguishing them from the heterotrimeric GTPases which are comprised of three separate subunits, α, β, and γ. Similar to the G α-subunits, the small GTPases bind GDP and GTP with high affinity and catalyze the hydrolysis of the covalent bond between the two terminal phosphate groups of bound GTP producing protein-bound GDP and a solvent-free phosphate group. The cleavage of this bond releases energy that changes the conformation of small GTPase from a GTP-bound one to an inactive GDP-bound conformation. An intrinsic GDP/GTP exchange activity allows the formation of the active GTP-bound conformation, which then interacts with one or more effector targets to promote a specific cellular response. This two-state structural change is fundamental for small GTPase biological function as a regulated binary switch.

Identification and Classification of Small GTPases

The first and prototypical small GTPases were identified as the proteins encoded by the retrovirus oncogenes of the Harvey and Kirsten rat sarcoma (Ras) viruses (Figure 1). Since its initial discovery, over 150 mammalian proteins have been identified which show varying degrees of both amino acid sequence and structural similarity with Ras proteins. Small GTPases were identified by a variety of approaches, some fortuitously (e.g., Rho), in yeast genetic studies (e.g., Rabs), and others by nucleic acid sequence homology (e.g., Ral, Rap). These 'Ras superfamily' proteins are further divided into families (e.g., the Rho and Rab families) by sequence identity and cellular function. The major branches are the Ras, Rho, and Rab families, with smaller branches including Arf, Sar, Rad/Gem, and Ran proteins (Figure 2). Small GTPases are well conserved in evolution, with structural and functional homologues of many of the mammalian proteins found in yeast, flies, worms, and plants.

In addition to the GTPase activity, a feature of many small GTPases is their modification by lipids. Members of the Ras, Rho, and Rab families are modified by either the C15 farnesyl or C20 geranylgeranyl isoprenoid lipid (Figure 3). Protein prenylation is signaled by carboxyl-terminal sequences. For Ras and Rho small GTPases, this lipid modification occurs at a particular motif, the CAAX tetrapeptide motif, consisting of a cysteine followed by any two hydrophobic amino acids, and a terminal X residue that determines whether the protein will be modified by farnesyltransferase (X = S, M) or geranylgeranyl transferase I (X = L, F). Rab proteins possess cysteine-containing carboxyl-terminal motifs (CC, CXC, CCX, CCXXX) that signal geranylgeranyl isoprenoid post-translational modification by a third prenyltransferase, geranylgeranyltransferase. Arf proteins are cotranslationally modified by the C14 myristate fatty acid at the amino-termini. Lipid modification is typically critical for protein function and facilitates the association of small GTPases with specific membrane compartments. Ran is unusual among the small GTPases and does not undergo any lipid modification.

GDP/GTP Regulation and Structure

THE SMALL GTPASE CYCLE AND ITS REGULATION

The "on" and "off" switching nature of small GTPases can be best described as a GDP/GTP cycle (Figure 1). The processes of nucleotide exchange and GTP hydrolysis are able to occur at basal intrinsic rates that are characteristic to each small GTPase. However, there are two classes of regulatory proteins which act to increase the rate of these processes by several orders of magnitude: guanine nucleotide exchange factors (GEFs) increase the rate of nucleotide exchange, and GTPase-activating proteins (GAPs) increase the rate of the GTP hydrolysis. The Rho and Rab families of small GTPases have further regulatory proteins called guanine nucleotide dissociation inhibitors (GDIs).
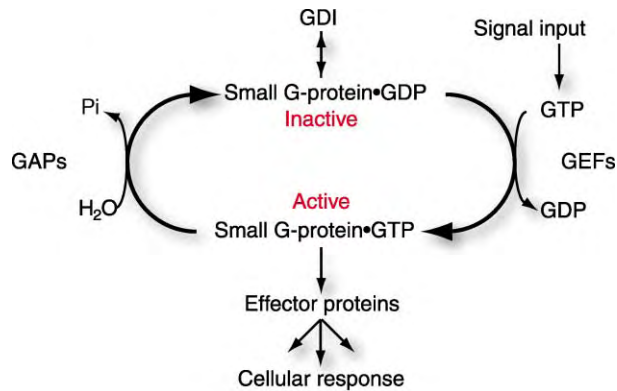
FIGURE 1   Small GTPases function as GDP/GTP-regulated molecular switches. The intrinsic GDP/GTP exchange activity is accelerated by guanine nucleotide-exchange factors (GEFs) to promote formation of the active, GTP-bound protein, whereas GTPase-activating proteins (GAPs) stimulate the intrinsic GTP hydrolysis activity to stimulate formation of the inactive, GDP-bound protein and the release of free phosphate (P_i). Guanine nucleotide dissociation inhibitors (GDIs) prevent nucleotide exchange as well as GAP stimulation. Small GTPase activation is most commonly mediated by a signal input that activates GEF function. The GTP-bound GTPase displays higher affinity for downstream effector targets, leading to stimulation of various cellular responses.

Guanine Nucleotide Exchange Factors

GEFs provide an important mechanism for regulated activation of small GTPases, and therefore their downstream effects. They are responsible for exchanging free cytosolic GTP for bound GDP. Exchange is not an active process; rather GEFs act to facilitate exchange of one nucleotide for another nucleotide, and it is the large excess of cytosolic GTP that favors the GTP exchange rather than GDP exchange. Furthermore, the binding of GTP to a small GTPase·GEF complex weakens the affinity of the GEF for the small GTPase causing their dissociation. The binding of GDP does not cause dissociation of the complex.

GEFs have been identified for many small GTPases. They are related by sequence within, but not between, the GEFs for specific small GTPase families. For example, there are currently two structurally distinct families of GEFs, which act upon either the Ras family (CDC25 homology domain-containing proteins) or the Rho family (Dbl homology domain-containing proteins) small GTPases. These two families act on the different Ras/Rho proteins since their catalytic domains are different, although their overall mechanism appears to be the same.

Different GEFs are Activated by Different Signaling Cascades   Specific members of the Ras and Rho GTPases can be activated by multiple GEFs; for example, there are at least 30 characterized Ras family GEFs, and over 150 Rho family GEFs. GEFs often contain a large number of different signaling protein domains which allow them to respond to a wide variety of signals and thus allow divergent signals to converge on a common small GTPase (Figure 4). In contrast, only a handful of GEFs have been identified for Rab family proteins, the largest family of small GTPases.

A common mechanism for activation of GEFs is through their recruitment to the membrane. The classical model of small GTPase activation is that of Ras and its GEF, son of sevenless (Sos). Upon activation of a receptor protein tyrosine kinase (e.g., EGFR) and its autophosphorylation of specific tyrosine residues, an adaptor protein, Grb2, is recruited to the plasma membrane via its Src homology 2 (SH2) domain (Figure 4). Grb2 interacts
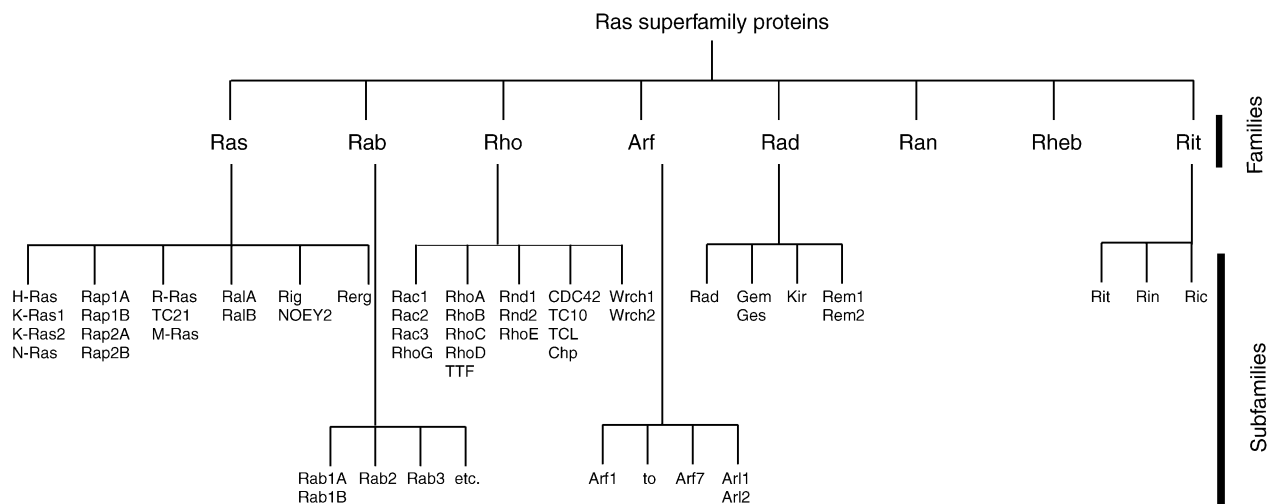
FIGURE 2   Small GTPase family.
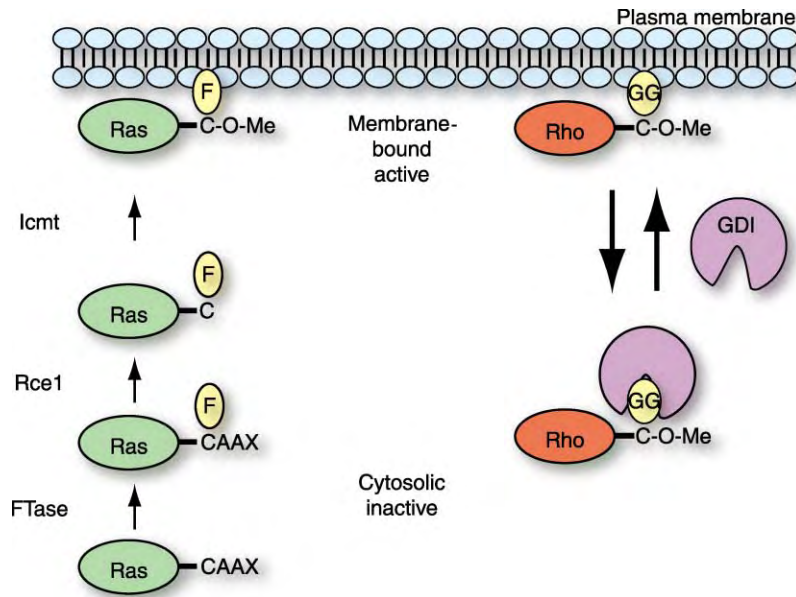
FIGURE 3 Post-translational modification by prenylation is important for the function of some small GTPases. Members of the Ras and Rho family of small GTPases terminate with CAAX tetrapeptide sequences that signal a series of post-translational modifications that promote association with membranes that is critical for function. For example, Ras proteins are synthesized initially as inactive, cytosolic proteins. Ras proteins are first modified by farnesyltransferase (FTase) which catalyzes the covalent addition of a C15 farnesyl isoprenoid (F) lipid to the cysteine residue of the CAAX motif. This is followed by Rce1-mediated proteolytic removal of the AAX peptide and Icmt-catalyzed carboxylmethylation (O-Me) of the now terminal farnesylated cysteine residue. Rho GTPases undergo the same series of modifications, with the first step catalyzed by geranylgeranyltransferase I (GGTaseI) and addition of the geranylgeranyl isoprenoid (GG). Rho GDIs recognize the prenylated form of Rho GTPases and prevent its association with membranes, thus leading to Rho GTPase inactivation.

via Src homology 3 (SH3) domain interaction with specific sequences on Sos, and recruits Sos to the membrane. The plasma membrane-bound Ras family proteins are now activated. Regulation of the production of membrane-bound lipids, such as diacylglycerol (DAG) or phosphatidylinositol 3,4,5-trisphosphate (PIP$_3$), are other mechanisms that regulate GEF membrane association and activation (Figure 4).
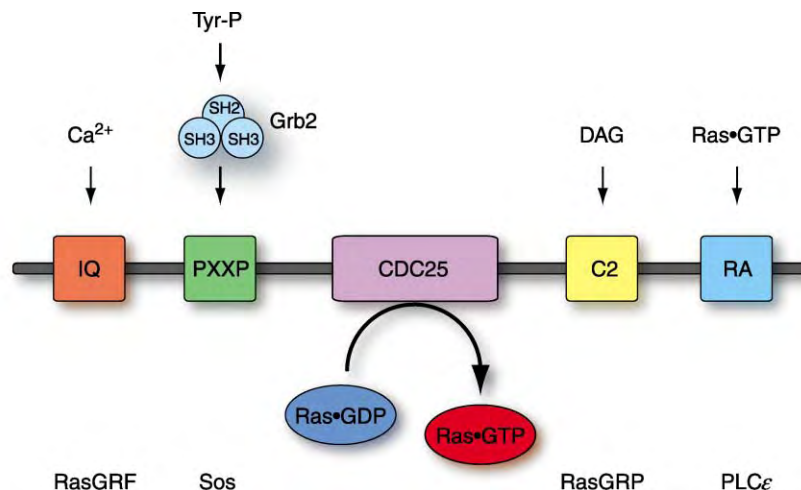
FIGURE 4 Ras guanine nucleotide exchange factors are activated by diverse upstream signals. Diverse extracellular signals promote Ras activation by stimulating the activity of different Ras GEFs. All Ras GEFs share a common CDC25 homology domain that catalyzes nucleotide exchange. The different Ras GEFs can be activated by calcium (RasGRF) through an IQ GAP homology domain, by tyrosine phosphorylation and interaction with the Grb2 adaptor protein and proline-rich sequences (PXXP) in Sos, by diacylglycerol (DAG) association with a C2 domain (RasGRP) or by association of Ras Association domains (RA) with other activated Ras family members.

GTPase Activating Proteins

The intrinsic hydrolysis rate of the majority of small GTPases is surprisingly slow, and sufficiently slow to be of little consequence in signaling cascades. Signals would not only be transduced too slowly, but the "on" state would persist for an excessive period of time. GEFs act to accelerate small GTPase activation, whereas GAPs act to accelerate the small GTPase hydrolysis and therefore make the signaling more transient (Figure 1).

In comparison to the vast number of Ras and Rho GEFs, the number of GAPs which act upon small GTPases is significantly smaller, fewer than ten for all the Ras family GAPs. There are structurally distinct GAPs for Ras and Rho family GTPases. Similar to the GEFs, these GAPs also typically contain additional sequences beyond the GAP catalytic domain. These sequences are involved in regulation, although much less is known about GAP regulation. Additionally, GAPs may possess functions, such as the effector function described for some Ras GAPs. Finally, the mutated Ras proteins found in human cancers contain single amino acid substitutions at glycine 12 (G12) or glutamine 61 (Q61), which render these mutants insensitive to GAP stimulation (Figure 5). These GTPase-deficient mutants are persistently GTP-bound in the absence of upstream stimulation and consequently cause deregulated effector activation. The experimental introduction of analogous mutations at residues corresponding to G12 and Q61 also renders other small GTPases GAP-insensitive and constitutively activated. Some wild-type small GTPases, however, possess naturally occurring sequence variation at these two residues, and are persistently GTP-bound proteins (e.g., RhoE/Rnd3). Thus, while GDP/GTP regulation is the major mode of small GTPase regulation, for some GTPases, regulation by other mechanisms (e.g., gene transcription, membrane association) is also important.

Other Small GTPase Regulatory Mechanisms

Guanine Nucleotide Dissociation Inhibitors (GDIs) GDIs have been identified for Rho and Rab proteins. Two distinct negative regulatory functions have been ascribed to GDIs. First, they bind to and mask the isoprenoid lipid modification in the carboxyl terminus of the small GTPase, thus preventing the association of the small GTPase with membranes (Figure 4). Second, their binding perturbs GAP and GEF regulation, preventing small GTPase activation. Three Rho GDIs and one Rab GDI have been described and can recognize multiple family members. The mechanisms of regulation of GDI activity are unclear, but may involve phosphorylation.

SMALL GTPASE STRUCTURE

Small GTPases and Conserved Structural Elements

The structure of many forms of Ras proteins has been determined and has established the general rules and requirements for all small GTPases. H-Ras was the first small GTPase to have its structure solved, and
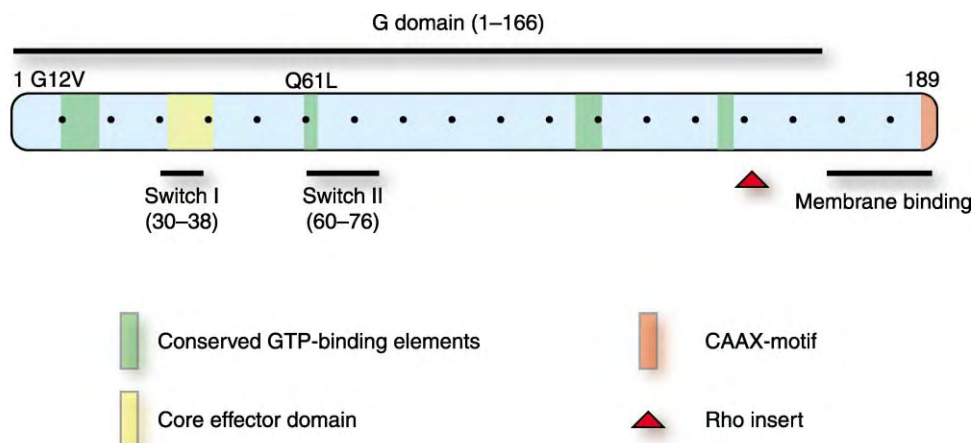
FIGURE 5 Small GTPase functional domains. Members of the small GTPase family share sequence similarities that define distinct functional domains. The residue numbers corresponding to H-Ras are shown. The G domain is comprised of a set of four conserved sequence elements involved in GTP-binding it alone is sufficient for the guanine nucleotide binding and GTPase activity, and is structurally similar among small GTPases. Mutated forms of Ras proteins are found in human cancers and possess single amino acid substitutions (e.g., G12V or Q61L) that render the protein insensitive to GAP stimulation and result in constitutively activated proteins. Similar mutations in other GTPases also result in constitutively activated proteins. Ras and Rho membrane association is facilitated by COOH-terminal sequences that include the CAAX motif as well as sequences upstream of this sequence. Effector binding involves the core effector domain (residues 32–40) and residues that change in conformation in the GDP- and GTP-bound states (switches I and II). Rho GTPases possess additional sequences (Rho insert) not found in other small GTPases. Some small GTPases contain additional NH_2- or COOH-terminal sequence extensions.

subsequent structures of related proteins have demonstrated a conserved overall canonical structural fold (designated the G domain) shared with all GTPases, but with variations in features for different small GTPase-family members (Figure 6).

All small GTPases possess a set of conserved sequence elements shared with other GTPases which represent the GDP/GTP nucleotide-binding pocket and GTP hydrolytic machinery (some, however, are GTPase deficient due to key residue changes), since the binding and hydrolysis of guanine nucleotides is the uniting element throughout small GTPases. A significant observation from Ras structural studies was the discovery of structural differences between the GDP- and GTP-bound forms, which are localized in two regions, termed switch I (residues 32–38) and switch II (residues 59–67). Similar switch-I and switch-II conformation changes have also been identified for the GDP- and GTP-bound states of other small GTPases.

Switch I and switch II show sequence divergence between families, although their loop-like structures are conserved. In addition to reflecting the nucleotide-bound state of the small GTPase, the switch regions are also involved in the interaction of small GTPases with effectors as well as GAPs and GEFs. The divergent sequence composition of these two regions contributes to the specificity of different small GTPases for these different binding partners.

Members of the Rho-family small GTPases possess a unique structural feature absent on all other small GTPases. The Rho family contains an insertion of 13 amino acids, called the insert region, positioned between Ras residues 122 and 123, which forms a surface-exposed loop (Figure 5). The primary sequence and secondary conformation of this insert region varies within the Rho family. Although the conformation of the insert region is not influenced by nucleotide binding, there is evidence that it is involved in effector interaction and activation.

Finally, other small GTPases possess additional amino- or carboxyl-terminal extensions important for function. For example, Arf and Sar1 proteins contain an amino-terminal extension necessary for insertion into and interaction with the membrane, whereas Ran has an elongated carboxyl-terminus that is crucial for its function in nuclear transport.

Diverse Biology and Function of Small GTPase

Despite their strong structural and biochemical similarities, small GTPases facilitate a remarkably divergent spectrum of cellular functions. This diversity is accomplished by the input signals that regulate the distinct GAPs and GEFs to modulate GDP/GTP cycling, and by the unique set of effectors that are recognized by the GTP-bound forms of small GTPases within and between families, resulting in many different cellular responses.

RAS PROTEINS AS SIGNALING NODES AND REGULATORS OF CELL PROLIFERATION

The frequent mutation of Ras proteins in human cancers has made these small GTPases the most intensely studied
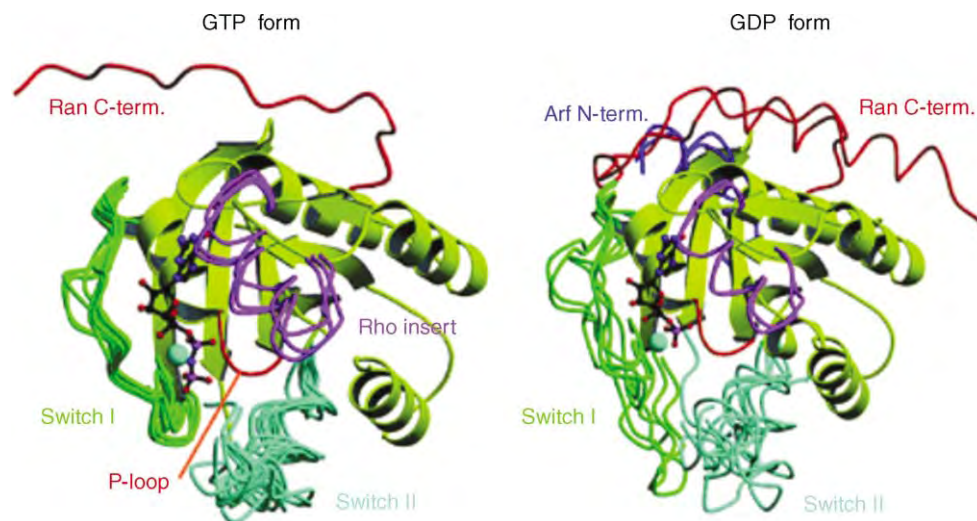
FIGURE 6 Canonical GTP conformation of small GTPases. Superimposition of a selection of Ras superfamily proteins on the G domain show the similar conformation changes in the switch-I and switch-II regions in the GDP- and GTP-bound forms. Rho GTPases possess additional sequences (Rho insert) not present in any other Ras superfamily proteins. Some small GTPases possess additional NH$_2$- and COOH-terminal sequences: the COOH-terminus of Ran is in red, the Rho insert in magenta and the Arf NH$_2$-terminal helix in blue.

and best characterized. Ras proteins serve as signaling nodes, where a wide diversity of extracellular signals such as growth factors (epidermal growth factor, platelet-derived growth factor), hormones (insulin), cytokines (interleukin-1), and the extracellular matrix proteins (via integrins) converge on and cause activation of Ras. Activated Ras in turn interacts with and regulates the activities of downstream effectors with highly divergent biochemical functions. Recent reviews have provided detailed discussions of Ras effector utilization. Therefore, we have highlighted various themes.

The Raf serine/threonine kinases are important effectors of Ras and facilitate activation of the ERK mitogen-activated protein kinase cascade (Figure 7). Ras promotes Raf activation by promoting the association of the normally cytosolic Raf with the plasma membrane, where a complex set of events, which includes protein phosphorylation, activates Raf kinase function. Activated Raf then phosphorylates and activates the MEK1 and MEK2 dual specificity protein kinases, which phosphorylate and activate ERK1 and ERK2. Activated ERK translocates to the nucleus and phosphorylates a variety of targets that include Ets family transcription factors. The Raf/MEK/ERK kinase cascade contributes significantly to the growth-regulatory functions of Ras.

The second best-characterized class of Ras effectors is the phosphatidylinositol 3-kinases (PI3K), a family of lipid kinases (Figure 7). A major activity of PI3K is the conversion of membrane-associated phosphatidyl-inositol 4,5-bisphosphate (PIP_2) to phosphatidylinositol 3,4,5-trisphosphate (PIP_3). PIP_3 in turn can regulate the activity of a diverse range of targets, including activation of the Akt serine/threonine kinase and Rac GEFs, via PIP_3 interaction with pleckstrin homology domains present in these proteins.

One significant theme that has emerged is that a number of Ras effectors are GEFs that link Ras with other small GTPases (Figure 7). First, there is a family of Ral GEFs that binds activated Ras and promotes activation of RalA and RalB, members of the Ras family of small GTPases. Second, Rin1 functions as a GEF for Rab5, thus linking Ras activity with regulation of vesicular trafficking. Third, Ras binds and activates Tiam1, which functions as a GEF for Rac. Fourth, Ras binds phospholipase C epsilon, which also contains a CDC25 homology domain that may activate the R-Ras small GTPase.

While Ras GTPases function as positive regulators of cell proliferation, some small GTPases appear to possess opposing functions and may function as tumor suppressors rather than oncogene proteins. This includes Rerg, NOEY2/ARH1, Rig, and DBC2. Thus, while these proteins share significant sequence and biochemical functions with Ras, clearly their effector functions are quite distinct. Additionally, whereas mutational activation of Ras proteins is associated with oncogenesis, it is the loss of gene expression of these proteins that accounts for their loss of function in tumor cells.
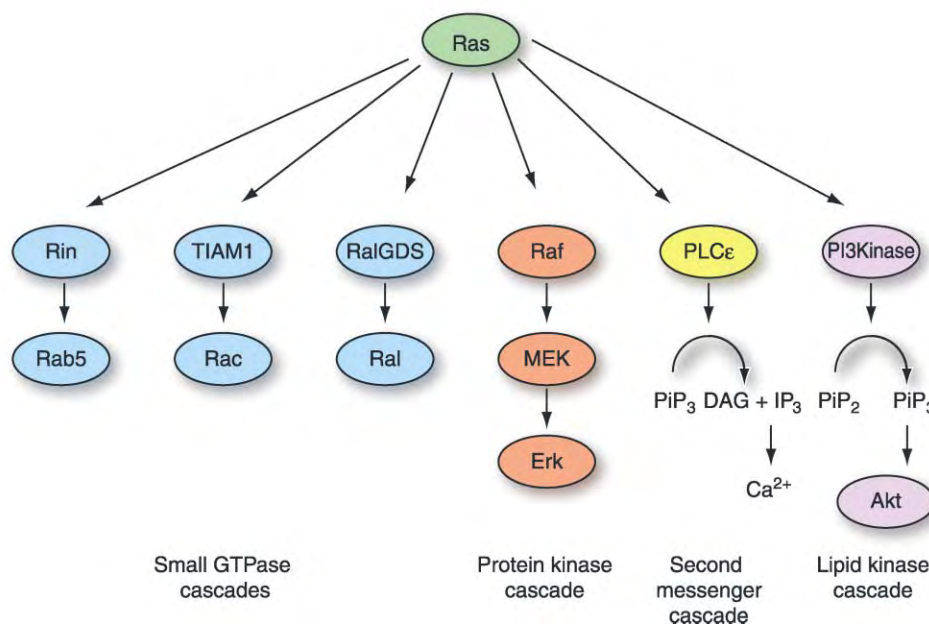
FIGURE 7 Ras effectors mediate diverse signaling outcomes. Activated, GTP-bound Ras can bind and regulate functionally diverse downstream effector targets. These include activation of a protein kinase cascade (Raf), activation of the PI3K lipid kinase, and activation of PLCE and the production of second messengers (DAG and calcium). Several Ras effectors are GEFs that activate other members of the Ras superfamily, including Rab5 (Rin1), Ral (RalGEFs), and Rac (Tiam1). Thus, Ras functions as a signaling node, where diverse signals converge and cause its activation (Figure 4), and once activated, results in activating of diverse downstream signaling pathways.

RHO-FAMILY PROTEINS AND ACTIN CYTOSKELETAL ORGANIZATION, CELL MORPHOLOGY, AND CELL MOVEMENT

Similar to Ras proteins, Rho-family small GTPases also function as signaling nodes activated by diverse extracellular signals. Perhaps the best-characterized cellular function of these proteins is their regulation of actin cytoskeletal organization, which in turn influences cell morphology, cell–matrix and cell–cell interactions, and cell movement. For example, RhoA causes actin stress fiber formation, Rac causes actin accumulation at the leading edge of a motile cell and lamellipodia formation, whereas Cdc42 promotes actin microspikes and filopodia formation. These changes in actin organization can influence cell shape, cell motility, and cell–cell interactions.

SMALL GTPASES AND MEMBRANE TRAFFICKING

Rab proteins constitute the largest subfamily of small GTPases, with more than 60 mammalian members. Rab proteins are involved in the modulation of specific steps in eukaryotic membrane trafficking in the secretory and endocytic pathways. Arf and Sar small GTPases are also involved in the regulation of cytoplasmic vesicular trafficking. By contrast, Ran is a regulator of nucleocytoplasmic transport. The regulation of the GDP/GTP cycle is also critical for the functions of these small GTPases in their distinct transport function.

SEE ALSO THE FOLLOWING ARTICLES

ARF Family • G$_{12}$/G$_{13}$ Family • Phosphatidylinositol Bisphosphate and Trisphosphate • Rab Family • Ran GTPase • Ras Family • Rho GTPases and Actin Cytoskeleton Dynamics

GLOSSARY

cytoskeleton A dynamic network of filaments, microfilaments, microtubules, and intermediate filaments.

effector proteins Proteins that bind preferentially to the GTP-bound form of small GTPases and facilitate the biological response of small GTPases.

GTPase High-affinity GDP-, GTP-binding, and GTP-hydrolyzing proteins that act as molecular switches and timers that cycle between inactive GDP-bound and active GTP-bound states.

GTPase activating protein Negative regulatory protein that accelerates the intrinsic GTP hydrolysis activity of small GTPases.

guanine nucleotide exchange factors Regulatory protein that accelerates the intrinsic GDP/GTP exchange activity of small GTPases.

guanine nucleotides Amine nucleotide bases composed of a guanine moiety attached to one (guanosine monophosphate, GMP), two (guanosine diphosphate, GDP), or three (guanosine triphosphate, GTP) phosphate groups.

isoprenylation Post-translational, covalent modification of proteins at carboxyl-terminal cysteine residues, by either C15 farnesyl or C20 geranylgeranyl isoprenoid lipids.

kinase An enzyme with phosphorylating activity on either proteins or lipids.

superfamily, family, and subfamily Small GTPases are classified into hierarchical phylogenies on the basis of structural, sequence, and functional similarity between members. Members of the Ras family share ~50% amino acid identity with the four Ras proteins whereas members of the Rho, Rab, and other Ras superfamily GTPases share ~25–30% amino acid identity with Ras proteins.

FURTHER READING

Der, C. J. (2002). Rho family proteins. In *Encyclopedic Reference of Cancer* (M. Schwab, ed.) pp. 799–804. Springer, Heidelberg, Germany.

Hall, A. (ed.) (2000). *GTPases*. Oxford University Press, New York.

Pruitt, W. M., and Der, C. J. (2002). Ras proteins. In *Encyclopedia of Cancer* (J. R. Bertino, ed.) 2nd edition, Vol 4, pp. 41–48. Academic Press, Orlando, FL.

Seabra, M. C., Mules, E. H., and Hume, A. N. (2002). Rab GTPases, intracellular traffic and disease. *Trends Mol. Med.* **8**, 23–30.

Vetter, I. R., and Wittinghofer, A. (2001). The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299–1304.

BIOGRAPHY

Adam Shutes graduated from Oxford University before receiving his Ph.D. from University College London in 2001. He is currently a postdoctoral Fellow at the Lineberger Comprehensive Cancer Center, UNC, where his interests lie in small GTPases and their regulation.

Channing Der is a Professor of Pharmacology at the University of North Carolina at Chapel Hill, where his research is focused on Ras-related proteins and their roles in cancer. He graduated from University of California, Los Angeles, before receiving his Ph.D. from University of California, Irvine, in 1981.

Somatostatin Receptors

Agnes Schonbrunn
The University of Texas Health Science Center, Houston, Texas, USA

Somatostatin receptors (sst receptors) comprise a family of five distinct plasma membrane receptors that bind the neuroendocrine peptides somatostatin and cortistatin. These receptors exhibit the seven-transmembrane domain structure typical of G protein-coupled receptors (GPCRs) and signal primarily through pertussis toxin-sensitive G proteins. The different sst receptor subtypes are found in specific endocrine and exocrine tissues, including the pituitary and the pancreas, in addition to being widely distributed in the central and peripheral nervous systems and the gastrointestinal tract. They have important physiological roles in inhibiting hormone secretion, particularly the secretion of growth hormone, insulin, glucagon, and gastrin, inhibiting exocrine secretion by the pancreas and stomach, and modulating neuronal excitability and smooth muscle contraction. Additionally, many neuroendocrine tumors express high levels of somatostatin receptors, and, because of their ability to inhibit tumor cell secretion as well as proliferation, somatostatin analogues are now used to target these receptors for both cancer therapy and diagnosis.

Physiological Somatostatin Receptor Ligands

Somatostatin (SS), originally called somatotropin-release inhibiting factor or SRIF, was discovered accidentally over 30 years ago while investigators were hunting for the brain peptide responsible for stimulating the release of growth hormone, or somatotropin. Surprisingly, they found that extracts of the hypothalamus, a specialized brain region that regulates pituitary hormone secretion, inhibited rather than stimulated the secretion of growth hormone. Using this inhibition as an assay, R. Guillemin, A. Schally, and their co-workers purified the active factor from hypothalamic extracts and identified the 14-amino-acid form of somatostatin, a discovery that contributed to their receiving the Nobel Prize in 1977 for studies of hormone production in the brain.

SS is now known to exist in two biologically active isoforms, 14 (SS14) and 28 (SS28) amino acids in length, which are produced by alternate proteolytic processing from a common precursor of 116 amino acids. These cyclic peptides are synthesized by specific endocrine, gastrointestinal, immune, and neuronal cells, as well as

some tumors, and act either locally as paracrine, autocrine, or neuronal modulators or through the bloodstream as hormones.

The cDNA for the related peptide cortistatin was discovered much later, in 1996, in the process of characterizing region-specific rat brain mRNAs. Cortistatin was named for its predominant expression in the brain cortex and its ability to depress cortical neuronal activity. Cortistatin is also synthesized as a precursor and is proteolytically processed into two biologically active products: a short (rat CST-14 and human CST-17) and an amino-terminally extended (CST-29) form. Although cortistatin is produced from a different gene than somatostatin, the two peptides have 10 of their 14 carboxy-terminal amino acids in common (Figure 1). Cortistatins are produced primarily by central nervous system neurons but have also been found in immune cells, lymphoid tissues, bone marrow, and the pancreas.

Because all forms of somatostatin and cortistatin bind to the different somatostatin receptor subtypes (sst's) with similar high affinities, it is not surprising that the peptides share many functional properties. However, they also produce some distinct biological effects. For example, intracerebroventricular injection of cortistatin in rats increases slow-wave sleep but not REM sleep, whereas somatostatin injection increases REM sleep. Thus, the identification of a human cortistatin selective G protein-coupled receptor (MrgX2, or Mas-related gene X2) fulfills prior predictions for the existence of distinct cortistatin-specific receptors. However, the rodent orthologue of this receptor has not been identified, and the relative roles of somatostatin receptors and cortistatin receptors in mediating the physiological actions of these peptides remain to be determined.

Biochemical and Molecular Characterization of Somatostatin Receptors

SOMATOSTATIN RECEPTOR SUBTYPES

Five somatostatin receptor genes, located on different chromosomes, have been identified and named *sst1* to
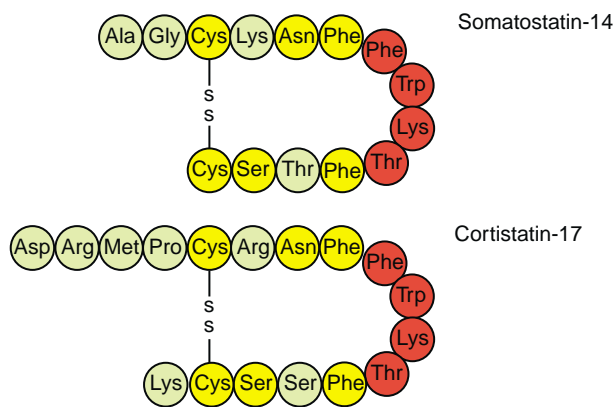
FIGURE 1 Structure of mammalian somatostatin-14 and human cortistatin-17. The amino acids in the red circles are required for high-affinity binding to somatostatin receptors. The amino acid differences between the two peptides are shown by the green circles.

sst5 in the order of their discovery (Table I). The coding regions of sst1, sst3, sst4, and sst5 all occur within a single exon. However, at least in some mouse and rat tissues, the mRNA for sst2 undergoes alternative splicing at the $3'$ end, producing two protein products. The sst2A variant is the product of the unspliced mRNA, whereas the sst2B variant contains an alternative exon that encodes a different and somewhat shorter receptor carboxy terminus. Only the sst2A variant has been detected in normal human tissues.

Somatostatin receptors belong to the G protein-coupled receptor (GPCR) family and are predicted to contain seven α-helical transmembrane domains. Receptors for somatostatin have been cloned from a variety of mammals as well as from several nonmammalian species, including a number of fish. The human sst receptors exhibit between 40 and 57% amino acid identity with each other, and sequence identity among the rat, mouse, and human orthologues of each receptor

subtype is even higher. The greatest sequence similarity between sst receptor subtypes occurs within the transmembrane domains; these domains, therefore, are thought to be involved in ligand binding. Sequence differences in the intra- and extracellular domains are presumably responsible for the unique signaling and trafficking properties of individual sst receptor subtypes.

Based on sequence similarity, sst receptors have been subdivided into two groups: sst1 and sst4 receptors form the SRIF2 subgroup, and sst2, sst3, and sst5 constitute the SRIF1 subgroup. In addition to a closer phylogenetic relationship (Figure 2), members of each group share several functional properties, such as their affinity for short synthetic somatostatin analogues, including octreotide and lanreotide, and sensitivity to agonist-induced receptor internalization (Table I).

The GPCRs most closely related to the sst receptor family are GPR7 and GPR8, which encode receptors for two paralogous brain peptides, neuropeptide B (NPB) and neuropeptide W (NPW) (Figure 2). The next most closely related family is the opioid receptor family, which also shares substantial sequence identity with the cortistatin receptor MrgX2 (Figure 2).

SOMATOSTATIN RECEPTOR STRUCTURE

Somatostatin receptors contain many of the structural features characteristic of GPCRs, including consensus sites for Asn-linked glycosylation within the amino terminus and multiple Ser and Thr phosphorylation sites in the intracellular loops and carboxy-terminal tail (Table I). All sst receptor subtypes except sst3 also contain a conserved Cys residue in the carboxy-terminal region, which provides a site for receptor palmitoylation. Lipid modification has yet to be demonstrated biochemically for any of the sst subtypes. However, sst1, sst2, sst3, and sst5 are known to be glycosylated from

TABLE I

Properties of Human Somatostatin Receptor Subtypes

Property	sst1	sst2A/B	sst3	sst4	sst5
Chromosomal localization	14q13	17q24	22q13.2	20p11.2	16p13.3
Reference sequence[a]	NM 001049	NM 001050	NM 001051	NM 001052	NM 001053
Amino acids	391	369/356	418	388	364
Asn glycosylation sites	3	4	2	1	3
Receptor glycosylation	+	+	+	−	+
Cys palmitoylation site	+	+	−	+	+
Receptor phosphorylation	+	+	+	−	nd[b]
Octreotide/lanreotide binding affinity	Low	High	Moderate	Low	Moderate
SS-stimulated receptor internalization	Slow	Fast	Fast	Slow	Fast

[a] Accessible at http://www.ncbi.nlm.nih.gov/.
[b] nd, not determined.
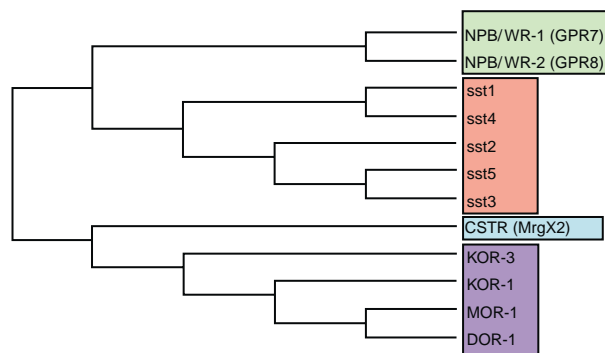
FIGURE 2 Sequence relationships between somatostatin receptors and closely related peptide receptors. The dendrogram was generated with the Clustal W program using MacVector's default parameters to align the peptide sequences. Human sequences were used in all cases and were obtained either from SwissProt or RefSeq. Somatostatin receptors: hsst1, P30872; hsst2A, P30874; hsst3, P32745; hsst4, P31391; hsst5, P35346. Cortistatin receptor: MrgX2, NP_473371. Opioid receptors: KOR-1, P41145; MOR-1, P35372; DOR-1, P41143; and KOR-3, P41146. Neuropeptide B/W receptors: NPBWR-1 or GPR7, P48145; and NPBWR-2 or GPR8, P48146. The plot shows that hsst1 and hsst4 form one subgroup and hsst2, hsst3, and hsst5 form a second subgroup within the somatostatin receptor family and that the cortistatin receptor is not a member of the sst receptor family.

the observation that treatment of these receptors with peptide-N-glycosidase F (PNGaseF), an enzyme that catalyzes the cleavage of N-glycosidically linked carbohydrate chains from Asn, causes a large decrease in their apparent molecular weights. Such a shift in molecular weight was not observed for sst4, either because it is not glycosylated or because glycosylation at its single consensus site does not significantly affect its apparent molecular weight.

Receptor phosphorylation, a covalent modification known to be important for GPCR regulation and signaling, has been examined biochemically for some but not all sst receptor subtypes. Somatostatin stimulates the incorporation of $^{32}PO_4$ into sst1, sst2A, and sst3 receptors, and protein kinase C activation also increases phosphorylation of sst1 and sst2A. However, phosphorylation of sst4 was not detected following agonist stimulation, and sst5 phosphorylation has not been investigated biochemically.

Pharmacology of Somatostatin Receptor Subtypes

The native somatostatin peptides exhibit little selectivity among the different receptor subtypes and have a very short lifetime in plasma (half-life < 3 min). Thus, analogues with increased metabolic stability and greater receptor specificity have been developed for therapy.

The earliest pharmaceuticals to target somatostatin receptors were short peptides, 6 to 8 amino acids in length, containing amino acids 7 to 10 of SS14, the region required for receptor binding (Figure 1). These agonists, including octreotide (SMS201-995), lanreotide (BIM 23014), and seglitide (MK-678), also contain D-amino acid substitutions and carboxy- or amino-terminal modifications to increase metabolic stability. These truncated peptides bind selectively to the SRIF1 group of receptors, exhibiting subnanomolar affinity for sst2 and nanomolar affinities for sst3 and sst5 (Table I). Further, octreotide does not bind to the MrgX2 cortistatin receptor. Both octreotide and lanreotide are used clinically to treat hormone-secreting pituitary adenomas and gastroenteropancreatic (GEP) tumors and comprise the only FDA-approved pharmaceuticals targeting sst receptors for therapy.

In an effort to generate receptor subtype specific analogues that are resistant to proteolytic degradation, S. Rohrer and co-workers at Merck screened combinatorial libraries to successfully identify the first highly selective, high-affinity, nonpeptide agonists for all sst subtypes. Such reagents offer the possibility of oral bioavailability, although they are not presently approved for clinical use. Nonetheless, they supply valuable tools for identifying the physiological roles of individual sst receptor subtypes.

The recognition that many tumors express several different sst receptors concomitantly has stimulated a search for stable somatostatin analogues that can bind to multiple somatostatin receptors. The hexapeptide analogue SOM230 shows high affinity for four of the five somatostatin receptor subtypes (sst1, sst2, sst3, and sst5), whereas the nonapeptide analogue KE108 appears to be truly universal since it binds to all five sst subtypes with nanomolar affinity. Because of their excellent metabolic stability, these broad-spectrum agonists have significant therapeutic potential.

The first somatostatin receptor antagonist, CYN-154806, described in 1996, showed high selectivity for the sst2 receptor. Since then, antagonists have been identified for all five somatostatin receptor subtypes, although none are used clinically.

In addition to their therapeutic applications, analogues of somatostatin have been developed to localize and stage sst receptor-positive tumors *in situ*. The radiolabeled derivatives used for tumor diagnosis contain a chelating group covalently attached to a stabilized somatostatin peptide, such that its receptor-binding properties are not compromised. The chelating group is then bound to a short-lived radioisotope, such as the gamma emitter ^{111}In, and injected into patients. After 24 h, accumulation of the radiolabel can be visualized in tumors expressing high levels of sst receptors by gamma camera scintigraphy, thus permitting localization of the tumors and their metastases. The analogue in current

clinical use is an indium-labeled octreotide derivative, [^{111}In-DTPA-D-Phe^1, Tyr^2]octreotide (OctreoScan), and hence is selective for the SRIF1 group of sst receptors, showing the most sensitive detection of sst2-expressing tumors. Unfortunately, this method cannot be used to localize tumors expressing sst1 or sst4 receptors, as appropriate somatostatin analogues are not yet available.

Somatostatin Receptor Signaling

Somatostatin receptors regulate a number of diverse signaling effectors, including adenylyl cyclase, phospholipases C and A2, calcium and potassium channels, protein and lipid kinases, and tyrosine and serine/threonine phosphatases. All sst receptors inhibit adenylyl cyclase via pertussis toxin-sensitive G proteins and thus decrease intracellular cyclic adenosine monophosphate (cAMP). However, other signal transduction pathways modulated by somatostatin receptors vary both with the receptor subtype and with the target cell.

The mechanism by which somatostatin receptors inhibit secretion in endocrine cells and neurons is understood in some detail. In addition to reducing intracellular cAMP levels, several sst receptors have been shown to reduce intracellular calcium levels in excitatory cells, also via pertussis toxin-sensitive G proteins. The reduction in cytosolic calcium can result either from a stimulation of various potassium channels, which hyperpolarize the cell membrane and thereby decrease influx through voltage-dependent calcium channels, or from direct calcium channel inhibition. The decrease in intracellular cAMP and calcium concentrations together contribute to the inhibitory action of somatostatin on secretion: when either signaling pathway is blocked, the magnitude of somatostatin's inhibitory effect is reduced.

Somatostatin stimulates contraction of intestinal smooth muscle cells by inhibiting adenylyl cyclase and activating phospholipase C-β3 via the α- and $\beta\gamma$-subunits, respectively, of pertussis toxin-sensitive G proteins. The activated phospholipase catalyzes the hydrolysis of the membrane lipid phosphatidylinositol 4,5-bisphosphate to form the second messengers inositol trisphosphate (IP$_3$) and diacylglycerol (DG). The binding of IP$_3$ to receptors on the sarcoplasmic reticulum results in the release of calcium from intracellular stores, producing a rise in cytosolic calcium concentrations. The released calcium forms a complex with the protein calmodulin, and this complex then activates myosin light-chain kinase (MLCK) to phosphorylate the light chain of myosin, leading to smooth muscle contraction. Because protein kinase A (PKA) decreases the sensitivity of MLCK to calcium, somatostatin inhibition of cAMP formation facilitates the calcium effect by reducing the activity of PKA. These pathways are also activated by somatostatin in aortic vascular smooth muscle cells. Interestingly, somatostatin does not activate phospholipase C-β3 in pituitary cells, even though the level of this enzyme in the pituitary appears to be similar to that in smooth muscle cells. The explanation for the signaling differences in these tissues remains to be elucidated, but could be due to differences either in the sst subtypes or in the signaling machinery present.

In endothelial cells, somatostatin inhibits cell migration, stress fiber assembly, and cytoskeletal reorganization produced by thrombin and other stimulators. In humans, these effects are mediated by the sst1 receptor and have been implicated in somatostatin's antiangiogenic actions. Although the molecular steps involved have not yet been identified, the mechanism includes an unusual pertussis toxin-independent inhibition of Rho, a low molecular mass GTPase that plays a central role in regulating cytoskeletal organization.

The mechanisms by which somatostatin inhibits cell proliferation and stimulates apoptosis are also poorly understood and appear to vary in the different cell types in which they have been examined. In most (though not all) cells, somatostatin activates the mitogen-activated protein kinase (MAPK) pathway and increases extracellular signal-related kinase (ERK)1/2 phosphorylation by a pertussis toxin-sensitive mechanism. However, this activation is often observed whether somatostatin inhibits or stimulates cell proliferation. Thus, it is likely that some of the other effectors activated by the sst receptors contribute to the final biological response. For example, in pancreatic acinar cells, which express sst2 receptors endogenously, somatostatin-induced growth arrest involves enhanced expression of the cyclin-dependent kinase inhibitor p27Kip and results from inhibition of the phosphatidylinositol-3-kinase (PI3K) pathway. In contrast, in sst2 transfected CHO cells, somatostatin induction of p27Kip appears to be dependent on stimulation rather than inhibition of PI3K, in that PI3K inhibitors block the effect of somatostatin analogues on ERK2 phosphorylation, which in turn is required for p27Kip up-regulation. These and other potential signaling pathways are under intense investigation in order to elucidate the clinically important actions of somatostatin to inhibit cell proliferation and stimulate apoptosis.

In summary, most but not all signaling by sst receptors involves the pertussis toxin-sensitive G$_i$/G$_o$ family. However, depending on the sst receptor subtype and the cellular environment, a spectrum of nonoverlapping signaling pathways can be activated. The link between the different effectors regulated by sst

receptors and particular biological responses is understood for some but by no means all of somatostatin's actions.

Somatostatin Receptor Regulation

Alterations in sst receptor responsiveness during somatostatin exposure vary dramatically between tissues, depending both on the response being measured and on the nature of the target cell. In some instances, desensitization occurs within minutes following initiation of somatostatin treatment, whereas in others no desensitization is detected even after years of somatostatin analogue exposure. For example, in chick sympathetic neurons, somatostatin inhibition of N-type Ca currents desensitizes with a half-life of 3 min. In contrast, desensitization to pharmacologic doses of the somatostatin analogue octreotide does not occur in months or years of therapy for pituitary tumors. The molecular basis for such dramatic differences in somatostatin receptor regulation remains poorly understood.

Studies of the different sst receptor subtypes in transfected cells indicate receptor-specific differences in their regulation. As has been shown for many GPCRs, hormone treatment leads to the rapid phosphorylation of several sst receptor subtypes (Table I). Interestingly, however, this phosphorylation appears to have different consequences for the different sst receptors. For example, in CHO cells, although both sst1 and sst2 receptors are phosphorylated and desensitized within minutes following somatostatin exposure, only the sst2 receptor is rapidly internalized.

The cellular environment also plays an important role in the observed differences in the regulation of somatostatin responsiveness. Interestingly, the sst2 receptor, which desensitizes within minutes of somatostatin exposure in cultured cells but is resistant to desensitization in human tumors, was recently shown to be phosphorylated *in situ* in a human tumor. Hence, at least this first step in the desensitization process appears to be intact in human tumor tissue. As the other molecular components of sst receptor desensitization are characterized, it will be interesting to determine whether their functions are altered in the desensitization-resistant tumors.

Summary

The five known somatostatin receptors play important physiological roles in the regulation of the endocrine, neuronal, gastrointestinal, and immune systems. In addition, they are recognized to be important targets in the diagnosis and therapy of a number of neuroendocrine tumors.

Although a great deal remains to be learned about many of somatostatin's functions, studies with sst1, sst2, and sst5 knockout mouse models, all of which are viable, have begun to delineate the biological importance of the individual receptor subtypes. Future focus on these receptors in their normal cellular environments can take advantage of the many new receptor-specific agonists and antagonists that have become available and will undoubtedly provide much-needed insight into the function of this physiologically and therapeutically important G protein-coupled receptor family.

SEE ALSO THE FOLLOWING ARTICLES

Adenylyl Cyclases • Mitogen-Activated Protein Kinase Family • Phospholipase A_2 • Phospholipase C • Rho GTPases and Actin Cytoskeleton Dynamics • Voltage-Sensitive Na^+ Channels

GLOSSARY

acromegaly A condition caused by the excess secretion of pituitary growth hormone after maturity, usually by a pituitary tumor. The disease is characterized by enlargement of the extremities, including the nose, jaws, fingers, and toes, as well as certain internal organs.

cortistatins The short (rat CST-14 and human CST-17) and long (CST-29) biologically active products produced by proteolytic processing of the precursor peptide encoded by the cortistatin gene.

G protein One of a family of related heterotrimeric proteins that bind GTP and GDP. The heterotrimeric forms, which are inactive, become activated at the plasma membrane by agonist occupied G protein-coupled receptors (GPCRs) that stimulate the binding of GTP to the G protein α-subunit and cause dissociation of the α-subunit from the $\beta\gamma$-subunit complex. When activated, the G protein subunits regulate downstream effectors, such as ion channels and enzymes that generate second messengers. G proteins become inactivated by hydrolyzing GTP to GDP, which then permits the reassociation of the α-subunit with the $\beta\gamma$-subunit complex.

somatostatins (SSs) The 14-amino-acid (SS-14) and 28-amino-acid (SS-28) peptides produced by alternative proteolytic processing from a single 92-amino-acid precursor called prosomatostatin.

FURTHER READING

Alliance for Cellular Signaling (2004). The Signaling Gateway. http://www.signaling-gateway.org/. ID numbers: A002204 for sst1, A002205 for sst2, A002206 for sst3, A002207 for sst4, and A000075 for sst5. Accessed May 2004.

Hofland, L. J., and Lamberts, S. W. (2003). The pathophysiological consequences of somatostatin receptor internalization and resistance. *Endocr. Rev.* **24**, 28–47.

Kreienkamp, H. J. (1999). Molecular biology of the receptors for somatostatin and cortistatin. *Results Probl. Cell. Differ.* **26**, 215–237.

Patel, Y. C. (1999). Somatostatin and its receptor family. *Front. Neuroendocrinol.* **20**, 157–198.

Reubi, J. C. (2003). Peptide receptors as molecular targets for cancer diagnosis and therapy. *Endocr. Rev.* **24**, 389–427.

Schonbrunn, A. (1999). Somatostatin receptors: present knowledge and future directions. *Ann. Oncol.* **10**, S17–S21.

Schonbrunn, A. (2001). Somatostatin. In *Endocrinology* (L. J. Degroot and J. L. Jameson, eds.) 4th edition, pp. 427–437. Saunders, New York.

Selmer, I., Schindler, M., Allen, J. P., Humphrey, P. P., and Emson, P. C. (2000). Advances in understanding neuronal somatostatin receptors. *Regul. Pept.* **90**, 1–18.

Slooter, G. D., Mearadji, A., Breeman, W. A., *et al.* (2001). Somatostatin receptor imaging, therapy and new strategies in patients with neuroendocrine tumours. *Br. J. Surg.* **88**, 31–40.

Spier, A. D., and de Lecea, L. (2000). Cortistatin: A member of the somatostatin neuropeptide family with distinct physiological functions. *Brain Res. Brain Res. Rev.* **33**, 228–241.

Weckbecker, G., Lewis, I., Albert, R., Schmid, H. A., Hoyer, D., and Bruns, C. (2003). Opportunities in somatostatin research: Biological, chemical and therapeutic aspects. *Nat. Rev. Drug Discov.* **2**, 999–1017.

BIOGRAPHY

Agnes Schonbrunn is a Professor in the Department of Integrative Biology and Pharmacology at the University of Texas–Houston, School of Medicine. She holds a Ph.D. in Biochemistry from Brandeis University and received her postdoctoral training at Harvard Medical School. Her principal research interests include the signaling and regulation of somatostatin receptors and the functional role that these receptors play in neuroendocrine cancers.

Spastic Paraplegia

Elena Irene Rugarli
Telethon Institute of Genetics and Medicine, Naples, Italy

Andrea Ballabio
University of Naples, Federico II, Naples, Italy

Hereditary spastic paraplegia (HSP) is a group of neuro-degenerative diseases characterized by weakness and spasticity of the lower limbs, loss of the vibratory sense, and urinary urgency. In 1983, Harding classified HSP in "pure" forms, when the symptoms are restricted to lower limbs weakness and spasticity, and in "complicated" forms, when the clinical picture is associated to a variety of neurological symptoms. These may include cerebral and cerebellar atrophy, optic atrophy, peripheral neuropathy, dementia, retinitis pigmentosa, amyotrophy. Age of onset is extremely variable, with a few forms arising in childhood, and most cases between 20 and 40 years of age. Usually the later the onset of the disease, the faster is the progression of its course. Although HSP is an invalidating disease, it never causes shortening of the life-span of patients.

Neuropathology

Hereditary spastic paraplegia (HSP) is characterized by the retrograde degeneration of the longest axons of the central nervous system: those of the crossed corticospinal tracts and of the fasciculus gracilis. In about half of the autoptic cases examined, the spinocerebellar tracts were also involved.

Axons of the corticospinal tracts arise from pyramidal neurons in layer V of the motor cortex and project through the internal capsule to reach the ventral surface of the medulla where they form two elongated swellings, the pyramids. These axons cross the midline at the junction between the bulb and the spinal cord and descend in the contralateral funiculus of the spinal cord. The crossed corticospinal tract conveys voluntary motor impulses to the secondary motor neurons located in the ventral horns of the spinal cord. The fasciculus gracilis is composed of the central branches of axons of pseudo-unipolar neurons located in dorsal root ganglia and ascend in the most medial part of the dorsal funiculus. These axons transmit deep sensory information from the lower extremities to secondary sensory neurons in the nucleus gracilis.

Characteristically, axonal degeneration in HSP involves preferentially the distal region of the axons. Degeneration of the corticospinal tracts is marked at lumbar levels, and less pronounced at cervical levels. In most cases the neuronal cell bodies are preserved. This pattern of axonal degeneration, also referred to as dying-back, is slowly progressive and occurs in several other conditions following a toxic, metabolic or genetic insult.

Genetics

It has become clear in the last decade that HSP is extremely heterogeneous from a genetic point of view. So far, 21 different loci have been mapped and it is expected that more will be identified. HSP can be transmitted as an X-linked (SPG1, SPG2, SPG16), autosomal dominant (SPG3A, SPG4, SPG6, SPG8, SPG9, SPG10, SPG12, SPG17, SPG19) or autosomal recessive trait (SPG5A, SPG7, SPG11, SPG14, SPG15, SPG20, SPG21). The molecular bases of the disease are also being unraveled with a quick pace. Eleven *HSP* genes have been identified, thus providing clues on the cellular pathways underlying the disease. Table I summarizes the current knowledge on the genetic bases of HSP. The following paragraphs describe the known *HSP* genes and classify them based on the existing functional information.

The Mitochondrial Forms

Two of the known HSP genes, SPG7 and SPG13, encode proteins that are involved in mitochondrial protein quality control. Their identification strongly indicates that at least a subgroup of HSPs should be regarded as mitochondrial diseases.

SPG7/Paraplegin

SPG7 was cloned in 1998 by Casari and colleagues, as the first autosomal recessive gene responsible for both pure and complicated forms of HSP. SPG7 encodes paraplegin, a 795 amino acid protein highly homologous to two yeast metalloproteases of the AAA family,

TABLE I

Genetic Bases of HSP

Acronym	MIM	Inheritance	Chromosomal location	Phenotype	Gene product	Proposed function
SPG1	312900	X-linked R	Xq28	Complicated	L1CAM	Development and guidance of the corticospinal tract
SPG2	312920	X-linked R	Xq21–22	Complicated	PLP	Axonal–glial interactions
SPG3A	182600	AD	14q11–q21	Pure	Atlastin	Guanilate-protein binding 1 homolog
SPG4	182601	AD	2p24-21	Pure	Spastin	Microtubule dynamics/nuclear role
SPG5A	270800	AR	8p21–q13	Pure		
SPG6	600363	AD	15q11.1	Pure	NIPA1	Transmembrane protein
SPG7	600146	AR	16q24.3	Pure/complicated	Paraplegin	Chaperone-like mitochondrial ATPase
SPG8	603563	AD	8q23–24	Pure		
SPG9	601162	AD	10q23–24	Complicated		
SPG10	604187	AD	12q13	Pure	KIF5A	Anterograde motor protein
SPG11	604360	AR	15q13	Pure		
SPG12	604805	AD	19q13	Pure		
SPG13	605280	AD	2q24	Pure	Hsp60	Mitochondrial chaperone
SPG14	605229	AR	3q27–28	Complicated		
SPG15	270700	AR	14q22–q24	Complicated		
SPG16	300266	X-linked R	Xq11.2	Complicated		
SPG17	270685	AD	11q12–q14	Complicated	Serpin	Endoplasmic reticulum protein, unknown function
SPG19	607152	AD	9q33–q34	Pure		
SPG20	275900	AR	13q12.3	Complicated	Spartin	
SPG21	248900	AR	15q22.31	Complicated	Maspardin	

Note: the loci SPG3B, SPG5B, and SPG18 are listed as "reserved" in the Hugo Gene Nomenclature Committee Database.

Yta10p (Afg3p), and Yta12p (Rca1p). Like its yeast homologues, paraplegin has an N-terminal mitochondrial leader sequence and two transmembrane domains that anchor the protein to the inner mitochondrial membrane. Paraplegin is composed of two functional domains: an ATPase domain typical of proteins of the AAA family (ATPases associated with diverse cellular activities) and a proteolytic domain with a consensus metal-binding site.

The yeast Yta10p and Yta12p assemble to form a high molecular weight complex of about 1 MDa, the m-AAA protease, which is embedded in the inner mitochondrial membrane but exposes proteolytic sites to the mitochondrial matrix (hence the prefix "m" in the name). The m-AAA protease ensures protein quality control system for inner membrane proteins, by guaranteeing the removal of non-assembled or misfolded proteins, and by performing crucial steps in mitochondrial biogenesis. Yeast cells lacking the m-AAA protease exhibit deficiencies in the expression of mitochondrially-encoded polypeptides and in the posttranslational assembly of respiratory chain complexes. These defects can be reproduced in yeast cells harbouring proteolytically inactive subunits, indicating that the phenotype is largely due to the loss of proteolytic activity. A recent study by Atorino et al. has shown that paraplegin co-assembles with the highly homologous AFG3L2 protein in human mitochondria to form a high-molecular-weight complex very similar to the one described in yeast. This complex complements the respiratory deficiency of yeast lacking the m-AAA protease, indicating functional conservation of the m-AAA protease across species and assigning proteolytic activity to the paraplegin/AFG3L2 complex.

Advancement in the understanding of the pathogenic basis of this form of HSP comes from the development by Ferreirinha et al. of a mouse model in which the Spg7 gene has been knocked-out. Paraplegin-deficient mice recapitulate the human disease, being affected by a late-onset progressive distal axonopathy, characterized by axonal swelling and degeneration, and involving the longest axons in the central and peripheral nervous systems. Axons swell because of accumulation of membranous organelles and neurofilaments, suggesting that axonal transport is impaired. Neurotracer studies in the peripheral nervous system of symptomatic Spg7 −/− mice indicated that retrograde axonal transport is delayed, suggesting that impairment of axonal trafficking processes may play a crucial role in the pathogenic process.

However, the first ultrastructural pathological sign that can be observed in paraplegin-deficient mice occurs several months before any evidence of axonal swelling and degeneration, and consists in the appearance of

abnormal mitochondria, first in synaptic terminals and then in the distal regions of the long spinal axons, that will later on undergo degeneration. Abnormal mitochondria include hypertrophic mitochondria, gigantic mitochondria, and mitochondria with disrupted or abnormal organization of cristae. This mitochondrial phenotype correlates with the onset of a measurable neurological impairment of the paraplegin-deficient mice on the rotarod apparatus, and is therefore the primary trigger of the pathological process. Since nothing is known so far about the substrates of paraplegin, the question of how lack of paraplegin is affecting mitochondrial morphology still awaits an answer. It is possible that abnormal mitochondria result from the accumulation of misfolded polypeptides that are not correctly cleared away by the paraplegin–AFG3L2 complex. Alternatively, this phenotype could underlie the abnormal processing of a regulatory molecule involved in some aspect of mitochondrial morphology.

Recent studies by Atorino *et al.* on fibroblast cell lines obtained from HSP patients with SPG7 mutations found a reduction of complex I activity and an increased sensitivity to oxidative stress, suggesting a potential mechanism for mitochondrial damage. This could be especially important in mitochondria located in nerve terminals where high energetic demands and increased concentration of free radicals are present.

SPG13/HSP60

Further support to the role of mitochondria in the pathogenesis of HSP came from the identification by Hansen *et al.* in 2002 of a missense mutation in the mitochondrial chaperonin heat shock protein 60 (HSP60) in one French family with an autosomal dominant form of pure HSP, mapping to chromosome 2q24–34. Hsp60 proteins belong to a conserved subgroup of chaperone proteins, termed chaperonins, which promote protein folding in the mitochondrial matrix, often in cooperation with the cochaperonin Hsp10. The mutation identified in the HSP family results in the V72I substitution. An elegant complementation assay performed in *E. coli* showed that only wild-type Hsp60, but not Hsp60–V72I, together with the cochaperonin Hsp10, can support the growth of bacteria in which the homologous chromosomal *groESgroEL* chaperonin genes have been deleted, indicating that the V72I mutant is functionally inactive. It is unclear whether this mutant leads to haplo-insufficiency through the formation of functionally inactive mixed chaperonin rings made up by both normal and mutated subunits, or may exert a genuine dominant-negative effect.

MITOCHONDRIAL PROTEIN QUALITY CONTROL AND AXONAL DEGENERATION: A PATHOGENIC HYPOTHESIS

The reason why the loss of ubiquitous mitochondrial proteins such as paraplegin and HSP60 causes the specific degeneration of a subset of axons, such as the corticospinal axons, is still to be unraveled. The most accepted hypothesis states that primary motor neurons would be particularly susceptible to any impairment of mitochondrial function due to their peculiar cellular homeostasis. These neurons have extremely long axons that can reach the length of more than 1 m in humans and depend for their life on efficient transport of organelles, molecular cargoes, and synaptic vesicles from the cell body to the distal tip of the axon, and back to the cell body. This process requires energy and is supported through the coordinated movement of mitochondria themselves along axons. Mitochondria located very distal to the cell body could be required to endure for longer period of time, and could be much less efficient in adapting to inefficient protein quality control mechanisms.

Genes Involved in Cellular Trafficking Events

A second group of genes involved in HSP appears to play a more or less direct role in cellular trafficking. Their identification supports the notion that axonal transport represents a key event for axonal homeostasis, and that any derangement from normal trafficking processes may lead with time to axonal degeneration. The long axons composing the corticospinal tracts and the fasciculus gracilis would be a privileged pathological target.

SPG10/KIF5A

The best example of the pathogenic role of impaired axonal transport in HSP comes from the identification by Reid *et al.* of a mutation in the gene encoding a specific neuronal kinesin heavy chain, KIF5A, in a family with an autosomal-dominant pure form of HSP. Kinesins are multisubunit complexes that work as motors for anterograde transport (from the cell body to the distal end of the axons) of membranous organelles and other macromolecular cargoes along microtubules. KIF5A is abundantly expressed in neurons, which is found in cell bodies, dendrites, and axons. The gene for the heavy chain of KIF5A is located within a large interval of chromosome 12 in which the SPG13 locus was mapped by cross-overs defined in a single large family affected by an autosomal-dominant pure form of HSP. A N256S missense mutation in KIF5A was identified in the

original family, and found to segregate with the disease. This mutation occurs at an invariant asparigine residue within the motor domain of the protein. Intriguingly, mutations at the corresponding residue in both the yeast Kar3 protein, a motor of the mitotic spindle, and the *Drosophila* Ncd kinesin motor, were found to block the microtubule-dependent stimulation of motor ATPase activity, thus acting with a dominant-negative mechanism. So far, the N256S mutation is the only one identified in HSP families.

SPG4/Spastin

Hazan *et al.* identified in 1999 the most frequent form of autosomal-dominant spastic paraplegia. SPG4 maps to chromosome 2p21-p22 and encodes a 616-amino-acid-protein, named spastin. Similarly to paraplegin, spastin belongs to the AAA (*ATPases Associated with various cellular Activities*) family, which is characterized by a conserved domain of 230 amino acids with ATPase activity. Based on sequence homology and phylogenetic analysis, spastin belongs to the subfamily-7 of AAA proteins, whose members are implicated in completely divergent cellular processes, such as microtubule severing and endosomal morphology and trafficking. The N-terminal part of the protein contains a newly recognized domain, the EPS or MIT domain, which is present in molecules involved in endosome trafficking, such as Vps4 and SNX15 (sortin nexin 15), and in spartin, another protein involved in HSP.

Mutations in SPG4 account for at least 40% of all autosomal-dominant HSP families. Missense, nonsense, and splice-site mutations as well as deletions or insertions have all been observed in the spastin gene. Notably, all the missense mutations fall into the AAA domain, with the exception of the S44L substitution that appears to be disease-causing only in the homozygous state, underlying the functional significance of this domain. The other mutations are scattered along the coding region of the gene and lead to premature termination codons, and mRNA instability, suggesting that haplo-insufficiency is the molecular cause of the disease.

The functional data available on spastin point to a complex cellular role. Recently, Charvin *et al.* found endogenous spastin to be localized in the nucleus, by using polyclonal antibodies raised against synthetic peptides. However, transient transfection experiments by Errico *et al.* suggested that the onset of spastin expression may be in the microtubule-organizing center and that, upon longer periods of expression, spastin may accumulate in cytoplasmic aggregates. The reason for the discrepancy between endogenous and exogenously expressed spastin is still to be determined.

Experimental evidence obtained in overexpression system suggests that spastin may interact dynamically with the microtubule cytoskeleton. In fact a stable association with a subset of microtubules and the formation of thick perinuclear bundles of microtubules are achieved when spastin mutants in the AAA domain are expressed in eukaryotic cells. A microtubule-binding domain was mapped to the N-terminal region of spastin. To explain these results, a hypothesis has been put forward that binding of spastin to microtubules may be transient *in vivo* and regulated through ATPase cycles. Mutations in the AAA domain would alter the ability of spastin to bind or hydrolyse ATP, and therefore entrap the protein in a microtubule-bound state.

To reinforce the idea that spastin is involved in microtubule dynamics, overexpression of wild-type spastin was found to promote microtubule disassembly in transfected cells. Although these data suggest that the degeneration of corticospinal axons, in HSP patients, could be due to impairment of fine regulation of the microtubule cytoskeleton, more studies are needed to demonstrate an interaction of endogenous spastin with the neuronal cytoskeleton, and to unravel its function in the nucleus.

SPG3A/Atlastin

The second most common gene involved in autosomal-dominant pure HSP with juvenile onset (~10% of the cases) has been linked to the SPG3A locus and found to encode a novel protein, atlastin, by Zhao *et al.* Four different missense mutations have been identified so far in SPG3A, all clustering in exons 7 and 8. Their mechanism of action is still to be determined. Although very little is known about the function of atlastin, this molecule shares very interesting homologies with members of the dynamin family of large GTPases, particularly with guanilate-binding protein-1. Dynamins are involved in important trafficking events in axons, including recycling of synaptic vesicles and maintenance and distribution of mitochondria, again suggesting that abnormal axonal transport may be at the basis of this form of HSP.

SPG20/Spartin

SPG20 is the gene involved in Troyer syndrome, an autosomal recessive form of HSP complicated by disarthria, distal amyotrophy, mild developmental delay, and short stature, that occurs with high frequency in the Amish population. The cloning of this gene by Ciccarelli *et al.* in 2002 has revealed that its protein product, spartin, shares homology in the N-terminal region with spastin within the ESP/MIT domain. This has led to the hypothesis that spartin may be somehow involved in ensodome trafficking. This theory still awaits experimental confirmation.

SPG21/MASPARDIN

Mast syndrome is mutated in an autosomal recessive complicated form of HSP associated with dementia present at high frequency in the Old Order Amish population. Simpson *et al.*, in 2003, have mapped this condition to chromosome 15q22.31 and identified the causative gene, SPG21. SPG21 encodes an acid-cluster protein of 33 kDa (ACP33), renamed maspardin, previously shown to localize in the endosomal/trans-Golgi network. This data again emphasize the causative role of proteins involved in cellular sorting and trafficking in the pathogenesis of HSP.

Developmental Genes

HSP may result from mutations in genes involved in the development of the corticospinal tracts. This is the case of SPG1, which has been linked to mutations in the cell adhesion molecule L1CAM. Consistently, the available data from neuropathological studies found absent or severely reduced pyramids. In this form of HSP, spastic paraplegia begins in the first two decades of life, with delayed acquisition of motor milestones and slow progression of symptoms. Although this form may manifest as pure HSP, it is more often observed in association with a complex disorder, referred to either as MASA syndrome (*m*ental retardation, *a*dducted thumbs, *s*pasticity, and *a*phasia) or CRASH syndrome (*c*orpus callosum hypoplasia, mental *r*etardation, *a*dducted thumbs, *s*pastic paraplegia, and *h*ydrocephalus). L1CAM is a transmembrane glycoprotein, which is expressed during development on the surface of long axons and growth cones, including those of the corticospinal tracts. L1 mediates cell adhesion, neurite outgrowth, axon pathfinding, and fasciculation through homophilic and heterophilic binding with a variety of extracellular and transmembrane molecules, and plays an important role in mediating guidance of corticospinal axons through the pyramidal decussation.

Myelin-Associated Genes

Disruption of intimate glia to axon interactions underlies the axonal degeneration observed in SPG2 patients. SPG2 maps to Xq22 and results from mutations in the
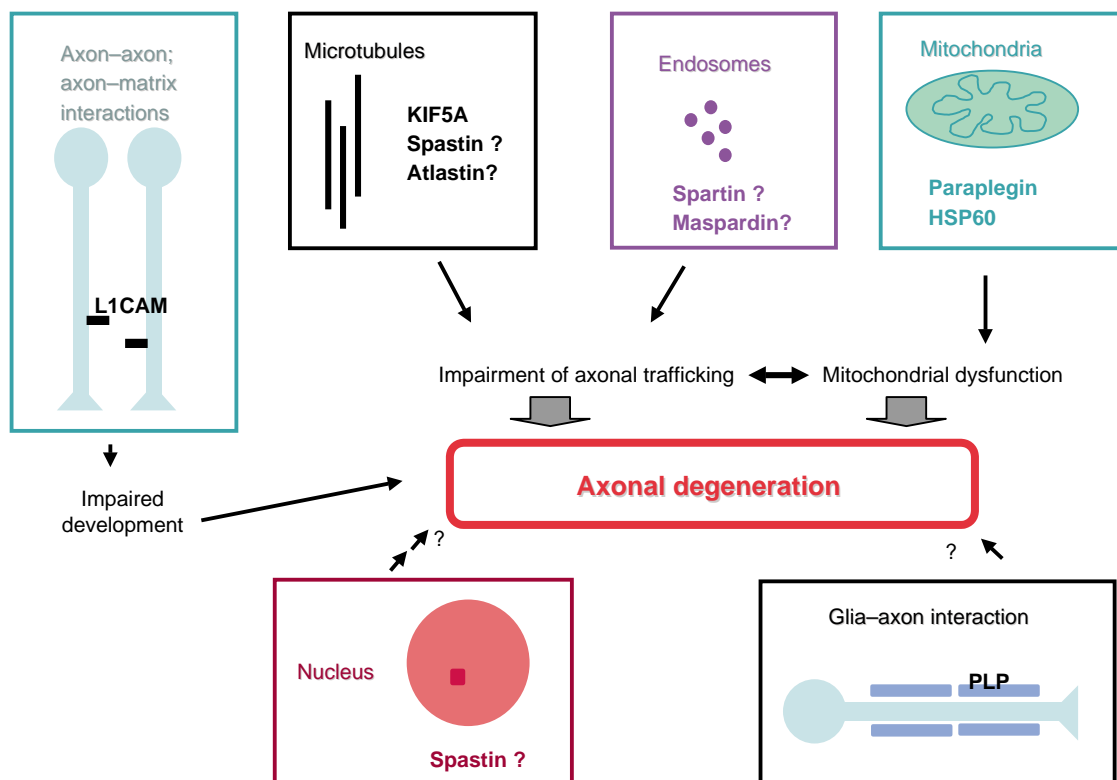
FIGURE 1   Proposed pathogenic mechanisms underlying axonal degeneration in HSP. Known proteins involved in HSP belong to different cellular compartments. Several of them, however, are believed to cause axonal degeneration either through mitochondrial dysfunction (paraplegin and Hsp60) or impairment of axonal transport (KIF5A, spastin, atlastin, spartin, and maspardin). Since mitochondria travel along axons, and play a major role in providing the energy necessary for axonal transport, the two pathways are likely to communicate. HSP can also be the result of defective development of the corticospinal tracts, as in the case of mutations in L1CAM, or of aberrant communications between glial cells and axons due to mutations in PLP1 protein. The role of spastin in the nucleus is still to be elucidated.

proteolipid protein (PLP), one of the major protein components of myelin in the central nervous system (CNS). PLP is a four-helix-spanning membrane protein that stabilizes the structure of the CNS myelin, by forming the intraperiod line. SPG2 comprises both pure and complicated forms of HSP, and is allelic to the severe hypomyelinating Pelizaeus–Merzbacher disease (PMD). A genotype-phenotype correlation between the nature of PLP mutation and the severity of the clinical picture has been described. In general, the milder phenotypes, such a pure spastic paraplegia, are observed when the gene mutations lead to the formation of protein products that are still able to traverse the secretory pathway and reach the cell surface. Surprisingly, knockout mice with complete absence of PLP, assemble compact myelin sheaths, but subsequently develop widespread axonal swelling and degeneration, most likely secondary to impaired axonal transport. It is still a mystery how the correct expression of PLP in olygodendrocytes provides support for myelinated axons. The future identification of putative signaling molecules would probably lead to a better understanding of this pathological cascade.

SPG6/NIPA1

In 2003, Rainier *et al.* reported mutations in the NIPA1 gene in a family with autosomal-dominant HSP linked to chromosome 15q11–q13 (SPG6). NIPA1 encodes for a predicted transmembrane protein, highly expressed in the brain. The function of NIPA1 is still unknown.

A Common Pathogenic Mechanism for HSP

The recent identification of several genes involved in HSP is beginning to shed light on the pathogenesis of this disorder. Although the known genes all belong to different families and have distinct subcellular localization (Figure 1), two common pathogenic themes have begun to emerge: (1) impairment of mitochondrial protein quality control and (2) defective axonal trafficking. These two routes likely intersect, since mitochondria themselves need to be transported along axons, and ATP is required to support transport of endosomes, cytoskeletal elements, synaptic vesicles, and other organelles. Future functional studies of the known

HSP proteins, together with the identification of new genes, will provide new exciting insights in this chapter of neuro-degenerative diseases.

SEE ALSO THE FOLLOWING ARTICLES

Chaperonins • DNA Oxidation • Kinesin Superfamily Proteins • Kinesins as Microtubule Disassembly Enzymes • Mitochondrial Genes and their Expression: Yeast

GLOSSARY

haplo-insufficiency A locus shows haplo-insufficiency when the amount of the gene product produced by a single allele is not sufficient to achieve a normal phenotype.
paraplegia Paralysis of both lower limbs.
spasticity Muscular weakness associated with increased stiffness and overactive reflexes.

FURTHER READING

Casari, G., and Rugarli, E. (2001). Molecular basis of inherited spastic paraplegias. *Curr. Opin. Genet. Dev.* **11**, 336–342.
Crosby, A. H., and Proukakis, C. (2002). Is the transportation highway the right road for hereditary spastic paraplegia? *Am. J. Hum. Genet.* **71**, 1009–1016.
Harding, A. E. (1983). Classification of the hereditary ataxias and paraplegias. *The Lancet* **21**, 1151–1154.
Langer, T., Kaser, M., Klanner, C., and Leonhard, K. (2001). AAA proteases of mitochondria: Quality control of membrane proteins and regulatory functions during mitochondrial biogenesis. *Biochem. Soc. Trans.* **29**, 431–436.
Reid, E. (2003). Science in motion: Common molecular pathological themes emerge in the hereditary spastic paraplegias. *J. Med. Genet.* **40**, 81–86.

BIOGRAPHY

Elena Irene Rugarli is a Group Leader at the Telethon Institute of Genetics and Medicine (TIGEM) in Naples, Italy. Her principal research interest is in the field of neuro-degeneration, in particular, the pathogenic mechanisms of hereditary spastic paraplegia. She holds an M.D. from the University of Milan, Italy, and received her postdoctoral training at Baylor College of Medicine in Houston, TX.

Andrea Ballabio is Director of the Telethon Institute of Genetics and Medicine (TIGEM) and Professor of Medical Genetics of the Federico II University in Naples, Italy. He holds an M.D. and is board-certified in Pediatrics. His main scientific interests comprise the identification of the molecular basis and pathogenetic mechanisms of genetic diseases as well as the use of functional genomics in medicine.

Spectrophotometric Assays

Britton Chance
University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

This brief history describes the impact of technical developments upon the progress of biophysics and biochemistry and focuses upon the contributions of optical spectroscopy. The optical method (now termed spectroscopy) served Otto Warburg well, and in the mid-1930s he developed a unique spectrophotometer. The device was so complicated that it was likely never used. However, it had some nice features; particularly, instead of balancing with slits and neutral density filters, it balanced intensities in the two beams by means of rotating sector discs. Furthermore, it had a high-quality monochromator that allowed Warburg to identify the important absorption bands of NADH (nicotinamide adenine dinucleotide; called DPNH at the time) and furthermore to identify the fluorescence of the reduced state. The optical determination of NADH was one of the great triumphs of analytical biochemistry, as evidenced by the large number of assays through which it was developed, primarily in Germany, where it eventually becomes the basis of the Boehringer company, and providing Olle Lowry, Janet Passoneau, and collaborators with a unique handle on micro-enzymatic assays. This quantitative method far surpassed the visual spectroscopic studies of David Keilin. Indeed, even the keen eyes of Keilin failed to confirm the absorption bands of flavoprotein and NADH as respiratory carriers. Surely flavoprotein was detectable, but because both flavoprotein and NADH were considered "Warburg's territory," it is not surprising that Keilin did not concentrate on flavin.

Manual Spectrometers

The rightfully famous physicist R. W. Wood made a high-quality plastic replica grating and incorporated a pair of them in the Coleman spectrophotometer to make a double-grating instrument that was valid from 400 to 700 nm, had a very good f number, and had a resolution independent of wavelength ($\sim$ 20 nm). This was an essential element of the kinetic and spectroscopic study of the enzyme–substrate compounds of catalases and peroxidase, as carried out in the laboratory of Hugo Theorell. While it was an excellent commercial instrument for laboratory work in a visible region, Cary's quartz monochromator, however, together with Beckman's unreferenced exploitation of DuBridge's ingenious

electrometer (of the pH meter as well), made the Beckman DU instrument superior. However, if Charles Chaplin were to make a film of the biochemist similar to "City Lights," this instrument would serve the purpose. Dark current adjustment, slit-adjustment, cuvette holder shifting, and balancing to a null by hand were the earmarks of the instrument that most biochemists used from the 1940s to the early 1960s. Not that there were no alternatives. C. C. Yang invented what was called the "Yang machine," a very simple and robust instrument used by many in the Johnson Foundation laboratory, which used a vibrating mirror system to send light through a reference cuvette and a measure cuvette, together with dynode feedback, and gave a logarithmic output. Concurrently, Lenart Åckerman in the laboratory of Bo, Thorell in Stockholm invented the same machine that also provided spectral scans but that was held off the market for about a decade so the DU could have an orderly and prosperous senescence.

Dual-Wavelength Technology

The development of dual-wavelength systems began with Glen Millikan, who ingeniously scribed a barrier in the Weston photovoltaic cell and used green and red filters to measure myoglobin spectral changes in the visible region. However, he connected the output to a mirror galvanometer so that the system was intrinsically quite slow and thus unsuitable for the rapid measurements required for the flow apparatus. The dual-wavelength principle was continued in the Millikan ear oximeter used so effectively during World War II.

Rapid spectrophotometry required phototubes. The significant experience Britton Chance had acquired with them by inventing an automatic steering device while still in his teens led him to use cesium oxide on silver and then antimony as photocathodes. These were combined with the technology of high-gain DC amplification from the electrophysiologists Adrian, Gasser, Erlanger, and the Schmitt brothers, who all had to develop an entire field of electronic amplification for studying nerve action potentials. The three-wavelength system used to measure the kinetics of

enzyme–substrate compound formation in the Soret region, together with the overall formation of leucomalachite green to malachite green, provided measurements of enzyme–substrate compounds that required a mechanical differential analyzer for the transient solution of the Michaelis–Menten/Briggs–Haldane equations for enzyme kinetics (Figure 1). Indeed, the rapid flow apparatus led the field of fast spectrophotometric determinations, not only of the enzyme–substrate compounds of peroxidase but also, in the hands of Quentin Gibson, for measuring hemoglobin kinetics.

The introduction of the chopper-stabilized amplifiers in systems exploited during World War II by the Group 63 of Chance and colleagues (Precision Circuits Group at MIT RadLabs) led to a whole new generation of low-level amplifiers and was the basis for many commercial instruments. Furthermore, the basis for the integrated circuit was afforded by the subminiature triodes developed for the proximity fuse, for which the rejection rate was so high that thousands of these diodes suitable for peacetime applications became available and were made in the first step of integrated circuits, precursors of the integrated chips that were soon to come in the post-war era.

PHOTOCHEMICAL ACTION SPECTRA

One historic experiment, which now seems almost to have been forgotten, was performed by the Schmitt brothers in order to validate the Warburg hypothesis, namely, to obtain the photochemical action spectrum for the propagation of action potentials in nerves. The results did, in fact, coincide with Warburg's hypothesis and provided the much-needed link between prokaryote and eukaryote. The Schmidt brothers received very

---

Michaelis-Menten Theory

These reactions are represented by Briggs and Haldane as the bimolecular combination of the enzyme, E, and substrate, S, to form an intermediate compound, ES, followed by a monomolecular decomposition into free enzyme and activated or altered substrate, Q, representative of the products of the "over-all" enzyme action.

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + Q \tag{4}$$
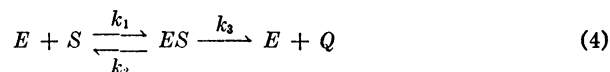
If e is the total molar enzyme concentration, x the molar substrate concentration, p the molar concentration of ES, k_1 the second order rate constant, and k_2 and k_3 the first order rate constants, then

$$\frac{dp}{dt} = k_1 x(e - p) - (k_2 + k_3)p \tag{5}$$

$$\frac{dx}{dt} = -k_1 x(e - p) + k_2 p \tag{6}$$

These two equations represent the rate of formation of the intermediate compound and the rate of disappearance of the substrate.

The solution of these equations has been already obtained by Briggs and Haldane for the special conditions of the steady state, when

$$p = p_{max.}, \qquad \frac{dp}{dt} = 0, \qquad \text{and} \qquad \frac{p_{max.}}{e} = \frac{x}{x - Km} \tag{7}$$

where

$$Km = \frac{k_2 + k_3}{k_1} = x \frac{(e - p_{max.})}{p_{max.}} \tag{8}$$

A further solution valid during the steady state is obtained by adding Equations 5 and 6,

$$\frac{dx}{dt} = -k_3 p_{max.} \tag{9}$$

where dx/dt is the rate of disappearance of substrate.

FIGURE 1   Michaelis–Menten theory.
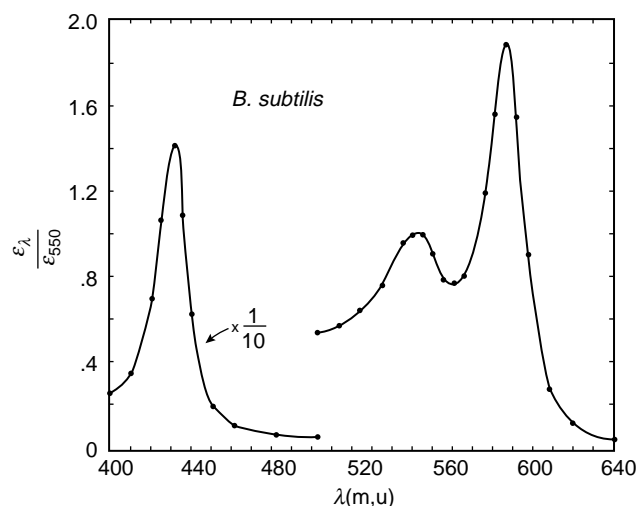
FIGURE 2 Photochemical action spectrum of cytochrome oxidase as obtained by the LeRoy Castor using the oxygen electrode to measure light activated respiration. It is seen that this spectrum is much better defined than Warburg's classic spectrum of yeast where mainly the lines of the mercury arc were employed. Here a continuous monochromated light would be used and the σ bands were clearly delineated as are the Soret bands.

little credit for their work, without which Warburg's and Keilin's work, based on yeast cells only, would have been less acceptable (Figure 2). At the time, there was no reason to believe that the eukaryotic yeast cell was a model for tissue studies. LaRoy Castor quickly improved the action spectral method and also discovered a heme prosthetic group oxidase (cytochrome o).

Tissue Spectroscopy

Glen Millikan was a pioneer in shining light through tissue and recording myoglobin deoxygenation in functional activity. During the same time, there were many microspectroscopic studies by Arvanitaki and Chalazonitis, and also, notably, by the laboratory of Casperson with his colleague Bo Thorell. Lundegårth was a pioneer in applying the spectroscopic technique to bundles of plant roots, but was unable to correct for the large changes of scattering due to the osmotic activity of the roots. However, Lou Duysens developed a sensitive single-beam method that he applied with diligence and effectiveness to cytochrome components of green leaves.

A time multiplex dual-wavelength system made possible by simply mounting on top of the "Brown" converter, or vibrating a small mirror at 60 Hz, allowed time sharing of two wavelengths through suspensions of highly scattered material derived from heart muscle or

from liver as intact mitochondria. This dual-wavelength machine was the precursor of the time multiplex systems that have been used for exploiting near infrared spectroscopy.

THE SCATTERED LIGHT PROBLEM

It was recognized early on that the scattered light would confound spectrophotometric measurements of pigments of tissue. Yet, in 1949, Chance recognized that the scattered light effect would be greatly diminished if the solid angle of the detector system were large instead of small, as in commercial spectrophotometers (for long cuvettes). Thus, instead of a 10-cm-long sample compartment, as used in most of the Beckman/Cary devices, Chance and colleagues placed a large area detector and an "end-on" photomultiplier almost in contact with the sample, particularly using the time-shared dual-wavelength system, from which very flat baselines were obtained for spectroscopic differences as long as the wavelength difference was not too great. The difference spectrum obtained from guinea pig liver mitochondria is shown in Figure 3 as an example of the data that were obtained with the dual-wavelength spectrophotometer, where the baseline was set at the suspected isosbestic points of 540 and 630 nm. The action spectrum for reduction of the oxidized form of the electron carriers was measured. This spectrum, although validating Keilin's microspectroscopy, clearly showed some anomalies. Cytochrome c_1 was a separate component. A trough due to the disappearance of oxidized flavin and peaks due to appearance of reduced NADH were not identified by Keilin's keen eye, which proved to be the most useful components of the respiratory chain because they were highly fluorescent. But most importantly, they were inhabitants of the mitochondrial matrix space where most of the citric acid cycle reactions were carried out. The instrument was surely very worthwhile. It had not only identified three components of the respiratory chain that had been previously underestimated or remained completely unobserved, but also, for example, the difference spectrum obtained by Lucille Smith and Aristid Lindenmeyer showed the P450 compound of carbon monoxide, an observation confirmed by G.R. Williams and used by P450 enthusiasts today.

CONTROL OF RESPIRATION

No other phenomenon was as important to physiologists as the control of respiration. The topic had received very little attention until Lardy and Wellman's historic paper, which showed that isolated prepared mitochondria very nearly completely ceased respiration in the absence of ADP and phosphate but restarted respiration upon addition of ADP and phosphate. This led Chance and
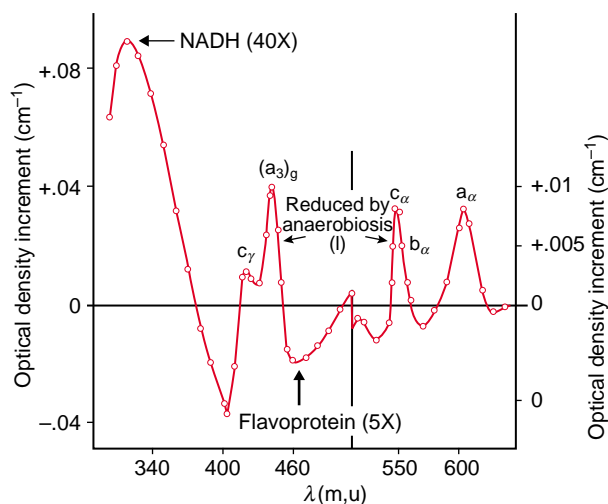
FIGURE 3 The difference spectrum of intact phosphorylating mitochondria. This was the first difference spectrum for oxidized minus reduced guinea pig liver mitochondria obtained with the dual wavelength method employing reference wavelengths of 540 nm for the alpha bands and 450 nm for the Soret band. The incremental absorption with respect to that reference is indicated giving the ratio of the Soret to the alpha bands to be 4 and the NADH peak to be 40 times that of the alpha peaks. The flavoprotein trough, i.e., oxidized flavoproprotein is the absorber, is shown at 460 nm. While the Soret peak of cytochrome oxidase was visible at 444 nm by Keilin's microspectroscope, the flavin trough and the NADH peak (the latter being responsible for the intense autofluorescence of cells) was not available to Keilin in Hartree's heart muscle preparation in which the absorbance of these two components was greatly diminished by the elimination of cytric acid cycle enzymes.

Williams to explore the nature of activation of the electron carrier in the respiratory chain. They found that that all spectroscopically identifiable components changed their steady-state values when ADP was added to the suspension of resting mitochondria. This phenomenon now provides the basis for all of the metabolic activations studied by magnetic resonance imaging (MRI) as the bulb effect, and for using the saturation of oxyhemoglobin *in vivo* as a measure of metabolic activity.

At the same time, Lehninger, Chance, and others studied the uptake of calcium into the mitochondrial matrix, and Chance and Williams found that calcium activated the suspensions of mitochondria to the same extent and with the same speed as did ADP suggesting, as recently discussed by Carafoli, the presence of a calcium carrier in the mitochondrial membrane.

LOW-TEMPERATURE TECHNOLOGY

Perhaps the most secure instrumental data were those obtained at low temperature in liquid nitrogen, where the spectra were more intense. The delineation of the absorption bands, particularly in the bc_1 region, turned out to be essential in showing the two kinds of cytochromes *b*.

LUBBERS' RAPID SCAN

Tissue absorption and scattering occupied the minds of a number of workers, particularly those of Lubber and Thews, who not only did much theoretical work on scattering but also perfected a galvanometer-driven wavelength scanning device, whose speed far exceeded that of anything else available at the time and which has only been perfected recently by the work of Desr and his colleagues.

PHOTOACTIVATION STUDIES

Perhaps the greatest advantage of the technique for observing tissue properties was the fact that it was possible to photoactivate the system under spectroscopic examination and observe the photoinduced absorption spectrum, a specialty of the Laboratory of Lou Duysems. This, of course, required a measure of common mode rejection, at which the dual-wavelength system excelled because the light flashes shared a single detector with an AC coupling, allowing the system to reject leakage of not only the photolysis light but also other disturbances that affected the common mode. Furthermore, illumination at subzero temperatures of the solid phase allowed the study of many photoactivated systems containing derivatives of chlorophyll or those rendered photoactivatable by combining with a ligand such as carbon monoxide. Thus, optical measurements in the region of the Soret band and illumination in the region of the alpha band allowed a measure of protection of the dual-wavelength detector system from very strong photolysis light, for example, that provided by intense sources such as the carbon arc. These studies bridged the gap between Otto Warburg's unbridgeable action spectrum experiment on yeast and Keilin and Hartree's microspectroscopical observations of cytochrome oxidase in yeast and in a pigeon heart muscle preparation.

Not only the cytochrome oxidase–CO complex CO, but also the splitting and recombination of myoglobin CO compounds were studied in detail, with optical spectroscopy as well as structural methods such as EXAFS (X-ray absorption spectroscopy), with results that make headlines even today for the subject of ligand docking concomitant with structural changes, confirming the X-ray structural studies in which the E7 histidine was altered by a ligation of myoglobin with CO.

THE RUBY LASER AND ELECTRON TUNNELING

Lacking powerful light sources of short duration, Chance and colleagues relied mainly upon xenon flash

lamps, particularly those of Edgerton and colleagues. However, construction of the first ruby laser used in biological studies by Bunkenberg and DeVault opened up a whole new time domain, not only because of the monochromaticity of the laser (which greatly simplified the filter leakages) but also because of the duration of the light pulse (in the microsecond region), which for the time was extraordinarily short. From this study came the totally unexpected observation that the photochemical oxidation of cytochrome *c* by photo-activated chlorophyll of *Chromatium* was temperature independent, not only from room temperature to liquid nitrogen temperature, but eventually at liquid helium temperatures, extending the time range of the spectro-scopic method to one of the primary or near primary results and/or reactions of photosynthesis (Figure 4), and also introducing the concept of electron tunneling in biochemical reactions, a concept pursued by Les Dutton and Harry Gray.

The history of spectroscopy is checkered with unexpected results and meaningful interpretations on the nature of electron transfer reactions. In fact, the study of reactions obtained from photolysis activation of CO compounds in the presence of O_2 caused ligand exchange to occur, especially in the case of CO. Together with Carlo Saronio, Chance discovered a number of inter-mediate steps involving higher oxidation states of iron mirroring those obtained with peroxidase and H_2O_2, but much faster and more complex. This was because copper oxidation was involved in the active site of CO, enabling donation of four electrons sequentially for the reduction of oxygen to water without significant amounts of free radical intermediates. This contrasted with the photo-activation of porphyrins in the absence of electron donors, resulting in the creation of singlet oxygen, a process used extensively in photodynamic therapies.

FLUOROCHROMES OF TISSUES

Spectroscopy of biological fluorochromes, such as NADH and flavoprotein compounds that exhibit
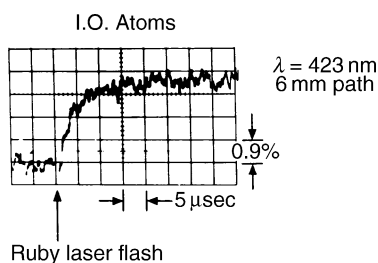
FIGURE 4 Typical recording of the oxidation of cytochrome c of photosynthetic bacteria at low temperatures as activated by a light flash of the ruby laser. Remarkably this reaction rate was affected very little by the transition from room temperature to liquid nitrogen temperature due to electron tunneling.

strong fluorescence in mitochondria, led Chance and co-workers to study them extensively *in vivo* by a simple method in which excitation was obtained by the very strong 366 and 436 lines of the mercury arc. This yielded emission spectra characteristic of the chromophores in a variety of functional states, characterizing electron transfer through the citric acid cycle as reductant and the cytochrome chain as oxidants, giving essential features of the redox state of the mitochondrial matrix. Interpretations of the importance of metabolic control and thermodynamic principles of the system were derived from these, particularly by R. L. Veech.

HEMOGLOBIN AND CYTOCHROME

It is perhaps ironic that the cell biologists and etymologists of the Molteno Institute focused on the cytochromes of hemoglobin-free cells and tissues and that the "real" physiology of oxygen delivery to tissues by hemoglobin and its utilization by CO had to wait for more sophisticated methods. Even now, the observation of cytochrome oxidase absorption bands in the presence of physiological concentrations of oxy and deoxy-hemoglobin is fraught with controversy, to the point that no reliable spectroscopic distinction of the copper component of cytochrome oxidase and the overlapping spectra of oxy-hemoglobin has been obtained. In fact, nearly all studies find that the so-called cytochrome oxidase Cu changes track those of HbO_2 due to the similarity of their spectroscopic absorption bands in the near infrared (NIR) region, and the great predominance of hemoglobin absorption over that of cytochrome oxidase.

NADH AS AN OXIMETER

It has been shown that the fluorescence of NADH and flavoprotein, when used in a ratiometric manner, can exhibit a reasonable immunity to changes of hemoglobin concentration, and has afforded standards for the independent changes of hemoglobin and NADH in transient hypoxia and re-oxidation. Such independence of the measures of cytochrome and hemoglobin has never been clearly demonstrated for the copper com-ponent of cytochrome oxidase, due to the overlapping spectra of oxyhemoglobin.

An alternate approach was based upon the fluor-escence of the newly discovered NADH and flavoprotein components of the mitochondrial respiratory chain, particularly by using the ratio of these two fluoro-chromes. Fortunately, NADH fluoresced in the reduced state while flavoprotein fluoresced in the oxidized state, so their ratio was a measure of the redox state of mitochondria, a signal that was found to be only marginally affected by changes in the oxygenation of hemoglobin in model systems. Using this criterion on the

very strong signal of NADH only, it was possible to show that the fluorescence of NADH was unchanged until the oxy/deoxy transition of hemoglobin was almost complete in a system in which functional activity was measured by the photoaction potential of an animal model, removing any doubt about the higher oxygen affinity of cytochrome *in vivo* as compared to hemoglobin (~ 20 Torr). This was a very important milestone for physiologists, who know that the deoxygenation of hemoglobin is very high and the critical pO_2 of mitochondrial function is compromised. Thus, the calibration of tissue oximeters in the region of intravenous saturation of hemoglobin (i.e., 20–30%) must be precisely measured to indicate critical tissue hypoxia *in vivo*. This indeed was subsequently validated by measurements of tissue energetics through phosphorus nuclear magnetic resonance (^{31}P NMR), particularly by measurements of the phosphocreatine:phosphate ratio.

NIR Spectroscopy

Jobsis used Kramer's technique to measure in the infrared, and developed a technology for the measurement of absorption of the copper component of cytochrome oxidase in the region of 830 nm based upon studies of the cat model and the heads of neonates, which he termed "transcranial spectroscopy." He further developed algorithms based upon fluorocarbon-perfused cat brain to give optical pathlengths that were believed to be transferable to the neonate brain and allow a deconvolution of blood volume and saturation changes from those of cytochrome oxidase signals using the full-length light algorithm. Delpy and co-workers avidly followed the lead of Jobsis and developed a close correlation between the decreased concentration of oxyhemoglobin and the so-called copper signal in a number of models, suggesting that the mitochondria in tissues contained a low-affinity cytochrome oxidase and responded to pO_2 (in a slice) in a way similar to that of hemoglobin. However, isolation of rat brain mitochondria failed to support this contention. Furthermore, the freeze-trapped hypoxic brain failed to show the absorption band of reduced cytochrome *c* in mild hypoxic stress that caused deoxygenated hemoglobin. In fact, the absorption band of reduced cytochrome *c* was not observed until the band of oxyhemoglobin was no longer detectable, according to the work of Bashford.

While attempts were made to detect the copper absorption band of hemoglobin in the NIR region, animal studies showed that the fluorescence of NADH and the flavoprotein could be used to detect anoxia in the presence of hemoglobin, particularly when the ratio of the fluorescent oxidized flavoprotein and the fluorescent reduced NADH were employed; this value was relatively insensitive to the hemoglobin concentration.

In fact, further demonstrations showed that the fluorescence of NADH was unaffected by the deoxygenation of hemoglobin in animal model brain. The NADH fluorescence increased in hypoxia only when the hemoglobin was already almost completely deoxygenated. This observation suggests that measurements of the critical pO_2 in hypoxia require accurate measurement of extreme values of hemoglobin desaturation, at the critical pO_2 for mitochondrial function.

NIR Spectroscopy of Brain and the BOLD Effect Measured by MRI

Much interest in the NIR method is based upon Ogawa's finding that changes in deoxyhemoglobin concentration (changes of the paramagnetic species of deoxyhemoglobin) enhanced water relaxation in the brain. This opened up the field of study of the activation phenomenon in the human brain, in which changes in deoxyhemoglobin levels are measured by NMR and by NIR tissue spectroscopy. The MRI changes are precisely imaged, while the NIR images, although crude, are measures of the rapidity of the changes. But in addition to incremental changes of deoxyhemoglobin, NIR could measure the saturation value of hemoglobin, which, for reasons involving Beer's law, originates mainly from the arteriolar/capillary/venolar bed. This feature, namely, the value of local oxygen extraction due to incremental changes of mitochondrial functional activity (i.e., localized activation), is not measured by MRI. The two techniques are now widely accepted as indicative of localized brain activation and have afforded the basis for in-depth studies of visual and sensory motor function. But, most importantly, NIR gives an excellent rendition of prefrontal cortex (PFC) signals without the difficulty of the large water content of the ocular system encountered with NMR (Figure 5). The use of activation images is appropriate to the NIR system, where baseline values may be somewhat variable and difficult to calibrate. The incremental changes of blood volume measured as changes of total hemoglobin, together with the aforementioned oxygen extraction measure, i.e., desaturation of hemoglobin, can be directly related to local metabolic activity, opening up a new field of NIR study of the semi-quantitative nature of the hemoglobin signals.

NIR Imaging

While the above-mentioned studies used dual-wavelength technology stemming from that of Glenn Millikan, a completely new concept was introduced by the discovery that photon migration through tissues can
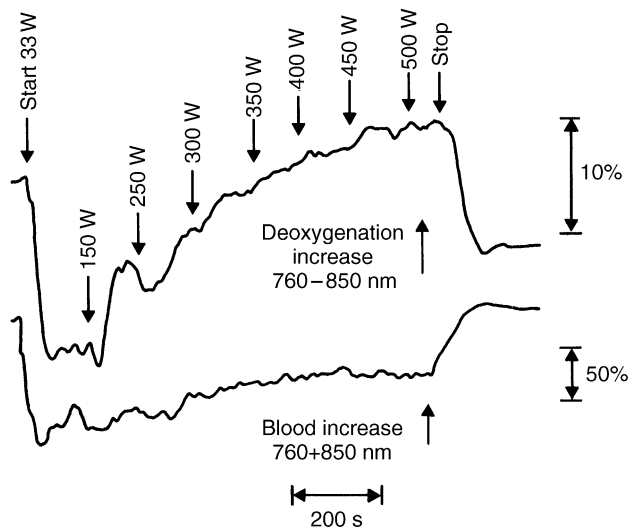
FIGURE 5 Evidence of activation of metabolism measured in the NIR (near infrared) by the dual wavelength method of 760–850 nm (equibestic wavelengths) which show progressive deoxygenation of the quadriceps muscle of a trained athlete during bicycle exercise up to the remarkable level 500 watts, at which nearly complete deoxygenation of hemoglobin can occur. As measured separately by the sum of signals at these two wavelengths. With appropriate coefficient, the increase of blood flow in the muscle also occurs giving a larger absorbance signal. At this level of exercise, it is probable that myoglobin is not deoxygenated as indicated by separate experiments with animal models.

be modeled by the diffusion equation. Multiple sources and detectors give very reasonable two- and three-dimensional imaging, and the propagation of light through tissue can be quantified by pulse time and phase and amplitude measurements, much as has been the case with measurement of fluorescence.

PHOTON MIGRATION IN TISSUES

The discovery that photons migrating through tissue followed the diffusion equation and that the tracks could be simulated by Monte Carlo methods, together with the adaptation of time-correlated single photon counting (TCSCP) to the task of measuring propagation times in tissues, opened up an entire field of NIR spectroscopy and imaging. This grew to be a field of medical science in a manner similar to NMR, but differing in the fact that the necessary equipment is not nearly as expensive, so the proliferation of the technique in research laboratories could be much more rapid. Because the profit margins did not match those of NMR, commercial production of NIR imagers has been restricted to two or three companies.

TISSUE OPTICAL PROPERTIES

Three techniques are outstanding in the measurement of tissue optical properties. The first, and still the foremost, is the pulse time method, in which photon delay is

caused by scattering and photon attenuation is caused by absorbers such as hemoglobin, water, and lipid. Because TRS (time-resolved spectroscopy) immediately deconvolutes scattering and absorption by time domain analysis, it is a preferred method. Similarly variable frequency modulated light will unravel by Fourier transformation exactly the same quantities as those obtained by TRS. However, the difficulty of stabilizing the phase shifts of electronic systems of variable frequency, together with the limitations of detection response to high-frequency radio waves, has limited this system to approximately 400 MHz. Nevertheless, many instruments have been made in the frequency range of 50 to 200 MHz that are used for quantifying absorption and scattering in multi-wavelength systems and that are capable of measuring hemoglobin saturation with significant precision. Such systems have adopted some cell phone components and are therefore compact and cheap. The most reliable and most used system modulates the light at very low frequencies in either, or time shares the light sources in a multiplex system, and appears to be the preferred system for many applications. The deconvolution of scattering and absorption can be obtained if sufficient data are taken at various source detector separations and optical wavelengths to include the scattering variations.

CANCER DETECTION

In cancer detection, scanning the breast for example, an activation signal based upon angiogenesis and hypermetabolism is given, causing more blood volume and more deoxygenated hemoglobin to be present. While this criterion may not be applicable to all cancers, it has given remarkably good scores in one breast cancer study.

MUSCLE STUDIES

A series of studies has been based upon activation measurements in which (e.g., in both muscle and brain), functional activation causes hemodynamic changes due to mobilization of blood flow and saturation changes due to varying metabolic activity (Figure 5). This has been used in muscle not only to evaluate exercise capability but also to quantify disability due to occlusive disease in the limbs.

BRAIN FUNCTIONAL ACTIVATION

In brain studies, whereas NMR and the optical method both measure hemodynamic activation due to functional activity, the NIR method is unique in that it measures changes of hemoglobin saturation caused by varying degrees of oxygen extraction from the capillary blood vessels of the brain. Thus, the convenience and economy of the optical method lends itself to studies of minimally
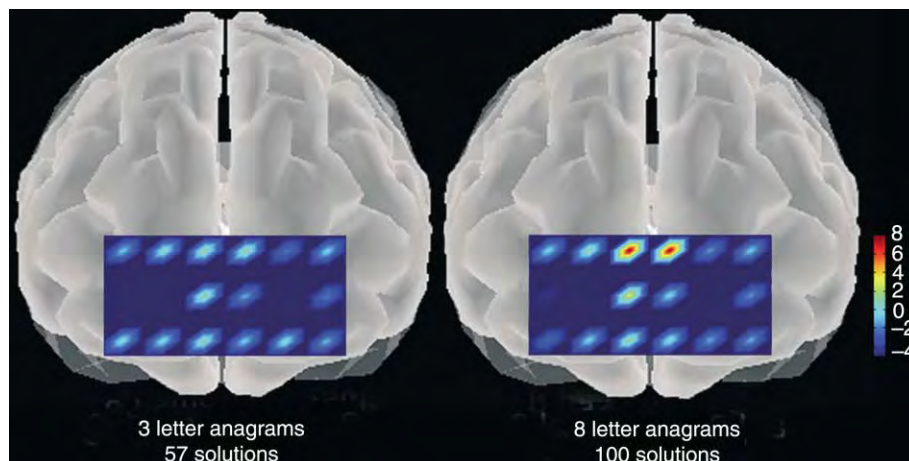
FIGURE 6 Illustrating the use of the NIR CW (continuous wave) dual wavelength system in localizing the particular voxels in the human forebrain at which brain functional activity is indicated to be increased by the oxygenation changes with respect to the baseline level of 3 letter anagrams. Each colored symbol represents the average of over 100 tests of a particular individual giving an average value of $8\,\mu M$ oxygenation under this condition of maximal stress. The region observed is the Broadman's 9 and 10 or ear to ear and hairline to eyebrow region of the projection of the prefrontal cortex. This separation was 4 cm in order to ensure signaling from the prefrontal cortex to accentuate the absorbance changes due to the prefrontal cortex and minimize those which might be associated with tissue layers at smaller depths.

perturbed human subjects, be they adults or neonates (Figure 6).

BRAIN STUDIES: THE NEURONAL SIGNAL

The early studies of David Hill, Tasaki, and Richard Keynes on transparency changes of axons upon stimulation suggest that a functional optical signal could be obtained in animal and human brain. While the studies of Salzberg did indeed verify the scattering changes in isolated preparations, studies using NIR to seek similar changes in the human brain suggest that only very small and somewhat irreproducible signals can be obtained. For example, the early experiments of Gabriele Gratton have not been duplicated either by himself or by others working in the field (Villringer and Franceschini, and Wolf), in part due to instrumental difficulties, and in part due to the very small size of the signal. The rise time of the signals appears to be in the range of 100 ms and the amplitude as small as one part in 10^4 or even smaller. However, the relatively robust signals obtained from the hemodynamic and metabolic activations discussed previously show the feasibility of measurements of functional activity in the human brain, particularly in the prefrontal region, and open up a reliable and economical method for human brain studies that predict a vibrant future of the optical methods in human studies.

Summary

The story of the development of optics in biochemistry and biophysics does not end here. In fact, some might say this is just the beginning of the transferability of optical tomography and optical biopsy to small animals on the one hand and human beings on the other. Perhaps spurred by the interest in online methods for evaluating the growth and recession of cancers under the influence of appropriate drugs in small animals, transferability of these principles to human subjects is becoming important. The optical method is taking its place along with MRI, PET, ultrasound, CT, and X-ray mammography as methods for studying pathologies in the human body.

SEE ALSO THE FOLLOWING ARTICLES

Cytochrome bc_1 Complex (Respiratory Chain Complex III) • Cytochrome c

GLOSSARY

Brown converter   An electrical chopper developed by the Leeds and Northrup Company that had very little contact potential variation.

cytochrome oxidase   The terminal enzyme of most oxygen-using systems.

dual-wavelength spectrophotometer   A spectrophotometer with time shared to adjacent wavelengths in order to minimize the effect of scattering changes upon absorbance changes, because scattering varies very slowly with wavelength, while cytochrome absorption varies sensitivity with wavelength.

DuBridge electrometer (DU)   A remarkable use of the suppressor grid of the pentode to control electron flow with very high input impedance, affording the basis of the Beckman pH meter and spectrophotometer.

electron tunneling   The transfer of electrons between two proteins at a distance without collision of their active sites.

EXAFS X-ray absorption spectroscopy used to obtain high resolution structures of Fe and Cu enzymes.

flavoprotein A widely spread pigment, but in this article the prosthetic group of ketoglutarate and pyruvate dehydrogenase.

mercury arc A very useful light source for biological studies, which gives light at exactly the correct wavelength for hemoglobin and in some cases cytochrome studies.

Molteno Institute A famous institute directed by David Keilin through the 30s, 40s and 50s. It became a Mecca for those working with cell respiration.

near infrared (NIR) imaging The use of the wavelengths in the red region just at the verge of invisibility between 700 and 900 nm to better penetrate tissue.

photochemical action spectrum The effect of light upon a biological system often used to activate a carbon monoxide inhibited cytochrome.

photomultiplier A highly sensitive light detector used in many spectrophotometers.

photon migration The phenomenon of photon diffusion through tissues used in great detail recently to image subsurface objects.

ruby laser One of the early forms of the laser, emitting in the red region.

scattered light Light that does not proceed directly through tissue.

spectrophotometer A device that measures the absorbance of materials as a function of wavelength or, in some cases, energy.

time resolve spectroscopy (TRS) Spectroscopy using sharp pulses of light to distinguish the scattering from the absorption of tissues.

X-ray absorption spectroscopy (EXAFS) Spectroscopy used to obtain high-resolution structures of Fe and Cu enzyme.

FURTHER READING

Carafoli, E. (2003). Historical review: Mitochrondria and calcium: Ups and downs of an unusual relationship. *Trends Biochem. Sci.* **28**, 175–181.

Chance, B. (ed.) (1989). *Photon Migration in Tissues*. Plenum Press, New York, NY.

Chance, B. (1991). The optical method. *Ann. Rev. Biophys. Biophys. Chem.* **20**, 1–18.

Keilin, D. (1966). *The History of Cell Respiration and Cytochrome*. Cambridge University Press, Cambridge, UK.

Slater, E. (ed.) (1966). *Flavins and Flavoproteins*. Elsevier, Amsterdam.

BIOGRAPHY

Britton Chance is Eldridge Reeves Johnson University Professor Emeritus of Biochemistry and Biophysics and Physical Chemistry and Radiologic Physics at the University of Pennsylvania School of Medicine. His current research interests focus on optical spectroscopic methods for the study of brain cognition, breast cancer detection, and physiological function. He has received numerous honors and awards and honorary Ph.D. and M.D. and is noted for his contributions in basic science research and technological development.

Sphingolipid Biosynthesis

Martina Leipelt and Alfred H. Merrill, Jr.
Georgia Institute of Technology, Atlanta, Georgia, USA

Sphingolipids are a complex family of compounds that perform diverse structural and regulatory functions for eukaryotes and some prokaryotes and viruses. They share a common structural feature, a sphingoid base backbone that is synthesized *de novo* from serine and a fatty acyl-coenzyme A, then converted into ceramides, phosphosphingolipids, glycosphingolipids, and other species, including protein adducts. Several diseases result from disruption of *de novo* sphingolipid biosynthesis by environmental factors or hereditary defects, but modulation of sphingolipid biosynthesis is also being explored as a means to control other diseases, including sphingolipid storage diseases and cancer.

Structures and Nomenclature

Sphingolipids can be divided into several major categories: the sphingoid bases and their simple derivatives, ceramides, and more complex sphingolipids (Figure 1). The International Union of Pure and Applied Chemists (IUPAC) has recommended a systematic nomenclature for sphingolipids. The root name "sphingosin," in reference to the sphinx, was given by J. L. W. Thudichum in 1884 "in commemoration of the many enigmas which it presented to the inquirer."

SPHINGOID BASES

The structure of sphingosine, the major sphingoid base of mammals, is (2S, 3R, 4E)-2-aminooctadec-4-ene-1, 3-diol (it is also called D-*erythro*-sphingosine and E-sphing-4-enine) (Figure 1). This is only one of many sphingoid bases found in nature, which vary in alkyl chain length and branching, the number and positions of double bonds, the presence of additional hydroxyl groups, and other features. The structural variation has functional significance; for example, sphingoid bases in skin have additional hydroxyls at position 4 and/or 6 that can interact with neighboring molecules, thereby strengthening the permeability barrier of skin.

Sphingoid bases function as intra- and extracellular signals and second messengers in the form of free sphingoid bases, sphingoid base 1-phosphates (Figure 1), and possibly other species. Nonetheless, sphingoid bases are present in cells primarily as the backbones of more complex sphingolipids.

CERAMIDES

Ceramides are fatty acid derivatives of sphingoid bases (Figure 1). The fatty acids are typically saturated or mono-unsaturated with chain lengths from 14 to 26 carbon atoms (or even longer in the special case of skin), and sometimes have a hydroxyl group on the α- or ω-carbon atom. These structural features favor the segregation of ceramides and some complex sphingolipids into specialized regions of the membrane (called "rafts" and "caveolae") that participate in cell signaling, nutrient transport, and other functions.

Ceramides also serve as second messengers that regulate cell growth, senescence, and programmed cell death (apoptosis). Their biologic activity depends on the type of sphingoid base and fatty acid; for example, dihydroceramides (i.e., without the 4,5-double bond of the sphingosine backbone) (Figure 1) are less potent than ceramides as inducers of apoptosis, whereas phytoceramides (i.e., with 4-hydroxysphinganine or "phytosphingosine" as the backbone) are more potent.

MORE COMPLEX PHOSPHO- AND GLYCO-SPHINGOLIPIDS

The major phosphosphingolipids of mammals are sphingomyelins (ceramide phosphocholines) (Figure 1), whereas insects contain mainly ceramide phosphoethanolamines and fungi have phytoceramidephosphoinositols and inositol phosphates. Some aquatic organisms also contain sphingolipids in which the phosphate has been replaced by a phosphono- or arsenate group.

Glycosphingolipids are classified on the basis of carbohydrate composition: (1) neutral glycosphingolipids contain one or more uncharged sugars such as glucose (abbreviated Glc, hence, glucosylceramide is GlcCer), galactose (Gal), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), and fucose (Fuc); and (2) acidic glycosphingolipids contain ionized functional groups (phosphate or sulfate) attached to neutral sugars, or charged sugar residues such as sialic acid (N-acetylneuraminic acid). The latter are called gangliosides, and the number of sialic acid residues is usually denoted with a subscript letter (i.e., mono-, di- or tri-) plus a number reflecting the subspecies within that
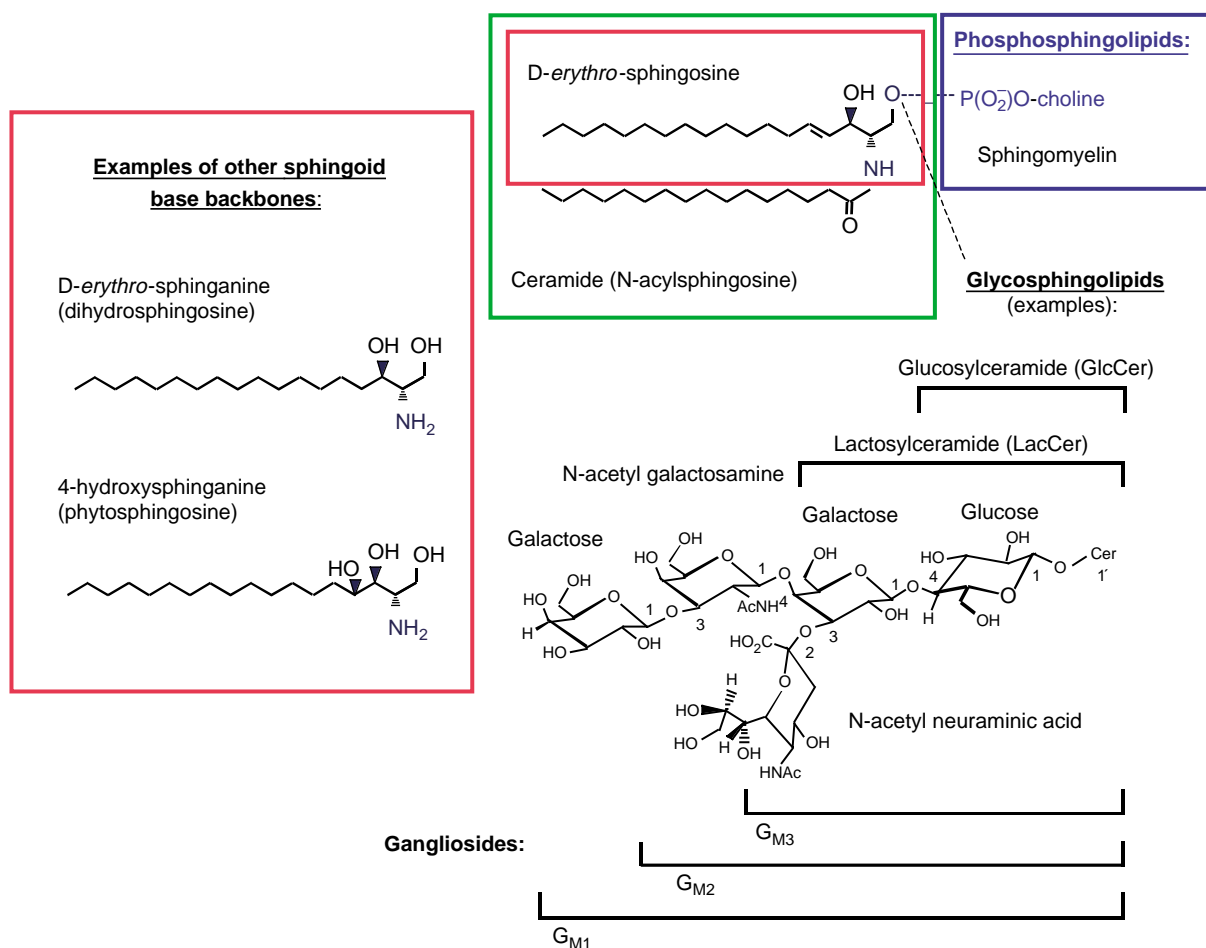
FIGURE 1 Structures of representative sphingolipids. Shown are several examples of sphingoid bases (sphingosine, sphinganine and 4-hydroxysphinganine, boxed in red), ceramide (in green), sphingomyelin (blue), and neutral (GlcCer and LacCer) and acidic (gangliosides G_{M1}, G_{M2}, and G_{M3}) glycosphingolipids.

category (see examples in Figure 1). For a few glycosphingolipids, historically assigned names as antigens and blood group structures are still in common usage (e.g., Lexis x and sialyl Lewis x).

PROTEIN ADDUCTS

Some sphingolipids are covalently attached to protein, e.g., ω-hydroxy-ceramides and -GlcCers are attached to surface proteins of skin and inositolphosphoceramides are used as membrane anchors for some fungal proteins, in a manner somewhat analogous to the glycosylphosphatidylinositol (GPI) anchors that are attached to proteins in other eukaryotes.

De novo Synthesis of the Ceramide Backbone

Sphingolipid biosynthesis is widespread among eukaryotic cells, and it appears that new synthesis (i.e., *de novo*) is relied upon more than reutilization of sphingolipids from exogenous sources, such as food. The biosynthetic

pathway for such a diverse family of compounds (conservatively estimated to be in the tens of thousands) is obviously complex; however, its fundamental features can be summarized in Figures 2 and 3.

SERINE PALMITOYLTRANSFERASE

Serine palmitoyltransferase (SPT) catalyzes the initial step of the pathway which, for many organisms, is the condensation of serine and palmitoyl-CoA to form 3-ketosphinganine (Figure 2). However, for organisms that produce sphingoid bases with other alkyl chain lengths (such as the C14 species of insects), the first enzyme of the pathway utilizes a different cosubstrate (dodecanoyl-CoA, in this example) and could be renamed "serine dodecanoyltransferase."

SPT is a pyridoxal 5′ phosphate-dependent enzyme comprised of two gene products (termed SPTLC1 and SPTLC2 for humans, and LCB1 and LCB2 for yeast); a third has also been identified in yeast, but does not appear to have a homologue in mammals. In most organisms, SPT is associated mainly with the
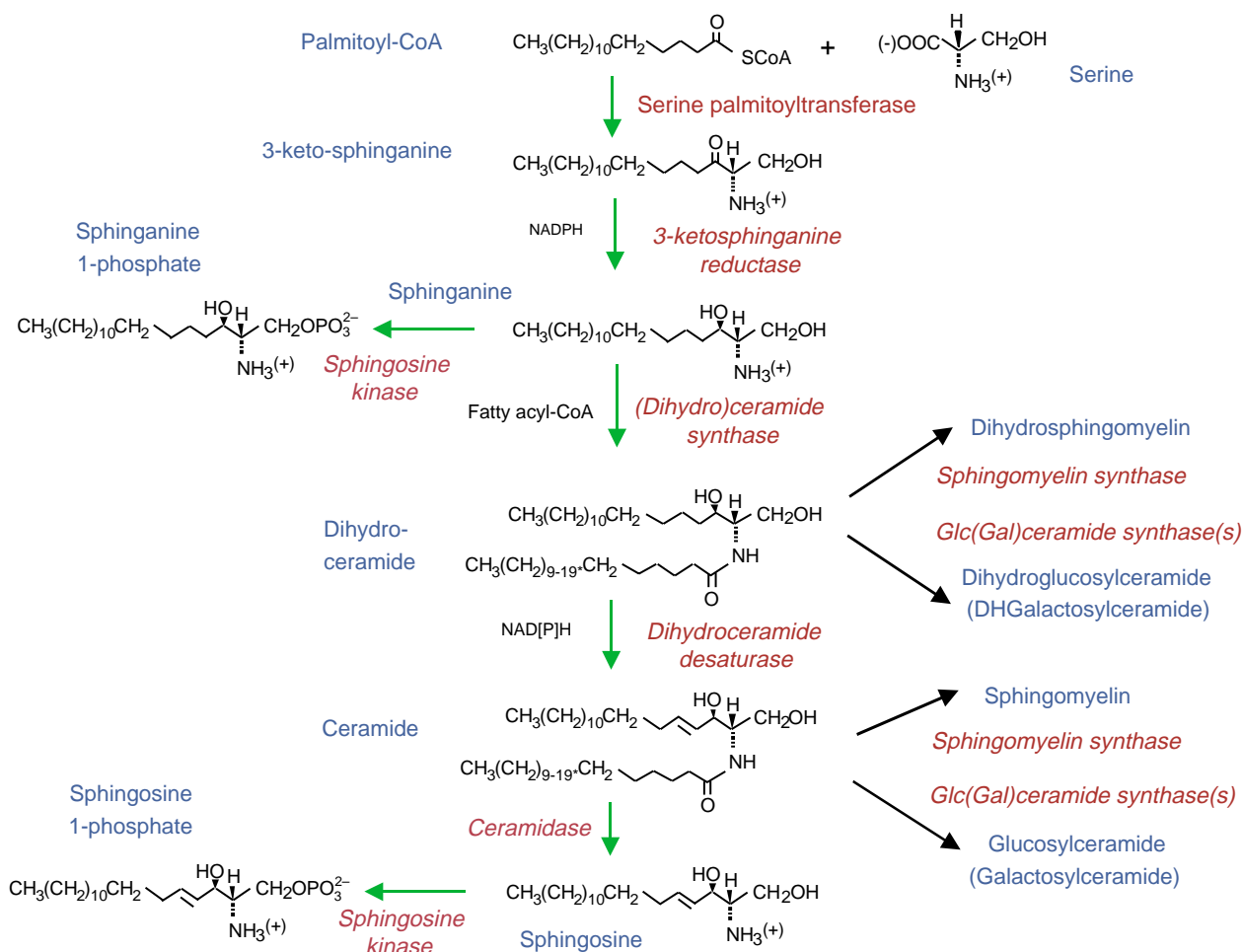
FIGURE 2 The *de novo* biosynthetic pathway for sphingoid bases, ceramide, (dihydro)sphingomyelins and (dihydro)glucosylceramides. The color coding distinguishes the enzyme names (in red) and the metabolites (in blue).

endoplasmic reticulum, as are the other enzymes of ceramide biosynthesis. SPT activity is affected by a wide range of factors: sphingosine 1-phosphate, endotoxin and cytokines, heat shock, UVB irradiation, cytotoxic drugs (including many cancer chemotherapeutic drugs), retinoic acid, and a number of small molecule inhibitors produced by microorganisms (one of which, ISP1 or myriocin, is often used to block *de novo* sphingolipid synthesis by cells in culture). The mechanisms of SPT regulation are not fully understood, but include (for example) both acute modulation by heat shock and increased expression of SPT mRNA by cytokines. Mutations in human SPTLC1 cause hereditary sensory neuropathy type I (HSN1), which is the most common hereditary disorder of peripheral sensory neurons.

CERAMIDE SYNTHASE

3-Ketosphinganine is rapidly converted to sphinganine by an NADPH-dependent reductase, then ceramide syntase(s) acylate sphinganine to dihydroceramides using fatty acyl-CoA's varying in length from C16 to C30 (and usually saturated or mono-unsaturated) (Figure 2). Ceramide synthase is actually a family of enzymes, each of which appears to arise from a different gene and to utilize a particular subset of fatty acyl-CoA's (e.g., TRH4 utilizes palmitoyl-CoA whereas UOG1 uses stearoyl-CoA).

Ceramide synthase is activated by a number of stimuli, including cancer chemotherapeutic drugs and irradiation, and the increased production of ceramide is thought to mediate the toxicity of these treatments. Ceramide synthase is also the target of a number of mycotoxins (fumonisins), which are produced by fungi that grow on corn and, when consumed, result in spectrum of diseases that are important to agriculture (equine leukoencephalomalacia and porcine pulmonary edema) as well as human cancer and possibly birthdefects.

DIHYDROCERAMIDE DESATURASE

Insertion of double bond(s) into the sphingoid base backbones occurs mainly after formation of dihydroceramide(s) (Figure 2). For mammals, introduction of the
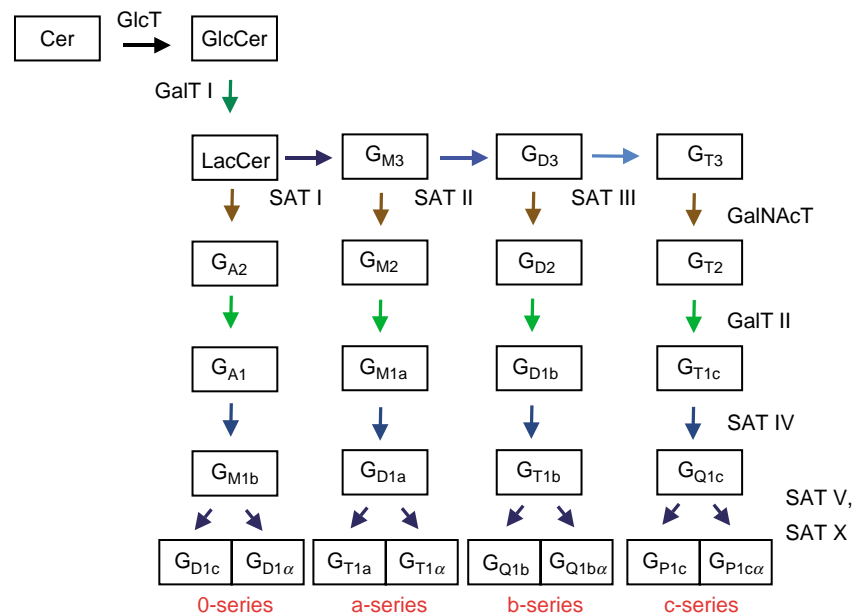
FIGURE 3 A representation of the combinatorial nature of glycosphingolipid biosynthesis. Shown are the reactions leading to the major ganglioside series and the enzymes involved. The abbreviations refer to ceramide (Cer), glucosylceramide (GlcCer), lactosylceramide (LacCer) and the different categories of gangliosides designated by "G" and subscripts for the number of sialic acids (M,D,T and Q representing 1,2,3 and 4, respectively) and other structural features. Abbreviations: GalNAcT, N-acetylgalactosaminetransferase; GalT, galactosyltransferase; GlcT, glucosyltransferase; and, SAT (sialyltransferase) with the Roman numerals reflecting the subtypes. (Modified from Kolter, T., Proia, R. L., and Sandhoff, K. (2002). Combinatorial ganglioside biosynthesis. *J. Biol. Chem.* 277, 25859–25862.)

4,5-double bond is catalyzed by two pyridine nucleotide-dependent desaturases (DES1 and DES2), one of which may also be responsible for addition of the 4-hydroxyl-group of phytoceramides.

Synthesis of More Complex Sphingolipids

Ceramides in their various forms (i.e., ceramides, dihydroceramides, phytoceramides, etc.) are at a key branch point of complex sphingolipid biosynthesis where these intermediates are partitioned into either phosphosphingolipids or glycosphingolipids. For cells that produce more than one category of glycolipid (for example, mammalian epithelial cells, which have both GlcCer and GalCer), the glycolipid arm can have multiple branches. The fate of a given intermediate is governed by the relative activities and selectivity of the enzymes at this branch point as well as by the subcellular localization of the participants.

SPHINGOMYELIN AND OTHER PHOSPHOSPHINGOLIPIDS

Sphingomyelins are synthesized by transfer of phosphorylcholine from phosphatidylcholine to ceramides (Figure 2). This reversible reaction links glycerolipid and sphingolipid metabolism and signaling, because

ceramides and diacylglycerols both function as metabolic intermediates and as intracellular second messengers. This may explain why cells produce dihydroceramides as the initial products of *de novo* sphingolipid biosynthesis since that allows a relatively innocuous intermediate to accumulate if later steps in the pathway slow.

Relatively little is known about the biochemistry of sphingomyelin synthase, including whether the activities in the Golgi apparatus and plasma membranes represent a single enzyme, or several different enzymes (two mammalian sphingomyelin synthase genes have been identified, *SMS1* and *SMS2*). The regulation of sphingomyelin biosynthesis is also intriguing – with changes in development, neoplasia, and other normal and abnormal cell states.

Ceramide phosphorylethanolamines are synthesized from phosphatidylethanolamine and ceramides in a reaction analogous to sphingomyelin synthase (i.e., transesterification with phosphatidylethanolamine), and once formed can be methylated to sphingomyelins in some species. Inositolphosphoceramides are also formed by transesterification (from phosphatidylinositols).

GLYCOSPHINGOLIPIDS

A pathway that is responsible for the biosynthesis of hundreds (to thousands) of different glycosphingolipids is obviously complex, but these compounds are nonetheless produced using surprisingly few

glycosyltransferases. Efficiency is achieved by a "combinatorial" biosynthetic pathway that directs precursors and intermediates toward the desired products by modulating the activities of key combinations of enzymes (see Figure 3 for an illustration).

The addition of the carbohydrate headgroups is catalyzed by glycosyltransferases that transfer a specific sugar from the appropriate sugar nucleotide (e.g., UDP-Glc, UDP-Gal, etc.) to ceramide or the nonreducing end of the growing carbohydrate chain attached to ceramide. GlcCer and GalCer are synthesized by UDP-Glc(or Gal):ceramide glucosyltransferases, hence, a major determinate of the types of glycosphingolipids made by a given cell type will be whether it expresses one or both of these genes. Factors that regulate these enzymes include cell type, the nature of the ceramide substrate (ceramides with α-hydroxy fatty acids are mainly utilized for GalCer synthesis), and exposure of the cells to agonists such as endotoxin and acute phase response mediators. A number of inhibitors of these glycosyltransferases are being tested for efficacy in sphingolipid storage diseases (caused by inherited defects in glycosphingolipid hydrolyases), based on the rationale that slowing biosynthesis may counterbalance these defects.

Additional glycosyltransferases are responsible for subsequent addition of sugars to make dihexosylceramides, trihexosylceramides, etc. as well as for addition of neutral sugars to gangliosides (Figure 3). Likewise, gangliosides are synthesized by the stepwise transfer of neutral sugars and sialic acids. In general, the enzymes responsible for these reactions are located in the lumen of the Golgi apparatus, and the region corresponds to the order in which the sugars are added. For example, the sialyltransferase catalyzing the synthesis of a simple ganglioside (ganglioside G_{M3}) is in the *cis*-Golgi, whereas enzymes involved in terminal steps of more complex gangliosides are located in the more distal *trans*-Golgi network.

Regulation of complex glycosphingolipid biosynthesis involves both transcriptional and posttranscriptional factors. For example, developmentally regulated, tissue selective variations in ganglioside amounts and types in mammalian tissues are under transcriptional control, but the activities of glycosyltransferases can be fine tuned by posttranslational modification.

The biosynthesis of sulfatides (i.e., sulfated glycosphingolipids such as 3′-sulfo-GalCer) is catalyzed by sulfotransferases (in this example: 3′-phosphoadenylylsulfate:GalCer 3′-sulfotransferase), which utilize the activated sulfate donor 3′-phosphoadenosine-5′-phosphosulfate.

OTHER SPECIES

Although once thought to be only intermediates of sphingolipid turnover, lysosphingolipids such as sphingosine 1-phosphate and sphingosylphosphocholine (lysosphingomyelin) are now known to be synthesized as important signaling molecules. Sphingosine 1-phosphate formation requires the release of sphingosine from ceramide (note that sphingosine is not a direct intermediate of *de novo* sphingolipid biosynthesis but first appears in ceramide) by ceramidase(s) followed by transfer of phosphate from ATP by sphingosine kinase(s) (Figure 2). Less is known about the origin of sphingosylphosphocholine, although it is plausible that this could be made by a phospholipase A_2-type cleavage of sphingomyelin, the transfer of phosphocholine to sphingosine, or both.

Sphingolipidomics

The large number and structural complexity of sphingolipids has made quantitative analysis of all of the molecular species technically difficult, and heretofore impossible with small samples such as cells in culture. However, it is now feasible to map the sphingolipid "metabolome" due to the relatively recent availability of tandem mass spectrometers of multiple configurations (e.g., tandem quadrupole, time-of-flight, and ion traps as well as hybrids of these technologies) and modes of ionization (such as electrospray and matrix-assisted laser-desorption ionization (MALDI)), especially when combined with high-performance liquid chromatography. When complemented by the tools of genomics and proteomics, the new field of "sphingolipidomics" will finally be able to answer the many riddles of how these molecules are made, and for what functions.

SEE ALSO THE FOLLOWING ARTICLES

Glycolipid-Dependent Adhesion Processes • Lipid Bilayer Structure • Lysophospholipid Receptors • Protein Palmitoylation • Sphingolipid Catabolism

GLOSSARY

ceramide An N-acyl-derivative of sphingosine that is both a metabolic intermediate and a cell signaling molecule. In some cases, the term is applied generically to any N-acyl-sphingoid base.

glycosphingolipid A compound with a carbohydrate bound to a sphingoid base (and most often, attached to position 1 of an N-acyl-sphingoid base).

glycosyltransferase An enzyme that transfers a carbohydrate from a donor (usually a UDP-sugar) to an acceptor which, in the case of sphingolipids, is either ceramide or a carbohydrate chain attached to ceramide.

phosphosphingolipid A compound with a phosphate or phosphodiester linked headgroup attached to a sphingoid base (or more often, to position 1 of an N-acyl-sphingoid base).

sphingoid base The backbone of more complex sphingolipids as well as a cell signaling molecule. Structurally, a long-chain alkane (or alkene) with an amino at position 2, and (usually) hydroxyl groups at positon 1 and 3 plus various alkyl chain lengths, degrees of unsaturation, and additional hydroxyl groups.

sphingosine 1-phosphate A bioactive metabolite that serves as an intracellular and an extracellular signal as well as an intermediate of sphingoid base catabolism.

FURTHER READING

Hannun, Y. A., and Obeid, L. M. (2002). The ceramide-centric universe of lipid-mediated cell regulation: Stress encounters of the lipid kind. *J. Biol. Chem.* **277**, 25847–25850.

IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) (1998). Nomenclature of glycolipids. Recommendations 1997. *Eur. J. Biochem.* **257**, 293–298.

Kolter, T., Proia, R. L., and Sandhoff, K. (2002). Combinatorial ganglioside biosynthesis. *J. Biol. Chem.* **277**, 25859–25862.

Merrill, A. H. Jr., (2002). *De novo* sphingolipid biosynthesis. A necessary, but dangerous, pathway. *J. Biol. Chem.* **277**, 25843–25846.

Merrill, A. H., Jr., and Sandhoff, K. (2002). Sphingolipids: Metabolism and cell signaling. In *New Comprehensive Biochemistry: Biochemistry of Lipids, Lipoproteins, and Membranes* (D. E. Vance and J. E. Vance, eds.) Chapter 14, Elsevier, Amsterdam.

Spiegel, S., and Milstien, S. (2003). Sphingosine-1-phosphate: An enigmatic signalling lipid. *Nat. Rev. Mol. Cell Biol.* **4**, 397–407.

Sullards, M. C., Wang, E., Peng, Q., and Merrill, A. H., Jr. (2003). Metabolomic profiling of sphingolipids in human glioma cell lines by liquid chromatography tandem mass spectrometry. *Cell Mol. Biol.* **49**, 789–797.

Thudichum, J. L. W. (1884). *A Treatise on the Chemical Constitution of Brain*. Bailliere, Tindall, and Cox, London.

BIOGRAPHY

Martina Leipelt holds a Doctorate in Genetics from the University of Hamburg, Germany. Her doctoral research with Professor E. Heinz provided a systematic functional analysis of the glucosylceramide synthase gene family with representatives of plants (*Gossypium arboreum*), animals (*Caenorhabditis elegans*), and fungi (*Magnaporthe grisea*, *Candida albicans*, and *Pichia pastoris*). She is currently a postdoctoral fellow with Dr. Merrill.

Al Merrill is the Smithgall Institute Chair in Molecular Cell Biology in the School of Biology and the Petit Institute for Bioengineering and Biosciences at Georgia Institute of Technology. Dr. Merrill's laboratory, in collaboration with Dr. Ron Riley at the USDA, discovered the inhibition of ceramide synthase by fumonisins and identified the first diseases caused by disruption of *de novo* sphingolipid biosynthesis His current research deals with sphingolipidomics (http://www.lipidmaps.org).

Sphingolipid Catabolism

Akira Abe and James A. Shayman
University of Michigan, Ann Arbor, Michigan, USA

Sphingolipids are composed of a variety of membrane-associated molecules that contain a long-chain sphingoid base. The base may be acylated, glycosylated, and phosphorylated to produce a variety of structures with important biological functions. The catabolic pathways responsible for the degradation of sphingolipids have been extensively studied. The constitutive degradation of sphingolipids occurs in lysosomes through a series of degradative enzymes known as sphingolipid-specific hydrolases. The inherited deficiencies of sphingolipid hydrolases may result in metabolic disorders that lead to the abnormal accumulation of sphingolipids within cells. Recently, some sphingolipid metabolites have been assigned functions as extracellular and intracellular signaling molecules. In particular, certain sphingomyelin metabolites such as ceramide, sphingosine, and sphingosine-1-phosphate, may play important roles in cellular processes such as cell growth, differentiation, apoptosis, stress, and inflammation. The degradation of sphingolipids involved in the response to extracellular stimuli occurs through both the lysosomal and nonlysosomal catabolic pathways.

Sphingomyelin Catabolism

Sphingomyelinase catalyzes the hydrolysis of sphingomyelin to form phosphorylcholine and ceramide. Ceramide may be further metabolized to form other sphingolipids or may function as a signaling molecule. Sphingomyelinases are categorized into four groups: acid sphingomyelinase (aSMase), secreted sphingomyelinase (sSMase), neutral sphingomyelinase (nSMase) and alkaline sphingomyelinase (bSMase).

ACIDIC SPHINGOMYELINASE

aSMase is a well-characterized sphingomyelinase with an optimal pH of 5 and localized to lysosomes. The enzyme primarily functions in the degradation of sphingomyelin. In humans, a genetic deficiency of lysosomal aSMase results in Niemann-Pick types A and B, autosomal-recessive lipid storage disorders. The enzyme was purified from urine as a 72 kDa monomeric glycoprotein with a 61 kDa polypeptide core. The human aSMase cDNA encodes a 629 amino acid polypeptide. Metabolic labeling studies in cell-lines cells transfected with the human aSMase cDNA reveal that a 75 kDa aSMase precursor is processed by extensive posttranslational modification during sorting to the lysosome.

There is some evidence that aSMase plays an important role in ceramide formation after stimulation with TNFα or CD95, or treatment with UV-A irradiation. However, the requirement of aSMase in apoptosis and differentiation induced by TNFα and Fas ligand is still debated. The involvement of aSMase in ceramide-meditated signal transduction may depend on the cell and tissue type and vary with the stimulus.

SECRETED SPHINGOMYELINASE

sSMase is a secretory form of the aSMase gene product and is secreted via a Golgi-apparatus-dependent pathway into the extracellular space. The enzyme is activated in the presence of physiological concentrations of Zn^{2+} and demonstrates maximum activity under acidic conditions. However, the enzyme is able to hydrolyze sphingomyelin from atherogenic lipoproteins, LDL extracted from arteriosclerotic lesions, and oxidized at neutral pH. LDL treated with sSMase forms aggregates that are retained on extracellular matrix and stimulates macrophage foam cell formation. Thus, sSMase may serve a normal physiological function in lipoprotein metabolism and a pathological function in atherogenesis.

NEUTRAL SPHINGOMYELINASE (nSMASE)

In mammalian cells, nSMase is a membrane-associated protein with an optimal neutral pH. Although the enzyme is expressed ubiquitously in mammalian tissues, the highest activity is predominantly found in brain. Mammalian tissues express two isoforms of nSMase. The enzyme activity of the higher molecular weight isoform form is Mg^{2+} dependent but that of the lower molecular weight isoform is Mg^{2+} independent. The low-molecular weight isoforms appear to be degradation products of the high-molecular isoforms. The purified enzyme is activated by phosphatidylserine,

Mg^{+2}, and Mn^{2+}, and inhibited by Cu^{2+}, Zn^{2+}, Ca^{2+}, Cd^{2+}, Hg^{2+}, glutathione, and asialo-ganglioside, GM3.

Recently, an nSMase candidate (nSMase 1) with molecular mass of 47.5 kDa was cloned from mouse and human based on multiple sequence alignments of bacterial nSMases. Bacterial nSMases are soluble proteins with an optimal pH between 4.2 and 8.0. They are Mg^{2+} dependent. The product of the nSMase 1 gene is localized in endoplasmic reticulum, Golgi, and/or the nuclear matrix of cells. The natural substrate of nSMase 1 is still not known and the enzyme appears to not be involved in ceramide formation after stimulation by TNFα. A second nSMase (nSMase 2) with a molecular mass of 71 kDa has been cloned using an improved database search method combined with phylogenetic analysis. nSMase 2 is a brain-specific nSMase with a different domain structure and only marginal sequence similarity to other SMases. nSMase 2 has the basic properties of rat and bovine brain nSMase and is activated in response to TNFα. nSMase 2 colocalizes with a Golgi apparatus marker in a number of cell lines. One more nSMase candidate that is a different gene from nSMase 1 and 2 was recently cloned by use of expression cloning. This nSMase cDNA encodes a 397 amino acid polypeptide. The enzyme is activated by Mg^{+2}, inhibited by Cu^{2+} and glutathione, and recognizes sphingomyelin as a preferred substrate. The deduced amino acid sequence indicates that the enzyme is a membrane-integrated protein and has a significant homology to the death domains of the TNF-α and Fas/AP-1 receptors. The overexpression of this recombinant nSMase in human aortic smooth muscle cells results in apoptosis and augmented oxidized LDL-induced apoptosis. The limited amino acid sequence information from purified nSMase from bovine brain shows no sequence homology to the nSMases cloned to date.

The activation of SMase by extracellular stimuli that induce differentiation, apoptosis, and stress and inflammation is associated with both aSMase and nSMase activities. Because nSMases are localized in the plasma membrane and cytosol, nSMases and not aSMase would appear to be, on topological grounds, the logical participants in sphingomyelin signaling pathways. Recently, an adaptor protein, FAN (factor associated to nSMase activation), was shown to link the TNF receptor to nSMase and act upstream of nSMase. In addition, several recent reports suggest that nSMase in lipid rafts contributes to TNFα signaling.

ALKALINE SPHINGOMYELINASE (bSMASE)

The enzyme activity of this SMase was initially found in the small intestine. Recently, bSMase was purified from rat intestine and required bile salt for the enzyme activity. The purified bSMase is 58 kDa, has an alkaline pH optimum, is Mg^{2+} independent, and is not inhibited by glutathione. The expression of this enzyme is specific to the intestinal mucosa. Another type of bSMase has an 85 kDa molecular mass and is found in human bile. These enzymes appear to function in the catabolism of dietary sphingomyelin.

Ceramide Catabolism

Ceramidase (CDase) catalyzes the hydrolysis of ceramide to fatty acid and sphingosine. Three types of ceramidase have been described based on their pH optima for activity. These include acid ceramidase, neutral ceramidase, and alkaline ceramidase. Sphingosine and its phosphorylated metabolite, sphingosine-1-phosphate, act as potent inhibitors of protein kinase C and potent effectors of cell proliferation and differentiation. CDase may change the balance of ceramide, sphingosine, and sphingosine-1-phosphate within cells in response to various stimuli. CDase may therefore regulate sphingolipid mediated signaling events.

ACID CERAMIDASE

Acid ceramidase (aCDase) can be characterized as an N-acylsphingosine deacylase with an acidic pH optimum. The enzyme is localized in lysosomal and endosomal compartments. The enzyme functions primarily in the degradation of ceramide. In humans, a genetic deficiency of lysosomal aCDase results in the lyososomal lipid storage disorder known as Farber disease. The enzyme, purified from human urine, is a 55 kDa heterodimeric glycoprotein consisting of two disulfide-linked polypeptide chains of 13 kDa (α) and 40 kDa (β). The human aCDase cDNA encodes a protein of 395 amino acids. The 13 kDa (α) and 40 kDa (β) subunits are derived from a common 55 kDa precursor encoded by the full length of aCDase-cDNA. Only the β-subunit is posttranslationally glycosylated during transport to acidic cellular compartments. aCDase activity is enhanced by an activator protein known as saposin D.

aCDase responds to extracellular stimuli. In rat hepatocytes, aCDase activity is bimodally regulated by IL-1β and is activated by tyrosine phosphorylation. In renal mesangial cells, aCDase is activated by TNFα but inhibited by nitric oxide. Inhibition of aCDase sustains the accumulation of ceramide induced by TNFα. Overexpression of aCDase in L929 cells suppresses TNFα induced ceramide formation and cell death.

NEUTRAL CERAMIDASE

In mammals, nCDase is present in a variety of tissues and cell types and its activity is mainly found in membrane fractions. nCDase has been purified from rat

brain, where the highest activity is found. nCDase is a 90 kDa membrane-bound, nonlysosomal protein. The enzyme has a broad pH optimum in the neutral to alkaline range and does not require divalent cations for activity. The enzyme is stimulated by phosphatidic acid and phosphatidylserine. Similar membrane-bound nCDases, purified from mouse liver and rat kidney, are 94 and 112 kDa monomeric glycoproteins, respectively. Partial amino acid sequences of each nCDase were used to clone the genes from mouse, rat, and human. The mouse, rat, and human nCDase cDNAs encode 761, 756, and 763 amino acid polypeptides, respectively. In rat kidney, nCDase is mainly localized to the apical membrane of the proximal tubules, distal tubules, and collecting duct. By contrast, liver nCDases are detected in endosome-like organelles within the hepatocytes. Human nCDase over expressed in HEK293 and MCF7 cells is found in mitochondria. nCDase activity is activated by IL-1β in rat hepatocytes, resulting in a decrease of ceramide concomitant with an increase of sphingosine. TNFα stimulates and nitric oxide inhibits nCDase activity in renal mesangial cells.

ALKALINE CERAMIDASE

Alkaline ceramidase activity has been described in human, rat, and mouse tissues. Two forms of membrane-bound alkaline ceramidases (bCDases) with molecular masses of 60 kDa (CDase-I) and 148 kDa (CDase-II) are present in skin. CDase-I and CDase-II have alkaline pH optima. Both enzyme activities are inhibited by sphingosine. Based on sequence homology to the yeast bCDase, a novel human bCDase was cloned. This bCDase cDNA encodes a protein of 253 amino acids. The enzyme has a pH optimum of 9.5 and is activated by Ca^{2+}, but is inhibited by Zn^{2+} and sphingosine. The enzyme hydrolyzes phytoceramide preferentially and when over expressed in COS-1 cells is localized to both the endoplasmic reticulum and Golgi apparatus.

Sphingoid Base Catabolism

At the penultimate step of sphingolipid catabolism the primary hydroxyl group of the sphingoid base is phosphorylated by sphingosine kinase. The phosphorylated sphingoid base is then cleaved between the vicinal carbons with amino and hydroxyl groups by sphingosine-1-phosphate lyase (SPLase). The predominant sphingoid base in mammalian cells is sphingosine. Sphingosine produced from ceramide by ceramidase is phosphorylated by sphingosine kinase and then degraded to form phosphoethanolamine and hexadecanal by SPLase. The sphingosine-1-phosphate, generated during sphingosine catabolism, is not only a catabolic intermediate but also functions as an extracellular and intracellular messenger. Sphingosine-1-phosphate is mitogenic. The extracellular actions of sphingosine-1-phosphate are mediated through the EDG receptor family of G protein-coupled receptors.

SPHINGOSINE-1-PHOSPHATE LYASE

SPLase is a pyridoxal phosphate-dependent member of a class of enzymes known as aldehyde lyases and cleaves 1-phosphorylated sphingoid bases into aliphatic fatty aldehydes and phosphoethanolamine with a neutral pH optimum. SPLase is a ubiquitous enzyme present in multiple species and mammalian tissues. However, platelets lack SPLase activity, suggesting that platelets may be the primary source of circulating sphingosine-1-phosphate. The enzyme is a membrane-bound protein and in rat liver is localized in the endoplasmic reticulum. The catalytic site and other domains essential for SPLase activity face the cytosol. Recently, SPLase cDNAs were cloned from mouse and human based on sequence homology to the *Caenorhabditis elegans* and yeast SPLases. Both cDNAs encode proteins of 568 amino acids. Hydropathy analysis indicates the presence of one-transmembrane region near the N terminus.

Sphingosine-1-phosphate levels in cells during signaling processes seem to be regulated not only by sphingosine kinase and SPLase but also by a lipid phosphate phosphohydrolase. Several sphingosine-1-phosphate specific phosphohydrolases have been identified in yeast and mammalian cells.

Glycosphingolipid Catabolism

Glycosphingolipids are components of cellular membranes and are comprised of one or more sugars linked to ceramide. Glycosphingolipids are tissue and cell-type specific. They form patterns that vary with the cell type, stage of growth and differentiation, viral transformation, and ontogeny. The biosynthesis of glycosphingolipids occurs within the Golgi apparatus where carbohydrates are added by membrane-associated glycosyltranferases. Glycosphingolipids are subsequently transferred to the plasma membrane. At this site they interact with membrane associated receptors and enzymes as well as with bacterial toxins and adhesion molecules and viruses. Following endocytosis, glycosphingolipids are transported to acidic intracellular compartments where they are degraded from the nonreducing ends by a set of lysosomal exoglycosidases that catalyze the stepwise cleavage of their component carbohydrates.

Glycosphingolipid storage disorders are caused by the deficiency of a specific glycosidase that is
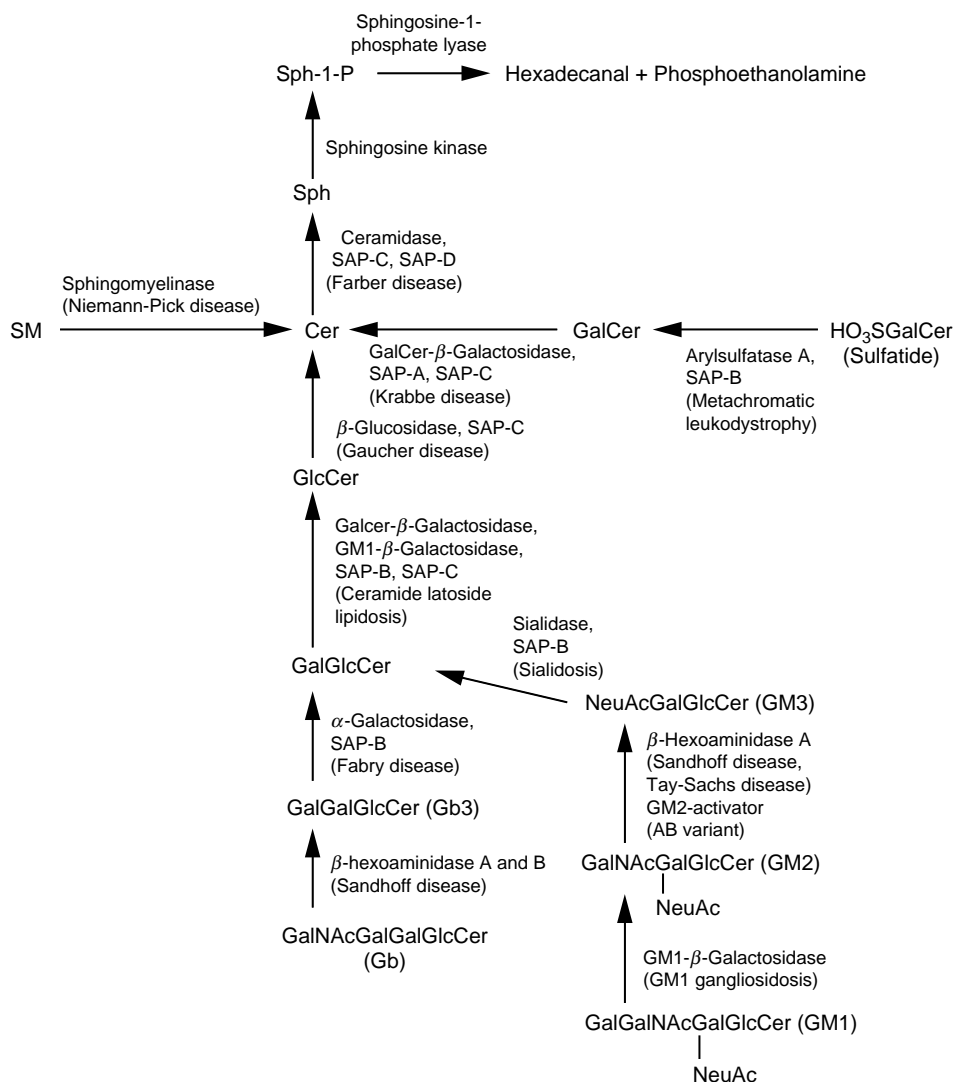
FIGURE 1 Degradation pathways of sphingolipids. The lysosomal diseases caused by genetic defects in a series of lysosomal enzymes and activator proteins for sphingolipid catabolism are indicated in parentheses. NeuAc, N-acetylneuramic acid; Cer, ceramide; Glc, glucose; Gal, galactose; GalNAc, N-acetylgalactosamine; Gb, globoside; Gb3, globotriaosyl ceramide; SM, sphingomyelin; HO$_3$SGalCer, sulfatide; Sph, sphingosine; Sph-1-P, spingosine-1-phosphate; SAP, sphingolipid activator protein.

involved in glycosphingolipid degradation (Figure 1). These deficiencies result in the abnormal accumulation of specific glycosphingolipids both in neural and nonneural tissues (Table I). These disorders vary in terms of their clinical presentations based on the affected organs.

The lysosomal hydrolases that degrade glycosphingolipids of the plasma membrane are water-soluble enzymes. Their activity is dependent on the presence of water-soluble activator proteins termed saposins. The four saposins (A through D) arise from a common precursor protein, prosaposin. Saposin B is an activator for arylsulfatase A, α-galactosidase A, and β-galactosidase. Saposin C is an activator of acid β-glucosidase. Saposin D is an activator of ceramidase. There is no

assigned function for saposin A. Glycosphingolipid storage disorders may also arise from inherited deficiencies of the prosaposin as well as from individual saposins.

ARYLSULFATASE A

Arylsulfatase A catalyzes the desulfation of 3-O-sulfogalactosyl residues in glycosphingolipids. The enzyme activity requires the presence of saposin B as an activator. The arylsulfatase A gene encodes a 507 amino acid precursor protein that undergoes posttranslational processing. In addition to N-linked glycosylation required for lysosomal sorting through the mannose-6 phosphate receptor pathway, there is a

TABLE I

Major Sphingolipidoses

Disease	Sphingolipid stored	Defective enzyme	Clinical phenotype
Farber	Ceramide	Acid ceramidase	Painful and deformed joints, subcutaneous nodules, hoarseness
Niemann-Pick A and B	Sphingomyelin	Acid sphingomyelinase	Neurodegeneration, hepatosplenomegaly
Metachromatic leukodystrophy	Sulfatide	Arylsulfatase	Blindness, quadrereparesis, seizures
Krabbe	Galactosylceramide	β-Galactocerebrosidase	Blindness, spastic paraparesis, dementia
Gaucher	Glucosylceramide	β-Glucocerebrosidase	Hepatosplenomegaly, bone infarctions and fractures
Fabry	Globotriaosylceramide	α-Galactosidase A	Parasthesias, renal failure, cerebrovascular disease
Sandhoff	Gangliosides GM2 and GA2, globoside	Hexosaminidase B	Neurodegeneration
Tay-Sachs	Ganglioside GM2	Hexosaminidase A	Neurodegeneration
GM1 gangliosidosis	Ganglioside GM1	GM1 β-galactosidase	Neurodegeneration, skeletal dysplasia, hepatosplenomegaly

unique oxidation that occurs for eukaryotic sulfatases. In human arylsulfatase A, a formylglycine residue is found in place of cysteine 69 and is due to the oxidation of a thiol group to an aldehyde. In addition to sulfatide, arylsulfatase A will cleave sulfate groups from other naturally occurring glycosphingolipids including lactosylceramide-3-sulfate and psychosine sulfate.

β-GALACTOSIDASES

β-Galactosidase catalyzes the degradation of galactosylceramide to galactose and ceramide within the lysosome. It also displays activity against galactosylsphingosine and lactosylceramide. The enzyme is more precisely referred to a galactocerebroside β-galactosidase because a second, genetically and enzymatically distinct β-galactosidase also exists with the lysosome, GM1 ganglioside β-galactosidase. GM1 ganglioside β-galactosidase catabolizes ganglioside GM1, GA1, lactosylceramide, and keratan sulfate. These enzymes have overlapping substrate specificities, but deficiencies in these enzymes result in distinct clinical disorders. Defects in the former enzyme cause Krabbe disease with the accumulation of galactosylceramide. Defects in the latter enzyme cause GM1 gangliosidosis.

Galactocerebroside β-galactosidase can be isolated as an 80 kDa polypeptide consisting of 50 and 30 kDa subunits. The pH optimum of the enzyme is acidic. The enzyme is active against galactosylceramides that vary in fatty acid chain length and α-hydroxylation. The enzyme is inhibited by sphingosine, ceramide,

and galactose. The human gene for galactocerebroside β-galactosidase maps to chromosome 14q24.3.

GM1 ganglioside β-galactosidase has an acidic pH optimum. It is activated by chloride ions. The enzyme is isolated as a large-molecular weight multimer with monomeric units of 65 kDa. This enzyme is also activated in the presence of saposin B.

β-GLUCOCEREBROSIDASE

Human acid β-glucocerebrosidase is a homomeric glycoprotein. The mature polypeptide is 497 amino acids in length. Both saposin A and saposin C activate the enzyme *in vitro*. However, only saposin C deficiency is associated with the clinical manifestations of Gaucher disease. Although saposin interacts directly with the enzyme, the association of glycosphingolipids and anionic phospholipids are required for its activity. The pH optimum for the glucocerebrosidase is 5.5. The enzyme is specific for D-glucosyl and not L-glucosyl forms of glucosylceramide. Activity is reduced by the absence of the C4−C5 *trans* double bond in sphingosine and is minimally affected by the acyl chain length of the ceramide moiety. The disease alleles in Gaucher disease are commonly missense mutations. This results in the synthesis of β-glucosidases with decreased catalytic activity or stability. These typically lead to the type I form of Gaucher disease that lacks central nervous system manifestations. The neuronopathic forms of Gaucher disease (types II and III) are more commonly associated with complete loss of catalytic activity.

α-GALACTOSIDASE A

α-Galactosidase A catabolizes glycosphingolipids with terminal α-galactosyl groups. These glycosphingolipids primarily include globotriaosylceramide, but are also present in galabiosylceramide and group B blood antigens. Initial studies on the enzyme activity suggested that there were two isozymes termed α-galactosidase A and B. Subsequent work demonstrated that these are two genetically distinct enzymes. α-Galactosidase B is an α-N-acetylgalactosaminidase. Deficiencies in α-galactosidase B are responsible for Schindler disease. Native α-galactosidase A has a molecular weight of 101 kDa and represents a homodimer with 49 kDa subunits. The enzyme is heavily glycosylated with asparagine-linked high mannose groups. The pH optimum for the glycosidase is acidic. The enzyme is activated in the presence of saposin B. The locus for human α-galactosidase A is found at Xq22.

HEXOSAMINIDASES A AND B

Ganglioside GM2 is degraded in the lysosome by β-hexosaminidase A. This enzyme removes the terminal N-acetylgalactosaminyl residue. The enzyme requires the presence of an additional protein termed GM2 activator encoded by the $GM2A$ gene. β-Hexosaminidase A is a heterodimer consisting of an α- and β-subunit. The α-subunit is a 55 kDa polypeptide encoded by the $HEXA$ gene found on chromosome 15q23-24. The β-polypeptide varies between 22 and 30 kDa and is encoded by the $HEXB$ gene found on chromosome 5q13. A second enzyme, β-hexosaminidase B is a homodimer consisting of two β-subunits. Thus $HEXA$ gene mutations result in loss of β-hexosaminidase A activity and $HEXB$ gene mutations result in loss of both β-hexosaminidase A and B activities. The β-hexosaminidase A and B enzymes are posttranslationally glycosylated for sorting to the lysosome through the mannose-6-phosphate receptor pathway. They display acidic pH optima as well.

Inherited mutations in $HEXA$, $HEXB$, and $GM2A$ all result in the accumulation of ganglioside GM2 and are associated with Tay-Sachs, Sandhoff, and GM2 activator deficiency diseases respectively. β-Hexosaminidase B deficiencies are associated with the accumulation of GA2 as well.

SEE ALSO THE FOLLOWING ARTICLES

Glycolipid-Dependent Adhesion Processes • Lipid Bilayer Structure • Lipid Rafts • Lysophospholipid Receptors • Protein Palmitoylation • Respiratory Chain and ATP Synthase • Sphingolipid Biosynthesis

GLOSSARY

cerebroside A glycosphingolipid containing a single carbohydrate, most commonly glucosylceramide and galactosylceramide.

ganglioside An acidic glycosphingolipid containing one or more sialic acid groups.

glycosphingolipid A sphingolipid covalently linked with one or more carbohydrates.

lysosome An organelle bounded by a single membrane bilayer in the cytoplasm in eukaryotic cells and having an internal pH of 4–5. Lysosomes contain several hydrolytic enzymes and serve as the site for the degradation and recycling of cellular metabolites.

sphingolipid A lipid with a backbone containing a long chain sphingoid amine.

FURTHER READING

Goni, F. M., and Alonso, A. (2002). Sphingomyelinases: Enzymology and membrane activity. *FEBS Lett.* **531**, 38–46.

Huwiler, A., Kolter, T., Pfeilschifter, J., and Sandhoff, K. (2000). Physiology and pathophysiology of sphingolipid metabolism and signaling. *Biochim. Biophys. Acta* **1485**, 63–99.

Pettus, B. J., Chalfant, C. E., and Hannun, Y. A. (2002). Ceramide in apoptosis: An overview and current perspectives. *Biochim. Biophys. Acta* **1585**, 114–125.

Siegel, G. J., Agranoff, B. W., Albers, R. W., Fisher, S. K., and Uhler, M. D. (eds.) (1999). *Basic Neurochemistry.* Williams and Wilkins, Lippincott.

BIOGRAPHY

Akira Abe holds the position of Research Investigator in the Division of Nephrology, Department of Internal Medicine at the University of Michigan Medical School. He holds a Ph.D. in biochemistry from Hokkaido University and received his postdoctoral training in the laboratory of Norman Radin at the Mental Health Research Institute at the University of Michigan.

James A. Shayman is Professor of Internal Medicine and Pharmacology at the University of Michigan School of Medicine. He also holds the position of Associate Chair for Research Programs for the Department of Internal Medicine. He was awarded his M.D. from Washington University and received his postdoctoral training in the Department of Pharmacology at Washington University.

Spliceosome

Timothy W. Nilsen
Case Western Reserve University School of Medicine, Cleveland, Ohio, USA

The spliceosome (splicing body) is a massive ribonucleoprotein complex that catalyzes the removal of noncoding intervening sequences (introns) from nuclear pre-mRNAs in the process known as splicing. Essential components of the spliceosome include five small nuclear RNAs (snRNAs) which function as RNA/protein complexes (snRNPs), and more than 100 non-snRNP-associated proteins. Despite the large number of proteins required for splicing, it seems clear that the spliceosome, like the ribosome, is fundamentally an RNA enzyme (ribozyme).

Discovery and Initial Characterization of the Spliceosome

Early analyses, primarily the comparison of cDNA and genomic clones, demonstrated that introns were ubiquitous in higher eukaryotic genes. These analyses also revealed the presence of consensus sequences that marked intron boundaries (5′ and 3′ splice sites) as well as conserved sequences ~30 nucleotides upstream of the 3′ splice site called the branch point region. However, biochemical dissection of splicing awaited the development of cell-free systems (extracts prepared from cells) that were capable of catalyzing efficient and accurate splicing of synthetic pre-mRNAs. The availability of such systems quickly led to elucidation of the chemistry of splicing and permitted the identification of cellular components required for splicing.

A variety of experiments including sedimentation analysis showed that, when incubated with extract, pre-mRNAs become incorporated into complexes similar in size to that of the ribosome. Several lines of evidence indicated that these complexes (dubbed spliceosomes) were relevant to splicing. First, pre-mRNAs in which the consensus sequences at 5′ or 3′ splice sites were altered by mutation failed to assemble into the large particles. Second, splicing intermediates were found exclusively within spliceosomes and third, inactivation of factors essential for splicing inhibited both the formation of spliceosomes and concomitantly, splicing. The discovery of spliceosomes led to an intense effort that continues today to identify its constituents and to elucidate the function of each component. Before considering the factors that comprise the spliceosome, it is necessary to briefly consider the chemistry of the splicing reaction itself.

Splicing Takes Place via Consecutive Transesterification Reactions

As noted, introns are characterized by three conserved sequence elements, the 5′ and 3′ splice sites, and the branch point region. Intron removal occurs in two chemical steps, both of which involve the exchange of one phosphodiester bond for another (transesterification). In the first step, the 2′ hydroxyl of an adenosine residue located within the branch point region (the branch point adenosine) attacks the phosphodiester bond at the 5′ splice site, breaking that 3′5′ phosphodiester linkage and replacing it with a new 2′5′ bond that connects the branch point adenosine and the first base of the intron. Accordingly, the products of the first step of splicing are liberated 5′ exon and the intron (in the form of a lariat) still linked to the 3′ exon. The second step of splicing is also a transesterification reaction. Here, the 3′ hydroxyl of the 5′ exon attacks the phosphodiester bond at the 3′ splice site breaking that bond and creating a new 3′5′ phosphodiester linkage that precisely connects the 5′ and 3′ exons. Thus, the products of the second step of splicing are ligated exons and the intron released in the form of a lariat. Because of the remarkable similarities between this reaction pathway and the pathway by which certain introns (group II) excise themselves independently of proteins, it is widely suspected that the two steps of nuclear pre-mRNA splicing are catalyzed by RNA. In considering how the spliceosome performs its task, it is important to keep the reaction pathway of splicing in mind. In this regard, the spliceosome must juxtapose the branch point adenosine with the 5′ splice site prior to the first step, anchor the 5′ exon following the first transesterification reaction and properly position the 3′ hydroxyl of the 5′ exon such that the second step

can proceed. It is now clear that these functions are performed in large part by the spliceosomal small nuclear ribonuclear proteins (snRNPs).

Required Spliceosomal Constituents: The snRNPs

Well before it became possible to analyze splicing *in vitro*, a family of small abundant stable RNAs localized to the nucleus had been identified and the sequences of individual RNAs had been determined. Because these RNAs appeared to be ubiquitous and were, in general, uridine rich they became known as U small nuclear RNAs. As they were characterized, they were given numerical designations; accordingly, the most abundant U snRNAs are called U1, U2, U4, U5, and U6. As with all cellular RNAs, the U snRNAs are complexed with proteins and are thus U snRNPs. Some of these proteins are common to many snRNPs, while some are specific to individual RNPs. For example, U1 snRNP contains seven common proteins and three U1 specific proteins.

When consensus signals demarcating introns were identified, it became obvious that specific sequences within the U snRNPs could potentially interact via base pairing with those sequences. In particular, the 5′ end of U1 snRNA was predicted to base pair with 5′ splice sites and a region of U2 snRNA could base pair with the branch point region. These and other observations suggested that the U snRNPs might be involved in splicing. This hypothesis could be tested once cell free systems became available. Using a battery of specific depletion strategies, it was shown that U1, U2, U4, U5, and U6 snRNPs were all required for splicing and all were constituents of the spliceosome; accordingly these RNAs are now known as spliceosomal snRNAs. The specific function of each spliceosomal snRNP has been the subject of intense investigation for many years and our current understanding of their roles is discussed subsequently.

Spliceosomal Constituents — Non-snRNP Proteins

While much attention was focused on snRNPs, other approaches, primarily genetic analyses in budding yeast and biochemical fractionation of extracts prepared from mammalial cells, revealed that a plethora of factors were required for splicing. Several techniques, including specific immunoprecipitation, showed that these factors were *bona fide* spliceosomal constituents and that they were not intrinsic components of snRNPs. At one point,

it was thought that there were 30–40 non-snRNP splicing factors. However, recent improvements in purification techniques coupled with dramatic advancements in the ability to determine the identity of proteins via mass spectrometry have revealed that the spliceosome contains at least 100 non-snRNP proteins and perhaps many more. Accordingly, the spliceosome contains, at a minimum, 200 distinct proteins and five distinct RNAs, making it among the most complex macromolecular machines in the cell. Some reasons for this remarkable complexity are discussed.

What are the non-snRNP splicing factors? While the role(s) of many non-snRNP proteins remain obscure, the functions of others have been elucidated at least in part. Among these are a group of enzymes known as RNA dependent ATPases (RNA helicases). At least six helicases are required for splicing and they are thought to catalyze the RNA/RNA rearrangements that occur during the splicing reaction. The only other known spliceosomal enzyme is a protein phosphatase, but a definitive substrate for this protein has not yet been identified. Other proteins with known functions are those involved in recognition of the pre-mRNA prior to assembly of the entire spliceosome; these proteins assist the snRNPs in engaging the intron.

Spliceosome Assembly and the Splicing Cycle

Unlike the ribosome, which is a stable preformed macromolecular complex, the spliceosome must assemble anew on each intron. An extensive variety of biochemical approaches has indicated that spliceosome assembly follows an ordered pathway (see Figure 1). First, U1 snRNP recognizes and binds to the 5′ splice site; this binding is mediated in part by base pairing between U1 snRNA and the pre-mRNA. Once U1 snRNP is bound, it promotes recognition of the branch point region by U2 snRNP, again this binding is mediated in part by base pairing. Once U1 and U2 snRNPs are bound, U4, U5, and U6 snRNPs join the growing spliceosome as a preformed unit (tri-snRNP). A host of non-snRNP associated proteins accompany tri-snRNP to the spliceosome. Upon completion of assembly, the spliceosome undergoes a dramatic conformational change which involves the release of proteins and the addition of new ones; it is not yet known what triggers this rearrangement. Most strikingly, during this first conformational change, U1 and U4 snRNPs are destabilized and released from the spliceosome; accordingly these snRNPs are not required for catalysis. After the precatalytic conformational changes are completed, the first chemical step of splicing occurs. Following this transesterification reaction, a second
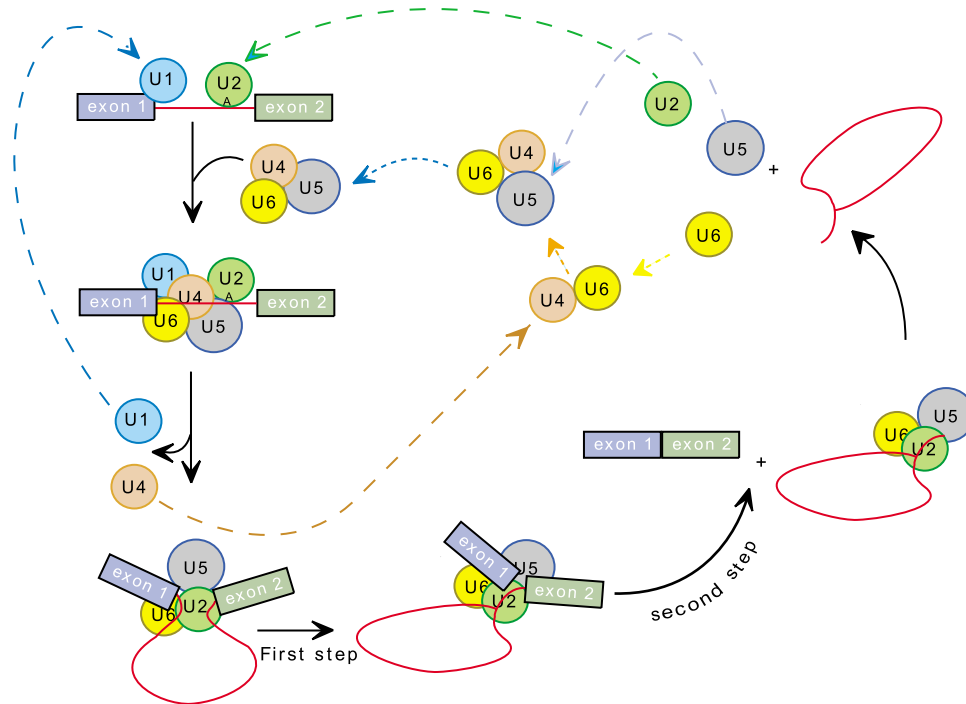
FIGURE 1   Schematic depiction of the splicing cycle (for details, see text).

conformational change ensues. Although not as well understood as the first, it is clear that this rearrangement also involves dramatic changes in spliceosomal composition; i.e., factors depart and new ones join. Subsequent to this conformational change, the second chemical step occurs and the spliceosome is disassembled. Disassembly is not well understood, but is it clear that the snRNPs and other splicing factors are recycled for subsequent rounds of assembly and catalysis (see Figure 1).

A Dynamic Network of RNA/RNA Interactions in the Spliceosome

While the precise role of many individual spliceosomal constituents remains to be determined, their main collective function is to orchestrate a complex series of snRNA/pre mRNA and snRNA/snRNA interactions that culminate in catalysis. Upon completion of spliceosome assembly, U1 snRNP is base paired to the 5′ splice site, U2 snRNP is base paired to the branch point region, U5 snRNP makes contact with the 5′ exon and U4 and U6 snRNPs are joined via extensive base pairing with each other. During the first rearrangement, the base pairing between U1 snRNP and the 5′ splice site is disrupted and replaced by base pairing interaction between U6 snRNA and the 5′splice site. Concomitantly, U4 and U6 are unwound and U6 snRNA forms extensive base pairing interactions with U2 snRNA. The net effect of these RNA/RNA arrangements is to juxtapose the

reactants of the first step of splicing, the 5′ splice site and the branch point adenosine. Following the first step, U5 snRNP tethers the 5′ exon and positions it for attack at the 3′ splice site (see Figure 2). A large body of evidence indicates that catalysis is mediated by the snRNAs themselves. Indeed, fragments of U2 snRNA and U6 snRNA can catalyze a reaction analogous to the first step of splicing in the complete absence of protein cofactors. Accordingly, the spliceosome, like the ribosome, is fundamentally an RNA enzyme (ribozyme).

Why is the Spliceosome so Complex?

The overall complexity of the spliceosome is, at first glance, puzzling because the reaction it catalyzes is chemically straightforward; as noted above, certain introns can be excised without the assistance of any proteins. So, why so many proteins? At least four considerations help to explain the multitude of splicing factors.

First, the spliceosome must recognize a huge number of substrates; hundreds of thousands of introns are present in a higher eukaryotic genome. Introns vary greatly in length and in the degree of conservation of splicing signals. Thus the spliceosome must be able to adapt to many different sequence contexts. Indeed, several splicing factors are essential for the splicing of some introns but dispensable for the splicing of others.
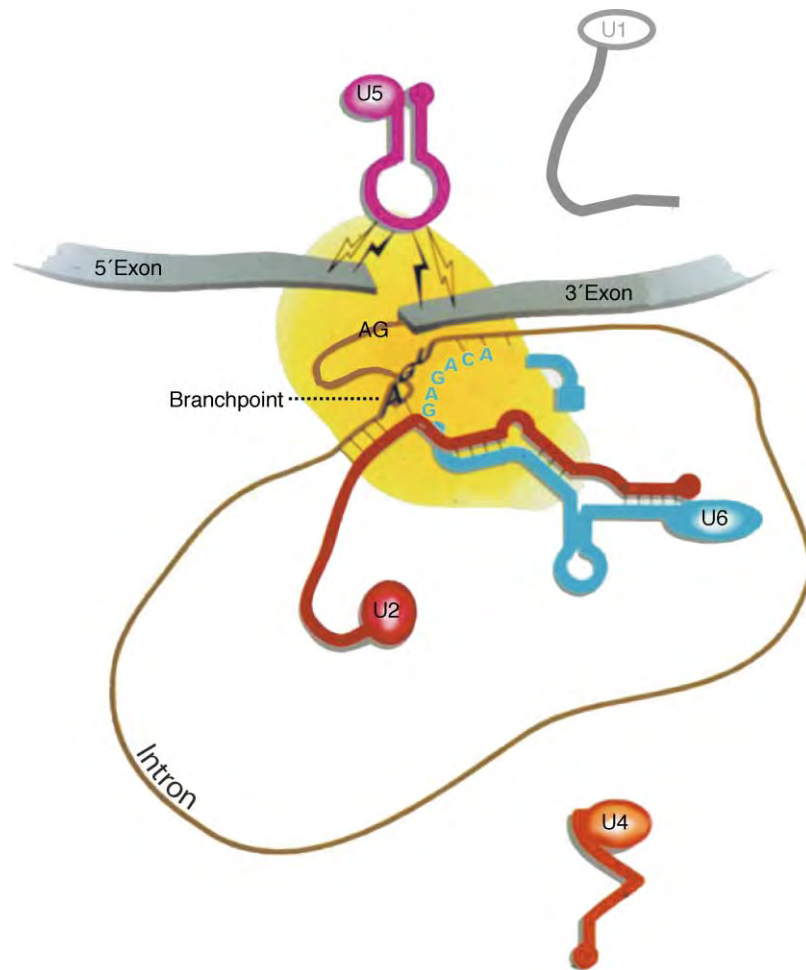
FIGURE 2 Schematic depiction of RNA/RNA interactions in the spliceosome after the first chemical step and prior to the second step (for details, see text).

Second, splicing must be extremely accurate. A mistake of only one base can lead to the production of a defective mRNA. The remarkable fidelity of splicing is achieved via the existence of multiple proofreading activities that monitor the correct alignment of each component prior to catalysis. For example, the 5' splice site is inspected by at least four distinct activities during different steps in spliceosome assembly. The existence of "fidelity factors" was revealed by mutations in splicing components that led to high levels of mis-splicing.

Third, accumulating evidence indicates that splicing is intimately linked to other cellular processes such as transcription and 3' end formation. It is now clear that many components of the spliceosome serve as molecular bridges between the transcription apparatus and the core splicing machinery.

Fourth, the spliceosome performs functions in addition to its catalytic role in splicing. Most strikingly in this regard, the process of splicing leaves a specific "mark" on spliced mRNAs. This "mark" known as the exon junction complex (EJC) is comprised of at least six proteins. While none of these proteins is essential for splicing itself, they serve multiple roles after splicing. The EJC facilitates transport of spliced mRNAs from the nucleus, is involved in mRNA surveillance (the process whereby defective mRNAs are identified and destroyed), and is important in the translation of spliced mRNAs.

These, and perhaps yet to be discovered activities, account in part for the remarkable complexity of the spliceosome. A major challenge in coming years will be the elucidation of the role(s) of many spliceosomal factors whose function is not yet known.

SEE ALSO THE FOLLOWING ARTICLES

Pre-tRNA and Pre-rRNA Processing in Bacteria • Pre-tRNA and Pre-rRNA Processing in Eukaryotes • Ribozyme Structural Elements: Group I Introns • Ribozymes and Evolution

GLOSSARY

branch point adenosine The intronic adenosine, whose 2' hydroxyl forms a 2'5' phosphodiester linkage with the 5' base of an intron during the first chemical step of spicing.

exon   Regions of pre-mRNAs that are retained in mature mRNAs following splicing.

exon junction complex   A collection of proteins that are deposited on the mature mRNA as a consequence of splicing.

introns   Noncoding regions of pre-mRNAs that are excised via splicing.

pre-mRNA   The primary transcript of a gene synthesized by RNA polymerase. The primary transcript extends from the promoter to beyond the 3′ end of the mature mRNA and is subject to numerous processing events including splicing.

transesterification   The chemical process in which one phosphodiester bond is broken and another is formed simultaneously.

FURTHER READING

Collins, C. A., and Guthrie, C. (2000). The question remains: Is the spliceosome a ribozyme? *Nat. Struct. Biol.* **7**, 850–854.

Hastings, M. L., and Krainer, A. R. (2001). Pre-mRNA splicing in the new millennium. *Curr. Opin. Cell Biol.* **13**(3), 302–309.

Jurica, M. S., and Moore, M. J. (2003). Pre-mRNA spicing: Awash in a sea of proteins. *Mol. Cell* **12**, 5–14.

Moore, M. J., Query, C. C., and Sharp, P. A. (1993). Splicing of precursors to mRNAs by the spliceosome. In *The RNA World* (R. F. Gesteland and J. F. Atkins, eds.) pp. 303–357. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Nilsen, T. W. (1998). RNA–RNA interactions in nuclear pre-mRNA splicing. In *RNA Structure and Function* (R. W. Simons and M. Grunberg-Manago, eds.) pp. 279–307. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Staley, J. P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell* **92**, 315–326.

Will, C. L., and Lührmann, R. (2001). Spliceosomal U snRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.* **13**, 290–301.

BIOGRAPHY

Timothy W. Nilsen is Professor and Director of the Center for RNA Molecular Biology at Case Western Reserve University School of Medicine. His principal research interest is in the mechanism of pre-mRNA splicing, with particular emphasis on the process of splice site recognition in higher eukaryotes and trans-splicing, an unusual splicing reaction that occurs in certain lower eukaryotes. He holds a Ph.D. from SUNY Albany and serves as Editor-in-Chief of the scientific journal, *RNA*.

Src Family of Protein Tyrosine Kinases

Jonathan A. Cooper

Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

The Src (sarcoma) family is a group of protein tyrosine kinases closely related to Src. They are nonreceptor kinases, meaning that they lack a transmembrane domain and lie totally within the confines of the cell, although they are associated with cell membranes. Src family kinase activity is regulated by phosphorylation sites within the kinase domain activation loop and at the carboxy terminus, and by inhibitory interactions between the kinase domain and two other conserved domains, the SH2 and SH3 scaffolding domains. These latter domains can either inhibit the kinase domain or bind other cell proteins, thus helping localize Src to specific sub-cellular regions and protein complexes while simultaneously regulating Src kinase activity. There are eight Src family members in vertebrates, with different but overlapping expression patterns and, potentially, distinct functions within a given cell. However, there is also extensive genetic redundancy within the family, so inactivating mutations reveal only a subset of the potential cellular and developmental roles of family members. On the other hand, any of a number of different mutations can activate Src kinases, and the resulting deregulated kinase activity has profound effects on cell biology, including malignant transformation of some cell types and increased differentiation of others. At least one Src-like kinase gene is found in every metazoan, including sponges and Cnidarians, but is absent, like other tyrosine kinases, from yeasts, protists, plants, and non-eukaryotes. Src kinases are thus one of the ancestral tyrosine kinases and have basic, fundamental roles in cell biology.

History of Src

The history of Src is a story of firsts, and research on Src has spawned wider areas of research. Src was the first protein tyrosine kinase activity to be identified, as a consequence of its role in malignant transformation. In the early 20th century, P. Rous demonstrated that a chicken sarcoma could be transmitted from bird to bird by a particle smaller than a bacterium, and the field of tumor virus research was born. Rous's virus, RSV, was later shown to contain separable genetic elements for replication and for cell transformation. RSV also was one of the first viruses shown, by D. Baltimore and H. Temin, to contain RNA-dependent DNA polymerase (reverse transcriptase) activity, consistent with the idea that transformation was caused by a DNA copy of a viral gene for sarcoma (*src*). D. Stehelin, H. Varmus, and J. Bishop, using nucleic acid probes, showed that a gene related to the viral src gene was present in uninfected vertebrate cells and was conserved over evolution. This provided evidence for the cellular origin of viral oncogenes. The viral src gene is now generally called v-*src*, while the cellular proto-oncogene is called c-*src*, or just *src*. The role of cellular proto-oncogenes in viral and non-viral cancers of animals and humans is now widely accepted.

In 1977, J. Brugge and R. Erikson found that certain antisera from RSV-infected newborn rabbits precipitated a 60,000 M_r protein in addition to the virus structural proteins. The gene encoding this protein was shown to be v-*src*, and the protein was named pp60v-*src* or vSrc. It was soon found that this protein contained, or was associated with, a protein kinase activity that would transfer phosphate to antibody molecules or other model substrate proteins. While this kinase was first thought to phosphorylate threonine residues, T. Hunter and B. Sefton soon found that only tyrosine residues were phosphorylated. Indeed, the first tyrosine kinase to be identified was an activity, now known to be due to cellular Src family kinases, associated with the tumor-causing protein (middle T antigen) of polyoma virus (a DNA tumor virus). Both Src and vSrc are tyrosine kinases, but vSrc is 20+-fold more active than Src due to mutations that interfere with the negative autoregulatory interactions that restrain the activity of Src in normal cells. Different strains and derivatives of RSV have vSrc proteins with different numbers of such activating mutations, which have been very informative in understanding how Src is regulated.

The ability of vSrc to cause malignant transformation is thought to be due to unrestricted phosphorylation of substrate proteins that are normally phosphorylated, in a controlled fashion, by Src and other Src family kinases. In addition, the high kinase activity of vSrc

may lead to the phosphorylation of other cell proteins that are not normally phosphorylated by Src, but it seems unlikely that these additional substrates are involved in transformation. Similarly, transformation by polyoma virus involves the deregulation of Src, by physical association between middle T antigen and Src and several other proteins. The consequent unregulated activity of Src is critical for polyoma virus transformation.

While activated vSrc transforms fibroblasts, it causes other effects in other cell types. In some cases it can induce differentiation and in others increase cell migration or morphology changes. These dramatic effects of activated mutant Src or vSrc led to an expectation that endogenous Src would be vital for many fundamental cellular processes.

Functions of Src in Development

The unique functions of Src were revealed when P. Soriano knocked out the mouse *src* gene. Surprisingly, considering the profound effects of activated v-*src* on vertebrate cells, absence of *src* had remarkably mild effects on development *in utero*. However, some defects were detected after birth. Teeth failed to erupt through the gum, and the cranium grew to form a domed shape. Both of these phenotypes are due to increased bone density, attributable to altered osteoclasts, the macrophage-related blood cells that resorb and remodel bone. Detailed studies of *src* mutant osteoclasts indicate that they differentiate normally but have a reduced ability to form specialized adhesion contacts through which they stick to bone and initiate bone resorption. However, blood platelets and neurons, where Src is highly expressed, are not noticeably altered in *src* mutant mice.

Other Vertebrate Src Family Kinases: Redundant and Specific Functions

The mild effects of *src* deletion are partly due to redundancy with other *src* family genes. The *yes* and *fgr* genes were first identified, as with *src*, as oncogenes in various cancer-causing retroviruses. Upon sequencing, they were found to be related to *src*. *lck, hck, blk, lyn*, and *fyn* were cloned by homology. Of the Src family proteins, Src, Fyn, and Yes are widely expressed in many different cell types in the body, whereas Fgr, Lck, Hck, Blk, and Lyn are restricted to, and are more important in, hematopoietic cells. However, alternative splicing and alternative promoter usage (with different coding exons) creates additional diversity. For example,

one splice form of Fyn is restricted to hematopoietic cells while another is more ubiquitous. All the Src family proteins contain canonical features of Src that are critical for regulation and localization: the SH3, SH2, and kinase domains, amino-terminal lipid modification sites, and carboxy-terminal and activation loop tyrosine phosphorylation sites. Other subfamilies of non-receptor protein tyrosine kinases include the Csk, Abl, Fes, Tec/Btk, JAK, Syk, and FAK families, which lack one or more features of Src or contain other distinctive domains. The common features of Src family members allow considerable overlap in function, as shown by targeted knockouts and characterization of multiple mutants.

Individual mutation of *fyn* and *yes*, like *src*, causes only mild effects. For example, *fyn* knockouts have several alterations in the brain, including misorientation of pyramidal neurons in cortex and hippocampus and reduced myelination. In addition, maturation of *fyn* mutant T lymphocytes is impaired. However, more phenotypes are manifested when *src, fyn*, and *yes* are mutated in combination. Double mutants have reduced perinatal survival, and triple knockouts survive poorly beyond mid-gestation (embryonic day 9 in the mouse). These triple mutant embryos resemble embryos mutant for the extracellular matrix protein fibronectin or some integrins, suggesting redundant roles for these three Src relatives in adhesive interactions between cells and the extracellular matrix.

The hematopoietic Src family kinases also have redundant and nonredundant functions. Consistent with their restricted expression, none of these kinases is needed for normal development: mice lacking *hck, fgr*, and *fyn* are moderately healthy and fertile, as are mice lacking *blk, fyn*, and *lyn*. However, some hematopoietic defects are detected. Mutation of *lck* alone blocks thymocyte development at a stage prior to a proliferative burst, with a resulting large decrease in T-cell numbers. Macrophages from *hck* mutant mice have impaired phagocytosis. Combined absence of *fyn* and *lyn* leads to autoimmune disease, absence of *blk, fyn*, and *lyn* inhibits late-stage B-cell development, and absence of *hck* and *fgr* leads to decreased resistance to bacterial infection. Blood platelets lacking *src, hck, fgr*, and *lyn* fail to spread on fibrinogen. Thus, important roles for Src kinases in signaling from immunoreceptors and integrins become evident when multiple Src kinases are absent.

The preceding examples indicate that there is considerable redundancy in the Src family (i.e., one kinase can take the place of another if one is missing). However, whether different family members have distinct functions in a given cell remains unknown, since one kinase may normally be dedicated to a specific function but may perform others if another family member is missing.

Cellular Functions of Src Family Kinases: Signaling Adhesion and Immune Responses

The Src family kinases present in hematopoietic cells are activated by a number of stimuli, including integrin ligands, cytokines, and, notably, stimuli that act via immunoreceptors, such as the T-cell antigen receptor on thymocytes and circulating T lymphocytes, B-cell antigen receptor on pre-B cells, and immunoglobulin (Fc) receptors on macrophages. Cell-based assays confirm essential (but redundant) roles for Src kinases in immunoreceptor responses. Current models envisage weak association of Src kinases with the cytoplasmic tails of subunits associated with the immunoreceptors. Clustering of the receptors, possibly associated with altered proximity to protein tyrosine phosphatases (PTPs) and movement into or out of lipid rafts, activates the Src kinase, allowing phosphorylation of tyrosine-containing activation motifs in the immunoreceptors and recruitment of other SH2 domain-containing nonreceptor tyrosine kinases of the Syk and Tec/Btk families. In this situation, the key Src substrate is the immunoreceptor, but Src also contributes to phosphorylation of downstream signaling molecules involved in calcium mobilization and gene expression.

Src, Fyn, and Yes are also activated by adhesive stimuli and by soluble ligands that act via transmembrane receptors of the tyrosine kinase, G protein-coupled, cytokine, and other classes. The substrates phosphorylated vary according to the type of stimulus and the other proteins recruited by the respective receptor. For example, integrin stimulation causes Src activation and phosphorylation of another tyrosine kinase, FAK. The exact relationship between FAK activation and Src activation is not clear, but probably FAK autophosphorylates and activates Src, which further phosphorylates FAK. Both proteins contribute to phosphorylation of other proteins associated with the integrin. Functionally, absence of Src kinases alters cytoskeletal dynamics, for example, reducing turnover of focal adhesion complexes and slowing cell migration. However, Src activation by integrins also contributes to tyrosine phosphorylation of signaling molecules responsible for stimulating mitogen-activated protein kinase cascades and regulating gene expression. Src kinase activity is also required for cell proliferation induced by soluble growth factors acting via tyrosine kinase receptors. Dissecting the effects of Src activation by the growth factor receptor from the effects of Src activation by integrins is complex, since many cell cycle events elicited by growth factors depend on cell adhesion via integrins. Src kinase activity can also regulate protein traffic to and from membranes. In relaying such diverse signals, Src family kinases probably have many important substrates, more than one of which may be necessary for an observed biological response. Src kinases thus represent branch points for activation of multiple signaling pathways.

Regulation of Src Kinase Activity

The combined results of mutational and crystallographic analysis have led to a picture of Src as a machine that integrates many different inputs to regulate its conformation and kinase activity (see Figure 1). There are at least two conformational states, simply thought of as on and off. In the off state, the kinase domain is relatively inactive, and the SH2 and SH3 domains are engaged in low-affinity intramolecular interactions. In the on state, the intramolecular interactions are absent, the kinase domain is at least partially active, and the SH2 and SH3 domains are completely accessible for binding other proteins. Thus, in its on state, Src can also potentially act as a scaffold, independent of its kinase activity. For example, Src may bind one protein through its SH3 domain and another through its SH2 domain, bringing the two proteins into a complex that conveys a signal to the cell. However, the extent to which Src kinases have kinase-independent functions *in vivo* when expressed at normal levels is not yet clear.

Src in the on state is only partially active as a kinase, and the kinase domain becomes fully active only when the activation loop is phosphorylated, a reaction that may be intramolecular but that can also be catalyzed by nearby tyrosine kinases, including other Src molecules. Clustering of partially activated molecules may thus increase activation loop phosphorylation and lead to full activity. Such clustering may be important in regulation of Src family kinases by integrins and immunoreceptors. Conversely, dephosphorylation of the activation loop would reduce Src activity in the cell.

The on and off states are not stable structures, however. Molecular motions at physiological temperature cause these states to be somewhat flexible, allowing the equilibrium to be disturbed by interactions with ligands for the SH2 and SH3 domains. Because the individual intramolecular interactions stabilizing the off state are each rather weak, reduction of any one interaction leads to a concerted transition to the on state. Thus, increasing the local concentration of either an SH2 or an SH3 ligand would be expected to shift the equilibrium toward the on state. This is thought to occur in some situations in which viral proteins (such as polyoma virus middle T antigen and human immunodeficiency virus Nef) are highly expressed. Particularly potent activators would have binding sites for both the SH2 and SH3 domains. Because phosphorylation of substrates by the Src kinase can create binding sites for
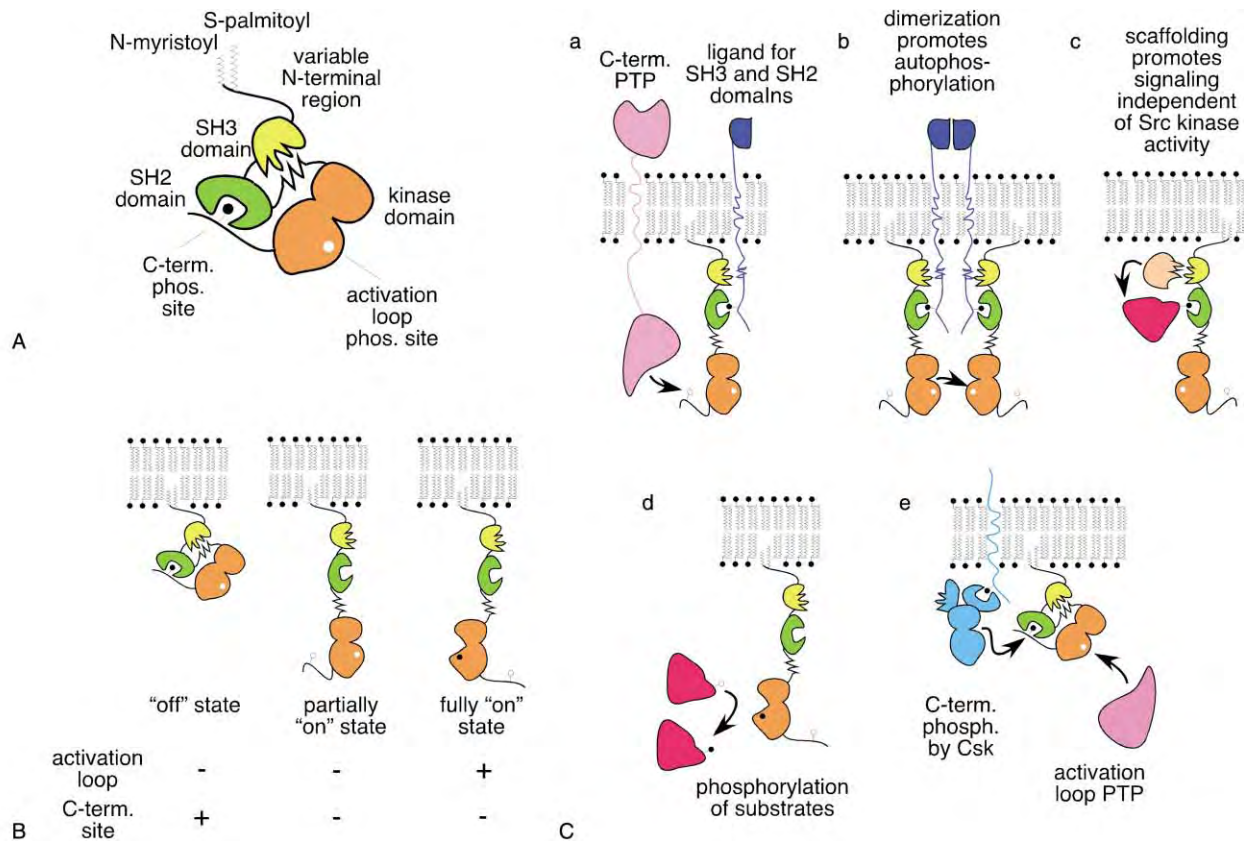
FIGURE 1 (A) Features of Src family kinases. The N terminus is modified by N-myristoylation and (for most but not all family members) by S-palmitoylation. Conserved structural domains are named SH3, SH2, and kinase (or SH1) domains, proceeding from N terminus to C terminus. In the off state conformation shown, the SH3 domain is engaged with a linker between the SH2 and kinase domains, and the SH2 domain is engaged with a phosphorylated tyrosine (black circle) near the C terminus. A tyrosine in the activation loop is not phosphorylated (white circle). (B) Different activity states of Src kinases. In both the partially and fully on states, the C-terminal phosphorylation site is dephosphorylated, and the molecule is opened into a flexible conformation with SH3, SH2, and kinase domains available for interactions with other proteins. In the fully on state, the activation loop phosphorylation site is phosphorylated, the kinase domain undergoes additional conformational changes, and substrates can be phosphorylated with high efficiency. (C) Regulatory interactions of Src kinases with other molecules in the cell. (a) Src can be partially activated by the combined or independent actions of PTPs that dephosphorylate the C-terminal site and protein ligands that can associate with the SH3 and/or SH2 domains. (b) Full activation may occur when partially activated molecules are brought together by clustering of a receptor or by decreased access to a PTP that dephosphorylates the activation loop (not shown). (c) Fully or partially active Src may also act as a scaffold, bringing proteins together but not phosphorylating them. (d) Active Src phosphorylates nearby cell proteins. Partially active Src may also phosphorylate substrates, but less efficiently. Substrates may be associated with the Src SH2 and/or SH3 domains. (e) Src is inactivated by the combined effects of Csk, which phosphorylates the C terminus, PTPs that dephosphorylate the activation loop, and inhibited access to SH2 and SH3 ligands (not shown). Csk may be recruited to Src by interactions with an anchor protein, such as Cbp/PAG. The PTP may also be specifically recruited (not shown). Interactions with regulatory molecules may be affected by movement of Src or the regulatory molecules between different microdomains in the lipid bilayer.

its SH2 domain, Src can be activated by the products of its own kinase activity. Clustering of Src with substrates containing SH3 domain-binding sites may thus activate Src under certain conditions. Removal of such ligands would allow Src to revert to the off state.

In addition to control by clustering and by phosphorylation of the activation loop, another control is provided by phosphorylation of a tyrosine residue in the carboxy terminus of Src. When phosphorylated, this tyrosine serves as a ligand for the Src SH2 domain, stabilizing the off state. Dephosphorylation of this site allows opening of the off state structure and hence activation. The carboxy-terminal region plays no apparent

role when Src is active, and many mutationally activated forms of Src family kinases, found in various oncogenic retroviruses, have deleted or mutated this region. In the cell, a major source of Src regulation is provided by phosphatases that dephosphorylate the carboxy terminus and kinases that rephosphorylate this site. Genetic evidence suggests that the PTP CD45 is key to activating Lck in antigen-stimulated T cells, and that the PTPs RPTPα and Shp2 are involved in regulating Src in fibroblasts. However, to what extent these PTPs are regulated catalytically, and whether they provide a permissive environment in which Src family kinases can be activated by other means, is not clear.

Rephosphorylation of the carboxy terminus is catalyzed by a dedicated protein kinase, Csk, first identified by M. Okada and H. Nakagawa in brain extracts. Csk has few, if any, substrates other than the carboxy termini of Src family kinases. Genetic disruption of *csk* causes hyperactivation of Src family kinases, with consequent developmental abnormalities and death at mid-gestation stage. In *csk* mutant fibroblasts, Src family kinases are activated and the cells are transformed. A close Csk relative, Chk, may also be involved. In *Caenorhabditis elegans* and *Drosophila melanogaster*, Csk relatives can be recognized by sequence, and they are likely to function as Src inhibitors. Csk is structurally similar to Src family kinases but lacks the activation loop and carboxy-terminal tyrosine phosphorylation sites and an amino-terminal lipid modification site. It is thus not regulated like Src family kinases; its location inside cells, however, is regulated. Recent evidence indicates that Csk can bind to, and is activated by, one or more Src substrates (one named Cbp or PAG) that colocalize with activated Src family kinases in membrane subdomains. When these substrate molecules are phosphorylated, Csk is recruited by its SH2 domain, and the Csk is then brought close to the activated Src family kinases. Csk may also be associated, through its SH3 domain, with PTPs that target the activation loop tyrosine of Src. This mechanism would ensure that the Src family kinases are then turned back off by Csk-catalyzed phosphorylation of their carboxy termini and PTP-catalyzed dephosphorylation of the activation loop.

The activation state of an individual Src kinase molecule in the cell is thus the result of a balance between protein–protein interactions and phosphorylation and dephosphorylation of the activation loop and carboxy-terminal tyrosines. However, the activity state (on or off) also affects access by the modifying enzymes. Therefore, there is a complex balance of interactions that allows Src to integrate different inputs. In addition, Src molecules are subject to other post-translational modifications that regulate activity or localization. In mitosis, Src becomes phosphorylated by Cdk1/Cdc2-cyclin complexes. These serine phosphorylations partially activate Src by reducing the intramolecular interactions that stabilize the off state. In cells responding to agonists that increase cyclic AMP levels, the cAMP-dependent protein kinase phosphorylates Src and activates it.

Subcellular localization may also be regulated. All Src family kinases are associated with the plasma membrane and intracellular membranes by amino-terminal sequences. The amino-terminal methionine is removed and the next residue (glycine) is subject to N-myristoylation. This modification is cotranslational and not regulated, but the majority of Src family kinases are also subject to reversible, regulated lipid modification by palmitoylation. This modification affects one or more cysteine residues near the amino terminus. Palmitoylation ensures that most Src family kinases (with the exceptions of Src and Blk) are concentrated in microdomains within the plasma membrane that are enriched in phosphatidylinositol-4,5-bisphosphate, cholesterol, and glycosphingolipids, so-called lipid rafts. These microdomains are also enriched in many signaling molecules, placing the Src kinases close to potentially important substrates and effectors. Cbp/PAG, the aforementioned Csk-recruiting molecule, is also in the lipid rafts. Receptors bearing carboxy-terminal glycosylphosphatidylinositol anchors are localized in the outer leaflet of lipid rafts, and clustering these molecules activates Src family members associated with the inner leaflet. It seems probable that slight changes in protein distribution within lipid rafts will be very important for regulating Src kinase activity and directing phosphorylation activity to the appropriate substrates.

Nonvertebrate Src Kinases and their Functions

Genetic studies in the fruit fly *D. melanogaster* and nematode *C. elegans*, where there are fewer *src* genes (two in *D. melanogaster* and one in *C. elegans*), have not revealed distinctive mutant phenotypes. Thus, unlike several other signal transduction pathways, genetic analysis in these organisms has not led the way in understanding how Src is regulated and functions. It is possible that Src proteins in these organisms (as in vertebrates) serve as integrators of many different inputs, and mutations lead to pleiotropic phenotypes. However, one identified function of Src64B in *D. melanogaster* is to regulate cytoskeletal structures in the cells that nurse the developing oocyte, and it does so in concert with a Btk/Tec-related tyrosine kinase. In this regard, it resembles the role of Lck in activating Btk/Tec-related tyrosine kinases in mammalian lymphocytes. Src42A also regulates the actin cytoskeleton in the developing embryo. In *C. elegans*, Src is important in the early embryo for correct cell–cell interactions, which may implicate Src in cytoskeletal organization. Further study in these systems may provide insights into Src regulation and functions in vertebrates.

Summary: Integration of Many Inputs and Regulation of Many Outputs

The many ways to regulate Src family kinases, and the many substrates implicated in different downstream events, position Src family kinases as servants with many

masters. In the cell, Src kinases are regulated by the balance of kinases and phosphatases acting at both inhibitory and activating phosphorylation sites, and by proteins that bind to their SH2 and SH3 domains. The proteins phosphorylated depend on the subcellular localization and the protein complexes in which the active Src is found, and relay signals to regulate cytoskeletal, membrane, and nuclear events.

See Also the Following Articles

B-Cell Antigen Receptor • Glycosylphosphatidylinositol (GPI) Anchors • Immunoglobulin (Fc) Receptors • Integrin Signaling • Lipid Rafts • Mitogen-Activated Protein Kinase Family • Protein N-Myristoylated • Protein Palmitoylation • Reverse Transcriptase and Retroviral Replication • T-Cell Antigen Receptor

Glossary

C-terminal Src kinase (Csk) Kinase that phosphorylates the C-terminal tyrosine residue present on all Src-related kinases, thus inhibiting Src kinase activity.

protein tyrosine kinase A protein kinase that phosphorylates tyrosine residues in substrates.

SH2 domain A protein module that binds with high affinity to phosphotyrosine residues contained in a defined peptide sequence, with primary specificity usually being conferred to the three residues C-terminal to the phosphotyrosine.

SH3 domain A protein module that binds to sequences that adopt a left-handed type II polyproline helix, typically with a PXXP core.

Further Reading

Abram, C. L., and Courtneidge, S. A. (2000). Src family tyrosine kinases and growth factor signaling. *Exp. Cell Res.* **254**, 1–13.

Blume-Jensen, P., and Hunter, T. (2001). Oncogenic kinase signaling. *Nature* **411**, 355–365.

Lowell, C. A., and Soriano, P. (1996). Knockouts of Src-family kinases: Stiff bones, wimpy T cells, and bad memories. *Genes Dev.* **10**, 1845–1857.

Martin, G. S. (2001). The hunting of the Src. *Nat. Rev. Mol. Cell. Biol.* **2**, 467–475.

Sicheri, F., and Kuriyan, J. (1997). Structures of Src-family kinases. *Curr. Opin. Struct. Biol.* **7**, 777–785.

Thomas, S. M., and Brugge, J. S. (1997). Cellular functions regulated by Src family kinases. *Annu. Rev. Cell. Dev. Biol.* **13**, 513–609.

Weiss, A., and Littman, D. R. (1994). Signal transduction by lymphocyte antigen receptors. *Cell* **76**, 254–263.

Biography

Jonathan A. Cooper is a member of the Division of Basic Sciences at the Fred Hutchinson Cancer Research Center and an affiliate Professor in the Department of Biochemistry at the University of Washington in Seattle. He earned his Ph.D. at the University of Warwick and did postdoctoral work with Bernard Moss at the National Institutes of Health and with Tony Hunter at the Salk Institute. His principal research interests are the signaling pathways regulated by Src family kinases and the signaling and traffic functions of adaptor proteins.

Starvation

Oliver E. Owen and Richard W. Hanson
Case Western Reserve University School of Medicine, Cleveland, Ohio, USA

There is a paradoxical relationship between morbid obesity and starvation because most of the reproducible data regarding fasting metabolism were derived from obese people undergoing prolonged fasting periods to reduce their body weights. During starvation, blood glucose concentration falls and this decrease is paralleled by serum insulin concentrations. Proteolysis, lipolysis, gluconeogenesis, and ketogenesis ensue. During starvation there develops an orchestrated interplay among the organ systems to produce, select, and conserve fuels needed for body metabolism. Steady-state metabolic processes develop after $\simeq 18$ days of total starvation. Urinary excretion of nitrogen declines to $4-5\,\mathrm{g\,day^{-1}}$, and ammonium becomes the primary nitrogenous excretory product. Blood substrates and hormones are near constant concentrations. Weight losses and energy requirements are comparable day after day. The brain derives the majority of its energy requirements by extracting β-hydroxybutyrate and acetoacetate from the blood. Glucose oxidation by the brain is suppressed. The liver supplies the large quantity of ketone bodies needed for brain metabolism. Muscle probably reduces acetoacetate to beta-hydroxybutyrate. The kidney permits a limited amount of ketone bodies to be excreted in the urine to maintain near-electrical neutrality with urinary ammonium. The kidney shares the essential but limited role of gluconeogenesis with the liver. Ketogenesis, gluconeogenesis, ammoniagenesis, ureagenesis, and ketonuria are tightly interrelated during prolonged starvation.

Starvation or hunger in humans is the deprivation of any or all of the elements necessary for their nutrition. Several extensive review chapters and books have been written about the physiological, biochemical, and psychological observations and measurements made during semi-starvation or total starvation. However, there is little reliable scientific information relating to weight loss, energy requirements, changes in the concentrations of hormones and substrate in the blood, and organ metabolism from healthy humans during or after prolonged periods of starvation. Nonetheless, reasonable conclusions regarding the metabolic effects of starvation can be made.

Food deprivation contrasts with conditions in most industrialized nations today, where obesity is epidemic and threatening the health of an increasing proportion of the population. There is a strange relationship between starvation and obesity in which overweight people often starve to lose weight.

The cells of the various organ systems require a continuous supply of energy to function. The biochemical mechanisms that maintain the constant availability of fuel to support metabolic functioning are part of a complex system of fuel homeostasis. All mammals store energy when food is available and mobilize the stored fuels during fasting.

This article focuses on the last century of metabolic studies that range from fasting overnight to prolonged starvation. It covers topics such as the loss of weight and energy requirements, the release of stored fuels, the conversion of free fatty acids (FFAs), glycerol, and amino acids into ketone bodies and glucose by the liver, the selective utilization of fuels by the brain and muscle, the conservation of fuels by the kidneys and their promotion of gluconeogenesis, and maintenance of acid–base balance.

Background

The resting metabolic rate (RMR) among humans varies widely. However, general principles of bioenergetics for adult humans, with body masses varying over a sixfold range, have been defined. First, the RMR increases as body weight increases; however, the energy requirement per unit weight decreases as body weight increases. Thus, as the weight of a human increases from 40 to 160 kg, the RMR per kg body weight decreases from 30 to 15 kcal per kg per 24 h. Second, based on total body weight, the RMR of men of a given weight is greater than it is for women. This difference disappears when the RMR is based on fat-free mass (FFM), measured by densitometry. Nonetheless, there is a broad range of energy requirements for people of identical gender and body weight, which is independent of body fat content. It is thus understandable why the quantity and nature of fuels oxidized under near-standard conditions among tens of thousands of human subjects show discrepancies for specific body sizes. However, there is general agreement when the fundamental rules of bioenergetics
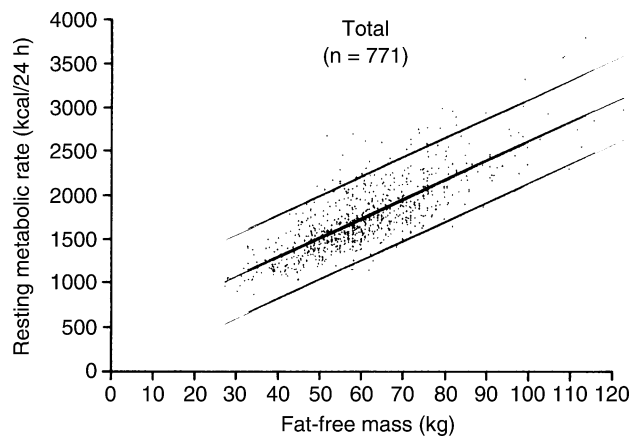
FIGURE 1 The relationship between fat-free mass, determined by hydrodensitometry, and the resting metabolic rate, as determined by indirect calorimetry, in 771 men and women.

are applied. This is illustrated in Figure 1 which shows the relationship between FFM and RMR in 771 men and women. It should be easier to understand why these data are diverse when the foregoing principles are applied to the analyses of the studies in Figure 1. A major advance in our understanding of the mechanisms of survival occurred with the advent of convenient and accurate methods for measuring β-hydroxybutyrate (β-OHB$^-$) and acetoacetate (AcAc$^-$). These two water-soluble, four-carbon compounds are synthesized from acetyl CoA derived from incomplete hepatic oxidation of long-chained fatty acids. Analytical techniques were developed in the laboratories of H.A. Krebs and coworkers in England. These fuels, along with acetone, are collectively referred to as ketone bodies and are immensely important for sustaining life during starvation.

Detailed, clinical studies of the biology of starvation began after 1965 when G.F. Cahill, Jr. and colleagues in the United States and elsewhere advanced the concept of fuel homeostasis in humans.

Obesity and Starvation, and Clinical Features of Starvation

It is somewhat paradoxical to note that mechanisms for maintaining the flux of fuels to provide the energy requirements to sustain life during starvation were derived primarily from normal and obese volunteers. These individuals were recruited from the general population or from hospitalized patients. Most of the volunteers and patients were housed in closely supervised, general clinical research units in hospitals. Those who underwent prolonged periods of starvation received water and, usually, salt tablets and multivitamins. All were studied after an overnight fast or during 3 to 63 days of total starvation.

A word about the courageous volunteers who made these important studies possible seems in order before we discuss the results. Prolonged starvation of the length mentioned above (up to 63 days) is a difficult physical and emotional experience and the volunteers required an enormous amount of support to complete these studies. Some developed headaches and postural hypotension during the first few days of total food deprivation. Everyone was preoccupied with hunger as starvation progressed. They did, however, usually develop a "team spirit" of trying to contribute the knowledge needed to understand what metabolic adaptations developed with weight loss during starvation. Understandably, some individuals became cantankerous and withdrew from the studies, and about one-half would manage to rarely sneak bits of food while under close observation, thus spoiling some of their data. Still, most of the volunteers struggled valiantly to comply with the research protocols. A sense of accomplishment pervaded the people working in facilities where these studies were conducted. We owe these individuals, whose names will never be known to us, a great debt of gratitude.

The compensated state of adult obese humans starving for experimental reasons is in sharp contrast to children, adults, and the elderly dying from inanition during famine or confinement in prisons. In the latter individuals, malnutrition is often complicated by other diseases. Understandably, their behavior varies from hostile selfishness to apathy and stupor. Superimposed bacterial, fungal, verminous, helminthic, and protozoal infections, with diarrhea, cough, and protuberant or contracted abdominal walls modify the clinical characteristics of starving people. Muscles can be "boggy or stringy." The loss of subcutaneous fat and muscle wasting accentuate the size of joints, especially the knees. The calf muscles are among the first to visibly shrink due to limited mobility, and the buttocks shows marked atrophy. Hair is broken and brittle and nails are split, pitted and spoon-shaped. The skin looks old and pale from anemia, but has regions of pressure hyperpigmentation, fistulae, and decubitus ulcers. Staring, sunken eyes search to identify anything, and the facial musculature is wasted and sometimes taut. Similar features are seen among patients hospitalized with severe malnutrition. Death occurs when about one-third to one-half of the lean body mass is lost in children, or lean and obese adults. In lean individuals practically all the body fat is mobilized before the irreversible state of malnutrition occurs.

The data from controlled starvation studies may not mimic many of the findings observed in patients from the harsh environments described above. Nonetheless, the data gathered from research centers provide the best available and reproducible data regarding starvation metabolism and fuel homeostasis. This data was collected largely over a period of about four decades.

As would be expected, the interpretation of some of the biomedical data changed with the accumulation of more information. An example of this is the changing attitude toward how the kidneys handle ketone bodies. As will be discussed in detail later, these ketone bodies play several metabolic roles and are essential for survival during prolonged starvation.

Metabolic Alteration Related to Starvation: Fed to Fasted to Fed

FED STATE

When food is presented to an average hungry human who has undergone a mere 12 h overnight fast, he/she can readily eat at a rate of 150 times the individual's resting metabolic rate. Only minor perturbations occur in the concentrations of fuels in the blood but the postprandial concentration of insulin, the major anabolic hormone, can increase 50-fold in obese individuals. It may take several hours to digest, absorb, oxidize, and store the excessive caloric content of the food. Some of the nutrients that are consumed in overabundance are transiently stored as glycogen and intracellular amino acids or protein during the early postabsorptive period. However, the dominant storage form of energy in the body is triglyceride in adipose tissue. A human can consume about as much as 10 000–12 000 kcal per day for many days. If this excessive caloric intake continues over a long period of time, the adipose tissue mass can become huge, accounting for as much as 500 000–1 000 000 kcal of energy present as triglycerides. However, the maximum body fat content is usually limited to about one-half of the total body mass. Lean body mass also increases to support the fat mass. Thus, a 150 kg (330 lb) man or woman can have a triglyceride (fat) mass of $\simeq 75$ kg (165 lb). The stored fat, along with some of the body protein and glycogen, is mobilized during semi-starvation or total fasting, usually when approximately 20–50 g of carbohydrate and 15–30 g of protein are eaten. However, during total, prolonged starvation among morbidly obese people, when amino acids are continuously mobilized from vital organs, (e.g. the heart), death may occur before the entire fat mass of the body is depleted. Thus, grossly obese humans subjected to prolonged starvation can die with body fat remaining.

STARVED STATE

In going from the fed state to the starved state the body shifts from storing fuels to mobilizing fuels. The factors that control production and utilization of specific fuels change rapidly during this transition. In contrast, starving volunteers enter a near steady-state of metabolism after $\simeq 2–3$ weeks of starvation, where day-to-day changes are minimal. During starvation, catabolism is tightly regulated and the supply of fuels from various depots adjusts to meet the body's energy requirements. Hepatic glycogen breakdown is brief (only 600 Kcal are available as hepatic glycogen in a 70 kg human), while proteolysis is continuous to supply gluconeogenic amino acid for energy and to maintain acid–base balance. The triglyceride present in adipose tissue is the major and persistent fuel reserve that supports metabolic process during starvation; it can account for as much as 90–93% of the body's energy needs during periods of prolonged starvation.

REFEEDING

During a period of semi-starvation, a constant awareness of hunger dominates the thought processes of humans. During refeeding after periods of starvation or undernutrition, emaciated people may ingest as much food as possible, sometimes vomiting because of excessive intake. This craving for food persists until weight gain occurs largely from the deposition of fat in white adipose tissue. Eventually, weight gain plateaus and weight loss may occur until the body weight of an individual reaches an equilibrium that is at or near their weight before semi-starvation. On the other hand, after total starvation, the obese volunteers who become outpatients experience mild abdominal pain and may pass fecaliths. During the first 2–3 weeks of refeeding, these individuals routinely retain fluid and become edematous and are encouraged to maintain a low caloric intake until diuresis develops. However, the success rate for preventing the re-accumulation of body fat is low. Some clinical investigators hypothesize that the body possesses a "set-point" or has a body fat "lipostat." This is questionable because it is just as likely that individuals eat more than they need to replace body fat, and gain additional weight predominantly as adipose tissue.

Weight Loss, Body Composition, and Energy Requirements

The measurement of body weight is usually obtained accurately and easily. However, edema can complicate the interpretation of this simple measurement. Measuring body composition and energy requirements has several shortcomings. FFM or "active tissue" (lean body mass) is a conceptualized mass and difficult to measure accurately. This is because there are differences in the loss of mass among different organ systems in both the lean and obese individuals during starvation. The skeletal muscles and subcutaneous fat lose the greatest

quantity of weight. Among organs, the liver and spleen lose at least 50% of their usual mass. The gastrointestinal tract becomes atrophic; the diameter of the lumen of the small intestine becomes reduced to $\simeq 50\%$ of its normal diameter, and the villae become flattened. Gut motility also decreases. The heart decreases in size and the pulse rate and blood pressure decrease. The loss of bone mass is less than the loss of adipose tissue and skeletal muscle during experimental periods of starvation.

The energy needs of the body are determined by the sum of the metabolic requirements of various tissues (e.g. brain, liver, muscle, heart, adipose tissue, spleen, leucocytes, etc.) that depend on the mass and activity of individual organs. During the resting state, there is a 50- to 100-fold variation per unit mass for energy requirements among different tissue types (e.g. brain and adipose tissue). The energy requirements of skeletal muscles can change 12-fold in the transition from the resting state to strenuous exercise.

The proposed "metabolic efficiency" during starvation has long been claimed in lean people, but it has been difficult to demonstrate in starving, obese humans. Studies by Owen and colleagues differ from those of Leibel and associates in this regard. There is no question that as body size decreases in lean and obese humans; their metabolic requirements also decrease. Nonetheless, metabolic adaptation to starvation, where the energy requirements per unit of measurement decrease out of proportion to changes in unit of measurement, remains a perplexing issue. This is partly due to the method by which the data are calculated or expressed. The nonresting metabolic rate is more difficult to measure and is more inaccurate than estimates of the resting energy expenditure. Therefore, if there is an adaptive metabolic efficiency induced by food deprivation, it should be demonstrable by the (more) standardized RMR measurement. However, this has not been consistently demonstrated in all studies. Therefore, the increase in efficiency may be more apparent than real.

Thus, there remains an inadequately defined relationship between weight loss and metabolic rate. Contradictory data have been presented as to whether there is a decreased RMR per unit mass in adult men produced by prolonged semi-starvation. However, the normal RMR varies widely from low, normal or high values for humans of identical weight, gender and age. Therefore, it is not surprising that the literature on RMR for humans, based on the standard of measurement, is inconsistent. The differences among the various thyroid hormones, T_4, T_3, and rT_3, in lean and obese people subjected to protracted food deprivation may influence the RMR. It is well known that after about a week of total starvation, when the diuretic phase of starvation is completed, body weight loss was only $\simeq 0.32\%$ per day. Much later it was reported that the RMR of obese, starving volunteers, based on oxygen consumption

per kg fat-free mass per day, and corrected for urinary nitrogen loss during starvation, is constant. This was truly disheartening for obese people who desperately wanted to lose body fat. It also showed clinical investigators how limited total starvation really is as a method for weight reduction for morbidly obese patients. Nonetheless, useful knowledge was gathered about starvation from these noble volunteers.

The Nature and Quantity of Fuels Oxidized during Starvation

Indirect calorimetry closely reflects the nature and quantity of the fuels oxidized. The results obtained using this method are influenced by dietary intake, total metabolic requirements, and state of physical activity. When lean individuals fast overnight, their nonprotein respiratory quotient (npRQ) is $\simeq 0.84$. This matches the previous day's intake of 12–20% protein, 40–45% fat, and 40–45% carbohydrate. When obese people eat the same balanced diet, but with a greater caloric intake to match their greater metabolic needs, the npRQ is significantly lower and decreases as body weight increases. Figure 2 shows the trend in the nature and quantity of fuels oxidized after an overnight fast in lean and obese humans. The data show that as weight and body fat increase, the npRQ falls. The greater the RMR, the greater the quantity of fat that is mobilized to meet the energy requirements of fasting humans. Due to body energy demands, obese humans shift into an accelerated rate of fat mobilization. However, this tendency to mobilize and oxidize stored fat becomes more obvious in
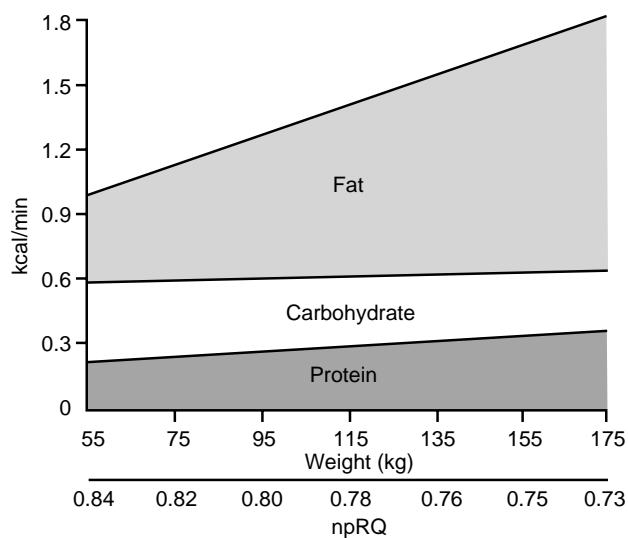
FIGURE 2 The nature and quantity of fuels oxidized by humans after a 12–14 h fast. Fat oxidation increases as body size increases.

both lean and obese people as fasting is prolonged. As starvation is extended to 2–3 days and beyond, lipids furnish 90–93% of the resting energy requirements in both lean and obese humans. These changes in the nature and quantities of fuels oxidized by tissues such as muscle and brain, spare carbohydrate (glucose) from oxidation; this provides the physiologic reasoning for insulin-resistance or insensitivity when carbohydrate oxidation must be curtailed for survival during food deprivation.

The absolute minimal rates for the major fuel utilization before death occurs have not been defined. The minimum requirement for fat and protein oxidation in grams/day/kg FFM has been approximated as follows: $\simeq 2.98 \pm 0.21$ g per kg per day FFM for fat; $\simeq 0.52 \pm 0.10$ g per kg per day FFM for protein. Glucose is derived from glyceride-glycerol, amino acid, and recycled lactate and pyruvate. About 1.91 ± 1.04 g per kg per day FFM are synthesized. However, most of the glucose that is synthesized during starvation is recycled from lactate and pyruvate (the Cori cycle), glycerol, alanine, and glutamine. The quantity of glucose that undergoes oxidation to CO_2 and H_2O is primarily determined by the catabolism of glucose in the central nervous system. This quantity is not closely related to the FFM because the size of the brain is unrelated to the FFM. The terminal oxidation of glucose is $\simeq 40–45$ g per day, which is derived primarily from gluconeogenesis from amino acids (alanine and glutamine) and glycerol.

In obese humans, during starvation there is no decrease in the metabolic requirement per unit mass, based on measurements of oxygen consumption. However, when the npRQ decreases below 0.7, the theoretical minimum for fat oxidation, there is a difficulty in extrapolating the respiratory exchange rates of O_2 and CO_2 and urinary nitrogen excretion into energy expenditure. The production of CO_2 falls disproportionately to O_2 consumption, creating a mysterious situation characterized by a calculated npRQ of 0.62–0.65. This unusual finding cannot be eliminated by correcting for ketonuria. However, it has been postulated that some of fatty acid released from triglyceride stores may undergo desaturation before being recycled to storage depots. The process of desaturation consumes oxygen and produces heat, but releases no carbon dioxide. Desaturated-FFAs in the blood have been identified in starving animals. If this process occurs in humans, the respiratory quotient (RQ) should rise before death, when unsaturated fatty acids are released and oxidized for energy. The RQ does rise in animals and humans during the pre-mortal period of starvation. This phenomenon has not been fully explained but is usually attributed to the mobilization and oxidation of amino acids. It is just as likely that it is due to oxidation of desaturated fatty acids.

Changes in the Concentration of Substrates and Hormones in the Blood

The concentration of glucose in the blood after an overnight fast is $\simeq 4.5$ mM; this value falls to $\simeq 3.6$ mM after 3 days of starvation. Thereafter, the concentration of glucose in the blood reaches a plateau. Anaerobic glycolysis in tissues such as muscle, brain, red blood cells, and kidney medulla converts glucose into lactate and pyruvate, and the liver extracts lactate and pyruvate to produce glucose. The concentration of lactate in the venous blood is less than 1.0 mM, while the concentration of pyruvate is less than 0.1 mM and is constant in the resting state. The concentration of FFA in the plasma rises from $\simeq 0.6$ mM, to $\simeq 1.4$ mM during the first few days of starvation. Thereafter, the concentrations of these fuels remain elevated and relatively constant. The rate of uptake and disposal of FFA is largely determined by its concentration in the blood. Insulin regulates the release of FFA from adipose tissue, primarily by influencing lipolysis, and thus the availability of FFA. The concentration of glycerol in the blood derived from lipolysis is less than 0.1 mM after an overnight fast and rises to 0.15 mM on the third day of starvation. Thereafter, glycerol remains constant because uptake, primarily by the liver, to synthesize glucose matches its release by adipose tissue. The concentration of triglycerides in the blood is less than 1.0 mM and remains constant during fasting.

A characteristic of fasting in humans is the presence of increased concentrations of ketone bodies in the blood (Figure 3) and urine. These water-soluble, short-chained compounds are synthesized primarily in the liver from the acetyl CoA derived from fatty-acid oxidation and serve as alternate fuels for tissues such as the brain. Ketone bodies replace glucose as the dominant energy source for the brain during starvation. Physical activity augments catabolism during starvation and promotes hyperketonemia. There are no other fuels in the blood that can change as markedly as the concentration of ketone bodies during starvation. This is partly due to the low concentration of ketone bodies in blood during the postprandial period, after a mixed meal containing adequate carbohydrate. After an overnight fast, lean people have blood $AcAc^-$ and β-OHB^- concentrations of $\simeq 0.05$ mM, while this value is slightly higher in obese individuals. Acetone is virtually absent after an overnight fast, unless the individual is regularly consuming a high-fat diet. There is an exponential rise in the concentration of $AcAc^-$ and β-OHB^- in the blood during starvation, until new steady levels develop. During the first 3 days of fasting, the $AcAc^-$ concentration in the blood increases to $\simeq 1.5$ mM, while β-OHB^- continues
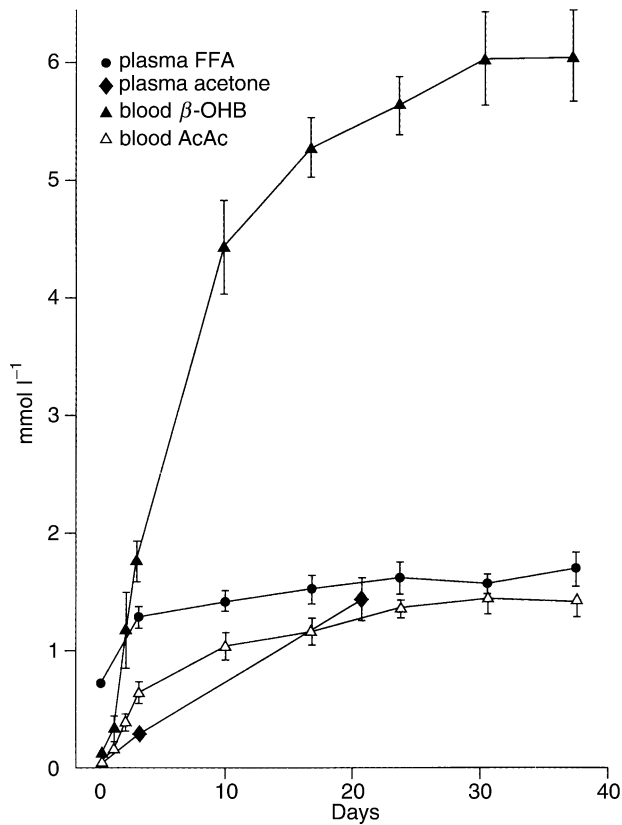
FIGURE 3 Changes in the concentration of ketone bodies and free fatty acids (FFA) during starvation.

to rise to $\simeq 4.5$ mM after day ten of starvation. AcAc$^-$ plus β-OHB$^-$ reaches a plateau at $\simeq 6$–8 mM after 18 days of total starvation (water, salt, and vitamins were provided). The smell of acetone halitosis becomes evident after $\simeq 2$–3 days of starvation when the blood concentration is $\simeq 0.25$ mM. Acetone slowly increases to $\simeq 0.35$ mM after 21 days of starvation. A minimum estimate of the change in the concentration of ketone bodies in the blood from an overnight fast to 18 or more days of total starvation is 75–160 fold (0.1–0.2 mM to 15.0–16.0 mM). Blood AcAc$^-$ and β-OHB$^-$ are anions and the hydrogen produced during ketogenesis combines with bicarbonate to form CO_2 and water. CO_2 is exhaled thus decreasing the concentration of bicarbonate in the blood; this creates the typical anion gap of ketonemia. Blood pH also falls appropriately (7.4 to 7.3), and a mild metabolic acidosis of starvation becomes evident.

The concentration of total amino acids in the plasma declines slowly from an overnight fasting value of $\simeq 4.6$ mM to $\simeq 3.7$ mM after 40 days of starvation. However, there are four basic patterns of change among the amino acids during total starvation, represented by alanine, glycine, valine, and glutamine. Alanine rapidly decreases to $\simeq 30\%$ of its overnight concentration. Glycine does the opposite; it rises nearly threefold,

before reaching a plateau. The concentration of valine doubles in the plasma during the first 7–10 days of starvation, and then slowly and progressively falls to a value below its overnight fasting value. Venous glutamine, the predominant plasma amino acid, remains relatively stable during 5–6 weeks of fasting.

Blood urea nitrogen mimics the concentration of alanine in the blood. The concentration of creatinine in the blood slowly drifts downward, reflecting the decrease in muscle mass.

Hormonal Changes

Insulin is the main hormone that controls the anabolic processes that maintain fuel homeostasis in humans. Its secretion is primarily regulated by the blood glucose concentration, but the levels of amino acid, FFA and ketone bodies also modulate insulin release by the β-cells of the pancreas. The influence of insulin on the metabolism of the major organs is readily demonstrated in adipose tissue, skeletal muscle, liver, white blood cells, and other tissues. Insulin also inhibits glucagon secretion, a significant counter-balance catabolic hormone.

The concentration of insulin in the blood parallels the changes in the levels of glucose. The fall in the concentration of insulin diminishes some of its inhibitory effects on peripheral proteolysis and lipolysis. In the starved state, the concentration of insulin falls, resulting in a decrease in the rate of uptake of glucose, amino acids, and fatty acids in peripheral tissues and a subsequent increase in the rates of gluconeogenesis, lipolysis, and proteolysis. After an overnight fast, the concentration of insulin in the portal blood is 2–3 times greater than the peripheral venous concentration. After 2 to 3 days of starvation, the portal-peripheral insulin concentration gradient is small. During starvation, there is still a high enough concentration of insulin in the blood to limit the maximal rates of glycogenolysis, gluconeogenesis, and ketogenesis. In addition, insulin has at least two indirect effects on hepatic glucose and ketone body production. Insulin decreases the delivery of gluconeogenic precursors and FFA from the extra-hepatic-tissue stores to the liver. As the concentration of insulin in the mesenteric blood decreases, the inhibitory effect of insulin on the β-cells of the pancreas curtails the secretion of glucagon. Thus, the low concentration of insulin in the blood promotes glucagon secretion.

Insulin has dual roles in controlling the release of glucose and ketone bodies from the liver (and kidney cortex), and the release of amino acid from muscle and FFA from adipose tissue. Glucagon has the opposite effect. It augments hepatic (and renal) gluconeogenesis and ketogenesis, and promotes peripheral lipolysis and proteolysis. A relatively low blood insulin concentration

and relatively high glucagon concentration creates a blood insulin/glucagon ratio that promotes fuel availability. These pancreatic hormones serve collectively to finely balance the fuel needs of various tissues. The catabolic role of glucagon is augmented by catecholamines, glucocorticoids, and growth hormones. Fuel homeostasis is critical for survival so there are multiple levels of hormonal control of this process.

Appetite Regulation

During starvation there are a number of factors that regulate the control of appetite. Ghrelin, an orexigenic hormone, primarily secreted by the stomach and duodenum, normally increases in the blood before meals and falls after meals. Ghrelin is thought to stimulate hunger, increase body weight, and decrease metabolic rate. Ghrelin rises with semi-starvation but its physiologic roles regarding hunger need further elucidation. Adipose tissue produces a satiety hormone, leptin, which suppresses appetite. The leptin concentration in the blood falls in parallel with insulin during the first three days of starvation. As body fat decreases during starvation, the blood leptin concentration decreases. A fall in the levels of leptin in the blood increases the hypothalamic concentration of neuropeptide Y, greatly stimulating appetite. However, the control of appetite is not a simple process, but rather involves the interaction of a number of neuroendocrine systems. There is no single hormonal neuropeptide that controls hunger in a starving human. Prolonged starvation causes exaggerated hunger, a psychosomatic state that dominates the conscious mind.

Insulin plays an important role in the regulation of appetite. Acting together with leptin, insulin circulates at concentrations proportional to the body fat content. Both hormones enter the CNS and bind to specific receptors in neurons involved in controlling food intake. The administration of leptin and insulin directly to the hypothalamus decreases appetite and suppresses food intake. It is probable that leptin is the more important of the two hormones since leptin deficiency, not a lack of insulin, results in obesity. The mechanism by which these hormones influence the CNS is the subject of intensive study. A detailed discussion of this area is beyond the scope of this article. However, it should be clear that during starvation, as the mass of adipose tissue decreases and the secretion of insulin by the pancreas is depressed due to a lack of food intake, the appetite centers of the brain would be stimulated by the lack of satiety signals such as leptin, insulin, and ghrelin.

Adipose tissue secretes a number of other regulatory peptides, besides leptin, that control fuel deposition. These include adiponectin, resistin, adipsin, acetylation-stimulating protein, angiotensinogen, and cytokines. The most abundant adipose specific hormone is adiponectin. The concentration of adiponectin in the blood is positively correlated with insulin sensitivity. Adiponectin promotes glucose uptake in muscle and fatty tissue. Its concentration in plasma is inversely related to adipose tissue mass, especially abdominal fat mass. The levels of adiponectin in the blood fall during semi- and total starvation, when glucose intolerance has been documented. Resistin is another protein secreted into the blood by adipose tissue, but it has an effect that is opposite to that of insulin. Its concentration in the blood of humans after prolonged starvation has not been defined. The behavior of adipose tissue cytokines in humans during weight reduction caused by semi-starvation is under investigation.

Biochemical Changes during Starvation

Starvation induces a number of specific biochemical changes that permit an individual to survive in the absence of food for a prolonged period of time. These include alterations in glucose synthesis, fatty acid utilization by specific tissues, and the preservation of nitrogen to insure the maintenance of lean body mass. Starvation is a catabolic state in which fuel reserves are mobilized to support metabolic needs. To understand the metabolic adaptations to starvation, it is first necessary to describe the available energy reserves of a human as starvation begins. It has been estimated that a 70 kg human contains 144 000 kcal as fat (largely as triglyceride), 24 000 kcal as protein (lean body mass), and only 900 kcal as carbohydrate (liver and muscle glycogen). The relatively low level of carbohydrate reserve in humans necessitates a number of biochemical adaptations to insure the appropriate energy supply to specific tissues. For example, the brain uses $\simeq 600$ kcal of glucose each day; this means that the total carbohydrate reserve is depleted in about one day of starvation! It is clear that mechanisms must exist for the use of fuels other than glucose (fatty acids and ketone bodies) by tissues such as the muscle and the brain, to insure survival during starvation. This process is called "fuel sparing."

CONTROL OF FATTY ACID RELEASE DURING STARVATION

Since fat is the major caloric reserve in humans, the regulation of mobilization of FFA from adipose tissue during starvation is a critical factor for survival. The breakdown of triglyceride via lipolysis is under the control of a number of hormones, including

epinephrine, glucagon, glucocorticoids, and insulin. The concentration of insulin in the blood falls by $\simeq 50\%$ after 3 days of starvation. At the same time there is an increase in the concentration of glucagon. This change in the insulin to glucagon ratio is a critical factor in the increase in lipolysis and the enhanced rate of hepatic and renal gluconeogenesis that occurs during starvation. Insulin is the major antilipolytic hormone and its decrease is the major factor in insuring the increased availability of the fatty acids needed for energy metabolism during starvation.

Adipose tissue contains a hormone-sensitive lipase that is sensitive to the state of its phosphorylation. An increase in the concentration of cAMP in the tissue (caused by a rise in epinephrine, glucagon and growth hormone, and a drop in the level of insulin) activates protein kinase A (PKA) that in turn phosphorylates, and thus activates the hormone-sensitive lipase; this results in the breakdown of triglyceride and generation of FFA. Insulin counters this process, partly by decreasing the levels of cAMP in the adipose tissue, and by increasing the activity of a phosphoprotein phosphatase that dephosphorylates the hormone-sensitive lipase, thereby inactivating it. There is also evidence that insulin increases the activity of a phosphodiesterase in the adipocyte, causing a fall in the concentration of cAMP in the tissue. The net result of prolonged starvation is an enhanced rate of FFA release from adipose tissue that is used as a fuel by a number of tissues.

About 50% of the plasma FFA presented to the liver during starvation is extracted and $\simeq 50{-}80\%$ of the extracted fatty acids undergo partial oxidation to synthesize equal quantities of $AcAc^-$ and $\beta\text{-}OHB^-$. Most of the remaining quantity of FFA extracted by the liver is converted to triglycerides and recycled to adipose tissue as VLDL. It is interesting that such a large fraction of the FFA released by lipolysis is re-esterified back to triglyceride in adipose tissue or in the liver and other tissues, and returned to the adipose tissue for the resynthesis of triglyceride. Fatty acid recycling via this so-called triglyceride–fatty acid cycle can account for as much as 60% of the fatty acid released after 3 to 4 days of starvation in humans. The synthesis of triglyceride requires the generation of 3-phosphoglycerol, usually from glucose. During starvation, when glucose is at a premium, the 3-phosphoglycerol is generated from pyruvate, lactate, and amino acids via an abbreviated version of gluconeogenesis termed *glyceroneogenesis*. The triglyceride/fatty acid cycle is a "futile cycle" that uses energy for the synthesis of triglyceride (6 molecules of ATP per molecule of triglyceride synthesized), so it must have a role in preserving the FFA that was released by adipose tissue to be used later as a fuel by peripheral tissues.

PROTEOLYSIS AND AMINO ACID METABOLISM

After a meal containing proteins, amino acids largely escape hepatic extraction and are carried by the blood to extrahepatic tissues. Insulin promotes the active transport of amino acids into cells, primarily skeletal muscle. Normally, amino acids are in a state of flux; they are precursors for protein synthesis and then appear as free amino acids after protein breakdown. Within the first day of starvation, however, protein catabolism dominates the metabolic flux. As starvation progresses, relatively more proteolysis occurs and amino acids are mobilized from protein depots. The dominant fate of the carbon skeletons and the amino and amide groups derived from the breakdown of amino acids in muscle, is conversion to alanine and glutamine that is mobilized from muscle and other depots and transported to liver and kidney to be utilized to synthesize glucose, urea, and ammonia.

Protein catabolism in the splanchnic tissues is somewhat less responsive to insulin than it is in skeletal muscles. In the periphery, the basal concentration of insulin that is present during starvation, limits proteolysis. Nonetheless, glucose must be continually synthesized from amino acids by the liver and kidney cortex during starvation; amino acids from muscle protein are a major source of carbon for gluconeogenesis. The first proteins degraded during starvation are the digestive enzymes secreted from the stomach, pancreas, and small intestine. The liver also loses various enzymes needed to process incoming nutrients into plasma protein, e.g., albumin. Thereafter, the largest protein mass of the body, striated muscle, begins to be drained, not only of intracellular proteins, such as enzymes, but also the contractile elements. The disintegrating muscles can easily be seen as skeletal muscle atrophy during prolonged starvation in humans.

The control of proteolysis in muscle is a complex process. Insulin retards proteolysis and enhances protein synthesis. In addition, metabolic acidosis promotes protein breakdown to insure the generation of ammonia to titrate the acidity of the tubular urine. The low plasma insulin concentration and mild metabolic acidosis of starvation are conducive to proteolysis. The best characterized pathway for protein catabolism is an ATP-independent system of acid proteases (cathepsin) and hydrolases contained in cellular lysosomes. In addition, there are calcium-dependent proteases, as well as cytosolic ATP-dependent and independent pathways. The most important muscle proteolytic system employs the ubiquitin proteasome pathways.

Amino acids generated from proteolysis undergo deamination and/or deamidation before entering the

tricarboxylic acid (TCA) cycle to participate in further transformations to alanine or glutamine which efflux from cells. Alanine plus glutamine account for $\simeq 80\%$ of the amino acids released from muscle after prolonged starvation, and are the principal vehicles for transporting nitrogen from peripheral tissues to liver and kidney. These two amino acids together account for only 10% of the composition of skeletal muscle protein, but they provide the majority of the carbon for the synthesis of glucose by the liver and kidney during starvation. Thus, there is a net conversion of other amino acids into alanine and glutamine in the muscle. Alanine is the major nitrogenous compound released from muscle and extracted by the liver for gluconeogenesis and ureagenesis. There is a "glucose–alanine" cycle between muscle and liver. Glutamine is the other amino acid released in high levels by muscle, but it is extracted from the blood by the kidney cortex to produce glucose and ammonia. During the catabolism of muscle, branched-chain amino acids provide the amino groups to α-ketoglutarate to form glutamate via glutamate dehydrogenase; glutamine is derived from glutamate and ammonia by the action of glutamine synthase.

The TCA cycle is the key metabolic pathway for energy production: acetyl CoA is oxidized to carbon dioxide. During starvation, the TCA cycle has another fundamental role in metabolism. Amino acids are catabolized to 4- (aspartate, asparagines, phenylalanine, tyrosine, methionine, isoleucine, and valine) and 5- (glutamine, glutamate, histidine, proline, and arginine) carbon intermediates that enter the TCA cycle by a process termed *anaplerosis*. The TCA cycle cannot serve as a carbon sink; therefore, the carbon atoms that enter the TCA cycle must leave the TCA cycle by a process called *cataplerosis*. Anaplerosis and cataplerosis are reciprocal reactions that must be balanced. These reactions have been reviewed in detail elsewhere in this volume.

During starvation, more fuels are released from storage depots or synthesized in liver and kidney than are oxidized to generate energy. About 60% of the fuels that enter the bloodstream are recycled, e.g., FFA, glucose and amino acids, especially alanine and glutamine.

KETOGENESIS

The liver not only stores glucose as glycogen, it also converts fuels such as FFA, amino acids, lactate, pyruvate, and glycerol to glucose and ketone bodies; it also detoxifies the ammonia from amino acid breakdown by converting it to urea. After an overnight fast, hepatic glycogenolysis, gluconeogenesis, and ketogenesis provide $\simeq 50\%$ of the total energy-yielding fuels for the body in the resting state. Lipolysis of triglyceride in adipose tissue supplies FFA and glycerol; these substrates become precursors for ketone body (fatty acids) and glucose (glycerol) syntheses. Amino acids released primarily by skeletal muscles complement glycerol as gluconeogenic precursors. As fasting is prolonged, the kidney cortex also contributes to fuel homeostasis by conserving substrates and sharing the gluconeogenic role with the liver.

Ketone bodies are the only fuels synthesized in the body that do not recycle. Unlike FFA, amino acids and glucose, ketone bodies are either oxidized or excreted in the urine and/or the breath (acetone). A small amount of acetone can be converted to glucose. β-OHB$^-$ and AcAc$^-$ are synthesized in the liver (and kidney) mitochondria by the following reactions. Two molecules of acetyl CoA condense head-to-tail to form acetoacetyl CoA; this reaction is catalyzed by acetoacetyl CoA thiolase. Another molecule of acetyl CoA is joined by β-hydroxy-β-methylglutaryl CoA (HMG CoA) synthase to form HMG CoA, also generating a hydronium ion (H$^+$). HMG CoA lyase cleaves HMG CoA into AcAc$^-$ and acetyl CoA. AcAc$^-$ is then reduced to β-OHB$^-$ by β-hydroxybutyrate dehydrogenase. Acetone and CO_2 are formed from the degeneration of AcAc$^-$. Ketone bodies are synthesized from the acetyl CoA that is derived from the β-oxidation of fatty acids in the mitochondria. A small quantity can be synthesized from ketogenic amino acids during starvation. In addition, acetoacetyl CoA can be formed from FFA and cleaved to AcAc$^-$ in the kidneys. The hepatic production of β-OHB$^-$ and AcAc$^-$ are about equal. During hyperketonemia, the concentration of β-OHB$^-$ in the blood is $\simeq 3$ times greater than AcAc$^-$. This is due to the preferential removal of AcAc$^-$ or conversion of AcAc$^-$ to β-OHB$^-$ by skeletal muscle. The brain removes β-OHB$^-$ and AcAc$^-$ in a concentration-dependent manner.

The metabolism of ketone bodies during starvation is a critical element in the control of fuel homeostasis in humans. The demonstration that β-OHB$^-$ and AcAc$^-$ could serve as major fuels for the metabolism of the brain during starvation provided the information needed to evaluate the roles of fatty acid oxidation, amino acid mobilization, glucose conservation, and urinary nitrogen excretion during prolonged starvation. From this research it became clear that the abundance of FFA stored in humans provides a substantial reserve for the synthesis of ketone bodies. The metabolism of ketones by the brain during starvation greatly limits the need to use amino acids to make glucose to support the metabolism of this tissue. This is an important example of "fuel sparing."

There are few synthetic processes that are quantitatively as large as the daily rate of ketogenesis during starvation. After an overnight fast, hepatic ketogenesis amounts to $\simeq 10$ g per day. After 2–3 days of starvation the liver synthesizes over 100–150 g per day of ketone bodies. Over this period, the average resting human

oxidizes a minimum of $\simeq 3$ g of fat per kg FFM per day. Thus, a large, but not obese adult man, weighing 80 kg, with a body composition of 80% FFM (64 kg) and 20% fat mass (16 kg) oxidizes a minimum of $\simeq 192$ g of fat per day. About 60 g of the FFA derived from the lipid stores undergo β-oxidation in the liver to yield an estimated 113 g per day of ketone bodies. Ninety percent of these water-soluble fuels undergo terminal oxidation, primarily by the brain and muscle. Most of the remainder is excreted in the urine. It is interesting to note that about half of the molecular weight of AcAc$^-$ and β-OHB$^-$ is oxygen. Thus, if 10–12 g of ketone bodies are excreted in the urine, only 5–6 g of the carbon skeleton is derived from stored triglycerides. Since β-OHB$^-$ and AcAc$^-$ are excreted with near equimolar quantities of NH$_4^+$, ketosis is an energetically cheap way to excrete nitrogen (the synthesis of urea requires 4 molecules of ATP per molecule of urea). Finally, for each gram of nitrogen lost in the urine $\simeq 3.57$ g of glucose is synthesized.

GLUCONEOGENESIS

The *de novo* synthesis of glucose from nonhexose precursors (gluconeogenesis) occurs in the liver, kidney cortex, and perhaps to a minor extent in the small intestine. The precursor molecules for gluconeogenesis, and the organs involved in this process, change as starvation progresses. After an overnight fast, the liver is the primary, if not the sole, organ that makes a net contribution to glucose in the blood. The kidney extracts and releases glucose after an overnight fast but its overall contribution is minimal. After 3 days of starvation, renal gluconeogenesis increases and by 10–18 days of total starvation the kidney contributes $\simeq 50\%$ of the glucose added to the blood. However, during prolonged starvation, total glucose added to the blood from gluconeogenesis is only $\simeq 50\%$ of what it was after an overnight fast when gluconeogenesis is greatly complemented by hepatic glycogenolysis. In addition, during prolonged starvation only half of this glucose contributed to the blood is oxidized to CO_2 and H_2O. This is the result of the "sparing" of glucose by tissues that use more abundant fuels such as ketone bodies.

A schematic representation of the quantities of ketone bodies and glucose contributed to the blood by the liver and kidneys is shown in Figure 4. In general, glycogenolysis decreases as gluconeogenesis increases, and glucose production decreases as ketogenesis increases. An interesting aspect of gluconeogenesis during starvation is the relationship between renal gluconeogenesis and ammoniagenesis. Glutamine produced by the muscle is used by the kidney cortex as the major source of carbon for gluconeogenesis. The amino and amide groups are removed from glutamine to form α-ketoglutarate,
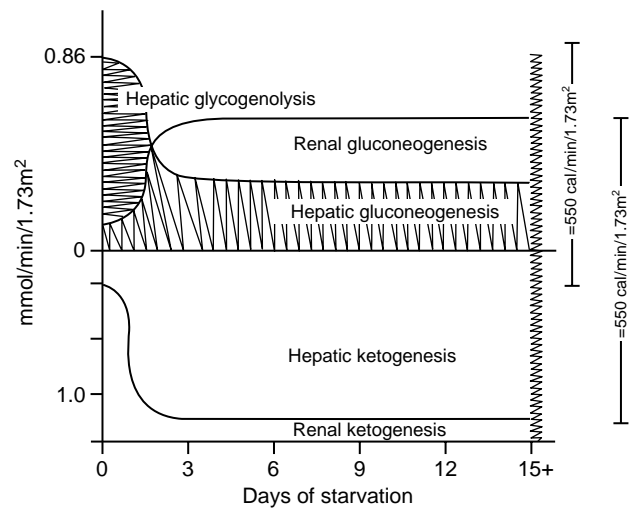
FIGURE 4 Schematic representation of fuel delivery by the liver and kidney during starvation. Hepatic gluconeogenesis, glycogenolysis and ketogenesis, and renal gluconeogenesis are expressed as mmol/min and converted to calories per 1.73m^2 body surface area. Hepatic glycogenolysis is rapidly dissipated during the first 1–3 days of fasting. The liver and kidney share the role of gluconeogenesis as starvation progresses. Hepatic ketogenesis becomes a dominant mechanism for supplying fuels to the blood after liver glycogenolysis is curtailed. Renal ketogenesis may contribute to ketonuria, but not to ketonemia.

which then enters the TCA cycle and is converted to glucose via renal gluconeogenesis. The kidney also has the ability to synthesize glucose from lactate, pyruvate, glycerol, amino acids and other precursors, but only glutamine has been closely associated with net renal gluconeogenesis.

The decreased rate of glucose production during starvation is linked to the increased ketone body production. Hyperketonemia is accompanied by ketonuria and ammoniagenesis. Renal NH$_4$ production is coupled to renal gluconeogenesis, while the release of NH$_3$ by the kidney is linked to hepatic urea production. In summary, gluconeogenesis, ketogenesis, ammoniagenesis, acid–base balance, and ureagenesis are all tightly interdependent during starvation.

Urinary Nitrogen Excretion

Total, daily urinary excretion undergoes an asymptotic, progressive decrease during starvation (Figure 5). Humans that consume a diet with 100 g protein (16.0 g nitrogen) excrete 15.5 g of nitrogen in the urine. Most of the nitrogen (85%) is excreted as urea. Ammonium, uric acid, and creatinine account for most of the remaining urinary nitrogen. Starvation initially induces an acute decrease in total nitrogen excretion. After 3–6 days of starvation, total urinary nitrogen decreases gradually, so that after 36 days of starvation
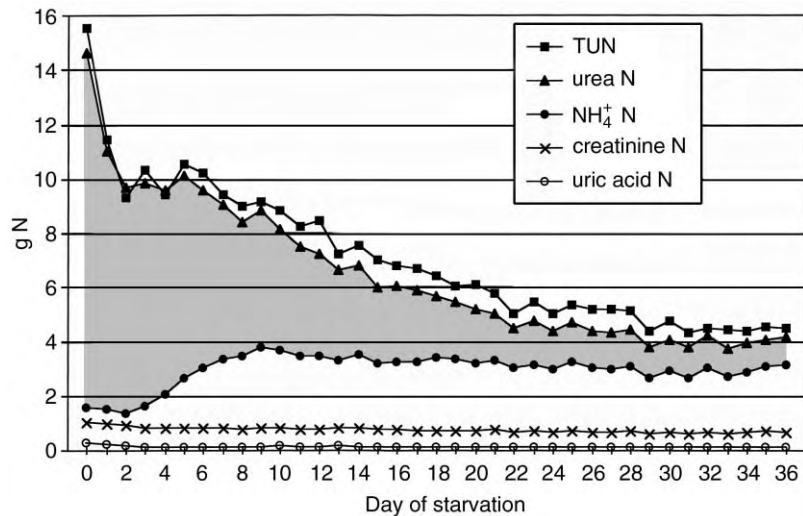
FIGURE 5   The total quantity and various components of urinary nitrogen excretion during prolonged starvation when five men and five women volunteers were given water, salt, and vitamins. TUN is an abbreviation for total urinary nitrogen. The nitrogen contents of urea, ammonium, creatinine, and uric acid are displayed.

the nitrogen excretion is $\simeq 4.6$ g per day. The biggest decrease is in the exponential decay in urea nitrogen excretion. The excretion of ammonium ion (NH_4^+) parallels the excretion of $AcAc^-$ and β-OHB$^-$; it peaks at $\simeq 120$–140 mmol per day at 6–9 days of starvation. After 15–18 days of drinking only water, NH_4^+ overtakes urea as the dominant urinary nitrogen compound. Urinary NH_4^+ gradually decreases as starvation is prolonged. NH_4^+, $AcAc^-$, and β-OHB$^-$ urinary excretion rates are sensitive to small quantities of ingested glucose. In early studies of metabolism during starvation, subjects were allowed to drink beverages that contained carbohydrates. This distorted both urea and NH_4^+ urinary excretion. The excretion of creatinine decays in a linear manner that is dependent upon diminishing body muscle mass that occurs as starvation progresses. The output of uric acid in the urine drops acutely during the first 4 days of starvation as ketonuria rises. Thereafter uric acid nitrogen excretion plateaus at $\simeq 2$ mmol per day (120 mg per day).

The rate of urea excretion provided the first clue that the concept that brain metabolized only glucose for energy had to be modified. Research in the first half of the twentieth century established that for each gram of urinary nitrogen excreted $\simeq 3.5$ g of glucose could be synthesized from amino acids. Excluding the glycerol that is present in triglycerides, only trivial quantities of glucose could be derived from triglyceride. The brain requires $\simeq 100$–125 g of glucose or equivalent energy sources daily. Plasma FFAs cannot pass the blood–brain barrier to provide a source of energy. In 1967 it was shown that $AcAc^-$ and β-OHB$^-$ were the primary fuels for brain metabolism during starvation. Therefore,

gluconeogenesis is clearly important during starvation, but the quantity of glucose formed is less than the quantity needed to provide energy for the brain. However, it should be pointed out that tissues such as the red blood cells, kidney medulla, and lens of the eye use glycolysis for energy production. The lactate and pyruvate generated by these tissues are recycled to the liver for the production of glucose.

Fuel Consumption during Starvation

While the liver and kidney release glucose into the blood during starvation, only the liver makes a net contribution of ketone bodies. However, more urinary excretion of $AcAc^-$ and β-OHB$^-$ occurs than is extracted across the renal vascular beds, suggesting that the kidneys synthesize some of the ketone bodies excreted in the urine. Ketonuria augments ammoniagenesis. The deamination of glutamine to form α-ketoglutarate generates two molecules of NH_4^+, which are released into the tubular urine to maintain electroneutrality and acid–base balance. As mentioned above, the α-ketoglutarate produced in this process enters the TCA cycle via anaplerosis, and is subsequently converted to glucose that is released into the blood. The kidney can also convert lactate, pyruvate, other amino acids, and perhaps glycerol into glucose. Clearly the kidney shares the role of glucose production with the liver, but the kidneys make a net contribution to blood only after several days of starvation.

Early in starvation, the small intestine may provide gluconeogenic amino acids to the splanchnic (portal) system, but by 3 days of fasting there is no easily detected release of alanine or glutamine into the portal blood in fasting humans.

After $\simeq 3$ days of starvation skeletal muscles derive $\simeq 50\%$ of their energy requirements from ketone bodies and, as fasting progresses, muscle preferentially extracts a small quantity of $AcAc^-$ and may release $\beta\text{-OHB}^-$. At this time in starvation, FFA provides the great majority of fuels for muscle.

After an overnight fast the brain derives its energy from glucose, consuming $\simeq 100-125\,g$ per day. As starvation progresses and blood glucose concentration decreases and the concentration of $AcAc^-$ and $\beta\text{-OHB}^-$ increases, ketone bodies become the major fuels for brain metabolism. These water-soluble fuels have access to neurons and can supply $\simeq 2/3$ of the total energy requirements during prolonged starvation.

Concluding Statement

Humans are capable of surviving long periods without food, but there are major limitations in this ability. When about one-half of lean body tissue is consumed to provide energy, death occurs. Morbidly obese people subjected to prolonged starvation can deplete the vital organs of structural proteins and die fat. The adaptations that permit humans to survive in the absence of food provide a fascinating look at the underlying biochemical adaptations to starvation. Much of what we know about the biological response to long-term starvation came from research with human volunteers who deserve our heartfelt thanks.

SEE ALSO THE FOLLOWING ARTICLES

A-Kinase Anchoring Proteins • Amino Acid Metabolism • Fat Mobilization: Perilipin and Hormone-Sensitive Lipase • Fatty Acid Oxidation • Glycogen Metabolism • Insulin- and Glucagon-Secreting Cells of the Pancreas • Insulin Receptor Family • Ketogenesis • Leptin • Pyruvate Carboxylation, Transamination and Gluconeogenesis • Tricarboxylic Acid Cycle

GLOSSARY

fuel homeostasis   The steady-state maintenance of fuels.
gluconeogenesis   The synthesis of glucose from nonhexose precursors.
glycogenolysis   The breakdown of glycogen.
ketogenesis   The synthesis of acetoacetate and β-hydroxybutyrate and the generation of acetone.
starvation   The deprivation of any or all of the elements for nutrition.

FURTHER READING

Balasse, E. O. (1979). Kinetics of ketone body metabolism in fasting humans. *Metabolism* **28**(1), 41–50.
Cahill, G. F. Jr. (1970). Starvation in man. *N. Engl. J. Med.* **282**, 668–675.
Felig, P., and Bergman, M. (1995). The endocrine pancreas: Diabetes mellitus. In *Endocrinology and Metabolism* (P. Felig, J. D. Baxter and L. A. Frohman, eds.) 3rd edition, pp. 1107–1250. McGraw-Hill, New York.
Keys, A. J., Brozek, J., Henschel, A., Mickelsen, O., and Taylor, H. L. (1950). *The Biology of Human Starvation*, Vols I and II, University of Minnesota Press, Minneapolis.
Leibel, R. M., Rosenbaum, M., and Hirsch, J. (1995). Changes in energy expenditure resulting from altered body weight. *N. Engl. J. Med.* **332**, 621–628.
Owen, O. E., and Reichard, G. F. Jr. (1971). Human forearm metabolism during progressive starvation. *J. Clin. Invest.* **50**, 1536–1545.
Owen, O. E., Morgan, A. P., Kemp, H. G., Sullivan, J. M., Herrera, M. G., and Cahill, G. F. Jr. (1967). Brain metabolism during fasting. *J. Clin. Invest.* **46**, 1589–1595.
Owen, O. E., Licht, J. H., and Sapir, D. G. (1981). Renal function and effects of partial rehydration during diabetic ketoacidosis. *Diabetes* **30**, 510–518.
Owen, O. E., Smalley, K. J., D'Alessio, D. A., Mozzoli, M. A., Knerr, A. N., Kendrick, Z. V., Kavle, E. C., Donohoe, M., Tappy, L., and Boden, G. (1990). Resting metabolic rate and body composition of achondroplastic dwarfs. *Medicine* **69**(1), 56–67.
Owen, O. E., Kalhan, S. C., and Hanson, R. W. (2002). The key role of anaplerosis and cataplerosis for citric acid cycle function. *J. Biol. Chem.* **277**(34), 30409–30412.

BIOGRAPHY

Richard W. Hanson is Professor of biochemistry at Case Western Reserve University School of Medicine in Cleveland, Ohio. For many years he has been an Associate Editor of the *Journal of Biological Chemistry* and was the President of the American Society of Biochemistry and Molecular Biology. His research interests include the dietary and hormonal control of gene transcription and the regulation of gluconeogenesis and glyceroneogenesis in mammals.

Oliver E. Owen was previously Professor of medicine and Co-Principal Investigator and Program Director, General Clinical Research Center, Temple University School of Medicine and Hospital, and formerly Professor and Chair of medicine, Southern Illinois University School of Medicine.

Steroid/Thyroid Hormone Receptors

Ramesh Narayanan and Nancy L. Weigel
Baylor College of Medicine, Houston, Texas, USA

Steroid and thyroid hormones are important in regulating a wide variety of normal physiological processes, including development, metabolism, and reproduction. The receptors for these hormones are members of the nuclear receptor superfamily of ligand-activated transcription factors. For the most part, each hormone interacts with a unique receptor although the receptor may have multiple forms derived from the same gene by various mechanisms. Exceptions include estradiol, which activates two receptors derived from independent genes, estrogen receptor α (ERα) and the recently discovered estrogen receptor β (ERβ), thyroid hormone receptors (TRα and TRβ), and retinoic acids, which activate the retinoic acid receptors (α, β, and γ). In response to their cognate hormones, nuclear receptors bind to specific DNA sequences altering transcription. In addition to their best-characterized actions as DNA-binding transcription factors, the receptors also influence transcription and resulting cell function through direct interactions with other transcription factors, as well as through alterations in cell signaling. These functions and the structures of the receptors are described in this article.

Overview of Nuclear Receptor Ligands and Mechanism of Action

Figure 1 shows the structures of some of the ligands for members of the nuclear receptor family. Estradiol is the primary ligand for both estrogen receptors, ERα and ERβ. Testosterone, which is closely related to estradiol, is one of the two major ligands for the androgen receptor (AR). The ligands for the thyroid hormone receptor (T_3 and T_4) and the vitamin D receptor ($1,25(OH)_2D_3$) are also shown. The five major classes of steroids are synthesized from pregnenolone, which is derived from cholesterol through the actions of the cholesterol side-chain cleavage enzyme (P450scc). The synthesis of the hormones is complex, with a number of alternate pathways leading to the same hormone. Testosterone and estradiol are derived from 17α-hydroxy pregnenolone. Testosterone, the major circulating androgen, is synthesized in the testes of males

and in the ovaries of females. It is important for development of the male reproductive tract, fertility, and secondary male characteristics. Estradiol, the major circulating estrogen, is produced in the ovaries of females. Estradiol is important in the female reproductive tract, playing roles in breast and uterine development, fertility, and also in bone and other tissues. Progesterone is synthesized directly from pregnenolone, and the major site of synthesis in females is the ovary. Progesterone, acting through the progesterone receptor, plays important roles in the breast and uterus and in the maintenance of pregnancy. Progesterone is a precursor of corticosterone and aldosterone, which are synthesized in the adrenal glands. The mineralocorticoid, aldosterone, is important for salt retention in the kidney. The glucocorticoid, cortisol, is produced in the adrenals from 17α-hydroxy pregnenolone and is important for regulation of carbohydrate metabolism; it also plays a role in suppressing immune responses. The secosteroid, $1,25(OH)_2D_3$, is derived from cholesterol through a UV-catalyzed reaction in the epidermis followed by sequential hydroxylations in the liver and kidney. The action of $1,25(OH)_2D_3$ is important for calcium homeostasis and also plays a role in the differentiation of a variety of tissues. Thyroid hormones are produced from tyrosines and iodide in the thyroid gland. Thyroid hormones have multiple actions in regulating metabolism, typically increasing oxidation rates. The hormones are transported through the blood to their sites of action.

Figure 2 depicts the general mechanism of action of nuclear receptors. The ligands are all lipophilic compounds, which enter the cells by passive diffusion. They bind to their cognate intracellular receptors located either in the cytoplasm or the nucleus. The receptors can be divided into two classes – those that do not bind DNA in the absence of hormone (Figure 2A) and those that bind to DNA in the absence of hormone (Figure 2B). The regulation of their activities differs somewhat. Classical steroid receptors such as AR, glucocorticoid receptor (GR), progesterone receptor (PR), mineralocorticoid receptor (MR), and
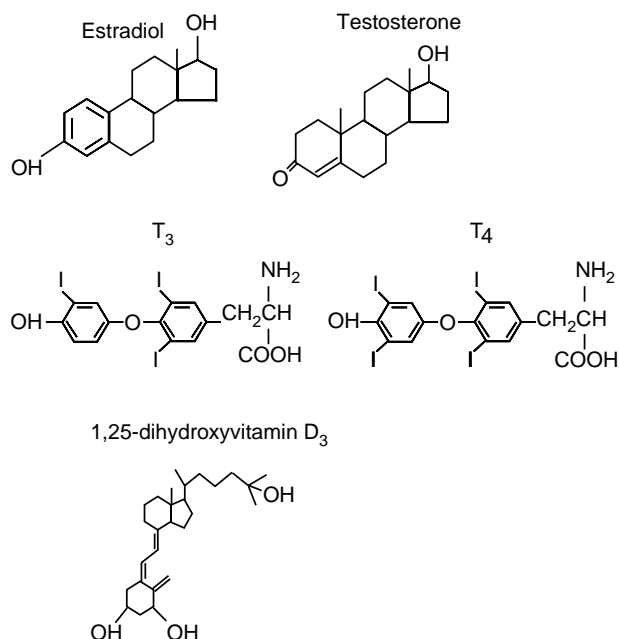
FIGURE 1   Structure of steroid and thyroid hormones. Estradiol is the ligand for estrogen receptor, testosterone is the ligand for androgen receptor, T_3-3,5,3-L-triiodothyronine, T_4-thyroxine are ligands for thyroid receptor, and 1,25-dihydroxyvitamin D_3 is the ligand for the vitamin D receptor.

ER belong to the first class; they are maintained in an inactive conformation capable of ligand binding by a complex of chaperone proteins including hsp90 and p23. Whether these complexes are nuclear or cytoplasmic depends upon the receptor. Upon ligand binding, the receptor changes its conformation, no longer binding to heat-shock proteins, homodimerizes and, if the unliganded receptor is localized in the cytoplasm, translocates to the nucleus. There, the receptor binds to DNA containing specific sequences called hormone response elements (HREs) that are most often found in the $5'$ flanking region of target genes. The receptor recruits components of the basal transcription machinery, as well as proteins termed coactivators that perform a variety of functions including histone acetylation that enhance transcription. Unlike the receptors for classical steroids, thyroid receptors (TRs) are not bound to heat-shock proteins in the absence of hormone and, instead, are bound to the DNA as a heterodimer with the retinoid X receptor (RXR), another member of the steroid/thyroid hormone receptor superfamily (Figure 2B). In the absence of ligand, TR binds corepressors, which in turn bind histone deacetylases resulting in lower levels of histone acetylation and repression of target gene transcription. Hormone binding releases corepressors and promotes binding of coactivators, and subsequent transcription follows.

The Steroid/Thyroid Hormone Receptor Superfamily

The steroid/thyroid receptors are the largest family of ligand-activated transcription factors. In addition to the well-characterized steroid, thyroid, retinoid, and vitamin D receptors, the family contains receptors for numerous lipophilic metabolites and xenobiotics as well as receptors, termed orphans, for which ligands have not yet been identified.

RECEPTOR STRUCTURE

Despite some evolutionary and functional differences, the steroid receptor family members have many similarities especially in their structure. As shown in Figure 3, there are multiple domains in the receptors, with all receptors containing domains A–E and only a subset containing the additional F domain.

The N Terminus

The N terminus or the A/B domain of the receptor is the least conserved domain among the family members. This region is the most variable in length ranging from a few amino acids to more than 500. This region has an activation function, AF-1, which contributes to the transcriptional activity of the receptor through binding of coactivators.

The DNA-Binding Domain (DBD)

The DNA-binding domain, region C, is important for the binding of receptor to the DNA and is the most highly conserved domain. This region has two type-2 zinc finger motifs, which are responsible for DNA recognition and dimerization (Figure 3B). Each finger is composed of four cysteines that coordinate with one zinc atom. Amino acids in this region also participate in receptor dimerization.

The Hinge Region

Downstream of the DBD is the hinge region (D), which contains a nuclear localization signal. This is a short lysine-rich region, with a high homology to the simian virus 40 T antigen nuclear localization signal. Additional functions of this region are receptor specific.

The Ligand-Binding Domain

The ligand-binding domain (E) is essential for the binding of ligand. The primary interaction site for the hsp complex is also in this domain. Also located in this region is the second activation function domain, AF-2,
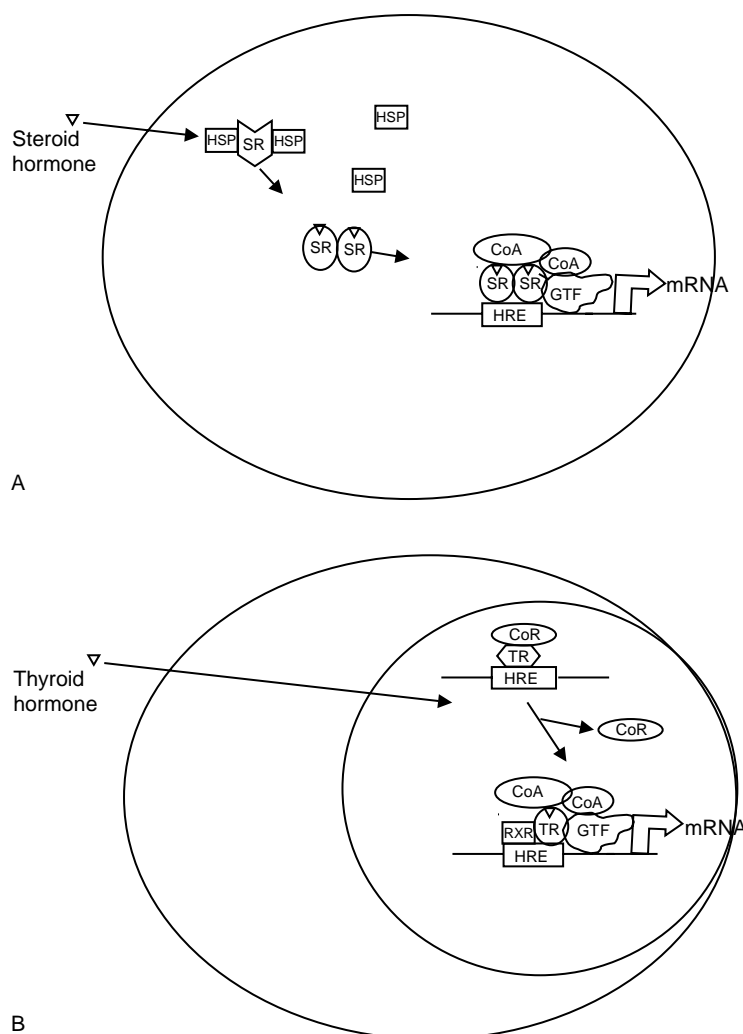
FIGURE 2 Mechanism of steroid (A) and thyroid (B) hormone action. The two classes are distinguished by whether they are associated with heat-shock proteins (A) like classical steroid receptors or are bound to DNA in the absence of hormone (B) like thyroid receptor. In both cases, binding of agonist causes dissociation of proteins that repress activity and promotes a conformation that induces recruitment of coactivators stimulating transcription of the target gene. SR, steroid receptor; HRE, hormone response element; GTF, general transcription factor; RXR, retinoid X receptor; TR, thyroid receptor; CoA, coactivator; CoR, corepressor.

which is responsible for ligand-mediated transcription of target genes. The relative importance of AF-2 and AF-1 in inducing transcription is receptor- and cell-type specific. The structures of the hormone-binding domains of several receptors have been determined using X-ray diffraction. The hormone-binding domain consists of a series of 12 α-helices. Binding of hormone causes a substantial conformational change in the receptor exposing AF-2 for interactions with coactivators. This domain also contains the strongest dimerization interface in most steroid receptors. The function of the F domain, located at the C terminus of some receptors such as the ER is not well defined.

RECEPTOR BINDING TO DNA

All of the receptors bind to their cognate HREs as dimers. The consensus binding sequence for AR, PR, GR, and MR, shown in Figure 3C, contains two half-sites separated by three nucleotides with the sites oriented to form a palindrome. ER recognizes a related pair of half-sites with the same spacing and orientation. Each monomer binds to a half-site. The class-II receptors, including VDR and TR, bind to pairs of half-sites whose sequences are identical to the ER half-site, but whose orientation (direct or inverted repeats) and spacing (0–6 nucleotides) determines the specificity of binding. TR and VDR each heterodimerize with
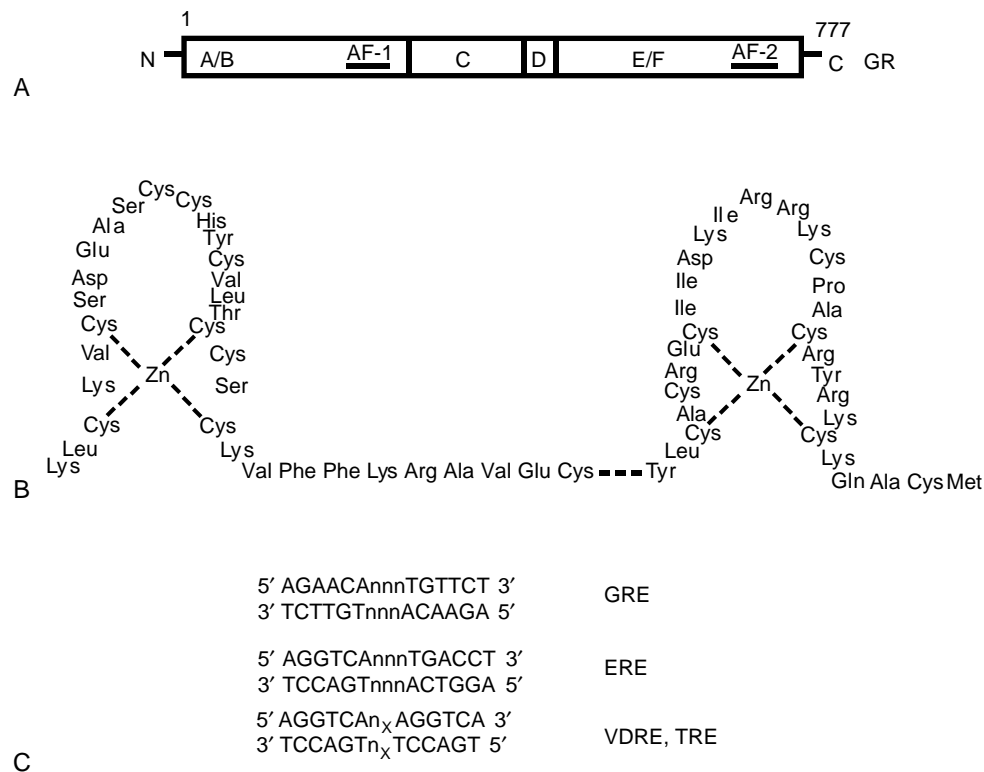
FIGURE 3 Receptor structure and DNA binding elements. Panel (A) shows the common structural features of nuclear receptors. The A/B region contains AF-1, a region important for transcriptional activation. C is the DNA-binding domain, the most conserved region in the nuclear receptors. D contains a nuclear localization sequence. E contains the hormone-binding domain and second activation function AF-2. Some receptors also contain a C terminal extension, termed the F domain, whose physiological function is not well described. Panel (B) shows the sequence of the DNA-binding domain of GR. Panel (C) shows sequences of consensus hormone response elements. The consensus sequence for a GRE (binds GR, AR, PR) and an ERE (binds ER) are shown. Vitamin D receptor and the thyroid hormone receptors bind to direct repeats separated by three and four nucleotides, respectively. Other receptors bind to direct or inverted repeats with a spacing of 0–6 (n_x indicates that the half-site may be separated by 0–6 nucleotides). GRE, glucocorticoid response element; ERE, estrogen response element; VDRE, vitamin D response element; TRE, thyroid response element; GR, glucocorticoid receptor; AF, activation function.

RXR and bind to the 3′ end (half of the HRE), whereas RXR binds to the 5′ end (half of the response element). Although these sequences represent the consensus binding sites, natural sequences may differ significantly and promoters may contain combinations of HREs as well as individual half-sites all of which contribute to the final activity.

Steroid Receptor Coregulators

When the receptor binds to the DNA, it recruits proteins in the basal transcription machinery such as TFIIB and RNA polymerase II. The receptors also bind additional proteins or protein complexes that modulate receptor activity; these are termed coactivators and corepressors. Coactivators are defined as proteins that interact with the receptors and increase their ability to transactivate the target gene. The mechanism by which individual coactivators achieve this can vary.

More than 100 candidate coactivators have been identified. While some coactivators function only with steroid receptors and a small subset of other transcription factors, others are used by many transcription factors. The best characterized of the steroid receptor coactivators (SRCs) is the p160 family of coactivators: SRC-1, SRC-2 (GRIP1, TIF2), and SRC-3 (Rac3, AIB1). These bind to the receptor recruiting additional coactivators including CBP (CREB-binding protein), p/CAF (CBP-associated factor) and CARM-1 (coactivator-associated arginine methyltransferase). CBP/p300, p/CAF and some of the p160 proteins are histone acetyl transferases (HATs) and their binding increases local histone acetylation. Other coactivators include the DRIP/TRAP (D receptor interacting protein/thyroid hormone receptor-associated protein) complex. Many coactivators interact with AF-2 located in the LBD. In other cases, coactivators interact with the AF-1 region and some interact with both domains. Interactions with

AF-2 are typically mediated by LXXLL (L = Leucine and X = any amino acid) motifs in the coactivator.

Another class of proteins, termed corepressor, reduces the activation of target gene transcription through interaction with the receptors. The best-characterized nuclear receptor corepressors are nuclear receptor corepressor (NcoR) and silencing mediator of retinoid and thyroid (SMRT) receptors. These proteins bind histone deacetylase complexes and also interact with class II receptors in the absence of hormone resulting in local reductions in histone acetylation. Corepressors do not bind to unliganded steroid receptors. However, steroid receptor antagonists cause changes in the conformation of the hormone-binding domain that induce binding of the corepressors.

Steroid Receptor Agonists and Antagonists

Although steroid receptor family members are important for normal physiological processes, there are a number of instances in which it is desirable to block the actions of selected steroid receptors. These include breast cancer (ER) and prostate cancer (AR). Thus, although natural antagonists of steroid receptor action have not been identified, much effort has been devoted to identifying compounds that will antagonize hormone action. These compounds compete with the natural ligand for binding to the hormone-binding domain of the receptors. Although some antagonists block dissociation from heat-shock protein complexes or destabilize the receptor, most of the antagonists promote dissociation from heat-shock proteins and cause the receptors to bind to DNA. However, the conformation induced by the antagonist differs from that induced by agonist. This prevents recruitment of coactivators to AF-2 and, instead, promotes recruitment of corepressors. In some cases, it is desirable to maintain the activity of a receptor in some tissues while inhibiting activity in other tissues. The most common example of this is the need for tissue-specific regulation of ER activity. Estradiol is important for maintaining bone mass and postmenopausal women frequently develop osteoporosis. However, estradiol can promote uterine cancer and may also be detrimental in breast. Thus, a great deal of effort has been devoted to developing selective estrogen receptor modulators (SERMs) which have tissue-specific agonistic and antagonistic activities. Tamoxifen (a SERM) has been used in the treatment of breast cancer, but increases the risk of uterine cancer. A newer SERM, raloxifene, is an antagonist in both breast and uterus, but acts as an agonist in bone.

Crosstalk between Nuclear Receptors and Cell Signaling Pathways

PHOSPHORYLATION OF RECEPTORS AND COACTIVATORS

The nuclear receptors and their coactivators are phosphoproteins; phosphorylation regulates various functions of these proteins. In some cases, enhanced cell signaling is sufficient to induce the transcriptional activity of the receptor. The ability to be activated by cell signaling pathways alone is receptor specific, although changes in cell signaling modulate the activity of all of the receptors. Estrogen receptors are activated both by growth factor pathways and by activation of protein kinase A. Other receptors, such as GR, require hormone for activity.

FUNCTIONAL INTERACTIONS BETWEEN NUCLEAR RECEPTORS AND SIGNAL-REGULATED TRANSCRIPTION FACTORS

In addition to altering transcription through direct binding to DNA, nuclear receptors alter transcription through interactions with other transcription factors. In some cases these protein–protein interactions prevent binding of the transcription factor to its DNA target site. In other cases, the receptor binds to the factors on their DNA target sites influencing (either + or −) the transcription of a target gene. Many of the anti-inflammatory actions of GR are a result of these protein–protein interactions. Figure 4 shows the
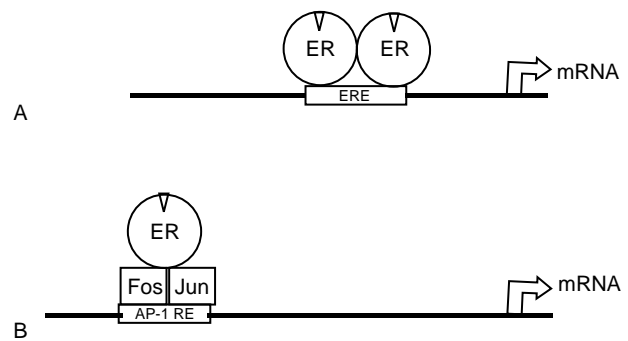
FIGURE 4 Mechanisms of transcriptional activation by nuclear receptors. Panel (A) shows the classical pathway for transcriptional activation with a steroid receptor dimer binding directly to a hormone response element. Panel (B) shows an alternative mode of activation. In this case, a receptor such as the estrogen receptor binds to another transcription factor and influences transcription through its interactions with the transcription factor and recruitment of coregulator proteins to the complex. ER, estrogen receptor; AP-1 RE, AP-1 response element; ERE, estrogen response element.

comparison between two of these pathways for ER. In the upper panel is the classical DNA-binding-dependent induction of transcription. We next depict the ability of ER to induce transcription through interactions with AP-1 complexes. In this instance, both estradiol as well as SERMs will stimulate the activity of AP-1.

NUCLEAR RECEPTOR STIMULATION OF CELL SIGNALING PATHWAYS

Both of the pathways above can be considered genomic pathways in that the nuclear receptor acts through altering transcription. Nuclear receptors can also act through stimulating kinase activity, although the final downstream target may be a change in transcription. These actions are rapid (minutes) and are termed non-genomic. The pathways are less well characterized, but there is evidence that activation of nuclear receptors can lead to downstream activation of mitogen-activated protein kinase (MAPK) In some cases, this is through activation of Src kinase and in others through generation of a ligand for a growth factor receptor. There are numerous other examples of these rapid actions. Induction by estradiol of nitric oxide synthase activity in endothelial cells is a rapid response that does not require transcription. Thus, steroid/thyroid hormones alter cellular activities through multiple mechanisms.

SEE ALSO THE FOLLOWING ARTICLES

A-Kinase Anchoring Proteins • Mitogen-Activated Protein Kinase Family • Thyroid-Stimulating Hormone/ Luteinizing Hormone/Follicle-Stimulating Hormone Receptors

GLOSSARY

agonists Natural or synthetic ligands that bind to the hormone-binding domain of the receptor stimulating activity.

antagonists Ligands that compete with agonists for binding to the hormone-binding domain, but do not cause activation of the receptor.

nuclear receptors Ligand-activated transcription factors characterized by conserved zinc-finger motifs in their DNA-binding domains and smaller conserved regions in the hormone-binding domain.

phosphorylation A posttranslational modification of an amino acid in which a phosphate group is added by a kinase.

FURTHER READING

Falkenstein, E., Tillmann, H. C., Christ, M., Feuring, M., and Wehling, M. (2000). Multiple actions of steroid hormones—a focus on rapid, nongenomic effects. *Pharmacol. Rev.* **52**(4), 513–556.

McDonnell, D. P., Connor, C. E., Wijayaratne, A., Chang, C. Y., and Norris, J. D. (2002). Definition of the molecular and cellular mechanisms underlying the tissue selective agonist/antagonist activities of selective estrogen receptor modulators. *Recent Prog. Horm. Res.* **57**, 295–316.

McKenna, N. J., Lanz, R. B., and O'Malley, B. W. (1999). Nuclear receptor coregulators: Cellular and molecular biology. *Endocr. Rev.* **20**(3), 321–344.

Tsai, M. J., and O'Malley, B. W. (1994). Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu. Rev. Biochem.* **63**, 451–486.

Weigel, N. L., and Zhang, Y. (1998). Ligand-independent activation of steroid hormone receptors. *J. Mol. Med.* **76**(7), 469–479.

Whitfield, G. K., Jurutka, P. W., Haussler, C. A., and Haussler, M. R. (1999). Steroid hormone receptors: Evolution, ligands, and molecular basis of biologic function. *J. Cell Biochem.* **32–33**(suppl.), 110–122.

BIOGRAPHY

Nancy Weigel is a Professor of Molecular and Cellular Biology at Baylor College of Medicine. Her interests are in steroid receptor action and crosstalk with cell signaling pathways as well as the role of steroid receptors in prostate cancer. She received her Ph.D. from Johns Hopkins University and received her postdoctoral training at Baylor College of Medicine.

Ramesh Narayanan is a postdoctoral fellow in the Department of Molecular and Cellular Biology at Baylor College of Medicine. He studies steroid receptor action and crosstalk with cell signaling pathways. He received his Ph.D. from the University of Madras, India.

Store-Operated Membrane Channels: Calcium

Indu S. Ambudkar

National Institute Of Dental and Craniofacial Research, Bethesda, Maryland, USA

Ca^{2+} entry via plasma membrane Ca^{2+} influx channels regulates a wide array of physiological functions such as neurotransmission, muscle contraction, secretion, and gene expression. A number of different types of Ca^{2+} channels have been identified in excitable and nonexcitable cells, including voltage-gated Ca^{2+} channels, primarily found in neuronal and various types of muscle cells; receptor-operated channels that are activated by extracellular ligand; and second-messenger-activated channels that are activated by intracellular "ligands" such as cGMP. Over the last decade considerable interest has been focused on store-operated Ca^{2+} channels (SOCCs), which mediate store-operated Ca^{2+} entry (SOCE, also referred to as capacitative Ca^{2+} entry, CCE). SOCE is activated following stimulation of cell-surface receptors that lead to phosphatidylinositol bisphosphate (PIP_2) hydrolysis, generation of diacylglycerol (DAG) and inositol-1, 4, 5-trisphosphate (IP_3), and IP_3-mediated release of Ca^{2+} from internal Ca^{2+} stores via the inositol trisphosphate receptor (IP_3R). The concept of SOCE was proposed by Putney in 1986 according to which depletion of Ca^{2+} in intracellular Ca^{2+} store(s) acts as a trigger for activation of plasma membrane Ca^{2+} influx. Ca^{2+} entering the cells via this pathway not only achieves refilling of the intracellular Ca^{2+} stores but also provides a sustained elevation of cytosolic $[Ca^{2+}]$ ($[Ca^{2+}]_i$) that is critical in the regulation of a variety of cellular functions. Despite the large number of studies that have been directed towards SOCE, the molecular composition of these channels as well as the mechanisms that activate or inactivate them have not yet been elucidated.

Characteristics of SOCE

Store-operated Ca^{2+} entry (SOCE) was originally identified in nonexcitable cells, although it has now been shown to be present in excitable cells as well. Thus, our knowledge of the characteristics of SOCE is primarily based on studies with nonexcitable cells that span over two decades. Early studies using radioactive Ca^{2+} and subsequent studies using Ca^{2+}-sensitive fluorescent probes together demonstrate that neurotransmitter stimulation of cells leads to a biphasic increase in cytosolic Ca^{2+}. An immediate increase that is not substantially altered by the removal of external Ca^{2+} suggesting that it is due to internal Ca^{2+} release, primarily mediated by IP_3. This is followed by a relatively sustained elevation of Ca^{2+}, completely dependent on the presence of external Ca^{2+}, that is due to influx of Ca^{2+} from the external medium. Although other Ca^{2+} influx pathways might contribute to this sustained $[Ca^{2+}]_i$ elevation, SOCE accounts for a major part, or all, of this Ca^{2+} influx. Several key studies lead to the conclusion that this Ca^{2+} influx is triggered by the depletion of Ca^{2+} in the internal Ca^{2+} store. (1) Ca^{2+} influx is not activated by IP_3 or its metabolites. (2) Ca^{2+} influx remains active even after the receptor-coupled signaling is stopped by addition of the antagonist. (3) Ca^{2+} influx is inactivated after Ca^{2+} is reintroduced into the cell and allowed to refill the internal Ca^{2+} store. (4) The same type of Ca^{2+} influx is activated by treating cells with the SERCA inhibitors such as Tg which induce a rapid and specific block of Ca^{2+} uptake into the ER and unmask a "leak" of Ca^{2+} from the ER.

The first channel activity associated with SOCE was measured in RBL cells by Penner and co-workers in 1992. This channel CRAC has been extensively studied and has also been found in lymphocytes and megakaryotes. CRAC is characterized by a high Ca^{2+}/Na^+ permeability ratio (>500), as well as a relatively rapid Ca^{2+}-dependent feedback inhibition. The channel displays strong anomalous mole fraction behavior suggesting that under normal physiological conditions external Ca^{2+} blocks the entry of Na^+ via the channel, thus Ca^{2+} is the favored ion to permeate this channel. Since 1992, store-operated Ca^{2+} influx channels have been measured in many different cell types, including cell lines and primary cell cultures. It is now clear that, although store-operated Ca^{2+} channels are all activated by the same, presently unknown, mechanism associated with internal Ca^{2+} store depletion they are not homogeneous. They display distinct biophysical properties, e.g., selectivity to Ca^{2+}, which suggest possible molecular diversity in their composition as well as differences in their modulation. An interesting question

that arises from such data is whether these distinct properties of store-operated Ca^{2+} channels (SOCCs) reflect their cell-specific physiological functions.

Mechanism of SOCE

Although agonist-stimulation of Ca^{2+} influx was first recognized in secretory cells almost three decades ago by Douglas and Poisner, the molecular mechanisms that activate or inactivate this Ca^{2+} influx have not yet been established. Unraveling this mechanism has been a major challenge in the field of Ca^{2+} signaling. Since SOCE was first identified, several mechanisms have been proposed to describe how it is activated. The earliest model proposed that Ca^{2+} in the internal store negatively regulates Ca^{2+} influx. When the store Ca^{2+} content is decreased, Ca^{2+} influx is activated and external Ca^{2+} is somehow directly accumulated into the ER, bypassing the cytosol. This model was primarily based on experiments which revealed that during refilling of the internal Ca^{2+} store (e.g., after agonist stimulation and addition of antagonist) by Ca^{2+} entry via the SOCE pathway, there is no detectable increase in $[Ca^{2+}]_i$. This model was disproved by studies in which external Ca^{2+} was substituted by divalent cations such as Mn^{2+}, which enter the cell via SOCE but are not pumped into the ER by the SERCA. These studies lead to the proposal that Ca^{2+} first enters the cytosol from where it is rapidly taken up into the ER lumen by the SERCA activity, and thus does not produce any substantial increase in $[Ca^{2+}]_i$. This results in refilling of the internal Ca^{2+} stores which leads to inactivation of SOCE. Thus, there is reciprocal regulation of the ER Ca^{2+} store and plasma membrane Ca^{2+} channels.

Later models addressed the nature of the signal that conveys the status of the internal Ca^{2+} store to either activate or inactivate SOCE in the plasma membrane. While a number of different models have been proposed in an effort to explain this "ER–PM coupling," three major mechanisms have garnered the most attention; (1) conformational coupling, (2) a diffusible factor, and (3) regulated recruitment of channels by fusion of intracellular vesicles (Figure 1). Conclusive data are presently lacking to either support or rule out any of these proposed mechanisms for SOCE activation. A major hurdle in these efforts has been the lack of knowledge regarding the molecular identity of the SOCC channel.

Molecular Candidates for SOCC

The relatively recent discovery of mammalian homologues of the *Drosophila Trp* (transient receptor potential) gene has propelled the field of SOCE in a new direction. TRP proteins form a large functionally diverse
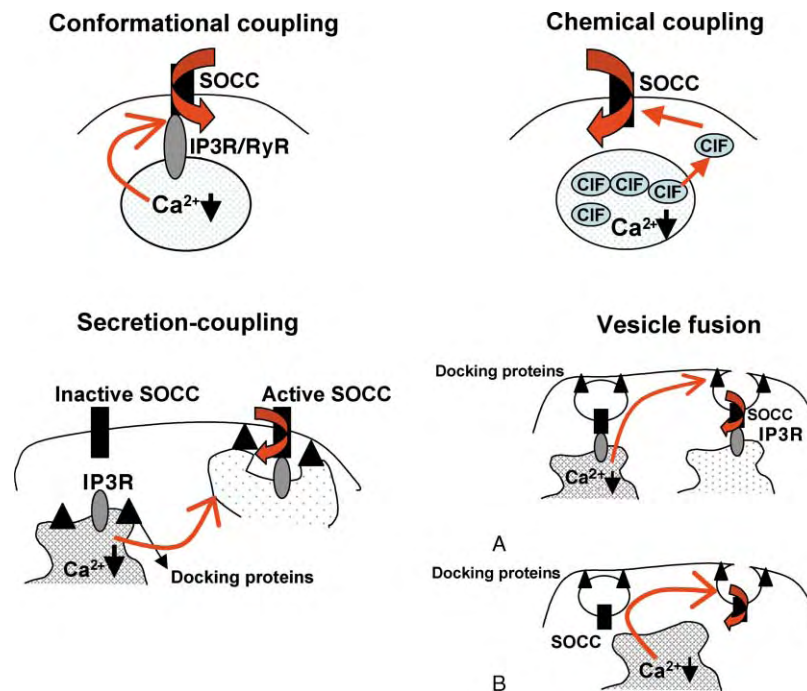
FIGURE 1 Proposed models for activation of SOCE. All components are labeled in the figures. See text for description. Two mechanisms are shown for the vesicle fusion model. (A) "Kiss and Run" model predicts that SOCC-containing vesicles might be docked to the ER and therefore sense depletion which initiates fusion to PM. Ion channel is retrieved during inactivation. (B) SOCC-containing vesicles are present in the subplasma membrane region and sense depletion of store that leads to vesicle fusion and channel insertion into the PM. Inactivation and retrieval could be independent events.

superfamily of ion channel proteins and are found in all excitable and nonexcitable tissues. Members of the TRPC subfamily and some members of the TRPV subfamily form channels that are activated in response to receptor-coupled Ca^{2+} signaling events (Figure 2 shows the proposed structure of TRPC1). TRP channels appear to fall into two general classes; nonstore operated such as TRPC3, TRPC6, TRPC7, which are activated by agonists and exogenous addition of diacylglycerol (DAG), but not by thapsigargin. Thus, it is suggested that this family forms channels that are activated by DAG generated in response to PIP_2 hydrolysis. The other group which includes TRPC1, TRPC4, TRPC5, TRPV6 has been shown to form store-operated channels. In this case, studies include heterologous expression, knockdown of endogenous proteins, and site-directed mutagenesis. However, there are exceptions. For example, TRPC3 has been shown to form SOCC and also to be regulated by the internal Ca^{2+} store via interaction with inositol trisphosphate receptor (IP_3R). Further, TRPCs have also been shown to interact to form heteromultimeric channels which display a wide range of biophysical characteristics. Thus, it is possible that TRPCs could be involved in the formation of a diverse range of SOCCs. Although further work is required to conclusively establish that TRP proteins are molecular components of SOCCs, presently they are the only viable candidates for these channels. Further, TRPs provide useful tools to test the validity of the models proposed for SOCC activation.

Signaling from ER to the Plasma Membrane

CONFORMATIONAL COUPLING

This hypothesis was originally proposed by Irvine in 1990 and was based on functional analogy between IP_3R and ryanodine receptors (RyR) in muscle cells. RyR are Ca^{2+} release channels in the muscle sarcoplasmic reticulum (SR) and have been suggested to physically couple to the L-type Ca^{2+} channels in the T-tubule plasma membrane. During excitation–contraction coupling, Ca^{2+} inflow via the plasma membrane channels regulates Ca^{2+}-induced Ca^{2+} release via RyR in the SR. Although in the case of SOCE, the flow of information can be predicted to occur in the reverse direction, i.e., from ER to the plasma membrane, the homology between IP_3R and RyR channeled the hypothesis that regulation of SOCC could be mediated via a direct physical association between the IP_3R and the SOCC in the plasma membrane. The model proposes that there are preformed IP_3R–SOCC complexes and that store depletion is detected by the IP_3R, leading to a conformational change that results in activation of SOCC. A caveat to the IP_3R requirement in the regulation of SOCC is its activation by SERCA inhibitors or by depletion of Ca^{2+} stores by loading the cytosolic or ER with Ca^{2+} buffers. To explain this discrepancy, it was proposed that only a certain pool of IP_3R's interacts with SOCC, and that these are localized
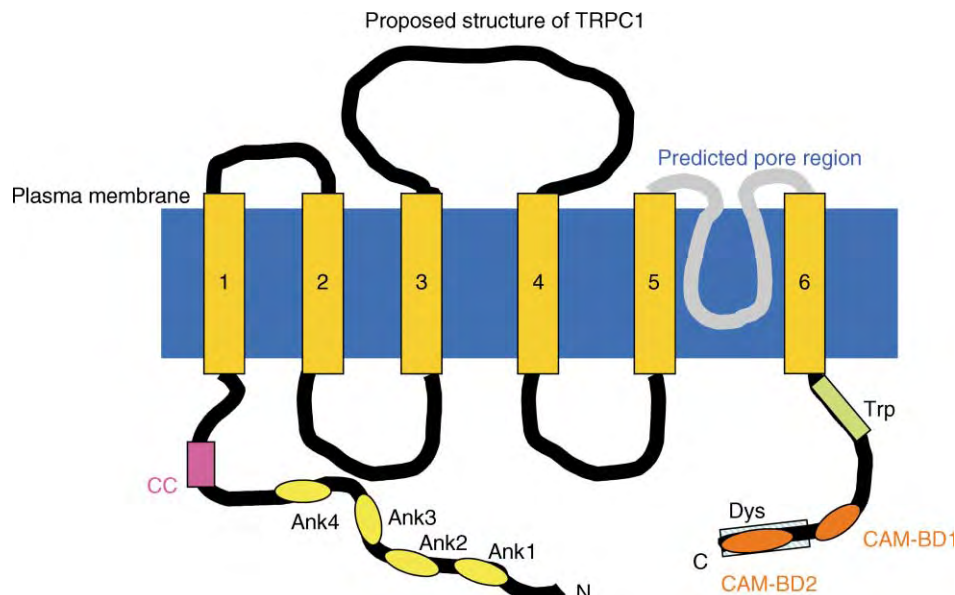
FIGURE 2   Proposed structure of transient receptor potential canonical 1 (TRPC1): The structure includes six-transmembrane domains (1–6), extracellular as well as intracellular domains (black lines) and a pore-domain between the fifth and sixth TMs (gray). The N terminus of the protein contains ankyrin repeats (ank) as well as a coil–coil domain (CC). The C terminus has the conserved TRP motif (Trp, EWKFAR), calmodulin-binding domains (CAMBD), and a dystrophin domain (Dys).

in ER membranes situated in close proximity to the PM. This also implied that this pool of IP_3R is not involved in internal Ca^{2+} release, but only in SOCC regulation. Currently, there are conflicting data regarding the role of IP_3R in the regulation of SOCC. Studies using a gene knockout approach have shown that IP_3Rs are not required for thapsigargin-stimulated Ca^{2+} entry, although they are clearly required for IP_3-mediated internal Ca^{2+} release. Other studies suggest that RyR, which are present in several nonmuscle cell types, can also couple with SOCC and regulate its function. Thus, it is possible that RyR could replace IP_3Rs in cells where IP_3R expression has been down-regulated or eliminated. However, further studies will be required to rule out or provide conclusive evidence for the conformational coupling hypothesis.

SECRETION-LIKE COUPLING/VESICLE FUSION

The activation of SOCE is a relatively slow process. A lag time of about 10 s has been detected between internal Ca^{2+} release and Ca^{2+} influx. Thus, it has been proposed that vesicle trafficking and fusion events could be involved in activation of SOCE. Two possible processes could occur. The first is a variation of the conformational coupling model and suggests that the ER–PM interaction is a dynamic, reversible, process and that the ER membrane moves towards and docks with the PM upon stimulation. The docking enables proteins in the PM and ER to interact, thus resulting in activation of SOCC. Although the ER protein is considered to be IP_3R or RyR, other proteins could also be involved in the ER–PM signaling. Since the ER and PM are apposed to each other at the site of interaction, there is no particular requirement that the ER or PM protein should have very long cytosolic domains. Support for the secretion-coupling hypothesis has been mainly provided by studies using reagents to disrupt the cytoskeleton and alter the spatial arrangement of cellular organelles. Further, TRPC1–IP_3R interaction was shown to be disrupted by reagents that induce cortical actin formation. However, several other studies have refuted this model as a possible activation mechanism for SOCC. The second mechanism that can be suggested involves vesicle trafficking and exocytotic insertion of the channel proteins. Here again, there are data to both support and refute the model. Experiments have shown that disruption of the SNARE proteins involved in exocytosis, inhibits activation of SOCE. However, in other studies, such maneuvers did not affect SOCC activation. An important point that needs to be considered when assessing the possible mechanisms for activation is whether the different SOCCs that have been detected in various cell types are activated by the same mechanism or does internal Ca^{2+} store-depletion induce

a variety of cellular signals which can then activate different channel types. For example, if different TRP channels are involved in the SOCCs in the different cell types, can that account for differences in their regulation? Voltage-gated Ca^{2+} channels represent a family of proteins that are activated by various thresholds of membrane potential. They are also regulated differently, exhibit distinct characteristics and carry out specific physiological functions. Analogous to this, we might have to consider SOCC channels as a family of channels that sense the same fundamental signal, but are regulated by subtly distinct mechanisms. What these mechanisms are, presents a challenging question for future studies in this field.

METABOLIC COUPLING

Another hypothesis that has received sporadic attention is that an as yet unknown diffusible factor, referred to as CIF, is either released from the ER with Ca^{2+} or is generated during this process. CIF can reach the PM SOCC channels and either activate it directly or bind to a regulatory protein and enable channel activation. Evidence in support of this shows that extracts from stimulated cells can increase Ca^{2+} influx in unstimulated cells. However, these findings have not held up for all types of cells. Other metabolites that have been shown to regulate SOCC are of the cytochrome P450 epoxygenase pathway. Modifiers of the lipoxygenase pathway has been shown to affect I_{CRAC} in RBL cells. A role for arachidonic acid has also been suggested. Thus, further studies are needed to establish whether CIF is involved in SOCC activation. It should be noted that the requirement for CIF and secretion-like coupling need not be mutually exclusive, since dynamic trafficking of ER to the PM would decrease the diffusion restraints for CIF. In addition, reassembly of the cortical actin can also play a role in the access to the PM and diffusion of CIF. Interestingly, the status of the actin is controlled by the PIP_2 levels in the plasma membrane. Thus, the hydrolysis of PIP_2 not only initiates Ca^{2+} signaling but also remodeling of the actin cytoskeleton in order to facilitate the regulation of cellular function. In fact, modulation of PIP_2 metabolism, i.e., inhibition of PI-3 kinase has been shown to alter SOCE in some cells. Thus, it is becoming exceedingly evident that regulation SOCE is a highly orchestrated process with several orders of complexity that might be determined by the particular SOCC that is present, and the specific physiological function that it contributes to, in any given type of cell.

Ca^{2+} Signaling Microdomains

Recent studies have highlighted spatio-temporal aspects of Ca^{2+} signaling in cells. It has been demonstrated that agonist-stimulated Ca^{2+} influx occurs within specific

spatially restricted microdomains. During Ca^{2+} influx-dependent refill of internal Ca^{2+} stores there is minimal diffusion of Ca^{2+} in the subplasma membrane region due to rapid uptake into the ER by the SERCA pump, indicating close apposition between the ER and plasma membrane at the site of Ca^{2+} influx. SERCA, mitochondria, and PMCA together determine the efficiency of SOCC by removing Ca^{2+} from the vicinity of the channel, thereby decreasing Ca^{2+}-dependent feedback inhibition of the channel. PMCA, mitochondria, and SERCA have been functionally localized to the microdomain where SOCE is occurring. Furthermore, PMCA and SERCA have been shown to be immunoprecipitated with TRPC channels. Thus, the architecture of such Ca^{2+} signaling microdomains facilitates direct physical, or functional, coupling between the molecular components that are involved in regulating plasma membrane SOCC. Current models support the idea that SOCC and functionally associated proteins are strategically localized by the action of scaffolding proteins which allow dynamic regulation of cellular function (Figure 3).

A candidate for such a microdomain is caveolar lipid raft domains which have been shown to be involved in the regulation of SOCE in several cell types. These domains are functionally and biochemically distinct microdomains formed by the lateral packing of glycosphingolipids and cholesterol within the membrane bilayer. It has been proposed that arrangement of the

lipids and scaffolding proteins within LRD forms a platform for the assembly of a number of proteins into signaling complexes. This compartmentalization of the signaling molecules can increase the rate of interactions and enhance crosstalk networks. Importantly, key protein and nonprotein molecules involved in the Ca^{2+} signaling cascade, such as PIP_2, $G_{\alpha q/11}$, PMCA, several TRPC proteins, and IP_3R-like protein, and Ca^{2+} signaling events such as receptor-mediated turnover of PIP_2 have been localized to caveolar microdomains in the plasma membrane. Studies have also revealed that (1) agonist-stimulated Ca^{2+} signal in endothelial cells originates in specific areas of the PM that are enriched in caveolin-1, and (2) intact lipid rafts are required for activation of SOCE. Thus, it has been proposed that caveolae might regulate the spatial organization of Ca^{2+} signaling by contributing to the localization of Ca^{2+} signaling complex as well as the site of Ca^{2+} entry.

It is now well recognized that Ca^{2+} signaling proteins are assembled in multiprotein complexes. This finding is again supported by studies carried out with TRP proteins. It was earlier shown that the well-studied TRP prototype, the *Drosophila* TRP, is assembled in a signaling complex by the scaffolding action of INAD, a multi-PDZ domain containing protein. Mammalian TRPC channels are also assembled in multimeric protein complexes that are associated with key Ca^{2+} signaling proteins such as $G_{\alpha q/11}$, IP_3R, PLCβ, PMCA, and SERCA as well as scaffolding proteins such as
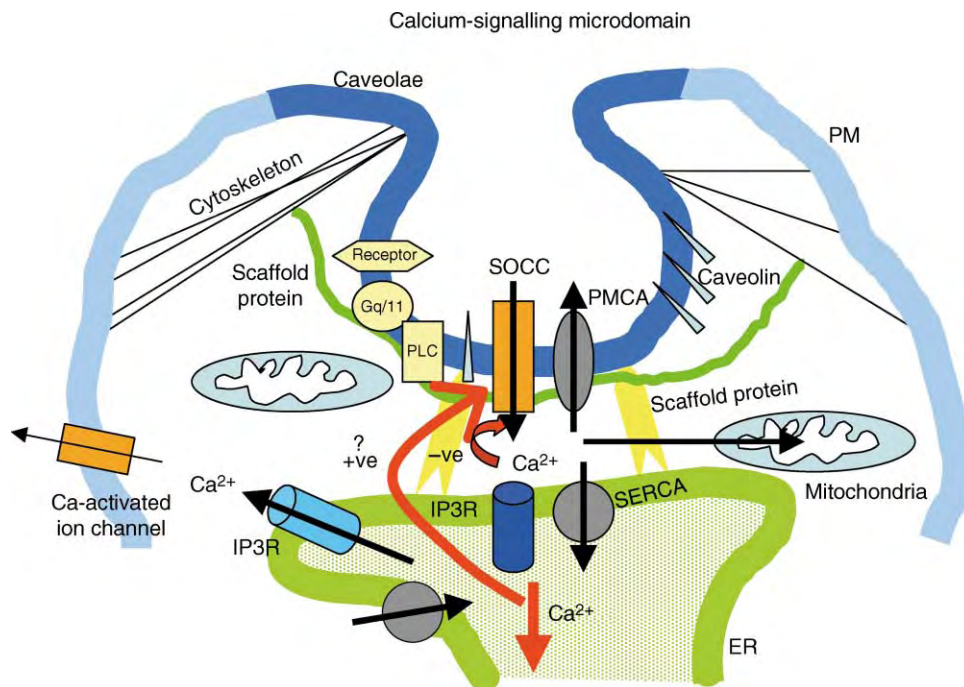
FIGURE 3 Proposed calcium-signaling microdomain: Functionally associated proteins are localized in distinct structures and assembled together by scaffolding proteins as well as cytoskeletal interactions. The architecture of the region enables the coupling between ER and PM that is involved in activation of SOCE. Protein components are labeled in the figure.

caveolin-1, homer, and NHERF. Accessory proteins have been identified not only for TRPC channels, but also for PMCA, SERCA, and IP$_3$Rs. These proteins encompass a variety of functions such as trafficking, phosphorylation, dephosphorylation, scaffolding, etc. Thus, further information regarding the protein components and assembly of the mammalian Ca^{2+}-signaling complex will provide a better understanding of SOCE. Such knowledge will impact several important areas of physiology and pathophysiology.

SEE ALSO THE FOLLOWING ARTICLES

Calcium Signaling: Cell Cycle • Lipid Rafts • Plasma-Membrane Calcium Pump: Structure and Function • Voltage-Sensitive Ca^{2+} Channels

GLOSSARY

Ca^{2+} channels   Protein(s) that mediate passive, electrochemical gradient driven, flux of Ca^{2+}.

Ca^{2+} pump   Protein(s) that utilize energy of ATP hydrolysis to drive Ca^{2+} flux against its gradient.

Ca^{2+} signaling   Events that initiate elemental changes in cellular levels of Ca^{2+} which are then decoded into a physiological response.

cellular microdomains   Morphological and functionally distinct subcellular domains.

neurotransmitter   Chemical messengers secreted by nerve endings.

FURTHER READING

Berridge, M. J. (1995). Capacitative Ca^{2+} entry. *Biochem. J.* **312**, 1–11.

Isshiki, M., and Anderson, R. G. W. (2003). Function of caveolae in Ca^{2+} entry and Ca^{2+}-dependent signal transduction. *Traffic* **4**, 717–723.

Montell, C., Birnbaumer, L., and Flockerzi, V. (2002). The TRP channels, a remarkable functional family. *Cell* **108**, 595–598.

Muallem, S., and Wilkie, T. M. (1999). G-protein dependent Ca^{2+} signaling complexes in polarized cells. *Cell Ca^{2+}* **26**, 173–180.

Parekh, A. B. (2003). Store-operated Ca^{2+} entry: Dynamic interplay between endoplasmic reticulum, mitochondria, and plasma membrane. *J. Physiol.* **547**, 333–348.

Putney, J. W., Jr., Broad, L. M., Braun, F.-J., Lievremont, J.-P., and Bird, G. J., St. (2001). Mechanisms of capacitative Ca^{2+} entry. *J. Cell Sci.* **114**, 2223–2229.

Venkatachalam, K., von Rossum, D. B., Patterson, R. L., Ma, H.-T., and Gill, D. L. (2002). The cellular and molecular basis of store-operated Ca^{2+} influx. *Nat. Cell Biol.* **4**, 263–272.

BIOGRAPHY

Indu Ambudkar is Chief of Secretary Physiology Section at the National Institute of Dental and Craniofacial Research, NIH, in Bethesda, Maryland. She holds both Masters and Ph.D. degrees in Biochemistry from Lucknow and Madurai Kamaraj Universities in India and received postdoctoral training at the University of Maryland, School of Medicine. Her primary research interest is regulation of cellular Ca^{2+} homeostasis and more recently store-operated Ca^{2+} influx in salivary epithelial cells. Her work has contributed significantly to the understanding of SOCE, and the role of TRPC proteins in this process.

Substrate Binding, Catalysis, and Product Release

W. Wallace Cleland
University of Wisconsin, Madison, Wisconsin, USA

Enzymes catalyze reactions by forming complexes with their substrate(s), accelerating the chemical reaction by lowering the activation energy for the reaction, and then releasing the product(s). The enzyme has to have sufficient affinity for a substrate to form a complex with it at the physiological concentration of this molecule. Conversely, however, too great an affinity for the substrate usually results in too great an affinity for the product, so that product release becomes rate limiting. The maximum rate of an enzyme-catalyzed reaction is limited, in fact, by the relationship between the equilibrium constant and the kinetic constants called the Haldane relationship, and one element of this restriction is the affinity of the substrate.

The Haldane Relationship

As noted above, the Haldane relationship allows calculation of the maximum rate of an enzyme-catalyzed reaction. For the simple conversion of substrate A into product P, the initial rate of the forward reaction in the absence of P is

$$v_f = -d[A]/dt = V_1[A]/(K_a + [A]) \qquad [1]$$

and that in the back reaction is

$$v_r = -d[P]/dt = V_2[P]/(K_p + [P]) \qquad [2]$$

where V_1 and V_2 are maximum velocities in forward and reverse reaction, and K_a and K_p are Michaelis constants for A and P (i.e., apparent dissociation constants in the steady state).

The Haldane for this mechanism is the ratio of the V/K values in forward and reverse directions:

$$K_{eq} = [P]_{eq}/[A]_{eq} = (V_1/K_a)/(V_2/K_p)$$

$$= V_1 K_p/(V_2 K_a) \qquad [3]$$

V/K, with units of reciprocal time, is the apparent first-order rate constant for reaction at low substrate concentration, so eqn. [3] is analogous to the one for a nonenzymatic reaction where K_{eq} would equal k_1/k_2,

i.e., the ratio of rate constants in forward and reverse directions.

To optimize the turnover number of an enzyme (V_1/E_t) in this case, also known as k_{cat}, with units of reciprocal time, we can raise the two V/K values in constant ratio until one V/KE_t value (with units of $M^{-1} s^{-1}$) reaches the limit set by diffusion. V/KE_t (or k_{cat}/K) is a second-order rate constant for productive combination of enzyme and substrate and subsequent reaction to give product, and the combination of enzyme and substrate can only go as fast as these two can diffuse together. This limit is $\sim 10^9 M^{-1} s^{-1}$ for small substrates, and somewhat less for larger ones like ATP.

Once the $V_1/K_a E_t$ value in eqn. [3] is as high as it can go, the only way to increase V_1/E_t is to increase K_a as well. But K_a cannot exceed the physiological level of A, or the enzyme will not be at least half-saturated. Thus V_1/E_t has an upper limit, and enzymes that are part of metabolic pathways are optimized to operate at this limit. The actual turnover numbers vary, depending on K_{eq} and the physiological level of the substrate, but evolution has brought these enzymes to the limit set by the Haldane relationship. On the other hand, enzymes that are involved in control, as opposed to ones in metabolic pathways, often sacrifice speed for control characteristics, and do not have turnovers numbers at the Haldane limit.

Catalysis

The catalytic process on an enzyme starts once the substrate is bound. As noted above, the enzyme has to have sufficient affinity to bind the substrate at its physiological concentration. The enzyme then has to lower the activation energy of the reaction sufficiently for it to go at a rate approaching the maximum given by the Haldane relationship. It is commonly said that the enzyme accomplishes this by binding the transition state structure more tightly than the ground state of the substrate, but this is really a definition of catalysis. Molecules resembling the transition state structure are

often bound much more tightly than the substrate (see entry on transition state analogues). This is either because the geometry of the transition state is a better fit, or includes inherently tighter interactions (stronger hydrogen bonds, for example; see entry on low barrier hydrogen bonds). Alternatively, the differential binding results because the substrate is destabilized when it is bound (sterically deformed, placed in an unfavorable electrostatic environment, etc.), but this destabilization is relieved in the transition state. Of course, destabilization in one part of the substrate must be matched by tight binding in another part, or the substrate will not have sufficient affinity for the enzyme at its physiological level. OMP decarboxylase is a classic example, where a substrate analogue with $-COO^-$ replaced by $-O^-$ binds 10^5-fold more tightly to the enzyme. The difference is ascribed to charge repulsion between the carboxyl group of the substrate and an aspartate on the enzyme.

Free-Energy Profiles

The energetics of what happens during an enzymatic reaction is illustrated by a free-energy profile. Figure 1 shows the free-energy profile at equilibrium for the nonenzymatic reaction of A to P. The activation energy, $\Delta G^\ddagger$, corresponds to the height of the barrier, and ΔG for the reaction ($= -RT \ln K_{eq}$) is the difference between the levels marked A and P. For a corresponding enzymatic reaction, one must decide what levels of A and P to use. Figure 2 shows an equilibrium free-energy

FIGURE 2 Free-energy profile at equilibrium for an enzymatic reaction. The symbols have the same meaning as in Figure 1.

profile where the concentration of A is above its dissociation constant, so that formation of EA has an equilibrium constant greater than unity, but where P is less than its dissociation constant, so that EP dissociation is favorable. $\Delta G^\ddagger$ is given by the height of the barrier above the level of EA, not E + A. The turnover number (V/E_t or k_{cat}) is determined by $\Delta G^\ddagger$.

If one picks different levels of A or P, the difference between E + A and EA, or E + P and EP, will differ as shown in Figure 3. The dissociation level of A (K_{ia}) is convenient, but then the level of P is determined by K_{eq} if an equilibrium free-energy profile is plotted, and will not usually match K_{ip}, its dissociation constant. Free-energy profiles may, of course, be plotted for nonequilibrium concentrations of reactants, but in any case it is critical to state the concentrations of reactions that are assumed in plotting the profile. When there are two or more substrates or products involved in the reaction, one must pick suitable levels of all reactants (and state them!) in order to plot a free-energy profile.

So far we have considered only simple free-energy profiles with a single transition state connecting

FIGURE 1 Free-energy profile at equilibrium for an uncatalyzed reaction. $\Delta G^\ddagger$ is the activation energy and $\Delta G = -RT \ln K_{eq}$, the energy difference between product and substrate.
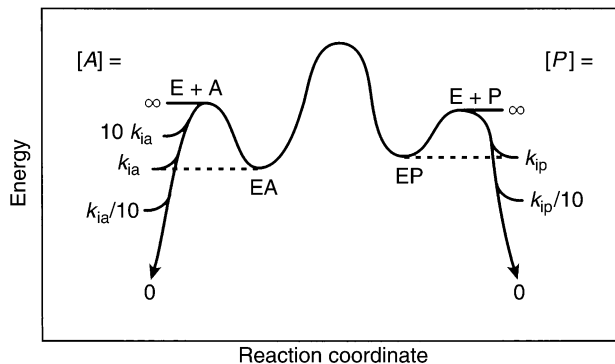
FIGURE 3 Free-energy profile showing the difference in the levels of (E + A) or (E + P) as the concentrations of A or P are changed. K_{ia} and K_{ip} are the dissociation constants of A from EA and of P from EP.
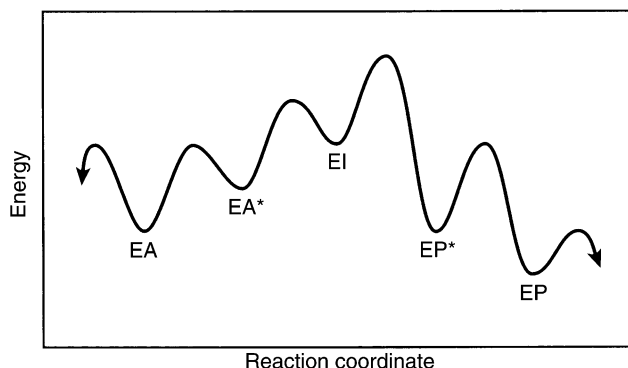
FIGURE 4 Free-energy profile showing conformation changes in the EA and EP complexes, as well as an intermediate between the two activated complexes. The levels of (E + A) and (E + P) are not shown.

EA and EP. In practice, enzymes have different conformations at the various stages of the catalytic cycle. Thus, they have open forms where the reactants are free to bind and dissociate, and closed forms in which catalysis takes place. The chemistry may also take place in more than one step, so that intermediates are present. Figure 4 shows a profile where a conformation change converts the open EA form to a closed, precatalytic EA* form. This undergoes a catalytic reaction to give an EI intermediate complex, which then is further converted to the closed EP* form. A final conformation change gives the open EP form, from which P can dissociate.

Rate-Limiting Step for *V/K*

The two independent kinetic constants for an enzymatic reaction are the maximum velocity, *V*, and the ratio of *V* and the Michaelis constant, *V/K*. There is a separate *V/K* for each substrate. Each of these parameters varies separately with pH, ionic strength, temperature, and the concentrations of other substrates, products, or inhibitors. The Michaelis constant, *K*, while it measures the apparent dissociation constant of the substrate in the steady state, is not an independent constant, but just the ratio of *V* and *V/K*. Thus when one speaks of "rate-limiting steps" one must specify whether one is referring to *V* or *V/K*.

The rate-limiting step for *V/K* is the one with the highest barrier in the free-energy profile. *V/K* involves the combination of enzyme and substrate and reaction through the first irreversible step, which normally is release of the first product. Later steps that involve further conformation changes in the enzyme and release of other products do not affect *V/K*. Thus the slow release of NADH from dehydrogenases, which often limits the maximum velocities of these enzymes, has no effect on *V/K*. As a result, the *V/K* value often is limited more by the chemistry of the reaction than *V*, and

isotope effects on the chemistry are more fully expressed (see entry on kinetic isotope effects). Except for some slow mutants, full rate-limitation by the chemical reaction is unusual, however, and the conformation changes that precede and follow the chemical reaction are usually partly rate-limiting. The equation for an isotope effect on *V/K* is:

$$^x(V/K) = (^xk + c_f + {}^xK_{eq}c_r)/(1 + c_f + c_r) \qquad [4]$$

where *x* defines the isotope effect (D, T, 13, 15, 18 for deuterium, tritium, ^{13}C, ^{15}N, ^{18}O) and the leading superscript indicates the ratio of the parameters for light and heavy isotopes. ^xk is the intrinsic isotope effect on the chemical step, and $^xK_{eq}$ the equilibrium isotope effect on the reaction. The constants c_f and c_r are forward and reverse commitments to catalysis and are the ratio of the rate constant for the isotope sensitive step to the net rate constant for release from the enzyme of either the substrate (for c_f) or the first product (for c_r). For the simple mechanism

$$E + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} EA \underset{k_4}{\overset{k_3}{\rightleftharpoons}} EA^* \underset{k_6}{\overset{k_5}{\rightleftharpoons}} EPQ^* \underset{k_8}{\overset{k_7}{\rightleftharpoons}} EPQ$$

$$\overset{k_9}{\rightarrow} EQ + P \overset{k_{11}}{\rightarrow} E + P + Q \qquad [5]$$
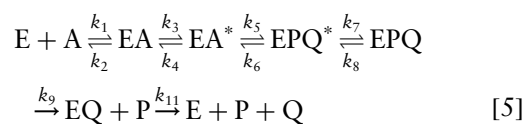
$$c_f = (k_5/k_4)(1 + k_3/k_2) \quad c_r = (k_6/k_7)(1 + k_8/k_9) \qquad [6]$$

The commitments thus consist of partition ratios of intermediates and these ratios correspond to the differences in barrier heights in a free-energy profile. In the profile shown in Figure 4, c_f and c_r will be small if the isotope effect measured involves reaction of EI to EP* because the barrier between EI and EP* is the highest one. But if the isotope effect measured is on reaction of EA* to EI, c_f will be small, but c_r will be large. Thus one will see largely the equilibrium isotope effect, as the EA* to EI step approaches equilibrium.

In the definition of c_f in eqn. [5], k_5/k_4 is the internal part of the commitment ($c_{f\text{-in}}$) and $k_3k_5/(k_2k_4)$ is the external part ($c_{f\text{-ex}}$). A sticky substrate is one where the net rate constant for reaction of the initial collision complex to give products is faster than the rate constant for dissociation (k_2). This ratio is called the stickiness ratio (S_r) and is given by $S_r = c_{f\text{-ex}}/(1 + c_{f\text{-in}} + c_r)$ in terms of the parameters of eqn. [5].

Rate-Limiting Steps for *V*

The rate-limiting step for the maximum velocity depends on the definition of rate limiting. One definition is the "slowest" step, or the one with the highest individual barrier in the forward direction. Another definition is the "least-conductive" step, which is the one that sees the highest total barrier to reach an irreversible step. A third definition is the "most-sensitive" step. This is the one

which causes the greatest percentage change in the turnover number for a given percent change in the forward rate constant. An isotope effect on this step is more fully expressed than one on any other step. The isotope effect on V for mechanism 3 is

$$^{x}V = (^{x}k + c_{v_{f}} + {}^{x}K_{eq}c_{r})/(1 + c_{v_{f}} + c_{r}) \qquad [7]$$

where ^{x}k, $^{x}K_{eq}$, and c_{r} have the same meaning as in eqn. [5], but

$$c_{v_{f}} = [k_{3}k_{5}/(k_{3} + k_{4})][1/k_{3} + (1/k_{7})(1 + k_{8}/k_{9}) \\ + 1/k_{9} + 1/k_{11}] \qquad [8]$$

Note that $c_{v_{f}}$, unlike c_{f}, does not consist of partition ratios. Rather k_{5}, reduced by the factor $k_{3}/(k_{3} + k_{4})$, is compared to k_{3}, and to each of the net rate constants after the chemical step. A low value of k_{11}, for example, can result in a large value of $c_{v_{f}}$, and a greatly reduced expression of the isotope effect on V.

The "slowest," "least-conducting," and "most-sensitive" steps may be the same in an enzymatic reaction, but need not be, so one must define what one means when using the term "rate-limiting" for the maximum velocity.

SEE ALSO THE FOLLOWING ARTICLES

Enzyme Inhibitors • Enzyme Kinetics • Enzyme Reaction Mechanisms: Stereochemistry • Kinetic Isotope Effects • Low Barrier Hydrogen Bonds • Metabolite Channeling: Creatine Kinase Microcompartments

GLOSSARY

Haldane relationship   An equation relating the equilibrium constant to the kinetic constants of an enzymatic reaction.

isotope effect   The effect of isotopic substitution, expressed as the ratio of the parameter for the light isotope to that for the heavy one.

Michaelis constant   The concentration of a substrate that gives half of the maximum velocity of an enzymatic reaction. It is the apparent dissociation constant in the steady state.

substrate   A molecule that is converted to a product during an enzyme-catalyzed reaction.

turnover number   The number of substrate molecules converted to product per enzyme molecule per unit time. The units are usually reciprocal seconds.

FURTHER READING

Cleland, W. W. (1982). An analysis of haldane relationships. *Meth. Enzymol.* **87**, 366–369.

Cleland, W. W. (1986). Enzyme kinetics as a tool for determination of enzyme mechanisms. In *Investigation of Rates and Mechanisms of Reactions* (C. F. Bernasconi, ed.) 4th edition, Part 1, pp. 791–870. John Wiley & Sons, New York.

Cleland, W. W., and Northrop, D. B. (1999). Energetics of substrate binding, catalysis, and product release. *Meth. Enzymol.* **308**, 3–27.

BIOGRAPHY

W. Wallace Cleland is M. J. Johnson Professor of Biochemistry at the University of Wisconsin and a co-director of the Institute for Enzyme Research. His research interest is in the use of kinetic methods for the determination of enzyme mechanisms, and in particular the use of isotope effects. He has a Ph.D. from the University of Wisconsin and was a postdoctoral Fellow at the University of Chicago. He is a member of the National Academy of Sciences and the American Academy of Arts and Sciences.

Sugar Nucleotide Transporters

Carlos B. Hirschberg
Boston University School of Dental Medicine, Boston, Massachusetts, USA

Sugar nucleotide transporters are proteins in the membrane of the Golgi apparatus. Their role is to translocate, from the cytosol into the Golgi lumen, nucleotide substrates for the glycosylation of proteins and lipids. Mutants in the above transporters have biochemical and developmental phenotypes, resulting in diseases such as leukocyte adhesion deficiency II.

A Requirement for Sugar Nucleotide Transport into the Golgi Apparatus Lumen

Approximately half of all proteins in eukaryotic cells are secreted or membrane-bound. These proteins are processed through the secretory pathway where 80% undergo post-translational modifications such as glycosylation, sulfation, and phosphorylation in the lumen of the Golgi apparatus. Intact nucleotide sugars, nucleotide–sulfate, and ATP, the substrates for these reactions, enter the lumen of the Golgi apparatus via specific transporter proteins (Figure 1). These are multi-transmembrane spanning proteins which function as antiporters with the corresponding nucleoside monophosphates and thereby concentrate the nucleotide derivatives in the Golgi lumen relative to their concentration in the cytosol (Figure 1). Evidence for this mechanism has been obtained through biochemical studies with rat liver-derived Golgi vesicles as well as with different species of yeast. In the latter, gene disruption of a Golgi lumenal GDPase, which generates GMP (the antiporter molecule coupled to entry of GDP-mannose into the yeast Golgi lumen), results in severe undermannosylation of proteins and lipids *in vivo*, reduced transport of the GDP-mannose into Golgi vesicles *in vitro*, as well as in a failure to make hyphae in *Candida albicans*.

Mutants in Nucleotide Sugar Transport—Biochemical and Developmental Phenotypes Including Diseases

Mutant mammalian cells grown in tissue culture have been described as having 95% reduced transport of specific nucleotide sugars into their Golgi apparatus, while their proteins and lipids have a drastic reduction in the specific sugar transported by the corresponding nucleotide derivative transporters. Multicellular organisms such as *C. elegans*, *Drosophila melanogaster*, plants, and humans that have mutations in the above proteins also have distinct morphological phenotypes affecting development of limbs, wings, brain, etc. Studies of these mutants have shown that some of the transporters can translocate more than one nucleotide sugar, whereas other transporters retain high substrate specificity such as differentiating between sugar epimers.

Mutants in nucleotide sugar transport have also been described in *Leishmania donovani* and in yeasts such as *S. cerevisiae*, *K. lactis*, and *C. albicans*.

Structure of Golgi Nucleotide Sugar Transporters

So far the primary amino acid sequence, together with the substrate specificity of approximately two dozen Golgi membrane nucleotide sugar transporters, has been determined. In most cases this was done by correcting the phenotypes of mutant cells or mutant multicellular organisms with libraries containing wild-type DNA or cDNA from homologous or heterologous organisms. In most cases after phenotypic correction, the specific DNA (or cDNA) was expressed in yeast or mammalian cells, Golgi vesicles were isolated, and transport of different nucleotide sugars assayed *in vitro*. Yeast is a particularly attractive system for
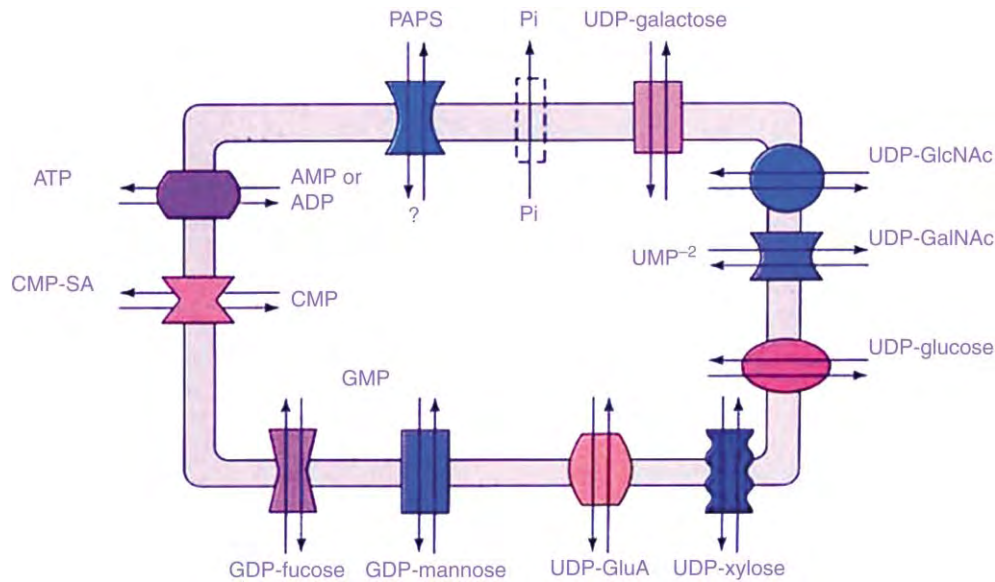
FIGURE 1   Golgi membrane nucleotide sugar, PAPS, and ATP transporters have been detected in mammals, yeast, protozoa, nematodes, insects, and plants. They are antiporters with the corresponding nucleoside monophosphate, except for PAPS, where the antiporter is not known, and for ATP, where it is either AMP or ADP or both.

heterologous expression of nucleotide sugar transporter genes, because it has few endogenous nucleotide sugar transport activities. This approach led to the identification of multisubstrate nucleotide sugar transporters from *C. elegans*. In some instances a particular cDNA was also tested for its ability to correct the phenotype of mutant cells where the substrate specificity of the defective nucleotide sugar is known.

The fact that the above approaches for determining the substrate specificity of particular transporters have worked is based on the following points:

1. the transporters DNA (or cDNA) can express the proteins in a heterologous system;

2. the targeting signals allowing the protein to localize to the Golgi apparatus are functional in different cell types; and

3. the targeted transporter is active once it reaches the Golgi apparatus of the recipient cell and can therefore affect phenotypic correction.

To date, the relationship of amino acid sequences and substrate specificities of nucleotide sugar transporters can be summarized as follows:

1. substrate specificities of nucleotides sugar transporters cannot be inferred even when amino acid sequences are $\sim 50\%$ identical;

2. only sequence identities of over 80% allow accurate predictions of substrate specificity; and

3. transporters from different organisms that have the same substrate specificity may have only 20% overall amino acid sequence identity.

Due to the above reasons the reader is cautioned that annotations about the identity (substrate specificity) of nucleotide sugar transporters in databases should be looked at very carefully!

Topography of Nucleotide Sugar Transporters in the Golgi Membrane

So far the only detailed topographic study of a nucleotide sugar transporter has been done with that for CMP-sialic acid in mammalian cells. Using a combination of tagged epitopes and immunocytochemistry, the transporter was found to have ten transmembrane domains and both its amino and carboxy termini facing the cytosol. A study on the topography of the *Kluyveromyces lactis* UDP-acetylglucosamine transporter using membrane accessibility of a peptide antibody and amino and carboxy termini epitope tags also found that both termini face the cytosol and that this transporter has either six or eight transmembrane domains. It is important to stress that none of the existing algorithms in the literature are designed specifically to address the issue of protein topography in the Golgi apparatus membrane which is somewhat thinner than the plasma membrane; most algorithms are designed for predicting topography in the plasma membrane.

Regulation of Macromolecular Glycosylation by Nucleotide Sugar Transporters

Studies with mutant mammalian cells in nucleotide sugar transporters grown in tissue culture as well as with fibroblasts and lymphoblasts from patients with leukocyte adhesion deficiency (LAD) syndrome II showed that decreased concentrations of nucleotide sugar in the Golgi lumen result in selective hypoglycosylation of proteins rather than across the board reduction. Thus, mutant MDCK cells which have a 95% reduction in transport of UDP-galactose into their Golgi lumen have a corresponding decrease in galactosylation of bulk glycoproteins, glycolipids, and keratan sulfate but not in heparan and chondroitin sulfates. LAD II patients appear to have normal O-fucosylation of Notch protein, whereas N-fucosylation of bulk proteins is drastically reduced. Although the detailed mechanism for these biochemical phenotypes is not understood, an attractive hypothesis is that when supply of nucleotide sugars is limited due to impairment in their transport, glycosylation reactions with lower K_ms would proceed, whereas those with higher K_ms would not.

SEE ALSO THE FOLLOWING ARTICLES

Glycosylation, Congenital Disorders of • Golgi Complex • Protein Glycosylation, Overview

GLOSSARY

glycosylation The covalent addition of sugars to each other or to proteins and lipids.

Golgi apparatus A membrane-enveloped organelle in the cytoplasm of all eukaryotes composed of several cisternae and vesicles; virtually all membrane-bound and secreted proteins of eukaryotes traverse through this organelle in route to their final location within or outside the cell.

nucleotide sugar Phosphodiesters, usually between nucleoside $5'$ diphosphates and carbon 1 of the sugar.

phenotype The visual or physical characteristics of individual cells or a multicellular organism.

FURTHER READING

Abeijon, C., Mandon, E. C., and Hirschberg, C. B. (1997). Transporters of nucleotide sugars, nucleotide sulfate and ATP in the Golgi apparatus. *Trends Biochem. Sci.* **22**, 203–207.

Hirschberg, C. B., Robbins, P. W., and Abeijon, C. (1998). Transporters of nucleotide sugars, ATP and nucleotide sulfate in the endoplasmic reticulum and Golgi apparatus. *Ann. Rev. Biochem.* **67**, 49–69.

Lodish, H., Berk, A., Zipusky, S. L., Matsudaira, P., Baltimore, D., and Darnell, J. (2000). *Molecular Cell Biology*. W.H. Freeman, New York.

Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., and Marth, J. (1999). *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, New York.

BIOGRAPHY

Carlos B. Hirschberg is a Professor and Founding Chairperson of the Department of Molecular and Cell Biology at the Boston University Goldman School of Dental Medicine. He holds a Ph.D. in Chemistry from the University of Illinois at Urbana-Champaign and received his postdoctoral training at the Harvard Medical School and MIT. His principal research interests are mechanisms of posttranslational modifications in eukaryotes. His laboratory co-workers discovered the transport/antiport system for nucleotide derivatives in the Golgi apparatus and endoplasmic reticulum membranes and were the first to purify and clone proteoglycan sulfotransferases.

SUMO Modification

Frauke Melchior and Andrea Pichler
Max-Planck Institute of Biochemistry, Martinsried, Germany

Small ubiquitin-related modifier, SUMO-1, and its homologues are 10–12 kDa eukaryotic proteins that serve to regulate protein function. Like their relative ubiquitin, they are covalently coupled to many different target proteins in the cell. Specific enzymes required for the formation or cleavage of isopeptide bonds between SUMO and its targets ensure specificity and dynamics of this posttranslational modification. Functional consequences reach from changes in protein-protein or protein–DNA interactions, alteration in subcellular localization, or enhanced stability, to changes in biological activity.

The SUMO Family

EXPRESSION

Proteins of the small *u*biquitin-related *mo*difier (SUMO) family appear to exist in all eukaryotic cells. For nomenclature see Table I. The number of distinct family members varies from one (e.g., in baker's and fission yeasts, nematodes, and fruit fly) to several (three in mammals; eight in *Arabidopsis thaliana*).Where present, different family members appear to have at least partially distinct functions. Interestingly, mammalian SUMO1 is mainly conjugated to substrates, while mammalian SUMO2/3 primarily exists in its free form under normal growth conditions, but is conjugated rapidly after stress stimuli. Whether SUMO proteins are differentially expressed, and their expression levels are regulated, is currently unknown. Available knockout data suggest that reversible SUMO conjugation is essential for life in most organisms. Fission yeast grows without a functional SUMO gene, but the cells are severely impaired.

STRUCTURE

SUMO proteins are small acidic proteins with distant homology to ubiquitin. At primary amino acid sequence, they are 10–20% identical to ubiquitin. On a structural level, their relatedness is much more pronounced. As shown in Figure 1, ubiquitin and SUMO share the classical ubiquitin-superfold. Characteristic for all members of the SUMO family, and absent from ubiquitin or other ubiquitin-related proteins, is an N-terminal flexible extension of 10–30 amino acids. The function of this extension is however currently unknown.

All SUMO proteins are expressed as precursors that require trimming of their C-terminus. Proteolytic removal of several amino acids, which is accomplished by specific enzymes (SUMO isopeptidases), results in exposure of a C-terminal glycine–glycine motif. Maturation is rapid, and there is currently no evidence for regulation of this process.

Enzymology

Covalent interaction between SUMO and its targets is a reversible, often dynamic, process. This is accomplished by formation and subsequent cleavage of an isopeptide bond between the carboxy-terminus of mature SUMO and a lysine of the acceptor protein. Isopeptide bond formation requires ATP and involves several (two or three) distinct enzymes (see Figure 2). In contrast, cleavage is accomplished by a single enzyme and does not require energy.

CONJUGATION

Overview

The first step towards isopeptide bond formation is activation of the SUMO carboxy-terminus. This is accomplished by ATP-dependent thioester bond formation between a cysteine residue in the E1 activating enzyme and the C-terminal glycine residue in SUMO. Subsequently, SUMO is transferred to a cysteine residue in the E2 conjugating enzyme and again the resulting bond is a thioester. Finally, an isopeptide bond is formed by transferring SUMO to the ε-amino group in a lysine of the acceptor protein. Depending on the specific acceptor protein, this step can be carried out by the E2 conjugating enzyme alone, or it may require an additional component, a so-called E3 ligase.

TABLE I

Nomenclature for SUMO and its Conjugating Enzymes

	Human	Baker's yeast
SUMOs	**SUMO1**/PIC 1/Ubl1/sentrin1/ GMP1/hSMT3	Smt3
	SUMO2/sentrin3/Smt3a	
	SUMO3/sentrin2/Smt3b	
E1 activating enzyme		
Subunit 1	**Aos1**/SAE1/Sua	**Aos1**
Subunit 2	**Uba2**/SAE2	**Uba2**
E2 conjugating enzyme	**Ubc9**/UBE2I	**Ubc9**
E3 ligating enzymes	**Pias1**/GBP/DEAD/H box-binding protein 1/ARIP	**Siz1**/UlI 1
	Pias3	**Siz2**/Nfi1
	Piasxα/ARIP3	
	Piasxβ/Miz1/Siz2	
	Piasy	
	RanBP2/Nup358	
	Pc2	

Known proteins involved in the SUMO conjugation pathway. Since most of them were independently identified by several groups, different names exist in literature.

SUMO Activating Enzyme (E1 Enzyme)

SUMO E1 activating enzyme is composed of two subunits, Aos1 and Uba2 (for alternative names see Table I), both of which are essential for the function. Aos1 and Uba2 share striking homology to the N- and C- terminal halves of ubiquitin E1 enzyme, respectively. SUMO E1 activating enzyme is present throughout the cell, with a strong enrichment in the nucleoplasm. A single SUMO E1 activating enzyme is expressed in most organisms, even those containing several distinct SUMO proteins. Current data suggest that mammalian SUMO E1 works with equal efficiency on each of the three mammalian SUMO proteins, although it has no affinity for other ubiquitin-related proteins.

SUMO Conjugating Enzyme (E2 Enzyme)

A single SUMO E2 conjugating enzyme, Ubc9, is expressed in cells (see Table I for alternative names). While it is strikingly similar to ubiquitin conjugating enzymes, it works exclusively on SUMO proteins. As for the E1 enzyme, Ubc9 does not appear to discriminate between different SUMO proteins expressed within a particular organism. Ubc9 localizes predominantly in the nucleus, but is also present in the cytoplasm and at nuclear pore complexes. A striking difference between the ubiquitin-conjugating and the SUMO-conjugating system is the ability of Ubc9 to directly bind to and modify SUMO targets. However, with a few exceptions, the efficiency of modification in the absence of E3 ligases is extremely poor.

SUMO Ligases (E3 Ligase)

Three types of SUMO E3 ligases are currently known (Table I). The first type is encoded by the
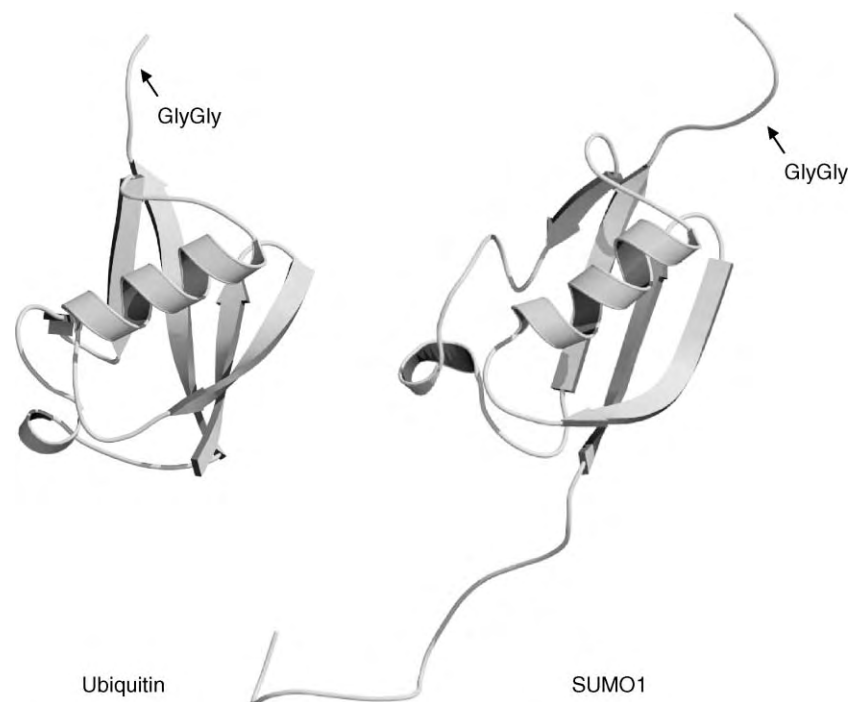
FIGURE 1 Ubiquitin and SUMO comparison. Although ubiquitin and SUMO1 share only 18% amino acid identity, the structural fold is very similar. The indicated C-terminal double glycine motif is the attachment site to substrates. The images (generated using Moscript 2.1.2 and Raster3D) were kindly provided by Dr. Peter Bayer.
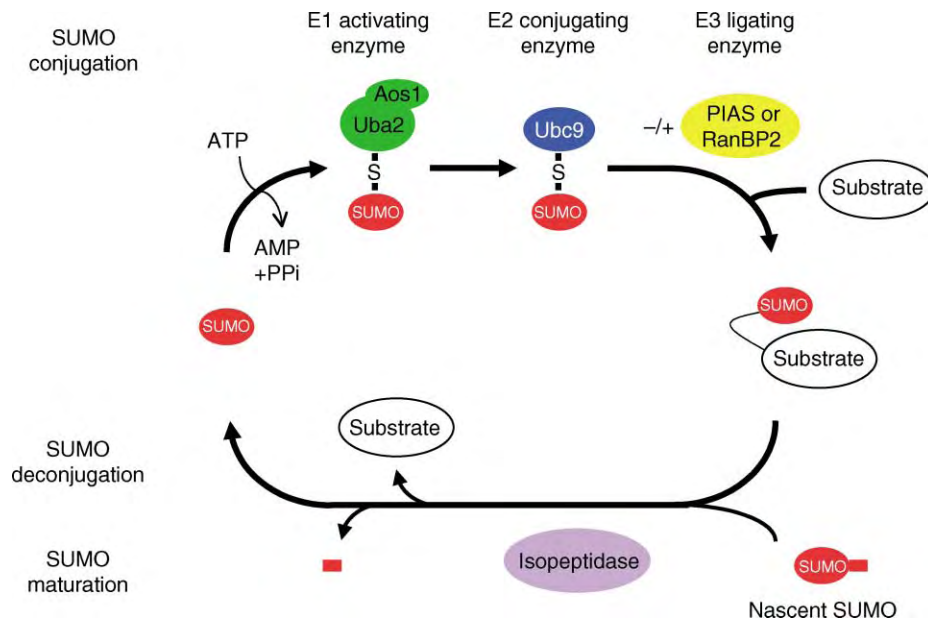
FIGURE 2 The SUMO modification pathway. SUMO is synthesized as a precursor that requires proteolytic processing. A C-terminal hydrolase/isopeptidase removes several C-terminal amino acids to expose the GlyGly motif essential for conjugation. SUMO conjugation involves an enzymatic cascade: In an ATP-dependent step, mature SUMO first forms a thioester with the E1 activating enzyme, the heterodimer Aos1/Uba2. Subsequently, it is transferred to the E2 conjugating enzyme, Ubc9. SUMO is then either directly or in an E3 ligase-dependent step conjugated to the substrate. SUMO de-modification is performed by isopeptidases.

"Protein inhibitor of activated STAT" (PIAS) family. These proteins contain a predicted RING finger-like structure that is essential for their function as E3 ligases. They bind directly to Ubc9 and to selected SUMO targets, and stimulate their modification both *in vivo* and *in vitro*. Baker's yeast contains two PIAS type E3 ligases, Siz1 and Siz2. These proteins seem to be responsible for most, but probably not for all, SUMO targets, and have partially overlapping target specificity. The second type is currently represented by a single (vertebrate specific) 358 kDa protein, RanBP2/Nup358 (Ran binding protein/nucleoporin). RanBP2 is a component of nuclear pore complexes, and thus may couple sumoylation and nucleocytoplasmic transport. Interestingly, the catalytic domain of RanBP2 does not contain a RING finger motif, and also shows no other homology to ubiquitin E3 ligases. The Polycomb group protein Pc2 appears to represent a third type of an E3 ligase, as it bears no similarity to other E3 ligases.

DECONJUGATION

Overview

Isopeptide bonds between ubiquitin-related proteins and their cellular targets are reversible in nature due to the presence of specific enzymes that recognize and cleave the bond. Currently, a single family of SUMO-specific isopeptidases is known (Table II). All members contain a 200 amino acid catalytic core that is distantly related to

viral cysteine proteases but shows no similarity to ubiquitin-specific enzymes.

Interestingly, SUMO isopeptidases serve two physiological functions *in vivo*, cleavage of isopeptide bonds, but also maturation of nascent SUMO (C-terminal hydrolase activity). This distinguishes the SUMO system from the ubiquitin system, where two distinct enzyme families serve to fulfill these functions.

SUMO Isopeptidases

The founding member of this family is *S. cerevisiae* Ulp1 (*u*biquitin *l*ike *p*rotease). Bakers yeast contains

TABLE II

Nomenclature for SUMO Deconjugating Enzymes

Human	Mouse	Baker's yeast
Senp1	SuPr-2	**Ulp1**
Senp2/KIAA1331	Senp2/**Smt3IP2**[*]/Axam2 and **Supr-1**[*]	**Ulp2**/Smt4
Senp3/Susp3/SSP3	Senp3/**Smt3IP1**/SuPr-3	
Senp5		
Senp6/**Susp1**/SSP1/KIAA0797		
Senp7/Susp2/SSP2/KIAA1707		

SUMO isopeptidases share a 200 amino acid core domain. Known proteins containing this core domain are listed in this table. The proteins characterized as SUMO isopeptidases are indicated in bold.

[*] Indicates different isoforms or alternative splice products.

two enzymes of this family, Ulp1 and Ulp2. Ulp1 is required for viability in yeast, whereas Ulp2 is not essential. They localize to different intracellular compartments, with Ulp1 being enriched at nuclear pore complexes, and Ulp2 being diffusely distributed throughout the nucleus. Mammals on the other hand have at least six distinct genes for SUMO isopeptidases (see Table II). Due to alternative splicing, the number of enzymes expressed is probably much larger. Only a few of these proteins have been characterized in some detail. They vary significantly in size and intracellular localization. It is plausible to assume that different isopeptidases exert different target specificity in large part through their physical localization. Whether they also differ in enzymatic properties, such as substrate specificity or kinetics of cleavage, remains to be seen.

SUMO Target Proteins

Sumoylation as a means to regulate protein function appears to be quite a common mechanism. Modification has been documented for more than 70 different target proteins, and this number is expected to increase significantly. Based on current knowledge, some generalizations can be made about the nature of the targets, motifs required for modification, consequences of modification, and regulation.

KNOWN TARGET PROTEINS

Based on immuno-fluorescence analysis, most targets for sumoylation are constitutively or transiently associated with the nuclear compartment. Consistent with this, most known targets can be associated with nuclear processes such as chromatin remodeling, DNA repair, transcription, or nucleocytoplasmic transport (amongst them are histone deacetylases, topoisomerases, thymine-DNA glycosylase, PCNA, p53, PML, heat shock factors, steroid hormone receptors, $I\kappa B\alpha$, RanGAP1, and many others). Other pathways to which SUMO has been linked are, e.g., progression through mitosis (mitotic arrest of yeast mutants defective in SUMO conjugation or deconjugation; yeast septins, topoisomerase II and Pds5 are modified specifically in mitosis), or infection with viruses (examples for viral SUMO targets are: cytomegalovirus immediate early proteins IE1 and IE2, adenovirus type 5 E1B-55 kDa, bovine papillomavirus E1).

CONSENSUS SITES FOR MODIFICATION

In contrast to ubiquitinylation, for which a unifying motif has not been identified, sumoylation of most targets seems to be specified by a short consensus sequence in target proteins. This motif consists of just four amino acids, $\Psi K \times E/D$, and includes the lysine residue that forms an isopeptide bond with SUMO. Ψ stands for a bulky aliphatic amino acid residue. Additional structural features are probably required to ensure accessibility to the conjugation machinery. Some motifs are present at the very N- or C-terminal end of a protein, others are flanked by proline residues that may induce a loop structure. Consistent with this, RanGAP1, which is an extremely efficient SUMO target, presents its sumoylation motif in an accessible loop.

FUNCTIONAL CONSEQUENCES FOR MODIFICATION

Similar to phosphorylation, sumoylation seems to have many different functional consequences that depend on the specific target protein. Considering SUMO's size, it is plausible to assume that conjugation can lead to masking of binding sites, generation of novel binding interphases, or conformational changes in the modified protein. Examples have been reported for changes in protein–protein or protein–DNA interactions, alteration in subcellular localization, enhanced stability through antagonizing ubiquitin/proteasome-mediated degradation, and changes in enzymatic activity.

REGULATION

While some SUMO targets appear to be modified constitutively, others are sumoylated only during a specific period of the cell cycle, upon stress, or upon a specific extracellular signal. Examples are sumoylation of yeast septins during mitosis, sumoylation of topoisomerase upon treatment with DNA-damaging agents, or *Dictyostelium* MEK1 sumoylation in response to chemoattractant. While this suggests the existence of elaborate regulatory mechanisms, current knowledge is rather poor. Increased modification of a specific target may for example be due to changes in the target, activation of a specific E3 ligase, or relocalization of a specific isopeptidase. Evidence for cell-cycle-dependent regulation of E3 ligases and isopeptidases is currently only available in yeast: (1) fission yeast ulp1 resides at the NPC during interphase, but is nuclear during mitosis; (2) baker's yeast E3 ligase Siz1 is intranuclear during interphase, but partially relocalizes to the bud neck in mitosis.

SEE ALSO THE FOLLOWING ARTICLES

Cysteine Proteases • JAK-STAT Signaling Paradigm • Nuclear Pores and Nuclear Import/Export • Ubiquitin System • Ubiquitin-Like Proteins • Zinc Fingers

GLOSSARY

cysteine protease   Protease that requires a cysteine for activity.

isopeptide bond   Any amide bond formed between a carboxyl group of one amino acid and an amino group of another amino acid with the exception of the conventional peptide bond that is formed between amino- and carboxy- groups in α-position.

nuclear pore complex   A large multiprotein complex embedded in the nuclear envelope that mediates both active transport of macromolecules and passive diffusion of small components to and from the nucleus.

RING finger proteins   A family of structurally related zinc-binding proteins containing the RING consensus sequence: $CX_2CX_{(9-39)}$ $CX_{(1-3)}HX_{(2-3)}C/HX_2CX_{(4-48)}CX_2C$. The cysteines and histidines represent the zinc-binding sites, whereby the first, second, fifth, and sixth of these complex the first zinc ion, and the third, fourth, seventh, and eighth complex the second.

FURTHER READING

Bernier-Villamor, V., Sampson, D. A., Matunis, M. J., and Lima, C. D. (2002). Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* **108**, 345–356.

Freiman, R. N., and Tjian, R. (2003). Regulating the regulators: Lysine modifications make their mark. *Cell* **112**, 1–17.

Goettsch, S., and Bayer, P. (2002). Structural attributes in the conjugation of ubiquitin, SUMO and RUB to protein substrates. *Front Biosci.* **7**, a148–a162.

Johnson, E. S. (2002). Ubiquitin branches out. *Nat. Cell Biol.* **4**, E295–E298.

Kim, K. I., Baek, S. H., and Chung, C. H. (2002). Versatile protein tag, SUMO: Its enzymology and biological function. *J. Cell Physiol.* **191**, 257–268.

Kurepa, J., Walker, J. M., Smalle, J., Gosink, M. M., Davis, S. J., Durham, T. L., Sung, D. Y., and Vierstra, R. D. (2003). The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is increased by stress. *J. Biol. Chem.* **278**, 6862–6872.

Melchior, F. (2000). SUMO – nonclassical ubiquitin. *Annu. Rev. Cell Develop. Biol.* **16**, 591–626.

Melchior, F., Schergaut, M., and Pichler, A. (2003). SUMO: ligases, isopeptidases and nuclear pores. *Trends Biochem. Sci.* **28**, 612–618.

Muller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001). SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* **2**, 202–210.

Schwartz, D. C., and Hochstrasser, M. (2003). A superfamily of protein tags: ubiquitin, SUMO and related modifiers. *Trends Biochem. Sci.* **28**, 321–328.

Seeler, J. S., and Dejean, A. (2003). Nuclear and unclear functions of SUMO. *Nat. Rev. Mol. Cell Biol.* **4**, 690–699.

Wilson, V. G., and Rangasamy, D. (2001). Viral interaction with the host cell sumoylation system. *Virus Res.* **81**, 17–27.

BIOGRAPHY

Frauke Melchior is a group leader at the Max-Planck Institute for Biochemistry in Martinsried, Germany. Her principal research interests are nucleocytoplasmic transport, and ubiquitin-related proteins of the SUMO family. She holds a Ph.D. from the University of Marburg and received her postdoctoral training at the Max-Planck Institute for Biophysical Chemistry in Göttingen and the Scripps Research Institute in La Jolla. She participated in the discovery of SUMO during her postdoctoral time, and has since contributed to the field with her own group.

Andrea Pichler obtained her Ph.D. from the University of Vienna and received a postdoctoral training at the Novartis Forschungsinstitut of Vienna. Since 2000 she is a member of Dr. Frauke Melchior's group at the Max-Planck Institute of Biochemistry in Martinsried.

Superoxide Dismutase

Irwin Fridovich
Duke University Medical Center, Durham, North Carolina, USA

A small fraction of the O_2 consumed by living cells is converted to superoxide O_2^-. This free radical, and its progeny, can damage a variety of biomolecules. This is the cause of oxidation stress. Defenses are provided by the superoxide dismutases that catalytically eliminate O_2^- and by the catalases and peroxidases that do the same for hydroperoxides. The varieties, distributions, and mode of action of the superoxide dismutases are presented herein.

Introduction

Molecular oxygen (O_2), while essential for the life of aerobes, is potentially toxic. This toxicity is due to the propensity of O_2 for reduction by a univalent pathway. This facile univalent pathway of O_2 reduction generates intermediates that lie between one O_2 and its four electron reduction products – two molecules of water – and it is the reactivity of these intermediates that is responsible for the toxicity of O_2. In order of their production, these intermediates are the superoxide anion radical (O_2^-), hydrogen peroxide (H_2O_2), and the hydroxyl radical ($HO\cdot$). Superoxide dismutases (SODs) catalyze the conversion of O_2^- into H_2O_2 plus O_2, i.e.,

$$O_2^- + O_2^- + 2H^+ \xrightarrow{SOD} H_2O_2 + O_2$$
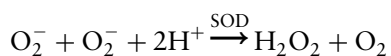
and in so doing provide an important defense. H_2O_2, in turn, is eliminated by the catalases and peroxidases. The concerted action of the SODs with the catalases and peroxidases prevents the formation of the very reactive $HO\cdot$. This is the case because $HO\cdot$ production *in vivo* requires both O_2^- and H_2O_2. O_2^- oxidizes the [4Fe–4S] clusters of dehydratases, such as the aconitases, causing release of Fe (II); the freed Fe (II) then reduces H_2O_2 to OH^- plus $HO\cdot$ in a reaction known as the Fenton reaction.

The SOD Family of Enzymes

The first truly photosynthetic organisms were the cyanobacteria and the oxygen they produced oxygenated a previously anaerobic biosphere. This imposed a common selection pressure on what must have been a varied anaerobic biota. It is not surprising that several different SODs were then evolved to deal with the toxicity of the accumulating O_2 and that these persist to this day. Thus, there are SODs based on Cu (II) plus Zn (II), Mn (III), Fe (III), and Ni (II). They all catalyze the dismutation of O_2^- into $H_2O_2 + O_2$.

All SODs work by a similar mechanism in which the metal at the active site is reduced by one O_2^- and then reoxidized by the next O_2^-. The active site metal thus acts as a mediator passing an electron from one O_2^- to the next. In this way the electrostatic repulsion, which would prevent close approach of one O_2^- to another, is bypassed by the SODs. All the SODs are very efficient catalysts and operate at close to the theoretical diffusion limit. Their rate constants for interaction with O_2^- are $\sim 3 \times 10^9$ $M^{-1}s^{-1}$.

THE CU, ZN SODS

These SODs have been found in the cytosols of eukaryotic cells, the intermembrane space of mitochondria, the periplasm of gram-negative bacteria, and in chloroplasts. The Cu (II) is linked to the Zn (II) at the active site by the imidazolate moiety of a histidine residue and this bridging imidazolate functions as a proton supply during the catalytic cycle. Thus, when the Cu (II)) is reduced by O_2^-, the Cu (I) releases the imidazolate that then protonates to imidazole. When the Cu (I) is reoxidized to Cu (II), the imidazole provides a proton to the nascent O_2^- converting it to HO_2^-, as the bridging to the Zn (II) is reestablished. The HO_2^- that leaves the active site protonates in the water to H_2O_2. This mechanism can be depicted as follows:

$$\text{Cu(II)} - \text{N} \diagup \overset{R}{\diagdown} \text{N} - \text{Zn(II)} \quad + \quad O_2^- \quad \xrightarrow{+H^+} \quad \left[\text{Cu(I)} \quad \text{NH} \diagup \overset{R}{\diagdown} \text{N} - \text{Zn(II)} \right] \quad + \quad O_2$$

$$\left[\text{Cu(I)} \quad \text{HN} \diagup \overset{R}{\diagdown} \text{N} - \text{Zn(II)} \right] \quad + \quad O_2^- \quad \longrightarrow \quad \text{Cu(II)} - \text{N} \diagup \overset{R}{\diagdown} \text{N} - \text{Zn(II)} \quad + \quad HO_2^-$$
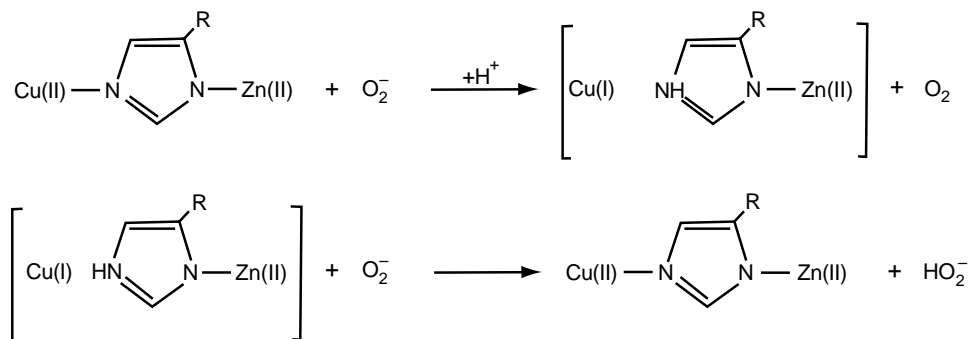
Both of these half-reactions of the catalytic cycle proceed with a rate constant of $\sim 3 \times 10^9 \, M^{-1} s^{-1}$. The enzyme surface contains charged groups that provide an electrostatic funnel guiding the O_2^- toward the active site crevice. This electrostatic facilitation contributes to the efficiency of these enzymes.

The Cu, Zn SODs found in eukaryotic cytosols are homodimeric proteins with subunit weights ~ 16 kDa. The corresponding enzyme from the periplasm of *Escherichia coli* is monomeric. There is an extracellular Cu, Zn SOD found in higher animals. It is usually homotetrameric and glycosylated and has a subunit weight of ~ 23 kDa.

THE MN SODS

These enzymes, whether from the matrix of mitochondria or bacteria, exhibit marked sequence similarity, reflecting a close evolutionary history and supporting an endosymbiotic origin for mitochondria. The bacterial enzyme is usually a homodimer, whereas the mitochondrial enzyme is a homotetramer. The subunit weight is ~ 23 kDa. Some bacterial Mn SODs are tetrameric. This is the case in Cryptococcus neoformans. In *E. coli* the biosynthesis of Mn SOD is under the control of the soxRS regulon, which coordinately up-regulates the expression of a number of genes in response to O_2^-. The constitutively expressed SOX R protein is transcriptionally inactive in its reduced form. It can be oxidized by O_2^- and then activates the expression of the SOX S protein, which in turn activates all the genes in the regulon. Thus, Mn SOD is not measurable in extracts of anaerobically grown *E. coli*, but exposure of cultures to aerobic conditions elicits production of Mn SOD. Increasing production of O_2^-, by raising $p\text{O}_2$, or by adding compounds such as viologens, which can mediate enhanced production of O_2^-, increases the level of Mn SOD. It has been possible to force *E. coli* to produce Mn SOD to 7% of its soluble protein by aerobic exposure to the viologen paraquat.

The nectar of tobacco flowers has been found to contain a stable Mn-protein named nectarin that appears to be an Mn SOD.

THE FE SODS

These SODs, which are highly homologous to the Mn SODs, are found in bacteria and in plants. Although usually homodimeric, homotetrameric Fe SOD has been found in Rhodococcus bronchialis and Mycobacterum tuberculosis. The Fe SOD of *E. coli* is constitutive and thus is found even in anaerobically grown cells. It can thus be viewed as a standby defense against O_2^-, which is always maintained to protect in the event of a sudden exposure to O_2.

It is possible to reversibly remove the metals from SODs. The apo enzymes so prepared are inactive but can be reactivated by restoring the active site metal. It is striking, that in spite of sequence and structural homologies, the Fe SOD is inactive when made to contain Mn in place of Fe and the Mn SOD is inactive when similarly reconstituted with the non-native metal Fe. There are, however, SODs that are active with either Mn or Fe at their active sites. These SODs, termed cambialistic SODs, are usually found in anaerobes such as *Propionibacterium shermanii* or *Streptococcus mutans*. These organisms produce the SOD with iron when grown anaerobically but, when exposed to low levels of O_2, produce the same SOD but with Mn. It is clear that even anaerobes have need of a defense against the O_2^- during transient exposure to O_2. Ferrous salts are soluble, and thus available anaerobically, but become insoluble when oxidized to the ferric state by O_2. Manganous salts remain soluble under both conditions – hence the wisdom of a cambialistic SOD for an anaerobe. An iron SOD has also been reported in the protozoan *Tetrahymena pyriformis*.

Deletions and Consequences

E. COLI

Mutants of *E. coli* lacking both the Mn SOD and the Fe SOD (SodA SodB) have been prepared. Anaerobically they grow as well as the parental strain. However, aerobically they grow slowly, exhibit a high rate of mutagenesis, and require branched-chain,

sulfur-containing, and aromatic amino acids. The null mutants are also hypersensitive towards paraquat. All of these phenotypic deficits disappear when a functional gene encoding any active SOD is inserted. The slow growth of the null mutant can be understood in terms of the energy devoted to repairing or replacing oxidatively damaged molecules and to the inactivation of such O_2^- sensitive enzymes as aconitase. The high rate of O_2-dependent mutagenesis reflects oxidative DNA damage, and repair, which is error-prone. The nutritional auxotrophies exhibited by the null mutant have several explanations. The need for branched amino acids arises because the penultimate step in the biosynthesis of these amino acids is catalyzed by a dihydroxy acid dehydratase that contains a [4Fe–4S] cluster that is rapidly oxidized, and inactivated, by O_2^-. The need for sulfur-containing amino acids evidently reflects leakiness of the cell envelope towards sulfite, which is an intermediate on the pathway from sulfate to cysteine. Finally, the aromatic amino acid auxotrophy reflects the inability to make erythrose-4-phosphate, which is one of the starting materials on the pathway leading to these amino acids. Erythrose-4-phosphate, in turn, depends on the sequential actions of transketolase and transaldolase, and the transketolase intermediate 1, 2-dihydroxyethylthiamine pyrophosphate is rapidly oxidized by O_2^-. The varied explanations for these phenotypic deficits reflect the variety of damage that can be caused by O_2^- and the multiple targets that are protected by the SODs.

Yeast

Mutants of the yeast *Saccharomyces cerevisia* lacking either the cytosolic Cu, Zn SOD or the mitochondrial Mn SOD have been prepared. Their phenotypic deficits, which are all oxygen-dependent, include slow growth, hypersensitivity towards paraquat or quinones, vacuolar fragmentation, sensitivity towards oxygen, and, in the case of the Cu, Zn SOD-null, lysine auxotrophy. All of these problems can be explained on the basis of O_2^- damage, as was the case for the *E. coli* SOD-null mutants.

Mice

Murine mutants lacking Mn SOD are severely affected and those that survive to term are low-birth-weight and only survive for 4–14 days. Their problems can be seen as a deficit of mitochondrial functions and that in turn can be explained on the basis of damage to mitochondrial components by O_2^-. The heterozygotes that have half the normal level of Mn SOD are viable but have been shown to be hypersensitive towards oxidative inactivation of aconitase and towards release of cytochrome *c* from mitochondria when challenged with tertiary butyl hydroperoxide. The deficit imposed by halving the Mn SOD was also seen as a greater susceptibility to reperfusion injury.

In contrast to the embryonic and neonatal fatality imposed by Mn SOD deletion, the homozygous Cu, Zn SOD-null mice were apparently normal under laboratory conditions. They are reported to be relatively intolerant of neuronal injury and prone to develop hepatoma at >9 months of age. It seems likely that there is some redundancy in scavenging O_2^- in the cytosol, so that the lack of Cu, Zn SOD may be compensated by up-regulation of another O_2^- scavenger, perhaps a superoxide reductase, such as has been found in anaerobes.

SOD and Amyotrophic Lateral Sclerosis

ALS or Lou Gherig's disease is a late onset, progressive, and fatal paralytic disease due to death of motor neurons. There are sporadic and familial types of ALS and these are not distinguishable on the basis of clinical symptoms. Approximately 20% of the familial cases of ALS have been associated with point mutations causing single amino acid replacements in Cu, Zn SOD. Transgenic mice expressing the wild-type human Cu, Zn SOD are normal, but those expressing one of the ALS-associated mutant Cu, Zn SODs develop progressive paralysis at ~ 3 months of age. Thus, the mutations in the SOD cause the death of motor neurons because of some as yet unknown gain of function, rather than because of a loss of SOD activity. This is certainly the case because: most of the mutant SODs retain full SOD activity; the transgenic mice retain the active murine Cu, Zn SOD; the disease is genetically dominant, i.e., heterozygotes develop the disease; and finally, Cu, Zn SOD knockout mice do not become paralyzed. The transgenic mice provide a useful model of the human disease and are being used to explore both the toxic gain of function of the mutant Cu, Zn SOD and treatment modalities.

Mimics

O_2^- has been found to play causative roles in numerous inflammatory diseases and reperfusion injuries. For this reason non-enzymic catalysts of the dismutation of O_2^- are being explored as pharmaceutical agents and appear to have promise.

See Also the Following Articles

DNA Oxidation • Free Radicals, Sources and Targets of: Mitochondria • Iron–Sulfur Proteins

GLOSSARY

apo enzyme   Enzyme whose prosthetic group has been removed.
Cu, Zn SOD   Superoxide dismutases with copper and zinc at the active site.
Fe SOD   Superoxide dismutases with iron at active site.
HO·   The hydroxyl radical.
homodimer   Protein composed of two identical subunits.
homotetramer   Protein composed of four identical subunits.
Mn SOD   Superoxide dismutases with manganese at an active site.
O_2^-   The superoxide anion radical.

FURTHER READING

Hassan, H. M. (1997). Cytotoxicity by oxyRadicals and the evolution of superoxide dismutases. In *Lung Biology in Health and Disease* (L. B. Clerch and D. J. Massaro, eds.) Vol 105, pp. 27–47. Marcel Dekker, New York.

Maier, C. M., and Chan, R. H. (2002). Role of superoxide dismutases in oxidative damage and neurodegenerative disorders. *Neuroscientist* **8**, 323–334.

McCord, J. M. (2002). Superoxide dismutases in aging and disease. *Methods Enzymol.* **349**, 331–341.

Melov, S. (2002). Therapeutics against mitochondrial oxidative stress in animal models of aging. *Ann. N.Y. Acad. Sci.* **959**, 330–340.

Storz, G., and Imlay, J. A. (1999). Oxidative stress. *Curr. Opin. Microbiol.* **2**, 188–194.

BIOGRAPHY

Irwin Fridovich is a James B. Duke Professor of Biochemistry, Emeritus. He is a member of the National Academy of Science and of the National Academy of Arts and Sciences. He and his students discovered and characterized the superoxide dismutase.

Syk Family of Protein Tyrosine Kinases

Andrew C. Chan
Genentech, Inc., San Francisco, California, USA

The Syk (spleen tyrosine kinase) family of protein tyrosine kinases (PTKs) encode essential signaling components required for normal immunity. Their functions have been most intensely studied within mammalian immune cells. While not found within the *Caenorhabditis elegans* genome, this family of PTKs is represented earlier within the phylogenetic tree, in the hydra vulgaris, as a single gene product expressed in epithelial cells and plays an important function in the recognition of foreign cells. In mammalian systems, this family of PTKs appears to have evolved from a gene duplication event to give rise to its two family members – ZAP-70 and Syk in mammalian cells. Genetic studies in humans and mice have demonstrated their essential roles in the function and development of T cells, B cells, mast cells, monocytes/macrophages, and the lymphatic system. Studies further underscore their importance in T cell antigen receptor (TCR), B cell antigen receptor (BCR), IgG and IgE receptors (FcRs) and integrin receptor signaling. This review will discuss our current understanding of this family of PTKs in mammalian immune cell function.

Zeta-Chain-Associated Protein of 70K Mr (ZAP-70)

ZAP-70 was initially identified as a 70K Mr tyrosine phosphorylated protein that associates with the TCR following receptor activation. Biochemical and molecular characterization revealed that ZAP-70 is a PTK with the characteristic Syk-family signature of tandem Src-homology (tSH2) domains at its amino (N) terminus and a carboxy (C)-terminal catalytic domain. The domain between the two SH2 domains has been termed Interdomain A, while Interdomain B spans between the C-terminal SH2 and the C-terminal kinase domain (Figure 1).

INTERACTION OF ZAP-70 WITH THE TCR ITAM

Binding studies and, ultimately, the solution of the crystal structure encoding the tandem SH2 domains revealed that both SH2 domains of ZAP-70 cooperate to bind the doubly phosphorylated immunoreceptor tyrosine based activation (dP-ITAM) motifs (Figure 2). The ITAMs consist of the consensus sequence D/E x x Y x x L/I X_{6-8} Y x x L/I that is represented in the cytoplasmic domains of integrins as well as each of the signaling subunits of the TCR, BCR, and FcRs. Both tyrosine residues within the ITAM are phosphorylated by Src-family PTKs that, in turn, provide the high-avidity binding sites for the tSH2 domains of ZAP-70. While the C-terminal SH2 domain binds the N-terminal ITAM tyrosine, the C-terminal SH2 domain forms a portion of the pocket for the N-terminal SH2 domain that correspondingly binds the C-terminal ITAM tyrosine residue. This interdependence of the two SH2 domains explains the rigid requirements for the tandem SH2 domains of ZAP-70 as well as the little flexibility of spacing between the two tyrosine residues within the ITAM sequence in binding ZAP-70. Mutation of either SH2 domains or mutation of either ITAM tyrosine results in >100-fold decrease in binding avidities and, in turn, a non-functional TCR.

While ZAP-70 was initially found to be inducibly associated with the TCR in Jurkat T cells, subsequent studies in human and murine thymocytes and peripheral T cells revealed that the TCR ζ-subunit is already phosphorylated in its basal unactivated state, though the degree of phosphorylation is further augmented following TCR activation. In turn, in resting thymocytes and peripheral T cells, ZAP-70 is constitutively associated with the TCR, though the degree of association appears to be augmented concomitant with the degree of ζ-phosphorylation following receptor activation. A model of sequential phosphorylation of the ζ-chain was proposed based on observations in the 3.L2 T cell hybridoma, but this sequence does not appear to apply to thymocytes nor normal peripheral T cells. In addition to localizing ZAP-70 to the TCR complex, the tSH2:ITAM interaction has also been proposed to regulate ZAP-70 enzymatic activity. *In vitro* binding of ZAP-70 to a dpITAM peptide results in enhanced ZAP-70 activity. This model of ZAP-70 activation, however, is at odds with the pre-existing ZAP-70:ITAM interaction
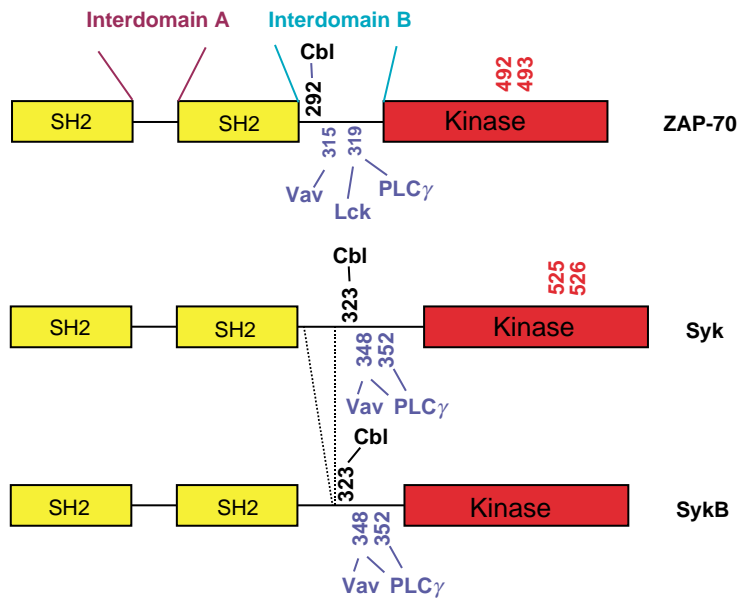
FIGURE 1 Schematic representation of ZAP-70 and Syk PTKs. Schematic structural representation of the structural domains within the Syk PTKs. T-loop tyrosine residues are represented in red; positive regulatory Interdomain B tyrosine residues are represented in blue; negative regulatory Interdomain B tyrosine residues are represented in black. Numbering utilizes the human protein sequences for both ZAP-70 and Syk.

described in normal human thymocytes, murine thymocytes, and peripheral T cells.

In contrast to the classical tSH2:dPITAM interaction, the crystal structure and NMR studies of the tSH2 domains in the absence of an ITAM peptide reveals two highly flexible independent SH2 domains. These additional forms of interaction may provide the structural basis for more non-classical interactions. For example, the N-terminal SH2 domain alone in conjunction with Interdomain A of both ZAP-70 and Syk bind the NXXY motifs within the cytoplasmic domain of the $\beta 3$ integrin. Of note, tyrosine phosphorylation of the NXXY motif inhibits binding to the SH2 domain. These more non-classical interactions may, in part, explain immunofluorescence studies that demonstrate targeting of ZAP-70 to the T cell cortex independent of the tSH2 domains. In contrast, studies utilizing fluorescence imaging and immunofluorescence of green fluorescent protein-tagged ZAP-70 is more consistent with a cytosolic distribution of ZAP-70 in resting cells with redistribution to the plasma membrane following TCR engagement.
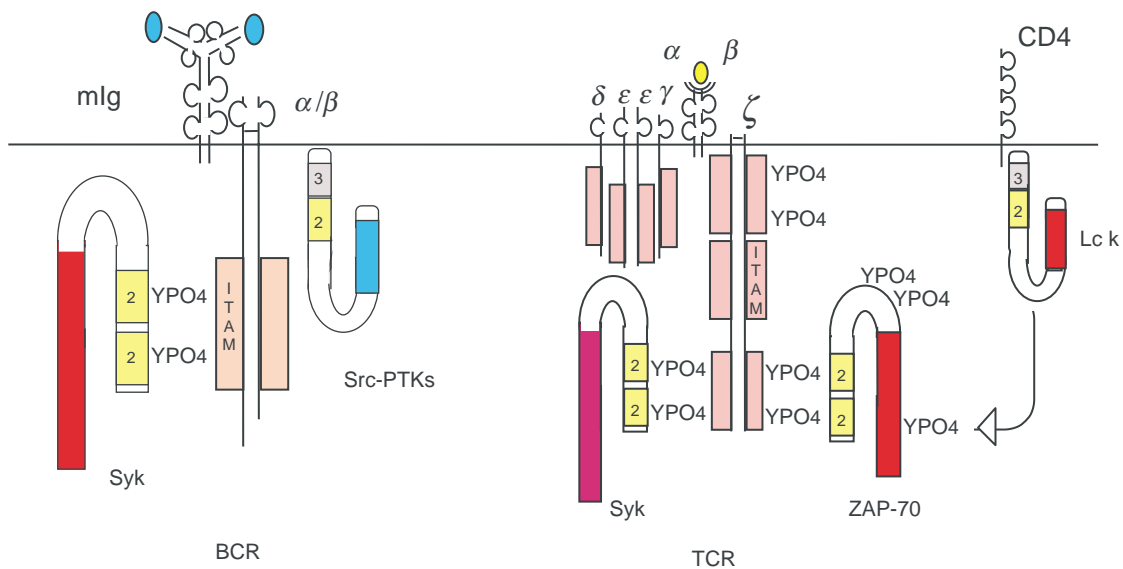
FIGURE 2 Schematic representation of T and B cell antigen receptors with Src- and Syk-PTKs.

Tyrosine Phosphorylation of ZAP-70

ZAP-70 undergoes a sequence of regulated phosphorylation events on multiple tyrosine residues that serve both positive and negative regulatory functions. The synergistic interactions between ZAP-70 and Src-PTKs was initially appreciated in overexpression studies in heterologous cell systems. While expression of ZAP-70, Lck or Fyn alone does not induce tyrosine phosphorylation of cellular proteins, co-expression of ZAP-70 with Lck or Fyn results in tyrosine phosphorylation of multiple cellular proteins. Studies in Jurkat T cells lacking either ZAP-70 or Lck further substantiate the requirement for both PTKs in efficient TCR function. Molecular dissection of the Lck/ZAP-70 interaction reveals multiple levels of regulation through phosphorylation of distinct tyrosines within ZAP-70.

Catalytic Domain: Tyr 492 and 493

Upon TCR engagement with major histocompatibility complex (MHC)-peptide complexes presented on antigen presenting cells (APCs), studies utilizing cellular fractionation, Forster resonance energy transfer analysis, chemical cross-linkers and microscopy have demonstrated the co-localization of CD4 and CD8 co-receptors with the TCR-centered "synapse". In this model, the synapse colocalizes a number of critical signaling molecules (e.g., TCR, ZAP-70, CD4 and Lck) to facilitate downstream signaling. As the cytoplasmic domains of both CD4 and CD8 interact with Lck, Lck is co-localized with ZAP-70 where it can trans-phosphorylate Tyr 493 within the trans-activation (T) loop of the ZAP-70 PTK (Figure 1). This initial phosphorylation results in the enzymatic activation of ZAP-70 and is required for the generation of second messengers (e.g., calcium mobilization and Ras activation). Within the T-loop, a hierarchy of phosphorylation occurs with the initial phosphorylation of Tyr 493, followed by phosphorylation of the neighboring Tyr 492. Mutation of Tyr 492 to Phe results in a hyperactive ZAP-70 PTK and heightened TCR functions and implicates a potential inhibitory function of Tyr 492 phosphorylation. Hence, the T-loop of ZAP-70 contains both positive and negative regulatory tyrosine residues that can finely modulate ZAP-70 enzymatic activity in a temporal fashion.

Interdomain B: Tyr 315 and 319

In addition to phosphorylation within ZAP-70's catalytic domain, ZAP-70 is phosphorylated on three tyrosine residues within its Interdomain B. Phosphorylation of Tyrs 315 and 319 serve scaffolding functions. The SH2 domains of Lck and PLCγ1 interact with Tyr 319 while the SH2 domain of Vav binds Tyr 315. Phosphorylation of these tyrosine residues is thought to serve scaffolding functions by which signaling complexes may be assembled and/or stabilized. Additionally, Tyr 315 may also have effects in determining the optimal tSH2 binding to the TCR ITAM sequences. Mutation of either tyrosine residues attenuates TCR-induced calcium and MAPK activation. Mice that express mutant ZAP-70 molecules in which Tyrs 315 or 319 are mutated to Phe demonstrate compromised T cell development. Given the homologous sequences surrounding these two tyrosines, it is likely that both tyrosine residues play overlapping functions.

Interdomain B: Tyr 292

Phosphorylation of Tyr 292 plays a negative regulatory function potentially through its interaction with the SH2-like (TKB) domain of the Cbl E3 ligase. Expression of a mutant ZAP-70 molecule in which Tyr 292 is mutated to Phe results in prolonged TCR signaling and is consistent with the inability of an activated and phosphorylated ZAP-70 to undergo ubiquitination and degradation. In turn, the TCR initiated signal lacks one of its normal extinguishing mechanisms and results in prolonged signaling. While Tyr 292 clearly plays a negative regulatory role in TCR activation, the mechanism through its interaction with Cbl has been called into question. The converse loss-of-function mutant in the Cbl TKB domain demonstrates altered activation of the Rac GTPase without any alteration to ZAP-70 enzymatic activity. Additional studies will be required to further define the molecular mechanism(s) by which Tyr 292 phosphorylation desensitizes TCR function.

Tyrs 474, 597, and 598

Given the precedent of the negative regulatory C-terminal tyrosine residues conserved within Src-PTKs, studies have also focused on a conserved series of tyrosine residues within the C-terminal of ZAP-70s catalytic domain (Tyrs 597 and 598). While these tyrosine residues have not been formally demonstrated to be phosphorylated *in vivo* following TCR engagement, mutation of either C-terminal tyrosine residues results in enhanced TCR-mediated IL-2 regulated promoter activity. In addition, Tyr 474 has been proposed to interact with the Shc adaptor protein. Likewise, expression of a mutant ZAP-70 with Tyr 474 mutated to Phe results in loss of IL-2 promoter activity without alterations in ZAP-70 activity or tyrosine phosphorylation of cellular proteins.

Posttranslational modification of ZAP-70 through tyrosine phosphorylation plays a major role in regulating ZAP-70 function. In addition to tyrosine phosphorylation, ZAP-70 is also phosphorylated on multiple

serine and threonine residues, the functions of which have not been well investigated. Hence, additional mechanisms likely exist to regulate ZAP-70 function.

Spleen Tyrosine Kinase (Syk)

Syk was initially identified as a 40K Mr (p40) peptide with intrinsic tyrosine kinase activity. Molecular characterization of Syk from a spleen cDNA library revealed that p40 represents a proteolytic fragment of a 72K Mr holoenzyme. While Syk is structurally homologous to ZAP-70 with tSH2 domains at its N terminus and a C-terminal catalytic domain, Syk has distinct regulatory mechanisms, when compared to ZAP-70, and has been implicated in a greater number of receptor signaling systems, in part, due to its more ubiquitous pattern of expression. Syk plays important functions in pre-TCR, BCR, FcR, IL-15R, and integrin signaling.

INTERACTION OF SYK WITH THE BCR ITAM

The tSH2 domains of Syk are similarly responsible for binding the dpITAM encoded within the Igα and Igβ signaling subunits of the BCR (Figure 2). Solution of the crystal structure of the Syk tSH2 domain complexed to a dpITAM revealed the lack of structural inter-dependence between the two SH2 domains. In fact, the two SH2 domains fold into independent SH2 domains that each bind a phosphorylated tyrosine residue within the dpITAM sequence. Moreover, a high degree of rotational flexibility was observed within Interdomain A, which may confer the ability of the Syk tSH2 domains to bind dpITAM sequences that have longer spacing between the two pY residues within the ITAM.

While ZAP-70 appears to be already associated with the dpITAMs within the TCR, Syk does not appear to be associated with the dpITAM in resting cells. In contrast, BCR or FcR engagement results in ITAM phosphorylation by Src-PTKs and the subsequent recruitment of Syk to the activated receptor. In addition to localizing Syk to the receptor, the dpITAM also plays an important role in Syk enzymatic activation. Hence, the tSH2 domains play both localizing and activation roles of the holoenzyme.

ALTERNATIVE SPLICING OF SYK: SYK AND SYKB

Syk is expressed in two different forms, Syk and SykB, as a result of alternatively splicing (Figure 1). SykB lacks a 23 amino acid sequence within the interdomain B region. This shortened form has comparable enzymatic activity and tyrosine phosphorylation events as Syk, but altered ability to bind the dpITAMs. As such, interdomain B can regulate Syk function though the *in vivo* significance of the SykB splice form remains unclear.

TYROSINE PHOSPHORYLATION OF SYK

Interdomain B: Tyrs 348 and 352

Three tyrosine residues within Interdomain B undergo phosphorylation following BCR engagement—Tyr 323, 348, and 352. Phosphorylation of Tyrs 348 and 352 are required for linking the BCR with PLCγ2 activation. Both sites can bind the SH2 domains of PLCγ while Tyr 348 can also bind the SH2 domain of Vav. These pTyr-SH2 interactions may contribute to signaling by localizing PLCγ and Vav effectors to the activated BCR complex, to facilitate the tyrosine phosphorylation and enzymatic activation of these effector molecules, and, in turn, facilitate membrane localization of these activated enzymes where their substrates (i.e., PIP2 and GTPases) normally reside. Mutation of these Interdomain B sites results in loss-of-function mutants of Syk. Hence phosphorylation of Tyrs 348 and 352 are thought to contribute a positive regulatory function for the Syk PTK.

Interdomain B: Tyr 323

While phosphorylation of Tyrs 348 and 352 results in gain of function, phosphorylation of Tyr 323 results in decreased receptor functions. Phosphorylation of Tyr 323 occurs in a Lyn-dependent fashion and facilitates the binding of Syk to the c-Cbl E3 ligase. Binding of c-Cbl to Syk initiates the ubiquitination pathway to down-regulate receptor-initiated signaling events including ubiquitination and degradation of Syk protein levels.

Catalytic Domain: Tyr 525 and 526

Similar to ZAP-70, Syk also has two tyrosine residues within the T-loop of its catalytic domain. However, unlike the hierarchy of phosphorylation within ZAP-70, both tyrosine residues within Syk are phosphorylated following receptor cross-linking and contribute to Syk enzymatic activation. Mutation of either tyrosine residues to phenylalanine results in attenuated activation of the resultant mutant enzyme. Also to be differentiated from ZAP-70 in which its activation results from trans-phosphorylation by the heterologous Src-PTKs, enzymatic activation of Syk appears to be mediated through trans-autophosphorylation. A recent study using a heterologous expression system suggests a slight variation in this model in which Syk functions as an allosteric enzyme that is positively regulated by ITAM phosphorylation. Within this model, Src-PTKs

phosphorylate the N-terminal ITAM tyrosine residue, while Syk can phosphorylate either ITAM tyrosine residues. Binding of Syk to the dpITAM induces a conformational change in the holoenzyme to induce amplification of its auto-activation and downstream signaling functions.

Tyrs 130, 629, 630, and 631

Tyrosine 130 within Interdomain A may play a role in regulating the release of Syk from the BCR. Mutation of Tyr 130 to Phe results in enhanced binding of Syk to the BCR; conversely, mutation of Tyr 130 to Glu reduced this interaction but results in enhanced Syk enzymatic activation. Finally, the C-terminal tyrosine residues 629, 630, and 631 of Syk, while also not having been demonstrated to be phosphorylated *in vivo*, can potentially serve as negative regulators of Syk function. Mutation of these C-terminal tyrosines results in an enhanced Syk PTK.

ZAP-70 and Syk PTKs in Hematopoietic Cell Function

The ubiquitous expression of the Syk PTKs amongst hematopoietic derived cells is consistent with the functional requirement for this family of PTKs in a multitude of receptor and cell type functions. While ZAP-70 has a more limited cellular expression, this PTK has been demonstrated to play important roles in $\alpha\beta$ TCR, $\gamma\delta$ TCR, integrin, and pre-BCR functions. The more ubiquitous expression of Syk is consistent with its demonstrated roles in pre-TCR, $\gamma\delta$ TCR, FcR, IL-15R, $\alpha2b\beta3$ integrin and collagen receptor-mediated functions.

T Lymphocyte Development

T cell development begins in the thymus where signaling through the pre-TCR promotes α-chain rearrangement and differentiation of CD4$^-$CD8$^-$ (double negative or DN) cells to immature CD4$^+$CD8$^+$ (double positive or DP) cells. Development of DP thymocytes into mature CD4$^+$ or CD8$^+$ single positive (SP) T cells is subsequently driven by selection events that require signals transduced through the $\alpha\beta$ TCR. A quantitative model of T cell selection has been proposed in which strong self-reactive T cells results in apoptosis through a process known as negative selection; T cells that do not recognize self-MHC have no signaling through the TCR and hence also undergo apoptosis through a process known as "death by neglect." Only T cells that recognize self-MHC in the absence of self-antigens are thought to have an "appropriate" level of TCR

signal strength to promote DP thymocytes to differentiate to SP thymocytes.

ZAP-70 in $\alpha\beta$ T Cell Development

The functions of ZAP-70 during both pre-TCR and $\alpha\beta$ TCR signaling have been elucidated through an elegant series of genetic studies of lymphocyte development in mice and natural mutations in humans. Patients lacking ZAP-70 were described in the early 1990s that present with a selective CD8$^+$ T cell deficiency. While CD4$^+$ T cells are present in the peripheral blood of these immunodeficient patients, these cells are non-functional and lack the ability to proliferate to TCR induced signals. Thymic histology revealed the presence of DP thymocytes within the medulla, but the absence of CD8$^+$ thymocytes in the cortex. The molecular basis of these mutations are multiple and include missense mutations within the catalytic domain, truncation mutations, and altered splice acceptor sites that give rise to unstable proteins.

In contrast to the selective developmental defect in humans, mice engineered to be deficient in *zap-70* through homologous recombination demonstrate a block at the transition of DP to mature CD4$^+$ and CD8$^+$ cells. In turn, *zap-70* $^{-/-}$ mice accumulate DP thymocytes without SP thymocytes or peripheral $\alpha\beta$ T cells. Analysis of *zap-70* $^{-/-}$ mice that express a transgenic TCR reveal an essential role for ZAP-70 in both positive and negative selection.

A spontaneously arising mutation in the DLAARN motif within the mouse ZAP-70 catalytic domain that abrogates kinase activity similarly results in an arrest in thymocyte development at the DP T cell stage. An identical mutation has been described in a SCID infant with non-functional peripheral CD4$^+$ and absent peripheral CD8$^+$ T cells. Hence, the identical mutation within ZAP-70 results in distinct T cell developmental phenotypes in humans and mice. This difference may be due to differential expression and regulation of the Syk PTK during human and mouse thymic development.

Syk in $\alpha\beta$ T Cell Development

While *syk* $^{-/-}$ mice were initially thought to have normal T cell development and hence no role in T cell function, recent studies demonstrated overlapping and potentially independent functions of Syk and ZAP-70 during T cell ontogeny. In contrast to *zap-70* $^{-/-}$ mice that are blocked at the DP thymocyte stage, mice deficient in both *zap-70* and *syk* do not develop any DP T cells and accumulate DN thymocytes. Moreover, the pre-TCR expressed in *zap-70* $^{-/-}$ *syk* $^{-/-}$ DN thymocytes is non-functional. Hence, either ZAP-70 or Syk can mediate pre-TCR function while ZAP-70, in part due to the down-regulation of Syk following pre-TCR

signaling, plays a unique function in the transition from DP to SP T cells. Consistent with this quantitative explanation, forced expression of Syk in *zap-70* $^{-/-}$ mice restores T cell development to the SP T cell stage. Hence, both ZAP-70 and Syk play overlapping functions during T cell development.

While T cell development appears normal in *syk* $^{-/-}$ mice, studies using hematopoietic chimeras with *syk* $^{-/-}$ hematopoietic stem cells (FL-HSCs) suggest a unique function of Syk in early T cell development. *Syk* $^{-/-}$ FL-HSCs demonstrate compromised ability to reconstitute T cells in *rag2* $^{-/-}$ mice at the CD44$^-$CD25$^+$ stage – the stage in which pre-TCR signaling is important. In addition, $\sim 50\%$ decrease in T cell reconstitution is observed with *syk* $^{-/-}$ FL-HSCs. Together, these studies demonstrate a potential role of Syk during early T cell development. Additional studies will be required to ascertain the potential roles of ZAP-70 and Syk during other stages of T cell differentiation.

ZAP-70 and Syk PTKs in $\gamma\delta$ T Cell Development

In addition to $\alpha\beta$ T cells, $\gamma\delta$ T cells also play critical roles in mucosal immunity. The Syk family of PTKs also play important roles in the development of various subsets of $\gamma\delta$ T cells. Studies in *zap-70* $^{-/-}$ and *syk* $^{-/-}$ mice demonstrate important roles for both ZAP-70 and Syk in the development of skin dendritic epithelial (DETC) and intestinal epithelial (IEL) $\gamma\delta$ T cells. Substantial reductions in these populations were observed in both strains of knockout mice. In addition, the remaining DETCs exhibited marked abnormalities in morphology in *zap-70* $^{-/-}$ mice. In contrast to the DETCs and IELs, lymph node and splenic $\gamma\delta$ T cells were found in greater abundance in *zap-70* $^{-/-}$ and *syk* $^{-/-}$ mice. Hence, while murine $\alpha\beta$ T cells demonstrate an absolute developmental requirement for ZAP-70, $\gamma\delta$ T cells demonstrate variable dependence upon ZAP-70 and Syk.

SYK AND ZAP-70 PTKs IN B CELL DEVELOPMENT

B cell development begins in the bone marrow through a program of developmental checkpoints regulated by signaling through the pre-BCR and subsequently the mature surface IgM (sIgM) receptor. In pro-B cells, heavy chain gene rearrangement begins through D_H to J_H genes followed by V_H to DJ_H genes. An in-frame rearranged heavy chain pairs with the $\lambda5$ and V-preB surrogate light chains to form the antigen-independent pre-BCR. Signaling through the pre-BCR expressed on pro-B cells induces cells to differentiate to pre-B cells. In pre-B cells, termination of heavy chain rearrangement (termed allelic exclusion), initiation of

gene rearrangement of K or λ light chains, and pairing of these resultant heavy and light chains gives rise to the sIgM receptor expressed on immature B cells. Immature B cells leave the marrow to the spleen and other lymphoid organs where a small minority of cells is selected to differentiate to mature B cells that express both IgM and IgD, a selection process that requires signaling through sIgM on immature B cells.

Studies in *syk* $^{-/-}$ mice demonstrate a requirement for Syk in pre-BCR function. *Syk* $^{-/-}$ mice demonstrate a significant, but partial, block at the pro- to pre-B cell transition. While a small number of immature B cells develop *in syk* $^{-/-}$ mice, adoptive transfer experiments utilizing radiation chimeras demonstrate an additional requirement for Syk in the developmental transition from immature to mature recirculating B cells. Despite migrating from the bone marrow to the spleen, *syk* $^{-/-}$ immature B cells are unable to mature into recirculating mature B cells and accumulate in the outer splenic T cell zones. Hence, Syk plays important roles in both pre-BCR and sIgM receptor signaling.

Similar to the overlapping roles of ZAP-70 and Syk in pre-TCR function, mice deficient in both *zap-70* and *syk* demonstrate an absolute block in pre-BCR function and, in turn, a failure of heavy chain allelic exclusion. Hence, similar to the overlapping roles of these two PTKs in early T cell development, both ZAP-70 and Syk play overlapping roles during early B cell development.

SYK PTK in FcR FUNCTIONS

While neither Syk nor ZAP-70 is required for monocyte or natural killer cell development, the functions of multiple receptors expressed on these cells are significantly compromised. While *syk* $^{-/-}$ macrophages form normal actin cups that oppose foreign particles, they are unable to phagocytose these particles. Additionally, *syk* $^{-/-}$ monocytes/macrophages exhibit defects in FcR-mediated antigen presentation and dendritic cell maturation.

In addition to FcRs, Syk is also required for signaling through the high affinity IgE receptor of mast cells. Mast cells derived from *syk* $^{-/-}$ bone marrow are unable to induce degranulation, synthesize leukotrienes, or screte cytokines when stimulated from the FcεRI receptor.

ZAP-70 AND SYK IN INTEGRIN-MEDIATED FUNCTIONS

A requirement for ZAP-70 has also been implicated in LFA-1-mediated functions. Inhibition of ZAP-70 through pharmacologic and genetic means demonstrate a role of ZAP-70 in LFA-1-dependent chemotaxis. An essential role of Syk has also been demonstrated for integrin-mediated functions in polymorphonuclear

(PMN) leukocytes. $Syk^{-/-}$ PMNs are unable to undergo degranulation, fail to generate a respiratory burst, and are unable to spread in response to $\beta1$, $\beta2$, or $\beta3$ signaling. Hence, in addition to the important roles of this family of PTKs in adaptive immunity, these PTKs also play important roles in innate immunity.

Syk PTK in Lymphatic Development and Platelet Function

Syk also plays requisite functions in collagen-mediated activation in platelets. $Syk^{-/-}$ platelets cannot induce increases in free cytoplasmic calcium in response to collagen or a collagen-related peptide. While $syk^{-/-}$ platelets are unable to function, bleeding times in $syk^{-/-}$ mice are normal. Hence, platelet dysfunction cannot account for the perinatal petechiae observed in these mice. $Syk^{-/-}$ mice develop peritoneal hemorrhage with chylous appearing ascites and the majority of mice die *in utero* or perinatally. This defect has been recently attributed to a critical function for Syk in a yet-to-be identified bone marrow derived endothelial cell required for separation of the lymphatic and vascular systems. Electron microscopy of the vasculature of $syk^{-/-}$ mice reveal not only decreased numbers of endothelial cells but also abnormal morphogenesis of these cells. Hence, Syk is required for a hematopoietic signaling pathway involved in the differentiation of the lymphatic from vascular systems.

Summary

Over the past decade, we have learned much about the regulation and functions of the Syk family of PTKs in mammalian immune cell function. While there is much to be still learned about the cell types and receptor systems that are regulated by these PTKs, there is emerging an interesting biology of these PTKs in a greater array of human diseases. Expression of Syk has been suggested to play a tumor-suppressive role in human breast cancer. The expression of ZAP-70 in human B cells has recently been identified as a marker for prognostication of survival in chronic lymphocytic leukemia. Clearly, much is still to be learned as to how aberrant expression of these PTKs may affect normal and abnormal immune and non-immune cell functions in health and disease.

See Also the Following Articles

Epidermal Growth Factor Receptor Family • Immunoglobulin (Fc) Receptors

Glossary

catalytic functions Enzymatic activity of a protein. In the case of protein tyrosine kinases, catalytic function measures the ability of the enzyme to phosphorylate itself or its substrates.

protein tyrosine phosphorylation Post-translational modification required of tyrosine residues that can modulate catalytic function as well as to mediate protein–protein interactions.

thymic ontogeny The process of T cell development from its immature stage in the thymus to differentiated T cells in circulation.

Further Reading

Chu, D. H., Morita, C. T., and Weiss, A. (1998). The Syk family of protein tyrosine kinases in T-cell activation and development. *Immunol. Rev.* **165**, 167.

Kurosaki, T. (2002). Regulation of B cell fates by BCR signaling components. *Curr. Opin. Immunol.* **14**, 341.

Singer, A. L., and Koretzky (2002). Control of T cell function by positive and negative regulators. *Science* **296**, 1639.

Turner, M., Schweighoffer, E., Colucci, F., and DiSanto, J. P. (2000). Tyrosine kinase SYK: essential functions for immunoreceptor signaling. *Immunol. Today* **21**, 148.

Van Leeuwen, J. E., and Samelson, L. E. (1999). T cell antigen-receptor signal transduction. *Curr. Opin. Immunol.* **11**, 242.

Biography

Andrew Chan is the Vice President of Research-Immunology at Genentech, Inc. His principal research interests are the signal transduction mechanisms in lymphocytes that regulate normal and aberrant immunity. He holds M.D. and Ph.D. from Washington University School of Medicine, St. Louis and received clinical training in Internal Medicine and Rheumatology.

T7 RNA Polymerase

Rui Sousa
University of Texas Health Science Center, San Antonio, Texas, USA

Upon infection of an *Escherichia coli* cell the T7 bacteriophage inactivates the host's transcriptional machinery and transcription of the phage genes is carried out by a phage encoded RNA polymerase. The T7 RNA polymerase is a single subunit protein with a molecular weight of 99 kDa and is therefore structurally simpler than the multisubunit cellular RNA polymerases which have molecular weights in excess of 500 kDa. This relative simplicity has facilitated study of this enzyme, so that T7 RNA polymerase is currently the best-understood RNA polymerase. T7 RNA polymerase exhibits structural similarity to DNA polymerases, reverse transcriptases, and RNA-directed RNA polymerases, and these structural similarities establish the existence of a vast polymerase superfamily that includes the majority of nucleic acid synthesizing enzymes. The high activity and stringent promoter specificity of T7 RNA polymerase have been used to develop systems for T7 RNA polymerase driven overexpression of heterologous genes *in vivo*, and to synthesize RNAs *in vitro* for a variety of purposes.

DNA-Directed RNA Polymerases

The enzymes which synthesize nucleic acids – the polymerases – are functionally defined on the basis of whether they use DNA or RNA as a template, and whether they synthesize RNA or DNA. The DNA directed RNA polymerases comprise two large families. One includes the large multi-subunit cellular RNA polymerases which synthesize the messenger and ribosomal RNAs of all cells. The other family includes the mitochondrial RNA polymerases and the RNA polymerases encoded by a variety of bacteriophage. These are simpler, typically single subunit, enzymes. The best-characterized representative of this family is the RNA polymerase of the T7 bacteriophage.

T7 RNA Polymerase

STRUCTURE

The three-dimensional structure of T7 RNA polymerase has been determined by X-ray crystallography ([Figure 1A](#)). This highly α-helical enzyme is comprised of an N-terminal domain (amino acids 1–312), and a C-terminal domain (amino acids 313–883). The C-terminal domain is further sub-divided into "thumb" (amino acids 330–410), "palm" (amino acids 411–448, 532–540, 788–838), and "fingers" (amino acids 541–737, 771–778) subdomains, which are so designated because together they form a structure similar in shape to a cupped right hand. Nucleic acids bind within the large cleft in this structure.

STRUCTURE–FUNCTION RELATIONSHIPS

The C-terminal domain contains the active site where RNA synthesis occurs. Phosphodiester bond formation is catalyzed by two Mg^{2+} ions complexed by two aspartates (D537, D812) of the palm subdomain ([Figure 2](#)). The thumb subdomain makes interactions with the RNA which stabilize the transcription complex, while the fingers subdomain binds the template strand and contains residues which bind the nucleotide triphosphate (NTP) and which discriminate ribo-NTPs from deoxyribo-NTPs. Residues 739–770 of the C-terminal domain form an extended "promoter recognition loop" which makes sequence specific interactions with the −7 to −11 base pairs (bp) of the promoter. The C-terminal domain also contains the binding site for the regulatory factor T7 lysozyme, and residues 839–883 form a "C-terminal loop" which functions in the allosteric mechanism by which T7 lysozyme regulates transcriptional activity. The N-terminal domain is involved in nascent RNA binding, in sequence specific binding to the promoter, and in separating (opening) the two strands of the promoter during initiation so as to expose one strand for templating RNA synthesis. Residues 93–101 of the N-terminal domain are rich in positively charged amino acids and make specific interactions with the AT-rich −13 to −17 bp of the T7 promoter, while residues 232–242 form an "intercalating hairpin" which inserts between the two DNA strands to open the promoter.

SIMILARITIES TO OTHER POLYMERASES

T7 RNA polymerase shares extensive sequence similarity with mitochondrial RNA polymerases and with
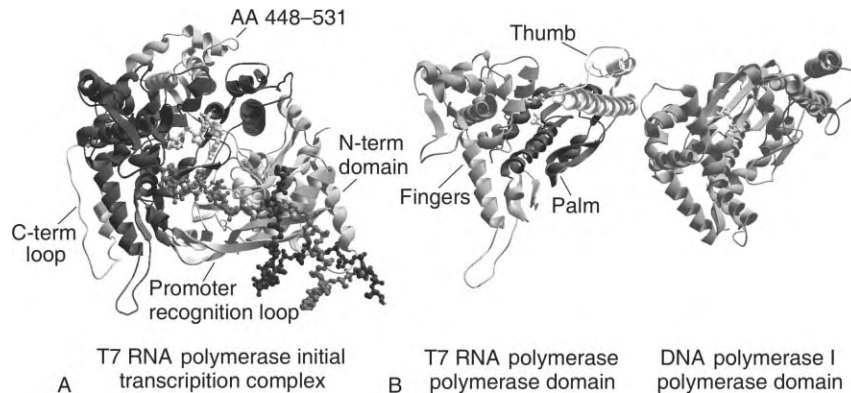
FIGURE 1 (A) Structure of a T7 RNA polymerase initial transcription complex. The template and nontemplate strands and the RNA are in medium, dark, and light gray, respectively. The core polymerase domain of the protein is in dark gray while the N-terminal domain, promoter recognition loop, C-terminal loop, and the subdomain formed by amino acids 439–531 are labeled and are colored light gray. (B) Comparison of the structures of the core polymerase domains of T7 RNA polymerase and DNA polymerase I. In the T7 RNA polymerase structure the thumb, palm, and fingers subdomains are in light, dark, and medium gray, respectively. Also shown are the side chains (colored light gray) of the pair of aspartic acids which bind the catalytic Mg^{2+} in the active site.

other bacteriophage-encoded RNA polymerases. T7 RNA polymerase is also structurally similar to the DNA-directed DNA polymerases of the pol I/ pol α family, to the RNA-directed DNA polymerases
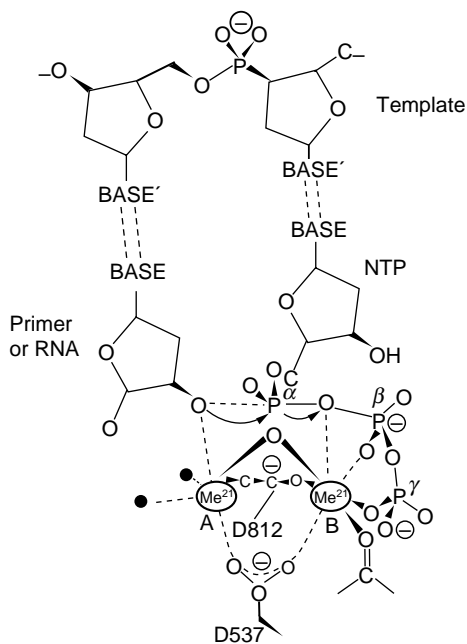
FIGURE 2 Bond formation in nucleic acid synthesis. The 3′-OH of the nucleotide at the end of the RNA or DNA primer attacks the α-phosphate of the NTP. A new bond is formed and the β- and γ-phosphates of the NTP are lost as pyrophosphate. The reaction is catalyzed by two metal ions which stabilize both the negative charge on the 3′-OH to enhance it nucleophilicity and the negatively charged pentacovalent α-phosphate intermediate which forms during the transition state. The two metal ions are bound by the side chains of two aspartic acids (D537 and D812 in T7 RNA polymerase) which are found in the active sites of all polymerases.

(reverse transcriptases), and to the RNA-directed RNA polymerases. However, unlike the extensive sequence similarities within the mitochondrial and phage RNA polymerase family, the identifiable sequence similarities between T7 RNA polymerase and the DNA polymerases, reverse transcriptases, and RNA-directed RNA polymerases are limited to a few well-conserved amino acids in a small number of sequence motifs. Conservation of these motifs correlates with the template and product specificity of the polymerase as shown in Figure 3. The structural similarity suggested by this limited sequence similarity is confirmed by comparison of the three-dimensional structures of these enzymes (Figure 1B). This comparison reveals further that functions specific to particular classes of polymerase are incorporated by accretion of structurally dissimilar domains or "modules" onto a structurally conserved core. For example, DNA polymerase I and T7 RNA polymerase both exhibit structurally similar thumb, palm, and fingers subdomains which together form a polymerase domain with the core function of processive template-directed nucleic acid synthesis. Attached to this core domain are structural elements which are responsible for functions displayed by T7 RNA polymerase but not by DNA polymerase I (and vice versa). The "accessory modules" of T7 RNA polymerase have no structurally similar counterparts in DNA polymerase I. They include the N-terminal domain, which is involved in nascent RNA binding, transcription termination, and promoter opening; the promoter recognition loop which, together with the N-terminal domain, is responsible for sequence specific binding of the T7 promoter; the C-terminal loop, which is involved in regulation by T7 lysozyme, and residues 449–531, which form a subdomain of as yet undefined function.

```
Motif designation          T/DxxGR      A            B              C

DNA-directed polymerases     |          |            |              |
                             |          |            |              |
        DNA polymerases    hT--GR     Dh--hEh      Khh----hYG     h-D
  (pol I-like, pol α-like)   |          |            |              |
                             |          |            |              |
       RNA polymerases    hDhRGRhY  Ph--D--C-GhQHh  R-h-K+-VMTh-YG  hHDSFGT
   (phage, mitochondrial)                |
                                         |
RNA-directed polymerases     |           |           |              |
                             |           |           |              |
       DNA polymerases     hDh---h--h  h-h-+hQG--SP  YHDDhhh   Gh-h●---K  h-hlGH
                             |           |           |              |
       RNA polymerases     Dh---hD     SG----●h    hh-GDD-hh   G--h----K   |

Motif Designation             A           B'          C              D    E
```

FIGURE 3 Patterns of sequence motif conservation in nucleic acid polymerases. h indicates a hydrophobic residue, + is a positively charged residue, − is any residue, and ● is a sequence gap. Invariant amino acids are in bold face. In T7 RNA polymerase the two invariant aspartic acids of motifs A and C correspond to D537 and D812, respectively, while the invariant arginine of motif T/DxxGR and the invariant lysine and tyrosine of motif B correspond to, respectively, R425, K631, and Y639.

T7 RNA Polymerase: Transcription Reaction

PROMOTER STRUCTURE, RECOGNITION, AND OPENING

The T7 promoter is 23 bp in length and has a tri-partite structure (Figure 4). The −17 to −6 sequence is important for specific binding of the polymerase via interactions between residues 746, 748, 756, and 758 and the −7 to −11 bp, and between residues 93–101 and the −13 to −17 bp (bases are numbered relative to the transcription start site at +1). The −17 to −6 bp remain base paired during transcription initiation. The −4 to −1 "TATA" element facilitates promoter opening, which begins at the −4 bp and extends downstream, driven by imposition of a sharp (∼50°) bend in the promoter when polymerase binds and by insertion of β-hairpin (formed by residues 231–242) between the DNA strands. The +1 to +6 initially transcribed sequence enhances the efficiency of the initial transcription reaction. Seven class III T7 promoters occur in the
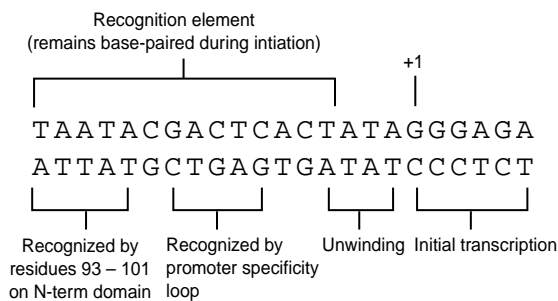
T7 genome and exhibit a perfect match to the sequence shown in Figure 4. There are also 16 class II T7 promoters in the T7 genome. The less active class II promoters typically exhibit a small number of base pair differences from the class III promoters.

INITIAL TRANSCRIPTION

Following promoter binding and opening, the polymerase initiates RNA synthesis. While the nascent RNA is small (<9 nucleotides), the transcription complex is unstable and 2–8 nucleotide RNAs are frequently released from the complex. After an RNA is released, the polymerase reinitiates, usually without releasing the promoter. For this reason this initial phase of transcription, which is also characteristic of multi-subunit cellular RNAPs, is also referred to as "abortive" transcription. Throughout this initial phase of transcription the polymerase retains the specific promoter interactions made with the −17 to −6 bp in the initial binding step. Transcription to +8 is achieved by threading the template strand through the active site and by compacting (scrunching) it within the template-binding cleft, as well as by conformational changes in the polymerase that accommodate a growing RNA:DNA hybrid.

PROMOTER RELEASE AND ELONGATION

When the RNA reaches 9 nucleotides in length a large conformational change is triggered in the polymerase causing it to release the promoter. In addition to breaking up the promoter-binding surface created jointly by the promoter recognition loop and N-terminal domain, this conformational change reorganizes the N-terminal domain, leading to formation of a tunnel

FIGURE 4 Structure of the T7 RNA polymerase class III promoter with the functions of different parts of the promoter indicated.

through which the emerging RNA passes. The disposition of the DNA immediately upstream and downstream of the RNA:DNA hybrid is also changed. The result of all these changes is a stable "elongation complex" which can move along the template, synthesizing thousands of nucleotides of RNA without releasing either transcript or DNA.

PAUSING AND TERMINATION

Certain sequences in the DNA can, however, act as pause or terminator sites and interrupt the progress of the elongation complex. Class I terminators cause the elongation complex to pause, and then to release the RNA. They contain a sequence which can form a G:C rich hairpin when transcribed into RNA (Figure 5). Immediately downstream of the hairpin is a U-rich sequence. It is believed that formation of a hairpin in the RNA may disrupt interactions normally made between the polymerase and single-stranded RNA 8–14 nucleotide away from the RNA 3'-end. It may also disrupt part of the RNA:DNA hybrid, which is usually ~7 bp in length in the elongation complex. Disruption of these interactions will weaken the association of the RNA with the elongation complex. The U-rich nature of the remaining RNA:DNA base pairs further weakens the RNA's association with the elongation complex, leading to release of the RNA and transcription termination. A class I terminator appears between genes 10 and 11 in the T7 genome, where it is important for attenuating transcription of downstream genes. Unlike class I terminators, class II terminators contain an invariant "ATCTGTT" sequence, which exhibits no obvious potential for forming secondary structure in the RNA. Termination at these sequences may involve a sequence-specific interaction with the RNA polymerase which alters the structure of the elongation complex. Encounter of the elongation complex with the class II terminator leads to a collapse in the size of the transcription bubble, from ~10 bases in the normal elongation complex, to ~5 bases in the complex which is paused at the class II site. A class II terminator occurs at the point at which T7 genomes are joined as concatemers before they are processed for packaging as mature phage particles, implying that pausing or termination of the polymerase at the concatemer junction is important for processing of the phage DNA.

REGULATION

During the T7 life cycle the transcriptional activity of T7 RNA polymerase is regulated by T7 lysozyme, which binds to the polymerase and inhibits transcription initiation. Inhibition is allosteric: lysozyme binding changes the conformation of the C-terminal loop (amino acids 839–883) of the polymerase. This conformational change weakens the affinity of the polymerase for NTPs and short RNAs. This, in turn, reduces the rate
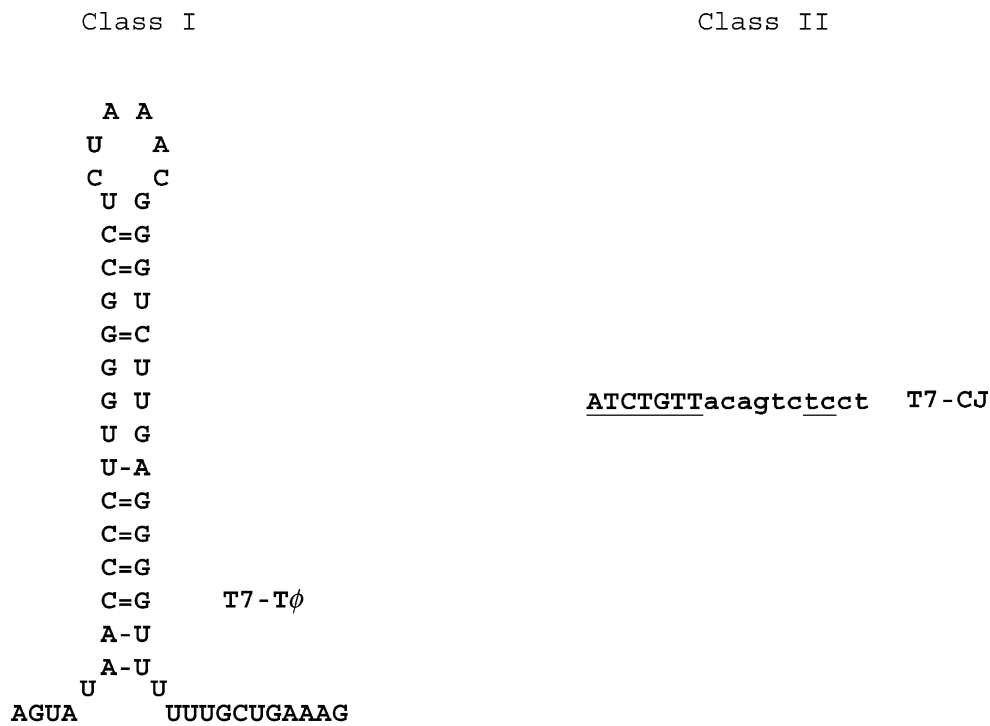
FIGURE 5   (Left) RNA structure of the T7 RNA polymerase Tφ class I terminator. Termination occurs at the underlined nucleotide. (Right) Nontemplate strand DNA sequence of the class II T7 RNA polymerase concatemer junction terminator.

of transcription initiation at limiting NTP concentrations, and decreases the efficiency of progression through initial transcription by increasing the frequency at which short RNAs are released from the complex during abortive transcription. These inhibitory effects are greater for class II promoters, which, relative to class III promoters, display intrinsically higher rates of RNA release during initial transcription and require higher NTP concentrations to achieve high rates of initiation. Class III promoters drive transcription of genes which encode proteins, such as phage coat proteins, which are required late in infection, while class II promoters drive transcription of genes which encode proteins required during the early and middle stages of phage infection, such as proteins involved in replication of the phage DNA. Thus, in a T7 infected *E. coli*, the accumulation of T7 lysozyme late in the phage life cycle leads to a disproportionate decrease in transcription from class II promoters and to an increase in the production of the late phage proteins required for assembly of the mature phage particles. Since T7 lysozyme is itself encoded by a gene transcribed from a class II promoter, an auto-inhibitory feedback loop is created which ensures that repression by T7 lysozyme is kept within an appropriate range.

Priming DNA Replication

In addition to its primary function of transcribing the T7 phage genes, T7 RNA polymerase also primes rightward replication of T7 DNA. Priming occurs within an A−T rich region immediately downstream of two T7 promoters (dubbed 1.1a and 1.1b) which are located at the T7 origin of replication. The mechanism by which the RNA primer initiated at these promoters is transferred from T7 RNA polymerase to T7 DNA polymerase is not understood.

T7 RNA Polymerase: Applications

The stringent promoter specificity and robust transcriptional activity of T7 RNA polymerase has been taken advantage of to overexpress proteins *in vivo* and to synthesize RNAs *in vitro*. In the most widely used embodiment of the former application the gene encoding T7 RNA polymerase is placed under the control of an inducible promoter and is then stably integrated into the genome of an *E. coli* cell. A plasmid carrying the gene of interest under the control of a T7 promoter is then introduced into *E. coli*. When the gene encoding the T7 RNA polymerase is induced, the expressed T7 RNA polymerase transcribes the gene of interest at a very high level, resulting in a high degree of overproduction of the desired gene product. Similar approaches are used to overexpress proteins in eukaryotic cells. Synthesis of specific RNAs *in vitro* is done by using purified T7 RNA polymerase and templates in which a sequence of interest is placed downstream of a T7 promoter. The only other required reaction components are a buffering agent, Mg^{2+}, and NTPs. Such *in vitro* synthesized RNAs are used for a wide variety of research purposes, and are also being evaluated as diagnostic and therapeutic agents.

See Also the Following Articles

DNA Polymerase I, Bacterial • RNA Polymerase Reaction in Bacteria • RNA Polymerase Structure, Bacterial

Glossary

downstream The direction in which an RNA polymerase moves along the DNA during transcription.
primer A DNA or RNA molecule, typically short, that is extended by a DNA polymerase during DNA replication.
promoter A DNA from which an RNA polymerase initiates transcription.
template strand When a nucleic acid directs the synthesis of DNA or RNA, the template strand selects − by Watson−Crick base pairing − the nucleotides incorporated into the newly synthesized molecule.
transcription The synthesis of RNA using a DNA template.

Further Reading

Cheetham, G. M., and Steitz, T. A. (2000). Insights into transcription: Structure and function of single-subunit DNA-dependent RNA polymerases. *Curr. Opin. Struct. Biol.* **10**, 117–123.
McAllister, W. T. (1993). Structure and function of the bacteriophage T7 RNA polymerase (or, the virtues of simplicity). *Cell Mol. Biol. Res.* **39**, 385.
Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**, 8783–8798.
Mooney, R. A., Artsimovitch, I., and Landick, R. (1998). Information processing by RNA polymerase: Recognition of regulatory signals during RNA chain elongation. *J. Bacteriol.* **180**, 3265–3275.
Sousa, R. (1996). Structural and mechanistic relationships between nucleic acid polymerases. *Trends Biochem. Sci.* **21**, 186–190.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.

Biography

Rui Sousa is a Professor in the Department of Biochemistry of the University of Texas Health Science Center at San Antonio. He holds a B.A. from Harvard College and a Ph.D. from the University of Pittsburgh where he also received his postdoctoral training. His principal research interests are in transcription mechanisms and protein structural biology.

Tachykinin/Substance P Receptors

Mark D. Richardson and Madan M. Kwatra
Duke University Medical Center, Durham, North Carolina, USA

Tachykinins are small peptides, found in both vertebrates and invertebrates, which regulate many physiological processes. They are distributed mainly in the central nervous system (CNS), but are also important regulators of contractility in vascular smooth muscle and many areas of the gastrointestinal tract. Tachykinins (meaning "fast-acting") are characterized by an amidated C-terminus containing the amino acids F-X-G-L-M-NH$_2$, where X is a hydrophobic amino acid residue. Tachykinins act through receptors that are members of the G protein-coupled receptor (GPCR) superfamily. The best-known mammalian tachykinin is substance P (SP), a peptide of eleven amino acids. The preferred receptor for SP, substance P receptor (SPR), occurs as a full-length receptor and in a truncated form, which lacks the carboxyl-tail. The carboxyl-tail of SPR plays an important role in receptor desensitization, therefore the truncated form of SPR differs from full-length SPR in its interactions with proteins involved in receptor desensitization. Finally, SPR has an important role in pain and in several human disorders including depression, emesis, and glioblastoma.

Tachykinins

Tachykinins are peptides of 10–11 amino acids having the motif F-X-G-L-M-NH$_2$ at the C terminus; the $-$NH$_2$ group indicates that the C-terminal amino acid is amidated, a posttranslational modification necessary for biological activity.

There are currently four known mammalian tachykinins: SP, neurokinin A (NKA), neurokinin B (NKB), and hemokinin 1 (HK-1). As illustrated in Figure 1, these tachykinins are generated from three genes: preprotachykinin-A (PPT-A), preprotachykinin-B (PPT-B), and preprotachykinin-C (PPT-C). Alternative splicing of the PPT-A mRNA transcripts yields four products: α-PPT-A, β-PPT-A, γ-PPT-A, and δ-PPT-A. α-PPT-A and δ-PPT-A encode only SP, while β-PPT-A and γ-PPT-A encode both substance P (SP) and NKA. PPT-B yields only one product, NKB, and PPT-C encodes HK-1. The products of PPT-A transcripts (SP and NKA) are found in the central nervous system, CNS, and across a range of peripheral tissues, while a product of PPT-B (NKB) is found only in the CNS. PPT-C is widely distributed in many peripheral organ systems, but is not detected in the CNS.

The actions of mammalian tachykinins are mediated by three receptors: substance P receptor, SPR (also called NK1 receptor), neurokinin-2 (NK2) receptor, and neurokinin-3 (NK3) receptor. These receptors have the following binding preferences: SPR binds in the order SP > NKA > NKB; NK2 binds in the order NKA > NKB > SP; and NK3 binds in the order NKB > NKA > SP. Although SP, NKA, and NKB bind the tachykinin receptors with different affinities, they are all full agonists at all three receptors.

Mammalian Tachykinin Function

SP

SP is present in the outer laminae of the dorsal horn of spinal cord and in many areas of the brain, which is consistent with SP acting as a neurotransmitter. The hypothesis that SP has a role in pain transmission stems from its presence in primary afferent nerve fibers of the spinal cord such as C-fibers of laminae I and II of the dorsal horn. Efforts aimed at developing SPR antagonists to treat pain have met with limited success, leaving some doubt about the involvement of SP in pain transmission. However, recent experiments performed in genetically engineered mice, which lack expression of PPT-A have given new support for the role of SP in pain. While mice lacking PPT-A display normal responses to a variety of painful stimuli, their responses become blunted as the stimulus increases in intensity. However, above a certain pain threshold, mice lacking PPT-A seem to display responses that are similar to control mice. These results suggest that SP is indeed a pain neurotransmitter, and it functions within a discrete "window" of pain intensities. Other recent studies in mice lacking SPR also support the conclusion that SP plays a role in pain transmission.

SP also influences blood pressure. Intravenous injection of tachykinins in dogs and rabbits produces a strong hypotensive effect on mean arterial blood pressure, with SP showing much higher potency than NKA, NKB, or nonmammalian tachykinins. In human volunteers, intravenous SP causes a drop in diastolic, but not

Gene	mRNA	Product	Amino acid sequence
PPT A	α-PPT A	Substance P	RPKPQQ FFGLM-NH$_2$
	β-PPT A	Substance P	
		Neurokinin A	HKTDSF FVGLM-NH$_2$
	γ-PPT A	Substance P	
		Neurokinin A	
	δ-PPT A	Substance P	
PPT B		Neurokinin B	DMHDFF FVGLM-NH$_2$
PPT C		Hemokinin 1	TGKASQ FFGLM-NH$_2$

FIGURE 1 Synthesis of mammalian tachykinins. The mammalian tachykinins are generated from three genes (PPT-A, -B, and -C). Amino acid sequences of the human tachykinins are shown; shaded boxes highlight the F-X-G-L-M-NH2 motif.

systolic blood pressure, whereas NKA has no effect on either. Interestingly, the hypotensive effect is not diminished by autonomic blocking agents such as atropine, atenolol, or prazosin. This observation suggests that SP acts directly on vascular smooth muscle.

SP/NKA

The PPT-A mRNA transcripts which produce NKA (β- and γ-PPT-A) also produce SP (Figure 1); thus, SP and NKA are cosynthesized and coreleased from neurons. This suggests that NKA does not act alone, however, NKA can act as the primary signaling peptide in situations where the ratio of NKA to SP favors NKA, and where NK2 (the preferred receptor for NKA) is the predominant receptor in the target tissue. NKA is recognized as a very potent bronchoconstrictor, more potent than SP or NKB. Blockade of nerve-evoked bronchoconstriction by an antiserum having high affinity for NKA indicates that NKA has a more important role than SP in bronchoconstriction. Furthermore, experiments using NK2-selective antagonists indicate a predominance of NK2 over SPR in the mediation of noncholinergic bronchoconstriction responses. NKA is also more potent than SP in decreasing tracheal vascular resistance. Recent data indicate that NKA, like SP, binds to SPR with high affinity, thus, the effects of NKA may potentially be mediated by either SPR or NK2 receptor. In the CNS, which does not contain NK2 receptors, NKA produces biological effects similar to SP by interacting with SPR.

NKB

While SP and NKA frequently have overlapping tissue distribution, the distribution of NKB is distinct from that of SP and NKA. For example, in the spinal cord NKB is present in laminae III, while SP and NKA are abundant in laminae I and II but are not present in laminae III. NKB is most abundant in the hypothalamus and in the intestines. The physiological effects of NKB include contraction of rat portal vein and inhibition of gastric acid secretion. Interestingly, NKB potently decreases alcohol intake by a strain of alcohol-preferring rats. Also of potential clinical importance, NKB dilates the vasculature in human placenta and is thought to have an important role in the development of pre-eclampsia.

HK-1

Discovered in 2000, the tachykinin HK-1 is unique in that it is not of neuronal origin. HK-1 was first cloned from mouse B-lymphocytes, subsequent cloning of human HK-1 showed that it differs from mouse HK-1 in five out of eleven amino acid residues. Expression of human HK-1 has since been demonstrated in heart, skeletal muscle, thyroid, and skin. Weaker expression is detected in liver, lung, stomach, testes, placenta, and prostate. Receptor binding assays in cultured cells and tissues show that HK-1 has a strong binding preference for SPR over NK2 or NK3 receptors, and it is considered a full agonist of SPR.

HK-1 is similar to SP in many physiological effects, often with a potency that is very close to that of SP. Intravenous injection of HK-1 or SP into guinea pigs produces a dose-dependent hypotensive effect, with HK-1 and SP decreasing diastolic blood pressure at similar potencies of 0.2 nmol kg^{-1} and 0.1 nmol kg^{-1}, respectively. HK-1 and SP also induce salivary secretion in rats at essentially equal potencies. Studies are underway in several laboratories to develop a more complete understanding of the function of HK-1.

SPR, or NK1 Receptor

SPR and the other tachykinin receptors belong to the GPCR superfamily of receptors, which are characterized by the presence of seven hydrophobic transmembrane domains, designated TM1 through 7 (Figure 2). Cloning of human SPR from glioblastoma cells yields two forms of the receptor. One clone, with a stop codon after amino acid 407, encodes a full-length receptor and the other, with a stop codon after amino acid 311 (Δ311, Figure 2), encodes a truncated form of the receptor; the truncated form has also been cloned from guinea pig nervous system.

Full-Length SPR

SPR has three extracellular (EC) and three intracellular (IC) hydrophilic loops, a fourth IC loop formed by the palmitoylation of a cysteine residue, and in the case of the full-length receptor, a carboxyl-tail. The deduced
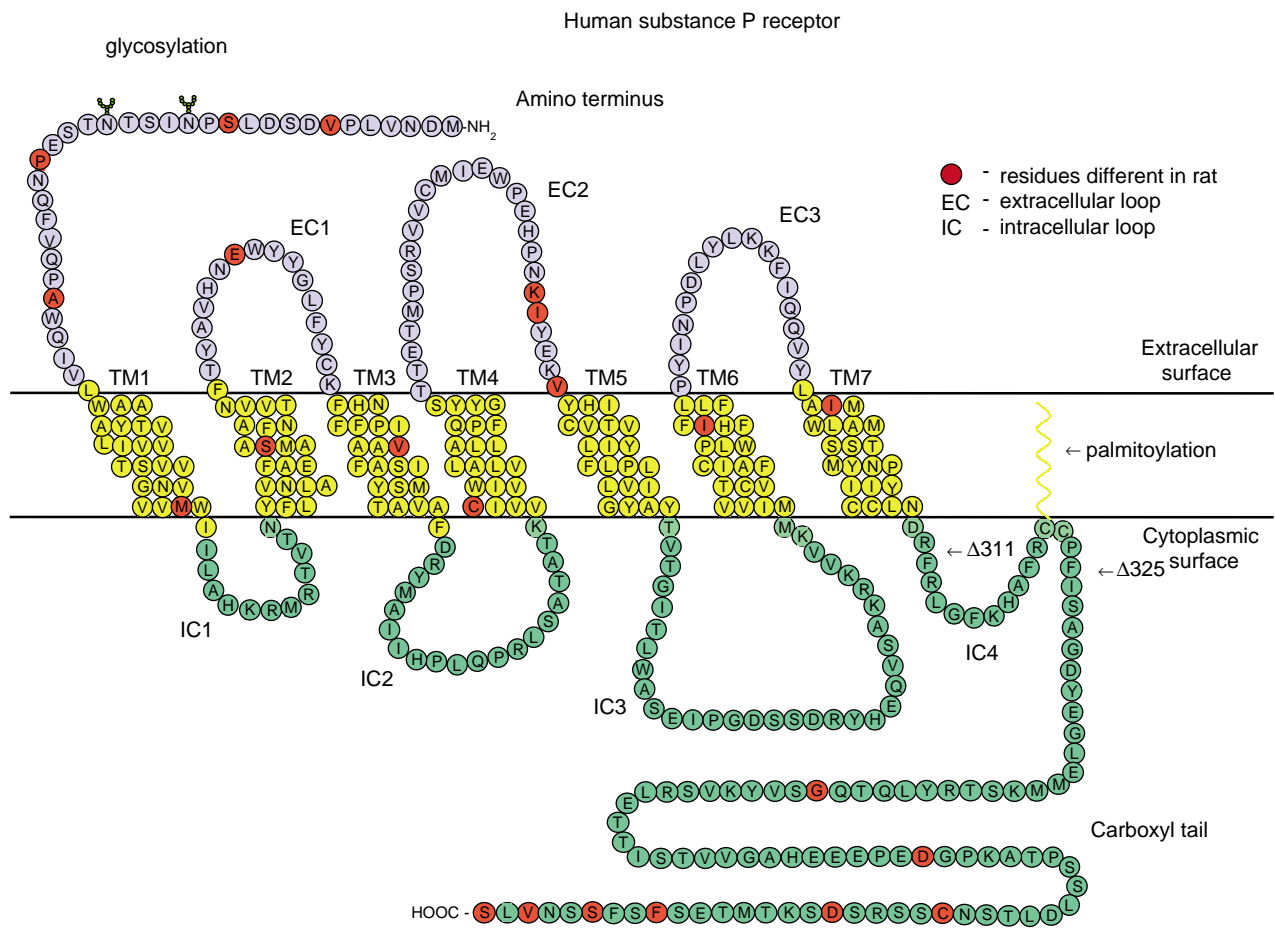
FIGURE 2 Two-dimensional representation of the human substance P receptor. The amino acid sequence of hSPR is shown. Seven putative transmembrane regions are labeled TM1–7, three extracellular loops are labeled EC1 through 3, and four intracellular loops are labeled IC1 through 4. Shaded circles show where amino acid residues in rat SPR differ from human SPR. Arrows indicate the termination site of the short form of SPR at Arg311, and truncation site of the mutant rat SPR at Phe325.

amino acid sequences of full-length rat and human SPR reveal that these receptors are 407 amino acids long, and share 95% amino acid identity; there are only 22 amino acids that differ between rat and human SPR (see shaded circles in Figure 2).

While rat and human SPR are very similar at the amino acid level, they interact differently with SP and several nonpeptide antagonists of SPR. For example, SP binds to rat SPR with a K_d of 3 nM whereas it binds to human SPR with a higher affinity characterized by a K_d of 0.7 nM. The difference between rat and human SPR pharmacology becomes even more striking for some of the nonpeptide antagonists. CP 96,345, the first nonpeptide antagonist of SPR to be synthesized, has 100-fold higher affinity for human SPR than for rat SPR, and the nonpeptide antagonist RP67580 has higher affinity for rat SPR than for human SPR.

Since rat and human SPR differ in only a few amino acid residues in the TM domains (a region believed to be involved in ligand binding), several groups have identified the amino acids responsible for differences in their pharmacology. These groups found that residue 290 in TM7 and the residues located on the second EC loop contribute to the selectivity of CP 96,345 for human SPR. Histidine 197 in TM5 is involved in binding of CP 96,345 whereas histidine 265 in TM6 is involved in RP67580 binding. These mutagenesis studies not only identify the residues involved in conferring species-dependent pharmacology to SPR, but also show that ligand binding to the receptor involves the TM domain as well as amino terminus and EC loops.

SPR activation in most cells stimulates phospholipase C, which hydrolyzes membrane phosphoinositides into two second messengers: inositol triphosphate (IP_3) and diacylglycerol. These molecules stimulate intracellular calcium release and protein kinase C (PKC) activation, respectively. More recent studies using U373 human glioblastoma cells, in which SPR occurs naturally, implicate several key molecules in SPR signaling including mitogen-activated protein kinases ERK1/2, mitogen-activated protein kinase p38, transcription factor NF-κB, and epidermal growth factor receptor. Of these molecules, ERK1/2 mediates SPR-stimulated proliferation of U373 MG cells.

TRUNCATED SPR

Functional studies in which the truncated and full-length forms of SPR were expressed in *Xenopus* oocytes revealed that the truncated receptor is 100-fold less active than full-length SPR. Ligand binding analyses of the full-length and truncated forms of SPR expressed in COS cells reveal that the truncated form binds SP with much lower affinity. Tissue distribution of the truncated form of SPR differs from that of full-length SPR, with the truncated form showing lower expression in most regions of the brain. However, in peripheral tissues the truncated form is more prevalent than full-length SPR, with bone and spleen expressing solely the truncated form. Potentially, these two receptors could mediate SP signaling differently in different tissues due to the presence or absence of the carboxyl-tail; thus it is important to understand the function of the carboxyl-tail in SPR biology.

ROLE OF THE CARBOXYL-TAIL IN SPR FUNCTION

One function of SPR in which the carboxyl-tail plays an important role is receptor desensitization. Like many other GPCRs, SPR undergoes agonist-induced desensitization, a phenomenon in which the responsiveness of the receptor diminishes in spite of the continued presence of the stimulus. Studies performed by Dr. Lefkowitz at Duke University Medical Center and by several other investigators have shown that agonist-specific or homologous desensitization of GPCRs proceeds through two cytosolic proteins, G protein-coupled receptor kinases (GRKs) and β-arrestins. GRKs phosphorylate agonist-occupied receptor at serine and threonine residues on the carboxyl-tail and IC loops, and this phosphorylation primes the receptor to bind β-arrestins. The binding of β-arrestin to the intracellular face of the receptor disrupts interactions between the receptor and G protein, resulting in a loss of signaling (Figure 3). β-arrestin binding also directs receptor internalization through clathrin-coated pits.

Since the carboxyl-tail plays a key role in GPCR desensitization, its absence would be expected to impair the ability of truncated SPR to desensitize. However, studies examining the role of the C terminus in rat SPR desensitization have yielded varying results. It has been reported that a C-terminally truncated mutant of rat SPR does not desensitize fully, while other studies have reported no loss of desensitization in the same truncation mutant. We have recently examined the effect of C-terminal truncation on the desensitization of human SPR. Full-length and truncated (Δ325) human SPR desensitize similarly, using two independent readouts of
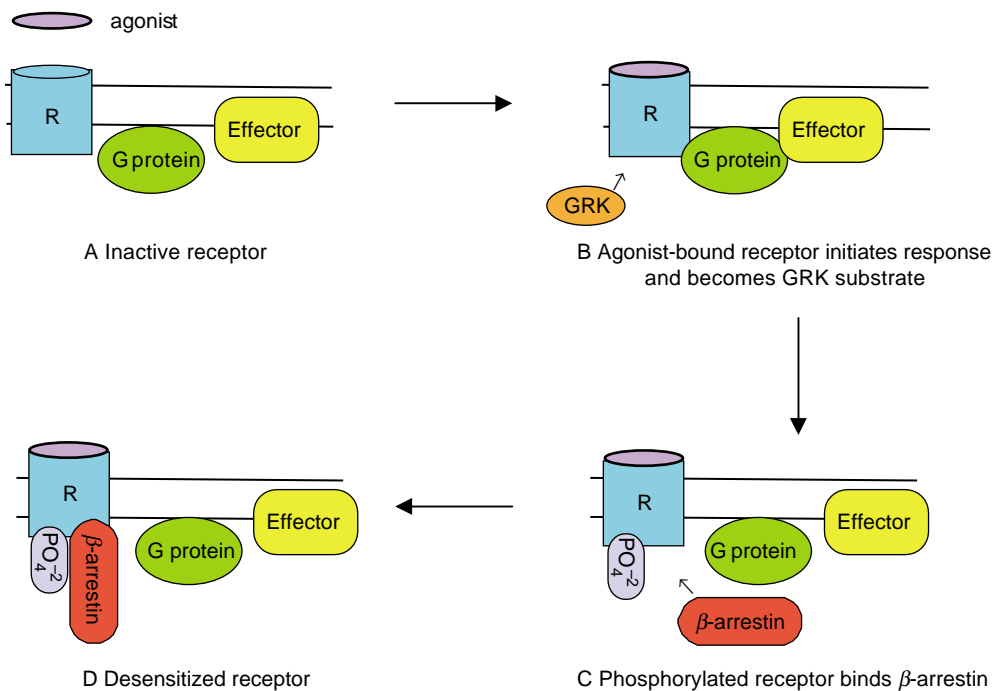
FIGURE 3   Model of homologous GPCR desensitization. (A) In the resting state, receptor, G protein and effector are inactive. (B) Occupancy of the receptor by agonist enables it to catalyze G protein activation, leading to activation of the effector by G protein. Receptor occupancy also makes it a target for phosphorylation by GRK. (C) Phosphorylation diminishes the ability of the receptor to activate G protein; the effector then reverts to the inactive state. Phosphorylation also facilitates the association of β-arrestin with the receptor. (D) Association of β-arrestin with phosphorylated receptor prevents any further G protein activations even though agonist is present.

receptor signaling, intracellular IP_3 accumulation and subcellular redistribution of fluorescently tagged PKC-βII in live cells. However, truncated and full-length SPR differ, in GRK-catalyzed phosphorylation and in their ability to form endocytic vesicles with two distinct β-arrestins, β-arrestin 1 and 2. Truncated SPR, unlike full-length SPR, does not undergo phosphorylation following exposure to SP. Also, full-length SPR forms endocytic vesicles equally well with either β-arrestin 1 or β-arrestin 2, while truncated SPR does not form any endocytic vesicles with β-arrestin 1 and forms fewer vesicles with β-arrestin 2. These data indicate that full-length and truncated SPR are likely to behave differently during endocytosis and receptor recycling.

The Role of SPR in Human Disease

While SP and SPR have been studied for several decades, renewed interest in SPR biology was generated following the report in 1998 that SPR antagonists have anti-depressant activity. This discovery generated so much enthusiasm that a commentary entitled "Reward for Persistence in Substance P Research" by Claes Wahlestedt appeared in *Science* along with a timeline highlighting important events throughout the history of substance P research (Figure 4). While an SPR antagonist has not yet been introduced into clinical medicine as an antidepressant, there is no doubt that SPR plays a major role in the pathophysiology of several diseases including depression, emesis, and glioblastomas.

SPR IN DEPRESSION

The distribution of SP and SPR in the brain is consistent with a role for these molecules in mood and depression. For instance, SP is the predominant tachykinin in human brain, and the expression of SP is particularly concentrated in areas of the brain that function in either affective behavior or stress responses, such as the hypothalamus, amygdala, habenula, periaqueductal gray, and dorsal raphe nucleus. Likewise, SPR is the most highly expressed tachykinin receptor in brain and is also concentrated in areas that are important for affective behavior.

The most convincing evidence of a role for SP in depression is found in a randomized, double-blind, placebo-controlled study of patients with moderate to severe major depression, which showed that the SPR antagonist, MK-869, has antidepressant activity equivalent to paroxetine, a current standard therapy. Importantly, MK-869 shows a significantly lower incidence of the side effects, which frequently cause patients to discontinue drug therapy. More recently, a second SPR antagonist, L759274, has shown similar results in clinical trials, further supporting a role for SPR in depression.

SPR IN EMESIS

An important role for SPR in emesis is indicated by clinical trials showing that the nonpeptide SPR antagonist, MK-869 (also called aprepitant), is effective against chemotherapy-induced nausea and vomiting (CINV). However, the molecular mechanism through which SPR influences emesis is not yet known. Clinical trial investigators noted that during the acute phase of CINV (the first 24 h), aprepitant was only effective in 37% of patients while the standard therapy, odansetron, was effective in 57% of patients. In contrast, during the delayed phase of CINV (the succeeding 6 days) aprepitant was effective in 72% of patients while odansetron was effective in only 7% of patients. This observation suggests that acute phase nausea and delayed phase nausea involve different molecular mechanisms and that SP is active in the delayed phase of CINV.

Another clinical trial tested a triple combination of aprepitant, granisetron, and dexamethasone for suppression of CINV. In this trial, 87% of patients receiving the triple combination reported that nausea did not affect their ability to perform daily functions, in comparison to 67% of patients receiving only granisetron and dexamethasone. Aprepitant has recently received Food and Drug Administration, FDA, approval for use in combination with other antinausea drugs to prevent CINV, which is a particular problem in high-dose cisplatin therapy. This result points to SPR as a promising target protein for future efforts to understand and control nausea.
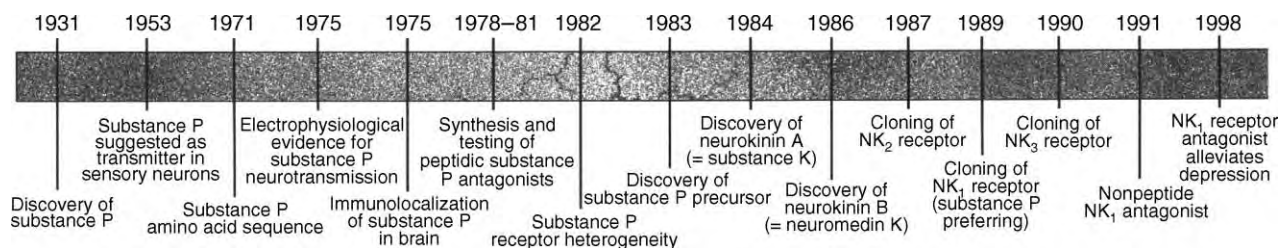
FIGURE 4 A timeline of landmark events in substance P research. (Reprinted from Wahlestedt, C. (1998). Substance P and related neuropeptides. *Science* **281**, 1624–1625, with permission of AAAS.)

SPR in Glioblastomas

Glioblastomas, classified as grade IV astrocytomas, are among the most aggressive and frequently occurring primary brain tumors in adults. Patients with glioblastomas have an extremely poor prognosis: with a 5-year survival rate of 1%, the currently available treatments of surgery, radiation therapy, and chemotherapy are inadequate. Current research focuses on understanding glioblastoma biology and identifying molecular targets for blocking tumor growth. SP has been observed to have a growth-promoting affect in a variety of cell types, and its role in glioblastoma growth is currently being studied. A recent study noted SPR expression in 9 out of 12 astrocytomas, and 10 out of 10 glioblastomas. Further, the density of SPR correlates with the degree of malignancy, with glioblastomas expressing more receptors than astrocytomas. At the molecular level, SPR stimulation in U373 MG human glioblastoma cells increases mitogenesis, cell proliferation, and release of interleukin-6 (IL-6). SPR-dependent release of IL-6 is noteworthy because IL-6 has been implicated in the progression of gliomas. Thus SPR likely plays an important role in the biology of glioblastomas. Consistent with this notion, SPR antagonists are reported to inhibit the growth of glioblastomas in nude mice.

See Also the Following Articles

G Protein-Coupled Receptor Kinases and Arrestins • Neurotransmitter Transporters • Phospholipase C

Glossary

carboxyl-tail The hydrophilic portion of a GPCR which follows the final transmembrane-spanning domain (TM7), extending from the cytosolic face of the plasma membrane and terminating in the cytosolic compartment of the cell.

G protein-coupled receptor (GPCR) Integral plasma membrane proteins having seven hydrophobic, transmembrane-spanning domains (designated TM1 through 7). GPCRs pass signals across the plasma membrane by transducing stimuli from extracellular signaling molecules to intracellular G proteins. G proteins, in turn, activate intracellular signaling cascades.

preprotachykinins The genes containing all the DNA sequence information needed to synthesize tachykinins. The actual production of tachykinins requires alternative splicing of mRNA transcripts derived from preprotachykinin, and post-translational modification of peptides generated from the mRNA transcripts.

phospholipase An enzyme that can hydrolyze the phosphodiester bond in phospholipids and produce soluble inositol lipids and membrane-bound diacylglycerol; type C phospholipases specifically use phosphoinositides as substrates.

tachykinin A peptide of 10-11 amino acids, having the amino acid sequence F-X-G-L-M-NH2 at the C-terminal end, where X is a hydrophobic amino acid residue. Tachykinins are released from neurons and act as neurotransmitters.

Further Reading

De Felipe, C., Herrero, J. F., O'Brien, J. A., Palmer, J. A., Doyle, C. A., Smith, A. J. H., Laird, J. M. A., Belmonte, C., Cervero, F., and Hunt, S. P. (1998). Altered nociception, analgesia and aggression in mice lacking the receptor for substance P. *Nature* **392**, 394–397.

Hökfelt, T., Pernow, B., and Wahren, J. (2001). Substance P: A pioneer amongst neuropeptides. *J. Intern. Med.* **249**, 27–40.

Holmes, A., Heilig, M., Rupniak, N. M. J., Steckler, T., and Griebel, G. (2003). Neuropeptide systems as novel therapeutic targets for depression and anxiety disorders. *Trends Pharmacol. Sci.* **24**, 580–588.

Kohout, T. A., and Lefkowitz, R. J. (2003). Regulation of G protein-coupled receptor kinases and arrestins during receptor desensitization. *Mol. Pharmacol.* **63**, 9–18.

Maggi, C. A. (2000). Principles of tachykininergic co-transmission in the peripheral and enteric nervous system. *Regulatory Peptides* **93**, 53–64.

Mantyh, P. W. (2002). Neurobiology of substance P and the NK1 receptor. *J. Clin. Psychiat.* **63**, 6–10.

Palma, C., and Maggi, C. A. (2000). The role of tachykinins via NK$_1$ receptors in progression of human gliomas. *Life Sci.* **67**, 985–1001.

Rittenberg, C. N. (2002). A new class of antiemetic agents on the horizon. *Clin. J. Oncol. Nurs.* **6**, 103–104.

Severini, C., Improta, G., Falconieri-Erspamer, G., Salvadori, S., and Erspamer, V. (2002). The tachykinin peptide family. *Pharmacol. Rev.* **54**, 286–322.

Wahlestedt, C. (1998). Substance P and related neuropeptides. *Science* **281**, 1624–1625.

Zimmer, A., Zimmer, A. M., Baffi, J., Usdin, T., Reynolds, K., Konig, M., Palkovits, M., and Mezey, E. (1998). Hypoalgesia in mice with a targeted deletion of the tachykinin 1 gene. *Proc. Natl. Acad. Sci. USA* **95**, 2630–2635.

Biography

Mark D. Richardson is a Research Associate in the Department of Anesthesiology, Duke University Medical Center. He is a protein biochemist with expertise in G proteins and his current focus is on signaling through substance P receptors. He holds a Ph.D. from the University of Texas (Houston).

Madan M. Kwatra is an Associate Professor in the Department of Anesthesiology and Assistant Professor in the Department of Pharmacology and Cancer Biology. His principal research interest is signaling through G protein-coupled receptors. His laboratory has been studying substance P receptor for the last 10 years and ongoing studies are directed toward understanding the role of substance P receptor and other G protein-coupled receptors in the biology of glioblastomas. He holds a Ph.D. from the University of Montreal, Canada.

Taste Receptors

John D. Boughter Jr.
University of Tennessee Health Science Center, Memphis, Tennessee, USA

Steven D. Munger
University of Maryland School of Medicine, Baltimore, Maryland, USA

Taste receptors are proteins that recognize taste stimuli of various types, thereby functioning as the initial component in the process of sensing and discriminating ingested material. Taste stimuli can be categorized as belonging to one of at least five classes, comprising qualities perceived by humans as sweet, salty, sour, bitter, and umami (the savory taste of L-amino acids such as glutamate). Recently, great progress has been made in the identification and functional characterization of mammalian taste receptors that respond to sweet, bitter, and umami (e.g., monosodium glutamate) stimuli. These receptors are expressed on the apical membranes of taste-receptor cells (TRCs) that extend into the oral cavity. The receptor–stimulus binding event initiates a transduction cascade in TRCs, leading to cell depolarization and neurotransmitter release onto afferent nerve fibers, and ultimately propagation of sensory information to taste processing areas in the central nervous system.

Taste-Receptor Cells

The initial events in taste processing occur in taste buds, structures found in the epithelia of the tongue, palate, larynx, and epiglottis of mammals, which contain ∼50–100 cells of neuroepithelial origin (Figure 1). Taste buds contain morphologically distinct cell types including elongate, spindle-shaped cells that express a variety of identified transduction-related proteins (including taste receptors) and therefore likely function as TRCs. TRCs possess apical processes with microvilli that protrude into the oral cavity and interact with taste stimuli. These microvilli contain taste receptors and taste-responsive ion channels. Tight junctions between the cells in a taste bud restrict the access of most stimuli to the apical membranes of the cells. TRCs synapse with afferent special sensory fibers from one of three cranial nerves (VII, IX, X), and taste information is subsequently relayed through brainstem nuclei to forebrain taste areas, or to local circuits controlling oromotor reflexes.

Multiple signaling cascades in TRCs have been described, and there are fundamental differences in the transduction pathways associated with each stimulus quality. For example, salt and acid stimuli directly permeate or gate apical ion channels, causing TRC depolarization, while umami-, sweet-, and bitter-tasting stimuli are thought to predominantly activate taste receptors.

G Protein-Coupled Taste Receptors

The largest gene superfamily in the mammalian genome encodes the group of proteins known as G protein-coupled receptors (GPCRs). GPCRs play critical roles in a variety of cellular functions, including neurotransmitter and hormonal signaling and the detection of sensory stimuli. Although they are quite diverse in amino acid sequence, all GPCRs share common structural and functional features. Their most striking structural motif is their seven helical transmembrane domains, i.e., the GPCR polypeptide passes 7 times through the plasma membrane, leaving an extracellular amino terminus and intracellular carboxy terminus. GPCRs also share a common signaling mechanism, as is evidenced by their name: activation of a GPCR by its ligand initiates an intracellular signaling cascade via the stimulation of an associated heterotrimeric guanosine triphosphate (GTP)-binding protein (G protein). Most chemosensory receptors are GPCRs, including the olfactory receptors, vomeronasal receptors, and taste receptors. In the taste system, three families of GPCR-type taste receptors (T1Rs, T2Rs, and taste-mGluR4) are implicated in the detection of sweet- and bitter-tasting stimuli and of certain amino acids.

T1R Receptors

The first two members of the T1R taste-receptor family were identified only a few years ago by Nicholas Ryba, Charles Zuker, and colleagues. Now called T1R1 and T1R2, they were cloned through differential screening of taste and nontaste tissues from the tongue. The gene encoding a third family member, T1R3, was identified
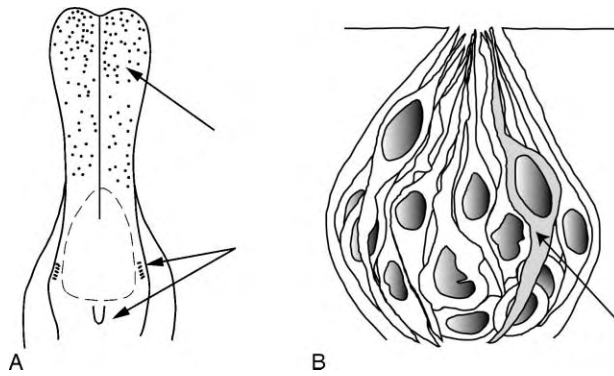
FIGURE 1 Taste-receptor cells. (A) Taste buds are specialized sensory structures located in the epithelia of the mammalian tongue (arrows indicate approximate locations of taste buds), as well as palate, larynx, and epiglottis (not shown). (B) Schematic of taste bud. Taste buds contain 50–100 cells of morphologically distinct types, including elongate, spindle-shaped cells (arrow) that express taste receptors.

by a number of investigators in a very different way: they determined that the *Tas1r3* gene corresponds to a genetic locus that confers taste sensitivity to saccharin and other sweeteners in mice. The T1Rs are class C GPCRs and share a large N-terminal extracellular domain, which comprises $\sim 50\%$ of the protein length, with other GPCRs of that class (Figure 2). The N-terminal domain contains the site of ligand binding in many class C GPCRs. However, it is unclear if the T1Rs interact with their ligands in a similar manner.

The mapping of the *Tas1r3* gene to a genomic locus conferring saccharin taste sensitivity suggested that T1Rs may play a role in the detection of sweet-tasting compounds. Further experimentation has supported this hypothesis. Experiments using either transgenic overexpression of T1R3 in mice or heterologous

expression of various T1Rs in cultured mammalian cells have shown that the T1Rs are receptors for sugars, artificial sweeteners, sweet proteins, and amino acids. Additionally, heterologous expression studies suggest T1Rs function as heteromultimeric complexes. Specifically, T1R1 and T1R3 together form a functional receptor for umami-tasting stimuli (e.g., L-amino acids) but are unresponsive to sweet-tasting stimuli, while T1R2 and T1R3 together comprise a receptor for sugars, artificial sweeteners, sweet proteins, and D-amino acids. The native T1R1 and T1R2 subunits are coexpressed with T1R3 in different subsets of taste-receptor cells (TRCs): T1R3 and T1R1 are found in the same TRCs of the anterior tongue, while T1R3 and T1R2 are found in the same TRCs of the posterior tongue. These observations have several implications. First, no T1R subunit forms a functional receptor alone. Second, T1R3 is a common subunit for receptors with different ligand specificities. Third, these ligand specificities are largely dependent on whether the receptor complex contains T1R1 or T1R2.

However, more recent studies from Robert Margolskee and colleagues suggest that T1Rs are not the only receptors for these stimuli. Mice in which the gene encoding T1R3 has been deleted from the genome display a normal sensitivity to the taste of glucose, a reduced sensitivity to sucrose, and are indifferent to several artificial sweeteners that are normally preferred by mice. Additionally, the taste sensitivity of these mice to monosodium glutamate is reduced at low concentrations of the stimulus when compared to wild-type (i.e., normal) mice. While these findings suggest that mammals have other receptors for sweet and amino acid stimuli, an alternative interpretation, supported by the work of Zuker and colleagues, is that T1Rs can function *in vivo* as homomeric receptors.
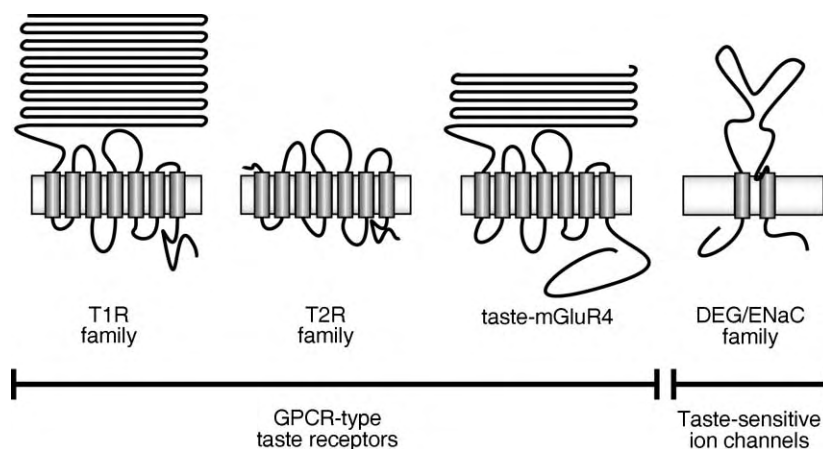
FIGURE 2 The proposed structures of taste-receptor proteins. Amino acid and sweet- and bitter-tasting stimuli are detected by GPCR-type taste receptors. Salt and acid stimuli activate TRCs through a GPCR-independent mechanism; they directly permeate or modulate ion channels, some of which may belong to the DEG/ENaC family.

T2R Receptors

A second class of taste receptors, the T2Rs, was identified by Ryba, Zuker, and colleagues, as well as by the laboratory of Linda Buck. They reasoned that genetic loci linked to bitter-taste sensitivities might contain genes encoding GPCRs involved in the detection of bitter tastants. Data mining of mouse and human genome databases bore out this hypothesis. To date, ~ 33 mouse and 25 human genes encoding T2Rs have been identified (each species also contains several apparent pseudogenes). T2Rs are class A GPCRs (the same class as olfactory receptors and rhodopsin), which are characterized by a short N terminus (Figure 2). Many class A GPCRs bind their ligands within a pocket defined by several of the transmembrane helices and/or the nearby extracellular loops; however, the site of ligand binding for T2Rs has not been determined.

To date, only a couple of T2Rs have been clearly linked to the detection of bitter-tasting compounds. The evidence implicating the mouse T2R5 receptor in the detection of cycloheximide is compelling: the genetic locus for cycloheximide taste sensitivity maps to a region of mouse chromosome 6 that contains the gene encoding T2R5; T2R5 specifically responds to cycloheximide in a heterologous expression assay; and polymorphisms in T2R5 that correlate with a reduced behavioral taste sensitivity to cycloheximide also confer a reduced sensitivity to this compound in an *in vitro* assay. In humans, the T2R38 receptor has been linked to phenylthiocarbamide (PTC) taste sensitivity. The gene encoding this receptor was recently mapped to a PTC-sensitivity locus in humans. Furthermore, polymorphisms in this gene correlate with variations in PTC-taste sensitivity between individuals in the sampled population.

As T2Rs display only a 30–70% sequence identity within the family, what is the evidence that other T2Rs are involved in bitter taste? First, many *Tas2r* genes are found in large clusters on one or two chromosomes where genetic loci for bitter-taste sensitivities have been mapped (two clusters are on mouse chr. 6 and one each on human chrs. 7 and 12). Second, localization studies have shown that numerous T2Rs are expressed in individual TRCs (primarily within bitter-responsive taste buds of the posterior tongue).

The Taste-mGluR4 Receptor

Another class C GPCR has been suggested to play a role in amino acid taste. Metabotropic glutamate receptors (mGluRs) are found throughout the nervous system, where glutamate serves as a major excitatory neurotransmitter. Nirupa Chaudhari and colleagues determined that one type of mGluR, mGluR4, is expressed in a subset of TRCs. Interestingly, the TRC-specific version of mGluR4 varies somewhat from the protein found in brain. The newly dubbed "taste-mGluR4", the product of an alternative transcript of the mGluR4 gene, displays a truncated N-terminal domain that is approximately half the length of other class C GPCRs. Heterologous expression of the taste-mGluR4 showed a glutamate dose-response curve to glutamate that is consistent with human taste thresholds, supporting a role for taste-mGluR4 in umami taste.

Ion Channels as Taste Receptors

Taste cells rely on the actions of a variety of ion channels to support stimulus-induced changes in membrane polarization and the synaptic release of neurotransmitter. However, several ion channel species can be thought of as receptors for taste stimuli: the channels interact directly with the taste stimulus and participate in the initial stage of taste transduction. Such channels play a central role in the detection of salty- and sour-tasting stimuli. A large portion of NaCl (salty) taste appears dependent on the influx of Na^+ through the epithelial sodium channel ENaC. The ENaCs, members of the degenerin (DEG)/ENaC superfamily of ion channels, are strongly Na^+ selective. DEG/ENaC family members display a common topology: they have two transmembrane domains, a large extracellular loop and a small pore-forming loop (Figure 2). They function as hetero-oligomeric complexes. Data implicating ENaCs in NaCl taste include: the presence of amiloride-blockable, Na^+-selective channels in TRCs that share a number of physiological properties with the ENaCs, and the expression of three homologous ENaC subunits, α, β, and γ, in TRCs.

Sour (acid) taste stimuli (i.e., H^+) depolarize TRCs either by activating cation channels or by directly permeating ion channels. As is the case for the transduction of Na^+, members of the DEG/ENaC channel family appear to play a major role in the transduction of acids: ENaC itself can conduct H^+, while ASIC-β acid-sensing ion channel-β (ASIC-β) and MDEG1/BNaC1 (mammalian degenerin-1/brain-type Na^+ channel-1), are activated directly by protons and are expressed in taste tissue.

Transduction Cascades

Unlike olfactory transduction, where a single intracellular pathway serves virtually all stimuli, the taste system appears to have a fairly diverse set of transduction mechanisms, both within and between stimulus classes. TRCs express many second-messenger components common to GPCR-coupled cascades such as cyclic nucleotide and phosphinositide signaling systems, and a transient receptor potential-related channel, TRPM5,

has also been implicated in taste transduction. TRPM5 appears to mediate capacitive calcium entry and is likely activated by the emptying of internal Ca^{2+} stores. The influx of calcium via TRPM5 may contribute to the receptor potential and/or mediate neurotransmitter secretion onto afferent fibers.

SEE ALSO THE FOLLOWING ARTICLES

G Protein-Coupled Receptor Kinases and Arrestins • Glutamate Receptors, Metabotropic • Neurotransmitter Transporters • Olfactory Receptors

GLOSSARY

taste quality   The perceptual quality attributed to a taste stimulus, such as salty, sour, sweet, bitter, or umami.

taste receptor   GPCR that binds sweet, bitter, or amino acid stimuli, leading to activation of a transduction cascade, receptor cell depolarization and transmitter release onto afferent nerve fibers.

taste-receptor cell   Elongate, spindle shaped cell in taste buds that expresses taste receptors and/or transduction components.

transduction   The process, often involving a biochemical cascade, by which one type of signal (e.g., a taste stimulus) is converted to another type (e.g., depolarization of the TRC).

FURTHER READING

Doty, R. L. (ed.) (2003). *Handbook of Olfaction and Gustation (Neurological Disease and Therapy)*, 2nd edition, Vol 57. Marcel Dekker, New York.

Gilbertson, T. A., and Boughter, Jr. J. D. (2003). Taste transduction: Appetizing times in gustation. *Neuroreport* **14**, 905–911.

Lindemann, B. (2001). Receptors and transduction in taste. *Nature* **413**, 219–225.

Margolskee, R. F. (2002). Molecular mechanisms of bitter and sweet taste transduction. *J. Biol. Chem.* **277**, 1–4.

Montmayeur, J. P., and Matsunami, H. (2002). Receptors for bitter and sweet taste. *Curr. Opin. Neurobiol.* **12**, 366–371.

BIOGRAPHY

John D. Boughter Jr. is an Assistant Professor of Anatomy and Neurobiology at the University of Tennessee Health Science Center. His principal research interests include the behavior genetics, physiology, and anatomy of the taste system. He holds a Ph.D. from Florida State University and received his postdoctoral training at the University of Maryland School of Medicine.

Steven D. Munger is an Assistant Professor of Anatomy and Neurobiology at the University of Maryland School of Medicine. His principal research interests include the molecular mechanisms of olfactory and taste transduction. He holds a Ph.D. from the University of Florida and received his postdoctoral training in the Howard Hughes Medical Institute at the Johns Hopkins Medical Institutes.

T-Cell Antigen Receptor

Andrea L. Szymczak and Dario A. A. Vignali
St. Jude Children's Research Hospital, Memphis, Tennessee, USA

The T-cell antigen receptor (TCR) is a highly organized, multi-molecular complex found exclusively on T and NK T cells. The majority of TCRs are composed of TCRα and TCRβ chains, while a small percentage contain TCRγ and TCRδ chains. The TCR chains heterodimerize and are associated with the invariant chains of the CD3 complex, CD3$\varepsilon\gamma\delta$ and CD247 (often referred to as the zeta chain, TCRζ or CD3ζ). Each chain is essential for TCR surface expression and consequently T cell development and function. This complex mediates the development of T cells that are able to mount a specific immune response against a wide variety of foreign antigens and pathogenic organisms. This unique capability is due to the unusual genomic organization and rearrangement of TCR genes, which generates a vast number of TCRs that recognize almost any antigenic peptide in the context of major histocompatibility (MHC) molecules.

TCR:CD3 Complex: Genes, Proteins and the Receptor Complex

T Cell Antigen Receptor

T cells can only recognize and bind antigens when they are presented by major histocompatibility (MHC) molecules, a process called MHC restriction. Initially, the T cell antigen receptor (TCR) was difficult to study because it was membrane bound and could only recognize antigen in the context of MHC. The first TCR complexes isolated were composed of a disulfide-linked heterodimer of TCRα and TCRβ. Monoclonal antibodies against this complex were either specific for a particular T cell clone or nonspecific, indicating that the TCR chains were similar to the immunoglobulin (Ig) molecules, containing both variable (V) and constant (C) regions. While the vast majority ($>95\%$) of T cells had TCR$\alpha\beta$ heterodimers, another type of TCR containing a TCRγ and TCRδ heterodimer was later identified. The TCR chains, in particular the variable regions, are responsible for recognizing and binding to peptide–MHC molecules presented on the surface of antigen presenting cells (APC).

Genes: Organization and Rearrangement

As with antibody molecules, T cells must be able to recognize a wide variety of pathogens, therefore a diverse repertoire of TCR must be generated that recognize such antigens. Many similarities exist between the Igs and the TCR. The four TCR loci (α, β, γ, and δ) have a germ-line organization similar to that of Ig. The α- and γ-chains are produced by rearrangement of V and J segments, while the β- and δ-chains are produced by rearrangement of V, D, and J segments (Figure 1). The segments for TCRδ are located between the Vα and Jα segments. When the TCRα gene is rearranged, the TCRδ segments are deleted, thereby ensuring the T cell does not express TCR$\gamma\delta$ and TCR$\alpha\beta$ at the same time. The basic domain structure is evolutionarily conserved and is believed to have arisen through gene duplication. Organization of the TCR genes in the mouse and human is similar but vary in the numbers of segments (Table I).

The TCR genes encode areas of hypervariability similar to the complementarity determining regions (CDRs) of antibody molecules. Embedded within the V regions are CDR1 and CDR2. There is an additional area of hypervariability (HV4) that does not appear to interact with antigen and is therefore not considered a CDR (Figure 1). CDR3 is a result of VJ or VDJ joining. This intentionally imprecise process frequently results in the addition or deletion of nucleotides, adding significantly to the variability in this region (N region diversity). Occasionally, a nonproductive rearrangement of the first gene locus occurs due to an out-of-frame rearrangement or insertion of a stop codon. This induces the rearrangement of the TCR locus on the other chromosome.

Rearrangement occurs through the action of recombination-activating genes (RAGs) 1 and 2 during T cell development. Flanking each gene segment in the germ line DNA are conserved recognition signal sequences. The RAG enzymes recognize these heptamer and nonamer sequences and catalyze the VJ and VDJ joining with a mechanism analogous to that used to rearrange the Ig locus. In addition, secondary rearrangements termed receptor editing or revision can occur in both B and T cells. The role of receptor editing and revision in
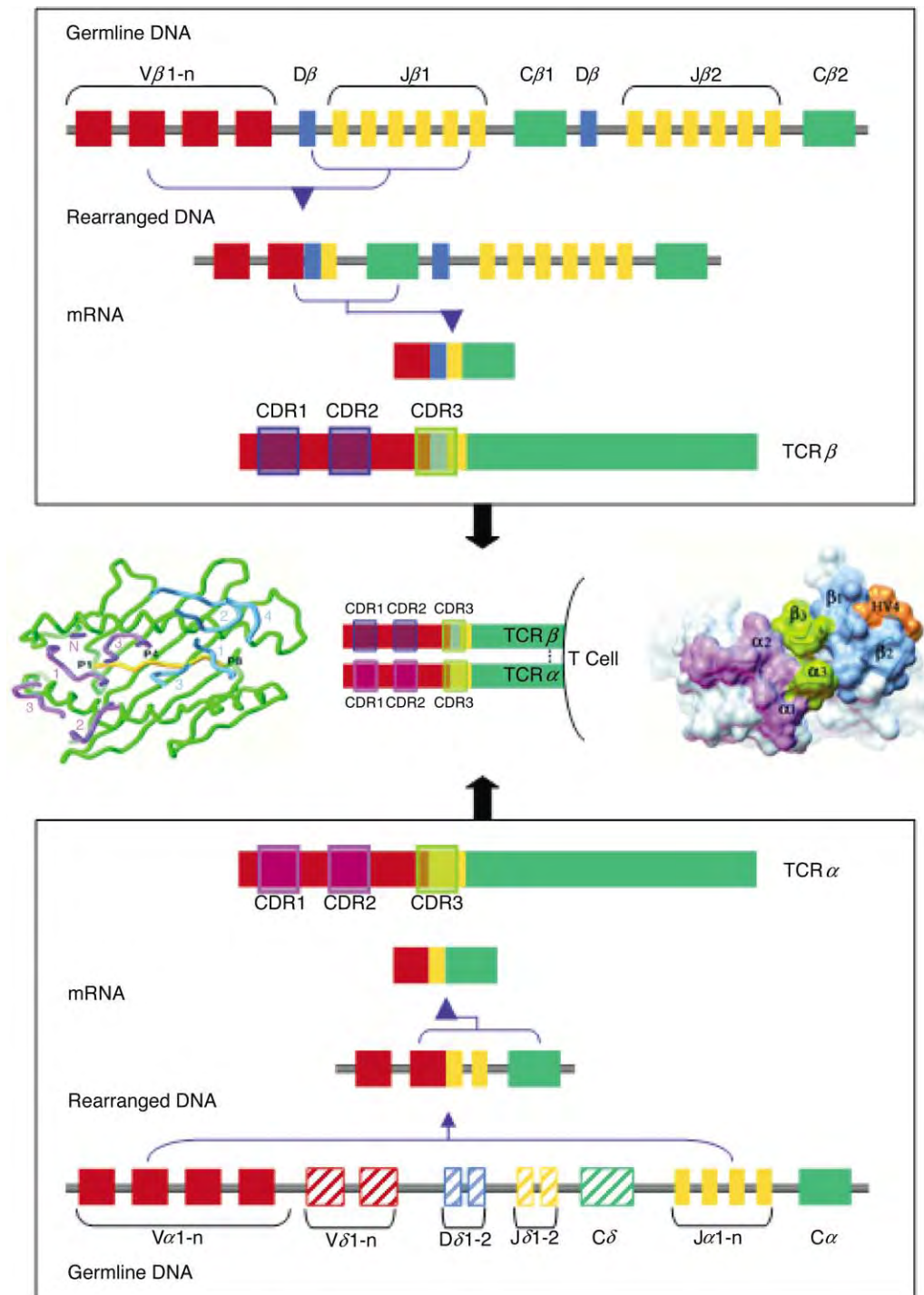
FIGURE 1 Rearrangement of TCRαβ genes. Example of TCRα (bottom panel) and TCRβ (top panel) rearrangement. TCRα undergoes V to J rearrangement while TCRβ undergoes D to J followed by V to DJ rearrangement. Primary transcripts of rearranged DNA are processed to mRNA that encode the mature TCRαβ chains expressed on the T cell surface (center panel, middle). Areas encoding the CDR regions are boxed. TCRδ segments in the TCRα locus are shown in stripes. The center left panel shows a view into the peptide–MHC combining site with peptide in yellow and MHC in green superimposed with the CDR regions (CDR1, 2 and 3; HV4) of TCRα (purple) and TCRβ (blue). CDR1 of Vα and Vβ and Vβ CDR3 are clearly positioned along the central axis of the peptide. N = the N-terminal residue of Vα. (Reprinted from Garcia, K. C., Degano, M., Pease, L. R., Huang, M., Peterson, P. A., Teyton, L., and Wilson, I. A. (1998). Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* **279**, 1166–1172, with permission of AAAS.) The center right panel shows the surface of the TCR-binding site. The surface of the loop trace of the Vα CDRs 1 and 2 are purple; CDRs 1 and 2 of TCRβ are blue; Vα and Vβ CDR3s are green; and the hypervariable region of TCRβ (HV4) is orange. (Reprinted from Garcia, K. C., Degano, M., Stanfield, R. L., Brunmark, A., Jackson, M. R., Peterson, P. A., Teyton, L., and Wilson, I. A. (1996). An αβ T cell receptor structure at 2.5Å and its orientation in the TCR-MHC complex. *Science* **274**, 209–219, with permission of AAAS.)

TABLE I

Chromosomal Location and Numbers of TCR Gene Segments

	Human				Mouse			
	TCRα (14)	TCRβ (7)	TCRγ (7)	TCRδ (14)	TCRα (14)	TCRβ (6)	TCRγ (13)	TCRδ (14)
V	40–50	30–57	11–14	3	50–100	20–50	7	10–11
D		2		3		2		2
J	60–70	13–14	5	3–4	60–100	12–14	4	2
C	1	2	2–3	1	1	2	4	1

Total number of gene segments for the human and mouse TCR. Numbers in parenthesis indicate the chromosome on which the gene is located. Reprinted from Allison, T. J., and Garboczi, D. N. (2001). Structure of γδ T cell receptors and their recognition of non-peptide antigens. *Mol. Immunol.* **38**, 1051–1061, with permission from Elsevier.

T cells is still unclear, but it has been suggested to contribute to peripheral T cell tolerance.

While many similarities exist between Ig and TCR, there is one important distinction: TCR genes do not undergo somatic hypermutation and thus there is no affinity maturation. While the number of TCR V region segments is dramatically reduced compared to the Ig genes, the number of J region gene segments is greater, increasing the potential diversity at the V–J interface (Table II). This is likely due to differences in the nature of the antigen recognized by TCR and Ig.

TCR Structure and Recognition of Peptide–MHC Complexes

The TCR is a member of the Ig super family. Structural analysis has shown that the Vα and Vβ domains pair via a conserved hydrophobic core, while the Cα and Cβ domains pair via a highly polar interface, with a skewed distribution of acidic residues in Cα and basic residues in Cβ. The two chains also pair via a conserved disulfide bond close to the membrane (Figure 2). In addition, the C region contains a connecting peptide, a transmembrane region and short intracellular domain. The anchoring transmembrane domain is unusual as it contains several highly conserved basic amino acid residues (TCRα–Lys/ and –Arg; TCRβ–Lys). These positively charged residues mediate interaction with the CD3 complex. The cytoplasmic domains are short, consisting of only 5–12 amino acids.

TCR specificity is mediated by the Vα and Vβ CDRs which interact with the peptide–MHC complex. MHC molecules are bound to the APC membrane and the antigenic peptide is bound in a groove between two α-helices (Figure 2). A number of TCR:peptide–MHC complexes have been crystallized. Although specific contact residues vary, all have a similar mode of binding in which the CDR2 regions contact the MHC surface and CDR1 and CDR3 contact both the peptide and MHC molecules (Figures 1 and 2). The V region of TCRα binds closer to the N-terminal region of the peptide while that of TCRβ is closer to the C terminus. The TCR binds in a diagonal orientation due to two peaks created by the MHC α-helices.

The binding properties for a number of TCR: peptide–MHC complexes have been described. In general, the on and off rate for TCR:peptide–MHC binding is fast and the affinity is low compared to other receptor–ligand interactions. These biophysical properties of the TCR contrast starkly with the generally high affinity of Ig:antigen interaction. However, in spite of these properties, TCR interaction is sufficiently stable to initiate signal transduction.

In most instances, only a very small fraction of MHC molecules on a cell contain the right antigenic peptide for an individual TCR. It has been proposed that T cells overcome this limited ligand supply by a process called serial ligation, where many TCRs (perhaps as many as 200) can be ligated by a single

TABLE II

Table Sequence Diversity in TCR and Ig Genes

	Ig		TCRαβ		TCRγδ	
	H	κ	α	β	γ	δ
V segments	250–1000	250	50	25	7	10
D segments	10	0	0	2	0	2
Ds read in all frames	Rarely			Often		Often
N-region addition	V–D, D–J	None	V–J	V–D, D–J	V–J	V–D1, D1–D2, D1–J
J segments	4	4	50	12	2	2
V region combinations	62 500–250 000		1250		70	
Junctional combinations	$\sim 10^{11}$		$\sim 10^{15}$		$\sim 10^{18}$	

(Reprinted from Davis, M. M. (1990). T cell receptor gene diversity and selection. *Annu. Rev. Biochem.* **59**, 475–496, with permission of Annual Reviews.)
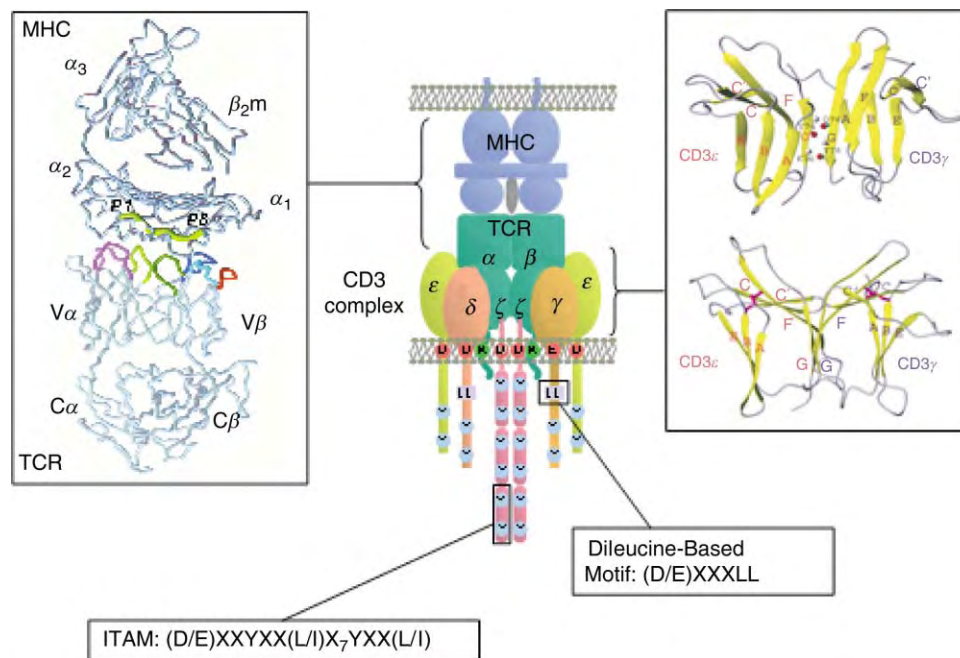
FIGURE 2 TCR:CD3 complex and its interaction with peptide–MHC. Schematic of the presumed stoichiometry of the TCR:CD3 complex. Highlighted are charged transmembrane residues, the ITAM sequence and the dileucine-based motif. In the upper left panel is a backbone structure of TCR:peptide–MHC complex. TCR is on the bottom and the MHC on top. The peptide (P1−8) is shown as a large tube in yellow. Vα CDRs 1 and 2 are in light and dark purple, respectively; αHV4 in white; Vβ CDRs 1 and 2, in light and dark blue, respectively; βHV4 in orange; and Vα and Vβ CDR3s in light and dark yellow, respectively. (Reprinted from Garcia, K. C., Degano, M., Stanfield, R. L., Brunmark, A., Jackson, M. R., Peterson, P. A., Teyton, L., and Wilson, I. A. (1996). An $\alpha\beta$ T cell receptor structure at 2.5Å and its orientation in the TCR-MHC complex. *Science* **274**, 209–219, with permission of AAAS.) In the upper right panel is a ribbon diagram of CD3$\varepsilon\gamma$. The beta strands are in yellow and the text colored red for CD3ε and blue for CD3γ. Three pairs of atoms involved in the hydrogen bond formation are designated with amide protons in gray and carbonyl oxygen atoms in red. Two disulfide-linked cysteine residues are shown. (Reproduced from Sun, Z. J., Kim, K. S., Wagner, G., and Reinherz, E. L. (2001). Mechanisms contributing to T cell receptor signaling and assembly revealed by the solution structure of an ectodomain fragment of the CD3γ heterodimer. *Cell* **105**, 913–923, with permission from Elsevier.) In the lower left and right panels are the ITAM and dileucine-based consensus motifs, respectively.

peptide–MHC complex. It is thought that the low affinity and rapid off rate of TCR:peptide–MHC interaction may be instrumental in mediating this process.

INVARIANT CHAINS: CD3, CD247, AND PRE-Tα

CD3 and CD247

The TCR is associated with the invariant chains of the CD3 complex.: CD3ε, CD3γ, and CD3δ are members of the Ig superfamily with one extracellular Ig domain followed by a transmembrane domain and cytoplasmic tail of ~40 amino acids in length. The zeta family of molecules are unique and have a short extracellular sequence and long cytoplasmic tail. This family consists of CD247 (otherwise known as CD3ζ or TCRζ), its splice variant η, and the γ-chain of the Fc receptor (FcRγ). The CD247 ζ- and η-chains are alternatively spliced gene products and are identical except for the carboxyl-terminal region of the cytoplasmic tail (113 and 155 amino acids long for ζ and η, respectively).

Analysis of CD3 and CD247 knockout mice demonstrates their importance in TCR expression and T cell

development and function. Mice lacking CD247 (CD3ζ) and CD3γ exhibit a substantial block in early stages of T cell development, whereas mice lacking CD3δ develop a block at a later stage. CD3ε knockout mice display a complete arrest in T cell development, likely due to its requirement to form heterodimers with CD3γ and CD3δ.

The CD3 and CD247 chains contain a number of amino acid residues and motifs that are important for assembly, signal transduction, and regulation of cell surface expression (Figure 2). The transmembrane regions contain negatively charged amino acids that interact with the positively charged residues of TCR$\alpha\beta$ (CD3δ, CD3ε, and CD3ζ-Asp; CD3γ-Glu). The cytoplasmic tails contain immuno-receptor tyrosine-based activation motifs (ITAM) that when phosphorylated provide docking sites for SH2 domain-containing proteins, which are important for downstream signal transduction. Each CD3$\varepsilon\gamma\delta$ chain contains one ITAM while CD247 (CD3ζ) contains three. CD3δ and CD3γ also contain a dileucine-based motif. Both the YXXL sequence in the ITAM and the dileucine sequence have been shown to play a role in internalization and downmodulation of many types of receptors from the cell

surface. It has been suggested that these motifs are also utilized for TCR transport. A number of studies using T cell lines *in vitro* have shown a role for the CD3γ dileucine-based motif in TCR down-modulation, however its role *in vivo* remains to be defined.

Pre-Tα

During T cell development in the thymus, TCR$\alpha\beta$ precursors express a pre-TCR. The pre-TCR is formed by the functionally rearranged TCRβ chain, the CD3 complex and a surrogate TCRα chain, pre-TCRα (pTα). There are two notable differences between pTα and TCRα. First, pTα only has one invariant extracellular Ig domain. Second, the cytoplasmic tail of pTα is longer and contains two potential serine and threonine phosphorylation sites and a potential SH3 domain-binding motif. The requirements for surface expression of pTα are still unclear. However, it is clear that a number of signaling events are initiated through the pre-TCR. Pre-TCR signaling confirms the successful rearrangement TCRβ and induces the suppression of further TCRβ locus rearrangement, a process called allelic exclusion. Rearrangement of the TCRα locus then occurs leading to progression of T cell development.

TCR:CD3 COMPLEX

Expression of the TCR on the cell surface requires all six chains of the complex. While the exact stoichiometry and make-up of the TCR:CD3 complex is unclear, it is generally accepted that the complex consists of four dimers: TCR$\alpha\beta$ or $\gamma\delta$, heterodimers of CD3$\varepsilon\gamma$ and CD3$\varepsilon\delta$, and a zeta family dimer. The majority of TCR:CD3 complexes (80–90%) contain a CD247 ζ homodimer. Heterodimers of ζ-η or ζ-FcRγ or homodimers of FcRγ have been observed in a small percentage of TCR complexes. An organizing principle has been proposed in which a single negatively charged amino acid in the TCR dimer interacts with two positively charged residues in the CD3$\varepsilon\gamma$, CD3$\varepsilon\delta$, and CD247 ζ dimers. In this model, the TCRα lysine interacts with the CD3$\varepsilon\delta$ dimer, the TCRβ lysine interacts with the CD3$\varepsilon\gamma$ dimer and the TCRα arginine interacts with the CD247 ζ dimer.

Role of the TCR in T Cell Biology and Signaling

T CELL DEVELOPMENT

T cells that exit the thymus have been through a selection process based largely on TCR affinity. T cells with TCR that recognize self peptide–MHC too strongly are negatively selected and deleted. T cells that have too low an affinity to productively interact with self-MHC are ignored and ultimately die of neglect. However, T cells that display a moderate affinity for the self-MHC molecules are positively selected and allowed to exit into the periphery.

CELL BIOLOGY

It is known that each of the six protein chains is required for correct assembly and surface expression of the TCR:CD3 complex. Once expressed on the cell surface, the TCR:CD3 complex is constitutively internalized and recycled back to the cell surface via the endosomal network. Once ligated by peptide–MHC complexes, the TCR is down-modulated from the cell surface and diverted to lysosomes for degradation.

Assembly and Surface Expression of the TCR:CD3 Complex

Assembly of the complex is a highly ordered process that takes place in the ER. Most of the evidence for assembly order supports a model in which CD3$\varepsilon\gamma$ and CD3$\varepsilon\delta$ heterodimerize followed by sequential addition of TCRα and TCRβ chains and the ζ_2 homodimer (or heterodimer). Once the correct stoichiometry is achieved, the intact complex is transported from the ER. It has been shown that the ζ-chain can exit the ER independently of the rest of the receptor complex and remain in the Golgi complex. CD3γ and CD3ε have been reported to partly exist as heterodimers in association with calnexin while the TCRα and β-chains have been found to associate with calreticulin. It has been suggested that these molecules may serve as chaperones to prevent the transport of partial complexes. A number of residues are present within the chains that serve as ER retention signals or to target partial complexes to lysosomes for degradation, thereby ensuring that only complete TCR:CD3 complexes are expressed on the cell surface. The charged residues in the transmembrane regions increase their susceptibility to ER degradation. A number of additional residues have been described as possible ER retention or degradation signals, however a detailed mechanism of complex assembly, including specific interactions between individual chains, remains to be defined.

Internalization and Down-Modulation of the TCR:CD3 Complex

The TCR:CD3 complex is constitutively internalized and recycled back to the cell surface. Although the exact mechanism of this process has not yet been defined, it is likely to involve the interaction of the CD3/CD247 molecules with adaptor protein (AP) complexes that are associated with clathrin at the plasma membrane and intracellular recycling vesicles of the endosomal network. The dileucine-based motif in CD3γ has been

shown to be capable of binding to a member of the AP family of complexes, AP-2. In addition to dileucine based motifs, AP-2 can recognize YXXL-based sequences in a number of receptor systems. Given that the TCR:CD3 complex contains 20 such sequences, it is possible that one or more may be utilized for TCR internalization.

Upon ligation with peptide–MHC complexes, the TCR:CD3 complex is down-modulated from the cell surface and recycling is prevented. While the exact mechanism of this important process is unknown, it may involve two related E3 ubiquitin ligases, c-Cbl and Cbl-b, as T cells from mice lacking both proteins fail to down-modulate their TCR following ligation. While both internalization and down-modulation are hallmarks of TCR biology, their physiological importance and function remains to be determined.

SIGNALING THROUGH THE TCR

Initiation of T cell activation occurs when the TCR recognizes peptide–MHC complexes. TCR$\alpha\beta$ consists of the ligand-binding unit while the CD3 complex transduces signals into the T cell. Clustering of TCR:peptide–MHC complexes brings in the coreceptors CD8 or CD4 which bind to MHC class I and II molecules, respectively. Both coreceptors are associated with Src-related protein tyrosine kinase (PTK) p56^{lck}, while another PTK, p59^{fyn}, interacts with the CD3 complex. In resting T cells, these kinases are inactive due to the interaction of the C-terminal phosphotyrosine residues binding to the N-terminal SH2 domain. This intramolecular interaction prevents substrate access to the kinase (SH1) domain. T cell:APC interaction induces the removal of these inhibitory phosphates by the transmembrane phosphatase CD45. Cross phosphorylation of active site tyrosine residues further potentiates p56^{lck} and p59^{fyn} kinase activity and results in the phosphorylation of the ITAM tyrosine residues in the CD3/CD247 cytoplasmic tails. Phosphorylation of both ITAM tyrosine residues is required for docking of a specialized PTK, zeta associated protein-70 (ZAP-70), which has two tandem SH2 domains. ZAP-70 kinase activity is further potentiated through phosphorylation by p56^{lck} and p59^{fyn}.

Activated ZAP-70 initiates a number of signaling pathways. A key target of ZAP-70 is the raft-resident linker for activated T cells (LAT) which is heavily phosphorylated and recruits a wide variety of downstream signaling molecules. Phosphorylation of phospholipase Cγ_1 (PLCγ_1) leads to the production of potent second messengers, diacylglycerol (DAG) and inositol triphosphate (IP$_3$), whose actions lead to protein kinase C (PKC) activation and NFκB nuclear translocation as well as Ca^{2+} release and nuclear factor of activated T cells (NFAT) translocation. These events lead to the transcription of genes required for T cell proliferation and interleukin-2 (IL-2) production. Signaling through ZAP-70 also initiates activation of the Ras pathway and the MAPK signaling cascade which also results in upregulation of genes required for proliferation. A wide array of additional signaling molecules and adaptor proteins have been shown to contribute to the signaling cascade initiated following TCR ligation and have been reviewed extensively elsewhere.

SEE ALSO THE FOLLOWING ARTICLES

Immunoglobulin (Fc) Receptors • Mitogen-Activated Protein Kinase Family • Protein Kinase C Family • Src Family of Protein Tyrosine Kinases

GLOSSARY

CD3 Complex of polypeptides containing three dimers: $\varepsilon\gamma$ heterodimers, $\varepsilon\delta$ heterodimers and, most frequently, $\zeta\zeta$ homodimer (CD247). It is associated with the TCR through charged transmembrane residues and is involved in transducing signals into the T cell upon TCR:peptide–MHC interaction.

complementarity-determining region (CDR) Areas in the variable regions of antibody and TCR genes. In the TCR, the CDR regions contact the peptide and MHC molecule on antigen presenting cells.

immunoglobulin superfamily Group of proteins that contain immunoglobulin-fold domains of $\sim$100 amino acids folded into two β-pleated sheets and stabilized by a central disulfide bond. Included in the family are MHC molecules, TCRs and a number of CD antigens.

major histocompatibility complex (MHC) A complex of polymorphic genes that encode histocompatibility antigens termed H2 in the mouse and HLA in humans. Two main classes of MHC antigens are found as surface glycoproteins on antigen presenting cells that bind and present peptides to T cells.

TCR Heterodimer of TCR$\alpha\beta$ or TCR$\gamma\delta$ expressed on the surface of T cells that is associated with the CD3 complex. The TCR binds to peptide–MHC molecules.

FURTHER READING

Allison, T. J., and Garboczi, D. N. (2001). Structure of $\gamma\delta$ T cell receptors and their recognition of non-peptide antigens. *Mol. Immunol.* **38**, 1051–1061.

Call, M. E., Pyrdol, J., Wiedmann, M., and Wucherpfennig, K. W. (2002). The organizing principle in the formation of the T cell receptor-CD3 complex. *Cell* **111**, 967–979.

Davis, M. M. (1990). T cell receptor gene diversity and selection. *Annu. Rev. Biochem.* **59**, 475–496.

Germain, R. N., and Stefanova, I. (1999). The dynamics of T cell receptor signaling: Complex orchestration and the key roles of tempo and cooperation. *Annu. Rev. Immunol.* **17**, 467–522.

Glusman, G., Rowen, L., Lee, I., Boysen, C., Roach, J. C., Smit, A. F. A., Wang, K., Koop, B. F., and Hood, L. (2001). Comparative genomic of the human and mouse T cell receptor loci. *Immunity* **15**, 337–349.

Goldsby, R. A., Kindt, T. J., and Osborne, B. A. (eds.) (2000). In *Kuby Immunology*. W. H. Freeman, New York.

Garcia, K. C., Degano, M., Stanfield, R. L., Brunmark, A., Jackson, M. R., Peterson, P. A., Teyton, L., and Wilson, I. A. (1996). An $\alpha\beta$ T cell receptor structure at 2.5Å and its orientation in the TCR-MHC complex. *Science* **274**, 209–219.

Garcia, K. C., Degano, M., Pease, L. R., Huang, M., Peterson, P. A., Teyton, L., and Wilson, I. A. (1998). Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* **279**, 1166–1172.

Hennecke, J., and Wiley, D. C. (2001). T cell receptor-MHC interactions up close. *Cell* **104**, 1–4.

Kruisbeek, A. M., Haks, M. C., Carleton, M., Michie, A. M., Zuniga-Pflucker, J. C., and Wiest, D. L. (2000). Branching out to gain control: How the pre-TCR is linked to multiple functions. *Immunol. Today* **21**, 637–644.

Samelson, L. E., Harford, J. B., and Klausner, R. D. (1985). Identification of the components of the murine T cell antigen receptor complex. *Cell* **43**, 223–231.

Sun, Z. J., Kim, K. S., Wagner, G., and Reinherz, E. L. (2001). Mechanisms contributing to T cell receptor signaling and assembly revealed by the solution structure of an ectodomain fragment of the CD3γ heterodimer. *Cell* **105**, 913–923.

Valitutti, S., Muller, S., Cella, M., Padovan, E., and Lanzavecchia, A. (1995). Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature* **375**, 148–151.

BIOGRAPHY

Andrea Szymczak is a graduate student in the Department of Pathology at the University of Tennessee in Memphis, under the guidance of Dr. Dario Vignali at St. Jude Children's Research Hospital. She received her B.Sc. degree in biology at the University of Tennessee at Chattanooga. At present, her research interest is in T cell biology, specifically TCR:CD3 complex internalization and down-modulation.

Dario Vignali is an Associate Member in the Department of Immunology at St. Jude Children's Research Hospital. His research interests are the TCR:CD3 complex, regulation of T cell function and type 1 diabetes. He holds a Ph.D. from the London School of Hygiene and Tropical Medicine in England, and was a postdoctoral fellow at the German Cancer Research Center, Heidelberg, Germany, and Harvard University. His laboratory has made important contributions to the understanding of how the TCR recognizes MHC class II:peptide complexes and how T cells traffic TCR:CD3 complexes.

Tec/Btk Family Tyrosine Kinases

Shuling Guo and Owen N. Witte
Howard Hughes Medical Institute, University of California, Los Angeles, California, USA

Nonreceptor tyrosine kinases (NRTKs) are cytoplasmic enzymes that phosphorylate tyrosine residues when activated and thereby play critical roles in many signal transduction pathways in multicellular organisms. These kinases are grouped into families including Src, Syk, Abl, and Fak families, according to their protein sequence and structure similarities. In 1993, mutations in Bruton's tyrosine kinase (Btk) were demonstrated to cause human X-linked agammaglobulinemia (XLA) and murine X-linked immunodeficiency (xid). Since then, more proteins similar to Btk were discovered and Tec/Btk family tyrosine kinases became a new NRTK subfamily. This family is now the second largest nonreceptor tyrosine kinase family after the Src family.

Tec Family Kinase Members and Expression Pattern

The Tec family is composed of five mammalian members: Btk, Tec, Itk, Txk, and Bmx. These kinases are differentially expressed and most of them are found primarily in hematopoietic cells. This family is also expressed in other species, including *Drosophila melanogaster*, *skate*, and zebrafish. In addition, a Btk orthologue designated NRTK3 has been identified in the sea urchin *Anthocidaris crassispina*.

1. Btk is also known as Atk, Bpk, or Emb. Btk is expressed in all stages of B cell development except plasma cells. Btk is also expressed in myeloid and mast cells as well as early erythroid and megakaryocytic precursors, but Btk is not expressed in T cells. In tissues, Btk is found in bone marrow, spleen, lymph node, and fetal liver.

2. Tec (*t*yrosine kinase *e*xpressed in hepatocellular *c*arcinoma) is expressed in bone marrow, spleen, thymus, and liver. In cell lines, Tec is primarily found in T cells, myeloid cells, and hepatocarcinoma cells.

3. Itk (*i*nterleukin-2 inducible *T*-cell specific *k*inase, also known as Tsk or Emt) is primarily expressed in T cells, natural killer (NK) cells, and mast cells. The expression of Itk in T cells is developmentally regulated. Its expression begins at early fetal thymus and

the expression level is higher in murine thymus than peripheral T cells.

4. Txk (*T* and *X* cell expressed *k*inase, also known as Rlk) is found in T cells, NK cells, as well as myeloid and mast cell lines.

5. Bmx (*b*one *m*arrow *k*inase gene on the X chromosome, also known as Etk) was originally identified from a bone marrow library and subsequently in prostate cancer cells. This kinase is the only member of the Tec family that is not primarily expressed in hematopoietic cells. Bmx is mainly found in endothelial, epithelial, fibroblast, neutrophil, and carcinoma cells.

Tec Family Kinase Domain Structure

The general domain structure for Tec family kinases is formed of the amino-terminal pleckstrin homology (PH) domain, a Tec homology (TH) domain, Src homology 3 (SH3) and Src homology 2 (SH2) domains, and the kinase domain which is adjacent to the SH2 domain through an SH2-kinase linker region. This differs from the Src family kinases that have a lipid modification motif in the amino terminus instead of a PH domain. Another difference between Src and Tec family kinases is that Src kinases contain a carboxyl-terminal tyrosine phosphorylation site as a negative regulation mechanism that the Tec kinases lack (Figure 1).

PH Domain

The core structure of PH domain is a 7-β-sheet structure that is mainly involved in protein-lipid interactions. For the PH domain of Btk, the high-affinity ligands are phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P$_3$) and inositol 1,3,4,5-tetrakisphosphate (I(1,3,4,5)P$_4$). PI(3,4,5)P$_3$ is the product of PI 3-kinase and acts as the second messenger to recruit cytoplasmic Btk to the plasma membrane. This recruitment is a critical step for the activation of Btk, so it is not surprising that many mutations are found in the PH domain in XLA patients.
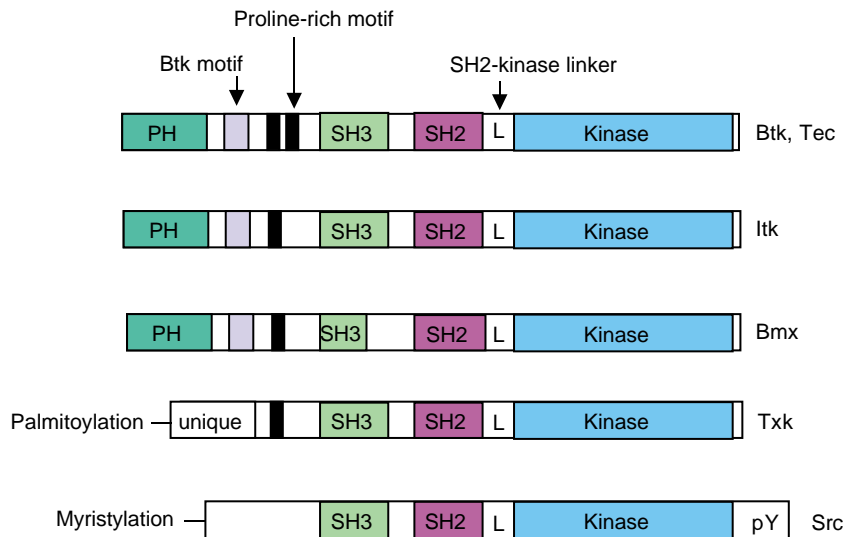
FIGURE 1   Domain structures of Tec family kinases. Src kinase is shown at the bottom for comparison.

The xid mutation (R28C) is also in the PH domain and the mutant PH domain has a greatly reduced ability to bind its ligand. On the other hand, a constitutively active mutant of Btk (E41K) showed improved membrane association capability. All Tec kinases contain the N-terminal PH domain except Txk. Txk has a unique region with a palmitoylated cysteine string, which serves a similar membrane-translocating function as the PH domain. As a result, Txk is not activated by PI 3-kinase.

TH DOMAIN

The TH domains in Tec family kinases consist of a Btk motif (BM) and proline-rich region. The Btk motif is a 27-amino-acid stretch containing a zinc-binding fold formed by conserved cysteine and histidine residues homologous to Ras GTPase activating protein (GAP). This Btk motif is absent in Txk. Following Btk motif are two consecutive proline-rich motifs in Btk and Tec, while only one proline-rich motif is found in Itk, Txk, and Bmx. These proline-rich motifs are able to bind SH3 domain, so they may participate in intermolecular or intramolecular interactions.

SH3 AND SH2 DOMAINS

SH3 domains in Tec family kinases are adjacent to the proline-rich motif(s). The Itk SH3 domain has been crystallized with the N-terminal proline-rich motif and the intramolecular proline-rich motif binds to the SH3 domain, suggesting that this interaction may serve as a mechanism for the regulation of enzyme activity. In contrast, Bmx has a truncated SH3 domain. SH2 domain binds phosphotyrosine, providing a docking site for regulatory proteins or effector proteins. In the Itk SH2 domain, there is a proline residue that is not

conserved in other Tec family kinases. This proline residue may undergo *cis–trans* conformational switch, possibly catalyzed by a peptidyl-prolyl isomerase cyclophilin A. This conformational change controls the orientation of the protein-binding surface of the SH2 domain and may affect the catalytic activity of Itk.

KINASE DOMAIN

The kinase domains of Tec kinases are highly conserved. The Btk kinase domain contains a small lobe and a large lobe, with the active site in between. This structure is very similar to Src family kinases and other tyrosine kinases. However, the 30-amino acid linker region between SH2 domain and the kinase domain is less conserved in the Tec family kinases and very different between the Tec and Src family kinases. This linker region has been shown to regulate the intracellular interaction of Src family kinases, yet the function of this linker region in Tec family kinases is not yet known.

Although these kinases are primarily cytoplasmic kinases, there has been evidence that Btk, Itk, and Txk can also be shuttled into the nucleus. Itk is translocated into the nucleus through the interaction with a nuclear import chaperone karyopherin α. A shorter form of Txk originated from internal initiation of translation gets into the nucleus via nuclear localization sequence (NLS)-dependent mechanism, while Btk is found in the nucleus through an NLS-independent way.

Tec Family Kinase Functions

To date, Btk is the only Tec family kinase that is involved in human disease when mutated. However, the physiological importance of all these kinases has been investigated through the murine knockout models.

Btk is essential for B-cell development and function, as demonstrated by XLA. XLA patients lack mature B cells in the periphery and do not have immunoglobulins as a result. However, inactivating Btk in the mouse causes only a mild defect in B cell development and function. This mimics the phenotype of xid mice, caused by a spontaneous mutation (R28C) in the Btk PH domain. In these mouse models, mature B-cell number is reduced and these B cells have a defect in response to B-cell receptor (BCR) stimulation. A minor B-cell population, B-1 cells, is absent in these mice. Moreover, the serum immunoglobulins IgG3 and IgM levels are greatly reduced in these mice and they are not able to respond to T-independent type-II antigens. Interestingly, even though Tec-/- mice have no major phenotypic alterations of the immune system, Btk-/-Tec-/- mice showed a severe defect in B-cell development, similar to human XLA. This suggests that Tec may compensate partially for the lack of Btk in murine B-cell development.

In T cells, three Tec family members coexist: Itk, Txk, and Tec. Inactivating Itk results in slightly decreased number of mature thymocytes, especially CD4$^+$ cells, while inactivating Txk does not affect either T-cell development or activation. Double mutants, on the other hand, showed improved mature T-cell number, while still maintaining a decreased CD4/CD8 ratio, compared to Itk-/- mice. However, T cells in the double knockout mice have a severe defect in T-cell receptor (TCR) induced proliferation and cytokine production.

When the Bmx gene was replaced by the LacZ reporter gene, the homozygous mice lacking Bmx activity showed no obvious phenotype. But the expression of the reporter gene is strong in endothelial cells of large arteries and in the endocardium from embryogenesis to adult mice. Moreover, Bmx is activated through endothelial receptor tyrosine kinases Tie-2, vascular endothelial growth factor receptor 1 (VEGFR-1) and tumor necrosis factor (TNF) receptor. These data suggest a redundant role of Bmx in endothelial signal transduction.

Tec Family Kinases in Signal Transduction

Tec family kinases not only play a critical part in T-cell receptor or B-cell receptor signaling, but are also involved in cytokine receptor signaling as well as mast cell Fcepisilon receptor I (FcεRI) signaling. Here we will use Btk as an example to discuss the signaling functions of these kinases.

Each domain of Btk is essential for the function of the kinase, as suggested by mutation analysis in XLA patients. To date, over 400 Btk mutations (missense mutation, frameshift, truncation) from 556 XLA families have been reported and these mutations cover all the domains. However, missense mutations are found in each domain except the SH3 domain, suggesting that the SH3 domain may tolerate such alterations. Each domain has been shown to bind regulatory and/or effector proteins. In addition to PI(3,4,5)P$_3$, Btk PH domain can also bind protein kinase Cβ (PKCβ), Fas, F-actin and a transcription factor TFII-I. TH domain binds G protein subunits and Src family kinases as Lyn or Hck. SH3 domain interacts with proteins as Cbl, WASP, and Vav, while SH2 domain binds B-cell linker protein (BLNK) through interaction with phosphotyrosine. The catalytic domain has also been shown to bind G protein $\beta\gamma$-subunit, and the kinase activity is activated by this interaction. Although some of the interactions have been confirmed in cellular context, the physiological importance of many still needs to be carefully evaluated.

B-cell development and activation is a tightly regulated process and Btk plays an important role. It is involved in a number of signaling pathways that are activated when B cells are stimulated through the BCR, accessory molecules such as CD19 and CD38, or cytokine receptors such as IL-5R and IL-10R. In normal B cells, cross-linking of the BCR activates PI 3-kinase leading to the translocation of cytoplasmic Btk to the lipid raft on the plasma membrane. Src family tyrosine kinase Lyn, which is also activated upon BCR stimulation, phosphorylates Y551 in the Btk kinase domain and activates Btk kinase activity. These serial activations lead to the assembly of a B-cell signalosome and proteins like Btk, Lyn, PI 3-kinase, BLNK, PLCγ2, BCAP (B-cell adapter for PI 3-kinase), and PKC are among the players in the B-cell signalosome. Once activated, Btk transduces signals to a number of effectors, including phospholipase C gamma (PLCγ), calcium response, transcription factors such as TFII-I, genes involved in apoptosis (bcl-2 and bcl-xl), and cell cycle control (cyclins). Interestingly, the Btk-deficient and xid phenotype is phenocopied by the deficiency of PI 3-kinase regulatory subunit p85 or the catalytic subunit p110δ, BCAP, PKCβ, BLNK, or PLCγ2. These data indicate that these proteins are involved in the same signaling pathway and likely in the same B-cell signalosome (Figure 2).

There are several mechanisms to down-regulate Btk. Autophosphorylation at Y223 in the SH3 domain has been implicated to negatively regulate Btk function, and phosphorylation of S180 in the TH domain by PKCβ also serves to down-regulate the membrane association of Btk. Additionally, a recent study identified an inhibitor for Btk (IBtk). IBtk directly binds the PH domain of
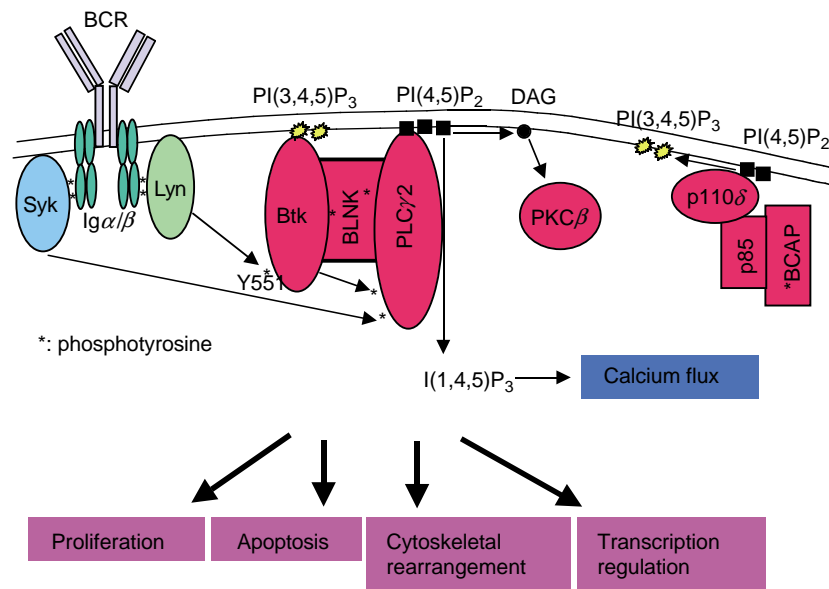
FIGURE 2   Signal transduction pathway after B cell receptor stimulation. Btk is activated through translocation to the plasma membrane through binding of $PI(3,4,5)P_3$ and phosphorylation at Y551 by Lyn. The B cell signalosome is assembled and downstream effectors will be sequentially activated. Proteins that have a xid-like phenotype when knocked out are shown in red. Adapted from Fruman, D. A., Satterthwaite, A. B., and Witte, O. N. (2000), Xid-like phenotypes: a B cell signalosome takes shape. *Immunity* **13**, 1–3.

Btk and down-regulates the kinase activity, calcium response and NF-κB-induced transcription.

In addition, Btk is involved in myeloid lineage cell signaling, including macrophages and mast cells. Although mast-cell development is normal in xid mice, Btk is phosphorylated and translocated to the plasma membrane after cross-linking of FcεRI in mast cells.

Future Directions

Extensive studies have been carried out on Tec family kinases since the 1990s, yet the detailed function of these kinases is far from resolved. In the coming years, further efforts will be put on the participation of these kinases in signalosome formation, the components of signalosome, and the regulation of the signal transduction. In addition, the complete structure of the proteins in this family will help the understanding of intramolecular domain–domain interaction as well as protein–protein interaction, and shed light on the regulation of kinase activity and function.

SEE ALSO THE FOLLOWING ARTICLES

c-*fes* Proto-Oncogene • Epidermal Growth Factor Receptor Family • Phosphatidylinositol Bisphosphate and Trisphosphate • Phosphatidylinositol-3-Phosphate

GLOSSARY

Bruton's tyrosine kinase (Btk) A member of the Tec/Btk tyrosine kinase family. Mutations of Btk cause X-linked agammaglobulinemia (XLA) in human and x-linked immunodeficiency (xid) in mouse. This protein was named after Dr. Ogden Bruton, who described XLA in 1950s.

signalosome Multiprotein signaling complex formed in the cell that is regulated both spatially and temporally.

Tec/Btk tyrosine kinase family A family of the non-receptor tyrosine kinases including Btk, Tec, Itk, Txk, and Bmx, which play important signaling roles primarily in hematopoietic cells.

X-linked agammaglobulinemia (XLA) A human genetic disease caused by the mutations in a proteins tyrosine kinase Btk. Patients lack mature B cells in the periphery and suffer from recurrent infections.

X-linked immunodeficiency (xid) An immunodeficiency found in CBA/N mice that is caused by a spontaneous point mutation in Btk (R28C). Xid mice have reduced mature B cells and defective B-cell response.

FURTHER READING

Fruman, D. A., Satterthwaite, A. B., and Witte, O. N. (2000). Xid-like phenotypes: a B cell signalosome takes shape. *Immunity* **13**, 1–3.

Lewis, C. M., Broussard, C., Czar, M. J., and Schwartzberg, P. L. (2001). Tec kinases: Modulators of lymphocyte signaling and development. *Curr. Opin. Immunol.* **13**, 317–325.

Miller, A. T., and Berg, L. J. (2002). New insights into the regulation and functions of Tec family tyrosine kinases in the immune system. *Curr. Opin. Immunol.* **14**, 331–340.

Pillai, S., and Moran, S. T. (2002). Tec kinase pathways in lymphocyte development and transformation. *Biochim. Biophys. Acta* **1602**, 162–167.

Qiu, Y., and Kung, H. J. (2000). Signaling network of the Btk family kinases. *Oncogene* **19**, 5651–5661.

Satterthwaite, A. B., and Witte, O. N. (2000). The role of Bruton's tyrosine kinase in B-cell development and function: A genetic perspective. *Immunol. Rev.* **175**, 120–127.

Smith, C. I., Islam, T. C., Mattsson, P. T., Mohamed, A. J., Nore, B. F., and Vihinen, M. (2001). The Tec family of cytoplasmic tyrosine kinases: Mammalian Btk, Bmx, Itk, Tec, Txk, and homologs in other species. *Bioessays* **23**, 436–446.

Vihinen, M., Kwan, S. P., Lester, T., Ochs, H. D., Resnick, I., Valiaho, J., Conley, M. E., and Smith, C. I. (1999). Mutations of the human BTK gene coding for bruton tyrosine kinase in X-linked agamma-globulinemia. *Hum. Mutat.* **13**, 280–285.

BIOGRAPHY

Owen N. Witte is an Investigator of Howard Hughes Medical Institute and a Professor at the University of California Los Angeles. His primary research interest is signal transduction in hematopoietic cells and related diseases. Dr. Witte is a pioneer in the study of tyrosine kinases. His group and others were the first to report Btk mutations cause Xid and XLA (Rawlings DJ *et al.*, Science 1993; Tsukada S *et al.*, Cell, 1993; Thomas JD *et al.*, Science 1993; Vetrie D *et al.*, Nature 1993).

Shuling Guo received her Ph.D. from Duke University and is currently a postdoctoral fellow in Dr. Witte's laboratory.

Telomeres: Maintenance and Replication

Alessandro Bianchi and David Shore
University of Geneva, Geneva, Switzerland

All linear eukaryotic chromosomes terminate in a specialized nucleoprotein structure, the telomere. Telomeres perform at least two essential functions: they provide a protective "cap" on chromosome ends that prevents their degradation or deleterious fusion, and they provide a special mechanism for replicating the DNA at chromosome ends. In most organisms, telomeres are composed of a tandem array of simple DNA repeats to which a large set of protein factors is bound. The telomeric DNA repeats are generated by a specialized reverse transcriptase, called telomerase that uses an endogenous RNA template. Defects in the maintenance of telomeric DNA, for example through inactivation of the telomerase enzyme, lead to the progressive loss of telomeric repeats and their bound factors, which eventually causes a catastrophic "uncapping" of telomeres that results in fusion of chromosome ends. The telomere complex is thus essential to ensure genome stability.

The Telomeric Complex, a Specialized Nucleoprotein Structure at Chromosome Ends

Pioneering studies in the fruit fly and in maize, carried out respectively by H. Muller and B. McClintock in the early 20th century, first identified the telomere as a special genetic entity that protects chromosome ends from degradation and fusion, a property absent in DNA ends that result from random chromosomal breakage. A wealth of subsequent genetic and biochemical studies have led to the understanding that telomeres are specialized nucleoprotein complexes composed of a large number of protein factors assembled onto telomeric DNA.

TELOMERIC DNA

In most eukaryotes, the terminal DNA sequences at each chromosome are composed of arrays of variable length of simple tandem repeats, as first recognized by Blackburn and Gall in the ciliate *Tetrahymena termophila*. These telomeric repeats are typically 5–8 nucleotides long, but can be up to 25 bp (Figure 1) and usually have a higher G content in the DNA strand that runs with a $5'$ to $3'$ polarity towards the end of the chromosome. In all species, the majority of the telomeric repeats appear to be in double-stranded form and their total length varies from a few nucleotides in some species of ciliates (28 bp of duplex telomeric repeats in *Euplotes*, for example) to tens of thousands of bp in some strains of laboratory mice. In both budding and fission yeasts, telomeres are ~300 bp in length, whereas in human cells telomeres are generally 6–12 kbp long. With the exception of several ciliate species in which the length of telomeric DNA is fixed, in most species the length of each telomere is variable from cell to cell, individual to individual, and during the life of the organism. In many species, it has been shown that telomeres terminate in a single-stranded overhang of the G-rich strand. This overhang can range from only a few nucleotides long (in some ciliates) to about 100–200 bases in mammals, and is believed to represent a general feature of telomeres. G-rich telomeric overhangs *in vitro* can adopt a variety of intra- and intermolecular paired structures involving most commonly four strands interacting through Hoogsteen-type G–G base-pairing (G-quartets), but the *in vivo* occurrence and significance of such structures remains unclear. In mammals, some ciliates, and Trypanosomes, telomeric DNA appears to fold back in a structure, the t-loop, where the single-stranded overhang is tucked into the duplex portion of the telomeric tract, by base pairing with the C-rich strand and thus displacing a short portion of the G-strand.

In most organisms, the short telomeric repeats described above are preceded by less conserved DNA elements, called subtelomeric repeats or telomere-associated sequences (TAS). These elements are of variable lengths but are generally at least a few hundred base pairs long and have been described in most studied organisms. In a particular species several classes of these repeats can be present, and elements of each class are associated with a particular chromosome with great variation in number and organization. The assembly of these subtelomeric repeats is highly dynamic and subject

Telomeric repeats	
Vertebrates	
Homo, Mus, Rattus, Gallus	TTAGGG
Arthropodes	
Insects	
Bombix mori	TTAGG
Nematodes	
Ascaris	TTAGGC
Caenorhabditis elegans	TTAGGC
Plants	
Arabidopsis thaliana	TTTAGGG
Green algae	
Chlamydomonas	TTTTAGGG
Fungi	
Ascomycota	
Hemiascomycetes	
Saccharomyces cerevisiae	T(G)$_{1-3}$
Kluyveromyces lactis	ACGGATTTGATTAGGTATGTGGTGT
Archeascomycetes	
Schizosaccharomyces pombe	TTACAGG(G)$_{0-4}$
Euascomycetes	
Neurospora	TTAGGG
Protists	
Alveolata	
Plasmodium	TT[T/C]AGGG
Ciliates	
Tetrahymena	TTGGGG
Oxytricha	TTTTGGGG
Euplotes	TTTTGGGG
Diplomonadida	
Giardia	TAGGG
Euglenozoa	
Trypanosoma	TTAGGG

FIGURE 1 Representative list of telomeric repeats in several eukaryotic organisms, including some of the most extensively studied ones with regard to telomere biology.

to active recombination, resulting in a large variation of subtelomeric regions between strains of yeast or between human individuals. TAS do not appear to carry out an essential telomeric function, as both human and yeast chromosomes that are devoid of them are replicated and segregated properly both through mitosis and meiosis.

TELOMERASE

DNA polymerases replicate DNA uniquely in a 5′ to 3′ direction. As a consequence, as the DNA replication "bubble" moves along the DNA, the two strands are replicated differently: one strand (the "leading" strand) is replicated in the same direction as the movement of the polymerase, while the other strand (the "lagging" strand) is replicated backwards in small installments, each one initiating from an RNA primer laid down by the primase enzyme (Figure 2A). Due to the requirement for an initiating RNA primer, which is later removed, replication of linear DNA molecules by the conventional DNA replication machinery would result in a terminal gap in the lagging strand end at the telomere (Figure 2A). The presence of overhangs at telomeres could in principle

mask this problem at the lagging strand, but a terminal gap would then occur at the strand replicated by leading strand synthesis after processing of the end to generate the single-stranded overhang (Figure 2B). Thus, in the absence of a specialized mechanism to replicate the ends, loss of genetic material is expected to occur at each cell division. The most general solution to this problem in eukaryotes is represented by the specialized replication of telomeric repeats by the telomerase enzyme.

Telomerase was first identified by Greider and Blackburn in the ciliate *Tetrahymena termophila* based on its ability to add telomeric repeats in a terminal transferase-like fashion to a telomeric DNA primer in an *in vitro* reaction. The enzyme is in fact a ribonucleoprotein, composed of a protein and an RNA moiety, both of which are required for catalytic function (Figure 2C). Isolation of the protein component from the ciliate *Euplotes* by Lingner and Cech in 1997 has revealed that telomerase shares extensive homology with reverse transcriptases. Thus telomerase is a specialized reverse transcriptase that utilizes an endogenous RNA template for the synthesis of telomeric repeats. The enzyme appears to be capable of synthesizing several repeats on
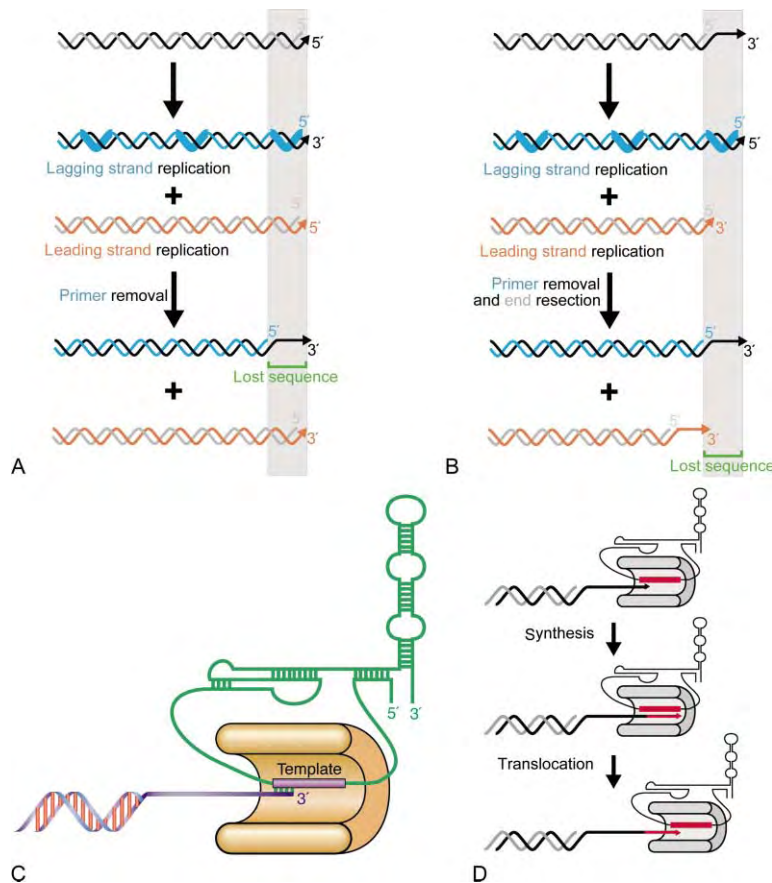
FIGURE 2   (A) Schematic view of the end replication problem of a linear DNA molecule. The RNA primers laid down by the primase enzyme for the synthesis of the so-called 'lagging' strand are indicated by thicker blue lines. After removal of the primers, replication of a double-stranded DNA molecule would result in a single-stranded terminal gap in the DNA strand replicated by lagging strand synthesis. (B) Schematic view of the end replication problem at a telomere terminating in a single-stranded overhang. Exonucleolytic resection of the ends is hypothesized to generate the overhangs after replication. In this scenario, a terminal gap would be created in the strand replicated by leading strand synthesis. (C) Representation of the structure of the core telomerase enzyme. The protein, based on its homology to known viral reverse transcriptases, is expected to be folded in a structure with a central cleft (palm) bearing the active site and flanked by two domains (thumb and fingers). The template region of the RNA is placed in this cleft, and the secondary structure of the RNA represented here is derived from ciliated Protozoa species and appears to be largely conserved in higher Eukaryotes. (D) Representation of the processive mechanism of action of telomerase, where a round of synthesis of telomeric repeats is followed by translocation of the enzyme on the DNA and a subsequent new round of synthesis.

a particular substrate through repeated steps of elongation and translocation (Figure 2D) and, like reverse transcriptases, appears to be a dimer.

TELOMERE PROTEINS

Telomeric repeats serve as binding sites for telomeric DNA-binding proteins, which act as a scaffold for the recruitment of additional protein factors to telomeres, including telomerase (Figure 3). The analysis of telomeric proteins in a wide range of organisms has revealed important similarities, though significant differences have also emerged.

The major double-stranded DNA-binding activity in budding yeast is Rap1, a protein also involved in transcriptional regulation at many other nontelomeric chromosomal sites. *S. cerevisiae* Rap1 binds directly to yeast telomeric repeats through two Myb-like DNA-binding domains. Rap1 is instrumental in recruiting to yeast telomeres a set of proteins involved in the regulation of telomere length (the Rap1 interacting factors 1 and 2, Rif1 and Rif2) and in the assembly of a complex (which includes the proteins Sir2, 3, and 4) that results in the transcriptional repression of genes positioned near telomeres (telomere position effect, or TPE). In contrast, mammalian and fission yeast Rap1 appears to be recruited to telomeres indirectly, through the binding to a different class of double-stranded DNA-binding factors, TRF1 and TRF2 in vertebrates and Taz1 in fission yeast. These proteins bind as dimers and contain a single Myb repeat that is essential for DNA binding. They share among themselves a large domain (TRFH, for TRF homology) that is responsible for homodimerization. Rap1 is recruited to mammalian and fission yeast telomeres through binding with TRF2 and Taz1, respectively. Human TRP2 also recruits the MRX
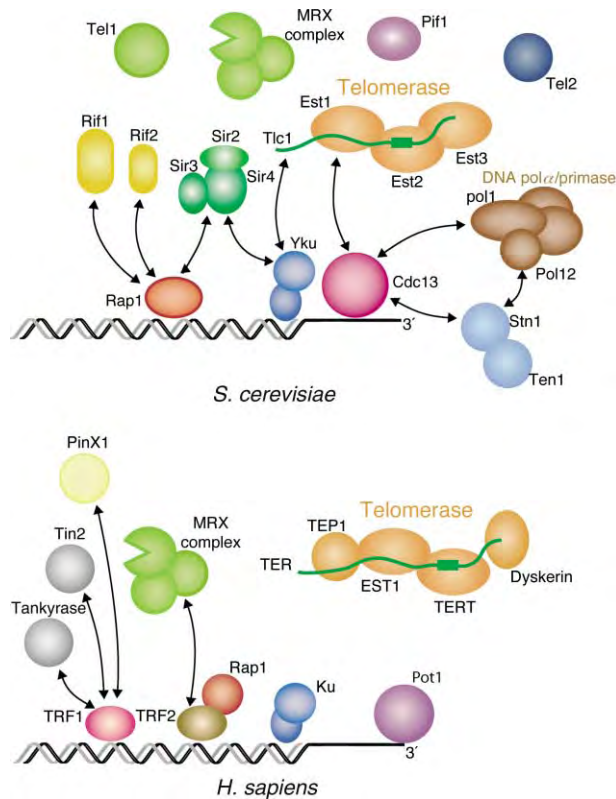
FIGURE 3 Representation of some of the characterized components of the budding yeast and human telomeric complexes. All proteins depicted have either been shown to affect telomere behavior and/or to interact with other telomere components.

complex to telomeres. TRF1 in mammals also is responsible for the recruitment of protein factors, including the telomere length regulators Tin2 and tankyrase, the telomerase inhibitor PinX1, and the single-stranded DNA-binding protein Pot1.

In addition to double-stranded DNA-binding activities, single-stranded DNA-binding factors also appear to be an essential and general feature of telomeres. Overhang-binding proteins were initially characterized in ciliates, where they form tenacious salt-resistant complexes with the single-stranded DNA termini. In budding yeast, the Cdc13 protein is bound to the single stranded overhangs created in S-phase in this organism and it carries out essential functions in telomere replication, through interaction with the telomerase and DNA replication complex, and in telomere protection, through interactions with the Stn1 and Ten1 proteins. In fission yeast and mammals, Pot1, an orthologue of the ciliate end factors, is involved in telomere protection in fission yeast and in telomere length regulation in mammals. Although Cdc13 does not share sequence homology with the ciliate, fission yeast and mammalian end factors, all these proteins bind DNA through an oligonucleotide/oligosaccharide-binding (OB) domain.

An additional factor, present at telomeres from yeast to humans, is the DNA repair protein Ku, which might bind to telomeres through its nonsequence-specific DNA end-binding activity. However protein–protein interactions between Ku and Sir4, and between Ku and TRF2 have been described in yeast and humans, respectively. Finally, the highly conserved telomere length regulator Tel2, has been shown (for the yeast protein) to be able to bind both single- and double-stranded telomeric repeats *in vitro*.

Telomere Replication

Telomeres are normally maintained through the action of the telomerase enzyme. Although telomerase activity in an *in vitro* assay requires only the reverse transcriptase-like protein motif and the RNA component, it is now clear that the action of the enzyme at telomeres is subject to complex regulation in *cis* at each individual telomere.

POSITIVE REGULATION OF THE TELOMERASE ENZYME

In budding yeast, a set of genes is required for telomerase activity *in vivo* in addition to the catalytic core components Est2 and Tlc1. These include Est1, which appears to bridge an interaction between telomerase and Cdc13 required to either recruit or activate the enzyme at the telomere, and telomerase-associated Est3, whose mode of action is unclear. The combined action of the DNA damage checkpoint kinases Tel1 and Mec1 is also necessary for telomerase activity in budding yeast, possibly by promoting a structural transition in the telomere complex from a "closed" to an "open" state that is accessible to telomerase. A similar requirement exists in fission yeast for the Tel1 and Mec1 homologues Tel1 and Rad3. In mammalian cells dyskerin binds the telomerase RNA and appears to affect *in vivo* telomerase activity by influencing accumulation of the RNA. Significantly, budding yeast telomerase action is strictly limited to the S-phase of the cell cycle and requires active DNA replication.

A NEGATIVE REGULATORY LOOP STABILIZES TELOMERE LENGTH

In cells expressing telomerase, telomere length is maintained around a constant average value that is species specific, as mentioned above. Evidence from a variety of organisms, including the budding and fission yeasts, and human cells, has led to the proposal that telomere length is controlled by a protein-counting mechanism. In this
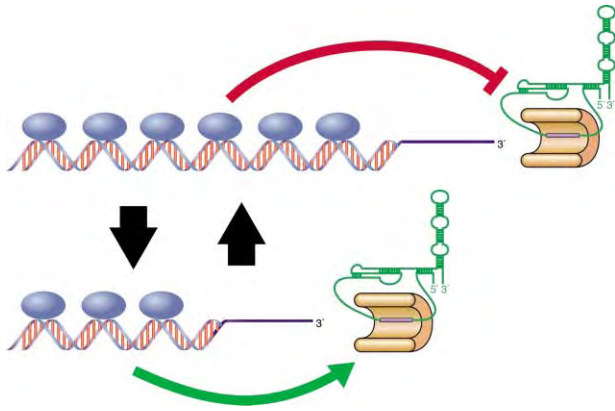
FIGURE 4 Depiction of the protein-counting mechanism for telomere length regulation. In this model, longer telomeres have higher numbers of inhibitory factors bound, resulting in repression of telomerase activity in *cis*. Diminished action by telomerase results in telomere shortening, loss of the bound inhibitor, and telomerase activation. In this manner telomeres are maintained in a state of dynamic equilibrium. The organism in which our understanding of the telomeric complex is the most complete, both with regard to the identification of protein components and to their function, is the budding yeast *Saccharomyces cerevisiae* (Figure 1).

scenario, the number of double-stranded DNA-binding factors that are associated with the telomere at any given time is somehow sensed by the telomere length regulation machinery. An increase in the number of these negative regulators results in the inhibition of telomerase activity in *cis* at the telomere (Figure 4). The inhibitory function of the DNA-binding factors (Rap1 in budding yeast, Taz1 in fission yeast, and TRF1 and 2 in mammals) is apparently mediated by a set of interacting factors (Rif1 and 2 in budding yeast, Pot1 in humans). The precise mechanism of the proposed *cis* inhibition of telomerase is unknown, but might, at least in some organisms, involve formation of the t-loop structure described previously.

TELOMERASE-INDEPENDENT MAINTENANCE OF TELOMERES

In the absence of telomerase, alternative pathways for the maintenance of telomeric repeats exists that are based on homologous recombination. In budding yeast, telomere-associated sequences can contribute to this recombinational process. In mammals, several tumor cell lines have been shown to maintain their telomeres without detectable telomerase activity through a mechanism named ALT, for alternative lengthening of telomeres. Finally, the fruit fly *Drosophila melanogaster* dispenses altogether with short telomeric repeats and forms telomeres through the repeated insertion of telomere-specific transposable elements (HeT-A and TART) onto the ends of its chromosomes.

Telomerase Repression in Somatic Cells: Implications for Replicative Senescence and Maintenance of Genome Integrity

Mice engineered to lack the gene encoding telomerase RNA (mTRT) exhibit a gradual (~sixth generation) loss of fertility, defects in cell proliferation, and an increase in the frequency of end-to-end chromosome fusions. These phenotypes result from telomere repeat erosion and consequent loss of telomere function. Since active telomerase enzyme is primarily detected in mouse germline cells, these data indicate that telomerase-dependent telomere length "resetting" in the germline is essential for long-term survival in the mouse, and by inference in all mammals. In humans, the gene encoding the catalytic subunit of telomerase is also repressed in most somatic cells, causing telomeres to undergo progressive shortening with each cell division. This process leads to cellular senescence in culture and can be reversed, at least in some primary cell lines, by the ectopic expression of an active telomerase gene in the senescing cells. These observations in cultured primary cells have led some researchers to propose that telomerase repression in humans, and consequent telomere erosion, acts as a kind of "mitotic clock" that determines aging of the whole organism.

An attractive alternative view is that the telomere "mitotic clock" in humans might instead represent a critical barrier against malignant transformation. Recent experiments suggest that either normal telomere erosion, or uncapping caused by mutational alteration of key telomere proteins, will induce a DNA damage checkpoint and apoptosis (programmed cell death). This mechanism may place a limit on continued cell proliferation, which is a prerequisite for tumor formation. However, telomere uncapping also leads to telomere-telomere fusions and rampant genomic instability, which is thought to be a driving force in oncogenesis. Telomere dysfunction might therefore represent an important step in the early stages of tumorigenesis, particularly in cells with compromised checkpoint (DNA damage surveillance) function. The proper regulation of telomerase and telomere capping thus appear to be remarkably critical cellular functions.

SEE ALSO THE FOLLOWING ARTICLES

Chromosome Organization and Structure, Overview • DNA Replication: Eukaryotic Origins and the Origin Recognition Complex • Homologous Recombination in Meiosis • Meiosis • Mitosis

Glossary

telomerase A nuclear-encoded ribonucleoprotein complex homologous to viral reverse transcriptases that adds single-stranded telomeric repeats to the $3'$ ends of duplex DNA using its own RNA template; responsible for telomere maintenance in most eukaryotic organisms.

telomere The DNA–protein complex at the extreme ends of linear eukaryotic chromosomes that orchestrates telomere replication and capping.

telomere associated sequence (TAS) The highly variable, often repeated DNA sequences found immediately internal to the terminal simple telomere repeat sequences.

telomere capping Refers to the ability of the telomere DNA–protein complex to prevent chromosome end degradation or fusion (DNA joining) reactions.

telomeric repeats Simple, short (usually 5–8, but as long as 25 nucleotides) DNA repeat sequences (T_2AG_3 in vertebrates and minor variants in many other eukaryotes) that constitute the DNA component of telomeres. Synthesized by the telomerase enzyme, with the TG-rich sequence extending towards the $3'$ DNA end.

Further Reading

Blackburn, E. H., and Greider, C. W. (eds.) (1995). *Telomeres*. Vol 29, Cold Spring Harbor Monographs, Cold Spring Harbor Press, Plainview, NY.

Chakhparonian, M., and Wellinger, R. J. (2003). Telomere maintenance and DNA replication: How closely are these two connected? *Trends Genet.* **19**(8), 439–446.

de Lange, T. (2002). Protection of mammalian telomeres. *Oncogene* **21**(4), 532–540.

Evans, S. K., and Lundblad, V. (2000). Positive and negative regulation of telomerase access to the telomere. *J. Cell Sci.* **113**(Pt), 3357–3364.

Kass-Eisler, A., and Greider, C. W. (2000). Recombination in telomere-length maintenance. *Trends Biochem. Sci.* **25**(4), 200–204.

Lingner, J., and Cech, T. R. (1998). Telomerase and chromosome end maintenance. *Curr. Opin. Genet. Dev.* **8**(2), 226–232.

Lundblad, V. (2002). Telomere maintenance without telomerase. *Oncogene* **21**(4), 522–531.

Maser, R. S., and DePinho, R. A. (2002). Connecting chromosomes, crisis, and cancer. *Science* **297**(5581), 565–569.

McEachern, M. J., Krauskopf, A., and Blackburn, E. H. (2000). Telomeres and their control. *Annu. Rev. Genet.* **34**, 331–358.

Wright, W. E., and Shay, J. W. (2001). Cellular senescence as a tumor-protection mechanism: the essential role of counting. *Curr. Opin. Genet. Dev.* **11**(1), 98–103.

Zakian, V. A. (1996). Structure, function, and replication of Saccharomyces cerevisiae telomeres. *Annu. Rev. Genet.* **30**, 141–172.

Biography

David Shore is a Professor in the Department of Molecular Biology at the University of Geneva (Switzerland) and member of the NCCR program "Frontiers in Genetics." His principal research interests are in gene regulation, chromatin structure, and telomere biology. He holds a Ph.D. from the Department of Biochemistry at Stanford University Medical School and did postdoctoral work at the MRC Laboratory of Molecular Biology in Cambridge (UK).

Alessandro Bianchi received his Ph.D. at the Rockefeller University (New York) and is currently a postdoctoral fellow at the University of Geneva.

Thyroid-Stimulating Hormone/Luteinizing Hormone/Follicle-Stimulating Hormone Receptors

Deborah L. Segaloff, Dario Mizrachi and Mario Ascoli
The University of Iowa, Iowa City, Iowa, USA

The thyroid-stimulating hormone receptor (TSHR), luteinizing hormone receptor (LHR), and follicle-stimulating hormone receptor (FSHR) are collectively referred to as the glycoprotein hormone receptors because they bind the structurally similar glycoprotein hormones. The glycoprotein hormones consist of the pituitary hormones thyroid-stimulating hormone (thyrotropin, TSH), luteinizing hormone (lutropin, LH), and follicle-stimulating hormone (follitropin, FSH) as well as the placental hormone chorionic gonadotropin (choriogonadotropin, CG). They are each composed of two dissimilar subunits (α and β) that are noncovalently associated. Within a given species the α-subunit is identical, and the β-subunits are distinct but homologous. Due to the nearly identical nature of LH and CG, the LHR binds either LH or CG. However, the FSHR binds only FSH and the TSHR binds only TSH. Because the LHR and FSHR are localized primarily to the gonads, these two receptors are also referred to as the gonadotropin receptors. The glycoprotein hormone receptors are members of the superfamily of G protein-coupled receptors (GPCRs) and, despite their different physiological roles, share a similar structural organization and mechanism of action. In addition, the glycoprotein hormone receptors are also members of a subfamily of GPCRs known as the leucine-rich repeat-containing GPCRs (LGR). This subfamily of GPCRs is characterized by the presence of a large extracellular domain composed of several leucine-rich repeats.

Expression and Physiological Roles of the Glycoprotein Hormone Receptors

THE THYROID-STIMULATING HORMONE RECEPTOR

TSHR is expressed primarily in the follicular cells of the thyroid. In more recent years, it has also been detected in lymphocytes, thymus, pituitary, testis, kidney, brain, adipose tissue and fibroblasts, heart, and bone.

In response to thyroid-stimulating hormone (TSH), the TSH receptor (TSHR) stimulates the synthesis and secretion of thyroid hormone by the follicular cells of the thyroid. Autoantibodies to the TSHR that are stimulatory are present in individuals with Grave's disease and stimulate the TSHR causing excessive and unregulated secretion of thyroid hormone. Conversely, inhibitory autoantibodies to the TSHR are found in individuals with Hashimoto's thyroiditis. These antibodies bind to the TSHR and inhibit the binding of TSH, thus blocking the synthesis and secretion of thyroid hormones.

The TSHR present in the adipose and connective tissue of the orbit of the eye is thought to play a role in the development of exopthalmos found in individuals with Grave's disease. The physiological roles of the TSHR in other nonthyroid tissues is not yet known.

THE LUTEINIZING HORMONE RECEPTOR

The luteinizing hormone receptor (LHR) is expressed primarily in the ovaries and the testes. Within the ovary, the LHR is present on theca and interstitial cells and on mature granulosa cells. After ovulation, the granulosa and theca cells of the ruptured follicle differentiate into the luteal cells, and express higher levels of LHR. In males, the LHR is expressed on the Leydig cells. In both males and females, nongonadal expression of the LHR has also been reported in the reproductive tract and many other tissues.

In the nonpregnant postpubertal female, ovarian theca cells respond to luteinizing hormone (LH) with increased synthesis of androgens, which are used as substrates for estrogen production by follicle-stimulating hormone (FSH)-stimulated granulosa cells. Mature granulosa cells respond to LH with increased

progesterone production. In response to the ovulatory surge of LH, the ovarian LHR receptors mediate ovulation. If pregnancy ensues, then the LHR of the corpus luteum responds to placental chorionic gonadotropin (CG) with increased progesterone synthesis. As such, the LHR is essential for the maintenance of pregnancy, particularly during the first trimester.

In males, the testicular LHR plays an important physiological role during fetal development. Maternal hCG stimulates the fetal Leydig cells to synthesize testosterone, which is required for the differentiation of the external male genitalia and for the descent of the testes into the scrotum. Testosterone levels in boys decrease after birth (due to the absence of maternal LH) and the LHR remains unstimulated until the time of puberty. After puberty the testicular LHR responds to pituitary LH with increased testosterone synthesis.

The physiological roles of nongonadal LHR are not known. It should be pointed out, though, that the only functional consequences of loss-of-function or gain-of-function mutations of the LHR described in males or females have been restricted to abnormalities in reproductive physiology.

THE FOLLICLE-STIMULATING HORMONE RECEPTOR

The follicle-stimulating hormone receptor (FSHR) is expressed in ovarian granulosa cells and in the Sertoli cells within the seminiferous tubules of the testes. In post-pubertal females, the FSHR mediates follicular growth and controls estrogen synthesis. In post-pubertal males, pituitary FSH facilitates spermatogenesis by stimulating the Sertoli cells that are adjacent to the developing sperm to synthesize and secrete components needed for spermatogenesis. Although it is accepted that optimal spermatogenesis requires the actions of FSH, it is controversial as to whether FSH is essential for this process.

Structural Organization of the Glycoprotein Hormone Receptors

SERPENTINE REGIONS

The glycoprotein hormone receptors are members of the rhodopsin-like family of GPCRs. As such, they all contain the seven membrane-spanning regions prototypical of the superfamily of GPCRs. Residues that are conserved within the transmembrane (TM) domains of the rhodopsin-like GPCRs are also generally conserved in the glycoprotein hormone receptors. The recent solving of a high resolution crystal structure of rhodopsin has provided a template for creating models of the TM regions of the glycoprotein hormone

receptors and permitting investigators to envision the interhelical interactions maintaining the receptors in their inactive states. The crystal structure of rhodopsin also revealed the presence of an eighth α-helix that extends from TM7 and lies parallel to the inner face of the plasma membrane. In rhodopsin, helix 8 extends until a cysteine residue that is palimitoylated and serves to anchor the helix to the plasma membrane. This cysteine is conserved as a single residue or as a pair in the glycoprotein hormone receptors. An alignment of the human TSHR, LHR, and FSHR is shown in Figure 1. The amino acid identity is greatest between the glycoprotein hormone receptors in the transmembrane regions and helix VIII.

EXTRACELLULAR DOMAINS

A unique feature to the glycoprotein hormone receptors is their relatively large (i.e., 300–400 amino acids) extracellular domains. This is the receptor domain that is responsible for the selective recognition and high-affinity binding of each of the glycoprotein hormones. The extracellular domains are N-glycosylated and a fully conserved tyrosine residue has been shown to be sulfated in the TSHR. The TSHR has the largest extracellular domain which is clipped once the receptor is inserted at the plasma membrane. This proteolytic cleavage results in the formation of an α-subunit containing a portion of the N-terminal extracellular domain and a β-subunit containing the remaining of the N-terminal extracellular domain and the transmembrane and C-terminal domains. Although the α- and β-subunits are initially bound by disulfide bonds, these are reduced and the α-subunit is released from the membrane bound β-subunit.

The extracellular domains of the glycoprotein hormone receptors can be subdivided into a short, N-terminal cysteine-rich region which is followed by nine leucine rich repeats (LRR) and a C-terminal cysteine-rich region. LRR motifs are found in a variety of proteins and are composed of 20–30 amino acids. Based on the known three-dimensional structure of LRR motifs present in the ribonuclease inhibitor, each LRR is proposed to be formed by a β-strand and an α-helix joined by short loops and positioned in a nearly antiparallel orientation. Tandem arrays of these units are believed to form a horseshoe-like structure with consecutive β-strands forming a parallel β-sheet at the concave surface of the horseshoe. By analogy with the known structure of the ribonuclease–ribonuclease inhibitor complex, it is assumed that hormone binding occurs mostly through contact points with the β-sheets present at the concave surface.

Homology cloning and data mining have now uncovered additional GPCRs with large extracellular domain containing a variable number of LRRs. Four of these (designated LGR4–8) are found in mammals.

Signal peptide

hTSHR	1	M R P A D L L Q L V L L L D L P R D L G G - - -	21
hFSHR	1	M A L L L V S L L A F L S L G S G - - - - - - -	17
hLHR	1	M K Q R F S A L Q L L K L L L L L Q P P L P R A	24

N-flanking Cysteine-rich sequence

hTSHR	22	- - - M G **C** S S P P **C** E **C** H Q E E D - - F R V T **C** K D I Q R I P S L P P S T Q T - - -	56
hFSHR	18	- - - - - **C** H H R I **C** H **C** S N R V F L - - - - - **C** Q E S K V T E I P S D L P R N A I E	50
hLHR	25	L R E A L **C** P - E P **C** N **C** V P D G - - A L R - - **C** P G P T A G L T R - - - - - - - -	53

Leucine-rich repeats

hTSHR	57	**L** K L I E T H L R T I P S H A F S N L P N I S R I Y V S I - D V T **L** Q Q L E S H S F Y N **L** S K V T H I E I R N T R N **L** T	115
hFSHR	51	**L** R F V L T K L R V I Q K G A F S G F G D L E K I E I S Q N D V - **L** E V I E A D V F S N **L** P K L H E I R I E K A N N **L** L	109
hLHR	54	**L** S L A Y L P V K V I P S Q A F R G L N E V I K I E I S Q I D S - **L** E R I E A N A F D N **L** L N L S E I L I Q N T K N **L** R	112

hTSHR	116	Y I D P D A L K E **L** P L **L** K F L G I F N T G L K M F P D L T K V Y S T D I F F I **L** E I T D N P Y M - T S I P V N A F Q G	174
hFSHR	110	Y I N P E A F Q N **L** P N **L** Q Y L L I S N T G I K H L P D - V H K I H S L Q K V **L** **L** D I Q D N I N I H T - I E R N S F V G	167
hLHR	113	Y I E P G A F I N **L** P R **L** K Y L S I C N T G I R K F P D V T K V F S S E S N F I **L** E I C D N L H I T T - I P G N A F Q G	171

hTSHR	175	**L** C N E T L T **L** K **L** Y N N G F T S V Q G Y A F N G T K **L** D A V Y **L** N K N K Y **L** T V I D K D A F G G V Y S G P S L **L** D V S	234
hFSHR	168	**L** S F E S V I L W **L** N K N G I Q E I H N C A F N G T Q L D E L N **L** S D N N N **L** E E L P N D V F H G A S G - P V I L **L** D I S	226
hLHR	172	M N N E S V T **L** K **L** Y G N G F E E V Q S H A F N G T T **L** T S L E **L** K E N V H L E K M H N G A F R G A T G - P K T L **L** D I S	230

hTSHR	235	Q T S V T A **L** P S K G **L** E H **L** K E **L** I A R N T W T **L** K K **L** P L S L S F L H **L** T R	274
hFSHR	227	R T R I H S **L** P S Y G **L** E N **L** K K **L** R A R S T Y N **L** K K **L** P T L E K L V A L M E	266
hLHR	231	S T K L Q A **L** P S Y G **L** E S I Q R **L** I A T S S Y S **L** K K **L** P S R E T F V N **L** L E	270

C-flanking Cysteine-rich sequence

hTSHR	275	A D L S Y P S H **C C** A F K N Q K K I R G I L E S L - - - - - - - - - M C N E S S M Q S L R Q R K S V N A L N S P L H Q E	325
hFSHR	267	A S L T Y P S H **C C** A F A N W R R Q I S E L - H P I C N K S I L R Q E V D Y M T Q T R G Q R S S L A E - - - - - - - - -	316
hLHR	271	A T L T Y P S H **C C** A F R N L P T K E Q N F S H S - - - - - - - - - I S E N F S K Q C E S T V R K V S - - - - - - - - -	312

hTSHR	326	Y E E N L G D S I V G Y K E K S K F Q D T H N N A H Y Y V F F E E Q E D E I I G F G Q E L K N P Q E E T L Q A F D S H Y	385
hFSHR	317	- D N E S S Y S - - R G F D M T Y T E F	333
hLHR	313	- N K T L Y S S M L A E S E L S G W D Y	331

hTSHR	386	D Y T I **C** G D S E D M V **C** T P K S D E F N P **C** E D I	411
hFSHR	334	D Y D L **C** N E V V D V T **C** S P K P D A F N P **C** E D I	359
hLHR	332	E Y G F **C** - L P K T P R **C** A P E P D A F N P **C** E D I	356

FIGURE 1 Alignments of the amino acid sequences of the human TSHR, FSHR, and LHR. The deduced amino acid sequences for each of the three glycoprotein hormone receptors of human origin are shown.

Transmembrane region

```
                    I                                              II
hTSHR  412  M G Y K F L R I V V W F V S L L A L L G N V F V L L I L L T S H Y K L N V P R F L M C N L A F A D F C M G M Y L L L I A  471
hFSHR  360  M G Y N I L R V L I W F I S I L A I T G N I I V L V I L T T S Q Y K L T V P R F L M C N L A F A D L C I G I Y L L L I A  419
hLHR   357  M G Y D F L R V L I W L I N I L A I M G N M T V L F V L L T S R Y K L T V P R F L M C N L S F A D F C M G L Y L L L I A  416

                           III
hTSHR  472  S V D L Y T H S E Y Y N H A I D W Q T G P G C N T A G F F T V F A S E L S V Y T L T V I T L E R W Y A I T F A M R L D R  531
hFSHR  420  S V D I H T K S Q Y H N Y A I D W Q T G A G C D A A G F F T V F A S E L S V Y T L T A I T L E R W H T I T H A M Q L D C  479
hLHR   417  S V D S Q T K G Q Y Y N H A I D W Q T G S G C S T A G F F T V F A S E L S V Y T L T V I T L E R W H T I T Y A I H L D Q  476

                 IV                                                           V
hTSHR  532  K I R L R H A C A I M V G G W V C C F L L A L L P L V G I S S Y A K V S I C L P M D T E T P L A L A Y I V F V L T L N I  591
hFSHR  480  K V Q L R H A A S V M V M G W I F A F A A A L F P I F G I S S Y M K V S I C L P M D I D S P L S Q L Y V M S L L V L N V  539
hLHR   477  K L R L R H A I L I M L G G W L F S S L I A M L P I V G V S N Y M K V S I C F P M D V E T T L S Q V Y I L T I L I L N V  536

                                                     VI
hTSHR  592  V A F V I V C C C Y V K I Y I T V R N P Q Y N P G D K D T K I A K R M A V L I F T D F I C M A P I S F Y A L S A I L N K  651
hFSHR  540  L A F V V I C G C Y I H I Y L T V R N P N I V S S S S D T R I A K R M A M L I F T D F L C M A P I S F F A I S A S L K V  599
hLHR   537  V A F F I I C A C Y I K I Y F A V R N P E L M A T N K D T K I A K K M A I L I F T D F T C M A P I S F F A I S A S F K V  596

                 VII
hTSHR  652  P L I T V S N S K I L L V L F Y P L N S C A N P F L Y A I F T  682
hFSHR  600  P L I T V S K A K I L L V L F H P I N S C A N P F L Y A I F T  630
hLHR   597  P L I T V T N S K V L L V L F Y P I N S C A N P F L Y A I F T  627
```

Helix VIII

```
            Hx VIII
hTSHR  683  K A F Q R D V F I L L S K F G I C  699
hFSHR  631  K N F R R D F F I L L S K C G - C  646
hLHR   628  K T F Q R D F F L L L S K F G C C  644
```

C terminus

```
hTSHR  700  K R Q A Q A Y R G Q R V P P K N S T D I Q V Q K V T H D M R Q G L H N M E D V Y E L I E N S H L T P K K Q G Q I S E E Y  759
hFSHR  647  Y E M Q A Q I Y R T E T S S T V H N T H P R N G H C S S A P R V T N G S T Y I L V P L S H L A Q N - - - - - - - - - - -  695
hLHR   645  K R R A E L Y R R K D F S A Y T S N C K N G F T G S N K P S Q S T L K L S T L H C Q G T A L L D K T R Y T E C - - - - -  699

hTSHR  760  M Q T V L  764
hFSHR       - - - - -
hLHR        - - - - -
```

Conserved cysteine   Transmembrane domains   Intracellular loops   Extracellular loops
Conserved leucine   Highly conserved residues within rhodopsin-like GPCRs

FIGURE 1   (*continued*)

Interestingly, the ligands for LGR7 and LGR8, relaxin and/or Leydig cell insulin-like peptide, are structurally homologous to insulin rather than to the glycoprotein hormones.

Genomic Organization of the Glycoprotein Hormone Receptors

The genes for these three receptors have a similar structure. The LHR gene has 11 exons and the TSHR and FSHR genes each have 10 exons. The last exon codes for a small portion of the N-terminal extracellular domain and all of the serpentine region, including the C-terminal cytoplasmic tail. The other exons are small and they are spliced together to form most of the N-terminal extracellular domain.

The TSHR, LHR, and FSHR genes map to human chromosomes 14q31, 2p21, and 2p21–16, respectively.

Activation of the Glycoprotein Hormone Receptors

G PROTEINS AND CELLULAR PATHWAYS ACTIVATED BY THE GLYCOPROTEIN HORMONE RECEPTORS

The identity of the G proteins activated by the FSHR has not been directly examined. The TSHR has been shown to activate all four families of G proteins and the LHR activates at least three or perhaps all four families as well. All of these receptors activate the cAMP and inositol phosphate/diacylglycerol signaling cascades. For each of the glycoprotein hormones, the EC_{50} for cAMP accumulation is lower than the EC_{50} for inositol phosphate accumulation. Moreover, activation of the cAMP pathway is detectable even at low receptor densities, whereas activation of the inositol phosphate/diacylglycerol pathway is generally detectable only at high receptor densities.

It is generally agreed that cyclic AMP is the principal mediator of the actions of the glycoprotein hormones on the differentiated functions of their target cells. The functional consequences of the activation of the inositol phosphate/diacylglycerol pathway are not fully understood. In addition to controlling the differentiated functions of target cells, the glycoprotein hormones also control their proliferation. The three glycoprotein hormones can each activate the extracellular regulated kinase 1/2 cascade (ERK1/2) in their target cells and, at least in the case of the LHR and the TSHR, this activation seems to be mediated by cAMP.

MECHANISM OF ACTIVATION

Although it is clear that binding specificity is provided by the large extracellular domains of the receptors and the β-subunits of the hormones, little else is known about how this event leads to receptor activation. Three models have been proposed. In one model the common α-subunit of the glycoprotein hormones is thought to directly interact with and activate the transmembrane region of the receptor. In another model hormone binding is thought to induce a conformational change in the extracellular domain of the receptor, thus allowing it to interact with and activate the transmembrane region. The last model proposes that the extracellular domain constrains a constitutive activity intrinsic to the transmembrane domain and that hormone binding to the extracellular domain relieves this inhibition. This latter model has received strong experimental support when tested with the TSHR.

Naturally Occurring Mutations of the Glycoprotein Hormone Receptors

Both activating (gain-of-function) and inactivating (loss-of-function) mutations of the glycoprotein hormone receptor genes have been identified in patients with particular endocrine disorders. Activating mutations of these receptors results in the constitutive activation and unregulated elevation of intracellular cAMP in the absence of bound agonist. Most activating mutations have been found in the serpentine regions of the receptors; however, activating mutations of the hTSHR have also been found in the extracellular domain. Germ-line mutations resulting in constitutive activation of the glycoprotein hormone receptors are inherited in an autosomal dominant manner.

Loss-of-function mutations refer to a number of different types of mutations that ultimately cause decreased target cell responsiveness to hormone. These include mutations that impair coupling to G protein, hormone binding, and/or cell surface expression. The majority of loss-of-function mutations of the glycoprotein hormone receptors result in decreased cell surface expression of the mutant receptor due to the intracellular retention of the misfolded mutant. Therefore, loss-of-function mutations have been identified in all regions of the glycoprotein hormone receptors. Mutations of the glycoprotein hormone receptors resulting in a loss-of-function are inherited in an autosomal recessive manner.

TSHR

A large number of activating mutations of the hTSHR have been identified in individuals with autonomously functioning thyroid adenomas (somatic mutations) or autosomal dominant nonautoimmune hyperthyroidism (germ-line mutations).

A number of loss-of-function hTSHR mutations have been identified in individuals with nonautoimmune congenital hypothyroidism, where the degree of hypothyroidism correlates with the extent of target cell unresponsiveness to TSH.

LHR

Numerous activating mutations of the hLHR have been identified in boys with gonadotropin-independent precocious puberty (testotoxicosis). In these cases, the constitutively active hLHRs inappropriately stimulate testosterone production in Leydig cells under conditions of prepubertal levels of LH. Only one hLHR mutation, D578H, has been found to also cause Leydig cell adenomas. Unlike the other hLHR activating mutations, D578H was somatic rather than germ-line in origin. Although it was initially postulated that D578H may activate a pathway(s) different from those activated by the other constitutively active hLHRs, such differences have not yet been observed. Therefore, as with the hTHSR, tumor formation by D578H may be due to its somatic origin.

Many different loss-of-function hLHR mutations have been identified in genetic males with micropenis (partial loss-of-function) or pseudohermaphroditism (severe loss-of-function). In the latter case, the individual presents with female external genitalia but fails to undergo breast development or menstruation at the time of puberty. Female siblings with homozygous or compound heterozygous loss-of-function LHR mutations are infertile.

FSHR

Only one naturally occurring activating mutation of the hFSHR has been reported. This mutation was found in a hypophysectomized male receiving testosterone supplementation who exhibited normal spermatogenesis. Studies of this mutant in cell culture, however, have led to disparate results on whether or not it is truly activating. Surprisingly, activating FSHR mutations have not been found in granulosa cell tumors examined thus far. The paucity of naturally occurring activating mutations of the hFSHR may reflect the lack of a readily detectable pathophysiological state in males or females arising from a constitutively active hFSHR and/or to the relative resistance of the hFSHR to mutation-induced constitutive activity.

The first loss-of-function hFSHR mutation was identified in a population of Finnish women with primary amenorrhea due to ovarian failure. Cells expressing the recombinant form of this hFSHR mutant display no responsiveness to FSH. Surprisingly, men homozygous for this mutation were not infertile, although their fertility may have been impaired somewhat. Therefore, although FSH plays an important role in spermatogenesis, these observations raised ongoing debates regarding its absolute requirement. Since then, other loss-of-function mutations of the hFSHR have been identified in women with varying degrees of FSH resistance.

See Also the Following Articles

G Protein-Coupled Receptor Kinases and Arrestins • Glycoproteins, N-linked

Glossary

chorionic gonadotropin (CG)   A hormone secreted by the placenta that is nearly identical to the pituitary hormone LH and is, therefore, recognized by the LH receptor. Also, choriogonadotropin.

follicle-stimulating hormone (FSH)   A pituitary hormone recognized by the FSH receptor. Also, follitropin.

glycoprotein hormones   The structurally related hormones TSH, LH, CG, and FSH. The glycoprotein hormone receptors are the LH, FSH, and TSH receptors.

gonadotropins   LH, CG, and FSH; members of the glycoprotein hormone family that act upon the gonads. The gonadotropin receptors are the LH and FSH receptors.

G protein-coupled receptors (GPCRs)   A class of membrane receptors that contain seven transmembrane regions and mediate their actions through the activation of G proteins.

luteinizing hormone (LH)   A pituitary hormone recognized by the LH receptor. Also, lutropin.

thyroid-stimulating hormone (TSH)   A pituitary hormone recognized by the TSH receptor. Also, thyrotropin.

Further Reading

Ascoli, M., Fanelli, F., and Segaloff, D. L. (2002). The lutropin/choriogonadotropin receptor, a 2002 perspective. *Endocr. Rev.* **23**, 141–174.

Davies, T., Marians, R., and Latif, R. (2002). The TSH receptor reveals itself. *J. Clin. Invest.* **110**, 161–164.

Hsu, S. Y., and Hsueh, A. J. (2000). Discovering new hormones, receptors and signaling mediators in the genomic era. *Mol. Endocrinol.* **14**, 594–604.

Latronico, A., and Segaloff, D. L. (1999). Naturally occurring mutations of the luteinizing-hormone receptor: Lessons learned about reproductive physiology and G protein-coupled receptors. *Amer. J. Hum. Genet.* **65**, 949–958.

Paschke, R., and Ludgate, M. (1997). The thyrotropin receptor in thyroid diseases. *N. Engl. J. Med.* **337**, 1675–1681.

Rapoport, B., Chazenbalk, G. D., Jaume, J. C., and McLachlan, S. M. (1998). The thyrotropin (TSH) receptor: Interaction with TSH and autoantibodies. *Endocr. Rev.* **19**, 673–716.

Simoni, M., Gromoll, J., and Nieschlag, E. (1997). The follicle-stimulating hormone receptor: Biochemistry, molecular biology, physiology, and pathophysiology. *Endocr. Rev.* **18**, 739–773.

Themmen, A. P. N., and Huhtaniemi, I. T. (2000). Mutations of gonadotropins and gonadotropin receptors: Elucidating the physiology and pathophysiology of pituitary-gonadal function. *Endocr. Rev.* **21**, 551–583.

BIOGRAPHY

Deborah L. Segaloff is a Professor of Physiology and Biophysics and Mario Ascoli a Professor of Pharmacology in the Roy J. and Lucille A. Carver College of Medicine at the University of Iowa in Iowa City, Iowa Dr. Segaloff's laboratory cloned the cDNA for the rat LHR, the first glycoprotein hormone receptor to be cloned. Since then, Dr. Segaloff's and Dr. Ascoli's laboratories have performed extensive studies on the regulation and the mechanisms of activation and desensitization of the gonadotropin receptors.

Dario Mizrachi is a Postdoctoral Associate in Dr. Segaloff's laboratory. As a Lalor Fellow, Dr. Mizrachi developed homology models of the glycoprotein hormone receptors for use in predicting the structural arrangements of the receptors.

Tight Junctions

Shoichiro Tsukita
Kyoto University, Kyoto, Japan

The tight junction (TJ) or *zonula occludens* is one mode of cell-to-cell adhesion in vertebrate epithelial and endothelial cellular sheets, and is located at the most apical part of their lateral membranes. The existence of separate fluid compartments with different molecular compositions is of particular importance for the development and maintenance of multicellular organisms. These compartments are delineated by various cellular sheets, which function as barriers to maintain the distinct internal environment of each compartment. Within these sheets, individual cells are mechanically linked to each other to maintain the structural integrity of the sheet, and the intercellular space between adjacent cells is sealed to prevent the diffusion of solutes through the intercellular space. TJs are directly involved in this intercellular sealing.

Structure

ELECTRON MICROSCOPIC IMAGE

On ultrathin section electron microscopy, tight junctions (TJs) appear as a zone where plasma membranes of neighboring cells focally make complete contact (Figure 1D). On freeze-fracture electron microscopy, TJs are visualized as a continuous, anastomosing network of intramembranous particle strands (TJ strands or fibrils) and complementary grooves.

THREE-DIMENSIONAL IMAGE

Detailed electron microscopic observations led to our current understanding of the three-dimensional structure of TJs (Figure 1C); each TJ strand is associated laterally with another TJ strand in apposing membranes of adjacent cells to form "paired" TJ strands, where the intercellular space is completely obliterated.

Molecular Architecture

INTEGRAL MEMBRANE PROTEINS

Three distinct types of integral membrane proteins are localized at TJs: claudin, occludin, and JAM (Figure 2). Claudin with a molecular mass of ~23 kDa is a major constituent of TJ strands, and bears four transmembrane domains. In mice and humans, claudins comprise a multigene family consisting of 24 members. When each claudin species is overexpressed in mouse fibroblasts, claudin molecules are polymerized within the plasma membranes to reconstitute paired TJ strands. As many distinct species of claudins are coexpressed in individual cells, heterogeneous claudin species are copolymerized to form individual TJ strands as heteropolymers, and between adjacent TJ strands claudin molecules adhere to each other in both homotypic and heterotypic manners, except in certain combinations.

Occludin, a ~65 kDa integral membrane protein, also bears four transmembrane domains, but does not show any sequence similarity to claudins. Occludin is incorporated in TJ strands *in situ*, but TJ strands can be formed without occludin. JAM with a single transmembrane domain associates laterally with TJ strands, while not constituting the strands *per se*. The physiological functions of occludin and JAM remain elusive.

PERIPHERAL MEMBRANE PROTEINS

Claudins are packed densely within the TJ strands. Therefore, their cytoplasmic surface appears similar to a brush consisting of numerous short carboxyl-terminal cytoplasmic tails of claudins. Interestingly, most of these tails end in valine at their carboxyl termini, suggesting that these carboxyl termini directly bind to PDZ domains. Therefore, the cytoplasmic surface of TJ strands can be regarded as a magnetic bar that strongly attracts and recruits many PDZ-containing proteins.

Indeed, three related PDZ-containing proteins, ZO-1 (~220 kDa), ZO-2 (~160 kDa), and ZO-3 (~130 kDa), are concentrated at TJs. These molecules all have three PDZ domains (PDZ1 to 3), one SH3 domain, and one GUK (guanylate kinase-like) domain in this order from their NH_2-termini indicating that they belong to membrane-associated guanylate kinase-like homologues (MAGUKs). Among these three PDZ domains, PDZ1 domain binds directly to the COOH termini of claudins.
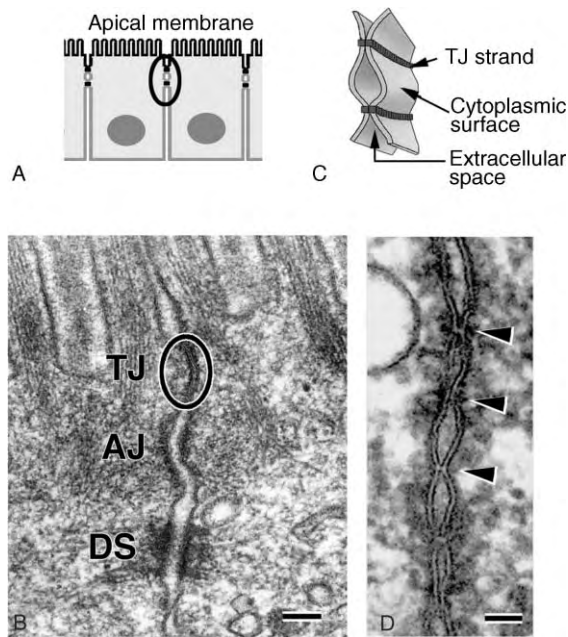
FIGURE 1 Structure of tight junctions. (A) Schematic drawing of simple epithelial cells. The junctional complex is located at the most apical part of lateral membranes (encircled). (B) Electron micrograph of the junctional complex in mouse intestinal epithelial cells. TJ, tight junction (encircled); AJ, adherens junction; DS, desmosome. (C) Schematic drawing of tight junctions. Individual tight junction strand within plasma membranes associates laterally with another strand in apposing membranes to form a paired strand. (D) Ultra-thin sectional view of tight junctions. At kissing points of TJs (arrowheads), the intercellular space is obliterated. Bars, 200 nm (b); 50 nm (d).

In addition to these three TJ-specific MAGUKs, several other PDZ-containing proteins are recruited to the cytoplasmic surface of TJ strands. These proteins are multidomain proteins, and may function as adapters at the cytoplasmic surface of TJs together with non-MAGUK proteins such as cingulin and symplekin, which recruit various proteins including cytoskeletal as well as signaling molecules.
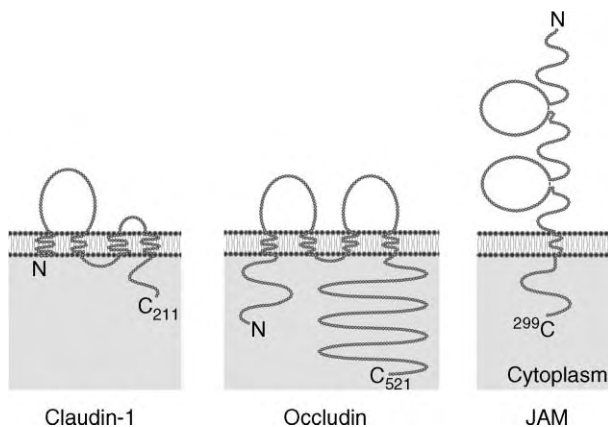
FIGURE 2 Integral membrane proteins localized at tight junctions. Membrane folding models of mouse claudin-1, occludin, and JAM.

Functions

BARRIER FUNCTION

Claudins are directly involved not only in the formation of TJ strands but also in their barrier function in simple epithelia, not only for solutes but also for pathogens such as bacteria. Furthermore, TJs are not simple barriers; they show ion and size selectivity and their barrier function varies significantly in tightness depending on the cell-type and physiological requirements. Such regulated and diversified permeability of TJs is required for dynamically maintaining the environment of each compartment. To explain the material transport across TJs, aqueous pores (or paracellular channels) have been postulated to exist within paired TJ strands. It is now widely believed that the combination and mixing ratios of claudins within individual paired strands determine the tightness and selectivity of these aqueous pores.

FENCE FUNCTION

TJ strands are heteropolymers of integral membrane proteins, claudins, within plasma membranes, and continuously encircle the top of individual epithelial/endothelial cells to delineate the border between the apical and basolateral membrane domains. Therefore, it is believed that TJ strands act as a "fence," limiting the lateral diffusion of lipids and proteins between the apical and basolateral membrane domains. Indeed, when fluorescently labeled lipids are inserted into the outer leaflet of the apical membrane of cultured epithelial cells, these lipids are retained on the apical surface. In contrast, the importance of TJs in the asymmetric distribution of integral membrane proteins remains controversial.

OTHER POSSIBLE FUNCTIONS

As previously mentioned, various types of proteins including PDZ-containing proteins are recruited to the cytoplasmic surface of TJ strands to form a huge macromolecular complex. This macromolecular complex is expected to play central roles in the intracellular signaling of epithelial cells, being involved in the regulation of their proliferation and morphogenesis, i.e., contact inhibition of cell growth and epithelial cell polarization, respectively.

SEE ALSO THE FOLLOWING ARTICLE

MDR Membrane Proteins

GLOSSARY

apical surface   The free surface of epithelial cells facing the luminal space or external environment.
endothelial cells   Thin, flattened cells of mesoblastic origin that are arranged in a single layer lining the blood vessels and some body cavities (e.g., those of the heart).

epithelial cells Closely packed cells, arranged in one or more layers, that cover the outer surfaces of the body or line any internal cavities or tubes (other than the blood vessels, heart, and serous cavities).

freeze-fracture replica electron microscopy An electron microscopic method that uses metal replicas to visualize the interior of cell membranes. This technique provides a convenient way of visualizing the distribution of large integral membrane proteins as intramembranous particles on the plane of a membrane.

PDZ domain Protein–protein interaction domain first described in proteins PSD-95, DLG, and ZO-1.

SH3 domain Src homology region 3, a protein domain present in various signaling proteins.

FURTHER READING

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell*. Garland Publishing, New York.

Cereijido, M., and Anderson, J. (2001). *Tight Junctions*. CRC Press, New York.

Gonzalez-Mariscal, L., Betanzos, A., Nava, P., and Jaramillo, B. E. (2003). Tight junction proteins. *Prog. Biophys. Mol. Biol.* **81**, 1–44.

Schneeberger, E. E., and Lynch, R. D. (1992). Structure, function, and regulation of cellular tight junctions. *Am. J. Physiol.* **262**, L647–L661.

Tsukita, Sh., Furuse, M., and Itoh, M. (2001). Multifunctional strands in tight junctions. *Nat. Rev. Mol. Cell. Biol.* **2**, 285–293.

BIOGRAPHY

Shoichiro Tsukita is a Professor in the Department of Cell Biology at Kyoto University, Kyoto, Japan. His principal research interests are in cell–cell adhesion and cytoskeletons. He holds M.D. and Ph.D. from the University of Tokyo. His group established a new isolation procedure for intercellular junctions and identified occludin and claudins.

Toll-Like Receptors

Himanshu Kumar, Kiyoshi Takeda and Shizuo Akira
Osaka University, Osaka, Japan

Toll-like receptors (TLRs) are the family of receptor molecules that recognizes microbe-specific conserved molecular motifs. Recognition of these molecular motifs by innate cells of innate immune system activates downstream signaling cascades to induce secretion of cytokines, chemokines, and expression of costimulatory molecules. These effects subsequently instruct the major component of adaptive immune system that is B and T cells to exhibit pathogen-specific antimicrobial activity. At the end of the twentieth century, this family of receptors has been discovered and till date ten TLR molecules are known.

Innate Immune System

Innate immunity acts as the first line of host defense against microbial pathogens. It is conserved from flies to mammals. Innate immunity is dependent on both cellular components such as monocytes, macrophages, neutrophils, dendritic cells, and endothelial cells, and humoral components such as C-reactive protein (CRP), lysozyme, and complements. The cellular component of innate immunity recognizes a wide range of pathogen-associated molecular motifs by virtue of pattern recognition receptors (PRR). There are two types of PRRs. One type of PRRs, such as CR3 and C-type lectins, facilitate receptor-mediated phagocytosis. Other PRRs, such as TLRs, trigger the activation of innate immunity. It has recently been shown that toll-like receptor (TLRs) are the most important PRRs in triggering the activation of innate immunity.

TLRs and their Ligands

Toll receptor was first described in context of the dorso-ventral development of *Drosophila* embryo. Toll and its mammalian homologues (TLRs) are type I membrane proteins, harboring an ectodomain consisting of leucine-rich repeats (LRRs) and one or two cysteine rich regions. The intracellular domain of toll receptors is structurally similar to that of the IL-1 receptor. Therefore, it is called Toll/IL-1R (TIR) domain. This region of TLRs is required for homophilic and heterophilic interaction of TLRs. The TIR domain further interacts with several TIR domain-containing adaptors such as MyD88, TIRAP/Mal, and TRIF. To date, ten TLRs (TLR1–10) have been identified and shown to recognize different classes of pathogen-derived molecular motifs (Table I).

Distribution of TLRs

Distribution of these TLR molecules in different cells and tissue type is shown in Table II.

Intracellular Signal Transduction through TLR

The molecular mechanisms by which TLRs trigger activation of innate immunity have been elucidated through analysis of signaling pathways. Intracellular signal transduction initiates from the TIR domain of TLRs. The TIR domain of TLRs interacts with the TIR domain (present in the C terminus) of MyD88 adaptor protein. MyD88 contains a death domain at the N terminus. When ligands are bound to TLRs, MyD88 recruits IRAK to TLRs through the interaction of death domains and leads to phosphorylation of IRAK, which has serine/threonine kinase activity. IRAK then binds TRAF-6 to form the IRAK–TRAF6 complex. This complex activates the NFκB transactivator that induces transcription of inflammatory cytokine genes. This complex also activates the JNK pathway and generates the active AP-1 transactivator as shown in Figure 1. In addition to this MyD88-mediated component of signaling pathway, a MyD88-independent component leading to IFN-β production has been identified in TLR3 and TLR4 signaling pathways.

MyD88-Dependent Signaling Pathway

Stimulation of MyD88-deficient macrophages with known TLR ligands induced no production of inflammatory cytokines. This indicates that MyD88 plays a pivotal role in the TLR-mediated production of inflammatory

TABLE I

Description of Toll-Like Receptor

TLR	Gene (Human)	Functional receptor	Ligands of TLR	Immune response
TLR1	4p14	Heterodimers of TLR2-TLR1	Tri-acyl lipopeptides (bacteria, Mycobacteria), soluble factors, *Neisseria meningitis*	Induce secretion of inflammatory cytokine(s)
TLR2	4q32	Heterodimers of TLR2-TLR1 and heterodimer of TLR2-TLR6	Lipoprotein/lipopeptides (different pathogen), peptiodoglycan (gram-positive bacteria), lipoteichoic acid (gram-positive bacteria), lipoarabinomannan (Mycobacteria), a phenol soluble modulin (*Staphylococcus aureus*), glycoinositolphospholipids (*Trypanosoma cruzi*), glycolipid (*Treponema maltophilum*), porins (*Neisseria*), zymosan (fungi), HSP70 (host), atypical LPS (*Leptospira interrogans*), atypical LPS (*Porphyromonas gingivalis*)	Induce secretion of inflammatory cytokine(s)
TLR3	4q35	Not known	Double-stranded RNA (mainly from virus)	Induce type I interferon (IFN-α/IFN-β)
TLR4	9q32-33	LPS-LPS binding protein complex associate with CD14 and MD-2, in B-cell TLR4 associate with RP105 and MD-1	LPS (gram-negative bacteria, taxol (plant), HPS60 (host), HPS70 (host), fusion protein (RSV), HPS60 (*Chlamydia pneumoniae*), fibronectin (host), envelope proteins (MMTV), type III repeat extra domain A of fibronectin (host), oligosaccharides of hyaluronic acid (host), polysaccharides fragments of heparin sulfate (host) and fibrinogen (host)	Induce secretion of inflammatory cytokine(s) and chemokines
TLR5	1q33.3	Not known	55 kDa monomer obtained from bacteria flagella (flagellin)	Induce inflammatory cytokine(s)
TLR6	4p14	Heterodimer of TLR2-TLR6	Di-acyl lipopeptides (mycoplasma)	Induce secretion of inflammatory cytokine(s)
TLR7	Xp22	Not known	Synthetic compounds like imidazoquinoline (imquimod (R-848), resiquimod (R848)), loxoribine and bropirimine	Induce secretion of inflammatory cytokine(s)
TLR8	Xp22	Not known	Ligand(s) not known	Not known
TLR9	3p21.3	TLR9 (localized in endosomal compartment)	Unmethylated CpG DNA (Mycobacteria and Bacteria)	Activate Th1 type immune response, Induce proliferation of B-cell, activate macrophage and DCs
TLR10	Not known	Not known	Ligand(s) not known	Not known

TABLE II

Distribution of TLRs

Cell/tissue type	Types of Toll-like receptor									
	TLR1	TLR2	TLR3	TLR4	TLR5	TLR6	TLR7	TLR8	TLR9	TLR10
1 Monocytes	+	+	−	+	+	+	+	+	+	+
2 Macrophage	+	+	−	+	+	+	+	+	+	+
3 Plasmacytoid dendritic cell	+	+	−	+	+	−	+	+/−	M(+) H(+)	NK
Myeloid dendritic cell	NK	+	NK	+	NK	+	H(+/−) M(+)	+	M(+) H(−)	NK
4 Immature dendritic cell[a]	+	+	NK	+	+	NK	NK	NK	NK	NK
Mature dendritic cell[a]	NK	NK	+	NK	NK	NK	NK	NK	NK	NK
5 Mast cell	NK	+	NK	+	−	+	NK	+	NK	NK
6 Intestinal cell (Apical side)	NK	NK	NK	+(low)	NK	NK	NK	NK	NK	NK
Intestinal cell (Basolateral side)	NK	NK	NK	NK	+	NK	NK	NK	NK	NK
7 Renal epithelia	NK	+	NK	+	NK	NK	NK	NK	NK	NK
8 Pulmonary epithelia	NK	NK	NK	+	NK	NK	NK	NK	NK	NK
9 Corneal epithelia	NK	NK	NK	+	NK	NK	NK	NK	NK	NK

[a] Initially immature dendritic cells (DCs) express TLR1, 2, 4, and 5 as soon as immature DC encounter with antigen the expression of TLR1, 2, 4, and 5 regresses and expression of TLR progresses.

Abbreviations: M, mouse; H, human; NK, not known.

cytokines. Four family members of IRAK have been reported in mammals (IRAK-1, IRAK-2, IRAK-4, and IRAK-M). IRAK-1 and IRAK-4 are ubiquitously expressed active kinases, whereas IRAK-M is preferentially expressed in macrophage and monocytes. Stimulation of IRAK-4-deficient macrophages with ligands of TLR2, TLR3, and TLR4 induced almost no production of inflammatory cytokines. Inflammatory cytokine production was also significantly diminished in IRAK-1-deficient macrophages stimulated with LPS (the TLR4 ligand). Thus, IRAK-1 and IRAK-4 are involved in the TLR-mediated signaling pathways. On the other hand, IRAK-M-deficient macrophages showed overproduction of inflammatory cytokines in response to TLR ligands, indicating that this molecule is involved in the negative control of cytokine production. Stimulation of TRAF6-deficient macrophages with LPS led to impaired production of inflammatory cytokines, suggesting that TRAF6 is also important in the TLR-mediated signaling pathway.

In addition to MyD88, TIR domain-containing adaptor protein (TIRAP)/MyD88-adaptor-like (Mal) has been identified as the second TIR domain-containing molecule. Stimulation of TIRAP/Mal-deficient mice with LPS (the TLR4 ligand) and lipopeptide (the TLR2 ligand) induced no production of inflammatory cytokines. Thus, TIRAP/Mal is essential for the MyD88-dependent pathway of TLR2 and TLR4.

MyD88-Independent Signaling Pathway

As described, stimulation of MyD88-deficient macrophages with TLR ligands did not lead to inflammatory cytokine production. However stimulation with LPS induced delayed activation of JNK and NFκB in these cells, indicating that there is another bypass pathway which transduces signals in the absence of MyD88. Beside LPS, dsRNA (the TLR3 ligand) is also capable of activating NFκB in MyD88-deficient macrophages. In addition, it has been shown that LPS and dsRNA both induce the activation of IRF-3 transactivator and lead to the production of IFN-β in a MyD88-independent fashion.

Recently, a third TIR domain-containing adaptor protein called TIR domain-containing adaptor inducing IFN-β (TRIF) was identified. TRIF-deficient mice showed no activation of IRF-3 or production of IFN-β in response to the TLR3 and TLR4 ligands. Therefore, TRIF is an essential molecule for the activation of IRF-3 transactivator through TLR3 and TLR4. Furthermore, the lack of LPS-induced activation of NFκB and JNK in MyD88/TRIF double deficient cells indicates that TRIF is essential for the TLR4-mediated MyD88-independent pathway.

Structural and Functional Homologues of TLRs

RP105, a molecule containing an ectodomain with LRRs and one cysteine-rich region but lacking a cytoplasmic TIR domain, has been identified. RP105 is preferentially expressed in mature B cells. Mice deficient in RP105 or TLR4 does not show proliferation in
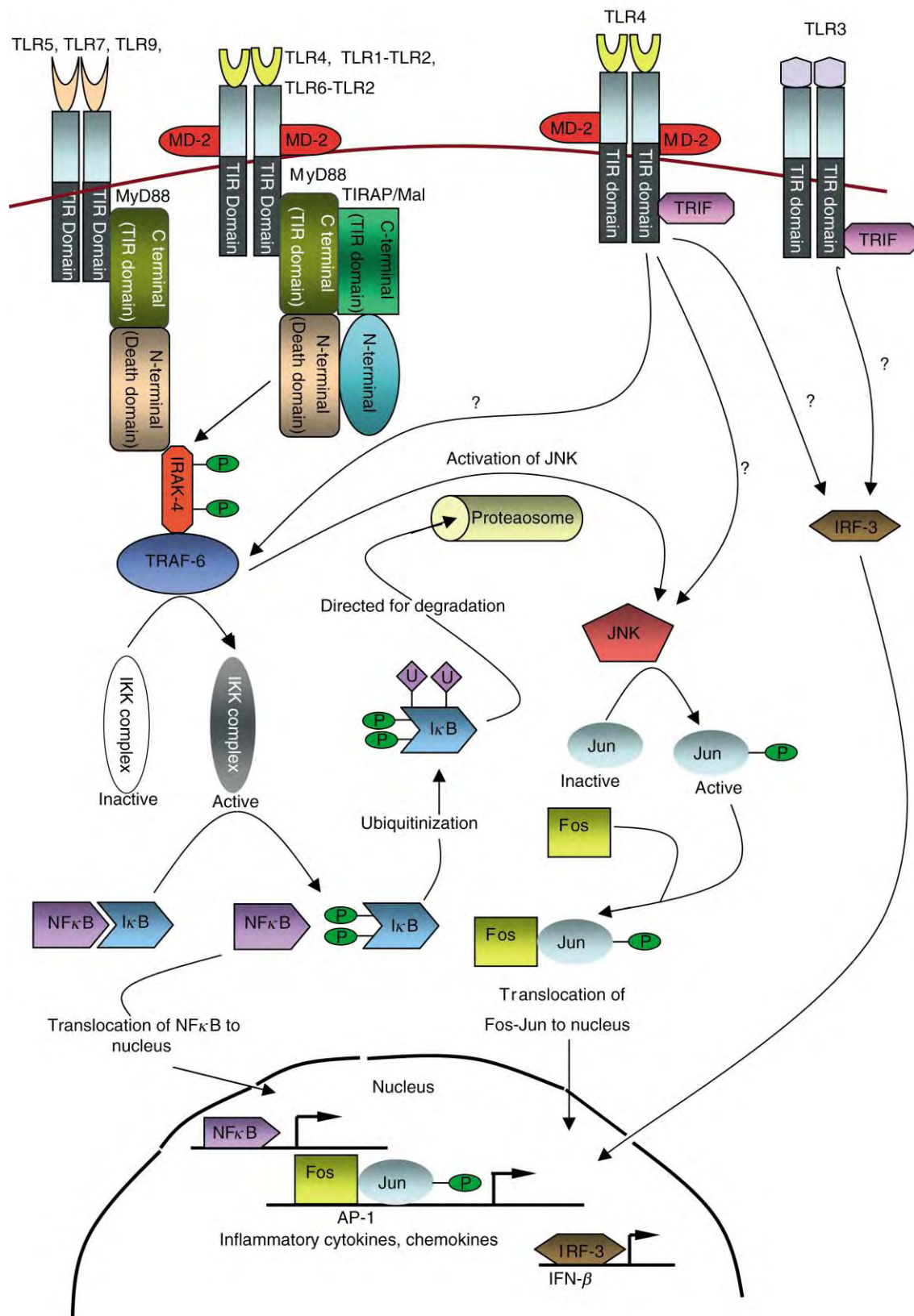
FIGURE 1 Intracellular signaling pathway through TLR. Binding of TLR ligands initiates TLR-mediated signaling pathways from the cytoplasmic TIR domain. The TIR domain-containing adaptor MyD88 is essential for the induction of inflammatory cytokines via all TLRs. Another TIR domain-containing adaptor TIRAP is involved in the MyD88-depdendent pathway of TLR2 and TLR4. In TLR3 and TLR4 signaling, the IRF-3 transcription factor is activated in a MyD88-independent fashion. The third TIR domain-containing adaptor TRIF mediates the MyD88-independent signaling, thereby inducing IFN-β.

response to LPS. This shows that RP105 is required for LPS-induced responses in B cells.

Besides mammals, plants also have more than 100 genes which encode proteins with the TIR domain. Some of these proteins also have LRRs and nucleotide-binding site (NBS). These proteins play an important role in defense against plant virus.

Future Perspective

Pathogens are dynamic entities, which contain many complex biomolecules such as LPS and lipoprotein. In recent years, intensive research has been conducted to elucidate the different TLR signaling pathways such as the MyD88-dependent and MyD88-independent pathway by using purified pathogen-derived biomolecules. However, it is still not known how all these pathogen-derived complex molecules interact with macrophages, what the net or final signals for clearance of pathogens are, or how different pathways communicate with each other *in vivo*. In the near future, all these questions will be clarified using several mutant mouse strains lacking the different components of TLR signaling pathways.

SEE ALSO THE FOLLOWING ARTICLES

Chemotactic Peptide/Complement Receptors • Nuclear Factor kappaB

GLOSSARY

adaptive immunity The response of antigen specific lymphocytes to an antigen, including the development of immunological memory. Adaptive immune responses are generated by clonal selection of lymphocytes.

complements The complement system is a set of plasma protein synthesized in the liver and act together on the surface of pathogens to form the membrane attack complex (MAC) in order to kill pathogens.

innate immunity The early phases of the host response to an infection, and is normally present in all individuals at all times. This immune system can discriminate between self and nonself (pathogen), but does not increase responses with repeated exposure to a given pathogen.

JNK A kinase that activates Jun by phosphorylation. Phosphorylated Jun associates with Fos and forms a complex known as AP-1, which transcribes many cytokine genes.

NFκB A transcription factor made up of 50 and a 65 kDa subunits. It is normally found in the cytosol, where it is bound to IκB, an inhibitor of NFκB. This transcription factor is involved in the transcription of many cytokine genes.

FURTHER READING

Akira, S. (2003). Mammalian toll-like receptors. *Curr. Opin. Immunol.* **15**, 5–11.

Akira, S., Takeda, K., and Kaisho, T. (2001). Toll-like receptor: Critical proteins linking innate and acquired immunity. *Nature* **2**, 675–680.

Medzhitov, R., and Kopp, E. B. (1999). The Toll receptor family and control of innate immunity. *Curr. Opin. Immunol.* **11**, 13–18.

Takeda, K., and Akira, S. (2003). Toll receptors and pathogen resistance. *Cell. Microbiol.* **5**, 143–153.

Takeda, K., Kaisho, T., and Akira, S. (2003). Toll-like receptors. *Annu. Rev. Immunol.* **21**, 335–376.

Yamamoto, M., Sato, S., Hemmi, H., Hoshino, K., Kaisho, T., Sanjo, H., Takeuchi, O., Sugiyama, M., Okabe, M., Takeda, K., and Akira, S. (2003). Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. *Science* **301**, 640–642.

BIOGRAPHY

Shizuo Akira is a Professor in the Department of Host Defense at Research Institute for Microbial Diseases, Osaka University in Osaka, Japan. His principal research interests are in mammalian innate immunity. He holds a Ph.D. from Osaka University and has received postdoctoral training at the University of California, Berkeley. He analyzes the molecular mechanism by which the innate immunity senses microbial invasion in mammals through generation of knockout mice.

Kiyoshi Takeda was an Assistant Professor in the Department of Host Defense at Research Institute for Microbial Diseases, Osaka University in Osaka, Japan. He received Ph.D. from Osaka University. Presently, he is a Professor in Department of Molecular Genetics at Medical Institute of Bioregulation, Kyusha University in Fukuoka, Japan. His research focuses on the mechanism for regulation of macrophage activity and the involvement of innate immunity in inflammatory and infectious diseases.

Himanshu Kumar is a graduate student, pursuing research in the Department of Host Defense at Research Institute for Microbial Diseases, Osaka University in Osaka, Japan. His research interest is elucidation of new molecules involve in TLRs signaling.

Transcription Termination

Thomas J. Santangelo and Jeffrey W. Roberts
Cornell University, Ithaca, New York, USA

Transcription termination is the process by which RNA synthesis by RNA polymerase (RNAP) is stopped and both the RNA and enzyme are released from the DNA template. Since most regulation occurs at the initiation of RNA synthesis, failure to terminate transcription at the ends of operons would lead to bypass of downstream gene regulation due to rogue elongation complexes continuing transcription beyond the end of the upstream operon. Two types of transcription termination in bacteria have been characterized in both biochemical reactions with purified RNA polymerase and in genetic experiments: (1) "intrinsic termination," which requires only special template sequences, and (2) "factor-dependent termination," which requires an additional enzyme that interacts with RNA polymerase. The RNA polymerase transcription elongation complex is extremely stable – it must persist over time and distance in order for gene expression to succeed – but it also must be disrupted efficiently at termination sites.

Models of terminator function have been strongly informed by the recently determined atomic structures of RNAP and the transcription elongation complex, which reveal the molecular sources of stability that are overcome by termination mechanisms.

Figure 1 shows a general elongation complex, highlighting the major contacts that provide stabilizing energy to the complex. The RNAP enzyme surrounds an 8 or 9 bp DNA–RNA hybrid that is held within the main channel of the enzyme. The RNA emerges from the enzyme through an exit channel where additional contacts are made between single-stranded RNA (ssRNA) and the protein. By far, the greatest sources of stabilizing energy to the complex are the hydrogen-bonding within the hybrid itself and the interactions between the hybrid and the surrounding protein. Terminators function by weakening these contacts, leading to dissociation of the hybrid and disruption of the complex.

Intrinsic Termination

Intrinsic terminators are composed of two critical elements: a short region of dyad symmetry in the DNA which encodes an RNA sequence capable of forming an RNA–RNA duplex or RNA "hairpin," and a sequence encoding a uridine-rich sequence in the RNA that immediately follows the hairpin (Figure 1). Transcript release occurs 8–10 bp downstream of the base of the RNA hairpin. Correct base pairing within the stem of the hairpin, the presence of the uridine-rich segment, and the correct spacing of these elements are critical to termination efficiency. Terminator efficiency also can be modulated by the DNA sequence immediately upstream and downstream of the termination position.

Current research is focused on the way these sequences and structures facilitate release of the transcript. The complex has been weakened just by sequence context alone; the final 8 or 9 bases of a terminated transcript occupy the hybrid-binding site and a dA:rU hybrid is the weakest possible combination of nucleic acids. In the absence of a terminator hairpin structure, the weak hybrid alone is sufficient to induce RNAP to pause at the end of the uridine-rich sequence. The role of the uridine-rich sequence in termination is likely twofold: (1) to pause the elongation complex immediately downstream of the region capable of hairpin formation; and (2) to minimize the energetic contribution of the hybrid to the overall stability of the complex.

The role of the hairpin structure and its effects on the elongation complex remain debated. Although their details differ, most models predict that hairpin formation disrupts the upstream bases of the hybrid, triggering collapse of the elongation complex. One model proposes that hairpin formation immediately upstream of the complex pushes RNAP downstream in the absence of further synthesis. Forward translocation of several base pairs would reduce the length of the weak RNA/DNA hybrid, and the loss of hybrid interactions would result in transcript release. An alternative model of termination evokes a relatively large conformational change in the enzyme upon interaction of the hairpin with RNAP. The hairpin structure formed in the RNA could physically interact with the complex, opening the enzyme and freeing the RNA transcript from its enclosure within the elongation complex, a reaction that again would be promoted by the weak hybrid. It is also possible that a large conformational change accompanies a forward translocation to result in both transcript and enzyme release.
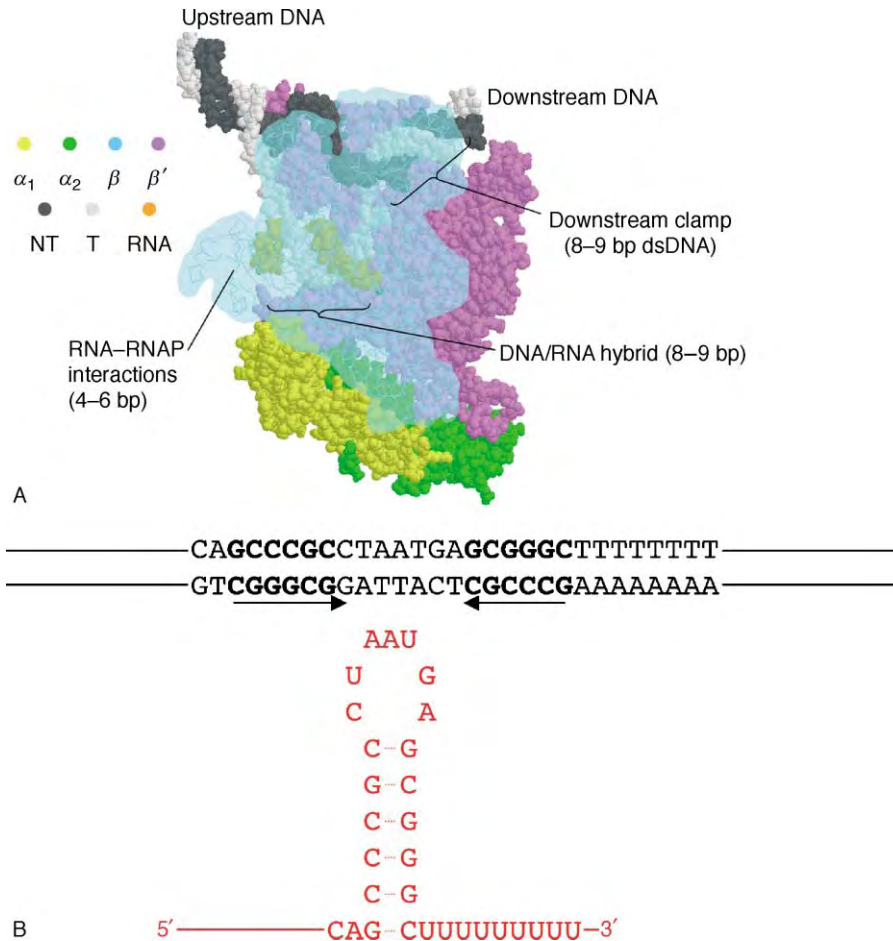
FIGURE 1 (A) The transcription elongation complex is shown, highlighting the protein–nucleic acid contacts that provide stabilizing energy to the complex. Subunits are colored according to the key at the left; β has been made partially transparent in order to reveal the main channel of RNA polymerase and the nucleic acid framework inside. NT, non-template DNA strand; T, template DNA strand. (B) Sequence of an intrinsic terminator (the *E. coli trpA* terminator) is shown in black with the corresponding RNA sequence and hairpin shown in red. The region of dyad symmetry in the DNA is shown in bold and is underlined and codes for the RNA hairpin.

All of the multisubunit RNAPs are predicted (and some are known) to share a common three-dimensional structure and thus are similar in terms of the contacts made between the enzyme and nucleic acid framework. Thus, termination of transcription by the specialized RNA polymerases in eukarya (RNA pol I, II, and III) and the archaeal RNAPs likely also involves mechanisms to initiate disruption of the RNA/DNA hybrid. Termination by pol I and pol III in eukarya, and the RNAP from archaea, does involve oligo-T tracts at or near the termination positions, although the mechanisms are unknown.

Factor-Dependent Termination

RHO-MEDIATED TERMINATION

Rho factor is a hexameric RNA-dependent ATPase that is found in most eubacteria, but not in archaea or eukarya. Its general mechanism of action is to bind the nascent transcript of an elongation complex, wrapping ~60 nucleotides of RNA, translocating toward RNAP in an ATP-dependent reaction, and finally dissociating the complex by an unknown mechanism. An atomic structure of the Rho hexamer shows a hexameric ring of monomers, each with two domains. The structural model proposes that a ring of RNA-binding domains engages most of the RNA, which then projects through the center of the ring to contact the ATPase domains that effect translocation. A break in the hexameric ring ("lock-washer" model) is proposed to allow RNA to penetrate the hexamer.

Rho function is largely determined by the very tight coordination between transcription and translation, because it requires free RNA immediately upstream of RNA polymerase in order to act. Translation in bacteria occurs as the transcript is being made, and the ribosome travels immediately behind the elongation complex,

so that free RNA is generally unavailable to Rho. However, when the ribosome stalls are released, as at a translation stop codon, transcription and translation become uncoupled. As RNA polymerase continues to transcribe, more RNA is exposed to solution, allowing Rho access to the RNA. Since RNAP elongates at 50–100 nucleotides per second, Rho provides a rapid mechanism to ensure that RNAP does not transcribe much beyond the point of uncoupling.

Although a major function of Rho is to abort futile transcription as described above, Rho-dependent termination also is the natural mechanism at certain sites, likely serving as the sole mechanism to terminate transcription of many operons in *E. coli*. Well-characterized examples of such sites have properties that favor Rho action during transcription: absence of translation, the absence of any secondary structure that would prevent Rho wrapping RNA, and a cytidine-rich transcript that favors Rho binding; furthermore, natural Rho-dependent terminators are transcription pause sites, a property that presumably gives Rho time to translocate to RNAP and carry out release. It is likely that these characteristics are sufficiently common that Rho could generally act after translation fails in most operons.

Mfd and Transcription-Coupled DNA Repair

DNA damage is repaired more rapidly in actively transcribed than in rarely transcribed genes, and the template strand is preferentially repaired over the non-template strand. These facts led to the discovery that an RNAP molecule stalled at a DNA lesion can serve as a signal that the DNA is damaged and lead to its rapid repair. This activity contributes to an essential function, because a lesion in a single copy gene would block production of an essential protein. In bacteria, RNAP stalled at a template strand DNA lesion is the target of the Mfd protein, also termed the Transcription Repair Coupling Factor. Mfd is an ATPase and a DNA translocase that mediates both removal of RNAP stalled at the lesion and recruitment of the nucleotide excision repair machinery to the site of the lesion. Similar to Rho activity, Mfd can terminate transcription of almost any obstructed RNAP. Mfd binds to the DNA immediately upstream of the stalled elongation complex and makes direct contact with RNAP; using its two translocase domains, Mfd then pushes RNAP downstream along the DNA in an ATP-dependent reaction, collapsing the upstream edge of the transcription bubble as it moves. In the case where elongation is blocked by a template strand DNA lesion, the rewinding and forward translocation are proposed to force the RNA from the elongation complex and drive RNAP downstream to mediate termination. The rewinding of upstream

sequences also can rescue complexes that are backtracked, forcing RNAP into an active, forward configuration in the case where continued elongation is possible.

General Elongation Factors – NusA and NusG

NusA and NusG are nearly universally conserved eubacterial proteins – both essential in *E. coli* – that affect the rate of elongation by RNA polymerase and modulate the efficiency of termination. Both NusA and NusG bind to the RNAP core enzyme and influence the activity of RNAP alone, in combination with each other, or in combination with specific regulatory factors. In a purified system, NusA generally slows elongation, enhances pausing, and increases intrinsic termination efficiency. NusA may interact with the RNA hairpin and enhance the proposed interactions between RNA hairpins and RNAP. NusG appears, in many ways, to function oppositely: in general, it enhances elongation and limits pausing. The molecular details of the interactions between these factors and RNAP are unknown.

Both NusA and NusG also influence Rho activity, and although strong Rho-dependent terminators are pause sites, the effects of NusA and NusG on Rho function are opposite to the prediction based on their effects on pausing: thus, NusA promotes pausing but inhibits Rho function, and NusG inhibits pausing and promotes Rho function. It has been suggested that the NusG effect may result from NusG inhibiting backtracking; this could stimulate Rho function if Rho is unable to release a backtracked complex. There may be a distinction between direct effects of the factors on RNA polymerase at a site where Rho-dependent termination occurs, and indirect effects due to NusA and NusG modulation of the overall rate of transcription. Thus, NusA inhibition could in fact favor termination by slowing RNAP so that Rho can catch it through translocation along the RNA. An important relation between the activities of NusA and Rho *in vivo* is shown by the ability of a NusA defective cell to survive if Rho is also defective, but not if Rho is normally functional.

Regulated Termination and Antitermination

Termination can serve as a means to regulate gene expression in two fundamentally different modes that involve antitermination. First, site-specific attenuator-based systems act at particular termination sites in operons to allow or disallow continued transcription, thus determining whether downstream genes are expressed. These terminators are usually intrinsic, but

are sometimes Rho-dependent. Second, generalized antiterminators modify RNA polymerase so that it ignores terminators that otherwise stop transcription, again allowing expression of downstream genes. Finally, a few examples are known of generalized modifications of RNA polymerase that cause termination.

SITE-SPECIFIC ANTITERMINATION: ATTENUATORS AND RIBOSWITCHES

Intrinsic termination is used to control expression of many operons through devices that alternatively permit or prevent termination in a leader sequence or initially translated region of a set of genes. Such operons contain alternative RNA structures in the beginning of the transcribed region; competition between alternative folding patterns determines whether or not intrinsic terminators are formed, and in turn, whether transcription of the remaining portion of the operon or gene continues. The competition between structures in the leader sequence is influenced in different examples by coordinated translation and ribosome positioning on the RNA, or by small molecules, including tRNAs, that bind to the RNA transcript and affect global folding.

Regulation of the *trp* operon is the classic example of translation-coupled attenuation. Stalling of the ribosome at tryptophan codons, because of limitation in charged-aminoacyl tRNAs, determines which sequences of the transcript can fold, and thus whether a terminator hairpin can be formed. Attenuation can also be mediated by interactions between uncharged tRNAs and leader sequences containing binding sites for these tRNAs, termed T-boxes. Riboswitches consist of the leader and initially translated regions of regulated operons, which undergo alternative folding in response to small regulatory molecules; both positive and negative riboswitches have been described.

TERMINATION AND ANTITERMINATION FACTORS

Specific regulatory factors, such as the λN and Q proteins, the bacterial RfaH protein, and the HK022 Nun, regulate expression of specific genes or operons by changing the termination capacity of RNA polymerase. In general, the terminator/antiterminator is recruited to particular transcribing RNAPs via *cis*-acting nucleic acid sequences. Details of the exact modifications these factors make to the complex and the effects of these modifications on the stability of the elongation complex are just beginning to be understood at a molecular level.

λN binds a specific site (*nut*) in the nascent RNA, and in a complex with the general elongation factors NusA and NusG, and the bacterial proteins NusB and S10, interacts with RNAP to stimulate elongation through downstream terminators. Both λQ and RfaH require *cis*-acting DNA sequences to pause RNAP at discrete sites and present a novel complex for interaction. λQ protein interacts with RNAP at a σ-dependent promoter proximal pause and becomes a subunit of the elongating RNAP, modifying it to prevent recognition of downstream terminators. Q likely remodels the active site and strengthens interactions between the enzyme and the hybrid to resist the action of a terminator. RfaH regulates expression of particular operons by interacting with RNAP at a regulatory pause (*ops* sequence) to stimulate expression of distal genes. HK022 Nun protein stimulates transcription termination after binding the same RNA sequences that N protein binds, thus blocking superinfection of its host by another lambdoid phage that requires N function for growth. Elongation complexes are halted by the action of Nun but not disrupted. Thus, Nun is not a release factor; Nun-arrested complexes require the activity of the Mfd protein to efficiently release the transcript and disrupt the elongation complex.

SEE ALSO THE FOLLOWING ARTICLES

Ribozyme Structural Elements: Hairpin Ribozyme • RNA Polymerase I and RNA Polymerase III in Eukaryotes • RNA Polymerase II and Basal Transcription Factors in Eukaryotes • RNA Polymerase II Elongation Control in Eukaryotes • RNA Polymerase II Structure in Eukaryotes • Transcription-Coupled DNA Repair, Overview

GLOSSARY

backtracking Reverse translocation of RNA polymerase along the DNA template. The RNA is re-threaded through the complex as RNA polymerase moves backward.

core RNA polymerase and holoenzyme Core RNA polymerase is composed of five subunits ($\alpha_2\beta\beta'\omega$) and is sufficient for RNA synthesis, but not promoter recognition. Binding of an additional subunit, σ, results in formation of holoenzyme; σ is necessary for promoter recognition but is dispensable at latter stages of transcription.

RNA/DNA hybrid An RNA and a DNA strand paired in an A-form 8 or 9 bp double helix within the main channel of RNA polymerase.

termination Release of the nascent transcript from RNA polymerase; whether transcript release and enzyme dissociation from DNA occur simultaneously is unknown.

FURTHER READING

Burns, C. M., Richardson, L. V., and Richardson, J. P. (1998). Combinatorial effects of NusA and NusG on transcription elongation and Rho-dependent termination in *Escherichia coli*. *J. Mol. Biol.* **278**, 307–316.

Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. *Mol. Cell* **3**, 495–504.

Komissarova, N., Becker, J., Solter, S., Kireeva, M., and Kashlev, M. (2002). Shortening of RNA: DNA hybrid in the elongation

complex of RNA polymerase is a prerequisite for transcription termination. *Mol. Cell* **10**, 1151–1162.

Martin, F. H., and Tinoco, I. Jr. (1980). DNA–RNA hybrid duplexes containing oligo (dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Res.* **8**, 2295–2299.

Park, J. S., Marr, M. T., and Roberts, J. W. (2002). *E. coli* transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* **109**, 757–767.

Selby, C. P., and Sancar, A. (1993). Transcription-repair coupling and mutation frequency decline. *J. Bacteriol.* **175**, 7509–7514.

Skordalakes, E., and Berger, J. M. (2003). Structure of the Rho transcription terminator: Mechanism of mRNA recognition and helicase loading. *Cell* **114**, 135–146.

Toulokhonov, I., Artsimovitch, I., and Landick, R. (2001). Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* **292**, 730–733.

Weisberg, R. A., and Gottesman, M. E. (1999). Processive antitermination. *J. Bacteriol.* **181**, 359–367.

Yarnell, W. S., and Roberts, J. W. (1999). Mechanism of intrinsic transcription termination and antitermination. *Science* **284**, 611–615.

BIOGRAPHY

Jeffrey Roberts is the Robert J. Appel Professor of Molecular and Cell Biology in the Department of Molecular Biology and Genetics at Cornell University. His principal research interest is the mechanism of termination and antitermination in bacteria. He obtained his Ph.D. from Harvard University, and he is a member of the National Academy of Sciences and the American Academy of Arts and Sciences, and a Fellow of the American Society for the Advancement of Science.

Transcriptional Silencing

Ann Sutton and Rolf Sternglanz
Stony Brook University, Stony Brook, New York, USA

Transcriptional silencing is the process by which large regions of a eukaryotic genome are rendered transcriptionally inactive due to a change in chromatin structure. The role of chromatin structure in the regulation of gene expression has become an exciting area of study in recent years. In all eukaryotes, DNA is packaged into nucleosomes to form a protein-DNA structure called chromatin. Each nucleosome consists of 146 base pairs of DNA wrapped around a protein octamer consisting of the four core histones, H2A, H2B, H3, and H4. Linker DNA separates adjacent nucleosomes. Microscopic analysis of chromatin from higher eukaryotic cells has revealed the presence of two types of chromatin, a highly condensed form called heterochromatin, and a less condensed form called euchromatin. In general, genes within heterochromatin are transcriptionally silenced, while those within euchromatin are transcriptionally active. The transcriptional silencing results from the inaccessibility of the DNA to components of the transcriptional machinery or from an inability of RNA polymerase to elongate through heterochromatin. Silenced domains are also less accessible to recombination, replication, and repair machinery than are euchromatic regions. Silencing differs from gene-specific repression in a number of ways. For example, silenced domains extend over large regions of DNA, while gene repression is more local and results from protein interactions within the promoter of a single gene. Furthermore, silenced chromatin can be maintained through many generations. The mechanism of this inheritance is not yet well understood.

Heterochromatic or silenced chromatin domains exist in all eukaryotes examined to date. Regions of the DNA that are known to exist as heterochromatin include telomeres and centromeres in fission yeast and metazoans, chromosome 4 in *Drosophila*, and one of the X chromosomes in female mammals.

Transcriptional Silencing in Yeast

Transcriptional silencing is best understood in two genetically tractable organisms, the budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*.

Budding Yeast, *S. cerevisiae*

The mechanism and regulation of formation of silent chromatin domains have been studied most extensively in the budding yeast, *S. cerevisiae*, which has four major silenced regions: the two silent mating loci, telomeres, and ribosomal DNA (rDNA). Haploid budding yeast exists as one of two mating types, **a** or α, determined by whether **a**- or α-specific alleles are expressed at the *MAT* locus. Yeast also have identical **a** and α alleles at two additional loci, *HML* (which usually carries the α genes), and *HMR* (which usually carries the **a** genes). However, the genes at *HML* and *HMR* are not expressed because of transcriptional silencing. If silencing is lost at these loci as a consequence of a mutation in a silencing factor, cells express both **a** and α mating type information and lose the ability to mate. Largely through the use of genetic screens for mutants defective in mating, most of factors required for silencing at the *HM* loci have been identified.

Silencing requires *cis*-acting regions that flank the mating type genes as well as *trans*-acting factors, in particular the four silent information regulator proteins (Sir1–Sir4). The *cis*-acting regions, termed silencers, contain binding sites for at least two of the following three proteins: Rap1, Abf1, and the origin replication complex (ORC) ([Figure 1](#)). All three proteins have essential functions in yeast in addition to their roles in gene silencing; Rap1 and Abf1 are transcription factors, and ORC is required for DNA replication. On the other hand, at the silent mating loci, these proteins function to recruit the Sir proteins to the silencer. Orc1, one of the six subunits of the ORCs, binds Sir1, which then recruits Sir4, which most likely binds as a Sir4–Sir2 complex. Rap1 and Abf1 recruit Sir3 and Sir4. Once the Sir proteins are recruited to the silencer, multiple protein-protein interactions between Sir2, Sir3, and Sir4, and between Sir3 and Sir4 and the amino terminal tails of histone H3 and H4, cause the spreading of these Sir proteins, and consequently silenced chromatin, from the silencers to the regions to be silenced, in this case the *HM* loci ([Figure 1](#)). The extent of the repressed domains can be determined by monitoring the expression of a reporter gene placed at increasing distances from the silencer. Silencing at *HMR* spans about 3.5 kbp of DNA, and silencing at *HML* also extends over several kbp.

Sir3 and Sir4 are believed to play a structural role in silent chromatin. Sir2, on the other hand, has recently
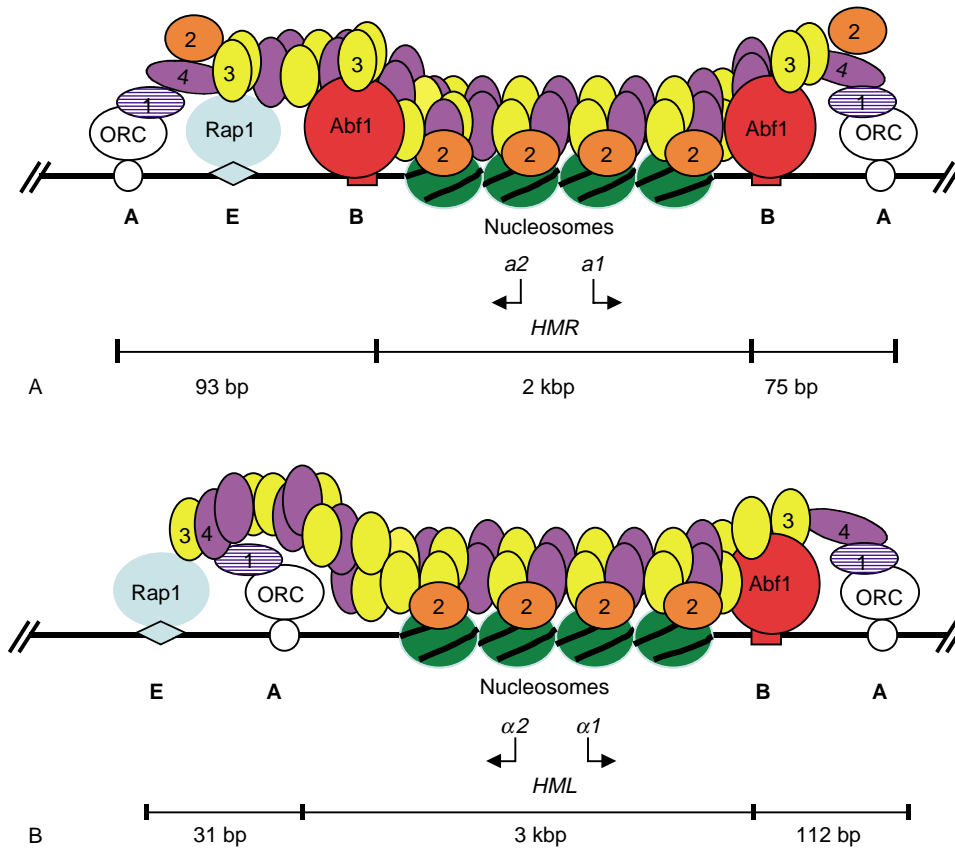
FIGURE 1   The silent mating loci in budding yeast, *S. cerevisiae*. The proteins that bind to the silencers are shown, as well as the four Sir proteins, Sir1 (cross hatched), Sir2 (orange), Sir3 (yellow), and Sir4 (purple). Nucleosomes are in green with diagonal hatch marks. Although the Sir proteins are depicted only between the two silencers, actually the proteins spread bidirectionally from the silencers. (A) The *HMR* locus. (B) The *HML* locus.

been shown to be an enzyme. It is an NAD^+-dependent histone deacetylase that removes acetyl groups from lysines in the amino terminal tails of histone H3 and H4. While this deacetylase activity is not required for the binding of the Sir proteins to the silencers, it is required for the spreading of the Sir proteins and the formation of silent chromatin. Hypoacetylation of the lysines in the histone tails has long been known to be a characteristic of heterochromatin and recent studies suggest that Sir3 preferentially binds to histone tails that are not acetylated. A model to explain the spreading of silent chromatin from the silencers is that Sir2, once recruited to a silencer, deacetylates the lysines on the tails of histones H3 and H4 on adjacent nucleosomes. Sir3 and Sir2–Sir4 complexes then bind to these histones, Sir2 deacetylates the histone tails on adjacent nucleosomes, Sir proteins then bind to them, and so on. The unusual NAD^+ requirement of Sir2 for its deacetylase activity may link silencing to cellular metabolism, and indeed, mutations in genes involved in a salvage pathway for NAD^+ biosynthesis have been shown to affect silencing.

The mechanism of transcriptional silencing at telomeres in *S. cerevisiae* is quite similar to that at the *HM* loci. Telomeres contain multiple binding sites for Rap1,

the same protein found at the *HM* silencers. Rap1 is able to recruit Sir3 and Sir4–Sir2 and these proteins spread several kbp from the telomeres into adjacent nucleosomes, causing silencing of genes in these regions (Figure 2). Sir1 is not involved in telomeric silencing, because there are no ORC-binding sites at telomeres and Sir1 is recruited by ORC. Silencing within the rDNA genes requires Sir2, but none of the other Sir proteins. At these loci, Sir2 is a component of the RENT complex that also contains Net1 and Cdc14. The role of these proteins in the establishment and spreading of silencing at rDNA remains to be elucidated. Since Sir2 is the only Sir protein required for rDNA silencing, the mechanism of silencing at this locus is considered fundamentally different from that at the *HM* loci and telomeres. However, the deacetylation activity of Sir2 is required for rDNA silencing, and in that sense the mechanism is similar at all silent loci in budding yeast.

While silent chromatin can extend over several kbp of DNA, mechanisms exist to stop the spreading. Assays have been developed to look for sequences that can function as boundary elements between euchromatin and heterochromatin. In some cases, boundary elements appear to consist of strong transcriptional promoter
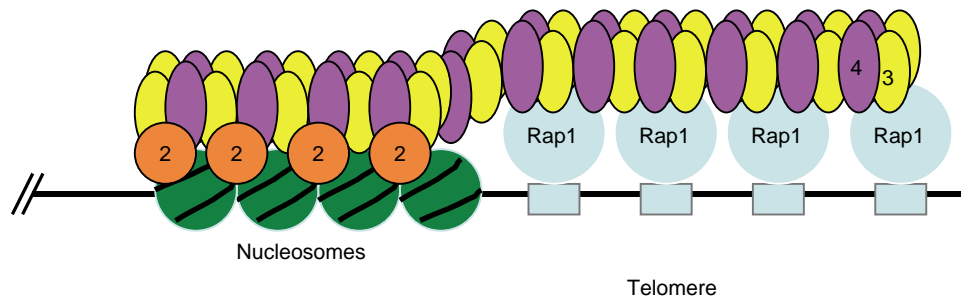
FIGURE 2   Silencing at a yeast telomere. A series of Rap1 proteins is shown binding to the repeated sequence at a telomere. Sir proteins and nucleosomes are designated and colored as in Figure 1.

elements. In these cases binding of transcription factors may sterically block the spread of the silencing proteins. The activity of histone acetyltransferases may also play a role in boundary function, perhaps by acetylating lysine residues on histone tails, and thus counteracting the Sir2-mediated deacetylation that is required for silencing. Mutations in the histone acetyltransferase Sas2 result in the spread of silencing and Sir proteins beyond the normal limits of silenced chromatin at telomeres, and tethering Sas2 to the DNA near *HMR* can limit the spread of silencing. Sas2 specifically acetylates lysine 16 of H4 and it is known that deacetylation of this residue, presumably by Sir2, is particularly important for the formation of silent chromatin.

A recently discovered histone modification that may also function to prevent heterochromatin formation in *S. cerevisiae* is methylation of lysine 79 of histone H3. The Dot1 protein is responsible for this methylation, and most of the H3 in yeast chromatin (but not that in silenced regions) is methylated. A current model is that Sir proteins do not bind well to nucleosomes with H3 methylated on lysine 79, and therefore this modification may function to restrict silencing to a few discrete regions of the genome. In summary, transcriptionally active (euchromatic) regions in yeast appear to be acetylated on H4 lysine 16 and methylated on H3 lysine 79, while silent regions (heterochromatin) lack these modifications.

FISSION YEAST, *S. POMBE*

Fission yeast also has silent mating type information, in this case within a 20 kbp region that is transcriptionally silent and completely devoid of recombination. There are significant differences between the two yeasts in the mechanism of silencing and the proteins involved. The most important difference involves a histone modification not found in budding yeast, namely, methylation of H3 lysine 9. This modification is also found in metazoan heterochromatin and is catalyzed by a conserved histone methylase called Clr4 in *S. pombe*, Su(var)3-9 in *Drosophila* and SUV39H1 in human.

In order for H3 lysine 9 to be methylated, it must be deacetylated. The *S. pombe* homologue of Sir2 plays a crucial role in this deacetylation. Another conserved protein called Swi6 in *S. pombe* and HP1 in metazoans binds specifically to nucleosomes with H3 methylated on lysine 9. Again, this protein is not found in budding yeast. The combination of lysine 9 methylation and Swi6/HP1 binding to nucleosomes with this modification spreads on the chromatin to create the silent heterochromatic regions. The nucleation point from which silent chromatin initiates and spreads in fission yeast is still under active investigation, but seems to involve repetitive DNA sequences and targeting of noncoding RNAs to them.

The 20 kbp silent region is flanked by inverted repeats that serve as boundary elements that isolate the silent heterochromatic domain. Studies of the histone modifications in and around the silent domain show a sharp demarcation at the boundaries. Within the silent domain, H3 lysine 9 is methylated whereas in the euchromatic regions it is acetylated. Also, H3 lysine 4 is acetylated in the transcriptionally active regions outside the boundary elements whereas it lacks this modification in the 20 kbp silent region.

In contrast to *S. cerevisiae*, the centromeres in *S. pombe* are very large (more than 100 kbp) and consist largely of repetitive sequences that are also heterochromatic. Centromeric heterochromatin is methylated on H3 lysine 9 and has HP1 bound to it, just as is the case at the silent mating locus.

Metazoans

As mentioned above, heterochromatin is found in all eukaryotes examined to date. In metazoans, the mechanism for silencing appears to be quite similar to that in *S. pombe*. In fact, the heterochromatin protein HP1 was first identified in *Drosophila*, and three different isoforms are found in mammals. Just as in *S. pombe*, HP1 binds to H3 methylated on lysine 9 in metazoans and is the fundamental building block of

heterochromatin. Interestingly, HP1 in *Drosophila* also binds to the N terminus of Orc1, just as Sir1 binds to that domain of *S. cerevisiae* Orc1. This is the case even though there is no sequence similarity between HP1 and Sir1, and between the Orc1 N termini of the two species. Thus, the function appears to have been conserved even though the sequence is not.

In summary, silent heterochromatic domains in metazoans appear to have the same histone modification (H3 lysine 9 methylation) and the same protein that binds to it (HP1) as is found in fission yeast. The mechanism for initiating heterochromatin formation in metazoans is not known yet.

In all eukaryotes, including budding yeast, heterochromatin is less acetylated on the histone N-terminal tails than is euchromatin. Presumably, this deacetylation is catalyzed by Sir2 homologues and other histone deacetylases. A common theme for silent heterochromatin in all species is the presence of specifically modified histones and proteins that bind uniquely to them.

SEE ALSO THE FOLLOWING ARTICLES

Chromatin: Physical Organization • DNA Replication: Eukaryotic Origins and the Origin Recognition Complex • Nuclear Organization, Chromatin Structure, and Gene Silencing • Telomeres: Maintenance and Replication

GLOSSARY

euchromatin Transcriptionally active chromatin, with a different structure than heterochromatin.
heterochromatin A specialized form of chromatin that is transcriptionally silent.
nucleosomes The basic building block of chromosomes, consisting of an octamer of histones H2A, H2B, H3, and H4 plus 146 base pairs of DNA.
Sir proteins Silent information regulator proteins that are required for silencing in the budding yeast, *S. cerevisiae*.
transcription RNA synthesis.

FURTHER READING

Gasser, S. M., and Cockell, M. M. (2001). The molecular biology of the SIR proteins. *Gene* **279**, 1–16.
Grewal, S. I. S., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. *Science* **301**, 798–802.
Moazed, D. (2001). Common themes in mechanisms of gene silencing. *Mol. Cell* **8**, 489–498.
Rusche, L. N., Kirchmaier, A. L., and Rine, J. (2003). The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* **72**, 481–516.

BIOGRAPHY

Ann Sutton is a Research Associate Professor and Rolf Sternglanz is a Distinguished Professor in the Department of Biochemistry and Cell Biology at Stony Brook University. Their research focuses on chromatin structure and function in yeast. Dr. Sutton received her Ph.D. in Molecular Biology from Stony Brook University and Dr. Sternglanz received his Ph.D. in Chemistry from Harvard University.

Transcription-Coupled DNA Repair, Overview

Isabel Mellon
University of Kentucky, Lexington, Kentucky, USA

Cells are continually exposed to a plethora of DNA-damaging agents formed within cells or present in the extracellular environment. To combat the harmful effects, cells possess an assortment of DNA repair pathways that recognize and remove damaged DNA. DNA can be structurally modified by the covalent addition of chemical adducts, the formation of crosslinks whereby two different bases on the same or opposite DNA strand become covalently linked, the introduction of UV light-induced photoproducts and an assortment of other alterations. Certain types of DNA damage inhibit transcription and pose blocks to RNA polymerase progression along the DNA template. Transcription-coupled repair (TCR) is a specialized feature of DNA repair that selectively removes transcription-blocking damage present in the transcribed strands of expressed genes.

Nucleotide Excision Repair

Nucleotide excision repair (NER) removes an assortment of different types of DNA damage. It removes chemical adducts introduced by exposure to chemical carcinogens and cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts produced by UV light. Given that this pathway removes structurally different types of lesions, it is likely that it recognizes the distortion of the DNA helix produced by the lesion rather than the lesion itself. It involves damage recognition, unwinding of the DNA at the lesion, two incisions, one on each side of the lesion, removal (excision) of a stretch of DNA containing the lesion, DNA synthesis to replace the excised DNA, and ligation of the newly synthesized DNA to the parental DNA. The general strategy has been conserved in *Escherichia coli* (*E. coli*), yeast, and mammalian systems.

TCR has been clearly demonstrated to be a subpathway of NER. It is usually measured as more rapid or more efficient repair in the transcribed strand of an expressed gene compared with the nontranscribed strand. This was first demonstrated studying the removal of UV light-induced CPDs from each strand of the DHFR gene in hamster and human cell lines. TCR of CPDs was subsequently demonstrated in *E. coli*

and yeast. In addition, certain bulky chemical lesions are substrates for TCR. Hence, this subpathway of NER has been conserved from bacteria to humans and operates on many different lesions. Repair in the nontranscribed strands of expressed genes and in unexpressed regions of the genome is referred to as global genome repair (GGR). Many of the same proteins are required for TCR and GGR. However, the two subpathways likely differ at the damage recognition step. For TCR, damage recognition is initiated by the stalling of RNA polymerase complexes at lesions in the transcribed strands of expressed genes. For GGR, damage recognition is initiated by other proteins.

E. COLI

NER in *E. coli* is understood in detail and has served as a paradigm for the investigation of other organisms. Damage recognition and processing is carried out by the UvrABC system. UvrA dimerizes (UvrA$_2$) and binds UvrB and the UvrA$_2$B complex binds DNA. The helicase activity of the complex may enable scanning for damage by translocation along the DNA and it unwinds DNA at the site of the lesion. UvrA$_2$ dissociates from the damaged site leaving an unwound preincision complex containing UvrB that is recognized and bound by UvrC. UvrBC produces an incision on each side of the lesion, both made by UvrC. UvrD unwinds and displaces the damaged oligonucleotide produced by the incisions. The resulting gap is filled in by DNA polymerase I and the repair patch is sealed by DNA ligase.

TCR in *E. coli* was first alluded to by studies of mutation frequency decline in tRNA operons. It was documented as more rapid removal of CPDs from the transcribed strand of the *lac* operon. Genetic and biochemical studies indicate that TCR and GGR require UvrA, B, C, and D. However, TCR also requires the mutation frequency decline (Mfd) protein. In addition, TCR in the *lac* operon requires transcription of the operon. CPDs in the transcribed strands of expressed genes pose blocks to RNA polymerase elongation while those in the nontranscribed strand are generally

bypassed. Hence, blockage of elongating RNA polymerase complexes at CPDs is an early step in TCR and the RNA polymerase complex and/or some feature of the transcription bubble likely play important roles.

Sancar and colleagues found that Mfd promotes the release of RNA polymerase complexes stalled at lesions in the transcribed strand of a gene expressed in a cell-free system. However, this provides somewhat of a conundrum in that, if the stalled polymerase complex or the transcription bubble is an important signal for TCR, presumably this signal is lost when the polymerase complex becomes displaced from the lesion. Recent studies have provided additional insights into possible mechanisms. First, certain lesions are bound more efficiently when present in "bubble" substrates and incision can occur in the absence of certain NER factors. In addition, bubble-like structures trigger the 3′ and 5′ endonuclease activities of UvrBC. Hence, it is likely that some aspect of the transcription bubble plays a key role in TCR. Second, a novel function for Mfd has been

recently defined by Parks and colleagues; it has the ability to reverse "backtracked" RNA polymerase complexes. Backtracking involves translocation of the RNA polymerase complex and the transcription bubble backward from the site of blockage. In fact, Hanawalt and colleagues proposed that RNA polymerase complexes backtrack at CPDs.

A model for TCR in *E. coli* is described in Figure 1 that incorporates backtracking and loading of NER factors onto the transcription bubble. The model is as follows: After UV-irradiation, RNA polymerase complex elongates until it encounters a CPD on the transcribed strand and stalls. The polymerase then translocates backwards. Mfd recognizes the backtracked complex and induces forward translocation of the polymerase until it re-encounters and perhaps even bypasses the lesion for a short distance. UvrA$_2$B or perhaps UvrB alone loads 5′ to the lesion (relative to the damaged strand). The loading of UvrB is facilitated by features of the transcription bubble brought about by
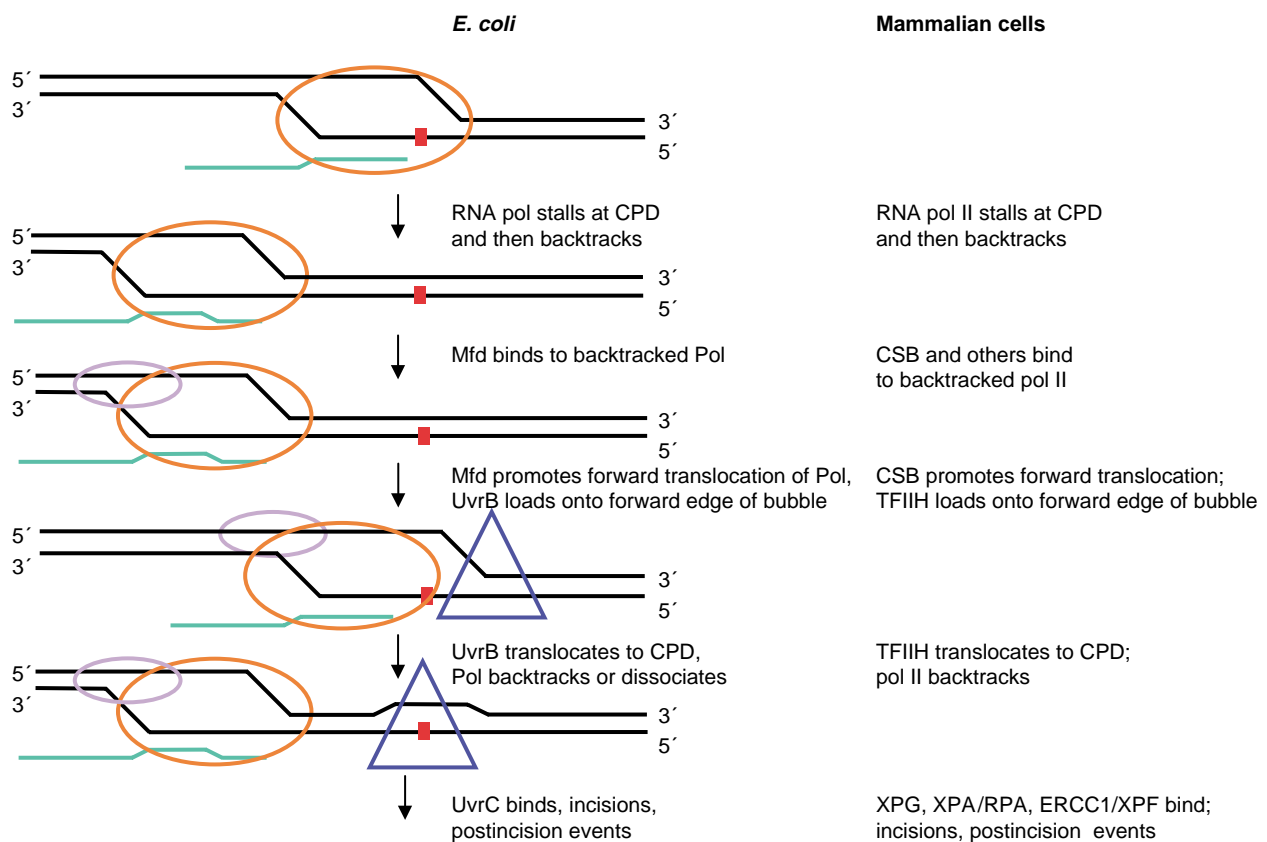
FIGURE 1 A model for TCR of CPDs. *E. coli*: Elongating RNA polymerase complex (orange oval) stalls at a CPD (small red square) in the transcribed strand. The polymerase complex, transcription bubble, and nascent RNA (green line) translocate backwards. Mfd (lavender oval) binds backtracked polymerase and DNA upstream of the bubble. Mfd promotes the forward translocation of the polymerase complex. UvrB (blue triangle) binds 5′ (relative to the damaged strand), loads onto the forward edge of the bubble, and translocates to the lesion. The polymerase complex backtracks or is dissociated by Mfd. Subsequent NER processing events continue as they would in nontranscribed DNA. Mammalian cells: Same as for *E. coli* except for the following: CSB (lavender oval) bind the backtracked polymerase complex and promote forward translocation. TFIIH (blue triangle) binds 5′ to the damage and loads onto the forward edge of the bubble. Subsequent NER events continue as they would in nontranscribed DNA.

the forward translocation induced by Mfd. At this point the polymerase may backtrack again or may be completely released by Mfd. UvrC then binds to the lesion-bound UvrB complex resulting in a stable preincision complex and this and subsequent downstream NER events continue as they would in nontranscribed DNA. The salient point is that the "coupling" of NER to transcription may be mediated by the correct positioning of the transcription bubble at the lesion. Mfd may serve two functions: one is to maintain the transcription bubble at the site of the lesion by reversing backtracked complexes and the other may be to ultimately displace the complex from the damaged site to allow incision and DNA synthesis.

Mammalian Cells

The general strategy of NER in mammalian systems closely parallels that of *E. coli*. However the repertoire of proteins required for mammalian NER is significantly more complex. There is considerable evidence that XPC/hHR23B complex is involved in an early step of damage recognition. TFIIH is then recruited which results in unwinding near the lesion by virtue of the helicase activities of XPB and XPD, components of TFIIH. XPG, XPA/RPA, and ERCC1/XPF assemble to form a stable preincision complex. Dual incisions are carried out; the 3′ incision by XPG and the 5′ incision by ERCC1/XPF, followed by postincision events. With the exception of XPC (and probably hHR23B) the same repertoire of proteins described above are required for TCR and GGR. It is likely that in TCR, the RNA polymerase complex replaces the function of XPC/hHR23B in damage recognition.

Genetic studies have indicated a requirement for additional genes in TCR. These include Cockayne syndrome group A and B (CSA and CSB) genes, genes involved in mismatch repair, UV-sensitive syndrome (UVSS), and XPA-binding protein (XAB2). Mutations in these genes result in a selective loss of TCR while repair in nontranscribed DNA is not effected or less effected. Biochemical studies have implicated the direct involvement of CSA, CSB, and XAB2 in TCR. As in *E. coli*, TCR in mammalian cells may be dependent upon the positioning of the transcription bubble at the lesion and not necessarily on direct interactions between NER proteins and transcription factors. CSA, CSB, TFIIH, and XAB2 may serve essential roles in remodeling the transcription bubble to facilitate TCR (Figure 1). Further investigation is required to determine their mechanism of action.

In mammalian cells TCR appears to be limited to genes transcribed by RNA pol II. The investigations of repair in genes transcribed by RNA polymerase I (pol I) and RNA polymerase III have found no evidence of TCR. However, the examination of ribosomal genes transcribed by pol I is complicated because only a subset of the genes are transcriptionally active. Recent studies by Smerdon and colleagues have fractionated active and inactive ribosomal genes in yeast and found TCR in the active fraction. Hence, TCR may not be limited to pol II genes in mammalian systems either and future studies are warranted to answer this important question.

Saccharomyces cerevisiae

The biochemical details of NER are not as well understood in *S. cerevisiae*. Genetic studies by several laboratories have demonstrated that *rad26* is required for TCR. More recent studies from the Smerdon laboratory have demonstrated that certain subunits of yeast RNA pol II, Rpb4, and Rpb9, also influence TCR. Deletion of the pol II subunit gene, *rpb9*, greatly reduces TCR of the GAL1 gene when cells are grown in log phase. Deletion of *rad26* greatly reduces TCR of GAL1 when cells are grown in stationary phase. In addition, deletion of a different pol II subunit gene, *rpb4*, restores TCR in the *rad26* mutant grown in stationary phase. Hence, in addition to providing novel information on the requirement of pol II in TCR, these studies also indicate that there are differences in TCR that depend on the growth state of the cell.

Base Excision Repair

Base excision repair (BER) represents a collection of repair pathways that operate on a variety of different lesions induced by oxidative damage, alkylation damage, and other types of damage. The broad substrate specificity is accomplished by a large number of different damage-specific glycosylases. Hence, this differs from NER where the broad substrate specificity is accomplished by assembling a multiprotein complex.

Alkylation Damage

Alkylating agents represent a broad class of DNA-damaging agents that are present in the environment, are used as chemotherapeutic agents and can be formed endogenously during cellular metabolism. N-methylpurines (NMPs) are the most abundant lesions produced by simple alkylating agents such as methyl methanesulfonate and dimethyl sulfate. 7-methylguaine and 3-methyladenine are the most abundant NMPs formed by these agents. NMPs are removed by BER in *E. coli*, yeast, and mammalian cells and repair is initiated by specific glycosylases. The removal of NMPs has been compared in the transcribed and nontranscribed strands of the DHFR gene in mammalian cells, the GAL1 gene in *S. cerevisiae*, and the lactose operon of *E. coli*. No significant difference was found in the repair of the transcribed and nontranscribed strands of these genes.

Hence, TCR does not appear to be a subpathway of methylation damage-specific BER.

OXIDATIVE DAMAGE

Oxidative damage is formed as a consequence of exposure to ionizing radiation and a variety of chemical agents and as byproducts of normal cellular metabolism. These agents introduce a large number of modifications to DNA including alterations of bases, the deoxyribose sugar, and cleavage of the phosphodiester backbone. Several studies have found more rapid removal of oxidative damage from the transcribed strands of

expressed genes in yeast and mammalian cells. Hence, TCR operates on oxidative damage. It has been proposed that this finding indicates that TCR is also a subpathway of BER since many forms of oxidative damage are substrates for BER. However, there has been no direct genetic or biochemical demonstration of a role of BER in TCR. Ionizing radiation and other forms of oxidative agents produce a wide spectrum of different lesions, including some that are substrates for NER. Moreover, the Cooper, Clarkson and Leadon laboratories have found that TCR of oxidative damage is abolished or significantly reduced in human cell lines with certain mutations in the NER genes XPG, XPB, and

RNA pol II stalls at oxidative lesion and then backtracks

CSB binds to backtracked pol II

CSB promotes forward translocation of pol II, TFIIH loads onto forward edge of bubble

TFIIH translocates to oxidative lesion, pol II backtracks or dissociates

XPG binds, incision is made adjacent to damage

XPG dissociates creating a flapped structure, the flapped structure is incised creating a small gap, the gap is filled in by repair synthesis

FIGURE 2 A model for TCR of oxidative damage. Elongating RNA pol II complex (orange oval) stalls at an oxidative lesion (small red square) in the transcribed strand. The pol II complex, transcription bubble, and nascent RNA (green line) translocate backwards. CSB (lavender oval) binds backtracked polymerase and promotes the forward translocation of the pol II complex. TFIIH (blue triangle) binds 5′ (relative to the damaged strand), loads onto the forward edge of the bubble, and translocates to the lesion. The polymerase complex backtracks. XPG binds (green square) and makes an incision at the oxidative lesion creating a flapped structure. The flapped structure is incised creating a small gap and the gap is filled in by repair synthesis.

XPD. XPG is involved in the incision process of NER and XPB, and XPD are components of TFIIH and unwind the helix at the site of damage. In addition, TCR of oxidative damage is abolished or reduced in human cell lines with mutations in CSA and CSB. The CSA and CSB genes are clearly required for TCR mediated by the NER pathway as described above.

A model for TCR of oxidative damage is described in Figure 2 that involves the RNA polymerase complex and components of the NER pathway. RNA polymerase II stalls at the oxidative lesion and then backtracks. CSB, the functional homologue of Mfd, promotes forward translocation of pol II. TFIIH translocates to the oxidative damage, unwinds the DNA as pol II backtracks. XPG incises the oxidative damage present in the unwound DNA creating a flapped structure. The flapped structure is incised 5′ to the damage creating a small gap that is filled in by DNA repair synthesis. As in the model for TCR of UV damage, the salient point is that the coupling of repair of oxidative damage to transcription may be mediated by the correct positioning of the transcription bubble at the lesion.

RNA Polymerase Turnover and Degradation

An interesting question from a teleologic viewpoint relates to why cells possess mechanisms that couple DNA repair and transcription. One reason may be that TCR serves to remove transcription-blocking lesions and hence, it facilitates a rapid recovery of transcription. However, transcription complexes are extremely stable when they are stalled at endogenous pause sites or at sites of damage. In the absence of a mechanism to specifically find transcription-blocking lesions, lesions would be shielded from the repair machinery by the RNA polymerase complex and hence, refractory to repair. Furthermore, stable arrested complexes would inhibit gene expression and perhaps interfere with or block the DNA replication machinery. Recent studies have found that RNA pol II complexes are degraded in response to DNA damage. Svejstrup has suggested that degradation of damage-stalled pol II complexes might be an alternative to TCR. Hence, the importance of removing stalled RNA polymerase complexes may be indicated by the development of specific repair mechanisms that remove transcription-blocking damage and if TCR fails to occur, then the RNA polymerase complex stalled at the damaged site may be actually degraded.

SEE ALSO THE FOLLOWING ARTICLES

Cell Cycle: DNA Damage Checkpoints • DNA Base Excision Repair • DNA Damage: Alkylation • DNA Mismatch Repair and Homologous Recombination • DNA Mismatch Repair and the DNA Damage Response • DNA Mismatch Repair: *E. coli* Vsr and Eukaryotic G–T Systems • DNA Mismatch Repair in Mammals • Nucleotide Excision Repair and Human Disease • Nucleotide Excision Repair, Bacterial: The UvrABCD System • Nucleotide Excision Repair: Biology • Nucleotide Excision Repair in Eukaryotes • RNA Polymerase I and RNA Polymerase III in Eukaryotes

GLOSSARY

cyclobutane pyrimidine dimer   The covalent linkage of two adjacent pyrimidines in DNA produced by exposure to ultraviolet light.

glycosylase   An enzyme that cleaves the N-glycosylic bond between a damaged base or inappropriate base and the deoxyribose sugar.

transcription bubble   The unwound DNA structure produced by RNA polymerase in the elongation mode of RNA synthesis.

FURTHER READING

Batty, D. P., and Wood, R. D. (2000). Damage recognition in nucleotide excision repair of DNA. *Gene* **241**, 193–204.

Copper, P. K., Nouspikel, T., Clarkson, S. G., and Leadon, S. A. (1997). Defective transcription-coupled repair of oxidative base damage in Cockayne syndrome patients form XP group G. *Science* **275**, 990–993.

Li, S., and Smerdon, M. J. (2002). Rpb4 and Rpb9 mediate sub-pathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *EMBO* **21**, 5921–5929.

Moolenaar, G. F., Monaco, V., van der Marel, G. A., van Boom, J. H., Visse, R., and Goosen, N. (2000). The effect of the DNA flanking the lesion on formation of the UvrB-DNA preincision complex. *J. Biol. Chem.* **275**, 8038–8043.

Park, J.-S., Marr, M. T., and Roberts, J. W. (2002). *E. coli* transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* **109**, 757–767.

Sugasawa, K., Ng, J. M. Y., Masutani, C. S. I., van der Spek, P. J., Eker, A. P. M., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. H. J. (1998). Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Mol. Cell* **2**, 223–232.

Svejstrup, J. Q. (2002). Mechanisms of transcription coupled DNA repair. *Nat. Rev.* **3**, 21–29.

Tornaletti, S., Reines, D., and Hanawalt, P. C. (1999). Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. *J. Biol. Chem.* **274**, 24124–24130.

Zou, Y., Luo, C., and Geacintov, N. E. (2001). Hierarchy of DNA damage recognition in *Escherichia coli* nucleotide excision repair. *Biochemistry* **40**, 2923–2931.

BIOGRAPHY

Isabel Mellon is an Associate Professor at the University of Kentucky, Lexington. Her principal research interests are in the field of DNA repair and include transcription-coupled repair and the roles of genetic alterations in DNA repair genes in the etiology of cancer. She holds a Ph.D. from the University of Illinois at Chicago. She was a postdoctoral fellow at Stanford University where she co-discovered transcription-coupled repair with Philip Hanawalt and colleagues.

Transforming Growth Factor-β Receptor Superfamily

Mark de Caestecker
Vanderbilt University, Nashville, Tennessee, USA

The transforming growth factor-β (TGF-β) superfamily consists of a large family of related growth factors. These can be divided into two main groups: the TGF-β/activin and bone morphogenetic protein (BMP)/growth and differentiation factor (GDF), and subdivided into several related subgroups based on their sequence similarity. Ligands are synthesized as precursor molecules that undergo cleavage, releasing the pro-domain from the active, receptor-binding, carboxy-terminal region of the molecule. The active, carboxy-terminal domains of these ligands have six intra-strand disulfide bonds that form a "cysteine knot" motif. A conserved seventh cysteine is required to form covalently linked dimeric structures that interact with their respective receptors. TGF-β is synthesized as an inactive precursor, cleaved into mature TGF-β and the latency associated peptide, LAP, which is then noncovalently linked to the mature TGF-β and prevents binding to the TGF-β receptors. Other members of the TGF-β superfamily are secreted as mature, active dimers that are inhibited locally through interactions with a variety of secreted antagonists including follistatin, noggin, chordin, and the DAN/cerberus family of proteins.

The TGF-β receptor superfamily have three-finger toxin folds in the ligand-binding extracellular domain, a single transmembrane and an intracellular serine–threonine kinase domain. These are divided into two main groups, the type I and type II receptors, based largely on sequence conservation within their kinase domains. In addition, the type I receptors have a conserved glycine-serine rich juxta-membrane domain (the GS box) that is critical for their activation. The nomenclature for the type I and II receptors is somewhat confusing. For simplicity I refer to the type I receptors by a common nomenclature, the activin-like kinases (Alks), and the type II receptors according to their dominant ligand interactions. Type I receptors are classified according to sequence similarities into three main groups: the Alk5 group which includes the TGF-β type I receptor Alk5, the activin receptor Alk4 and the nodal receptor Alk7; the Alk3 group comprising the BMP type I receptors Alk3 and 6; and the Alk1 group Alk1 and 2, which interact with different BMP/GDF and TGF-β/activin family ligands. Type II receptors include the TGF-β type II receptor, TGF-β RII, the activin and BMP/GDF receptors, activin RII and activin RIIB, the BMP/GDF receptor

BMP RII, and the Mullerian inhibitory substance (MIS) receptor MIS RII. The purpose of this review is to describe the biochemical properties of the TGF-β superfamily receptors in mammalian cells, focusing specifically on receptor–ligand interactions, accessory receptors, and receptor activation. The reader is referred to a series of reviews for a more detailed description of the functional properties of this these events *in vivo*.

Receptor–Ligand Interactions

The basic paradigm by which TGF-β superfamily ligands (Figure 1) interact with and activate the receptors has been largely established from studies of TGF-β. TGF-β binds to the constitutively active TGF-β type II receptor, TGF-β RII, which then recruits the type I receptor, Alk5, resulting in transphosphorylation of the type I receptor and activation of downstream signals. A single dimeric ligand interacts with two, type I and two type II receptors to form heterotetrameric signaling complex. A similar mechanism occurs with activin binding to the type II receptor Act RII and recruiting the type I receptor Alk4, and the BMP-ligands BMP6 and 7 interacting with the type II receptors Act RII and IIB and recruiting the type I receptors Alk2, 3, and 6. In contrast, receptor-binding affinity is reversed with BMP2 and 4, which preferentially bind to the type I receptors Alk3 and 6 and recruit type II receptors into heteromeric signaling complexes. These ligand–receptor interactions are inhibited through direct interactions with secreted antagonists including noggin, chordin, follistatin, and the DAN/cerberus family of proteins (Figure 2). Overlapping receptor usage by different TGF-β family ligands adds to the complexity of this system. For example, TGF-β itself has the capacity to interact with and activate at least two additional type I receptors Alk1 and 2, in the presence of TGF-β RII, while the related BMP-ligands BMP2 and 4, have the capacity to interact with the type I receptors Alk3 and 6 and recruit type II receptors BMP RII/Act RII and Act RIIB into active receptor complexes. While the
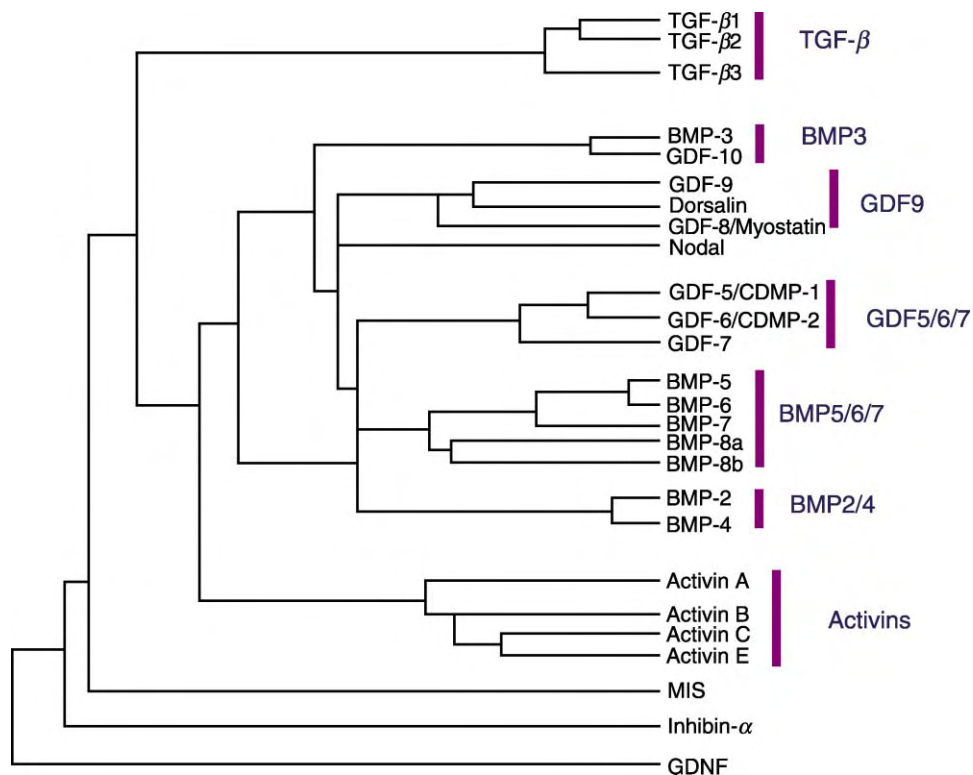
FIGURE 1   TGF-β superfamily of ligands.

functional significance of many different receptor–ligand combinations are poorly understood, the complexity of this combinatorial system is likely to explain the diversity of downstream responses that can be generated by engagement of these receptors in different cell types.

Accessory Receptors

Other cell surface proteins interact with and may be required for signaling. Some of these have been cloned and characterized and will be discussed in this article, while others have been identified from ligand–receptor cross-linking studies in certain cell types and are of uncertain significance. The first of these to be described in detail was the high-molecular-weight type III TGF-β receptor, TGF-β RIII or betaglycan. TGF-β RIII is a ubiquitous, highly glycosylated transmembrane protein with a large extracellular domain and short cytoplasmic tail lacking kinase activity. It was initially thought to function solely as a high-affinity TGF-β receptor, promoting TGF-β ligand binding to the signaling receptors, and is required for binding of TGFβ2 with TGF-β RII. However it is now known that TGF-β RIII is also a coreceptor for inhibin, promoting interactions between inhibin and other type II receptors, and giving rise to a functional inhibition of activin and BMP-dependent signaling. More recently it has been shown

that the cytoplasmic tail of TGF-β RIII is required to support TGF-β2-signaling but is not required to promote binding of TGF-β2 to the signaling receptor complex. This suggests that it plays an additional role in regulation of TGF-β signaling. The mechanisms underlying these effects are uncertain, although it has been proposed that in the presence of ligand, the cytoplasmic tail of TGF-β RIII selectively interacts with activated, autophosphorylated TGF-β RII, enabling enrichment of active TGF-βRI/II complexes.

Endoglin is a related, membrane-associated disulfide-linked dimer that was originally identified as a TGF-β1 and β3-binding protein in endothelial cells. It is now known that endoglin interacts with diverse TGF-β family ligands including BMP2, 7, and activin, and that it is selectively expressed in other cell types. Unlike betaglycan, ligand binding occurs indirectly through association with the respective type I, and II receptors. Furthermore, while endoglin and betaglycan form heteromeric complexes in endothelial cells, comparison between endoglin and betaglycan overexpressing cells indicate that endoglin can inhibit while betaglycan enhances TGF-β responsiveness. This suggests that the two accessory receptors may have distinct functional properties.

Recently, another family of accessory receptors has been identified that are required for TGF-β superfamily signaling. Cripto and cryptic are two members of the EGF-CFC family of membrane-anchored proteins.
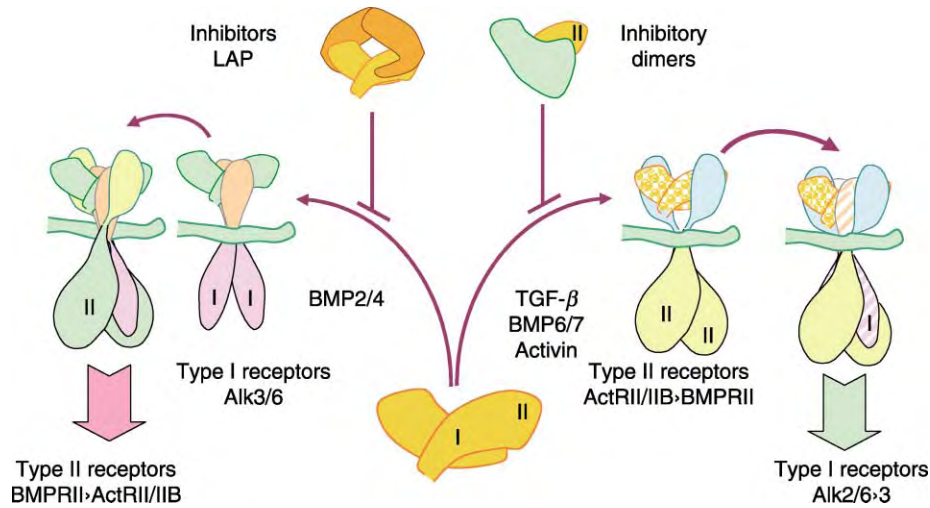
FIGURE 2 TGF-β superfamily receptor-ligand interactions. Ligands are secreted as disulfide-linked dimers, each presenting distinct hydrophobic type I and type II receptor interacting surfaces. Different ligands show variable affinities for type I and II receptors, promoting assembly of distinct hetero-tetrameric receptor–ligand complexes and downstream signals. These receptor interactions are inhibited through the formation of inhibitory complexes (e.g., LAP/TGF-β and noggin/BMP7), and/or heteromeric dimeric that block receptor–ligand interacting surfaces (e.g., inhibin-α/activin-βA subunits of inhibin A).

Genetic and biochemical evidence indicate that both proteins are required for signaling by the TGF-β family members, nodal and GDF1. These interact directly with nodal, GDF1, and the type I receptor, Alk4, and are required for the assembly of Alk4/activin type II receptor signaling complexes following engagement of these ligands. Interestingly, activin-dependent assembly of the ActRII and ActRIIB/Alk4 complexes are inhibited by overexpression of cripto, suggesting that EGF-CFC proteins may exert opposing effects on the assembly of different ligand-dependent TGF-β receptor complexes.

Receptor Activation

Activation of type I TGF-β receptors results from serine phosphorylation by the type II receptor kinase within the GS box immediately upstream of the catalytic domain. These phosphorylation events are associated with a conformational change in the GS box which forms an inhibitory wedge in the kinase domain of the inactive type I receptor, enabling ATP binding and phosphorylation of the downstream substrates, the receptor activated Smads (R-Smads). Basal activation of the type I receptor kinase domain is regulated through the interaction of a repressor protein, FKBP12, which binds to the unphosphorylated GS box, capping the type II receptor phosphorylation sites and stabilizing the receptor in an inactive conformation. In addition, ligand-dependent phosphorylation of the R-Smads is inhibited by direct competition for receptor binding by the inhibitory Smads 6 and 7,

which are themselves transcriptionally regulated by diverse signaling pathways.

Phosphorylation of the R-Smads activates the Smad signaling pathway, resulting in the nuclear translocation of an R-Smad/Smad4 complex. This regulates transcriptional responses through direct interaction with both *cis-* and *trans-* activating elements associated with a variety of different gene targets. Two main groups of R-Smads are activated by different sets of type I receptors and activate distinct downstream responses. The BMP receptor activated Smads 1, 5, and 8, and the TGF-β/activin activated Smads, Smad2 and 3. A variety of proteins have been identified that interact with both the receptor complexes and the respective R-Smads, and are thought to function as Smad-anchors and/or chaperones involved in the recruitment of Smad proteins to and from the receptor complex. These include SARA, Hgs, Axin, ELF, Disabled-2, TRAP-1, FTLP, and SANE. The most extensively characterized of these is the FYVE domain protein, Smad anchor for receptor activation (SARA), which recruits Smad2 to the activated receptor complex and is required for receptor-mediated Smad phosphorylation.

Specificity of these responses is determined by a cluster of residues within the L45 loop of the type I receptor kinase domain which interact with a matching set of residues in the L3 loop in the carboxy-terminal domains of the R-Smads. The L45 loop sequence of the Alk5 group of receptors are compatible with L3 loop residues in TGF-β/activin activated Smads 2 and 3, while the Alk3 group sequence is compatible with BMP-activated Smads 1, 5, and 8. The Alk1 group of receptors, which are activated both by TGF-β and

BMP-ligands, carry a cluster of L45 loop sequences that are distinct from the Alk3 group of receptors but are compatible with the BMP-activated Smads 1, 5, and 8. As different ligands can recruit different type I receptors into the signaling complexes, differential activation of Smad substrates by Alk1 and 5 groups of receptors may account for the activation of distinct downstream signals by the same ligand under different conditions.

Endocytic trafficking of activated receptors plays an important role in regulating receptor-dependent signaling. This may result from the activation of signaling by promoting association between activated receptors and various signaling intermediates in the endosomal compartments, or by dampening receptor-dependent signaling through degradation of the activated receptor complexes. Recent data suggests both mechanisms can regulate TGF-β receptor-mediated signaling. In the presence of R-Smads, the FYVE domain protein SARA promotes uptake of the SARA/Smad/receptor complex to early endosomal membranes, promoting clathrin-dependent endocytosis, receptor-mediated phosphorylation and activation of the R-Smad. In contrast, Calveolin-dependent endocytosis and proteosomal degradation of the activated receptor complex is mediated by recruitment of the HECT domain E3 ligase SMURF to the activated type I receptor by the inhibitory Smad7. This results in the uptake of the SMURF/Smad7/receptor complex to Calveolin positive endosomes, and promotes SMURF-mediated ubiquitination and proteosomal degradation of the signaling complex.

In addition to the canonical Smad-signaling pathways, there is considerable evidence of cell-type-dependent regulation of alternative signaling pathways by different members of the TGF-β superfamily. These include activation of MAPK signaling pathways ERK, JNK and p38 MAPK, PI3 kinase, PKC, and inhibition of p70^{S6K} signaling. These signals may be required to mediate and/or augment maximal Smad-dependent responses, or they may exert distinct downstream responses in different cell types. However, the precise mechanisms linking these noncanonical signaling pathways with the activated receptor complexes are incompletely understood. Some of these effects may be indirect, resulting from the induction of Smad-dependent target genes that themselves regulate these responses. For example, while JNK activation by TGF-β may result from activation of a rapid, Smad-independent response in a variety of cell types, a delayed Smad-dependent response has also been described. In other cell types, TGF-β signaling may repress p38 MAPK signaling through the induction of Smad-dependent MAPK phosphatases expression. In contrast, direct regulation of these alternative pathways may result from ligand-dependent protein interactions with the activated receptor complex. For example, interactions between the serine threonine phosphatase, PP2A and the activated TGF-β type I/II receptor complex, results in rapid, dephosphorylation-dependent inhibition of p70^{S6K} activity. Other downstream intermediates may link receptor activation to these pathways. These include the Rho family of small GTPases and the MAPKKK TAK1 and its upstream kinase HPK1, that are involved in regulating TGF-β-dependent activation of JNK, Ras GTPase-dependent activation of ERK MAPK, and TAK1-dependent activation of p38 MAPK by TGF-β. Direct links between these pathways and the activated receptors are less clear.

In addition to these effects, activation of TGF-β family receptors gives rise actin polymerization and cytoskeletal re-organization in a variety of different cell types. While the long-term effects of TGF-β on these responses may be indirect, rapid changes in cytoskeletal organization can involve direct modification of small GTPase-dependent actin polymerization by TGF-β receptor signaling. More recently, LIM kinase 1 (LIMK1), which regulates actin polymerization through phosphorylation-dependent inactivation of Cofilin, has been shown to interact with the cytoplasmic tail of the BMP type II receptor. BMP4-treatment is associated with re-distribution of LIMK1 to the cell periphery, phosphorylation of Cofilin, and active re-organization of the actin cytoskeleton. This suggests that BMP RII-dependent activation of LIMK1 may regulate additional signaling pathways, and raises additional questions regarding the role of the C-terminal tails of TGF-β family receptors in the regulation of downstream signaling.

Conclusions

This review outlines the large number of TGF-β ligands, the plasticity of receptor usage and the diversity of downstream pathways that can be regulated by these receptors. However, while considerable advances have been made in our understanding over the last decade, many of these discoveries have raised many new questions about the complex biology of these receptor-signaling pathways. Important areas of research in the future will provide a more detailed understanding of the factors contributing to the diversity of cell-type-dependent responses and links to other signaling pathways.

SEE ALSO THE FOLLOWING ARTICLES

Mitogen-Activated Protein Kinase Family • p70 S6 Kinase/mTOR • Protein Kinase C Family • Ras Family • Rho GTPases and Actin Cytoskeleton Dynamics

GLOSSARY

activins and inhibins   Proteins released from the gonads that stimulate and inhibit, respectively, the secretion of follicle-stimulating hormone, which is secreted by the pituitary and is a major regulator of reproductive function.

LIM kinase   A protein serine kinase that phosphorylates and inactivates cofilin, a protein that regulates actin depolymerization and hence influences cytokinesis (cell division), endocytosis (uptake of molecules by cells), chemotaxis (directed cell movement) and morphogenesis (cell shape change).

Smads   Proteins that are phosphorylated by transforming growth factor receptor family members and move as complexes into the nucleus to activate gene transcription.

transforming growth factors β   A secreted protein that acts locally to either stimulate or inhibit cell proliferation or differentiation, and which plays a role in development and wound healing.

FURTHER READING

Balemans, W., and Van Hul, W. (2002). Extracellular regulation of BMP signaling in vertebrates: A cocktail of modulators. *Dev. Biol.* **250**(2), 231–250.

Chang, H., Lau, A. L., and Matzuk, H. H. (2001). Studying TGF-beta superfamily signaling by knockouts and knockins. *Mol. Cell Endocrinol.* **180**(1–2), 39–46.

de Caestecker, M. P., Piek, E., and Roberts, A. B. (2000). Role of transforming growth factor-beta signaling in cancer. *J. Natl. Cancer Inst.* **92**(17), 1388–1402.

Kawabata, M., and Miyazono, K. (2000). *Skeletal Growth Factors*. Lippincott Williams and Wilkins, Philadelphia.

Massague, J., Blain, S. W., and Lo, R. S. (2000). TGFbeta signaling in growth control, cancer, and heritable disorders. *Cell* **103**(2), 295–309.

Miyazono, K., Kusanagi, K., and Inoue, H. (2001). Divergence and convergence of TGF-beta/BMP signaling. *J. Cell. Physiol.* **187**(3), 265–276.

Pangas, S. A., and Woodruff, T. K. (2000). Activin signal transduction pathways. *Trends Endocrinol. Metab.* **11**(8), 309–314.

Piek, E., Heldin, C. H., and Ten Dijke, P. (1999). Specificity, diversity, and regulation in TGF-beta superfamily signaling. *Faseb J.* **13**(15), 2105–2124.

Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113**(6), 685–700.

Zhao, G. Q. (2003). Consequences of knocking out BMP signaling in the mouse. *Genesis* **35**(1), 43–56.

BIOGRAPHY

Mark de Caestecker is an Associate Professor with the Departments of Medicine and Cell and Developmental Biology at Vanderbilt University. He graduated with a medical degree from the Universities of Cambridge and London, specializing in general internal medicine and nephrology, and Ph.D. from the University of Manchester in England. He began working on TGF-β signaling as a postdoctoral fellow with Anita Roberts in the Laboratory of Cell Regulation and Carcinogenesis at the NIH. His current research interests include the role of BMP-signaling in vascular remodeling and the regulation of epithelial cell fate in renal development and injury.

Translation Elongation in Bacteria

Oliver Vesper and Knud H. Nierhaus

Max-Planck-Institut für Molekulare Genetik, Berlin, Germany

Protein synthesis is one of the major processes in a living cell that translates the genetic information into protein structure and thus organizes and directs the life cycle and metabolism of a cell. The process of protein synthesis can be subdivided into four consecutive phases: (1) initiation, (2) elongation, (3) termination, and (4) ribosome recycling. All show features that are specific for each of the three main evolutionary domains, viz., bacteria, archaea, and eukarya, with only one exception, the elongation phase. This phase is at the heart of protein synthesis, where the codon sequence of an mRNA is translated into the corresponding amino acid sequence of proteins.

Introduction

Ribosomes translate the genetic information of mRNAs by using tRNA as adaptors. An acylated tRNA connects the decoding center on the small ribosomal subunit via the anticodon at the tip of the long arm of the L-shaped tRNA with the peptidyl-transferase (PTF) center on the large ribosomal subunit via its short arm, the amino-acid acceptor stem.

All ribosomes examined to date from all three evolutionary domains show three tRNA binding sites: (1) The A site, where the correct aminoacyl-tRNA is selected according to the codon present here. The A site tRNA binds in the form of a ternary complex (aa-tRNA·EF-Tu·GTP; aa, aminoacyl; EF-Tu, elongation factor Tu), thus providing the new amino acid for the growing peptide chain. (2) The P site, where the peptidyl-tRNA is located carrying the nascent peptide chain before peptide bond formation. And (3) the exit site (E site) that binds exclusively uncharged (deacylated) tRNAs. It is from this site that the tRNA is released from the ribosome. During the course of three elongation cycles a tRNA enters the ribosome at the A site, moves through the P site and leaves the ribosome from the E site. The only exception is the very first tRNA, termed initiator tRNA, which binds directly to the P site and selects the first codon to be translated thus determining the reading frame.

The Three Basic Reactions of Elongation

Elongation of the nascent peptide chain by one amino acid is performed in a cyclic manner, the sequence of reactions is termed elongation cycle. An overview of the elongation cycle is shown in Figure 1, where the three basic reactions are depicted: (1) A site occupation, (2) peptide bond formation, and (3) the translocation reaction.

A site occupation is separated into two subreactions: in the first step the correct (or cognate) ternary complex aa-tRNA·EF-Tu·GTP is selected via codon–anticodon interaction before the aa-tRNA fully occupies the A site. Successful decoding (decoding reaction) is sensed by the ribosome and leads to an as yet undefined conformational change within the ribosome that triggers hydrolysis of GTP to GDP by elongation factor Tu (EF-Tu). EF-Tu is a G protein, i.e., it can bind a GTP molecule and is now in its "on" conformation, where EF-Tu·GTP can bind an aa-tRNA thus forming the ternary complex. After the ternary complex has delivered its aa-tRNA to the ribosomal A site, the ribosome triggers the activation of the GTPase center on EF-Tu, the resulting EF-Tu·GDP snaps into the "off" conformation and falls from the ribosome.

In the second so-called accommodation step, the release of EF-Tu·GDP from the ribosome allows the tRNA to swing into the A site docking the aminoacyl-residue of the aa-tRNA into the PTF center of the 50S subunit. The aminoacyl-tRNA now occupies the A site and is ready to accept the peptidyl-moiety from the peptidyl-tRNA present at the adjacent P site. With A and P sites occupied with tRNAs the ribosome is in the so-called pretranslocational (PRE) state, although the ribosome is not yet ready for translocation. This is the case only after the next reaction.

In the second reaction peptidyl transfer occurs. The peptidyl residue from the donor (P site) is linked to the aminoacyl residue of the acceptor (A site) via a peptide bond, forming a peptidyl-tRNA at the A site (elongated by one amino acid) and leaving a deacylated tRNA at the P site. Note that the ribosome is still in the PRE state
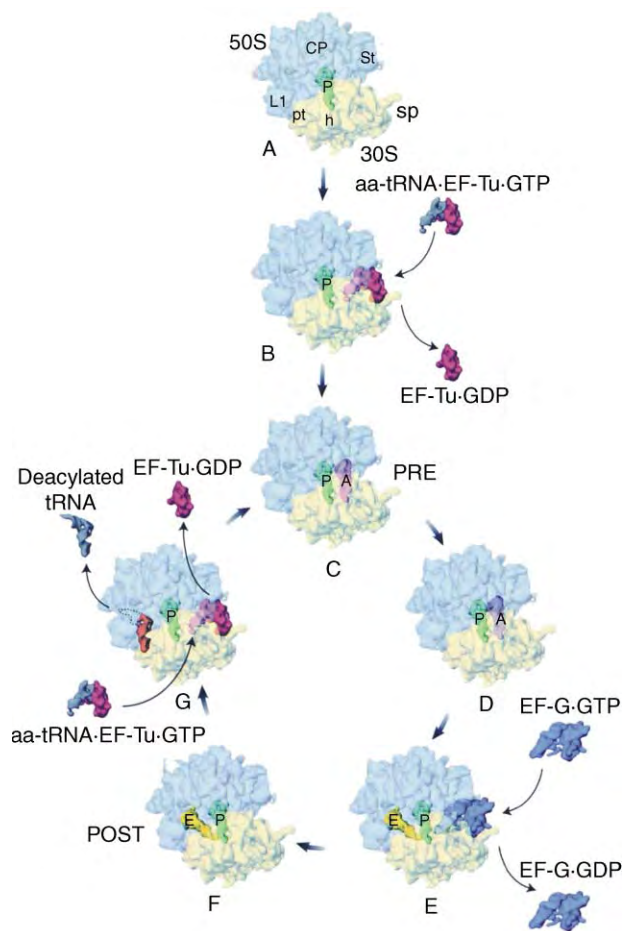
FIGURE 1 Overview of the translational elongation cycle. Multiple cryo-EM studies determined the tRNA and elongation factor binding positions on the 70S ribosome from *E. coli* during the different stages of the elongation cycle. The schematic view of the elongation cycle starts with an initiation complex (A) with the initiator fMet-tRNA in the P site that represents the last stage of initiation. A ternary complex aa-tRNA·EF-TU·GTP enters the vacant A site and after decoding and GTP hydrolysis the binary complex of EF-Tu·GDP leaves the ribosome (B). The A-site aa-tRNA is accommodated into the A site and a pretranslocation complex (PRE state) is formed that is characterized by occupied A and P sites (C). At this stage the peptidyl residue is linked to the aminoacyl-tRNA via a peptide bond. The result is a deacylated tRNA at the P site and a peptidyl-tRNA – prolonged by one aminoacyl-residue – at the A site. The tRNA positions do not change after peptide-bond formation (D). In the next step EF G·GTP binds to the PRE complex and facilitates the translocation of the A and P site tRNAs to the P and E sites, respectively (E). After hydrolysis of GTP EF-G·GDP dissociates from the ribosome, the ribosome is now in the posttranslocational state (POST state) (F). The POST complex is ready for the newly incoming aminoacyl-tRNA coming as ternary complex (aa-tRNA·EF-Tu·GTP). After decoding and GTP hydrolysis the binary complex of EF-Tu·GDP leaves the ribosome, the deacylated tRNA from the E-site is released via a reciprocal coupling between A and E sites and the PRE complex is formed (G → C). (Adapted from Agrawal, R. K., Spahn, C. M. T., Penczek, P., Grassucci, R. A., Nierhaus, K. H., and Frank, J. (2000). Visualization of tRNA movements on the *Escherichia coli* 70S ribosome during the elongation cycle. *J. Cell. Biol.*, **150**, 447–459.)

(A and P sites are occupied), and that the tRNAs have not changed their position after peptide-bond formation.

In the third reaction, the translocation reaction, tRNAs are moved from the A and P site to the P and E sites, respectively, shifting the ribosome into the post-translocational state (POST state, P, and E sites are occupied). The A site is vacant and ready for receiving the next incoming ternary complex. This movement of the tRNA$_2$·mRNA complex within the ribosome by one codon length is facilitated by elongation factor G (EF-G), which is also a G protein with its own GTPase center.

Functional Models of the Elongation Cycle

Various RNA modifying techniques have been used to probe the interactions of tRNAs within the ribosome and to identify tRNA-related functional centers on the ribosome. Distinct sets of rRNA bases have been assigned to contact tRNA in each of the classical binding site (A, P, or E site), which could be explained in the light of high-resolution structure as resulting from either direct contact between the tRNA and bases of rRNA or local conformational changes of the binding regions. In particular, the movement of tRNAs through functional sites during a single elongation was examined using foot-printing techniques. Two different approaches have been applied; interestingly both led to distinct models of the elongation cycle, although they are not mutually exclusive. The hybrid-site model proposed by Noller and colleagues is based on the protections of rRNA bases from chemical modification (kethoxal, dimethylsulfate (DMS) or hydroxyl radicals) by ribosome bound tRNA.

Nierhaus and co-workers applied the phoshorothio-ate technique for tRNA leading to the $\alpha-\varepsilon$ model of the elongation cycle. The latter model focuses not only on the path of tRNA through the ribosome but also on mechanistic features of the ribosome associated with decoding and maintenance of the reading frame.

THE HYBRID SITE MODEL

The essence of this model (Figure 2A) is a creeping movement of tRNAs through the ribosome. The diagnostic feature of this model is the movement of the tRNAs exclusively on the large subunit after peptide-bond formation and before translocation, while the 30S bound part of the tRNAs remain in the same site. This results in a hybrid site: The peptidyl tRNA moves after peptide-bond formation from an A/A site to an A/P site, and the deacylated tRNA from P/P site to P/E (the site before the slash indicates the site on the small subunit,
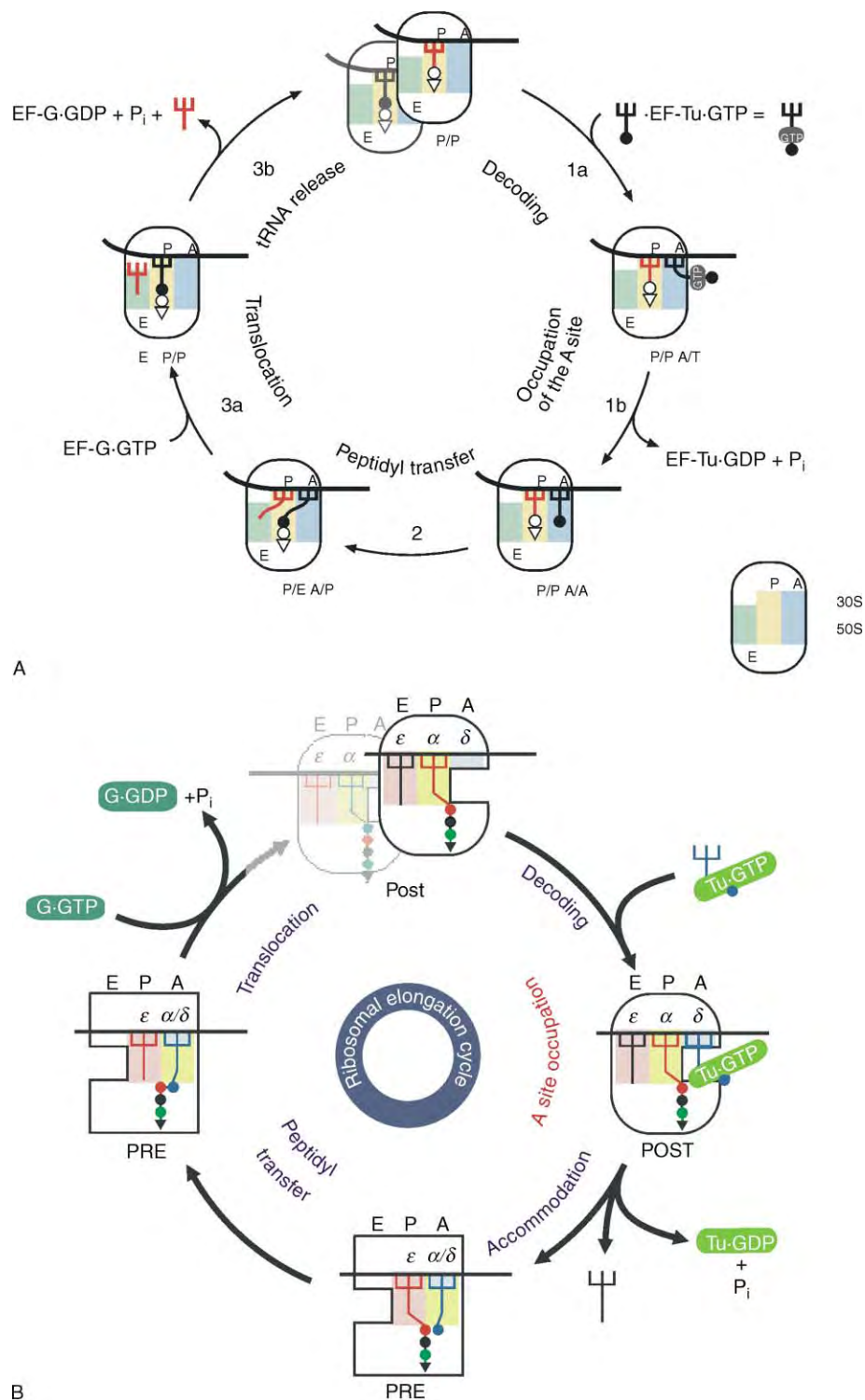
FIGURE 2 Models of the elongation cycle. (A) The hybrid-site model according to Moazed, D., and Noller, H. F. (1989). Intermediate states in the movement of transfer RNA in the ribosome. *Nature* **342**, 142–148; for explanation see text. The basic feature of this model is a creeping movement of tRNAs through the ribosome starting with a tRNA movement only on the 50S ribosome after peptide bond formation and before translocation. This movement is uncoupled from that on the 30S subunit. (B) The α–ε model of the elongation cycle according to Dabrowski, M., Spahn, C. M., Schafer, M. A., Patzke, S., and Nierhaus, K. H. (1998). Protection patterns of tRNAs do not change during ribosomal translocation. *J. Biol. Chem.* **273**, 32793–32800. The essential features are moveable ribosomal α- and ε-domains that connect both subunits through the intersubunit space, bind both tRNAs and carry them in concert from A and P sites to P and E sites, respectively, during the translocation step facilitated by EF-G. The model keeps all the features of the allosteric three side model (see text), but explains the reciprocal linkage between A and E sites by the fact that the moveable domain moves out of the A site during translocation leaving the decoding center alone at the A site in the POST state ribosome. The occupation of the E site generates a low-affinity A site, which is important for the selection of the newly incoming ternary complex at the A site. Yellow and pink, the two binding regions of the α–ε domain, blue the decoding center (δ) at the A site.

and after the slash that on the large subunit). In terms of the hybrid-site model, six different protection patterns were correlated to tRNA binding positions. The translocation reaction brings the peptidyl-tRNA from the A/P hybrid site to the P/P site and the deacylated tRNA from the P/E to the E/E site.

A number of criticisms can be raised, the most serious one is that a hybrid site was not observed when functional states of the elongating ribosome were systematically investigated by cryo-electron microscopy (cryo-EM). Note that a hybrid site can be easily detected, since 85% of a tRNA is in contact with the large subunit with the consequence that an even partial movement of a tRNA on the large subunit would result in a substantial change in the overall position of the tRNA. Even a movement of the CCA-ends of the tRNAs at the PTF center could not be observed in crystal structures of 50S complexes. This prompted a revision in the hybrid-site model so as to keep the tRNAs in the classical A/A and P/P sites after peptide-bond formation, and only after an undefined time span the tRNAs are then shifted to the hybrid sites A/P and P/E, respectively.

However, it seems likely that a peptidyl-tRNA at the A site moves via a transient hybrid position A/P into the P site and likewise a deacylated tRNA from P to E via hybrid position P/E during the translocation reaction, which has been resolved into a ratchet-like forth-and-back movement between the subunits. Analyses applying cryo-EM suggested indeed, that a tRNA moves from P to the E site via a hybrid P/E position during translocation.

The $\alpha{-}\varepsilon$ Model

The essential feature of the $\alpha{-}\varepsilon$ model of elongation cycle is a movable domain, called $\alpha{-}\varepsilon$ domain, which binds both tRNAs of an elongating ribosome and carries them from the A and P sites to the P and E sites, respectively, during translocation (see Figure 2B). But the model also includes the features of the previously described allosteric-three-site model.

1. The ribosome contains three tRNA-binding sites: A, P, and E site.

2. E and A site are coupled in a reciprocal manner: An occupied E site decreases the affinity at the A site and vice versa. The decreased affinity at the A site by an occupied E site might be important for preventing an interference of noncognate ternary complexes with the selection process of the cognate ternary complex from near cognate ternary complexes at the A site. During the accommodation of the aa-tRNA into the A site (stable A site occupation) the E site tRNA is released. The reciprocal interaction between the A and the E sites explains why statistically two tRNAs are found on polysomes that contain a mixture of both PRE and POST state ribosomes.

3. Both tRNAs on a ribosome are bound via codon–anticodon interactions in both the PRE and the POST states. Especially the codon–anticodon interaction at the E site seems to be essential for: first, establishing the POST state which is the proper substrate for the ternary complex, second, reducing the error rate of protein synthesis, and third, keeping the reading frame.

It is probable that the $\alpha{-}\varepsilon$ model will be modified as soon as higher resolution structures of the ribosomal PRE and POST states become available.

A comprehensive analysis of the translocation reaction by cryo-EM to date challenges an essential feature of the $\alpha{-}\varepsilon$ model: PRE states of ribosomes were analyzed by cryo-EM carrying an fMet-Phe-tRNA at the A site and a deacylated tRNA^{Phe} at the P site. In these complexes a tRNA is also seen at the E site, although no tRNA was specifically bound to this site. The authors assume that the tRNA binds from the free pool of deacylated tRNA in solution to the "high affinity" E site. Either way the presence of well populated A and E site is at odds with the $\alpha{-}\varepsilon$ model. However, neither the origin of the E-site tRNA nor the tRNA:70S stoichiometry is known, therefore it would be premature to disregard carefully controlled experiments upon which the $\alpha{-}\varepsilon$ model is based.

Selection of the Ternary Complex: Decoding and A Site Occupation

In *Escherichia coli*, there are 45 different species of tRNAs, where a species is defined as a tRNA with a unique anticodon sequence. The tRNA species can be separated into three classes with respect to the codon displayed in the A site. The "cognate" class contains one aminoacyl-tRNA with an anticodon complementary to the A-site codon. The "near-cognate" class contains four to six aminoacyl-tRNA that carry anticodons similar to that of the cognate one (never more than one mismatch). The "noncognate" class contains the bulk of aminoacyl-tRNAs with a dissimilar anticodon (usually more than one mismatch). As the tertiary structure of tRNA is highly conserved the ribosome has to distinguish between tRNAs that hardly differ. The problem is compounded by the fact, that the substrate for the A site is not an isolated aa-tRNA but rather the much larger ternary complex aa-tRNA·EF-Tu·GTP, i.e., EF-Tu $*$ GTP is twice as large as a tRNA. The ribosome must therefore discriminate between relatively large ternary complexes (72 kDa) that are extremely similar, on the basis of a small discriminatory region, namely the anticodon (1 kDa).

In view of the predominance of the nondiscriminatory fraction of free energy of binding over the discriminatory energy, protein synthesis must be either slow and accurate or fast and imprecise. This is not what

we see *in vivo*, where protein synthesis is fast and accurate, incorporating up to 10–20 amino acids per second with an accuracy of one misincorporation per 3000 amino acid incorporations. How is the ribosome solving this paradox?

A first hint gave the observation that the A site occupation occurs in two steps as already mentioned (see Figure 2B): a decoding step, where the selection of the cognate ternary complex takes place, is followed by an accommodation step. The α–ε model has integrated this observation in the following way: the decoding takes place at an A site with low affinity for tRNA, which reduces the binding energy of the ternary complex to mainly codon–anticodon interactions and excludes contacts of the tRNA outside the anticodon and of EF-Tu with the ribosome. In this state the free energy of binding is small, and since it is restricted to codon–anticodon interaction it is more or less identical with the discrimination energy. This feature explains why the majority of "noncognate" aa-tRNA ($\sim 90\%$ of the aa-tRNA species) do not interfere with the decoding process: their anticodon is different from that of the cognate aa-tRNA, and interactions outside the anticodon are prevented by the low-affinity A site.

This fast initial step is followed by an accommodation step, where the aminoacyl-tRNA is tightly bound and accommodated into the A site. This step is accompanied by some gross conformational changes, since during this step the E-site tRNA is released and the A site switches into its high-affinity state. Therefore, the second step is probably slow in comparison to the decoding the step. The A site occupation is therefore a coupled system of two reactions, the first of which is fast and the second slow. An important consequence of this arrangement is that the first runs at equilibrium even under steady-state conditions and thus can exploit the discrimination potential of the decoding process.

The reciprocal linkage between A and E site seems to be a universal feature of ribosomes and has been demonstrated not only in bacteria but also in eukarya (yeast).

After considering the competition cognate versus noncognate aa-tRNAs, still a discussion on how the ribosome discriminates between cognate and near-cognate tRNAs arises. Two models have been proposed: (1) the kinetic proofreading model, and (2) the Potapov model.

In the late 1970s, stability measurements of anticodon:anticodon interactions within a complex of two tRNAs have demonstrated that the corresponding energy cannot explain a selection accuracy of better than 1:10. Therefore, proofreading models have been developed according to which the stability energy is exploited several times in order to explain the observed accuracy of aa-tRNA selection at the ribosomal A site of about 1:3000. One proofreading event requires one EF-Tu dependent GTP hydrolysis, so that a measurement of the number of GTPs hydrolyzed by EF-Tu per incorporation of a near-cognate amino acid indicates the importance of proofreading for the selection process. Precise measurements revealed that the importance of proofreading is much less than originally thought, initial binding of the ternary complex gives a precision of about 1:300 up to 1:1000, whereas the corresponding proofreading factor is not better than 1:10.

How initial binding is able to achieve such an accuracy is explained by the Potapov model. This model suggests that the decoding center on the ribosome does not measure the stability of codon–anticodon interaction, but rather the stereochemical correctness of the three Watson–Crick base pairs, just as an enzyme recognizes its substrate. With this assumption the correct position of the sugar pucker contributes to the accuracy, and it could be demonstrated that indeed the 2′OH groups of the codon bases are of utmost importance for the accuracy of the selection process.

The detailed molecular mechanism could be unravelled by the Ramakrishnan group who determined the crystal structure of 30S subunits carrying either a cognate or near-cognate anticodon stem-loop structures. Indeed, the correct positions of the 2′OH groups of the codon–anticodon complex is checked by forming hydrogen bonds with universally conserved bases of the 16S rRNA (Figure 3). The first base pair of codon–anticodon interaction at the A site is analyzed via the so-called A-minor motif type I and the second by an A-minor motif type II (Figure 3B and 3C), whereas the third wobble position has more freedom to accommodate also non-Watson–Crick base pairs (Figure 3D). Furthermore, the head and shoulder of the 30S subunit move relative to each other defining an open and closed 30S configuration. In the "open" configuration binding of cognate (but not a near-cognate) substrate to the decoding center flips out the bases A1492 and A1493 from the helix 44, brings G530 from a "syn" into an "anti" conformation (Figure 3A), and shifts the subunit into a "closed" configuration providing a molecular basis for an understanding of mutations that increase or decrease accuracy. A molecular dynamic simulation agrees with the main conclusions and shows in addition that the kink between the A and P site codons of about 135° influences the accuracy pattern.

AN ADDITIONAL ROLE OF EF-TU

It is well known that EF-Tu binds an aa-tRNA at the amino acid acceptor stem thus shielding the labile ester bond between the aminoacyl residue and the tRNA, and delivers the aa-tRNA to the A site on the ribosome. However, a second function of EF-Tu was identified by Uhlenbeck and co-workers. Measuring the affinities of various cognate aa-tRNA (e.g., Val-tRNA^{Val}) and some mispairs (e.g., Ala-tRNA^{Val}) they recognized that either
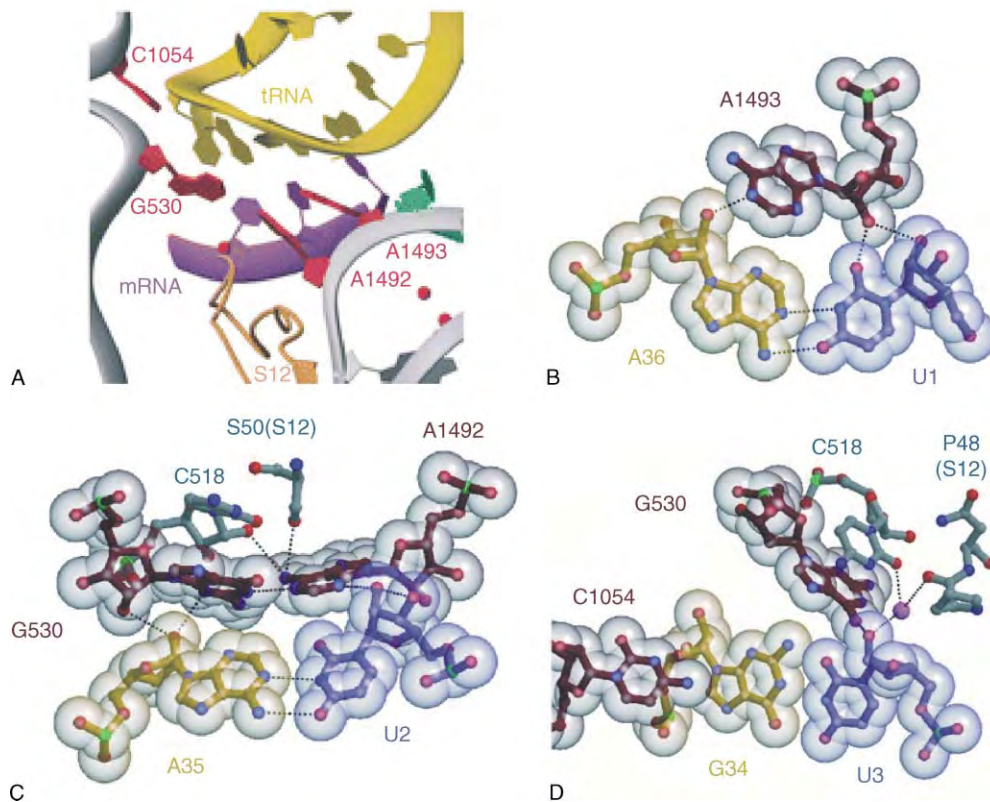
FIGURE 3 Molecular details of the decoding process, principles of decoding according to Ogle *et al.* 2001, *Science*, **292**, 897–902. (A) Conformational changes at the decoding center upon binding of a cognate anticodon-stem-loop (ASL) with permission. (B) Recognition of the Watson–Crick interaction at the first position of codon–anticodon by a type I A minor motive, A1493 is clinging into the minor groove of the first base pair of codon-anticodon interaction. (C) Recognition of the Watson–Crick interaction at the second position of codon–anticodon by type II A minor motive, A1492 and G530 are filling the minor groove at the second position, forming a hydrogen bond network to 2′OH groups, bases of codon–anticodon, and G518 and serine50 of S12. (D) The Watson–Crick geometry at position three is less restricted, G530, C518, and P48 are stabilizing the third codon position, but giving freedom, e.g., the G:U wobble base pair. Adapted from Ramakrishnan, 2002.

the amino acid or the tRNA should bind to EF-Tu with high affinity in order to form stable ternary complexes aa-tRNA·EF-Tu·GTP. For example, EF-Tu·GTP forms easily a ternary complex with Asp-tRNA^{Asp} (weak aa and strong tRNA) or Asn-tRNA^{Asn} (strong aa and weak tRNA), but not with Asp-tRNA^{Asn}, since in the latter case both moieties bind with low affinities. This observation explains an important case that was an enigma hitherto. In most organisms, there are not 20 different synthetases corresponding to the 20 natural amino acids, but only 19 or sometimes 18. For example, many organisms do not contain a synthetase specific for asparagine (Asn-RS). In this case the Asp-RS is charging also the tRNA^{Asn} with aspartic acid yielding Asp-tRNA^{Asn}, which is recognized by enzymes amidating Asp to Asn on the tRNA. The mis-charged Asp-tRNA^{Asn} does not form a stable ternary complex with EF-Tu·GTP and thus Asp is not incorporated at codons specifying Asn. This discrimination process via EF-Tu was termed thermodynamic compensation.

Peptide-Bond Formation

The PTF reaction, a central step in protein synthesis, is the catalytic activity of the large subunit. Even if the substrates are large, the reaction that occurs is quite simple–the aminolysis of an ester bond to form a peptide bond. The nucleophilic α-amino group of the amino acid moiety of the aminoacyl tRNA at the A site attacks the electrophilic carbonyl carbon atom of the ester-bonded peptide moiety of the P site tRNA. This forms an tetrahedral intermediate, which breaks down to an uncharged (deacylated) tRNA in the P site and a peptidyl-tRNA prolonged by one amino acid at the A site (Figure 4A). The PTF center of the 50S has been identified by using a transition state analogue of the PTF reaction, which was soaked into 50S crystals of the archaea *H. marismortui*. This analogue, CCdA-phosphate-puromycin (CCdApPmn), is a mimic of the CCA end of a tRNA in the P site attached to puromycin in the A site, where the
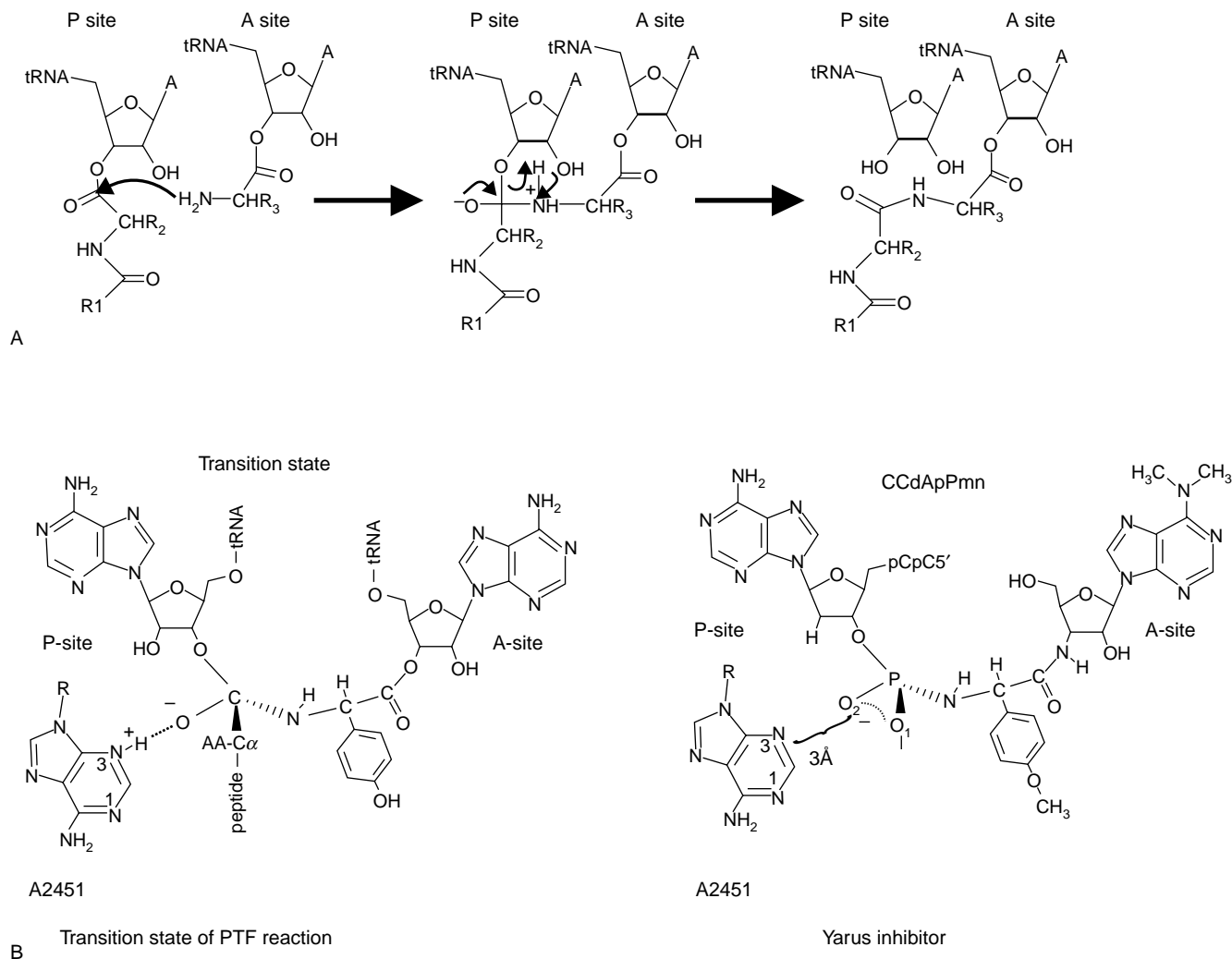
FIGURE 4 Mechanism of peptide bond formation. (A) PTF reaction: The α- amino group of the A site bond aa-tRNA attacks the electrophilic carbonyl carbon, which is attached via ester bond to the 3′OH of the adenosine residue of P site bond peptidyl-tRNA, forming a tetrahedral transition state that breaks down to a prolonged peptidyl-tRNA at the A site and a deacylated tRNA at the P site. (B) Comparison of the putative transition state of the PTF reaction with a possible transition-state analogue created by the Yarus group. From Parnell, K. M., Seila, A. C., and Strobel, S. A. (2002). Evidence against stabilization of the transition state oxyanion by a ρka-perturbed RNA base in the peptidyl transferase center. *Proc. Natl Acad. Sci. USA* **99**, 11658–11663.

3′-terminal deoxy-adenosine and the phosphate residue resemble the tetrahedral intermediate formed during peptidyl transfer (Figure 4B). It was shown previously by the Yarus group that this analogue is a strong inhibitor of the PTF reaction competitively inhibiting binding of the A-site substrate.

A long-standing discussion has been about the composition of the PTF center and the "players" involved in the PTF reaction, but now it is clear from the crystal structure – the ribosome is a ribozyme, i.e., no protein is directly involved in catalysis. The PTF center is tightly packed with rRNA, mainly derived from the domain V of 23S RNA (the so-called PTF ring), which are highly conserved over all domains of life. Although there are 15 proteins interacting with domain V, proteins are absent from the PTF within a distance of at least 18Å. The nearest proteins to the PTF in *H. marismortui* structure are L2, L3, L4, and L10e and interestingly, all four are also present in the vicinity of PTF in the eubacterial ribosome, when it is taken into account that eukaryotic L10e is evolutionary related to L16 in prokaryotes. These proteins have been identified previously, together with the 23S rRNA, as the major candidates for the PTF activity by single omission tests in a total reconstitution system of the large subunit. Although the structure reveals no direct involvement of protein in the catalysis, proteins might have a role in aligning the substrates and functioning as a "glue" to stabilize rRNA tertiary arrangement necessary for the peptide bond formation since complete removal of

proteins could only accomplished under conditions that unfolded the rRNA and totally abolished the PTF activity, whereas the removal of up to over 80% of proteins of the 50S subunit from *Thermus aquaticus* maintained activity. Detailed analysis of the neighborhood of the residues within close proximity to the CCA end analogues of tRNA led to the proposal for an acid-base catalysis mechanism for the PTF reaction involving the universal conserved A2451 in analogy to the back reaction of the mechanism used by serine proteases. This proposal was immediately under attack from a number of groups who presented biochemical and genetic data to the contrary.

Crystal structures of the 50S of *Deinococcus radiodurans* revealed some significant differences in the arrangement of nucleotides within the PTF center and the presence of a protein (L27), which is thought to play a role in placement of the CCA ends of both tRNAs. The debate on the mechanism of PTF reaction is on going. By re-examination of the transition state analogue it turned out that this analogue within the 50S structure can not answer all the mechanistic questions about the PTF reaction, since it is not a true mimic of the intermediate of the PTF reaction: modeling the missing oxygen atom at the 2′ position of the desoxy-adenosine caused a sterical clash with the phosphate group. Although the reaction seems to be quite simple, the kind of contribution of the ribosome in catalysis remains open.

Two main principles of enzymatic reactions can be distinguished and could be involved in PTF reaction. The first principle is the physical or template model, where the enzyme, here the ribosome, arranges the two substrates in optimal stereo-chemical positions for the reaction to proceed. Such an arrangement of substrates is sufficient to allow for a dramatic acceleration of the reaction rate by six to nine orders of magnitude. This strategy is certainly used by the ribosome, since there are binding sides for both tRNA substrates fixing the substrates in a defined position by interaction between 23S rRNA (A loop and P loop) and tRNAs (see Figure 5), placing the corresponding CCA ends into the PTF center.

In addition it is possible that a chemical concept is utilized by the ribosome, i.e., transiently covalent bonds are formed and broken between the enzyme and the substrates. There are three groups in close proximity to the reactive amino group within the PTF center in the *H. marismortui* crystal structure, that could form hydrogen bonds with it, namely (1) the 2′-OH of the peptidyl-tRNA, (2) the N3 of A2451 (*E. coli* nomenclature) of 23S rRNA, and (3) the 2′OH of A2451.

The hydrogen bonds these groups could form with the α-amino group of the aa-tRNA at the A site may help to fix and optimally align the reactive α-amino group within the PTF center. And if one of these groups have an elevated pK_a, its hydrogen bond would
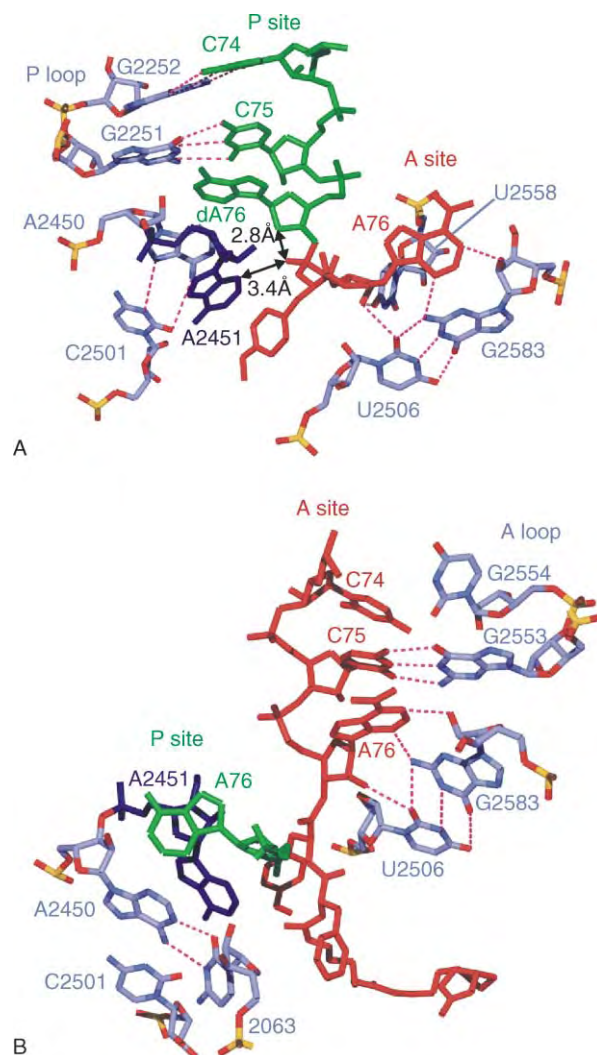
FIGURE 5 Tight fixation of the CCA ends of the P- and A-tRNAs observed in 50S subunit from *H. marismortui* in complex with (A) the Yarus inhibitor, and (B) the products following peptide bond formation. The CCA-ends of the tRNAs in the A and P sites are colored red and green, respectively. The N3 of A2451 (dark blue) is 3.4Å from the O2 of the Yarus inhibitor, while the same O2 is only 2.8Å from the 2′-deoxy of A76 (arrowed). Selected rRNA residues of domain V of the 23S rRNA are colored light blue, including the A- and P-loop bases that participate in A and P site CCA end fixation (*E. coli* numbering). In (B) the P site C74 and C75 have been omitted for clarity. Dashes indicate hydrogen bonding and rRNA nucleotides use the following color scheme: Oxygen, red; phosphorus, yellow; nitrogen, blue; carbon, dark blue. Adapted from Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930.

facilitate the nucleophilic attack from the α-amino group of the aminoacyl-tRNA at the A side to the carbon of the carbonyl group of peptidyl-tRNA at the P site. A major candidate for an enhancement of nucleophily of the α-amino group is N3 of A2451 (see Figure 6).
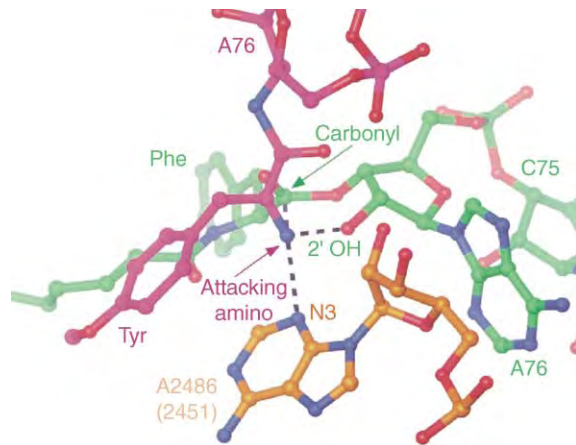
FIGURE 6 Putative arrangements of the residues involved in peptide-bond formation derived from crystal structures of the *H. marismortui* 50S with and without low-molecular weight substrate and transition state analogues. N3 of A2486 (A2451 in *E. coli*) and the 2′OH group of A76 of peptidyl-tRNA (green) forming hydrogen bonds to the α-amino group of the A site aa-tRNA (red) that may facilitate the nucleophilic attack of the α-amino group to the carbonyl carbon of peptidyl-tRNA at the P site. From Moore P. B., and Steitz, T. A. (2003). After the ribosome structures: how does peptidyl transferase work? *RNA* **9**, 155–159.

Translocation

The translocation step of elongation cycle is one of the most difficult to understand, since the ribosome must move two large molecules, the tRNAs (each ~25 kDa, ~75 × 50 × 30Å), and a mRNA in a concerted action through the ribosome precisely by one codon length (~10–15Å). Under certain *in vitro* conditions translation can occur spontaneously but with a rate that is three orders of magnitude slower than that in the presence of EF-G. Nevertheless, this indicates that the translocation reaction is a feature inherent in the ribosome itself. An impressive confirmation comes from the observation that the antibiotic sparsomycin that binds to the PTF center and blocks peptide-bond formation can trigger one round of translocation efficiently. Even EF-G dependent GTP hydrolysis is not needed in order to promote the translocation reaction, since EF-G·GDPNP, the noncleavable analogue of GTP, efficiently fosters translocation. Nevertheless, the proposal has been made that at least the energy of EF-G dependent GTP cleavage accelerates the translocation based on the observation that in the presence of the hardly cleavable GTP-γ-S analogue the translocation is somewhat slower than in the presence of GTP.

The most recent progress in the understanding of the translocation reaction comes from a detailed cryo-EM study of this reaction. EF-G induces a ratchet like movement between 30S and 50S subunits as first shown with empty ribosomes, but here the movement could be coupled to translocation of tRNAs at higher resolution yielding a more precise picture of this reaction:

1. After occupying the A site with an aa-tRNA and peptide bond formation, the peptidyl-tRNA is in the A site and a deacylated tRNA at the P site, no hybrid site is visible as was seen in cryo-EM studies.

2. EF-G·GTP binds and induces the first forth-movement of the ratchet motion of the 30S subunit, a turn of ~20°. During this movement the deacylated tRNA is shifted into a hybrid position P/E and it is hypothesized (although not yet observed) that the peptidyl-tRNA would also move in an analogous fashion to occupy a A/P hybrid position.

3. The ribosome triggers the EF-G dependent GTP hydrolysis and the authors postulate that the resulting EF-G conformational change into the GDP conformer completes the translocation reaction in that

- the 30S subunit moves back (second part of the ratchet movement),
- the tRNAs continue their movement to E and P sites, respectively, and
- EF-G·GDP dissociates from the ribosome.

A strong dislocation movement of domains 3-5 of EF-G·GDPPNP (a GTP analogue) on the ribosome was observed as compared to the crystal structure of EF-G·GDP. In particular, the tip of domain 4 moves ~35 Å away upon GTP cleavage. The authors speculate that this movement is responsible for the second half of the ratchet motion with the three consequences outlined in the preceding sentence. They also detect a substantial movement of the L1 protuberance during the translocation and postulate an active participation of this structure in the translocation of the deacylated tRNA from the P site to the E site. The translocational movement of the tRNAs from, e.g. , the A site to the P site via a A/P state can be well reconciled with the α–ε model of the elongation.

Note the difference to the hybrid-site model: this model postulates that after peptide-bond formation and before translocation the tRNAs move on the large subunit but stay on the 30S subunit, i.e., the peptidyl-tRNA at the A site moves from A/A site to the A/P site, and only the translocation movement brings the tRNA into the P/P site, whereas the cryo-EM study suggests that during translocation the peptidyl-tRNA moves from the A to the P site via a transient A/P position.

The ribosome research is an exciting upheaval phase where the structure begins to explain the function. For the first time it can be imagined how the ribosome is using its complicated structure to perform its yet more complicated function.

SEE ALSO THE FOLLOWING ARTICLES

EF-G and EF-Tu Structures and Translation Elongation in Bacteria • Ribosome Assembly • Ribosome Structure • Translation Initiation in Bacteria: Factors and Mechanisms • Translation Termination and Ribosome Recycling

GLOSSARY

A site Ribosomal tRNA binding site for aminoacyl-tRNA (A for aminoacyl-). At this site decoding of the displayed codon takes place; it is the entry site for aminoacyl-tRNA complexed with the elongation factor EF-Tu.

E site Ribosomal tRNA-binding site that exclusively binds deacylated tRNA before it dissociates from the ribosome.

EF-G An elongation factor that promotes translocation. It is a G protein that binds as binary complex EF-G·GTP to the ribosome. The homologue in eukaryotes and archaea is termed EF2.

EF-Tu An elongation factor that carries aminoacyl-tRNA to the ribosomal A site as a ternary complex aminoacyl-tRNA·EF-Tu·GTP. It is also a G protein. The homologue in eukaryotes and archaea is termed EF1.

P site Ribosomal tRNA-binding site for peptidyl-tRNA before peptide-bond formation (P for peptidyl-).

FURTHER READING

Agrawal, R. K., Spahn, C. M. T., Penczek, P., Grassucci, R. A., Nierhaus, K. H., and Frank, J. (2000). Visualization of tRNA movements on the *Escherichia coli* 70S ribosome during the elongation cycle. *J. Cell. Biol.* **150**, 447–459.

Fredrick, K., and Noller, H. F. (2003). Catalysis of ribosomal translocation by sparsomycin. *Science* **300**, 1159–1162.

Green, R., and Noller, H. F. (1997). Ribosomes and translation. *Annu. Rev. Biochem.* **66**, 679–716.

LaRiviere, F. J., Wolfson, A. D., and Ühlenbeck, O. C. (2001). Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* **294**, 165–168.

Moore, P. B., and Steitz, T. A. (2002). The involvement of RNA in ribosome function. *Nature* **418**, 229–235.

Noller, H. F., Yusupov, M. M., Yusupova, G. Z., Baucom, A., and Cate, J. H. D. (2002). Translocation of tRNA during protein synthesis. *FEBS Lett.* **514**, 11–16.

Ogle, J., Carter, A., and Ramakrishnan, V. (2003). Insights into the decoding mechanism from recent ribosome structures. *TIBS* **28**(5), 259–266.

Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M., and Frank, J. (2003). Locking and unlocking of ribosomal motions. *Cell* **114**, 123–134.

Wilson, D. N., and Nierhaus, K. H. (2003). The ribosome through the looking glass. *Angew. Chemie* **42**(30), 3464–3486.

BIOGRAPHY

Oliver Vesper has studied biochemistry at the Free University in Berlin and is currently pursuing his Ph.D. studies in the Max-Planck-Institut für Molekulare Genetik in Berlin.

Knud H. Nierhaus has studied medicine in Tübingen and Vienna, is a Professor of Biochemistry at the Technical University in Berlin and Group Leader at the Max-Planck-Institut für Molekulare Genetik in Berlin.

Translation Elongation in Eukaryotes

William C. Merrick and Anton A. Komar
Case Western Reserve University School of Medicine, Cleveland, Ohio, USA

The process of translation is defined by three steps: initiation, which effects mRNA binding and correct match of the initiator tRNA with the AUG start codon; elongation, which effects the correct matching of all the remaining aminoacyl-tRNAs in a codon specific manner; and termination, which senses a stop codon and leads to the release of the completed peptide chain. Of these three steps, elongation is the most predominant one as it accounts for all but one of the amino acids that end up in the completed polypeptide chain. Elongation itself is composed of three traditionally defined steps: eEF1A-directed binding of the aminoacyl-tRNA to the A site (aminoacyl site) of the ribosome; peptide bond formation triggered by the one enzymatic activity of the ribosome (the peptidyl transferase center); and eEF2-mediated translocation which moves the peptidyl-tRNA from the A site to the P site (peptidyl site) by precisely one codon (three nucleotides). Although the matching of aminoacyl-tRNAs to codons in the A site is dictated by the genetic code, the accuracy and speed of the elongation process is driven by the hydrolysis of GTP in the binding and translocation steps.

Introduction

The purpose of this article is to provide a basic understanding of the process of protein synthesis elongation. While the focus will be on eukaryotic systems, the general mechanism of protein synthesis elongation is essentially the same in eukaryotes and prokaryotes (Table I). This is perhaps not surprising given the ease with which one can trace the evolution of the eukaryotic factors from their bacterial predecessors. At present, the best biochemical and structural data on the elongation cycle are from bacterial systems and includes crystallographic structures of the ribosomal subunits (30S and 50S), EF1A·GDP, EF2·GDP, and EF1A·GDPNP·Phe-tRNA. The regulation of the elongation cycle has been best studied in the eukaryotic system and there appear to be several posttranslational modifications that influence synthetic rate.

The Traditional Elongation Cycle: The A and P Sites

At the end of the initiation cycle, the final product is an 80S ribosome that has an mRNA (or AUG codon) bound to the initiator tRNA (Met-tRNA$_i$) correctly based paired with AUG start codon, and with the Met-tRNA$_i$ in the peptidyl or P site (Figure 1). This placement of the initiator tRNA in the P site sets the reading frame for all subsequent aminoacyl-tRNAs. The next event is the binding of the ternary complex (eEF1A·GTP·aminoacyl-tRNA) to the aminoacyl or A site (Figure 1). This binding is directed by the interaction of the anticodon of the tRNA and the three nucleotides in the A site. The specificity is such that perfect Watson–Crick base pairs are observed for the first two nucleotides in the codon (i.e., only A-U or G-C base pairs), but altered base pairing is possible at the third position (as originally proposed by Crick). The genetic code dictates which aminoacyl-tRNA will bind to the A site in response to the 61 possible triplet codons that specify an amino acid (note that three codons of the possible 64 are stop codons: UAA, UAG, and UGA). With the correct base pairing, a conformational change occurs that triggers the hydrolysis of GTP leading to the release of eEF1A·GDP from the ribosome.

The next step is peptide bond formation. The synthesis of the peptide bond is catalyzed by the peptidyl transferase center in the large ribosomal subunit (60S). *In vivo*, this step appears to occur instantaneously and is likely the most rapid step in the entire process. As should be noted in Figure 1, peptide bond formation leads to the movement of the upper halves of the two tRNAs such that the tRNA originally in the P site is now half in the exit or E site (aminoacyl end) (Figure 2) and half in the P site (the anticodon end). In a similar manner, the tRNA originally in the A site is now half in the P site (aminoacyl end) and half in the A site (the anticodon end).

The final step in the cycle is translocation. This is driven by the binding of eEF2·GTP to the A site of

TABLE I

Eukaryotic and Prokaryotic Components of the Elongation Cycle

Prokaryotes	Eukaryotes	Function
30S, 50S subunits	40S, 60S subunits	Peptide bond formation
EF1A	eEF1A	Binding of aa-tRNA to ribosomes
EF1B	eEF1B	Nucleotide exchange factor
EF2	eEF2	Translocation
aa-tRNA	aa-tRNA	Carrier of the activated amino acid, recognition of the codon in the A site

the ribosome. Hydrolysis of the GTP in the eEF2·GTP complex triggers the movement of the peptidyl-tRNA in the P/A site into the P/P site. During this movement, it is thought that an "arginine finger" in the ribosome stabilizes the interaction between the anticodon of the tRNA and the codon in the mRNA. By moving the anticodon of the peptidyl-tRNA fully into the P site (P/P), the mRNA is moved along the ribosome precisely

by three nucleotides (or one codon). In this manner, the ribosome maintains the reading frame of the mRNA.

Although a complete elongation cycle has been completed, we are not back to the beginning. As noted above, eEF1A is released as an eEF1A·GDP complex. The binding affinity of eEF1A for GDP is ~0.1 μM. Thus, in biological terms, this is a rather stable complex. To enhance the conversion of eEF1A·GDP into the biologically active eEF1A·GTP form, the eEF1A·GDP complex binds to eEF1B. This leads to the release of GDP and the formation of an eEF1A·1B complex. Subsequent binding of GTP yields eEF1A·GTP + eEF1B. This exchange cycle is very similar to that of most "standard G proteins." In contrast, eEF2 does not require an exchange factor as it has a much poorer affinity for GDP, and thus GDP is rapidly released from eEF2 following GTP hydrolysis.

A few notes on the elongation cycle: first, the energy cost would appear to be two high-energy phosphates per cycle; however, the actual value is four when one considers that two high-energy phosphates are required for the attachment of the amino acid to the tRNA, which was then placed in the A site. Second, on careful
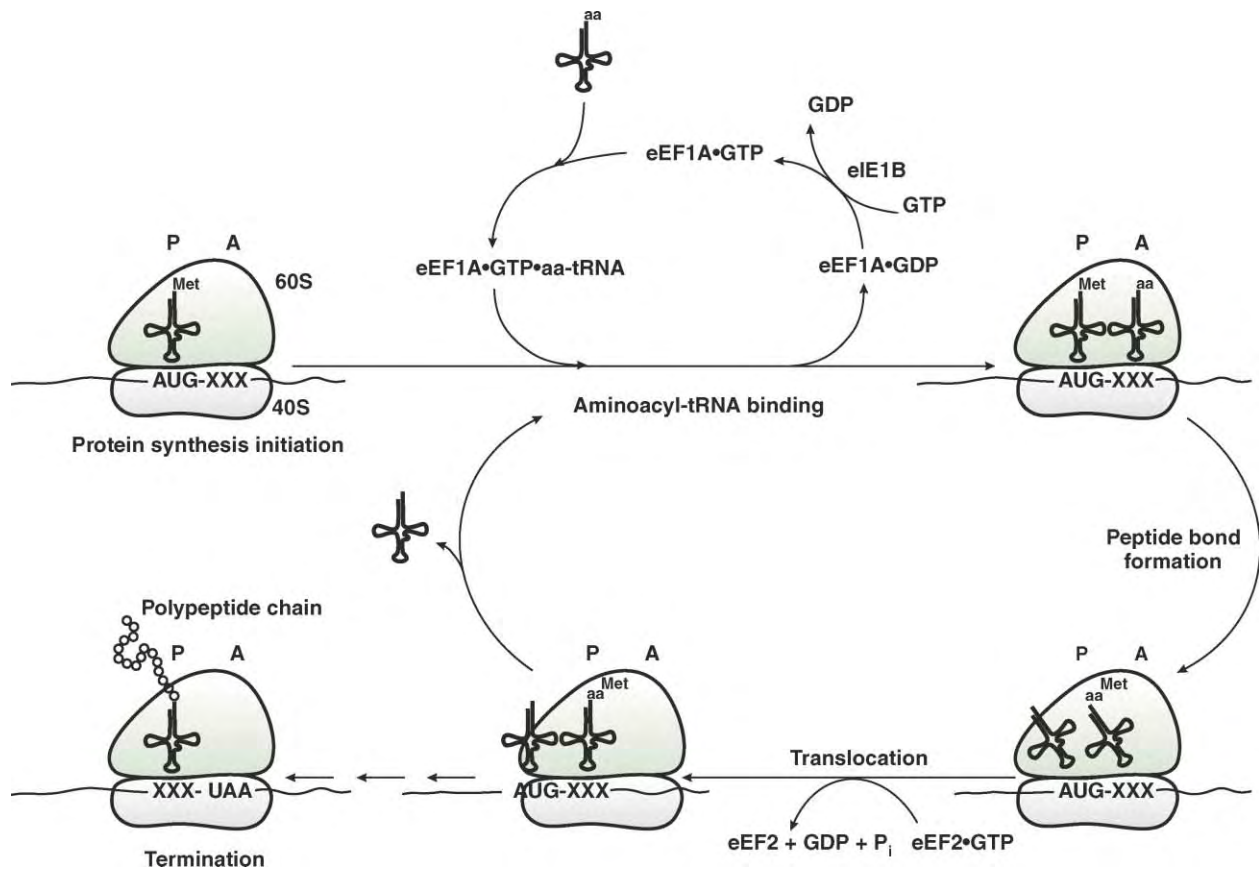
FIGURE 1 The traditional elongation cycle: The A and P sites. Shown above is the elongation cycle with the traditional A and P sites (aminoacyl and peptidyl sites, respectively). Following initiation, subsequent steps in the elongation cycle ensue with aminoacyl-tRNA binding, peptide bond formation, and then translocation. This process continues until a stop codon (UAA, UAG, or UGA) appears in the A site at which point termination occurs.
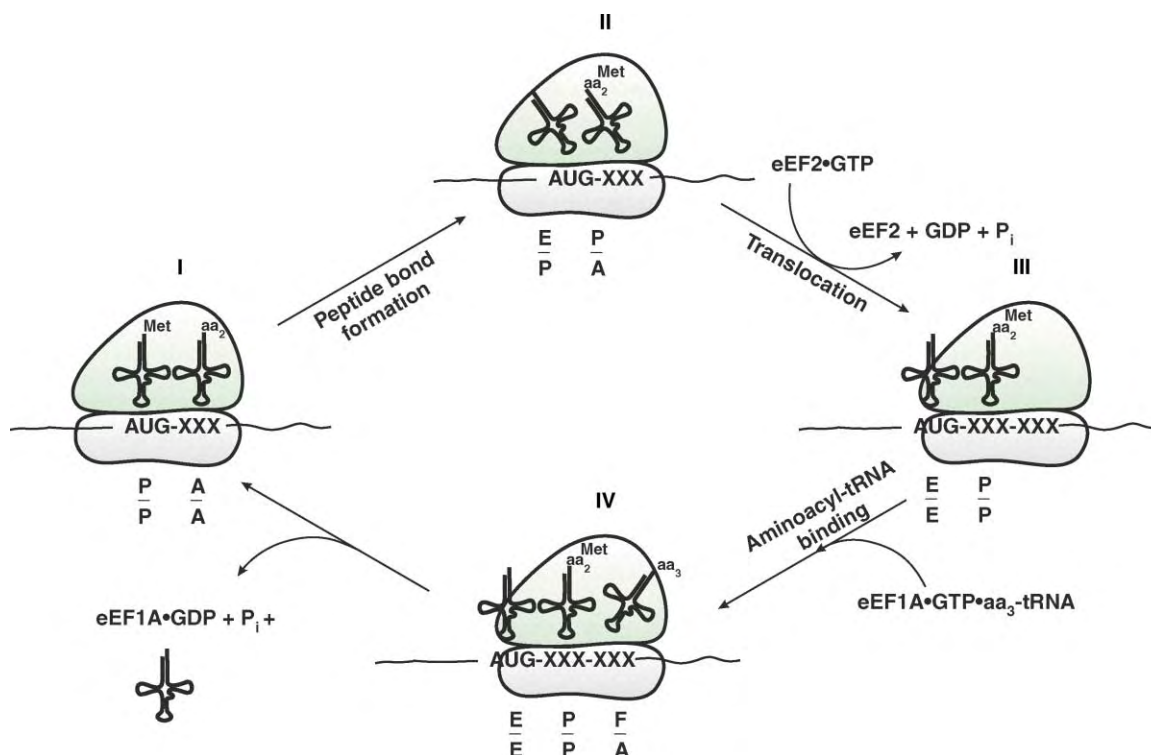
FIGURE 2   Influence of other sites. Illustrated in this figure are the four proposed sites in the eukaryotic ribosome. The A and P sites are the "traditional" tRNA binding sites on the ribosome (Figure 1). The E site (exit site) and the F site (entry site) are also shown. Note that a portion of each site is located on the small (40S) and the large (60S) ribosomal subunit. In the designation for the tRNA molecules, the upper letter indicates position on the 60S subunit and the lower letter indicates position on the 40S subunit.

examination, it should be apparent that the synthesis of the peptide/protein is from the amino terminus to the carboxy terminus while the mRNA codons are read in a $5'$ to $3'$ direction.

Elongation Cycle Chemistry

Figure 3 shows the "chemistry" that occurs with peptide bond formation. The initial step is the nucleophilic attack of the amino group of the second amino acid (R2) on the carbonyl of the initiating amino acid (methionine, R1) as is seen in Schiff's base formation (I). It is thought that the function of the peptidyl transferase center is to facilitate the abstraction of the H^+ in the intermediate (II) by providing a weak base (:B). With the release of this H^+ and a shift of electrons, the bond between the $3'$ oxygen of the terminal adenosine in the tRNA in the P site and the methionine is cleaved, thus completing the transfer of the methionine to the amino group of the second amino acid and formation of the peptide bond (III). The cycle is then completed with the reacquisition of a proton at the $3'$ position of the tRNA in the P site (IV). Finally, with eEF2-catalyzed translocation, the peptidyl-tRNA is moved completely into the P site (V).

In the overall thermodynamics of the reaction, the resulting peptide bond is rather low energy while the original aminoacyl linkage is high energy (equivalent to that in ATP), and thus the net ΔG (change in the Gibb's free energy) is quite negative. While most primary amines would be expected to have a pK_a of ~9–10, that of the "α" amino groups is ~7.5 due to the inductive effects of the ester linkage of the amino acid to the tRNA (in a similar manner, the pK_a of the amino group at the amino terminus of the peptide/protein is also ~7.5). Thus, at physiologic pH values, the nucleophilic amino group is positively charged only about 50% of the time.

Molecular Mimicry

The crystallographic structures of bacterial homologues of eEF1A·GDPNP·Phe-tRNA and eEF2 have been determined. Comparison of these two structures indicated that to a first approximation they had a similar size, shape, and electronic distribution. While this might have been expected for the GTP-binding domain, EF2 has a unique protein finger that appears to mimic the anticodon stem and loop of a tRNA by overall size and charge (rich in aspartic and glutamic acid residues).
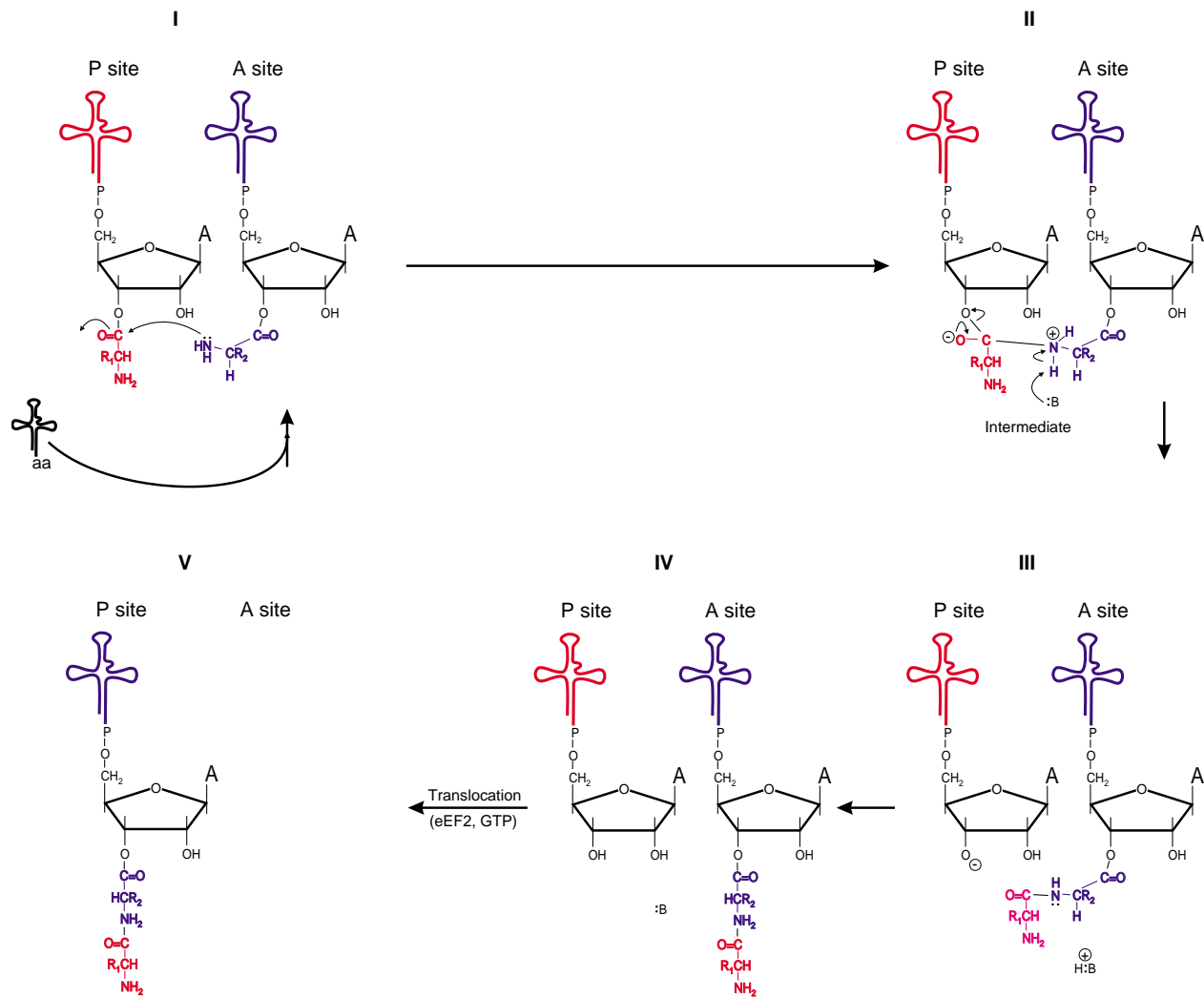
FIGURE 3   Elongation cycle chemistry. Shown above is the possible chemistry of the process of elongation. Nucleophilic attack by the amino group in the A site on the carbonyl bond of the amino acid in the P site leads to an amino-tetrahedral intermediate (II). The primary role of the ribosome in peptide bond formation is to facilitate the forward reaction by use of a general weak base (:B) although the group in the 60S subunit serving this function has not yet been identified. The chemistry is completed by subsequent cleavage of the ester bond to the tRNA in the P site (III). The cycle is then completed by the acquisition of a proton to the $3'$ position, and the loss of the proton from the weak base (IV) followed by translocation (V). Binding of a new aminoacyl-tRNA (as an eEF1A·GTP·aminoacyl-tRNA complex) starts the beginning of a new cycle (I).

The ability of a protein domain to take on the apparent characteristics of RNA has been termed "molecular mimicry." Although not discussed here, it has been proposed that all factors (initiation factors, termination factors, stringent factor, etc.) that bind to this region of the ribosome are also molecular mimics and will probably yield structures that are similar to those of the ternary complex and EF2.

Influence of Other Sites

In the discussion above, reference was made to an E site (Figure 2). This site, to the $5'$ side of the P site, has been visualized by cryoelectron microscopy and inferred by nuclease footprinting. The most important effect of the E site is that it appears to increase the stringency for the correct matching of the codon and anticodon in the A site of the ribosome. It should be noticed that, for the original binding of the first aminoacyl-tRNA, Met-tRNA$_i$, the E site was unoccupied. One might anticipate that this would allow for less stringent recognition of the initiating AUG codon; however, this strict recognition appears to result from the unique set of factors associated with the initiation process. A second proposed site is the F site or entry site which is just $3'$ of the A site. The suggestion here is that this may be an initial test-binding site (a check of codon/anticodon match)

such that correctly matching complexes would be pulled into the A site, while incorrectly matching complexes would dissociate. The advantage of the F site is that it would be more accessible to ternary complexes of eEF1A·GTP·aminoacyl-tRNA, as both the A site and the P site appear to be partially occluded at the interface of the large and small ribosomal subunits. However, the biochemical evidence supporting the existence of the F site is weak compared to that for the E site.

Assuming that all four of the ribosomal sites exist and that they are functional, they can be characterized by positions of the tRNAs on either the large or small ribosomal subunits. Thus, following initiation, the Met-tRNA$_i$ may be described as being in the P/P position (both the aminoacyl end and the anticodon end of the tRNA correspond to the P site (Figure 2) I or III). In the next step, the initial binding of the aminoacyl-tRNA is to the F/A site where the anticodon is in the A site, but the aminoacyl end is in the F site, and this end of the aminoacyl-tRNA is still bound to eEF1A·GTP (Figure 3, IV). With correct codon/anticodon recognition, GTP is hydrolyzed and eEF1A·GDP is released from the ribosome, concomitant with the movement of the aminoacyl end of the tRNA into the A site (now in the A/A configuration; Figure 2, I). At the same time, any tRNA in the E site (as would be true for most elongation steps) would be released from the ribosome. Subsequent reactions would cyclically yield the other states of the ribosome (P/P A/A, E/P P/A, and then E/E P/P).

The Other Elongation Factor, eEF3

A translation elongation factor unique to yeast and fungi is eEF3. This protein, which contains two-nucleotide-binding sites, appears to be required for the nucleotide-dependent release of the nonacylated tRNA from the ribosomal E site. As this protein is an essential gene product in yeast, it is surprising that an equivalent activity has not been identified in other eukaryotes. However, it has been noted *in vitro* that only elongation reactions using yeast ribosomes demonstrate the eEF3 requirement, and thus this requirement for eEF3 would appear to reflect unusual properties of the yeast ribosome compared to other eukaryotic ribosomes.

Regulation of the Elongation Cycle

The major regulation of gene expression at the translational level occurs through the covalent modification of translation initiation factors or regulatory proteins that bind to the initiation factors. However, there is ample evidence that regulation of translation also occurs at the level of elongation although the degree of regulation (the fold change in the elongation rate) is not as great as seen with regulation of initiation. While eEF1A, eEF1B, and eEF2 are all phosphorylated under different conditions, eEF1A and eEF2 also contain unique posttranslational modifications. eEF1A contains methylated lysines (mono-, di-, or trimethyl lysine) and glycerylphosphorylethanolamine. eEF2 contains a hypermodified histidine residue (histidine #714 in mammalian eEF2's) that is found in all eukaryotic eEF2's called diphthamide. The diphthamide residue is a known inactivation site for ADP-ribosylation catalyzed by either *Diphtheria* or *Pseudomonas* A toxins with NAD serving as the donor of the ADP group. There is tentative evidence that this modification may be part of the normal cellular regulation of eEF2 activity as well.

The best-studied regulation of any of the elongation factors is via phosphorylation. In general, phosphorylation of eEF1A and eEF1B correlates well with increases in the rate of elongation noted *in vivo* with either insulin or phorbol ester treatment. Additionally, most phosphorylations of translation factors are associated with an increased rate of protein synthesis. In contrast, the phosphorylation of eEF2 leads to its inactivation. Under most circumstances, changes in the elongation rate due to changes in covalent modifications of eEF1A, eEF1B, or eEF2 are associated with a coordinate change in the rate of initiation of protein biosynthesis.

SEE ALSO THE FOLLOWING ARTICLES

EF-G and EF-Tu Structures and Translation Elongation in Bacteria • Ribosome Structure • Translation Elongation in Bacteria • Translation Initiation in Bacteria: Factors and Mechanisms • Translation Initiation in Eukaryotes: Factors and Mechanisms • Translation Termination and Ribosome Recycling

GLOSSARY

E, A, P, and F sites   Physical locations on the surface of the ribosome that are occupied by aminoacyl- or peptidyl-tRNA.

elongation (of protein synthesis)   The sequential steps that lead to the addition of one amino acid at a time to the growing polypeptide chain.

elongation factor (EF)   A nonribosomal protein that facilitates the process of elongation only. (Note: eukaryotic elongation factors are designated eEF, where the lower case "e" signifies "eukaryotic.")

initiation (of protein synthesis)   The required steps that lead to the placement of the initiator tRNA in the P site of the ribosome, correctly base paired with the initiating AUG codon.

protein synthesis   The process of joining amino acids in a specific sequence through the a carbonyl and a amino groups via a peptide bond that is templated by an mRNA molecule.

termination (of protein synthesis)   The codon-directed (UAA, UAG or UGA) process of cleavage (and therefore release) of the polypeptide chain from the tRNA in the P site of the ribosome.

FURTHER READING

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B., and Steitz, T. A. (1999). Placement of a protein and RNA structures into a 5 Å-resolution map of the 50S ribosomal subunit. *Nature* **400**, 841–847.

Burkhardt, N., Junemann, R., Spahn, C. M. T., and Nierhaus, K. H. (1998). Ribosomal tRNA binding sites: Three-site models of translation. *Crit. Rev. Biochem. Mol. Biol.* **33**, 95–149.

Clark, B. F. C., Grunberg-Manago, M., Gupta, N., Hershey, J. W. B., Hinnebusch, A. G., Jackson, R. J., Maitra, U., Mathews, M. B., Merrick, W. C., Rhoads, R. E., Sonenberg, N., Spremulli, L. L., Trachsel, H., and Voorma, H. O. (1996). Prokaryotic and eukaryotic translation factors. *Biochimie* **78**, 1119–1122.

Clemons, W. M., Jr., May, J. L. C., Wimberly, B. T., McCutcheon, J. P., Capel, M. S., and Ramakrishnan, V. (1999). Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature* **400**, 833–840.

Merrick, W. C., and Nyborg, J. (2000). The protein biosynthesis elongation cycle. In *Translational Control of Gene Expression* (N. Sonenberg, J. W. B. Hershey and M. B. Mathews, eds.) pp. 89–125. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Moazed, D., and Noller, H. F. (1989). Intermediate states in the movement of tRNA in the ribosome. *Nature* **342**, 142–148.

Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B. F. C., and Nyborg, J. (1995). Crystal structure of the ternary complex of Phe-tRNA[Phe], EF-Tu, and a GTP analog. *Science* **270**, 1464–1472.

Proud, C. (2000). Control of the elongation phase of protein synthesis. In *Translational Control of Gene Expression* (N. Sonenberg, J. W. B. Hershey and M. B. Mathews, eds.) pp. 719–739. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Rodnina, M. V., and Wintermeyer, W. (2001). Fidelity of aminoacyl-tRNA selection on the ribosome: Kinetic and structural mechanisms. *Annu. Rev. Biochem.* **70**, 415–435.

BIOGRAPHY

William C. Merrick is a Professor in the Department of Biochemistry in the School of Medicine at Case Western Reserve University. For over 30 years he has studied the processes of protein synthesis initiation and elongation in eukaryotic systems. These studies have resulted in the identification, cloning, and characterization of approximately 15 different translation factors and have indicated a sequential utilization of these factors.

Anton A. Komar is a Senior Research Associate in Dr. Merrick's laboratory. Dr. Komar received his Ph.D. from Moscow State University. His research expertise is in molecular biology, yeast genetics, and protein chemistry. His efforts have led to the identification of an IRES (internal ribosome entry site) in the Ure2p mRNA, the first in yeast. He is continuing studies of the Ure2p mRNA, both as relates to IRES function and the ability of Ure2p to behave as a prion.

Translation Initiation in Bacteria: Factors and Mechanisms

Cynthia L. Pon and Claudio O. Gualerzi
University of Camerino, Camerino, Italy

Initiation of protein synthesis is a fundamental biological process which contributes greatly to fidelity, efficiency, and regulation of gene expression. Translation initiation is a multistep process in which start codon and consequently the reading frame of the mRNA are selected by the small ribosomal subunit (30S) through the decoding of initiator N-formyl-methionyl-tRNA (fMet-tRNA) by the initiation codon producing a "30S initiation complex" which subsequently either forms the "70S initiation complex" by its association with the 50S ribosomal subunit or—if incorrectly formed—dissociates into its individual components. The aminoacyl-tRNA encoded by the second codon of the mRNA is then bound, in a ternary complex with EF-Tu and GTP, to the A-site of the 70S initiation complex which bears fMet-tRNA in the P-site. The subsequent formation of the first peptide bond between N-formyl-methionine and the second aminoacyl-tRNA yields an "initiation dipeptidyl-tRNA" in the A-site. The first translocation event moves the dipeptidyl-tRNA to the P-site and the ribosome enters the elongation phase of translation. The entire translation initiation process is kinetically controlled by three proteins, the initiation factors IF1, IF2, and IF3.

Properties, Aminoacylation, and Formylation of the Initiator tRNA

In bacteria protein synthesis begins with a special brand of tRNA, the initiator tRNA$^{\text{fmet}}$. Apart from being recognized by the same Met-tRNA synthetase (the main recognition element being the CAU anticodon), the initiator tRNA is endowed with a number of structural and functional properties which make this particular RNA one of a kind. The main features distinguishing the initiator tRNA from elongator tRNAs is the extended G–C helix in its anticodon stem, which favors its interaction with the ribosomal P-site, and the unpaired 5'-end base which allows the recognition of Met-tRNA$^{\text{fmet}}$ by the transformylase. The recognition of fMet-tRNA by IF2 relies on the presence of a formyl group blocking the α-NH$_2$ group of the amino acid

and requires just a few bases of the 3' acceptor end of the molecule; indeed, fMet-ACCAAC has almost the same affinity for IF2 as the intact fMet-tRNA.

The Translation Initiation Region of Prokaryotic mRNAs

The "translation initiation region" (TIR) of the mRNAs invariably contains an initiation triplet from which the ribosomes begin to translate and which sets, at least provisionally, the reading frame of the template (a reading frameshift may occur during elongation). Whereas the most frequently used initiation triplet is AUG, ~10% of all known genes and open reading frames begin with rarer yet "canonical" codons such as GUG and UUG, and, even less frequently, with "noncanonical" triplets such as AUU, AUC, and AUA. However, regardless of its nature, the initiation triplet is always decoded by the initiator fMet-tRNA. The reason for this initiation codon degeneracy is not always clear but two facts should be recalled. First, AUG is the only triplet giving rise to a "best-fit" Watson–Crick base-pairing with the anticodon (CAU) of the initiator tRNA, while all other start codons are expected to yield either 3' or 5' wobbling interactions with the same anticodon. Second, if not resulting from neutral mutations, the rare initiation triplets must be considered (or suspected to be) important regulatory signals capable of controlling timing and/or efficiency of translation. It is not by chance that, as described below, initiation factor IF3 can discriminate against translation beginning at noncanonical start codons; the best known example of translation regulation based on the use of a rare codon is the translational autorepression of IF3 mRNA which begins with an AUU initiation codon.

Aside from some "leaderless" mRNAs, which are occasionally found in (mainly Gram-positive) bacteria, which begin with a 5' AUG, the TIR of most mRNAs (Figure 1) contains, upstream of the initiation triplet a 5' untranslated region (5' UTR) of variable length
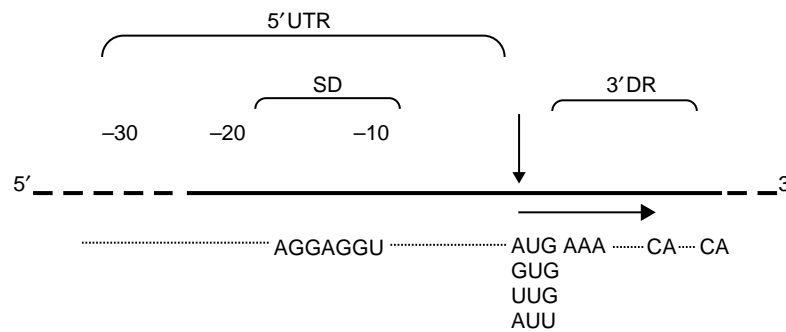
FIGURE 1 Schematic representation of the features characterizing a bacterial leader-containing mRNA translation initiation region (TIR). The arrows indicate the beginning of the coding region of the mRNA with four of the most common initiation triplets. To the left of the start point is the 5′ untranslated region (5′ UTR) with the approximate location of the Shine–Dalgarno sequence (SD) whose most extended consensus sequence is indicated below. To the right of the start point is the 3′ downstream region (3(DR)) with the most favorable second codon AAA and the repeated CA dinucleotides acting as translational enhancers in some mRNAs. Leaderless mRNA would start at the initiation codon.

and structure which generally includes a purine-rich Shine–Dalgarno sequence (SD sequence) complementary to the anti-SD sequence at the 3′-end of 16S rRNA. The SD sequence and the initiation triplet are separated by a spacer of variable length (optimally five nucleotides). Also the sequence at the 3′ side of the initiation triplet, the downstream region (DR), can have a strong influence on TIR selection. Highly expressed mRNAs seem to have a bias in favor of an AAA triplet as the second codon and a CA repeat sequence within the first 15 codons was shown to contribute to highly efficient initiation. Finally, important elements which contribute to the efficiency of translation initiation are also the secondary and tertiary RNA structures of the TIRs which may favor or disfavor formation of the 30S initiation complex and very often influence mRNA stability. In general, mRNAs whose TIRs are devoid of secondary structures tend to be translated with greater efficiency, although it is the optimal combinations of all the aforementioned TIR elements which can maximize translation of mRNA.

Initiation Complex Formation

Although the structural elements of the mRNA, such as the SD sequence of the TIRs, may contribute to the thermodynamic stability of the productive 30S–mRNA interaction, thereby favoring the selection of the correct start site of the mRNA, the initiation site is selected kinetically by the 30S ribosomal subunit with the aid of the initiator fMet-tRNA whose anticodon base pairs with the initiation codon of the mRNA forming a ternary complex (the "30S initiation complex") comprised of the 30S ribosomal subunit, fMet-tRNA and mRNA. Both conventional and fast kinetic analyses have contributed to the elucidation of the mechanistic

aspects of the translation initiation process schematically illustrated in Figure 2, which summarizes the steps leading to the formation of the 30S and 70S initiation complexes and the late events of translation initiation. In this pathway the 30S ribosomal subunit with a full complement of initiation factors interacts (steps A and B′ or B and A′) with mRNA and fMet-tRNA in stochastic order forming first two "binary complexes" and then an unstable "pre-ternary complex" in which both mRNA and fMet-tRNA are 30S-bound without interacting with each other. A first-order isomerization of this pre-ternary complex kinetically controlled by the three initiation factors (step C) causes the mRNA start codon to base pair in the P-site of 30S with the anticodon of fMet-tRNA to yield a "30S initiation complex." The specific role played by the individual initiation factors in this as well as in other steps of initiation will be described below. The "70S initiation complex" is generated by the joining of the large (50S) ribosomal subunit (step D), a process which induces a conformational change in the 30S subunit and thereby causes the ejection of IF1 and IF3. The intrinsic GTPase activity of IF2 (step E) is also triggered in this step, generating an IF2–GDP complex and inorganic phosphate. The latter is then released from ribosome in a step (step F) which possibly entails also the dissociation of IF2–GDP. The dissociation of IF2 leaves fMet-tRNA in the ribosomal P-site with the acceptor end near the peptidyl transferase center of the 50S subunit. The binding and the adjustment of the cognate EF-Tu·GTP·aa-tRNA to the ribosomal A-site is a multistep and overall rapid process (steps G, H, I, J) which is followed by the formation of the initiation dipeptide (step K) with the P-site-bound fMet-tRNA. Formation of the initiation dipeptide is a fairly slow process compared to the rate of transpeptidation during elongation.
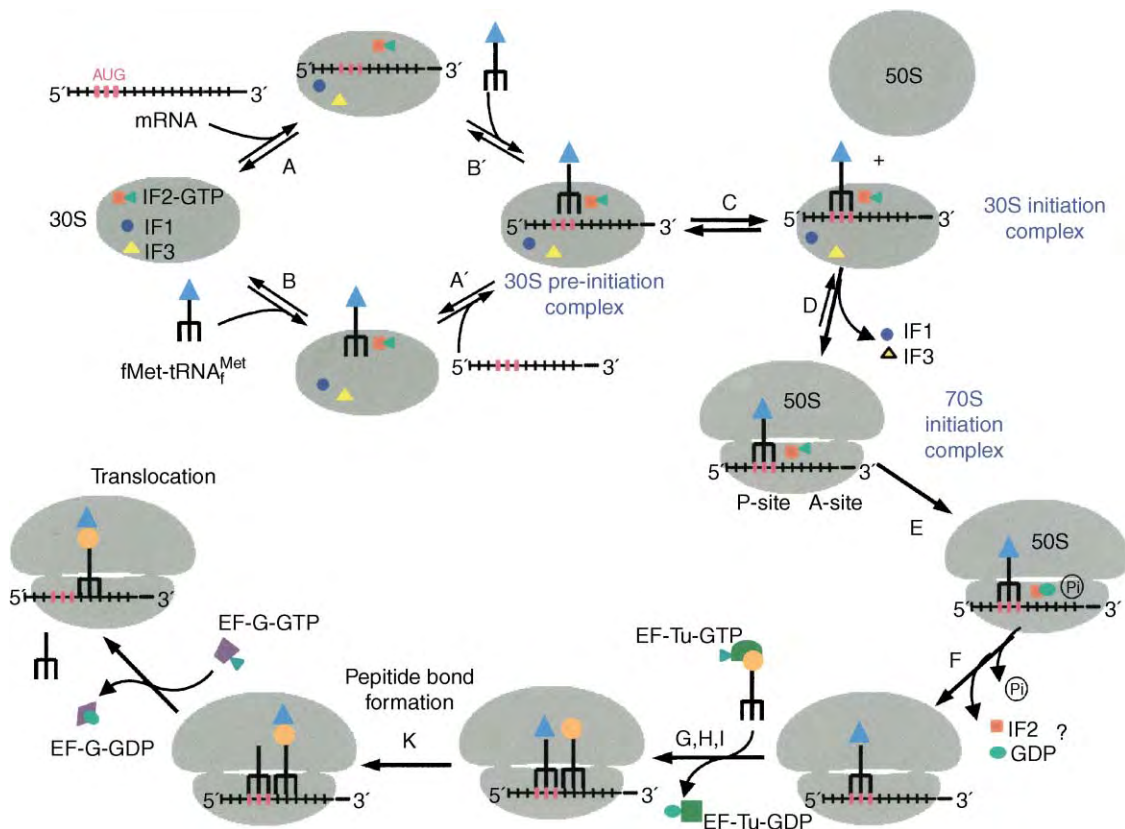
FIGURE 2   Initiation pathway. Scheme depicting the translation initiation pathway in bacteria as determined over the past two decades, mainly through traditional and fast kinetics analyses. Further details can be found in the text. The symbols and the denominations of the intermediate complexes are as indicated in the figure.

Structure and Function of the Initiation Factors

Overall, the three initiation factors ensure speed and accuracy in the initiation phase of protein synthesis. The properties and specific roles of these proteins are outlined below.

IF1

Structure

This protein (encoded by *infA*) consists of ∼70 amino acids (71 in *Escherichia coli*) and is characterized by a rigid five-stranded β-barrel structure from which protrude the short disordered and highly flexible N- and C-terminal tails (Figure 3A). This structural motif, known as the "OB fold," is characteristic of a class of proteins that interact with oligonucleotides and RNA (like IF1) or with oligosaccharides.

Topographical Localization and Function

IF1 binds to the 30S ribosomal subunit mainly through electrostatic interactions involving the positively charged surface of the protein and the phosphate backbone of specific regions of 16S rRNA. Earlier topographical studies and more recent chemical probing and crystallographic data indicate that IF1 binds in the A-site of the 30S ribosomal subunit. More precisely, IF1 fits in the cleft between ribosomal protein S12, the 530 loop and helix 44 of 16S ribosomal RNA (rRNA) establishing contacts with two functionally important bases (A1492 and A1493) which belong to this rRNA helix. This ribosomal localization could cause IF1 to block the premature access of aminoacyl-tRNA to the A-site during 30S initiation complex formation and supports the premise that IF1 contributes to the fidelity of translation initiation. The existence of a mutual influence of IF1 and IF2 on their respective interactions with the 30S ribosomal subunit is well established so that IF1 is regarded as being a modulator of IF2 recycling on and off the ribosomes. IF1 also increases the rates of association/dissociation of ribosomal subunits and this activity is probably a decisive factor in favoring the otherwise inefficient ribosome dissociation activity of IF3. Finally, IF1 can also increase the rate of 30S initiation complex formation, probably through a conformational change of the small ribosomal subunit. Indeed, there is compelling evidence from the
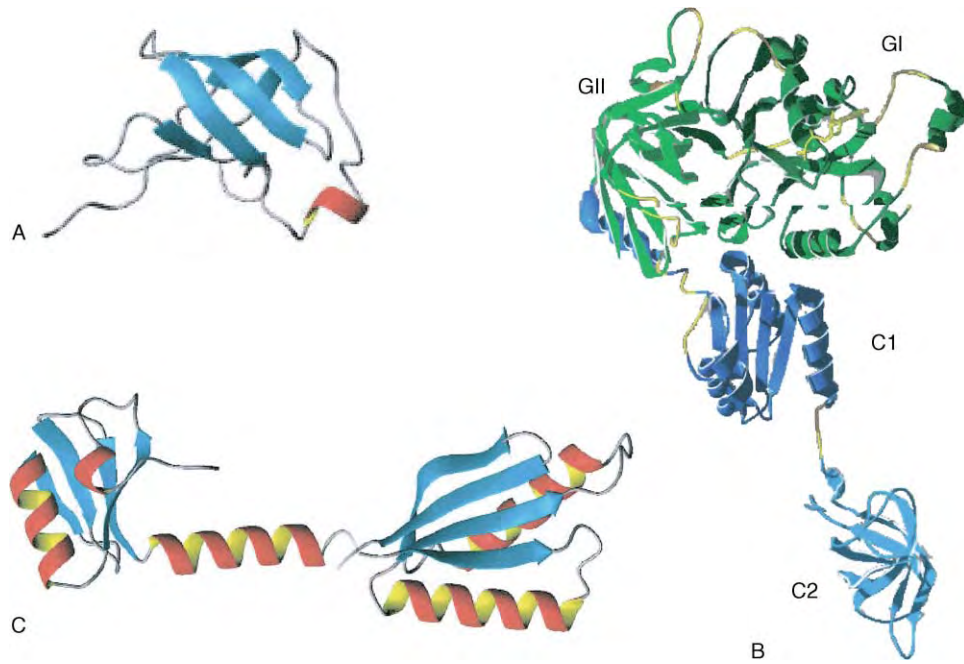
FIGURE 3 Three-dimensional structures of the three translation initiation factors. (A) β-Barrel structure of *E. coli* IF1 in solution as determined by NMR spectroscopy, (B) predicted structure of bacterial IF2 based on the crystal structure of archaeal eIF5b and on the solution structure of *Bacillus stearothermophilus* IF2C2 determined by NMR spectroscopy. The N-terminal domain is missing while the other domains are indicated with the following color patterns. Dark green, GI, light green GII, dark blue, C1, light blue, C2. (C) Crystallographic structure of *B. stearothermophilus* IF3 displaying the N domain (left), linker (center), and the multi-functional C-domain (right).

crystallographic data that IF1 binding induces several localized and long-range changes of the 30S structure. Overall, IF1 induces a rotation of head, platform and shoulder of the 30S subunit towards the A-site.

IF2

Structure

Bacterial IF2 (encoded by *infB*) is the largest initiation factor (890 residues in *E. coli*) and consists of three major parts: (i) a variable N-terminal region, (ii) a highly conserved 40 kDa region containing two domains, GI and GII, and (iii) the C-terminal region (25 kDa) which also consists of two domains, C1 and C2.

So far no three-dimensional (3D) structure has been directly determined for the whole bacterial IF2 molecule, and structural information concerning this factor relies on the NMR structure of the C2 domain of *Bacillus stearothermophilus* IF2 and on the crystal structure of *Methanobacterium thermoautotrophicum* eIF5B (Figure 3B), the archaeal homologue of IF2. However, although it is likely that the structures of the archaeal and bacterial factors are similar, it should be noted that the functions of these two proteins have very little in common. Overall, IF2 is an extended molecule containing four domains arranged in a unique architectural motif resembling a chalice whose cup is constituted

by GI, GII, and N-terminal part of C1, the stem by the C-terminal part of the same domain and the foot by the C2 domain. No structural information is available for the hydrophilic, positively charged and likely flexible N-domain which is not present in the archaeal protein as well as in some bacterial IF2 molecules. The likely function of this domain, which is dispensable for translation both *in vitro* and *in vivo*, is that of anchoring the factor on the 30S ribosomal subunit in a more or less specific way until the acceptor end of an fMet-tRNA molecule is "captured" by the C2 domain of the factor. The GI-domain binds GTP/GDP and is highly homologous to small GTPases and to equivalent domains of other GTP/GDP binding proteins such as elongation factors EF-Tu and EF-G. The GI domain is also responsible for the interaction of IF2 with the 50S ribosomal subunit and probably contains the GTPase center of the protein which is naturally activated by the interaction with the ribosomes.

The structure of this domain consists of an eight-stranded β-sheet flanked by six α-helices and a 3_{10} helix. As mentioned above, this domain contains the four conserved sequence elements characteristic of GTP-binding proteins (G1/P loop, which participates in phosphate binding, the G3 and G4 loops, forming the walls of a hydrophobic pocket where the guanine moiety of GTP or GDP is bound and G4). The Switch 2 region

of the GI domain is centrally located in the cup of the chalice where it makes extensive contacts to domains GII and C1 and can therefore cause conformational changes resulting in profound global rearrangements of the molecule that can reach as far as 90Å to the C-terminal region of IF2. The GII domain, which together with C1 is responsible for the interaction of IF2 with the 30S ribosomal subunit, has a β-barrel structure very similar to that of the C2 domain. Three β-strands of this domain interact closely with the GI-domain in the vicinity of the Switch 2 region. The GII domain is connected by a 17 residue α-helix to the C1-domain which is characterized by a unique α−β−α sandwich fold consisting of a four-stranded parallel β-sheet flanked on both sides by two α-helices. As mentioned above, the last 40Å long helix (H12) of the C1 domain extends from the cup to the C2 domain thus forming the "stem" of the chalice. The C2 domain of IF2, which is responsible for the specific recognition and binding of the acceptor end of fMet-tRNA is endowed with a structure which also consists of an eight-stranded β-barrel fold similar to that of GII and to domain II of EF-Tu and EF-G.

Topographical Localization and Function

IF2 is the only one of the three factors displaying a specific and fairly high affinity for both ribosomal subunits, and its interaction with the isolated 50S subunit is sufficient to elicit its GTPase activity. Recent experiments localize IF2 in a region of the 30S subunit topographically adjacent to the A-site, on a surface of the subunit's body facing the factor-binding region of the 50S subunit. With respect to the 50S subunit, IF2 was found to influence the chemical reactivity and/or the accessibility to nucleolytic cleavage of bases belonging to helix 89, to the sarcin-ricin domain (SRD) and to the L11/Thiostrepton-binding region of 23S rRNA leading to the conclusion that its topographical localization is on the right edge of the subunit interface site of the particle and at least partly overlaps that of elongation factors EF-G and EF-Tu.

The main function of IF2 is that of recognizing and binding (K_d in the μM range) the initiator fMet-tRNA and to stimulate (through an increase of the on-rate) its binding to the ribosomal P-site. Both specificity and thermodynamic stability of the IF2-fMet-tRNA interaction are properties of the C2 domain (~ 11 kDa). Additional IF2 functions include the stimulation of subunit association and the positioning of fMet-tRNA in the ribosomal P-site of the 70S initiation complex which favors the first transpeptidation. Furthermore, IF2 is a GTP/GDP-binding protein and a ribosome-dependent GTPase like EF-Tu and EF-G but, unlike these elongation factors, the function of the IF2-dependent GTP binding and hydrolysis is difficult to pin down.

Thus, since neither GTP/GDP-binding nor GTPase activity seems to be mandatory for any translational function of IF2 and since the "metabolic alarmone" ppGpp can bind in place of GTP and inhibit the IF2-dependent 30S initiation complex formation and initiation dipeptide synthesis, it has been postulated that IF2 uses its GDP/GTP-binding site as a receptor for GTP (under optimal growth conditions) or for ppGpp (during nutritional stress) and accordingly behaves like a sensor of the metabolic state of the cell. This raises the interesting possibility that, in addition to and because of its roles in translation initiation, IF2 might function as a global regulator linking translational activity to the transcriptional control of stable RNA synthesis by adjusting the translational rate of the cell as a function of the allowable growth rate.

IF3

Structure

The structure of this medium-sized protein encoded by *infC* (180 amino acids in *E. coli*) is characterized by the presence of two domains of approximately equal mass connected by a long (~ 45Å) lysine-rich linker (Figure 3C). Whereas considerable controversy exists between crystallographic and NMR data as to whether this linker is a long and rigid α-helix or unstructured and flexible, the 3D structures of both N-terminal (IF3N) and C-terminal (IF3C) domains seem to be well established. IF3N contains a globular α/β fold consisting of a single α-helix, packed against a mixed four-stranded β-sheet. IF3C possesses a two-layered α/β sandwich fold, comprising a four-stranded β-sheet which is packed against two parallel α-helices in a βαβαββ topology. The fold of IF3C is similar to that found in many eukaryotic RNA-binding proteins (such as U1A) and indeed IF3C interacts with the 30S subunit via a protein–RNA interaction involving primarily structural elements of this domain like strands β-7 and β-9 that contain consensus RNP motifs and two loops (L7 and L8). Regardless of the actual structure of the linker, several lines of evidence indicate that the two domains of IF3 do not interact with one another in the free or in the ribosome-bound protein and that they interact, independently of each other, with different sites of the 30S subunit.

Topographical Localization and Function

The main interaction of IF3 with the 30S subunit occurs via IF3C, the domain which encompasses all IF3 activities. Whereas there is good agreement that IF3C is localized on the platform of the 30S ribosomal subunit, the ribosomal localization of IF3N is more controversial, although it seems likely that its binding site is located somewhere on the head of the particle.

Binding of IF3 to the 30S ribosomal subunit interferes with subunit association thereby shifting to the left of the equilibrium $30S + 50S \rightleftharpoons 70S$. In turn, this increases the pool of free 30S which are amenable to initiate a new round of translation. Furthermore, the presence of IF3 on the 30S increases the on-rate of "pre-ternary complex" isomerization which leads to 30S initiation complex formation (Figure 2) and thereby stimulates overall mRNA translation. However, since IF3 increases also the off-rate of the isomerization, in this context it can promote initiation fidelity favoring the dissociation from the 30S subunits of both tRNA and template. Indeed, the dissociation occurs with different rates depending on whether the 30S complex is a "canonical," a "noncanonical," a "leaderless," or a "pseudo" initiation complex and is in kinetic competition with the association of the same 30S initiation complex with the 50S ribosomal subunit which yields a "70S initiation complex" (Figure 2). The potential 30S initiation

complexes recognized and accepted as "correct" by IF3 and those kinetically discriminated against as being "incorrect" are schematically represented in Figure 4. Contrary to earlier interpretations, IF3 is not required for mRNA binding to the ribosome but can promote a re-positioning of 30S-bound mRNA shifting it from the "stand by site" to the "P-decoding site" of the subunit. It is likely that this activity may be correlated to the kinetic selection of the correct initiation triplet by IF3. It is noteworthy that isolated IF3C not only binds to 30S subunit but is also capable of performing all the other known functions of the intact molecule while isolated IF3N has no autonomous function. Since the affinity of IF3C for the 30S subunits is approximately two orders of magnitude lower than that of the intact molecule and since secondary contacts established by IF3N stabilize the interaction, it has been suggested that the two-domain structure is required to modulate binding and release of IF3 to and from the 30S subunit.
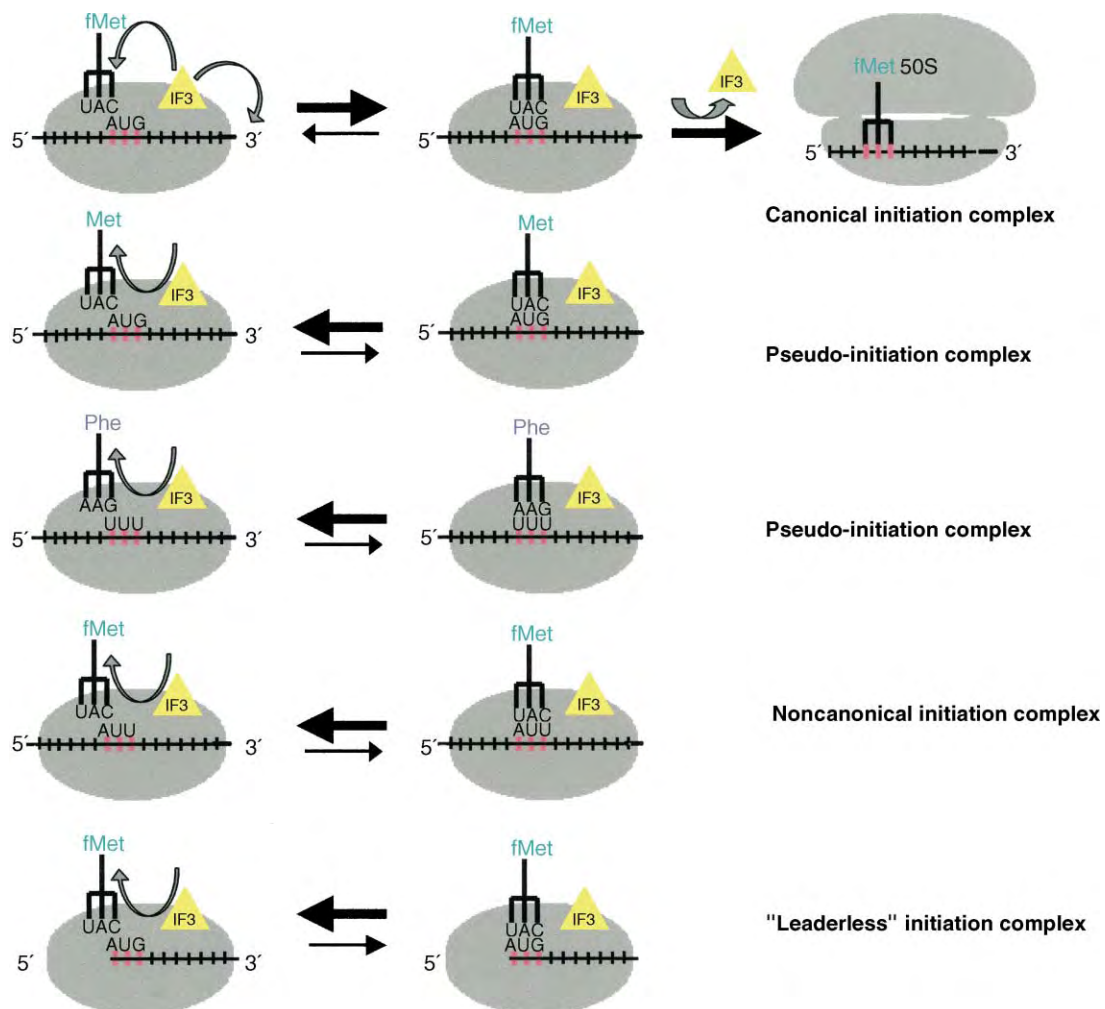
FIGURE 4 Initiation complexes recognized as "correct" and "incorrect" by IF3. The figure presents a scheme of the nature of the various types of complexes containing 30S ribosomal subunit, template and aminoacyl-tRNAs which are subjected to kinetic positive or negative discrimination by IF3. Further details may be found in the text.

The conformational change of the ribosome induced by subunit association would push platform and head of the 30S subunit away from each other thereby widening the gap between the IF3N- and the IF3C-binding sites. The loss of the stabilizing interaction established by IF3N would then facilitate the dissociation of IF3 from the 30S subunit.

SEE ALSO THE FOLLOWING ARTICLES

EF-G and EF-Tu Structures and Translation Elongation in Bacteria • Ribosome Assembly • Ribosome Structure • Translation Elongation in Bacteria • Translation Termination and Ribosome Recycling

GLOSSARY

A-site Aminoacyl-tRNA site, the binding site on the ribosome occupied by the tRNA carrying the next amino acid to be added to a growing polypeptide chain.

P-site Peptidyl-tRNA site, the binding site on a ribosome occupied by the tRNA carrying a growing peptide chain.

Watson–Crick base-pairing Association of two complementary nucleotides in a DNA or RNA molecule stabilized by hydrogen bonding between their base components.

wobbling interaction Ability of a tRNA to recognize more than one codon by unusual (non-G-C, non-A-U) pairing with the third base of a codon.

FURTHER READING

Boelens, R., and Gualerzi, C. O. (2002). Structure and function of bacterial initiation factors. *Curr. Protein Pep. Sci.* **3**, 107–119.

Gualerzi, C. O., Brandi, L., Caserta, E., Garofalo, C., Lammi, M., La Teana, A., Petrelli, D., Spurio, R., Tomsic, J., and Pon, C. L. (2001). Initiation factors in the early events of mRNA translation in bacteria. *Cold Spring Harbor Symp. Quant. Biol.* **66**, 363–376.

Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* **108**, 557–572.

BIOGRAPHY

Cynthia L. Pon is a Professor of Molecular Genetics at the University of Camerino, Italy. She holds a Ph.D. from Rutgers University and her current research interest concerns the regulation of gene expression during cold stress.

Claudio O. Gualerzi is a Professor of Molecular Biology in the Department of Biology at the University of Camerino, Italy. He holds a laurea degree from the University of Rome and his current research is directed toward the search for new antibiotics.

Translation Initiation in Eukaryotes: Factors and Mechanisms

Tatyana V. Pestova
Moscow State University, Moscow, Russia

Christopher U.T. Hellen
State University of New York Downstate Medical Center, Brooklyn, New York, USA

Translation initiation, the first stage in protein synthesis, is the process of assembly of large (60S) and small (40S) ribosomal subunits to form an 80S ribosome containing initiator tRNA (Met-tRNA$_i^{Met}$) that is base paired to the initiation codon of a mRNA in the ribosomal peptidyl (P) site. This process is mediated by at least 11 eukaryotic initiation factors (eIFs) and proceeds via the sequential formation of intermediate complexes. Initiation is the rate-limiting step of translation, and is a major focus for pathways that regulate gene expression.

The Structure of Eukaryotic mRNAs

Nearly all eukaryotic mRNAs have a 5′-terminal 7-methylguanosine (m^{7}G) "cap" and a 3′-terminal poly(A) tail that synergistically enhance the efficiency of translation initiation. Most eukaryotic mRNAs contain a single major open reading frame (ORF) that is translated into protein, and initiation usually begins at the first AUG triplet from the 5′-end of a mRNA, which follows a 5′ leader that is < 100 nucleotide long. An AUG triplet can be bypassed if its context deviates from the optimal sequence GCC(**A/G**)CC**AUGG** (in which the initiation codon is underlined and the nucleotides in bold have the greatest influence), if it occurs very close to the 5′-end of an mRNA or if the 5′ leader has little secondary structure. Stable structures in the 5′ leader reduce initiation efficiency whereas stable structures downstream of an AUG codon can enhance initiation at it.

The Mechanism of Translation Initiation

Translation initiation on most eukaryotic mRNAs begins with binding of Met-tRNA$_i^{Met}$ to a 40S subunit, followed by ribosomal attachment at the 5′-end of a mRNA, scanning to the initiation codon and joining with a 60S subunit to form an 80S ribosome. Initiation is mediated by at least 11 eIFs (Table I), many of which act at multiple stages in this process (Figure 1). Initiation on a few mRNAs occurs by noncanonical mechanisms, of which the most common is 5′-end independent internal ribosomal entry.

DISSOCIATION OF RIBOSOMES INTO FREE 40S AND 60S SUBUNITS

Initiator tRNA and mRNA initially bind to the 40S subunit rather than to the 80S ribosome. However, association of 40S and 60S subunits to form empty 80S ribosomes is favored under ionic conditions in the cytoplasm and a mechanism to maintain a pool of free subunits is therefore a prerequisite for initiation. eIF1A and eIF3 shift the equilibrium between ribosomes and their subunits towards dissociation. eIF3 dissociates 80S ribosomes and, with eIF1A, prevents subunit reassociation by binding directly to free 40S subunits.

RECRUITMENT OF INITIATOR tRNA TO THE 40S RIBOSOMAL SUBUNIT

The initiation codon is decoded by a unique initiator methionyl-tRNA. Its sequence and structural features distinguish it from the methionyl-tRNAs that decode AUG triplets during translation elongation. These features enable eIF2 to select Met-tRNA$_i^{Met}$ from the cytoplasmic pool of aminoacylated and deacylated initiator and elongator tRNAs, and also exclude Met-tRNA$_i^{Met}$ from translation elongation. eIF2 is a stable heterotrimeric protein consisting of α-, β- and γ-subunits that binds GTP and Met-tRNA$_i^{Met}$ to form a ternary complex. Binding of the ternary complex to a 40S subunit is strongly stabilized by eIF1A and eIF3 and yields a 43S preinitiation complex. eIF2's activity in binding Met-tRNA$_i^{Met}$ is regulated; its affinity for Met-tRNA$_i^{Met}$ is enhanced by prior binding of GTP

TABLE I

Mammalian Initiation Factors

Factor	Subunits	Mass (kDa)[a]	Functions
eIF1		13	Promotes scanning and the fidelity of initiation codon recognition
eIF1A		17	Ribosome antiassociation; stabilizes Met-tRNA$_i$ binding to 40S subunit; promotes scanning
eIF2	α, β, γ	36, 39, 52	Binds GTP and Met-tRNA$_i$ to 40S subunit; GTPase
eIF2B	$\alpha, \beta, \gamma, \delta, \varepsilon$	34, 39, 50, 58, 80	Guanine-nucleotide exchange factor for eIF2
eIF3	a–l	167, 105, 99, 64, 52, 38, 35, 40, 37, 29, 25, 67	Ribosome dissociation and ribosome subunit antiassociation; stabilizes binding of the eIF2/GTP/Met-tRNA$_i$ complex to 40S subunit; required for ribosomal binding to mRNA and scanning on the 5′ leader
eIF4A		44	ATP-dependent RNA helicase/RNA-dependent ATPase
eIF4B		69	mRNA-binding cofactor for eIF4A
eIF4E		25	m^{7}G "cap"-binding protein
eIF4F, eIF4E, eIF4A, eIF4G		25, 44, 176	Cap-binding complex comprising eIFs 4A, 4E, and 4G
eIF4G		176	Binds and coordinates the functions of mRNA, PABP, and eIFs 3, 4A, 4E
eIF4H		25	mRNA-binding cofactor for eIF4A
eIF5		49	GTPase-activating protein specific for eIF2
eIF5A		17	May enhance first cycle of translation elongation
eIF5B		139	GTPase; ribosome subunit joining
PABP		70	Binds the 3′ poly(A) tail and promotes ribosomal binding to mRNA

[a] Masses (kDa) correspond to those of human proteins, and where appropriate, to the largest isoform.

whereas hydrolysis of eIF2-bound GTP towards the end of each initiation cycle yields eIF2·GDP which dissociates from Met-tRNA$_i^{Met}$, leaving it in the peptidyl (P) site of the 40S subunit. Biochemical and genetic data suggest that eIF2's γ-subunit binds Met-tRNA$_i^{Met}$ and GTP, but despite eIF2γ having the sequence motifs and structure characteristic of a conventional GTP-binding protein, eIF2 does not have an intrinsic GTPase activity. Hydrolysis of eIF2-bound GTP is induced by eIF5, a GTPase activating protein specific for eIF2.

ATTACHMENT OF 43S PREINITIATION COMPLEXES TO MRNA

Binding of initiator tRNA to form a 43S complex is an obligatory first step before a 40S subunit can bind mRNA, which for most mRNAs is a 5′ end-dependent process. Attachment of 43S complexes is enhanced synergistically by the 5′-terminal cap, the 3′ poly(A) tail and the factors that associate with these structures. The m^{7}G "cap" is bound by eIF4E, the cap-binding subunit of the heterotrimeric eIF4F, which also contains eIF4A and eIF4G subunits. eIF4A is an ATP-dependent RNA helicase that cycles in and out of the eIF4F complex. eIF4G is a large polypeptide that binds mRNA and eIF4E, eIF4A, and eIF3 and the cytoplasmic poly(A)-binding protein PABP, thereby coordinating and in some instances enhancing their activities. eIF4B and the less abundant eIF4H enhance the helicase activities of eIF4A and of eIF4F; eIF4B enhances but is not essential for ribosomal attachment to mRNA. The 3′ poly(A) tail's influence is dependent on PABP which binds to it, and synergism between the cap and the 3′ poly(A) tail depends on the bridging interaction of eIF4G with the m^{7}G cap/eIF4E and 3′ poly(A)/PABP complexes.

The eIF4F complex is required on most mRNAs to unwind the cap-proximal region to prepare it for attachment of 43S complexes. Recent studies have shown that 43S complexes can bind directly to a mRNA in the absence of eIF4F if the 5′-terminal region is completely unstructured. Ribosomal attachment likely involves protein–protein interactions between the eIF3 component of 43S complexes and both eIF4G and eIF4B as well as direct binding of eIF3, eIF4G, and the 40S subunit to the mRNA. ATP-dependent restructuring of mRNA and ribosomal attachment to the unwound region of the 5′ leader are probably coordinated by interaction of the eIF4G component of the eIF4F/PABP/mRNA complex with eIF3.
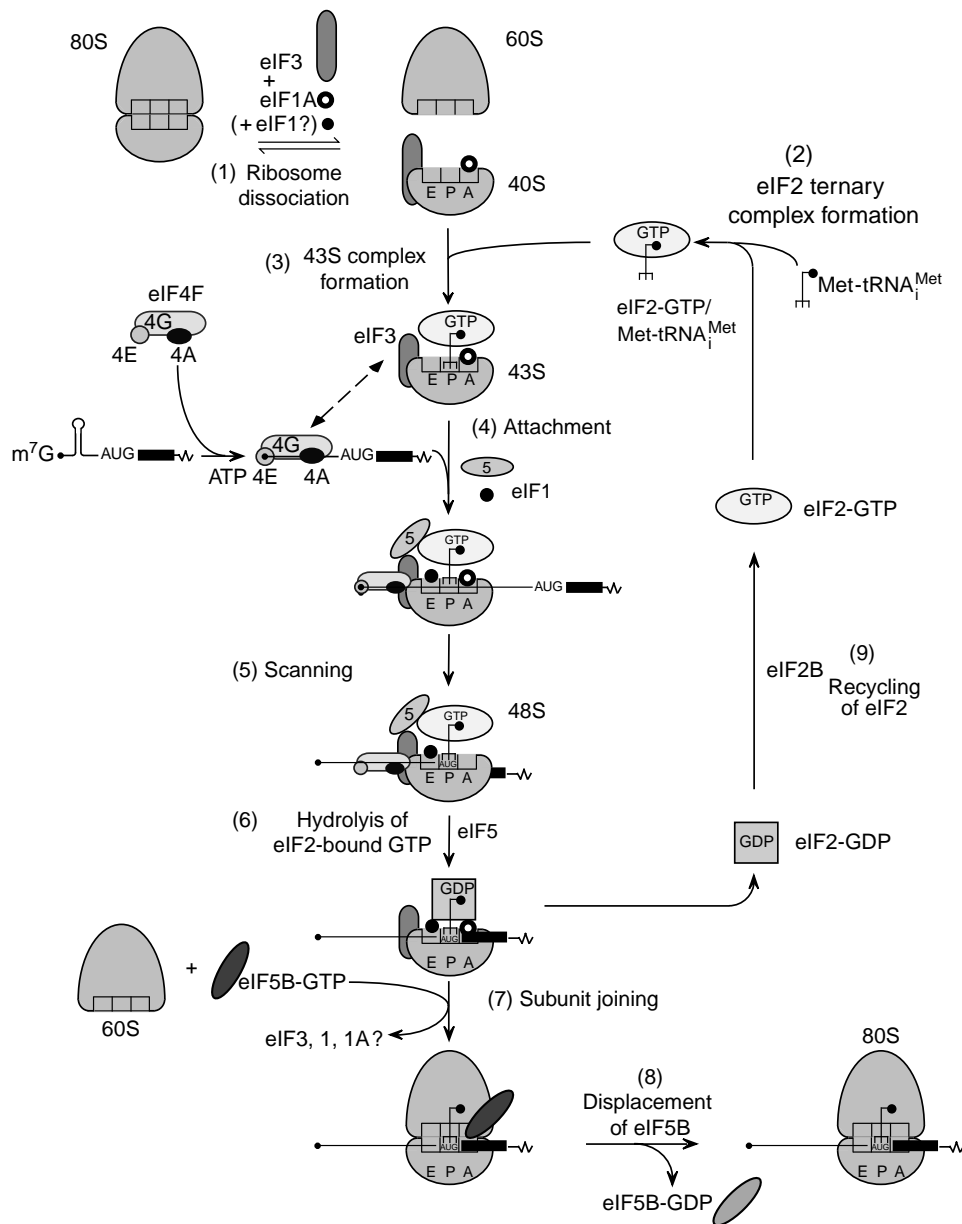
FIGURE 1   Schematic model of the pathway of 80S initiation complex formation on a model capped eukaryotic mRNA. eIF1A and eIF3 promote dissociation of the 80S ribosome into 40S and 60S subunits (Step 1). eIF2 binds aminoacylated initiator tRNA (Met-tRNA$_i^{Met}$) and GTP to form a ternary complex (Step 2). Binding of the ternary complex to a 40S subunit is stabilized by eIF1, eIF1A, and eIF3 to form a 43S complex (Step 3). The eIF4F complex binds to the 5′-terminal m^{7}G "cap" of an mRNA and with associated cofactors creates an unstructured cap-proximal site to which the 43S complex binds (Step 4). This step may be enhanced by the poly(A)-binding protein associated with the mRNA's 3′ poly(A) tail (not shown). The dashed, double-headed arrow indicates interactions that promote attachment of the 43S complex to mRNA, such as those of the eIF3 component of the 43S complex with the eIF4G subunit of cap-bound eIF4F and probably with the mRNA. The 43S complex scans downstream on the 5′-leader until it reaches the initiation codon (Step 5), which is base paired to the anticodon of Met-tRNA$_i^{Met}$ in the resulting 48S complex. Hydrolysis of eIF2-bound GTP is triggered by eIF5, and probably releases of eIF2-GDP (Step 6). The stages at which eIF5 joins and eIF1, eIF1A, and eIF3 are released from the 40S subunit are not known. The GTP-bound form of eIF5B mediates joining of a 60S subunit to the resulting complex (Step 7). eIF5B-GDP is released after hydrolysis of eIF5B-bound GTP, which is induced by both ribosomal subunits. The resulting 80S ribosome is then able to begin protein synthesis (Step 8). eIF2B recycles inactive eIF2-GDP to active eIF2-GTP, which can then bind Met-tRNA$_i^{Met}$ again (Step 9).

RIBOSOMAL SCANNING
ON THE 5′-LEADER

Following attachment to the 5′-terminal region of a mRNA, a 43S complex scans downstream until it locates the initiation codon, which is usually the first AUG triplet from the 5′-end. Scanning consists of ribosomal movement on the 5′-leader and inspection of it to identify the initiation codon. Recent studies have shown that ribosomal complexes containing only eIF1, eIF2, and eIF3 are intrinsically capable of ATP-independent movement on a 5′-leader if it is unstructured; eIFs 1, 1A, 4A, 4B, and 4F all contribute to the processivity of scanning. eIFs 4A, 4B and 4F, and ATP are involved in restructuring mRNA to permit scanning and are essential for scanning if the 5′-leader contains even weak-secondary structure. Many details of the mechanism of scanning remain poorly understood, including the identity of the set of factors associated with the scanning complex, the mechanism by which helicase-mediated unwinding of structured 5′-leaders is coupled to ribosomal movement, the stage at which eIF4F dissociates from the cap and from the 43S complex and the mechanism of action of eIF1 and eIF1A.

SELECTION OF THE INITIATION CODON

The scanning 43S complex stops when it encounters an AUG triplet that it recognizes as an initiation codon. This is primarily determined by base pairing between it and the anticodon of initiator tRNA, but in higher eukaryotes the sequence flanking an AUG triplet also influences its selection. Genetic analyses in yeast of mutants that permit initiation at an UUG triplet have indicated that eIF1, eIF2, and eIF5 all influence start site selection. Recent studies have shown that eIF1 plays the principal role in maintaining the fidelity of initiation codon selection, for example, by destabilizing complexes aberrantly arrested at non-AUG triplets or assembled on AUG triplets that have a poor context.

DISPLACEMENT OF eIF2
FROM THE 48S COMPLEX

Establishment of base pairing between the initiation codon and the anticodon of Met-tRNA$_i^{Met}$ stimulates eIF5-induced hydrolysis of eIF2-bound GTP in ribosomal complexes. eIF2·GDP does not bind Met-tRNA$_i^{Met}$ or the 40S subunit and is released from the 48S complex, leaving Met-tRNA$_i^{Met}$ base paired to the initiation codon. eIF5-induced hydrolysis of eIF2-bound GTP is involved in conversion of the scanning ribosomal complex to a complex that is arrested at the initiation codon. Mutations in eIF5 or eIF2 that alter the rate of this reaction therefore influence selection of the initiation codon. eIF2 has a much higher affinity for GDP than GTP, and GDP has a slow off-rate from eIF2, so eIF2B, a guanine nucleotide exchange factor specific for eIF2 is required to regenerate active eIF2·GTP from inactive eIF2·GDP.

60S SUBUNIT JOINING TO FORM
AN 80S RIBOSOME

The release of eIF2·GDP from 48S complexes that is induced by eIF5 is not sufficient to permit 60S subunits to bind to the 40S subunit/Met-tRNA$_i^{Met}$/factor complex assembled at the initiation codon. Subunit joining also requires eIF5B, which is a homologue of the prokaryotic initiation factor 2, and like it has a GTPase activity that is maximally stimulated by large and small ribosomal subunits. Subunit joining leads to the release of all initiation factors from the 40S subunit and leaves Met-tRNA$_i^{Met}$ in the ribosomal P site. The GTP-bound form of eIF5B is active in promoting subunit joining but does not dissociate from the ribosome and instead blocks the ribosomal A site. GTP hydrolysis leads to the release of eIF5B·GDP from the 80S ribosome, so that it is able to begin polypeptide synthesis. Translation initiation therefore requires hydrolysis of two GTP molecules in reactions catalyzed by eIF2γ/eIF5 and eIF5B, respectively. The initial round of elongation may be enhanced by eIF5A.

Regulation of Translation Initiation

The intrinsic efficiency of translation initiation on a mRNA is determined by properties such as the degree of secondary structure in the 5′-leader and the sequence context of the initiation codon. Differences in these properties account for many of the differences in the relative levels of translation of different mRNAs. Translation initiation is also regulated either selectively on a single species or a small subset of mRNAs, or globally on all mRNAs, in order to integrate protein synthesis with physiological demands. Translation of a single species of mRNA or of a related group of mRNAs can be repressed by binding of a protein to a specific site on the mRNA in a manner that interferes with initiation. For example, the iron regulatory protein that binds to the 5′ leader of ferritin mRNA sterically prevents it from binding to the 43S complex. More commonly, translation is regulated at a more global level by alteration of the activities of initiation factors, either as a result of phosphorylation or by binding to regulatory proteins. Phosphorylation of eIF2α causes eIF2 to bind to and inhibit eIF2B, ultimately reducing formation of the eIF2/GTP/Met-tRNA$_i^{Met}$ complex and thus down-regulating initiation globally. Initiation on most mRNAs is cap-mediated and is down-regulated by

a reduction in the level of active eIF4F, which can occur either by proteolysis of eIF4G (for example, during some viral infections) or by disruption of eIF4E's interaction with eIF4G by eIF4E-binding proteins that compete with eIF4G for binding to eIF4E.

Initiation of Translation by Internal Ribosomal Entry

Translation on most eukaryotic mRNAs is initiated by end-dependent ribosomal scanning, but initiation on a few viral mRNAs is end-independent and is instead mediated by an internal ribosomal entry site (IRES) that promotes binding of the 40S subunit to an internal site in the mRNA without scanning from the 5′-end. IRESs are in general large and contain significant secondary structure. IRESs from a single virus family are similar, but the size and structure of IRESs from unrelated virus families differ greatly from each other. Three groups of IRESs have been characterized in detail; each mediates initiation by a different mechanism but all involve direct, noncanonical interactions of the IRES with canonical components of the translation apparatus. They all have simpler initiation factor requirements than cap-mediated initiation, and therefore escape some mechanisms that regulate that process. IRESs have been identified in several cellular mRNAs, but little is known of the mechanisms by which they promote initiation.

SEE ALSO THE FOLLOWING ARTICLES

mRNA Polyadenylation in Eukaryotes • Pre-tRNA and Pre-rRNA Processing in Eukaryotes • Ribosome Assembly • Ribosome Structure

GLOSSARY

eukaryotic initiation factor   A protein that acts in one or more steps in the process of translation initiation.

initiator tRNA   The anticodon of initiator transfer RNA is complementary to the AUG initiation codon and its structural properties differentiate it from tRNAs that decode AUG triplets during elongation. Initiator tRNA is activated by covalent linkage to methionine to form methionyl-tRNA, which is the substrate used by ribosomes to initiate protein synthesis.

mRNA   The RNA template that is translated by ribosomes to synthesize a protein. Its coding sequence comprises consecutive triplet codons that are decoded by base pairing with the anticodon of aminoacylated tRNAs, and that therefore determine the amino acid sequence of the resulting protein.

ribosome   The complex macromolecule that catalyzes mRNA template-directed protein synthesis. Its two subunits both consist of ribosomal RNA and proteins. The 40S subunit binds mRNA and the anticodon end of tRNA; the 60S subunit aligns the aminoacyl ends of tRNAs and catalyzes peptide bond formation.

FURTHER READING

Dever, T. E. (2002). Gene-specific regulation by general translation factors. *Cell* **108**, 545–556.

Gingras, A.-C., Raught, B., and Sonenberg, N. (1999). eIF4 initiation factors: Effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* **68**, 913–963.

Hellen, C. U. T., and Sarnow, P. (2001). Internal ribosomal entry sites in eukaryotic mRNA molecules. *Genes Develop.* **15**, 1593–1612.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266**, 19867–19870.

Pestova, T. V., and Kolupaeva, V. G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Develop.* **16**, 2906–2922.

Sonenberg, N., Hershey, J. W. B., and Mathews, M. B. (eds.) (2000). *Translational Control of Gene Expression*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

BIOGRAPHY

Christopher U. T. Hellen is an Associate Professor in the Department of Microbiology and Immunology at the State University of New York Downstate Medical Center in New York. His principal research interests are in mechanisms of translation initiation by internal ribosomal entry. He holds a D.Phil. from Oxford University.

Tatyana V. Pestova is an Assistant Professor in the Department of Microbiology and Immunology at the State University of New York Downstate Medical Center in New York, and a senior research scientist in the A. N. Belozersky Laboratory of Physico-chemical Biology, Moscow State University, Moscow, Russia. Her principal research interests are in molecular mechanisms of translation initiation in eukaryotes. She holds a Ph.D. and a D.Sc. from Moscow State University.

Translation Termination and Ribosome Recycling

Nadja Koloteva-Levin and Mick F. Tuite
University of Kent, Canterbury, UK

The decoding of genetic information stored in an mRNA molecule is initiated by a ribosome locating the translation start codon AUG. The codons that define the encoded polypeptide chain are then read sequentially by the ribosome with the amino acids being delivered to the ribosome by tRNAs. The chain of amino acids is therefore assembled as directed by the order of codons in the mRNA which in turn are recognized by the anticodon sequence of the tRNA. Translation termination is the last stage in this process and results in the release of the completed polypeptide chain from the ribosome once the end of the coding sequence has been reached. The trigger for the termination process is the appearance of one of the three possible stop codons at the decoding center of the ribosome. Following termination the ribosome dissociates from the mRNA and is recycled to take part in a new protein synthesis cycle.

Translation Termination Apparatus in Prokaryotes and Eukaryotes

The facilitators of the final stage of the translation cycle are the protein release factors (RFs). Translation termination occurs when one or other of the three stop codons—UAA, UAG, or UGA—at the end of the mRNA's coding sequence are positioned at the ribosomal "acceptor" (A) site, i.e., at the decoding site within the small ribosomal subunit (Figure 1, step 1). The stop codon is recognized by an individual RF or by an RF complex that then triggers hydrolysis of peptidyl-tRNA from the nascent polypeptide chain positioned at the adjacent ribosomal "peptidyl" (P) site (Figure 1, step 2). This is mediated by the peptidyl-transferase center (PTC) of the large ribosomal subunit transferring the peptidyl group to water rather than to an aminoacyl tRNA.

RFs fall into one of two classes: class-I RFs recognize and bind to the stop codon while class-II RFs facilitate the release of the class-I factor after peptidyl-tRNA hydrolysis (Figure 1, step 3). In the final stage, a ribosome recycling factor (RRF) triggers the disassembly of the terminated ribosome from the mRNA and the release of the deacylated tRNA from the P ribosome (Figure 1, step 4). Table I summarizes the known protein factors involved in translation termination.

The ribosome also plays an active role in the process of termination: interaction of the small ribosomal subunit RNAs and ribosomal proteins with the stop codon and a class-I RF ensures that the stop codon is accurately decoded at the A site, whereas the interaction of large ribosomal subunit RNAs with the RF or the RF complex facilitates the catalysis of peptidyl-tRNA hydrolysis at the P site via the ribosomal PTC.

Release Factors

PROKARYOTIC POLYPEPTIDE CHAIN RELEASE FACTORS

In bacteria, translation termination is controlled by three different RFs (Table I). Two class-I protein release factors, RF1 and RF2, each decode two of three stop codons, UAA or UAG (RF1) and UAA or UGA (RF2). Recognition of the stop codon by the RFs is mediated via a conserved tripeptide motif: Pro-Ala-Thr (PAT) in RF1 and Ser-Pro-Phe (SPF) in RF2. These tripeptides are referred to as peptide anticodons. The other important functional domain of class-I RFs contains the highly conserved amino acid motif Gly-Gly-Gln (GGQ) and it is this domain that triggers hydrolysis of the protein-tRNA bond. An understanding of how RFs trigger the release of the completed polypeptide from the tRNA at the P site has come from a study of RF2 and its interaction with the terminating ribosome. When RF2 binds to the ribosome with a stop codon positioned at the A site, RF2 changes its three-dimensional conformation such that the domain with the conserved "peptide anticodon" (SPF) interacts with the mRNA at the decoding center, and the GGQ-containing domain comes in the contact with the ribosomal PTC to trigger hydrolysis of the peptidyl-tRNA linkage (Figure 2).
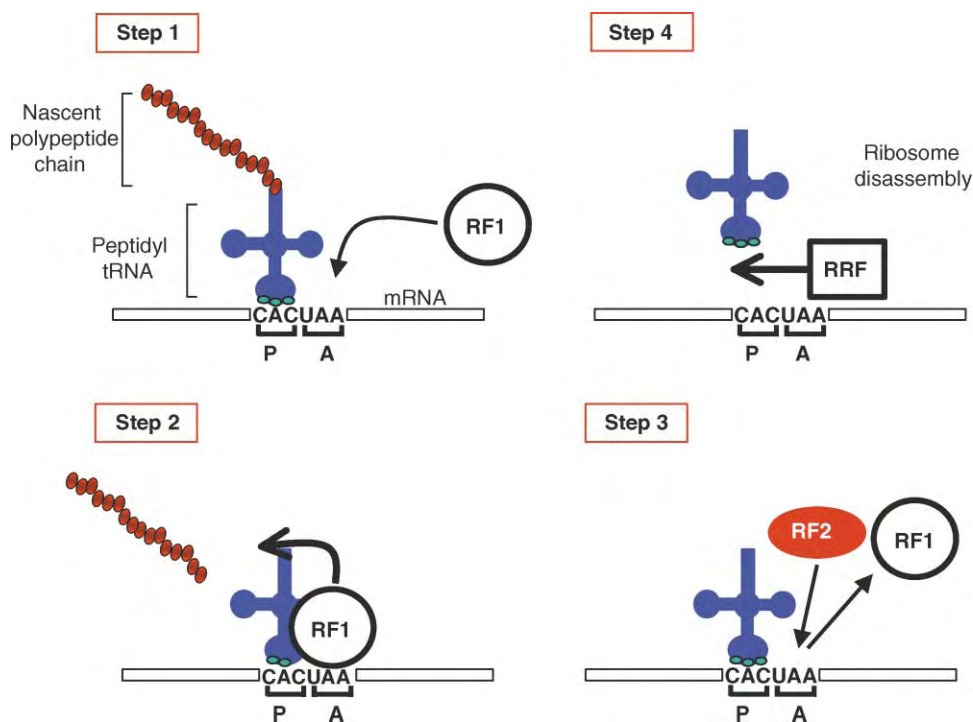
FIGURE 1   An overview of the mechanism of translation termination. The four key steps of the process triggered by the arrival of a stop codon at the ribosomal A site are shown. The peptidyl tRNA is positioned at the ribosomal peptidyl P site. The major protein factors associated with each step are indicated: RF1, class-I release factor; RF2, class-II release factor; RRF, ribosome recycling factor.

The single class-II RF in bacteria, namely RF3, is a GTP-binding protein that accelerates dissociation of either RF1 or RF2 from the ribosomal A-site after release of the completed polypeptide chain from the ribosome. RF3 bound to GDP accesses the ribosome

TABLE I

Protein Factors Involved in Translation Termination

Organisms	RF	Function
Archaebacteria		
	aRF1	Class-I RF, recognizes UAA/UAG/UGA
Prokaryotes		
	RF1	Class-I RF, recognizes UAA/UAG in mRNA
	RF2	Class-I RF, recognizes UAA/UGA in mRNA
	RF3	Class-II RF, GTPase, accelerates dissociation of termination complex from the ribosome
	RRF	Ribosome recycling factor
Eukaryotes[a]		
	eRF1	Class-I RF, recognizes UAA/UAG/UGA
	eRF3	Class-II RF, GTPase, increases termination efficiency, precise function unknown

[a] No RRF homologue has yet been described for eukaryotic cells other than a mitochondrially encoded homologue.

which, in complex with RF1 or RF2, acts as guanine nucleotide exchange factors (GEFs) and triggers dissociation of GDP from RF3. This leads to the formation of a stable ribosome-RF1 (or RF2-) RF3 complex. Hydrolysis of the peptidyl-tRNA linkage triggered by RF1 or RF2 allows GTP binding to RF3. This induces an altered RF3 conformation with a high affinity for the ribosome and leads to rapid dissociation of RF1 or RF2 from the termination complex. To leave the ribosome, RF3 requires GTP hydrolysis, which converts it to the GDP-bound form of RF3 which has a lower affinity for the ribosome. Once RF3 leaves the ribosome, it is ready to enter the next translation cycle.

Eukaryotic Chain Release Factors

In contrast to bacteria, in eukaryotic cells translation is terminated by a single heterodimer consisting of two different RFs, eRF1 and eRF3. eRF1 is a class-I RF that decodes all three stop codons and triggers peptidyl-tRNA hydrolysis by the ribosome to release the nascent polypeptide. In eRF1, the stop codon recognition site is located close to the amino terminus of the protein molecule in a region that contains an evolutionarily conserved tetrapeptide sequence, Asn-Ile-Lys-Ser (NIKS). This sequence may be functionally equivalent to the bacterial RF peptide anticodon. The GGQ motif found in bacterial RF1 and RF2 is located in a
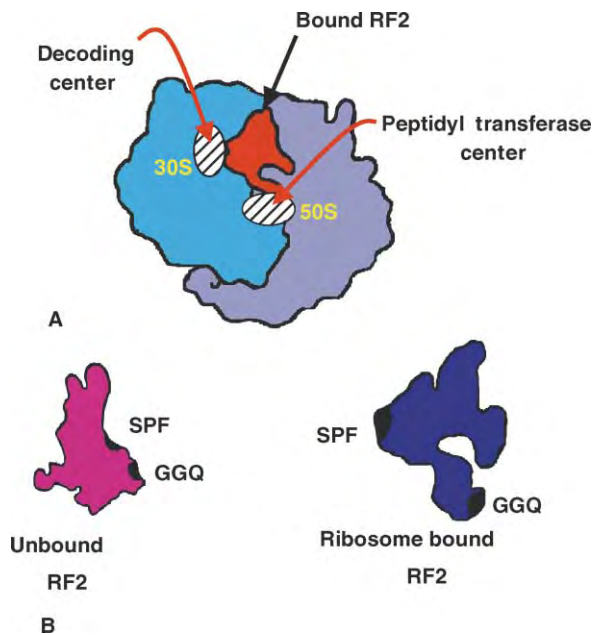
FIGURE 2   The three-dimensional conformation of a release factor changes once it is bound to the ribosome. (A) The RF2:bacterial ribosome complex indicating the points of contact between RF2 and the decoding center and the PTC of the ribosome. (B) The three-dimensional structures of the unbound and bound forms of bacterial RF2 indicating the location of the "peptide anticodon" sequence Ser-Pro-Phe (SPF) and the conserved Gly-Gly-Gln (GGQ) amino acid motif.

FIGURE 3   Many of the protein factors involved in translation elongation and termination have three-dimensional shapes that mimic a tRNA molecule. (A) Three-dimensional structure of a tRNA showing the location of the anticodon and the site to which the amino acid is covalently attached. (B) The three-dimensional structures of the mammalian RF eRF1 (left) and the bacterial translation elongation factor EF-G (right). The position of the conserved NIKS and GGQ amino acid motifs are indicated on eRF1, together with the domain to which the class-II factor eRF3 binds.

separate domain of the protein to the NIKS sequence. The carboxy-terminal part of the eRF1 binds eRF3. The crystal structure of human eRF1 shows that it is a Y-shaped molecule that resembles the structure of a tRNA (Figure 3). Since both bacterial and eukaryotic class-I RFs carry out essentially the same function, one would expect them to interact similarly with ribosomal A site. As with RF2 (see Figure 2), eRF1 must also undergo a conformational change after binding to the ribosome in order that it can interact with both the ribosomal decoding site and the ribosomal PTC.

The eukaryotic class-II RF, eRF3, forms a complex with eRF1, and via this interaction enhances the efficiency of the translation termination although its function remains to be fully defined. The GTPase activity of eRF3, which is triggered by stop codons, is both eRF1- and ribosome dependent. The carboxy-terminal half of the eRF3 molecule shows significant amino acid identity to the translation elongation factor eEF1A that is responsible for bringing the aminoacyl tRNAs to the ribosome during polypeptide chain elongation. This suggests that eRF3 may act in a manner analogous to eEF1A, a protein factor bringing the RF complex to the ribosome. A number of functions have been attributed to eRF3, e.g., it may displace eRF1 from the ribosome (i.e., the function assigned to prokaryotic RF3), or proofread the eRF1:stop codon interaction, or

stimulate ribosome disassembly (i.e., the function assigned to prokaryotic RRF; see below), although direct experimental proof for any one specific function is lacking. In mammalian cells, eRF3 binds to the poly(A)-binding protein (PABP), a protein associated with the poly(A) tails located at the $3'$ end of the majority of eukaryotic mRNAs. This interaction might be important for the regulation of both mRNA stability and translation initiation, since disruption of the eRF3::PABP interaction suppresses the recycling of the ribosome on the same mRNA.

In Bakers' yeast (*Saccharomyces cerevisiae*), the eRF3 protein (also known as Sup35p) has an additional remarkable property; it is a prion protein that acts as the protein determinant of a non-Mendelian genetic element called [PSI^+]. Like the mammalian prion protein PrP, eRF3 can exist in one of two different conformations with the aggregation-prone prion conformer catalyzing the conversion of the normal soluble form, to the prion form. In [PSI^+] cells the majority of eRF3 is found as an inactive, high molecular weight complex, and thus [PSI^+] cells have a defect in translation termination albeit without detriment to yeast cell growth. This mechanism may represent a novel means of regulating the efficiency of translation termination in yeast. There is no evidence that mammalian eRF3 is a prion.

In Archaebacteria, the single class-I RF found (aRF1) shares significant amino acid sequence and structural homology with eRF1, but not with either RF1 or RF2. Like eRF1, aRF1 is able to decode all three stop codons. No eRF3 homologue has been identified in any Archaebacterial genome.

Molecular Mimicry

RFs decode stop codons at the ribosomal A site and in that respect can functionally be compared to aminoacyl tRNAs which decode sense codons of the mRNA when positioned at the ribosomal A site during translation elongation. In so doing, RF1 and RF2 physically interact with the stop codons in the mRNA. That a common ribosomal binding site for tRNAs, translation elongation factors, and RFs exists has led to the suggestion that these protein and RNA molecules may also have similar three-dimensional structures, the so-called "molecular mimicry" model. The structural similarity—discovered by Clark, Nyborg, and colleagues in 1995—between bacterial elongation factor EF-G and the tRNA-EF-Tu-GTP complex provided direct evidence for this model. One of the structural domains of the EF-G protein strongly mimics the structure of the T-stem and anticodon loop of a tRNA (Figure 3). The "molecular mimicry" model can also be extended to include the RFs and confirmation of this came from the solution of the crystal structure of human eRF1, whose overall shape and dimensions resemble a tRNA molecule with the three clearly identifiable structural domains corresponding to the anticodon loop, the aminoacyl acceptor stem, and T-stem of a tRNA molecule (Figure 3). There is also significant amino acid sequence similarity between several RF domains and the carboxy-terminal portion of elongation factor EF-G, the region of the molecule that interacts with the ribosome.

Posttermination Events

Once the translation of an mRNA has been terminated and the completed polypeptide chain released, the ribosome must dissociate from the mRNA in order to be able to participate in a new round of protein synthesis. In bacteria, another protein factor, the RRF, is involved in this posttermination step together with the elongation factor that mediates ribosomal translocation, i.e., EF-G. The three-dimensional structure of RRF, like eRF1, also mimics a tRNA molecule which, in turn, is important for its activity in the dissociation of the posttermination ribosome. After RF3 catalyzes the release of RF1/RF2 from the ribosome, RRF enters the A-site of the ribosome, and it is then translocated by

EF-G to the P site, similarly to the translocation of the peptidyl-tRNA during translation elongation (Figure 1, step 4). This releases the deacylated tRNA from the P site which then moves to the ribosomal exit (E) site from where it rapidly dissociates from the ribosome. The release of the mRNA from the ribosome is accompanied by the release of RRF and EF-G. The final step before the ribosome can re-enter the translation cycle is dissociation of the ribosome into its two subunits: an event that requires an additional protein factor, the translation initiation factor 3 (IF3).

In eukaryotes, the posttermination step still remains to be characterized. A gene encoding a eukaryotic RRF-like protein is found in most eukaryotic genomes but encodes an RRF that only participates in translation in the mitochondria. It is conceivable that eRF3 fulfills this function.

Bypassing Stop Codons: The Causes and Consequences

The accurate recognition of a stop codon as a termination signal is important for cell viability, although cells can tolerate a certain level of stop codon translation by nonsense suppressor tRNAs. However, the efficiency with which RFs mediate translation termination at a stop codon can be affected by several factors. These include the nucleotide context of the stop codons in the mRNA (i.e., the identity of the bases directly preceding or following the termination codon), the cellular concentration of RFs, and the presence of naturally occurring, nonmutant tRNAs, which are able to recognize a stop codon as well as their usual normal sense codon(s) albeit at much lower efficiency. For example, in yeast the tRNA that decodes the CAA codon as a glutamine residue can also recognize the related stop codon UAA at low efficiency.

A number of nonstandard translational events are known to compete with the termination machinery in order to facilitate the translational bypassing of specific stop codons. Such a bypass can extend the decoding properties of a single gene and is an event that can be regulated. For example, certain animal viruses can cause ribosomes to shift into a different reading frame just before a stop codon is reached thereby generating a second extended form of the translation product – which in this case is the viral gag-pol protein. In certain plant viruses, e.g., Tobacco Mosaic Virus (TMV), host-encoded tRNAs can decode a viral stop codon as a sense codon if that stop codon is present in the appropriate nucleotide context.

A further example where stop codons are treated as sense codons is in the translation of genes encoding certain selenoproteins, i.e., contains one or more

selenocysteine residues. In these mRNAs the UGA codon is used both as a signal for termination and as a signal for selenocysteine incorporation. How the UGA codon is read is defined by both a specific translation factor (e.g., the SelB protein in bacteria) and a structural element within the mRNA molecule that is being translated (e.g., the SECIS element in eukaryotic selenoprotein-encoding mRNAs).

SEE ALSO THE FOLLOWING ARTICLES

Prions and Epigenetic Inheritance • Prions, Overview • Ribosome Structure

GLOSSARY

aminoacyl tRNA   A transfer RNA molecule that has an appropriate amino acid residue esterified to its $3'$-terminal adenosine.

anticodon   A contiguous sequence of three nucleotides in transfer RNA that are complementary to a specific codon in mRNA.

codon   A contiguous sequence of three nucleotides in mRNA that are used to specify a particular amino acid, e.g., CAG is the codon specifying the amino acid glutamine.

nonsense suppressor tRNAs   tRNAs that contain a mutation in their anticodon sequence such that they can recognize an mRNA stop codon positioned at the ribosomal A site by standard base pair interactions.

peptidyl-transferase center   The region of the ribosome that is responsible for two reactions important for protein synthesis: peptide bond formation during translation elongation and nascent peptide chain release during termination.

prion   A protein that can exist in a self-perpetuating conformationally altered isoform. Prions have been identified as infectious agents responsible for certain neurodegenerative diseases of animals and as epigenetic determinants in yeast.

reading frame   Any one of three ways in which an mRNA sequence can be translated one codon at a time. The "open" reading frame encoding a gene product is set by the initiation codon AUG.

FURTHER READING

Bertram, G., Innes, S., Minella, O., Richardson, J., and Stansfield, I. (2001). Endless possibilities: Translation termination and stop codon recognition. *Microbiology UK* **147**, 255–269.

Ehrenberg, M., and Tenson, T. (2002). A new beginning of the end of translation. *Nat. Struct. Biol.* **9**, 85–87.

Kisselev, L., Ehrenberg, M., and Frolova, L. (2003). Termination of translation: Interplay of mRNA, rRNAs, and release factors? *Eur. Molecul. Biol. Orgn. J.* **22**, 175–182.

Klaholz, B. P., Pape, T., Zavialov, A. V., Myasnikov, A. G., Orlova, E. V., Vestergaard, B., Ehrenberg, M., and van Heel, M. (2003). Structure of the *Escherichia coli* ribosomal termination complex with release factor 2. *Nature* **421**, 90–94.

Nakamura, Y., and Ito, K. (2003). Making sense of mimic in translation termination. *Trends Biochem. Sci.* **28**, 99–105.

Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B. F., and Nyborg, J. (1995). Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. *Science* **270**, 1464–1472.

Rawat, U. B., Zavialov, A. V., Sengupta, J., Valle, M., Grassucci, R. A., Linde, J., Vestergaard, B., Ehrenberg, M., and Frank, J. (2003). A cryo-electron microscopic study of ribosome-bound termination factor RF2. *Nature* **421**, 87–90.

BIOGRAPHY

Mick Tuite is a Professor of Molecular Biology at the University of Kent and the former Head of the Department of Biosciences at this institution. He has had a long-standing interest in the mechanism and regulation of translation termination using the yeast *S. cerevisiae* as a model system. This has recently led to a study of the prion-like properties of one of the release factors eRF3 (Sup35p). He has a D.Phil. from the University of Oxford and undertook postdoctoral research at both Oxford and the University of California, Irvine before joining the faculty at Kent. He has published over 150 research papers and reviews.

Nadja Koloteva-Levin is a Senior Postdoctoral Researcher in Mick Tuite's laboratory working on the prion behavior of eRF3 (Sup35p). She received her Ph.D. from University of Manchester Institute of Science and Technology (UMIST) in the United Kingdom and followed this with 4 years of postdoctoral research on posttranscriptional gene regulation at Tel Aviv University in Israel.

Translesion DNA Polymerases, Eukaryotic

Alexandra Vaisman and Roger Woodgate
National Institutes of Health, Bethesda, Maryland, USA

Translesion DNA polymerases are enzymes capable of copying templates containing DNA-distorting lesions. Unlike the highly processive and accurate replicative polymerases, translesion polymerases are relatively distributive and lack intrinsic $3'-5'$-exonuclease activity and are, therefore, unable to proofread any errors made during replication. As a consequence, they have low base substitution and frameshift fidelity on undamaged DNA templates. Translesion DNA polymerases are also characterized by an ability to extend mismatched and misaligned primer termini.

Lesion bypass is believed to occur in two mechanistically discrete steps, both of which can be either relatively accurate or highly error-prone. The first step is nucleotide incorporation opposite the damaged base(s) itself, and the second step is extension of the nascent chain beyond the damage. For some lesions, translesion replication is facilitated by a single enzyme, but for certain lesions two polymerases are required; one enzyme for incorporation opposite the lesion and another for extension. The polymerase that gains access to the lesion terminus largely determines the fidelity of translesion synthesis, i.e., whether it is accurate or mutagenic. After lesion bypass is completed, the inherently distributive translesion polymerases dissociate from the DNA template so as to allow replicative polymerases to resume chromosomal duplication.

Y-Family DNA Polymerases: polη, polι, polκ and Rev1

Most of the enzymes implicated in replication of imperfect DNA belong to the Y-family of DNA polymerases. Phylogenetic analysis of this family suggests that it can be subdivided into several discrete branches consisting of UmuC-, DinB-, Rev1-, and Rad30-like proteins. The DinB subfamily is the most diverse and is found in prokaryotes, archaea, and eukaryotes. Members of the UmuC subfamily are found exclusively in bacteria, and proteins from the Rad30 and Rev1 branches are found only in eukaryotes. Multiple Y-family orthologues are often present in one organism. Humans, for example, have four Y-family DNA polymerases, two *RAD30* paralogues, polη and polι, a DinB homologue polκ, and Rev1.

BIOCHEMICAL PROPERTIES

Enzymatic characterization of the Y-family DNA polymerases *in vitro* reveals that despite numerous similarities, each member of the family exhibits distinctive properties during DNA synthesis. For example, polη is truly unique among eukaryotic DNA polymerases in its efficiency and accuracy of replication past UV-induced cyclobutane pyrimidine dimers (CPDs). In contrast, Polι, a homologue of Rad30, differs from other Y-family DNA polymerases in that it has a $5'$-deoxyribose phosphate lyase activity, similar to X-family pols β and λ. A most unusual biochemical feature of polι is, however, its sequence context fidelity when replicating undamaged DNA templates. Remarkably, Polι prefers making dGMP·T and dGMP·U wobble mispairs – three- to tenfold greater than the correct dAMP·T and dAMP·U base pairs. Yet at the same time, polι discriminates against certain other mismatches quite well. Thus, errors made by polι at template A occur up to 100-fold less frequently than that made by human pol η or κ. Polκ on the other hand, readily catalyzes extension of misaligned primer termini resulting in -1 frameshift mutations and is considerably more proficient at extending terminal mismatches compared to polη and polι. Finally, unlike other Y-family DNA polymerases that utilize all four deoxynucleotide triphosphates (dNTPs) in a template-dependent manner, *Saccharomyces cerevisiae* and human Rev1 proteins appear to be specialized deoxycytidyl transferases, since they utilize the three other dNTPs only very sparingly.

STRUCTURAL FEATURES

Sequence alignment of the Y-family DNA polymerases shows a high degree of conservation in the N-terminal region containing the catalytic domain, while the

C-terminal part is unique for each family member. Even though the amino acid sequences of these polymerases show no similarity to classical DNA polymerases, the general structure of polymerases from all families appears to be remarkably conserved. Crystallographic data have shown that like polymerases from the A-, B-, and X-families, the Y-family polymerases adopt a right-hand structure with characteristic finger, thumb, and palm domains. The Y-family polymerases differ, however, from classical polymerases in having a fourth small domain that has been termed the "little finger." DNA is held in a cleft between the thumb and the little finger domains (Figure 1). The active site of Y-family polymerases is more solvent accessible and, due to shorter fingers and thumb domains, forms a more open structure than that of replicative polymerases. Therefore, if the geometric constraints on the active site of Y-family polymerases are less stringent, DNA lesions and mispairs can be more readily accommodated within their active site. Such observations thereby help explain their proficiency at lesion bypass and a concomitant decrease in fidelity on undamaged DNA.
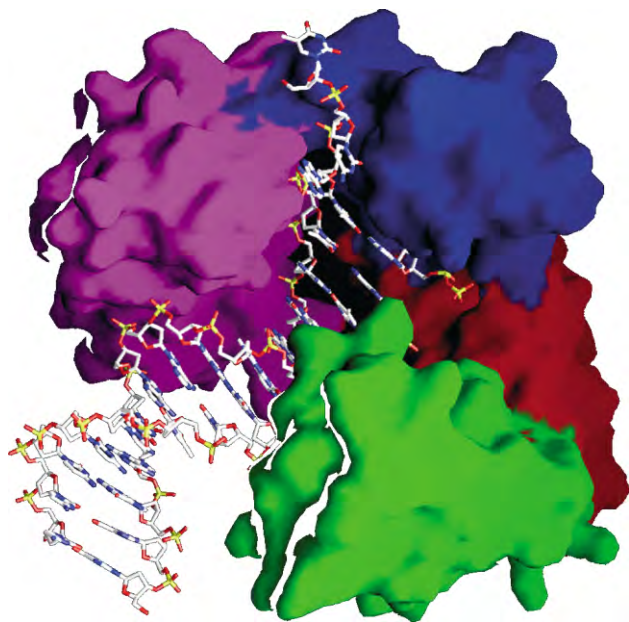
FIGURE 1 Crystal structure of Dpo4, an archaeal Y-family polymerase, in a ternary complex with a DNA template and an incoming nucleoside triphosphate. The surface of the enzyme is color-coded to indicate the four subdomains found in all Y-family polymerases. These consist of the palm (red), thumb (green), finger (blue), and little finger (purple) domains. The DNA molecule is drawn as sticks. Unlike replicative DNA polymerases, the catalytic cleft in Dpo4 is cavernous and solvent exposed, which explains in part why Y-family polymerases exhibit low-fidelity synthesis on undamaged DNA and can accommodate bulky DNA adducts/mismatched bases in their active sites.

B-Family DNA Polymerase: polζ

The ability to accommodate distorted DNA is not just limited to Y-family polymerases. Polζ, a eukaryotic B-family polymerase, is another important enzyme involved in the replication of damaged DNA. Catalytically active polζ comprises a heterodimer made from the REV3 and REV7 proteins. The fidelity of polζ on undamaged DNA is significantly higher than that of Y-family polymerases (Table I). However, polζ is extremely efficient at the extension of aberrant DNA primer ends, including lesion-containing sites and misaligned primer-template termini.

X-Family DNA Polymerases: polβ, polλ, and polμ

Although much attention in the last few years has focused on the Y-family polymerases, earlier studies raised the possibility that DNA polβ may play a role in translesion synthesis, even though its major function is in base excision repair. Polβ is a distributive enzyme in its mode of action, however, it is more processive than most Y-family DNA polymerases and the inherent ability of polβ to select correct rather than incorrect dNTPs for incorporation is considerably higher than that of the Y-family polymerases (Table I).

Lesion bypass with varying efficiencies can also be achieved *in vitro* by other X-family polymerases, such as the recently discovered pols λ and μ. Polμ is closely related to terminal deoxynucleotidyl transferase (TdT), but in contrast to TdT, it is a partially processive template-directed DNA polymerase. *In vitro*, polμ is highly error-prone and makes -1 frameshift errors in repetitive sequences. Polλ is a close homologue of polβ and like polβ is more accurate than polμ and possesses an intrinsic deoxyribophosphate lyase activity (Table I).

Lesion Bypass Specificity

DNA molecules are intrinsically unstable and are constantly subjected to the damaging effects of endogenous or exogenous agents. DNA synthesis past a wide variety of DNA lesions has been studied using both biological and biochemical approaches. Among the most common lesions that have been investigated are UV-induced *cis-syn* CPDs and the 6-4 pyrimidine-pyrimidone photoproducts (6-4PPs); abasic sites occurring spontaneously or from the action of DNA glycosylases; cisplatin adducts (CP) formed at adjacent guanines by anticancer drugs; and benzo[*a*]pyrene diol epoxide (BaP DE) adducts, a naturally occurring metabolite of carcinogens commonly found in cigarette

TABLE I

Eukaryotic DNA Polymerases Known to Participate in Translesion Replication

DNA polymerase	Gene or synonym	Family	Error rate on undamaged DNA	Additional proposed functions
β, beta	*POLB*	X	10^{-4}–10^{-6}	Base excision repair, double-strand break repair, synapsis and recombination, meiosis, neurogenesis
ζ, zeta	*POLZ (REV3L/REV7L)*	B	10^{-4}–10^{-5}	Somatic hypermutation, mismatch extension, developmental cell proliferation
η, eta	*POLH (RAD30A, XPV)*	Y	10^{-2}–10^{-3}	Somatic hypermutation
ι, iota	*POLI (RAD30B)*	Y	10^{1}–10^{-4}	Base excision repair, somatic hypermutation
κ, kappa	*POLK (DINB1)*	Y	10^{-3}–10^{-4}	Spontaneous mutagenesis
λ, lamda	*POLL (POL4)*	X	10^{-3}	Base excision repair, repair during meiosis
μ, mu	*POLM*	X	?	Nonhomologous end-joining pathway, lymphoid formation
Rev1	*REV1L*	Y		

smoke. The accuracy and efficiency of translesion replication varies not only on the chemical nature of the lesion and specific distortion that it imposes upon DNA, but also on the local sequence context within which the lesion is located and on the specific polymerase or combination of polymerases involved in a given translesion pathway. For example, in both *S. cerevisiae* and humans, polη is the most proficient and accurate enzyme at replicating past thymine–thymine CPDs. Pols ι, polμ and to a lesser extent polζ and polβ are also able to bypass the lesion but with a much lower efficiency and fidelity. By comparison, both polι and polη are proficient at incorporating a nucleotide opposite the $3'$ base of a thymine-thymine 6-4PP, and polζ can readily extend the resulting primer termini. Similarly, translesion replication past an abasic site might involve pols η, ι, κ, β, λ, or Rev1 to incorporate a base opposite the lesion and polζ for its subsequent extension. Polη and polβ are both very efficient and accurate when replicating past a CP lesion, whereas bypass by polμ is highly error-prone since it generally results in 2, 3, and 4 nucleotide deletions, polκ, polι, and polλ are completely blocked by CP adducts in template DNA.

The ability of the polymerase to facilitate lesion bypass also depends on the nucleotide at which a specific adduct is formed. For example, in human cells, efficient and accurate bypass of BaP DE adducts formed at deoxyadenosine (dA) can be achieved through the sequential action of polι and polκ, but not by polη or polβ. In contrast, the ability of translesion polymerases to bypass the same BaP DE adduct formed at deoxyguanosine (dG) base is quite different. Whereas human polη and polι bypass BaP DE dG adducts inefficiently and inaccurately, polκ bypasses the lesion with high efficiency and predominantly incorporates the correct base, dCMP, opposite the adducted dG.

Protein Partners

Despite extensive biochemical characterization of the polymerases involved in translesion synthesis, it is still unclear how the process physically occurs in living cells. It has been shown that at least three of the translesion polymerases, polη, polι, and Rev1 are associated with the replication machinery in an undamaged cell and that upon DNA damage the polymerase accumulates at sites of stalled replication forks.

The catalytic activity of pols η, ι, and κ on undamaged and lesion-containing DNA is substantially increased through protein–protein interactions with accessory proteins normally utilized by replicative polymerases. These include proliferating cell nuclear antigen (PCNA), which serves as a sliding clamp that helps recruit polymerases to their site of action, and a clamp loader, replication factor C (RFC). Given the sixfold symmetry of the PCNA complex, it is possible that multiple DNA polymerases connect simultaneously to a replicative single clamp. However, how a cell chooses between the various translesion polymerases it has at its disposal and how the switch from the replicative polymerase to a translesion polymerase and then back to the replicase occurs *in vivo* still remains to be determined.

Cellular Role(s)

Polη is clearly responsible for the accurate and robust replication of CPD-containing DNA *in vivo*. Although the lack of polη in human cells does not confer a substantial hypersensitivity to killing by UV light, humans with a defective polη do exhibit the *Xeroderma Pigmentosum Variant* (XP-V) syndrome and are predisposed to the manifestation of skin cancers.

To date, XP-V is the only known example of a human disorder caused by a defect in translesion replication. Apart from a clear role for polη in the bypass of CPDs, there is indirect evidence suggesting the participation of other enzymes in translesion replication *in vivo*. For example, cells with a reduced level of the deoxycytidyl transferase Rev1 show lower levels of UV-induced mutagenesis. *In vitro* assays suggest that polκ may play an important role in accurate bypass of certain lesions generated by polycyclic aromatic hydrocarbons and recent studies with transgenic mice devoid of polκ support such hypotheses.

Just as translesion replication often requires multiple DNA polymerases, so a particular translesion polymerase might have evolved to participate in one or more functional tasks in a cell. Indeed, various lines of evidence support a role for translesion polymerases η and ι in the somatic hypermutation of rearranged immunoglobulin genes. While *S. cerevisiae* strains with deletions in their *REV3* or *REV7* genes (encoding polζ) are viable, disruption of murine *REV3* unexpectedly leads to embryonic lethality in mice, suggesting that, in addition to lesion bypass, mammalian polζ may play additional role(s) in cellular development. Genetic and biochemical studies have shown that *S. cerevisiae* Rev1, in addition to participating in the bypass of UV-induced lesions and abasic sites, also has a second function that is independent of its dCMP transferase activity. Polβ whose primary function is in base excision repair, has also been implicated in double-strand-break repair, meiotic events associated with synapsis and recombination and in neurogenesis. On the other hand, owing to an associated 5′-deoxyribose phosphate lyase activity, it has been hypothesized that polι and polλ might possibly substitute for polβ in certain forms of specialized base excision repair. Finally, polμ's main cellular role may actually be to participate in a non-homologous end-joining pathway that repairs DNA double-strand breaks and its ability to traverse certain DNA lesions may, in fact, be rarely utilized *in vivo*.

SEE ALSO THE FOLLOWING ARTICLES

DNA polymerase α, Eukaryotic • DNA polymerase β, Eukaryotic • DNA polymerase δ, Eukaryotic • DNA polymerase ε, Eukaryotic • DNA Polymerases: Kinetics and Mechanisms

GLOSSARY

DNA polymerase families DNA polymerases have historically been classified into "families" based upon the phylogenetic relations of their primary amino acid sequence. A-family polymerases are related to *Escherichia coli* pol I; B-family polymerases to *E. coli* pol II; C-family polymerases to *E. coli* pol III; X-family polymerases to mammalian polβ; and Y-family polymerases to UmuC, DinB, Rad30, and Rev1 proteins.

fidelity The accuracy with which a polymerase replicates a DNA template.

polymerase processivity The number of nucleotides incorporated into a nascent DNA strand per polymerase–template binding event.

proofreading The removal of misincorporated nucleotides at a growing 3′ end by a 3′ exonuclease often associated with the polymerase.

replicative polymerases Enzymes involved in the accurate copying of genetic material. The active sites of these enzymes are usually much smaller and constrained than the translesion polymerases, so as to ensure that only the complementary "Watson and Crick" bases are incorporated into nascent DNA.

translesion replication Also referred to as lesion bypass or translesion synthesis (TLS), it is an inherently error-prone process that permits cells to tolerate the presence of persistent DNA damage and involves the direct replication through and beyond the DNA damaged site by DNA polymerases.

FURTHER READING

Friedberg, E. C., Wagner, R., and Radman, M. (2002). Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science* **296**, 1627–1630.

Goodman, M. F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* **71**, 17–50.

Ohmori, H., Friedberg, E. C., Fuchs, R. P. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001). The Y-family of DNA polymerases. *Molecul. Cell* **8**, 7–8.

BIOGRAPHY

Alexandra Vaisman is a Senior Research Fellow in the Laboratory of Genetic Integrity at the National Institute of Child Health and Human Development. Her principal research interests are in the field of mammalian DNA repair and replication. She holds a Ph.D. from the Moscow Institute of Fine Chemical Technology and received her postdoctoral training at the Russian Academy of Sciences and the University of North Carolina at Chapel Hill.

Roger Woodgate has devoted his scientific career to understanding the molecular mechanisms of damage-induced mutagenesis in prokaryotes and eukaryotes. He received his Ph.D. from the University of Sussex in Brighton, United Kingdom and for the past 15 years has worked at the National Institutes of Health in Bethesda, Maryland. At present, Dr. Woodgate is Chief of the Laboratory of Genomic Integrity in the National Institute of Child Health and Human Development.

Trehalose Metabolism

Alan D. Elbein
University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

Trehalose is a nonreducing disaccharide of glucose in which the two glucoses are linked in an α,α-1,1-linkage. This rather simple sugar is widespread throughout the biological kingdom, being found in bacteria, yeast, fungi, protozoa, nematodes, insects, plants, and so on. In these various organisms, this sugar can serve a number of key functions: as a storehouse of energy and carbon, as a protectant and stabilizer of cell membranes and proteins during stress, as a signaling molecule, and as a structural component of various glycolipids in bacterial cell walls. There are at least three different pathways for the synthesis of trehalose in various organisms, and several or all of these pathways may be present in the same organism. The fact that an organism may have multiple routes for the synthesis of trehalose strongly suggests that this sugar must be essential for the survival of the organism.

Pathways of Biosynthesis of Trehalose

The best-known and most-widely distributed pathway for synthesis of trehalose (Figure 1) involves two enzymes known as trehalose-phosphate synthase (TPS) and trehalose-phosphate phosphatase (TPP). The reactions catalyzed by these two enzymes are shown in Figure 2. TPS catalyzes the transfer of glucose from the sugar nucleotide, uridine diphosphate glucose (UDP-glucose), to glucose-6-phosphate to produce trehalose-6-phosphate and UDP. TPP then cleaves the phosphate from trehalose-6-P to give free trehalose and inorganic phosphate. These two enzymes have been isolated and characterized from the cytoplasm of many cells including bacteria, yeast, fungi, insects, and plants. The genes for these enzymes have been cloned from several different organisms and expressed in *Escherichia coli*. In several cases, the purified recombinant enzymes have been well characterized.

A more recently described pathway of trehalose formation, so far demonstrated in only a few bacteria, including *Mycobacterium tuberculosis*, involves the intramolecular rearrangement of maltose to convert the α1,4-glycosidic linkage of this disaccharide to the α,α1,1-glycosidic linkage of trehalose. This reaction is catalyzed by the enzyme trehalose synthase (TS) as shown in Figure 2. The gene for TS from *Pimelobacter* species was cloned to produce a 573 amino acid protein that had considerable homology to the glycosidase, maltase. The mechanism of action of TS is not known, but this enzyme can convert trehalose to maltose, or maltose to trehalose. Its physiological function in these cells is also not known.

A third pathway also only identified in a few bacteria involves several different enzymes, the first of which converts the glucose linked in an α1,4-bond at the reducing end of a malto-oligosaccharide to the α,α1,1-bond of trehalose. This enzyme is called malto-oligosyltrehalose synthase (TreY). The second enzyme then releases the trehalose to leave a malto-oligosaccharide i.e., two glucoses shorter (at the reducing end) than the starting oligosaccharide. This second enzyme is called malto-oligosyltrehalose trehalohydrolase (TreZ) (see Figure 2). These latter two pathways give rise to free trehalose without the involvement of the phosphorylated sugar, whereas the first pathway produces trehalose phosphate as the initial product. In some bacteria, at least two and possibly all three of these pathways may be present at the same time. The presence of these multiple pathways would suggest that trehalose is an essential component for these cells, and therefore alternate pathways have evolved to ensure that levels of trehalose are maintained.

Occurrence of Trehalose in the Biological World

The first report on the presence of trehalose in living cells was in 1832 when trehalose was tentatively identified in ergot of rye by Wiggers. Since that report, this sugar has been isolated from many different organisms, including yeast and fungi where it occurs in spores, fruiting bodies, and vegatative cells. In some of these organisms, such as the fungus, *Neurospora tetrasperma*, or the yeast, *Saccharomyces cerevesiae*, the amount of trehalose may be as high as 10% of the dry weight of the cells.
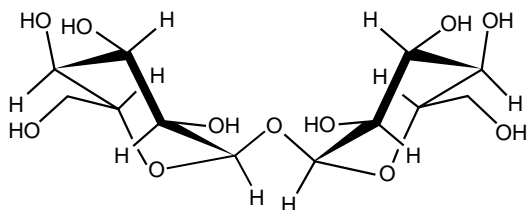
FIGURE 1 Structure of the naturally occurring isomer of trehalose, i.e., $\alpha,\alpha1,1$-D-glucopyranosyl-D-glucopyranoside.

Trehalose has also been isolated from lichens and algae, and is present in many higher plants, such as the resurrection plant (*Selaginella lepidophylla*), and *Arabidopsis thaliana*. It is also a component of the wound exudates of some plants, such as *Fraxsinus aras*. Many different species of bacteria, including *Streptomyces*, *Mycobacterium*, *Corynebacteria*, *Rhizobium*, *Arthrobacter*, and *Escherichia* have been shown to contain trehalose, and in most cases to also have the ability to synthesize this disaccharide. These many studies suggest that trehalose is probably present in many, if not most bacteria.

In the animal kingdom, trehalose was first reported in insects where it is present in hemolymph and also in larvae or pupae. In the adult insect, the levels of trehalose fall rapidly during certain energy-requiring activities such as flight, suggesting a role for this sugar as a source of glucose for producing ATP. In addition to insects, trehalose has also been identified in the eggs of the roundworm, *Ascaris lumbricoides*, where it may be present at levels as high as 8% of the dry weight. It is also found in a number of other invertebrates. On the other hand, trehalose has not been found in any mammals, although the enzyme trehalase is present in human intestine (intestinal villae membranes) and in human kidney (kidney brush border membranes). The intestinal enzyme is probably important in metabolizing ingested trehalose, since this sugar is present in various foods such as mushrooms which contain significant amounts of trehalose. On the other hand, the role of trehalase in human kidney remains a mystery.

Many lower organisms do contain the enzyme trehalase, and this activity may be important in maintaining physiological concentrations of trehalose in the cytoplasm. That is, as discussed below, trehalose plays a role as a protectant or stabilizer in some cells, and the levels of trehalose in these cells is significantly increased during times of stress. A high concentration of trehalose in the cytoplasm could upset the osmolar balance in the cell and cause serious problems. Thus, trehalase may be necessary to control such imbalances. This enzyme may also have an important function in the utilization of trehalose as an energy and carbon source for insect flight, for spore germination, and for other glucose-requiring processes.

In yeast, there are several different trehalases which apparently have different functions. The cytoplasmic trehalase, also referred to as the neutral trehalase because it has a pH optimum of 7.0, is a regulatory enzyme. This enzyme can exist in an inactive or zymogenic form, and this zymogen is converted by a cyclic-AMP-dependent phosphorylation to the active trehalase. Another trehalase in these cells is the acidic enzyme, which is found in the vacuoles and has a pH optimum of ~ 4.5. The acid enzyme may be involved in the utilization of trehalose as a carbon source by yeast since deletion of the gene for this protein leads to an inability of yeast to grow on trehalose. On the other hand, the trehalase that is regulated by phosphorylation might be involved in regulating the levels of trehalose in the cytoplasm, since trehalose levels can be significantly increased by various types of stress. However, the exact role of the various trehalase activities still remains to be determined.

Physiological Functions of Trehalose

Trehalose may have a number of different functions in nature, and its specific role may vary depending upon the system or organism being considered. The established functions of trehalose are outlined below.

AS A SOURCE OF ENERGY AND/OR CARBON

Trehalose levels may vary greatly in certain cells depending on the stage of growth, the nutritional state of the organism, and the environmental conditions prevailing at the time of measurement. Trehalose is the major sugar in the hemolymph and thorax muscle of insects, and it is converted to glucose and consumed during flight. It is also an important source of energy and carbon in fungal spores and is utilized during the germination process.

AS A STABILIZER AND PROTECTANT OF PROTEINS AND MEMBRANES AGAINST VARIOUS STRESSES

Various organisms, including plants, yeast, fungal spores, nematodes, brine shrimp, etc., can withstand dehydration and remain in this state (anhydrobiosis) for extended periods of time until water becomes available. These organisms generally contain high concentrations of trehalose. For example, when nematodes are slowly dehydrated, they convert as much as 20% of their dry weight into trehalose. Similar results have been
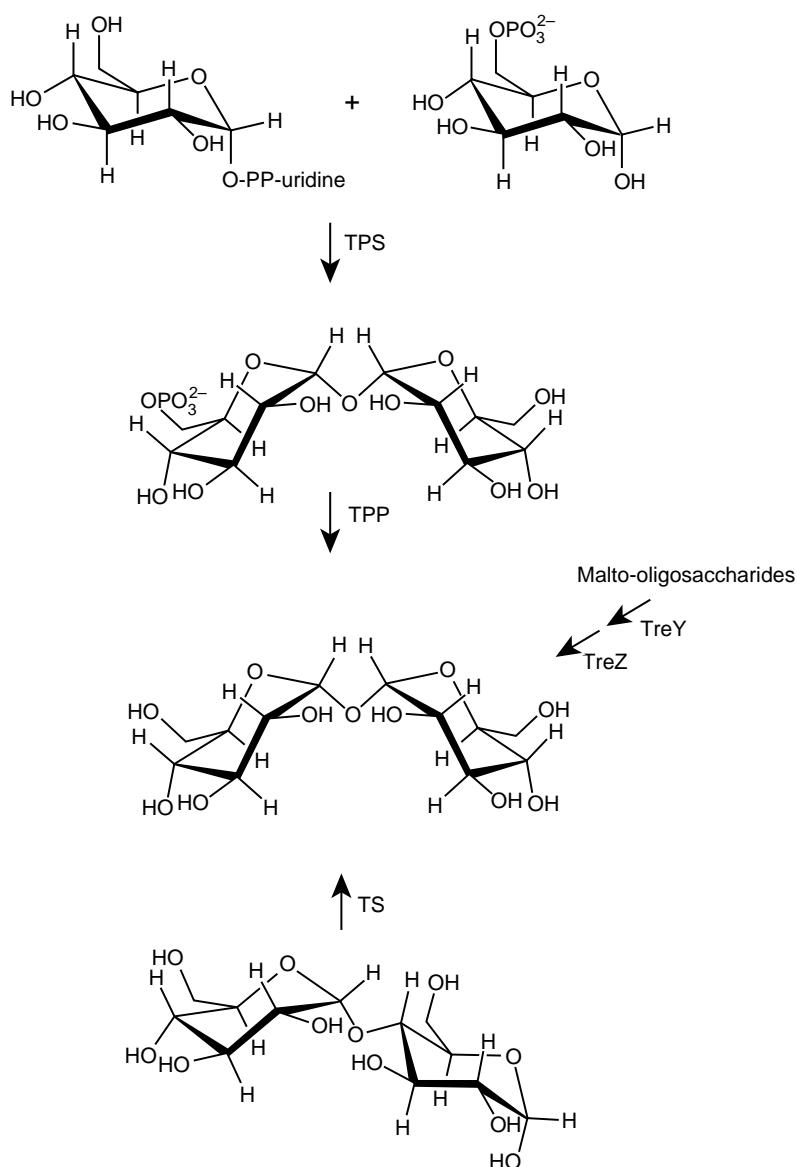
FIGURE 2 Pathways of biosynthesis of trehalose. The best-known pathway involves the transfer of glucose from UDP-glucose to glucose-6-phosphate to form trehalose-6-phosphate and UDP, catalyzed by trehalose-phosphate synthase (TPS). A second enzyme in this pathway, trehalose-phosphate phosphatase (TPP), removes the phosphate to give free trehalose. An alternate pathway involves the enzyme, trehalose synthase (TS) which catalyzes the interconversion of maltose and trehalose. Another alternate pathway involves two enzymes, TreY and TreZ, which catalyze the formation of trehalose at the reducing end of a malto-oligosaccharide (TreY) and then hydrolysis of this trehalose (TreZ) to give free trehalose.

demonstrated with other organisms. The use of trehalose to enable cells to survive dehydration may be an ancient adaptation since even *Archaebacteria* have been found to accumulate trehalose in response to stress.

Two primary changes are thought to cause destabilization of lipid bilayers during dehydration. These changes are fusion and lipid-phase transitions. Trehalose inhibits fusion and depresses the phase transition temperature of the dry lipids which maintains them in the liquid crystalline phase in the absence of water. This stabilizing effect of trehalose is a property of its structure and its stereochemistry. X-ray diffraction studies show

that trehalose fits well between the polar head groups of the lipids with multiple sites of interaction. Its stereochemistry allows it to have the most favorable fit with the polar head groups of the phospholipids of cellular membranes.

This ability to withstand stress by increasing the levels of trehalose has been used by researchers to bioengineer more stable plant and animal cells. For example, introducing the genes that code for the biosynthetic trehalose enzymes (i.e., TPS and TPP) into human fibroblasts allowed these cells to be maintained in the dry state for up to 5 days, whereas control cells were

rapidly killed by this treatment. Similar technology has also been used to produce more stable transgenic plants.

As a Protectant against Heat

In yeast, stimuli that trigger the heat-shock response also cause an accumulation of trehalose. Two of the subunits of the trehalose-6-P synthase complex are actively synthesized when yeast cells are subjected to heat shock. Yeast mutants that are defective in either of the genes coding for TPS or TPP are not able to accumulate trehalose, and as a result, they are much more sensitive to heat shock than wild-type cells. Physiological concentrations of trehalose, up to 0.5 M, were found to protect enzymes of yeast and other organisms from heat inactivation, *in vitro*, and trehalose was a considerably better stabilizer than a number of other polyols, sugars, or amino acids. These data strongly implicate trehalose as playing a key role in thermotolerance, as well as in other stress conditions.

As a Free Radical Scavenger to Prevent Oxidative Stress

Another role for trehalose is in protecting cells against oxygen radicals. Exposure of yeast to a mild heat shock, or to a proteosome inhibitor, induced the accumulation of trehalose but also markedly increased the viability of these cells during exposure to a free radical-generating system (such as H_2O_2/iron). However, when the cells were returned to the normal growth temperature, both the trehalose content and the resistance to oxygen stress decreased rapidly and returned to the normal level. Mutants that were defective in the enzymes of trehalose metabolism were much more susceptible to oxygen stress than the wild-type organism, but adding exogenous trehalose to the medium enhanced the resistance of these mutant cells to the oxygen stress. The major effect of oxygen radicals on these cells was to damage amino acids in cellular proteins, and trehalose was able to prevent this damage, suggesting that this disaccharide functions as a free-radical scavanger.

As a Structural Component of the Bacterial Cell Wall

In mycobacteria and corynebacteria (and perhaps other organisms), trehalose is the basic component of a number of cell wall glycolipids. The best known and most completely studied of these trehalose lipids is cord factor, a cell wall component of the tubercle bacillus, that contains the unusual fatty acid called mycolic acid, esterified to the hydroxymethyl group of each glucose to give trehalose-dimycolate. This lipid is considered to be one of the major toxic components of the *Mycobacterium tuberculosis* cell wall, and is also largely responsible for the low permeability of this cell wall, which provides the organism with considerable resistance to many drugs. However, the function of this lipid, besides its obvious structural role, is still uncertain.

There are other antigenic glycolipids in the mycobacterial cell wall that also have trehalose as the basic structure. For example, there are a variety of acylated-trehalose compounds that contain any of three major types of fatty acids attached to the 2 and 3 hydroxyl groups of the two glucose moieties. These fatty acids are unusual and may be either C_{16-19} saturated fatty acids, C_{21-25} α-methyl branched fatty acids, or C_{24-28} α-methyl branched, β-hydroxy fatty acids. *M. tuberculosis* and other mycobacteria also have trehalose lipids that contain sulfate, such as 2,3,6,6'-tetra-acyl-2-sulfate-trehalose (sulfatide 1), or other types of fatty acids such as phthienoic acids. This great variation in the types of fatty acids found in these organisms and as cell wall components suggests specific functions, but thus far these have not been demonstrated.

Finally, some mycobacteria, such as *M. kansasii*, are characterized by the presence of seven species-specific neutral lipooligosaccharide antigens. These oligosccharide structures all have a common tetraglucose core which is distinguished by an α,α-trehalose substituent to which are attached various sugars, such as xylose, 3-0-methylrhamnose, fucose and a novel N-acylsugar. The exact structures of these complex polymers have not been established, but specific lipooligosaccharides are present in a number of "atypical" mycobacteria.

See Also the Following Articles

Free Radicals, Sources and Targets of: Mitochondria • Glucose/Sugar Transport in Bacteria • Oligosaccharide Chains: Free, *N*-Linked, *O*-Linked • Sugar Nucleotide Transporters

Glossary

enzymes Proteins that are necessary to catalyze specific chemical reactions at low temperature within the cell.

free-radicals Compounds that have an unpaired electron and are very reactive; free radicals can oxidize proteins or cell membranes and inactivate them.

free-radical scavenger Any substance that can react with and neutralize free radicals.

maltose A sugar (disaccharide) that is composed of two glucose molecules linked together in an α1,4-glycosidic bond; maltose differs from trehalose in that it is a reducing sugar and has a different glycosidic linkage.

stabilizer A chemical compound that prevents (i.e., protects) proteins or membranes from being inactivated or denatured.

sugar nucleotide The activated (or "high-energy") form of a sugar that is necessary for enzymatic formation of a glycosidic bond.

FURTHER READING

Banaroudj, N., Lee, D. H., and Goldberg, A. L. (2001). Trehalose accumulation during cellular stress protects cells and cellular proteins from damage by oxygen radicals. *J. Biol. Chem.* **276**, 24261–24267.

Birch, G. G. (1963). Trehaloses. *Adv. Carbohyd. Chem.* **18**, 201–224.

Brennan, P. J., and Nikaido, H. (1995). The envelope of mycobacteria. *Annu. Rev. Biochem.* **64**, 29–63.

Crowe, J., Crowe, L., and Chapman, D. (1984). Preservation of membranes in anhydrobiotic organisms. The role of trehalose. *Science* **223**, 209–217.

Elbein, A. D. (1967). Carbohydrate metabolism in *Streptomycetes*. Isolation and enzymatic synthesis of trehalose. *J. Bacteriol.* **94**, 1520–1524.

Elbein, A. D., Pan, Y. T., Pastuszak, I., and Carroll, J. D. (2003). New insights on trehalose: A multifunctional molecule. *Glycobiology* **13**, 17R–27R.

Lederer, E. (1976). Cord factor and related trehalose esters. *Chem. Phys. Lipids* **16**, 91–106.

Leopold, A. C. (1986). *Membranes, Metabolism and Dry Organisms*, pp. 377. Cornell University Press, Ithaca, NY.

Pan, Y. T., Carroll, J. D., and Elbein, A. D. (2002). Trehalose-phosphate synthase of *Mycobacterium tuberculosis*. *Eur. J. Biochem.* **269**, 6091–6100.

Thevelein, J. M. (1984). Regulation of trehalose metabolism in fungi. *Microbiol. Rev.* **48**, 42–59.

BIOGRAPHY

Alan Elbein is Professor and Chairman of the Department of Biochemistry and Molecular Biology at the University of Arkansas for Medical Sciences in Little Rock. He received his Ph.D. in microbial biochemistry from Purdue University and did postdoctoral research at the University of Michigan and the University of California at Berkeley. He joined the faculty as Assistant Professor of biology at Rice University and then moved to the Department of Biochemistry at the University of Texas Health Science Center at San Antonio as Professor. His research focuses on the role of complex carbohydrates in glycoprotein function, and on novel carbohydrate target sites in chemotherapy of microbiol diseases.

Tricarboxylic Acid Cycle

Richard L. Veech
National Institutes of Health, Rockville, Maryland, USA

The tricarboxylic acid cycle is the central pathway in intermediary metabolism, which accounts for the final combustion of foodstuffs into CO_2 and water while conserving the reducing equivalents produced for transmission up the mitochondrial electron transport system, to form the mitochondrial proton gradient responsible for ATP generation. The cycle occurs in nearly all life forms, even the archebacteria, suggesting that during the early evolution of life its function was anaplerotic and its direction reversed from that which exists in an oxygen-containing atmosphere.

The Discoverer of the Cycle

The tricarboxylic acid (TCA) cycle was first correctly formulated by Han Krebs, who was born on August 25, 1900 in Hildesheim. He was the son of a surgeon, Georg, and mother Alma, who was to die of depression when Krebs was 19. After completing gymnasium in Hildesheim in September 1918, Krebs joined the German army signal corps until the armistice of November 11, 1918. After completing his training in internal medicine at Gottingen, Freiburg, Munich, and Berlin in December 1925, Krebs looked for ways to obtain biochemical training. The time spent in the clinics had convinced him that it was necessary to pursue medical research. Through the influence of Bruno Mendel, who served with Krebs in the Third Medical Clinic in Berlin, and who was an acquaintance of Albert Einstein, Professor of Physics at the University of Berlin, who was in turn a frequent dinner guest at the home of Otto Warburg's father, Mendel heard that Warburg was looking for an assistant and recommended Krebs. This circle of influence allowed Krebs to obtain his first paid position as an assistant to Otto Warburg at the Kaiser Wilhelm Institut für Biologie at Berlin-Dahlem (see Figure 1). This stint was to change Krebs from a physician to a biochemist. Of Warburg, Krebs wrote (1981):

> Otto Warburg was the most remarkable person I have ever been closely associated with. Remarkable as a scientific genius of the highest caliber, as a highly independent, penetrating thinker, as an eccentric who shaped his life with determination and without fear, according to his own ideas and ideals

The Kaiser Wilhelm Institute was the leading scientific research facility of its time. In addition to Warburg, its faculty included Emil Fischer, Richard Willstater, Max von Laue, Fritz Haber, Otto Hahn, Lisa Meitner, Michael Polanyi. Otto Meyerhof, a former student of Warburg, was working in the same building and elucidated the glycolytic pathway responsible for the breakdown of glucose to lactate. Students of biochemistry who passed through Kaiser Wilhelm included, among others, Fritz Lipmann and Severo Ochoa, both of whom were to play a critical role in showing the role of coenzyme A in the TCA cycle. In addition to the discipline and rigor required for scientific research, Krebs learned from Warburg the techniques he had developed: use of tissue slices, spectrophotometry and manometry. These techniques, most notably manometry, were fundamentally important in leading Krebs to formulate the concept of a metabolic cycle, first the urea cycle and later the TCA cycle. It is important to recognize in understanding Krebs' work that the Warburg dictum "methods are everything" was dominant. There was no theoretical master plan. Only during his later work on the thermodynamics linking of cellular redox and phosphorylation states would Krebs remark in relation to that work, "This is the first time I ever thought you could predict in biology." In his early work, Krebs was a thorough empiricist. He could measure the rates of reaction. He added various potential participants in a metabolic pathway, and if they increased the overall rate, then he concluded they were part of the pathway. Krebs' training is of central importance to understand the cycle. As his biographer Holmes has pointed out, "Hans Krebs took 31 years to become a well-trained independent scientific investigator, and only 9 months to make one of the most significant discoveries of his generation in his chosen field."

History of the Discovery of the TCA Cycle

After leaving Warburg's lab in 1931, Krebs had obtained a position as chief medical resident in the university

FIGURE 1 Hans Krebs at age 68 in his laboratory at Oxford standing next to his beloved Warburg apparatus, which he learned to use in the laboratory of Otto Warburg and which was the major instrument used in his discovery of both the urea and the citric acid cycles. The square box in the background is a power supply for a Zeiss PMQ spectrophotomer. The spectrophotometer was first divised by Otto Warburg who also first discovered pyridine nucleotide and observed their spectral shift during oxidation/reduction. This spectral change formed the basis of many of the assays Krebs undertook in his later years. With these simple instruments and a devoted staff he worked happily in this building at the Radcliffe Infirmary until he was 81, when he died in a hospital ward in the same building in 1981 after a 2 week illness.

hospital at Freiburg, and was in charge of caring for 44 patients. With his Warburg manometer and the enzyme urease, he could measure the production of urea in liver slices. In his "off time" working by himself with the assistance of a medical student, Kurt Henseleit within 1 year he had formulated a salts solution to mimic the composition of plasma, "Krebs–Henseleit saline," the basis for modern tissue culture fluids, and had worked out the ornithine cycle responsible for the hepatic synthesis of urea from ammonia. This established for the first time that cyclic, not linear, pathways could play

an important role in central metabolic processes. The process of this discovery is the subject of a detailed history dissecting the inventive processes entailed.

One of the great unsolved problems was to determine the reactions whereby foodstuffs undergo combustion to produce CO_2 and water. It was known that for glucose to undergo these reactions, it must first be broken down to pyruvate in the glycolytic pathway elucidated by Otto Meyerhof. Although in retrospect, it seems logical to assume that the idea of metabolic cycles would have been foremost in Krebs' mind when he began his

investigation into the oxidative combustion of foodstuffs, a careful analysis of his writing and notebooks gives no support for this idea. Rather his formulation of the TCA cycle was the result of the empirical methods learned in Warburg's lab. From Albert Szent-Gyorgyi, he had pigeon breast muscle minces which preserved intact, the then unknown mitochondria containing the enzymes of the TCA cycle, and could oxidatively decompose glucose to CO_2. From Warburg he had a manometer with which he could measure rates of CO_2 production. Again, he added possible intermediates and observed whether the rates of reactions increased. As he put it in his autobiography:

> One way of tackling the problem of the intermediate steps of the combustion process is to test which substances, apart from carbohydrate and fat, burn most readily. The logic is that if a substance is an intermediate then it must readily undergo combustion: if it proves to be non-combustible, it cannot be an intermediate.

In two sentences was his simple plan: the same as he had used in elucidating the urea cycle. There was no grand plan. In choosing the intermediates to be added, he was guided by others working on the same general problem. Earlier Szent-Gyorgyi had shown that addition of small amounts of succinate, fumarate, malate, or oxaloacetate could catalytically increase the rate of oxidation by pigeon breast muscle minces. Szent-Gyorgyi interpreted his findings as indicating that these compounds were not intermediates in a pathway, but rather that they served as hydrogen carriers transporting the H^+ and electrons from foodstuffs to cytochromes. By 1937, experiments by Martius and Knoop, Krebs' teacher at Freiburg, had shown in liver that citrate was converted to aconitate, then isocitrate and α-ketoglutarate, which in turn was known to be converted to succinate. It was also known from Thunberg's work that malonate could inhibit the conversion of succinate to fumarate. This presented a way for Krebs to separate the reactions of Carl Martius and Franz Knoop from those of Albert Szent-Gyorgyi. Krebs then hit on the crucial question:

> So I asked myself whether perhaps oxaloacetate, together with a substance derived from foodstuffs, might combine to form citrate again, after the manner of a cycle.

Krebs began his "Results" section on the discovery of the tricarboxcylic acid cycle with the experiments or data showing the catalytic effects of citrate on respiration in pigeon breast minces. His experimental results confirmed his hypothesis. In his mind this cycle could explain the complete combustion of foodstuffs to CO_2 and water.

Into the Wilderness and Back Again

After his triumph in discovering the urea cycle in 1932 and having been invited by Max Plank to discuss his findings, in the same year, on 19 June 1933, he was forced into exile from his own homeland, to which he and his father were devoted and whose army he had served, for reasons of racial identity which he had rejected. He found refuge in Hopkins Biochemistry Laboratory at Cambridge, dependent upon the kindness of strangers.

Triumph would again be followed by rejection, after Krebs formulation of the TCA cycle. This time, however, the rejection would not be by the Nazi's who took over his homeland, but rejection would come from his scientific colleagues. A revolutionary theory always has its critics. Krebs' tricarboxcylic acid cycle, published in 1937 after being rejected by *Nature*, was no exception. By 1940, Earl Evans, who had been trained in Krebs' lab at Sheffield, and Harland Wood, one of America's most distinguished biochemists then working at Iowa State, produced evidence that when C-labeled CO_2 condensed with pyruvate to form oxaloacetate it yielded α-ketoglutarate with the entire radioactivity confined to the carboxyl group adjacent to the carbonyl group. Wood suggested that this radioactive evidence was compatible with the condensation of pyruvate with oxaloacetate forming aconitate, then isocitrate and finally a-ketoglutarate at a rapid rate. If the formation of citrate from aconitate were only a slow side reaction, then this would account for the radioactive findings. For 7 years this reasoning was accepted and Krebs changed the name of his cycle from the more euphonious "citric acid cycle" (see Figure 2) to the more cumbersome TCA cycle, which it bears today to accommodate the generally accepted view in the biochemical community.

Krebs confined his work to practical nutritional problems for wartime Britain, which was literally starving from the depredations of Nazi submarine attacks. Then in 1948, Alexander Ogston published a short paper in Nature pointing out that the "three-point" attachment of the symmetrical citrate molecule to the citrate synthase enzyme would confer optical properties to the product of that chemically symmetrical molecule. The "citric acid cycle" was back in business as Krebs pointed out in his Harvey Lecture delivered on 17 March 1949. The final confirmation of the validity of Krebs' original postulate of the "citric acid cycle" came from Fritz Lipmann's study of the acetylation of sulfanilamides where he identified the so-called "active acetate" as acetyl coenzyme A. Soon thereafter, Stern, Ochoa, and Lynen showed that the intermediate through which pyruvate enters the TCA cycle was acetyl CoA which then combines with oxaloacetate to form citrate, confirming Krebs' original formulation of the cycle.
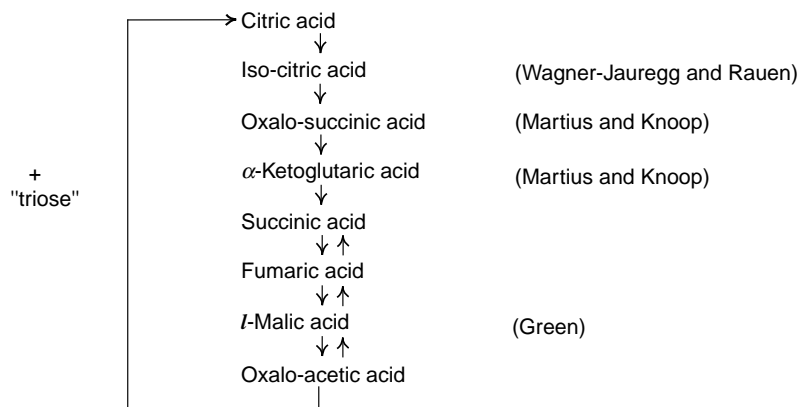
FIGURE 2 The original schematic of the citric acid cycle as it appeared in an article by Krebs and Johnson in 1937, after having been rejected by the editor of *Nature*. Krebs attempted to define the "reversible" and "irreversible" reactions of the cycle as he knew them at the time, but which would not be accepted as correct today. The nature of the "triose" combining with oxaloacetate was not known until the identification of acetyl CoA by Lipmann and Ochoa in the 1950s. The essential nature of the cycle as drawn in 1937 is however, correct to this day.

These workers also demonstrated that not only carbohydrate, but also fats and ketone bodies form acetyl CoA. This established that the Krebs citric acid cycle not only accounted for the complete conversion of carbohydrate to CO_2 and water, but that this cycle was the terminal pathway of the degradation of all foodstuffs, including fats and amino acids (see Figure 3). It may have been with some irony that Krebs entitled his Nobel prize speech of 11 December 1953, "the citric acid cycle" rather than the "TCA cycle" as it was called after the radioactive data were mistakenly interpreted to indicate that aconitate, and not citrate, was the condensation product of carbohydrate degradation. Subsequently, students of biochemistry would certainly have found the original name less fearsome than the "tricarboxcylic acid cycle," a name that evolved from erroneous interpretation of experimental results and was perpetuated by Krebs inherent modesty.

The TCA Cycle Today

It is now accepted that the TCA cycle is not only the metabolic pathway which accounts for the complete combustion of the product of glycolysis, pyruvate, to CO_2 and H_2O that is the pathway which accounts for the complete combustion of all foodstuffs, carbohydrates, fats, and amino acids in heterotrophic organisms. Much about the control of flux through the TCA cycle remains unknown even today. One persistent question is how cyclic pathways appear to defeat the laws of thermodynamics. It is easy to understand how a linear metabolic pathway, such as glycolysis, uses the chemical energy of the reactants to proceed from A to E. It is not so clear, however, how a pathway which goes from A to A accomplishes this feat in an apparent violation of thermodynamic imperatives.

The eight enzymes catalyzing the reactions of the TCA cycle and their equilibrium constants are listed in Table I.

The overall reaction of the TCA cycle is therefore:

$$AcetylCoA + 2H_2O + 3NAD^+ + coenzymeQ + GDP + Pi$$
$$\rightarrow 2CO_2 + CoA + 2H^+ + 3NADH$$
$$+ coenzymeQH_2 + GTP$$
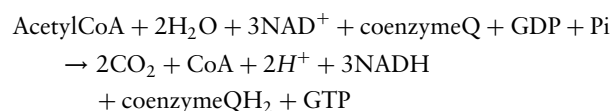
with sum standard free energies, ΔG^o at pH 7 of -11.7 kcal mol^{-1}. The sum of standard free energies of the reactions of the cycle is negative. While the sum of the standard free energies of the reactions of the cycle are negative in the direction of the cycle written, there is no complete information on the actual $\Delta G'$ of the cycle *in vivo*. The actual $\Delta G'$ of a reaction is of course, a sum of the standard free energy, ΔG^o, and a term representing the actual ratio of the concentration of products over reactants.

Without knowledge of the ratio of concentrations in mitochondria of the products over the reactants of the various eight steps of the TCA cycle, no definitive statement about the reversibility of any step can be made, although most biochemical textbooks assign one or another TCA cycle as "irreversible."

The synthesis of citrate from acetyl CoA and oxaloacetate, by citrate synthase, has the largest standard free energy of any reaction within the cycle and might therefore qualify for the irreversible designation. However, the distinguished biochemist, Paul Srere, who spent much of his career in biochemistry studying the kinetics of citrate enzymes, argued forcefully that the citrate synthase reaction achieved near equilibrium *in vivo*. Krebs himself argued that both the NAD- and NADP-dependent mitochondrial isocitrate dehydrogenases were in near equilibrium with the free mitochondrial [NAD$^+$]/[NADH] pool. A near-equilibrium reaction is hardly compatible with irreversibility.
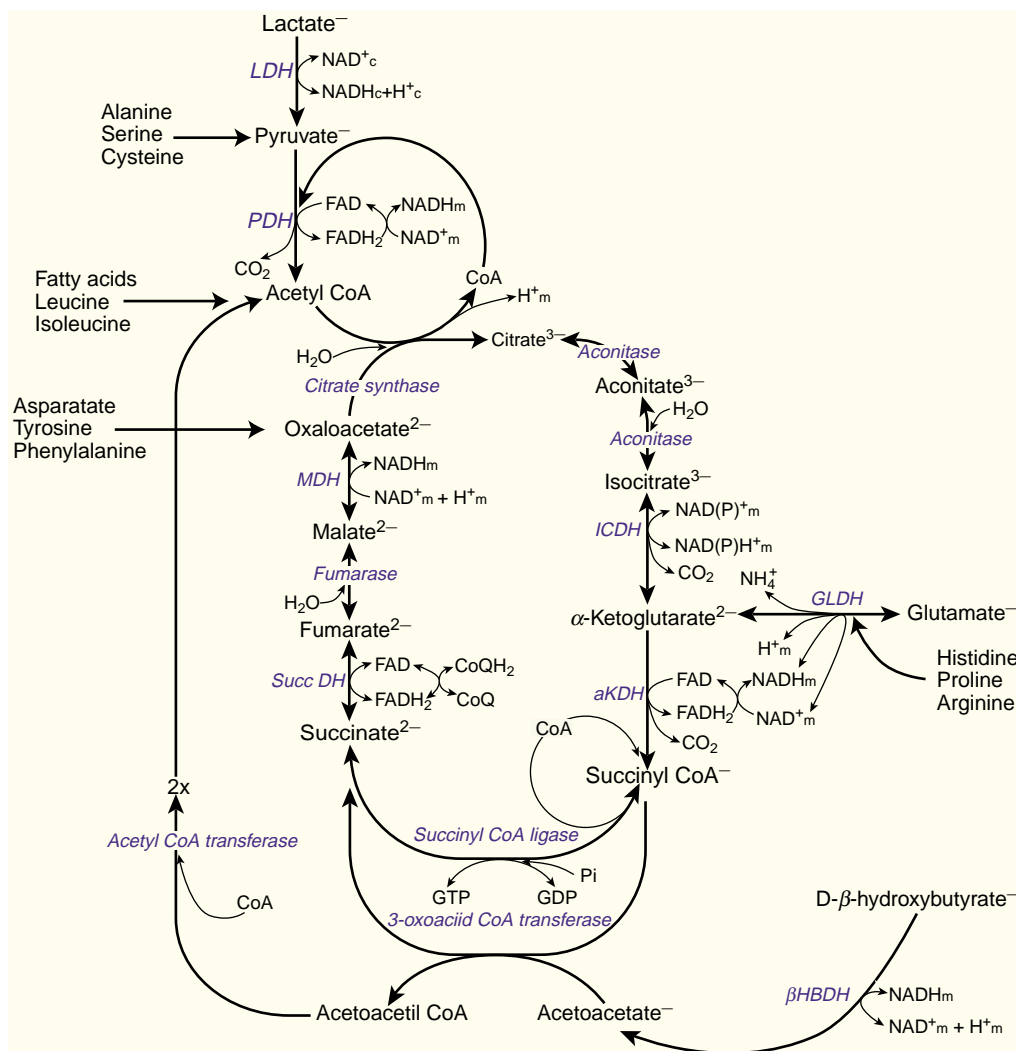
FIGURE 3   The TCA cycle is now recognized as the central pathway for the complete degradation of the major foodstuffs including carbohydrate, lipid, and proteins. It also captures the reducing equivalents which are then transferred to the electron transport system of mitochondria where the energy of their redox reactions generate the mitochondrial proton gradient used to power ATP synthesis.

The reaction catalyzed by the α-ketoglutarate dehydrogenase multienzyme complex could well be irreversible, but specific knowledge allowing one to come to a firm conclusion is lacking.

What is clear is that the disposal, of the reducing power produced in the reactions of the TCA cycle, mainly in the mitochondrial electron transport system critically affects the extent and possibly the direction of the reactions of the cycle. Within the last decade, it has been established that the TCA cycle exists in the earliest surviving life forms, the archebacteria which developed when Earth was covered in a reducing atmosphere of CO_2 and NH_3 and no O_2. Archebacteria contain a TCA cycle, which was proposed to serve as a primary anaplerotic source of metabolites and amino acids, not as the primary pathway for the degradation of foodstuffs in heterotrophs. When life forms were developing, in the

absence of O_2, the direction of the TCA cycle would have been reversed.

Whether one accepts this speculation or not, it emphasizes the intimate relationship between the TCA cycle and the production of reducing power in the form of reduced pyridine nucleotide, which serves as the primary substrate for the electron transport system. When formulated in 1937, Krebs was primarily concerned with the stoichiometry of reactions accounting for the transformation of foodstuffs into CO_2. In 1936, Otto Warburg had just discovered NADP in red cells, so the central role of these essential cofactors was not appreciated until much later when Krebs played a crucial role in defining the redox potential of the mitochondrial $[NAD^+]/[NADH]$ ratio using principles first defined Holzer, Lynen, Bucher, and Klingenberg in defining the free cytoplasmic $[NAD^+]/[NADH]$. Krebs and his

TABLE I

The Eight Enzymes Catalyzing the Reactions of the TCA Cycle and their Equilibrium Constants

Enzyme	EC number	Reaction	Keq at pH 7.0	ΔG° at pH 7 (kcal mol^{-1})
Citrate synthase	4.1.3.8	Acetyl CoA + oxaloacetate + $H_2O \rightarrow$ citrate + CoA	2.24×10^6	-9
Aconitase	4.2.1.3	Citrate $\leftrightarrow$ isocitrate	18	-1.8
Isocitrate dehydrogenase	1.1.1.41	Isocitrate + NAD(P)$^+$ $\leftrightarrow$ α-ketoglutarate + CO_2 + NAD(P)H	1.17 M	-0.01
α-Ketoglutarate dehydrogenase	Enzyme complex	α-Ketoglutarate + NAD$^+$ + CoA $\rightarrow$ succinyl CoA + CO_2 + NADH	1.17×10^3	-4.4
Succinyl CoA ligase	6.2.1.4	Succinyl CoA + GDP + Pi $\leftrightarrow$ succinate + GTP + CoA	3.5	-0.08
Succinate dehydrogenase	1.3.5.1	Succinate + ubiquinone $\leftrightarrow$ fumarate + ubiquinol	6.5	-1.2
Fumarase	4.2.1.2	Fumarate + H_2O $\leftrightarrow$ L-malate	4.4	-0.9
Malate dehydrogenase	1.1.1.37	Malate + NAD$^+$ $\leftrightarrow$ oxaloacetate + NADH	2.86×10^{-5}	6.5

colleagues then went on to show that the redox potentials of all of pyridine nucleotides were related to one another in a coherent manner. The free cytosolic NAD couple was about -0.19 V, and hence could accept reducing equivalents from glycolysis, the free mitochondrial NAD couple was a more negative -0.28 V, providing the energy necessary for mitochondrial ATP generation, while the free cytosolic NADP system was the most negative at -0.41 V, required for the completion of the reductive synthesis of fats. While it was recognized at the time that the pyridine nucleotides were related to the free [ATP]/[ADP][Pi] ratio, subtleties in the effects of free [Mg^{2+}] complicated this relationship and it took another 10 years to accurately determine the equilibrium constants required to quantitatively define that relationship.

In addition to accounting for the chemical decomposition of foodstuffs, it is now recognized that the TCA cycle produces reducing power in the form of three reduced pyridine nucleotides and one reduced coenzyme Q within the mitochondria. Each of these cofactors contains two electrons. Those contained in NADH feed into the mitochondrial electron transport system at complex I, the NADH dehydrogenase multienzyme complex. The reducing equivalents contained in reduced coenzyme Q, generated in the conversion of succinate to fumarate feed into the electron transport chain at a site with a higher redox potential than that of the mitochondrial [NAD$^+$]/[NADH] couple. Approximately four protons are ejected from mitochondria at each site creating a ΔG [H$^+$]cytosol/mitochondria, which provides the energy approximately equivalent to that of the $\Delta G'$ of ATP hydrolysis in a near equilibrium reaction catalyzed by the F1 ATPase.

It was originally thought that the TCA cycle was so fundamental that any defects in the cycle would be fatal.

It is now recognized that because of its central role, not only in the degradation of food stuffs, the TCA cycle plays a central role in generating the redox potentials not only of the mitochondrial NAD couple but also of the mitochondrial Q couple. The difference in the redox potential of these two couple, in turn, sets the magnitude of the energy of the proton gradient between the mitochondrial and cytosolic phases of the cell and thus determines the energy realized when the ATP formed by the mitochondria is hydrolyzed. The amount of realizable energy formed is a function of the amount and type of the substrate being utilized by mitochondrial TCA cycle.

See Also the Following Articles

Coenzyme A • Glycolysis, Overview • Pyruvate Carboxylation, Transamination and Gluconeogenesis • Urea Cycle, Inborn Defects of

Glossary

ΔG° The standard free energy of a reaction, a constant where the products and reactants are 1 M in activity or more practically concentration.

$\Delta G'$ The free energy of a reaction, which is not constant but varies with the ratio of activity of concentration of products divided by reactions and is defined by the formalism

$$\Delta G' = \Delta G^{\circ} + RT \ln \frac{[\text{products}]}{[\text{reactants}]}$$

TCA cycle The tricarboxylic acid cycle.

FURTHER READING

Holmes, F. L. (1991). *Hans Krebs, the Formation of a Scientific Life, 1900 to 1933*. Oxford University Press, Oxford.

Holmes, F. L. (1993). *Hans Krebs, Architect of Intermediary Metabolism, 1933 to 1937*. Oxford University Press, Oxford.

Krebs, H. A. (1954). The citric acid cycle. In *Les Prix Nobel en 1953*, pp. 139–150. Boktryckeriet P.A. Norstedt and Soner, Stockholm, Kungl.

Krebs, H. A. (1981). *Hans Krebs: Reminiscences and Reflections*. Clarendon Press, Oxford.

Krebs, H. A., and Henseleit, K. (1932). Untersuchungen über die Hamstoffbildung in Tierkörper. *Hope-Seyler's Z. Physiol. Chem.* **210**, 33–66.

Krebs, H. A., and Johnson, W. A. (1937). The role of citric acid in intermediate metabolism in animal tissue. *Enzymologia* **4**, 148–156.

Krebs, H. A., and Veech, R. L. (1969). Pyridine nucleotide interrelations. In *The Energy Level and Metabolic Control in Mitochondria* (S. Papa, J. M. Tager, E. Quagliariello and E. C. Slater, eds.) pp. 329–382. Adriatica Editrice, Bari.

Morowitz, H. J., Kostelnik, J. D., Yang, J., and Cody, G. D. (2000). The origin of intermediary metabolism. *Proc. Natl. Acad. Sci USA* **97**, 7704–7708.

Ogston, A. G. (1948). Mechanism of fixation of carbon dioxide in the Krebs cycle. *Nature* **162**, 963.

Wood, H. G., Werkman, C. H., Hemingway, A., and Nier, A. O. (1941). Interpretation of experiments on metabolic processes using isotopic tracer elements. *J. Biol. Chem.* **139**, 483.

BIOGRAPHY

Richard L. Veech is the Head of the Laboratory of Metabolic Control, NIAAA, NIH, DHHS, in Rockville, Maryland. He is an M.D. from Harvard Medical School, and has a D.Phil. in Biochemistry from Oxford University where his supervisor was Hans Krebs. He works on problems involving the control of the processes of intermediary metabolism, particularly in disease states.

tRNA Synthetases

Karla L. Ewalt and Paul Schimmel
Scripps Research Institute, La Jolla, California, USA

Protein biosynthesis is the culmination of the transfer of genetic information from DNA to proteins. It is a highly coordinated process that requires many cellular factors including RNAs, proteins, nucleotides, and amino acids. An essential component for the translation of genetic information from genes to proteins, is the group of enzymes known as aminoacyl-tRNA synthetases. These enzymes covalently link amino acids to transfer RNAs (tRNAs) and thereby establish the rules of the genetic code that pair amino acids to codons. The translation of codons on the messenger RNA into amino acids occurs at the ribosome – the center of protein biosynthesis.

Enzymatic Reaction

Accurate protein biosynthesis requires a collection of amino acids attached to tRNAs (aminoacyl-tRNAs), which are typically generated by 20 amino acid-specific enzymes called aminoacyl-tRNA synthetases. The aminoacylation reaction is carried out at the enzymes active sites in two-steps. In the first step, an amino acid (AA) is condensed with ATP yielding an aminoacyl-adenylate (AA-AMP) and pyrophosphate (PPi) (Scheme 1). In the second step, the aminoacyl group is transferred to either the 2′ or 3′-ribose hydroxyl on the terminal adenosine of tRNA (AA-tRNA). Certain tRNA synthetases contain a second active site (editing active site) for the hydrolysis of incorrectly paired amino acids and tRNAs. At this site, the tRNA synthetase corrects its errors.

Translational Accuracy

Aminoacyl-tRNA synthetases establish and determine the accuracy of the genetic code and of protein biosynthesis, as they pair amino acids with codons through selection of amino acids and tRNAs. Accuracy is achieved during selection of amino acids and tRNAs, as well as through editing reactions that correct errors.

Amino Acids

Each tRNA synthetase is specific for one of the 20 amino acids. A selective binding site for the amino acid accomplishes discrimination among the various amino acid substrates. However, amino acid substrates for a few enzymes are closely similar in size and functional groups to other amino acids. For example, isoleucyl-tRNA synthetase selects between isoleucine and valine, which are both β-branched aliphatic amino acids differing in length by only one methylene. The misactivation of valine to valyl-adenylate by isoleucyl-tRNA synthetase occurs at a low frequency. Threonyl-tRNA synthetase must select between threonine and serine, which differ by only a single methylene. The active site is a "course sieve" that allows the correct amino acid and occasionally similar, but slightly smaller amino acids to be transferred to tRNA.

tRNAs

tRNAs are adapters that covalently connect anticodons with amino acids. Sequences of tRNAs are 70–90 nucleotides (76 is most common) that form a common cloverleaf secondary structure with four major arms. This secondary structure in turn folds into an L-shaped tertiary structure made up of two domains (Figure 1). Amino acids are attached to tRNAs at the 3′-terminal adenosine at the end of a CCA trinucleotide sequence that is common to all tRNAs. In the second domain is a trinucleotide sequence known as the anticodon, located ∼75Å from the amino acid attachment site. During protein synthesis, the anticodon triplet binds through complementary base pairing to a codon on mRNA and thereby delivers an amino acid to the growing polypeptide chain.

Since the genetic code is degenerate, more than one tRNA may exist for a single amino acid. The 61 sense codons of the code are recognized by a combination of distinct tRNAs (one tRNA specific for each codon) and tRNAs that employ wobble-base pairing (one tRNA that recognizes sets of related codons). (A notably reduced set of tRNAs is found in the mitochondria of eukaryotes, where 22 tRNAs read the codons.) Although, tryptophan and methionine each have a single tRNA, there are as many as five tRNAs specific for the six leucine codons. The multiple tRNAs specifying a single amino acid are termed isoacceptors, and are substrates for a single

$$E + AA + ATP \rightleftharpoons E{\cdot}(AA{-}AMP) + PPi$$

$$E{\cdot}(AA{-}AMP) + tRNA \rightleftharpoons E + AA\text{-}tRNA + AMP$$
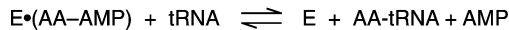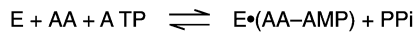
SCHEME 1 Two-step aminoacylation reaction catalyzed by aminoacyl tRNA synthetases (E). Amino acid (AA) and ATP react to form an activated aminoacyl-adenylate (AA-AMP) and pyrophosphate (PPi). The aminoacyl group is transferred to either the 2′ or 3′ ribose hydroxyl of the terminal adenine on tRNA.

aminoacyl-tRNA synthetase. Specific recognition of tRNA is primarily through a nucleotide determinant located in the acceptor stem. Nucleotides at other positions including the anticodon stem are also important in many instances. The specific recognition of tRNAs by synthetases based on the nucleotides present in the acceptor stem, has led to the idea that these features constitute a second genetic code that may be left over from a smaller, historical RNA, which predated the current tRNA structure.

EDITING

Several tRNA synthetases have an editing activity to ensure correct pairing of amino acids with tRNAs (Scheme 2). This editing, or proofreading, function prevents misincorporation of amino acids into proteins by hydrolyzing incorrectly aminoacylated tRNAs (post-transfer editing). The tRNA synthetases for isoleucine, leucine, valine, threonine, and alanine have been shown to have specific domains with editing active sites that are separate from the active site for aminoacylation. The editing site operates as a "fine sieve" to exclude the correct aminoacyl-tRNA while binding the smaller and incorrect aminoacyl-tRNA, which is then hydrolyzed. It is currently thought that tRNA synthetases can also hydrolyze misactivated amino acids prior to attachment to the tRNA (pre-transfer editing). Through the combination of proper substrate selection and editing, the general error rate for protein synthesis is estimated to be ≤ 1/3000.
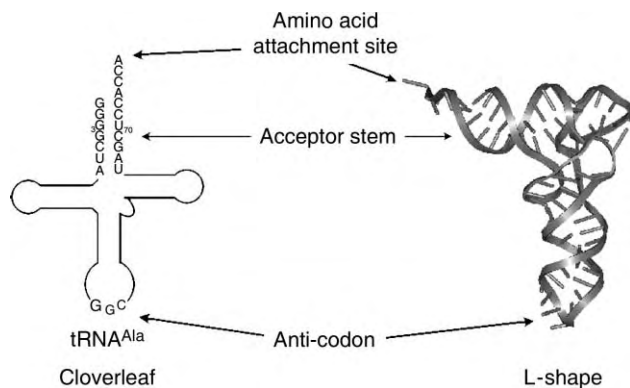
FIGURE 1 The cloverleaf secondary and L-shape tertiary structure of tRNA.

Two Classes of Enzymes

STRUCTURES

Aminoacyl-tRNA synthetases are divided into two groups of ten enzymes – class I and class II – based upon the sequences and structures of the active site domains. Ten different class I enzymes activate and transfer the amino acids Arg, Cys, Glu, Gln, Ile, Leu, Met, Tyr, Trp, and Val to tRNAs. All class I enzymes contain a Rossmann nucleotide-binding domain and an active site with two signature sequences, an 11 amino acid element that ends in HIGH and a KMSKS pentapeptide. The active site is interrupted by a polypeptide insertion called CP1 that encodes an editing domain. (Figure 2). Similarly, there are ten class II enzymes that activate and transfer the amino acids Ala, Asn, Asp, Gly, His, Lys, Phe, Pro, Ser, and Thr to tRNAs. These ten enzymes share an active site architecture formed by three conserved motifs (motif 1, 2, and 3) in an overall structure having a seven-stranded antiparallel β-sheet flanked by α-helices. Enzymes in both classes have additional motifs or domains that are idiosyncratic to each enzyme. Some of these domains provide RNA binding elements, while others have functions such as signaling in cytokine pathways.

MECHANISMS

The enzymes in the two classes are distinguished by the orientation with which they bind tRNA. Class I enzymes generally bind tRNA on the minor groove side of the RNA helix and attach the amino acid to the 2′-ribose hydroxyl of the terminal adenosine. Most class II enzymes bind tRNA on the major-groove side and attach the amino acid to the 3′-ribose hydroxyl. While aminoacylation is specific to a particular hydroxyl, the amino acid rapidly migrates between the 2′- and 3′-ribose hydroxyl groups.

Alternative Pathways and Enzymes

In a few cases, amino acids are modified while attached to specific tRNAs in order to generate different aminoacyl-tRNA species. This type of mechanism has evolved as a biosynthetic pathway for glutamine, asparagine, and selenocysteine.

tRNA DEPENDENT AMIDATION

Bacteria and archea have developed an alternate two-step route for producing aminoacylated tRNAs for glutamine and asparagine. In these organisms, glutamic acid is attached to the glutamine specific tRNA

$$\text{E} \bullet (\text{AA--AMP}) + \text{tRNA} \rightleftharpoons \text{E} \bullet (\text{AA--AMP}) \bullet \text{tRNA} \rightleftharpoons \text{E} \bullet (\text{AA--tRNA}) + \text{AMP}$$

Pretransfer editing \qquad Posttransfer editing

$$\text{E} + \text{AA} + \text{AMP} + \text{tRNA}$$
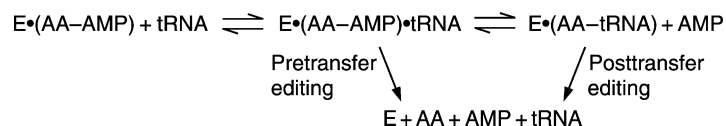
SCHEME 2   Pretransfer and post-transfer editing reactions catalyzed by aminoacyl-tRNA synthetases (E). Pretransfer editing occurs prior to the aminoacyl group transfer to tRNA. Posttransfer editing occurs after the aminoacyl group has been transferred to tRNA. The net result of both pathways is the enzymatic hydrolysis of the aminoacyl-group (AA-AMP or AA-tRNA) to the free amino acid (AA) with release of AMP and tRNA.

(tRNA$^{\text{Gln}}$) by glutamyl-tRNA synthetase. This system requires that the glutamyl-tRNA synthetase have a relaxed specificity that permits the acylation of glutamic acid onto both tRNA$^{\text{Glu}}$ and tRNA$^{\text{Gln}}$. In the second step, glutamine is produced as Gln-tRNA$^{\text{Gln}}$ from Glu-tRNA$^{\text{Gln}}$, by the enzyme glutamyl-tRNA$^{\text{Gln}}$ amidotransferase. In these two kingdoms, this system is the predominant mechanism for producing Gln-tRNA$^{\text{Gln}}$. Although some bacteria and archea have the gene encoding asparaginyl-tRNA synthetase, an analogous system also exists for the amidation of Asp-tRNA$^{\text{Asn}}$ to produce Asn-tRNA$^{\text{Asn}}$.

SELENOCYSTEINE

Sometimes called the 21st amino acid, selenocysteine is incorporated into proteins at the UGA codon that normally is a stop codon. However, a special tRNA (tRNA$^{\text{Sec}}$) exists for the incorporation of selenocysteine at this position. Selenocysteyl-tRNA$^{\text{Sec}}$ is produced by modification of Ser-tRNA$^{\text{Sec}}$. Serine is attached to the tRNA$^{\text{Sec}}$ by seryl-tRNA synthetase, and is then modified

by selenocysteine synthase. The enzyme incorporates selenide through a pyridoxal 5-phosphate-mediated reaction. Incorporation of selenocysteine into proteins at UGA is dependent on the presence of a selenocysteine-insertion-sequence in the mRNA that is immediately downstream of the codon.

Novel Functions of tRNA Synthetases

In addition to their essential role in protein biosynthesis, several tRNA synthetases are involved in other cellular processes. The novel functions that appear in tRNA synthetases are not common features to all of the enzymes for a particular amino acid, but rather are specialized features of the enzymes in selected organisms. For examples, mitochondrial leucyl- and tyrosyl-tRNA synthetase promote splicing in certain fungi. Bacterial threonyl-tRNA synthetase regulates its own translation by binding to an upstream region of the
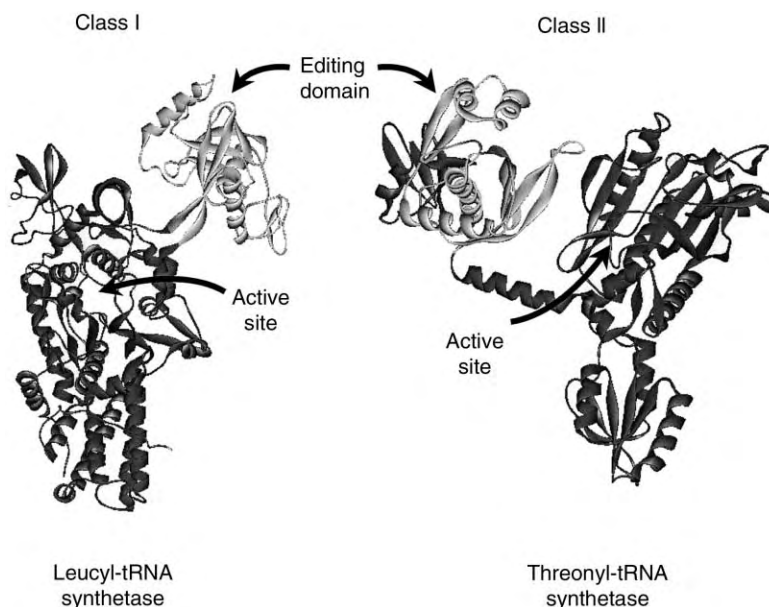
FIGURE 2   The crystallographic structure of a class I enzyme, leucyl-tRNA synthetase, and a class II enzyme, threonyl-tRNA synthetase, indicating the active sites and editing domains.

mRNA, and a histidyl-tRNA synthetase related protein called GCN2 regulates translation in eukaryotes. Increasingly, roles for tRNA synthetases as cell signaling factors are being discovered in higher eukaryotes. Fragments of tyrosyl- and tryptophanyl-tRNA synthetase are mediators of angiogenesis, while histidyl- and asparaginyl-tRNA synthetase specifically stimulate immune cells. These extra functions demonstrate how tRNA synthetases have been adapted and utilized for a wide variety of functions in addition to their enzymatic role in protein biosynthesis.

SEE ALSO THE FOLLOWING ARTICLES

Amino Acid Metabolism • Pre-tRNA and Pre-rRNA Processing in Bacteria • Pre-tRNA and Pre-rRNA Processing in Eukaryotes • RNA Editing

GLOSSARY

aminoacyl-tRNA synthetases   A family of enzymes that attach each amino acid to its appropriate tRNA.

anticodon   The three nucleotide sequence on tRNA that binds to the codon through complementary base pairing.

codon   A three nucleotide RNA sequence that specifies a particular amino acid.

translation   The process of protein biosynthesis from messenger RNA.

tRNA   A small RNA molecule that carries an amino acid to the ribosome during protein biosynthesis, recognizes the codon specifying the amino acid and delivers it to the growing polypeptide chain.

FURTHER READING

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (2002). *Molecular Biology of the Cell*. Garland Publishing, New York.

Schimmel, P., and Ribas de Pouplana, L. (2000). Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.* **25**(5), 207–209.

Soll, D., and RajBhandary, U. L. (1995). *tRNA Structure, Biosynthesis, and Function.* ASM Press, Washington, DC.

Voet, D., and Voet, J. G. (1995). *Biochemistry.* Wiley, New York.

BIOGRAPHY

Paul Schimmel is the Ernest and Jean Hahn Professor and member at the Skaggs Institute for Chemical Biology at the Scripps Research Institute. He holds a Ph.D. from the Massachusetts Institute of Technology, where he went on to become the John D. and Catherine T. MacArthur Professor of Biochemistry and Biophysics in the Department of Biology prior to joining the Scripps Research Institute. His major research activities have concentrated on the decoding of genetic information, with an emphasis on the rules of the universal genetic code established by aminoacyl-tRNA synthetases.

Karla L. Ewalt is a Staff Scientist in the Department of Molecular Biology at the Scripps Research Institute. She holds a Ph.D. from the University of California, San Diego. Her research focus is on the functional specializations found in components of eukaryotic protein translation machinery.

trp Operon and Attenuation

Paul Gollnick

State University of New York, Buffalo, New York, USA

Transcription attenuation can be defined as any mechanism that utilizes transcription pausing or termination to control expression of downstream genes. This mechanism of regulation is used to control expression of many genes in bacteria in responses to changes in their environment. Moreover, bacteria have evolved several elaborate and elegant variations on the attenuation theme, mainly based on the mechanism used to control formation the RNA structures that signal RNA polymerase to pause or terminate.

Attenuation

The first attenuation mechanism to be described was for the *Escherichia coli* tryptophan biosynthetic (*trpEDCBA*) operon. When Charles Yanofsky and his co-workers elucidated this mechanism, it was the first demonstration that organisms can utilize changes in RNA structure to regulate gene expression. Subsequently many other examples of transcription attenuation have been discovered. In each case the *cis*-acting genetic information is contained within a 150–300 bp region located after the start of transcription and prior to the start of the coding sequence for the first structural gene of the operon. This region is called the leader region. In this article, the features of transcription attenuation control of the *E. coli trp* operon are discussed, followed by the modifications in this classic system that allow it to be adapted to control other amino acid biosynthetic operons. Thereafter attenuation control of the *trp* operon in the gram-positive bacterium *Bacillus subtilis*, which involves a larger variation on the attenuation theme, is described. There are several recent reviews that cover the attenuation in more depth as well as describe attenuation control of other bacterial genes.

The *E. coli trp* Operon: Attenuation Based on Translation of a Leader Peptide

The *E. coli trpEDCBA* operon encodes the enzymes required to synthesize *L*-tryptophan from chorismic acid. Transcription of the *trp* operon is regulated in response to changes in intracellular tryptophan levels. When the cells contain adequate amounts of tryptophan, for example when it is present in the growth medium, transcription of the operon is down-regulated. In contrast, when tryptophan is limiting, the *trp* operon is actively transcribed in order to express the enzymes required for its synthesis. Initiation of transcription is regulated by the *trp* repressor, a DNA-binding protein encoded by the *trpR* gene. In addition, after transcription has initiated, the elongating transcription complex is subject to regulation by attenuation. Together, repression (80-fold) and attenuation (8-fold) serve to allow ∼600-fold overall control of transcription of the *trp* operon in response to various levels of tryptophan availability.

ATTENUATION CONTROL OF THE *E. COLI TRP* OPERON

The *E. coli trp* operon contains a 162 bp leader region prior to the start of the *trpE* coding sequence. The *trp* leader transcript contains several inverted repeats, composed of the segments labeled 1–4 in Figure 1, that can form three different overlapping base-paired RNA secondary structures. These structures include an intrinsic transcription terminator (3:4), an overlapping antiterminator (2:3), and a pause structure (1:2). In addition, the leader transcription contains a small open reading frame (ORF) that encodes a 14-amino acid leader peptide, which contains two critical tandem UGG Trp codons. The cell's ability to efficiently translate these two Trp codons determines which RNA structure forms in the nascent leader transcript, which in turn controls whether transcription halts in the leader region or continues into the structure genes of the operon.

Shortly after transcription initiates from the *trp* promoter, the 1:2 pause hairpin forms (Figure 1). This structure signals RNA polymerase to pause transcription after nucleotide 92. This pausing of RNA polymerase is a critical feature of the attenuation mechanism because it allows time for a ribosome to
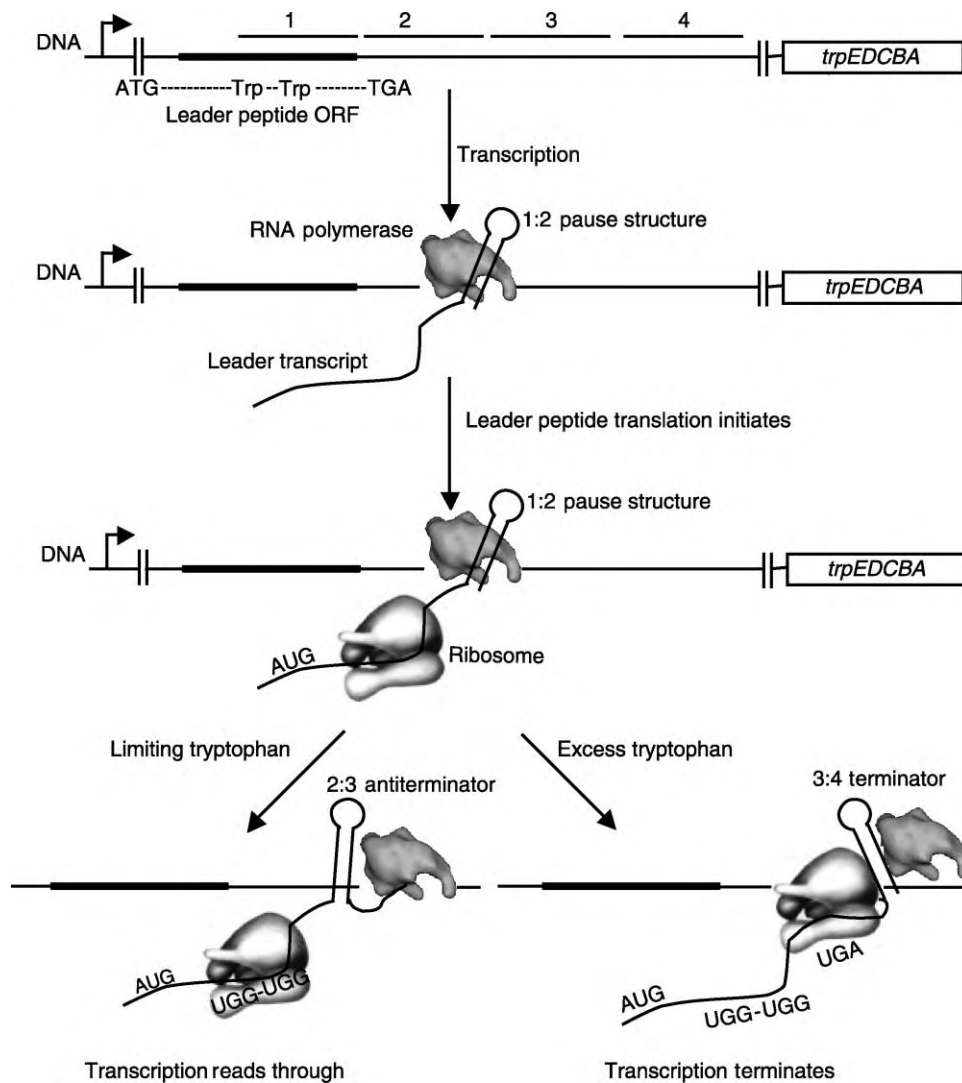
FIGURE 1 Model of transcription attenuation of the *E. coli trp* operon. RNA polymerase pauses following formation of the pause structure. This provides time for a ribosome to initiate translation of the leader peptide. Under tryptophan-limiting conditions the ribosome stalls at the tandem Trp codons, resulting in transcription read through. Under conditions of tryptophan excess the ribosome reaches the leader peptide stop codon. This ribosome position blocks formation of the antiterminator leading to terminator formation and transcription termination.

initiate translation of the leader peptide. When the ribosome begins translating the leader peptide this releases the paused RNA polymerase to resume transcription. Transcription and translation are now coupled, with the ribosome closely following RNA polymerase. This situation is essential to allow events involving the ribosome to affect transcription by the associated RNA polymerase.

At this point there are two possible pathways for the attenuation mechanism to follow depending on the level of tryptophan in the cell. The choice depends on how efficiently the tandem Trp codons in the leader peptide are translated. This efficiency reflects the availability of aminoacylated tRNA$^{\text{Trp}}$ in the cell.

Under conditions where tryptophan is limiting, the amount of charged tRNA$^{\text{Trp}}$ is low. As a result of this low concentration of tryptophanyl-tRNA$^{\text{Trp}}$ translation of the tandem Trp codons is inefficient and the ribosome stalls at one of these two codons. The associated RNA polymerase continues to transcribe through the *trp* leader region and transcription and translation become uncoupled. As RNA polymerase proceeds through segments 2 and 3 of the leader region, the antiterminator structure (2:3) forms, which prevents formation of the overlapping intrinsic terminator (3:4) structure. Hence transcription continues through the leader region and into the *trp* structural genes.

When tryptophan is plentiful, the level of charged tRNA^Trp in the cell is high. This allows efficient translation of the tandem Trp codons and hence the ribosome proceeds rapidly to the end of the leader peptide. When the ribosome reaches the leader peptide stop codon, it covers part of RNA segment 2 and thus prevents formation of the 2:3 antiterminator structure as transcription proceeds. This frees RNA segment 3 to base pair with segment 4 and form the terminator. Under these conditions transcription terminates in the leader region prior to the *trp* structural genes, which are therefore not expressed.

In this attenuation mechanism the regulatory signal is the level of charged tRNA^Trp and the sensory event is the efficiency with which the ribosome can translate the tandem Trp codons in the leader peptide. This system can easily be adapted to regulate other bacterial amino acid biosynthetic operons. The only needed modification is to change the identity of the critical codons in the leader peptide, which controls the amino acid the system will respond to. There are numerous examples of such adaptations of this attenuation mechanism, particularly in enteric bacteria, including the *his*, *phe*, and *leu* operons, which contain seven His, seven Phe and four Leu codons in their respective leader peptides.

The *Bacillus subtilis trp* Operon: Attenuation Mediated by an RNA-Binding Protein

The transcription attenuation mechanism that regulates the *trp* operon in the gram-positive bacterium *Bacillus subtilis* differs more dramatically from that the *E. coli trp* operon than those described above for other amino acid biosynthetic operons in gram-negative bacteria. Most notably there is no leader peptide, and ribosomes are not involved in this attenuation mechanism. Instead an RNA-binding protein senses the level of tryptophan in the cell and determines whether transcription terminates in the leader region or continues into the structural genes.

THE *B. SUBTILIS TRP* LEADER

Expression of the *trpEDCFBA* operon in *B. subtilis* and several related bacilli is regulated by the *trp* RNA-binding Attenuation Protein (TRAP). Transcription of the operon initiates from the *trp* promoter 203 nucleotides upstream of the start codon of the first structural gene, *trpE* (Figure 2). There is no evidence for any regulation of initiation of transcription from this promoter. Hence there is no homologue, either
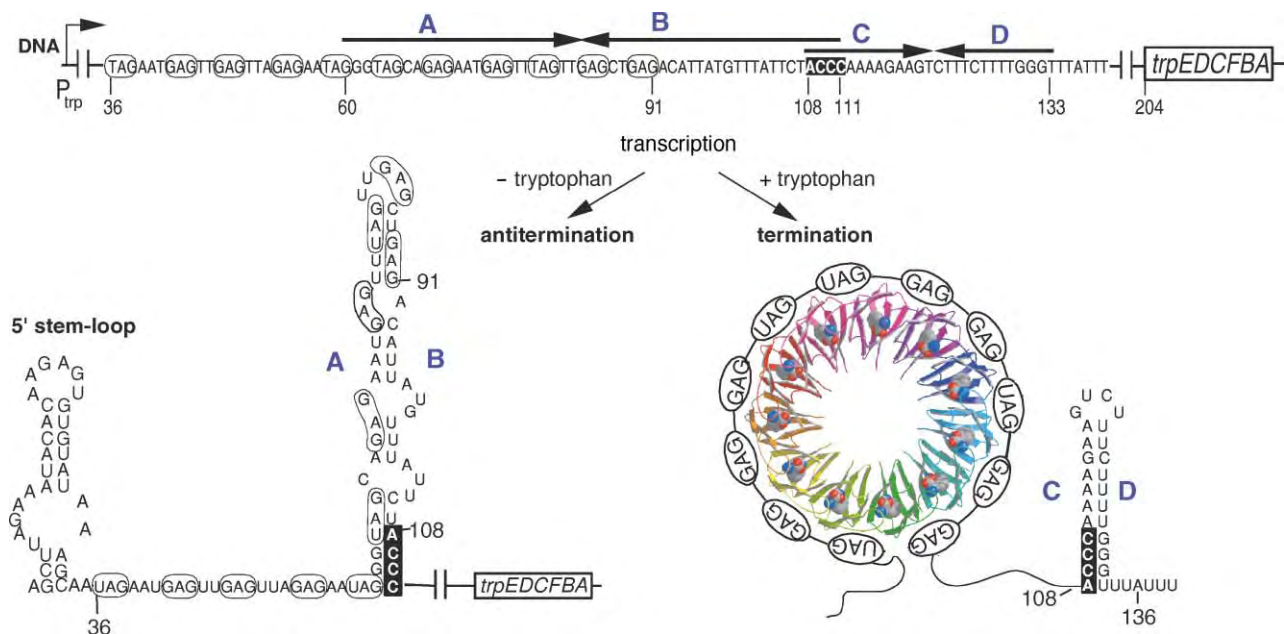
FIGURE 2 Model of transcription attenuation of the *B. subtilis trpEDCFBA* operon. The large blue letters indicate the complementary strands of the terminator and antiterminator RNA structures. The TRAP protein is shown as a ribbon diagram with each of the 11 subunits as a different color and the bound RNA is shown forming a matching circle upon binding to TRAP. The bound tryptophans are shown as space filling models. The GAG and UAG repeats involved in TRAP binding are shown in ovals and are also outlined in the sequence of the antiterminator structure. Numbers indicate the residue positions relative to the start of transcription. Nucleotides 108–111 overlap between the antiterminator and terminator structures and are shown as darkly outlined letters.

functional or based on amino acid sequence, of the *E. coli* DNA binding *trp* repressor in *B. subtilis*. The *B. subtilis trp* leader transcript also contains inverted repeats that can form an intrinsic transcription terminator (C:D) and an antiterminator (A:B) RNA secondary structure (Figure 2). These two structures overlap by four nucleotides and hence their formation is mutually exclusive.

In the presence of excess tryptophan, TRAP binds to a series of 11 trinucleotide repeats in the *trp* leader transcript consisting of seven GAGs and four UAGs (Figure 2). The (G/U)AG repeats are separated from each other by two or three "spacer" nucleotides, whose sequence is not conserved. These triplet repeats in part overlap the 5′ portion of the antiterminator structure and hence TRAP-binding interferes with formation of the antiterminator (A:B) thus favoring formation of the terminator (C:D) hairpin. Therefore under these conditions transcription halts in the leader region prior to the *trp* structural genes, which are not expressed.

When tryptophan is limiting, TRAP does not bind to the *trp* leader RNA transcript and formation of the A:B antiterminator is favored. This situation allows transcription to proceed into the *trp* operon structural genes, which can now be expressed to produce the enzymes to synthesize tryptophan.

THE TRAP PROTEIN

In this system the ability of TRAP to sense the levels of intracellular tryptophan and to then influence the structure of the nascent *trp* leader mRNA transcript is the basis for attenuation control of the *trp* operon. TRAP is a ring-shaped protein composed of 11 identical subunits (Figure 3). TRAP is activated to bind its RNA targets by binding 11 molecules of *L*-tryptophan in clefts between adjacent subunits. The detailed mechanism by which tryptophan binding activates TRAP to bind RNA is not known, however, recent studies suggest that tryptophan binding induces a conformational change in TRAP as well as reduces the flexibility of the protein. Upon activation by tryptophan, TRAP binds to its RNA target consisting of the 11 (G/U)AG repeats by wrapping the single-stranded RNA around the outer perimeter of the protein ring (Figure 3). The phosphodiester backbone is on the outside of the RNA ring and the bases point in toward the protein. The specificity of this interaction derives mainly from hydrogen bonds between several amino acids of each TRAP subunit including Glu36, Lys37, Lys56, and Arg58, and the bases on the (G/U)AG repeats. The spacer nucleotides do not contact the protein, and their identity is not crucial for TRAP recognition and binding. This mechanism of wrapping the RNA around the protein ring explains
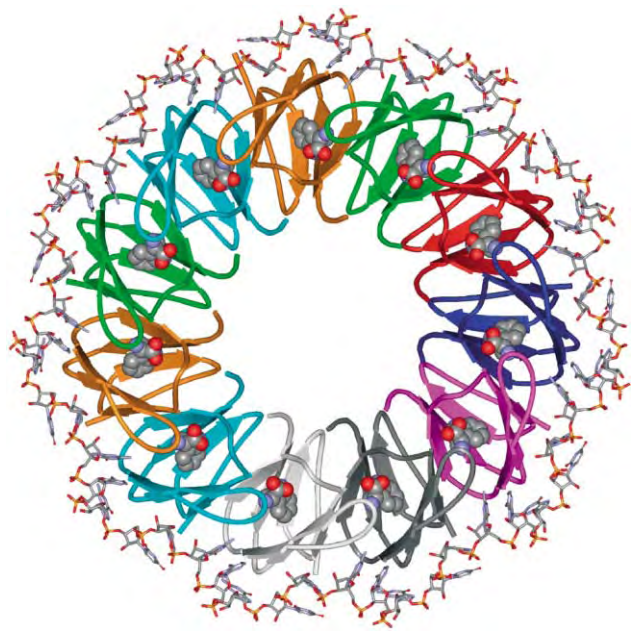
FIGURE 3 Ribbon diagram of TRAP complexed with an RNA containing 11 GAG repeats separated by AU spacers. The TRAP protein is shown in ribbon diagrams with each subunit depicted in a different color. The RNA is shown in stick models and the 11 molecules of bound *L*-tryptophan are shown in space-filling models.

both the specificity of the TRAP–RNA interaction as well as how TRAP binding alters the structure of the *trp* leader RNA to regulate attenuation.

THE ROLE OF PAUSING IN TRAP MEDIATED ATTENUATION

In addition to the terminator and antiterminator structures, there is an RNA hairpin at the 5′ end of the *trp* leader mRNA (Figure 2; 5′ stem-loop). Although this structure has been shown to be important for proper attenuation control of the *trp* operon, it is not a transcriptional pause signal analogous to the 1:2 structure of the *E. coli trp* attenuation system. The role of this 5′ stem-loop may be to enhance the rate of TRAP binding to the *trp* leader RNA. This is an important consideration because TRAP must bind to its target before RNA polymerase transcribes beyond the terminator structure, otherwise TRAP binding would not be able to influence attenuation. Recent studies published by Yakhnin and Babitzke indicate that RNA polymerase pauses at U107 (Figure 2) during transcription of the *trp* leader region, and that this pausing is dependent on the NusA transcription factor. This pause would provide additional time for TRAP to bind to the nascent *trp* leader RNA and promoter termination.

ANTI-TRAP AND THE ROLE OF tRNA^Trp IN REGULATING THE *B. SUBTILIS trp* GENES

In this attenuation system the regulatory signal is the level of free tryptophan available to activate TRAP to bind RNA. In the previously described *E. coli trp* attenuation system, free tryptophan is sensed by the DNA binding *trp* repressor protein and the regulatory signal for attenuation is the availability of charged tRNA^Trp. The level of charged tRNA^Trp also influences attenuation control of the *B. subtilis trp* operon, although by a very different mechanism than in the *E. coli* system. In *B. subtilis* the *rtpA* gene encodes a protein called anti-TRAP (AT). AT influences expression of the *trp* genes by binding to tryptophan-activated TRAP and preventing it from binding RNA, thus elevating expression of *trp* genes. Expression of *rtpA*, which is in a two-gene operon together with *ycbK* (unknown function), is regulated in response to changes in the level of uncharged tRNA^Trp by a mechanism known as T-box antitermination. High levels of uncharged tRNA^Trp induce expression of AT, which prevents TRAP from down-regulating expression of the *trp* genes. Hence even if free tryptophan levels are high enough to activate TRAP, if the level of charging of tRNA^Trp is low, then the *trp* genes will be expressed. Thus while both *E. coli* and *B. subtilis* regulate expression of the *trp* genes in response to changes in levels of both free tryptophan and aminoacylated tRNA^Trp, they have evolved very different mechanisms to do so.

OCCURRENCES OF TRAP MEDIATED ATTENUATION IN BACTERIA

TRAP mediated attenuation control of the *trp* operon has till date been only observed in *B. subtilis* and several related bacilli including *B. pumilus, B. stearothermophilus, B. caldotenax*, and *B. halodurans*, as well as in *Clostridium thermocellum*. In contrast to the leader-peptide-dependent mechanism described previously, the RNA-binding protein-dependent attenuation mechanism has never been characterized for any other amino acid biosynthetic operons beside tryptophan even though it would seem to be easily adaptable by simply changing the amino acid that activates the protein to bind its RNA target. Perhaps with the explosion of bacterial genomic information currently becoming available we might soon discover such a system.

SEE ALSO THE FOLLOWING ARTICLES

DNA Polymerase I, Bacterial • DNA Polymerase II, Bacterial • DNA Polymerase III, Bacterial • Ribosome Structure • RNA Polymerase Reaction in Bacteria • RNA Polymerase Structure, Bacterial • T7 RNA Polymerase

GLOSSARY

intrinsic terminator A signal in the nascent RNA transcript that signals RNA polymerase to halt transcription and dissociate from the DNA template. Intrinsic terminators consist of a short base-paired stem-loop structure followed by a short stretch of U residues in the RNA. Also called factor-independent terminator or Rho-independent terminator.

ribosome The large RNA–protein complex that translated mRNA into protein. It consists of two subunits, termed small and large.

RNA polymerase The enzyme that transcribes DNA into RNA. In bacteria there is only one version of this enzyme.

transcriptional pausing In response to signals in the RNA, RNA polymerase will pause and discontinue transcription but not dissociate from the DNA template. Either after some period of time, or in response to a signal, a paused RNA polymerase will resume transcription.

tRNA The adaptor RNA molecule that reads the genetic code in the mRNA by base-pairing with the codon triplets. An amino acid is attached to the 3′ end of the last base in the tRNA by an enzyme called an aminoacyl tRNA synthetase in a process called aminoacylation or charging. There are tRNAs corresponding each of the 20 amino acids.

FURTHER READING

Antson, A. A., Otridge, J. B., Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Wilson, K. S., Smith, T. M., Yang, M., Kurecki, T., and Gollnick, P. (1995). The three dimensional structure of *trp* RNA-binding attenuation protein. *Nature* **374**, 693–700.

Antson, A. A., Dodson, E. J., Dodson, G. G., Greaves, R. B., Chen, X.-P., and Gollnick, P. (1999). Structure of the trp RNA-binding attenuation protein, TRAP, bound to RNA. *Nature* **401**, 235–242.

Babitzke, P., and Gollnick, P. (2001). Posttranscriptional initiation control of tryptophan metabolism in *Bacillus subtilis* by the *trp* RNA-binding attenuation protein (TRAP), anti-TRAP, and RNA structure. *J. Bacteriol.* **183**, 5795–5802.

Valbuzzi, A., and Yanofsky, C. (2001). Inhibition of the B-subtilis regulatory protein TRAP by the TRAP-inhibitory protein, AT. *Science* **293**, 2057–2059.

Yakhnin, A. V., and Babitzke, P. (2002). NasA-stimulated RNA polymerase pausing and termination participates in the *Bacillus subtilis trp* operon attenuation mechanism *in vitro*. *Proc. Natl. Acad. Sci. USA* **99**, 11067–11072.

Yanofsky, C. (2000). Transcription attenuation: Once viewed as a novel regulatory strategy. *J. Bacteriol.* **182**, 1–8.

Yanofsky, C. (2001). Advancing our knowledge in biochemistry, genetics, and microbiology through studies on tryptophan metabolism. *Annu. Rev. Biochem.* **70**, 1–37.

BIOGRAPHY

Paul Gollnick is a Professor in the Department of Biological Sciences at the University of Buffalo, the State University of New York. His principal research interests are in RNA–protein interactions and regulation of gene expression. He holds a Ph.D. from Iowa State University and received postdoctoral training at Stanford University. He and his collaborator Dr. Alfred Antson at York University in England have determined the structure of TRAP and the TRAP:RNA complex.

Tubulin and its Isoforms

Eva Nogales
Howard Hughes Medical Institute, University of California, Berkeley, California, USA

Tubulin is an $\alpha\beta$ dimeric protein that self-assembles into microtubules and is present in all eukaryotes. Tubulin is highly conserved across species, reflecting the sequence constraints imposed by microtubule structure and function. Both α- and β-subunits exist in numerous isotypic forms and undergo a variety of posttranslational modifications. Tubulin assembly and disassembly, which are linked to GTP hydrolysis, make the microtubule network dynamic both in time and space to accommodate the needs of the cell during the cell cycle. Purified tubulin retains its self-assembling capabilities, allowing the biochemical and biophysical characterization of the microtubule polymerization and depolymerization processes. A variety of proteins interact with tubulin in the cell, affecting the stability of microtubules and their function. Numerous ligands bind to tubulin and influence its assembly properties, among them several drugs that have proven to have anticancer properties.

γ-tubulin is a rarer tubulin isoform involved in the nucleation of microtubules at microtubule organizing centers. Recently, additional tubulin isoforms of yet ill-defined function have been identified.

$\alpha\beta$-Tubulin

Physical Properties

General

Tubulin α- and β-subunits have molecular weights of ~ 50 kDa and are 36–42% identical and 63% homologous. Both tubulin subunits bind guanine nucleotides. The binding to α-tubulin at the N-site is nonexchangeable, while the binding to β-tubulin at the E-site is exchangeable. Magnesium increases the affinity of the β-subunit for GTP with respect to GDP. Nucleotide in oligomeric tubulin or in microtubules does not exchange with the solution, except for terminal subunits at microtubule ends.

Neuronal cells are particularly rich sources of tubulin because microtubules are required for axonal transport. Neural tissue contains sufficient tubulin to allow tubulin purification by repeated cycles of 37°-induced assembly and 0°-induced disassembly, with intervening centrifugation to alternately pellet microtubules or impurities.

The yield of tubulin from 1 kg of brain and yeast is ~ 150 and ~ 5 mg, respectively. The assembly of purified tubulin can be assayed by light scattering, X-ray scattering, centrifugation, and electron microscopy.

Isotypes and Posttranslational Modifications

Tubulin exists in different isotypic forms, the biological significance of which is still a matter of debate. The number of tubulin isotypes increases with the organism complexity. While yeast has only two α- and one β-isotypes, higher eukaryotes have up to seven β- and six α-tubulin isotypes. Certain isotypes have been found to be tissue specific, and differential expression of β-tubulin isotypes has been observed during the cell cycle. Some of these isotypes have been shown *in vitro* to have different relative stabilities, and such differences seem important for the response of the cell to anti-tubulin, anticancer drugs. The majority of differences between isotypes localize within the last 15 residues of the sequences, a region that has been identified as important in the interaction of microtubules with microtubule associated proteins (MAPs), pointing to a possible relevance for the functionality of microtubules in the cell.

Both tubulin subunits can be extensively altered by posttranslational modification, including detyrosination/tyrosination, acetylation/deacetylation, phosphorylation, polyglutamylation, and polyglycylation. All of these modifications, except for the acetylation of α-tubulin at Lys 40, occur at the divergent, highly charged C-terminal end of α- and β-tubulin. As for the different tubulin isotypes, the functionality of the posttranslational modifications is still a matter of debate. Certain modifications were initially identified as causing microtubule stability, and had been later described as the effect, not the cause of microtubule stability.

Structure

The structure of tubulin was solved by electron crystallography of zinc-induced two-dimensional tubulin sheets stabilized with the anticancer drug taxol. The α- and β-tubulin have basically the same secondary structure, each being made of a core of two β-sheets

surrounded by helices as shown in Figure 1. The N-terminal domain forms a Rossmann fold with the nucleotide-binding site at the C-terminal end of six parallel strands that are surrounded by five helices. An intermediate domain containing the binding site of taxol and colchicine is formed by a mixed β-sheet of four strands and five surrounding helices. The C-terminal domain contains two long helices that overlap the previous two domains and constitute the outside crest of the protofilaments in the microtubule to which motor molecules bind. The nucleotide sits at the interface between subunits along the protofilament. The N-site in

α-tubulin is buried within the dimer, while the E-site in β-tubulin is partially exposed in the dimer but occluded in microtubules. FtsZ, a ubiquitous protein in eubacteria and archebacteria essential for cytokinesis, is the only known structural homologue of tubulin.

Synthesis and Folding

Tubulin synthesis in cells is regulated by a process in which an increased subunit concentration leads to specific degradation of β-tubulin mRNA. In animal cells there is a mechanism to assure equivalent synthesis of α- and β-subunits. For its correct folding tubulin requires the citosolic chaperonin CCT (also referred to as TriC, TCP1, or Ct-cpn60), a hetero-oligomer formed by eight different subunits assembled into a hexadeca-mer of two double rings. Folding by CCT requires cycles of binding and full release, each cycle consuming one ATP by the chaperonin. In addition tubulin requires additional chaperonin cofactors that bind sequentially to α- and β-monomers and are necessary for the formation of the tubulin heterodimer. Folding cofactors are important also in regulating α/β tubulin ratios.

TUBULIN POLYMERIZATION

Microtubule Assembly and Structure

The tubulin sequence and structure contains the information required for its self-assembly into polar, dynamic microtubules, which in turn interact with a variety of cellular factors. Tubulin dimers bind head to tail making linear protofilaments, which associate in a parallel fashion giving rise to a polar microtubule. In a cell the so-called minus-end of microtubule, capped by α-subunits, is attached to the centrosome where γ-tubulin and related proteins nucleate microtubules. The more dynamic plus-end, capped by β-subunits, binds the kinetochore in mitosis.

The orientation of the tubulin subunits in the microtubule is such that the C-terminal helices form the crest of the protofilaments on the outside surface, making them an essential part of the binding site for motor proteins (kinesins and dyneins) and MAPs. The taxol-binding site is on the inside surface of the microtubule, and close to lateral interactions. The lateral contact between protofilaments is dominated by the interaction of the so-called M-loop in the second domain with loop H1–S2 and helix H3 in the N-terminal, nucleotide-binding domain.

The plus end of the microtubule is crowned by β-tubulin subunits exposing their nucleotide end to the solution, while the minus end is crowned by α-subunits exposing their catalytic end. When a dimer is added to a plus end, its catalytic end contacts the E-site nucleotide of the previous subunit forming the interface that should bring about hydrolysis.

FIGURE 1 Ribbon diagram for the structure of $\alpha\beta$-tubulin corresponding to a view from the inside of the microtubule with the plus end at the top. β-Tubulin (top) is bound to GDP while α-tubulin (bottom) is bound to GTP. The color scheme highlights the three domains in the structure of each monomer: green for N-terminal, nucleotide-binding domain (helix H7 or core helix is shown in lime); second or intermediate domain in blue, C-terminal domain in purple.

Dynamic Instability and GTP Cap Model

Microtubules are highly dynamic and can switch stochastically between growing and shrinking phases, both *in vivo* and *in vitro*. This nonequilibrium behavior, known as dynamic instability, is based on the binding and hydrolysis of GTP by tubulin subunits. Only dimers with GTP in their E-site can polymerize, but following polymerization this nucleotide is hydrolyzed by interactions with the previous tubulin dimer and then becomes nonexchangeable. In the GTP cap model the unstable body of the microtubule made of GDP–tubulin is stabilized by a layer of tubulin subunits at the ends that still retain their GTP. When this cap is stochastically lost, the microtubule rapidly depolymerizes. Depolymerization may occur by weakening of lateral contacts at the ends, and the consequent release of the constrained GDP subunits into a curved, lower energy, conformational state.

Microtubule assembly and stability are further modified in the cell by interaction with cellular factors that stabilize or destabilize microtubules at different points in the cell or at different stages in the cell cycle. Measurement of microtubule dynamics *in vitro* and *in vivo* by DIC or fluorescence microscopy yields the rate constants for addition of tubulin–GTP and loss of tubulin-GDP subunits at the two microtubule ends, as well as the rates of catastrophe (switch from growth to shrinkage) and rescue (switch from shrinkage to growth). A variety of drugs bind to tubulin and affect its polymerization. Microtubule-depolymerizing drugs, such as colchicine, nocodazole, vinca alkaloids, or podophyllotoxin, have a much higher affinity for the dimer than for tubulin in microtubules, so that disassembly is caused by their mass action effect. At substoichiometric concentrations tubulin-drug subunits bind to the microtubule ends and form caps that dramatically modify microtubule dynamics. Microtubule-stabilizing drugs such as taxoids, epothilones, or discodermolide act by binding preferentially to the polymerized form of tubulin.

An alternative microtubule behavior is treadmilling, a net flow of subunits from the plus to the minus end without a significant change in microtubule length.

Anti-Mitotic Tubulin Ligands

Recent years have seen the discovery of numerous tubulin ligands with anti-mitotic properties and anticancer potential. Concerning their binding site these agents can be classified into three main groups: those that bind tubulin at the colchicine binding site; those who bind it at the vinblastine site; and those who bind it at the taxol site. Functionally these antimitotic ligands can be separated into two classes: those that inhibit microtubule assembly (e.g., the colchicine and vinblastine families), and those that promote microtubule
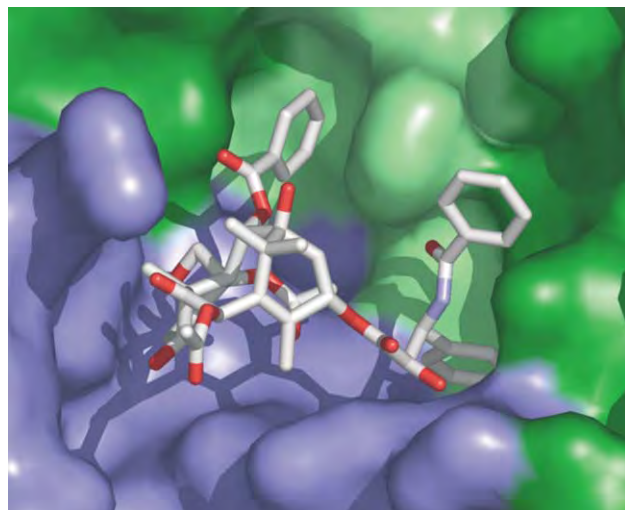
FIGURE 2 Anticancer drug taxol at its tubulin-binding site. Tubulin is shown on a surface representation with the same color scheme as in Figure 1. Both the N-terminal and second domain are part of the taxol-binding pocket.

assembly and stabilization (e.g., the taxol family). In spite of these differences, the main action of all these agents seems to be to cause mitotic arrest by inhibiting normal dynamic instability at very low concentrations.

The colchicine-binding site is located at the monomer–monomer interface within the dimer, in agreement with a model of colchicine action in which binding a distortion of the dimer structure that inhibits its polymerization. Vinblastine-like compounds are thought to bind at longitudinal polymerization contacts, resulting in a distorted protofilament structure. Finally, there is direct information on the binding site of taxol from the crystal structure of tubulin. Taxol binds to β-tubulin, near the M-loop and the site of lateral interactions, in a hydrophobic pocket that in α-tubulin is occupied by eight extra residues (see Figure 2). New microtubule-stabilizing agents with promise for cancer treatment, such as epothilones, discodermolide, eleutherobin, and the sarcodictins, while very different in structure, all seem to compete with taxol for the binding to a common site. It has been proposed that taxol stabilizes lateral contacts, or alternatively, that it acts as a bridge that holds the N-terminal and second domains in a relative orientation that favors longitudinal contacts between subunits.

γ-Tubulin and Microtubule Nucleation

Crucial to their dynamic behavior and function is the nucleation of microtubules in the cell at microtubule

organizing centers (MTOCs), to which they are attached by their minus ends. An essential role in microtubule nucleation is played by γ-tubulin, a protein with high homology to α/β-tubulins that localizes at MTOCs. γ-Tubulin forms ring structures that serve as templates for microtubule growth. The direct involvement of γ-tubulin in microtubule nucleation has been demonstrated *in vitro* using purified γ-tubulin-containing ring complex (γTuRC) containing at least seven different proteins. There are two main models in the literature. A model based on the shape and size of the γ-TuRC is that the ring forms the first helical turn of the growing microtubule, serving as a template for longitudinal interaction with the tubulin $\alpha\beta$ dimers. An alternative model, based on the similarity of rings structures formed by $\alpha\beta$-tubulin, γ-tubulin, and FtsZ, is that γ-tubulin forms a protofilament-like structure by longitudinal self-association that then serves as a template for lateral interaction with $\alpha\beta$-tubulin protofilaments.

RARER TUBULIN ISOFORMS: δ-, ε-, ζ-, AND η-TUBULINS

Four new tubulin isoforms have recently been discovered. The δ-tubulin was identified in *Chlamydomonas* mutants having abnormal basal bodies (these are microtubule structures at the base of cilia and flagella structurally similar to the centrioles in the centrosome at MTOCs). Human δ-tubulin was subsequently found in the human genome database, and shown to localize to the centrosome, where it partially colocalizes with γ-tubulin.

The ε-tubulin was identified from the human genome database on the basis of sequence similarity to other tubulins. Like δ-tubulin, ε-tubulin localizes to the centrosome, but in a cell-cycle-dependent manner: in cells with duplicated centrosomes ε-tubulin localizes only with the old centrosome.

Even rarer, ζ-tubulin has so far only been found on kinetoplastid protozoa where it localizes to the basal body, while η-tubulin has been found in paramecium where it may interact with γTuRC.

TUBULINS AND THE CELL

The involvement of the microtubule cytoskeleton in a large number of essential and diverse functions requires both reliability and flexibility from the system at the expense of biochemical and structural complexity. Dynamic instability is an inherent property of microtubules, built into the $\alpha\beta$-tubulin structure. The spatial and temporal organization of the microtubule network in the cell is obtained through the regulation of dynamic instability by an increasing number of factors that fine-tune the behavior of the microtubule system to accommodate the requirements of the cell.

Regulation may happen at many different stages, via transcription of different tubulin isotypes, the control of tubulin monomer folding, the formation of functional dimers, the posttranslational modification of tubulin subunits, the nucleation of microtubules, or the interaction of microtubules with numerous stabilizers and destabilizers. Tubulin-binding drugs can dramatically disrupt the finely tuned behavior of microtubules. Finally, while γ-tubulin is known to be essential for microtubule nucleation, additional tubulin isoforms, only recently discovered, have yet ill-defined functions.

SEE ALSO THE FOLLOWING ARTICLES

Centrosomes and Microtubule Nucleation • Chaperonins • Kinesins as Microtubule Disassembly Enzymes • Microtubule-Associated Proteins

GLOSSARY

$\alpha\beta$ tubulin dimer   Essential, highly conserved protein dimer present in all eukaryotes that self-assembles forming microtubules. It is the target of antimitotic drugs with anticancer potential.

dynamic instability   Nonequilibrium behavior of microtubules by which they can stochastically switch between phases of growth and shrinkage. It originates from the hydrolysis of GTP in β-tubulin and can be regulated by the interaction of tubulin/microtubules with cellular factors and antimitotic agents.

microtubules   Cytoskeletal polymers made of $\alpha\beta$-tubulin essential for cell transport and cell division. They are polar, dynamic, and regulated through the cell cycle by their interaction with stabilizers and depolymerizers.

γ-tubulin   Tubulin isoform most abundant at microtubule organizing centers where it is involved in microtubule nucleation. It forms higher-order complexes with associated proteins.

FURTHER READING

Desai, A., and Mitchison, T. J. (1997). Microtubule polymerization dynamics. *Ann. Rev. Dev. Biol.* **13**, 83–117.

Downing, K. H. (2000). Structural basis for the interaction of tubulin with proteins and drugs that affect microtubule dynamics. *Ann. Rev. Cell Dev. Biol.* **16**, 89–111.

Lewis, S. A., Tian, G. and Cowan, N. J. (1997). The α- and β-tubulin folding pathways. *Trend. Cell Biol.* **7**, 479–484.

Lowe, J., Li, H., Downing, K. H., and Nogales, E. (2001). Refined structure of alpha beta-tubulin at 3.5 A resolution. *J. Mol. Biol.* **313**, 1045–1057.

Mitchison, T., and Krischner, M. (1984). Dynamic instability of microtubule growth. *Nature* **312**, 237–242.

Nogales, E. (2000). Structural insights into microtubule function. *Ann. Rev. Biochem.* **69**, 277–302.

Nogales, E., Wolf, S. G., and Downing, K. H. (1998). Structure of the $\alpha\beta$ tubulin dimer by electron crystallography. *Nature* **391**, 199–203.

Nogales, E., Whittaker, M., Milligan, R. A., and Downing, K. H. (1999). High resolution structure of the microtubule. *Cell* **96**, 79–88.

BIOGRAPHY

Eva Nogales is an Assistant Professor in the Department of Molecular and Cell Biology at UC Berkeley, an Assistant Investigator at the Howard Hughes Medical Institute, and a Staff Scientist at Lawrence Berkeley National Laboratory. She was trained as a physicist in Spain, her country of origin, and obtained a Ph.D. in biophysics from Keele University for her work at the Synchrotron Radiation Source at Daresbury, UK. During her postdoctoral studies with Dr. Kenneth H. Downing at Lawrence Berkeley National Laboratory she obtained the structure of $\alpha\beta$-tubulin using electron crystallography. Her present research centers around the structural bases of microtubule dynamics and the structural characterization of protein complexes involved in eukaryotic transcription.

Tumor Necrosis Factor Receptors

Karen G. Potter and Carl F. Ware
La Jolla Institute for Allergy and Immunology, San Diego, California, USA

Tumor necrosis factor receptors (TNFRs) are a family of structurally similar cytokine receptors that act as transducers of cell death and induce the expression of genes involved in cellular differentiation and survival. Binding of specific ligands to their cognate TNFR initiates the recruitment of adaptor proteins, either death domain (DD)-containing or TNFR-associated factor (TRAF) family of adaptor proteins, to the cytosolic signaling domain of the receptor to initiate diverse effector functions. The most well-known function of the TNF superfamily is in immune regulation and development with specific roles in host defense, inflammation, cellular homeostasis, and lymphoid organogenesis. A critical role for TNFR in immunobiology is evidenced by the linkage of naturally occurring mutations in TNF family genes to human disease, as well as by the targeting of TNF family members by viruses as a mechanism of immune evasion. However, some TNF family members also act outside the immune system by regulating the development of hair follicles, sensory neurons, or bone-resorbing osteoclasts.

Features of Tumor Necrosis Factor Receptors

STRUCTURE

Tumor necrosis factor receptors (TNFRs) are identified by a highly conserved, cysteine-rich domain (CRD) in the extracellular portion of the protein that binds ligand (Figure 1). The CRD generally contains six cysteines that form three disulfide bonds typically recognized by the signature sequence motif CxxCxxC. Currently, there are 29 members of the cellular TNFR family in mammals, and several variants of TNFRs are found in herpesviruses and poxviruses. The cellular receptors are primarily type I transmembrane proteins (extracellular N terminus). Some receptors in this family lack transmembrane and cytoplasmic domains and are secreted, functioning as decoy receptors for the ligand. The TNFR family can be divided into two general groups – those that contain a death domain (DD) and those with a peptide motif that binds TNFR-associated factors (TRAF) adaptors – based on the structure of their cytoplasmic tails and the signaling adaptors they recruit to propagate signals to the cells.

Given the predominant role of the TNFR family in regulating immunity, this suggests that the evolution of the receptors in this family arose coincident with the evolution of adaptive immunity, also found exclusively in vertebrates. The size of the TNFR superfamily appears to have grown in a large part by gene duplication as many of the TNFR genes are linked to discrete loci reflecting their evolutionary derivation. Perhaps most obvious are those TNFRs found on chromosome (Ch) 12p13 (TNFR1, LTβR, and CD27), which likely underwent duplication and translocation events giving rise to the larger locus of TNFR on Ch 1p36 (TNFR2, HVEM, Ox40, CD30, AITR, 4-1BB, and DR3) (Figure 2A). Strikingly, NGFR is the only receptor that binds ligands structurally unrelated to TNF, representing a clear functional demarcation from the typical TNFR (Figure 2B). Another subtle branch in the TNFR family tree are those receptors that engage BAFF, the B cell survival factor, and related ligands APRIL and TWEAK (Figure 2C). These receptors have a single CRD and in the case of BAFFR only two disulfide bonds. Functional divergence is evident in the role some TNFRs play in bone (Osteoprotegerin-RANK-RANKL/TRANCE) and ectodermal (Ectodermal dysplasin EDA-EDAR), and angiogenesis (TWEAK-Fn14).

EXPRESSION

Expression patterns of the receptors are complex. Several members of the TNFR superfamily are expressed on cells of the immune system; for example, BAFFR expression is exclusive to B lymphocytes. Other receptors are found in hair follicles (e.g., EDAR and TROY) or the nervous system (e.g., NGFR), yet others have very broad tissue expression patterns, such as TNFR1. The expression of some TNFR is inducible and regulated and others are constitutive. For example, TNFR2 is absent on naïve T cells but is rapidly upregulated upon activation of T cells. In contrast, HVEM is highly expressed on naïve T cells, but shows diminished expression upon T-cell activation. Regulation of receptor expression can also be achieved through the
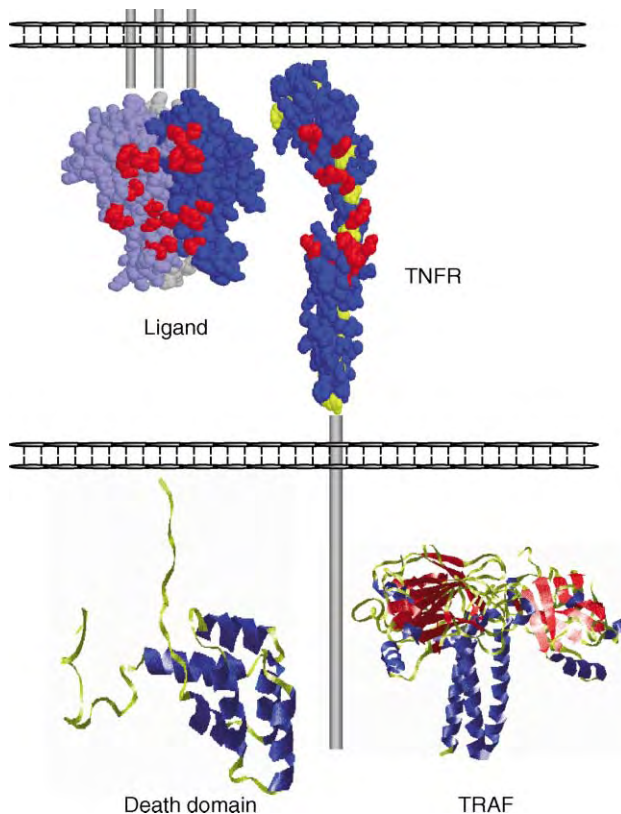
FIGURE 1   Structural features of TNFR family. Space filling model of a trimeric TNFR1 ligand, LTα, as it would exist as a membrane anchored ligand, and a single TNFR1 rotated 180° to reveal contact residues. Also shown are crystal structures of the cytoplasmic DD found in some TNFR and the TRAF that interacts with TNFR-containing TRAF binding sites.

generation of soluble receptors; soluble TNFRs can be generated by proteolytic processing (CD27, CD30, CD40, TNFR1, and TNFR2), alternative splicing of membrane forms (Fas, 4-1BB), or encoded in the genome without a transmembrane region (OPG, DcR3, TRAIL-R3).

LIGAND BINDING

TNF ligands are characterized as type II (intracellular N terminus) transmembrane proteins containing a "TNF homology domain" (THD) in their C terminus (Figure 1). The THDs of each ligand associate to form trimeric proteins; in most cases, the ligands form homotrimers, but LTα and LTβ form heterotrimers, and so do BAFF and APRIL, as has been shown recently. The TNF ligands are active in their membrane forms where cell-to-cell contact is required to initiate signaling; however, some ligands can be proteolytically cleaved from the surface creating soluble cytokines that affect cells distally.

The TNF ligand family currently includes 20 proteins that are able to pair off with one or more receptors.

In fact, some ligands display overlapping receptor recognition; for example, LIGHT and LTα1β2 bind LTβR, and LTα and LIGHT bind HVEM. The ligand-receptor pairing seems promiscuous, although comparative studies using mice deficient in a specific TNF ligand or receptor suggest that each cytokine-receptor system has a unique role in immune physiology.

TNFR Signaling

Each TNF-related ligand has three receptor binding sites that can cluster together two or three cell surface receptors, juxtapositioning the cytoplasmic tails of the receptors to initiate signal transduction. Recruitment of specialized signaling molecules (adaptors) to the cytoplasmic domain occurs following receptor clustering. Propagation of TNFR signals occurs through two distinct classes of cytoplasmic adaptor proteins: TRAFs or DD molecules.

TRAF SIGNALING

TRAF adaptor proteins are a small family of RING finger proteins that play a critical role in propagating signal transduction leading to the activation of latent transcription factors (Figure 1). There are six numerically named TRAF proteins (TRAF1–6) that function predominantly in TNFR-induced signaling, although TRAF6 is also a key player in signal transduction initiated by the interleukin-1 receptor and the Toll-like receptor (TLR) superfamily. Several TRAFs can bind directly to the cytoplasmic tails of most of the TNFR. The binding site found in CD40 or the LTβR is a short peptide sequence, PXQXT/S or IPEEGD, respectively. The binding site in TRAF for receptor binding is flexible, thus accommodating a variety of motifs. TNFRs with a DD bind TRAFs indirectly via other adaptor proteins, such as TRADD (TNFR-associated DD). Each TRAF interacts with several different receptors such that TNFRs display distinct interaction patterns with multiple TRAFs. Since each TRAF is believed to have distinct biological effects, the variation in TRAF binding by the TNFR may be able to direct the signaling pathway to distinct biological outcomes.

Functionally, binding of TRAFs to TNFR culminates in the activation of transcription factors that act to regulate gene expression. For example, nuclear factor κB (NFκB) is a small family of latent transcription factors that induces the expression of a large variety of genes involved in inflammatory and immune responses (Figure 3). Two distinct forms of NFκB are recognized: NFκB1 (RelA, p65) and NFκB2 (p100/p52), and each pair with itself or other proteins (e.g., p50, RelB, and cRel) that form active transcription factors that bind
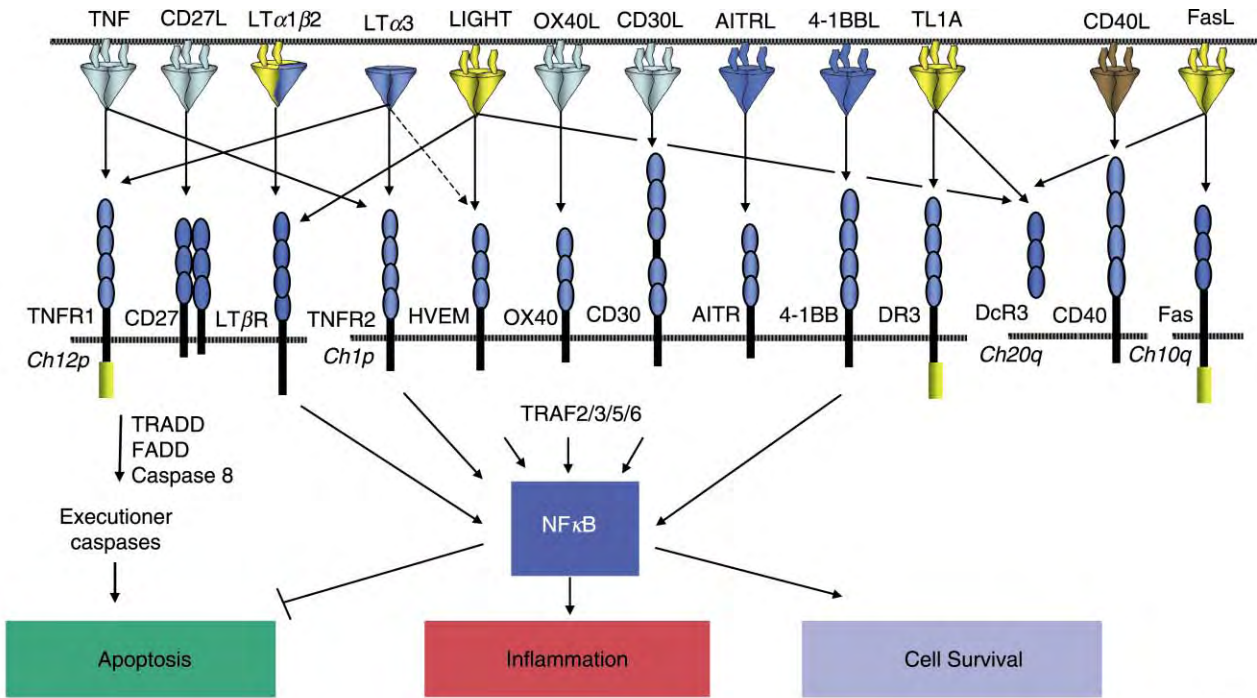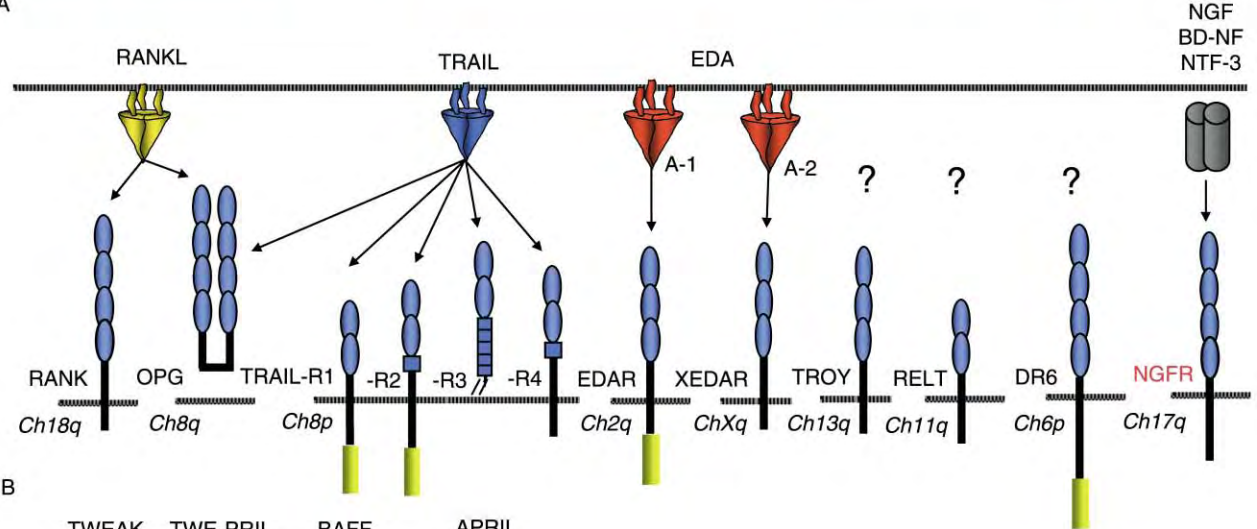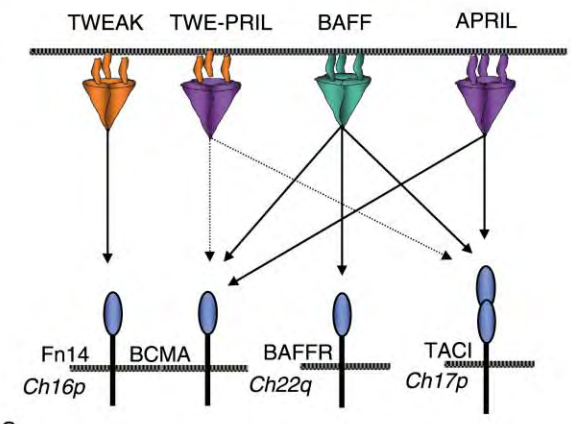
FIGURE 2   The TNF–TNFR superfamily grouped on the basis of chromosome (Ch) localization of TNFR. (A) TNFR and TNF ligands are shown with arrows connecting ligand–receptor pairs. CRDs are shown as small ovals. DD is denoted as a rectangular box in the cytoplasmic tail of appropriate receptors. The signaling and effector functions induced by either DD-containing or TRAF-binding TNFR is shown. (B and C) Same as (A), with remaining TNFR grouped on the basis of chromosome localization and ligand binding.
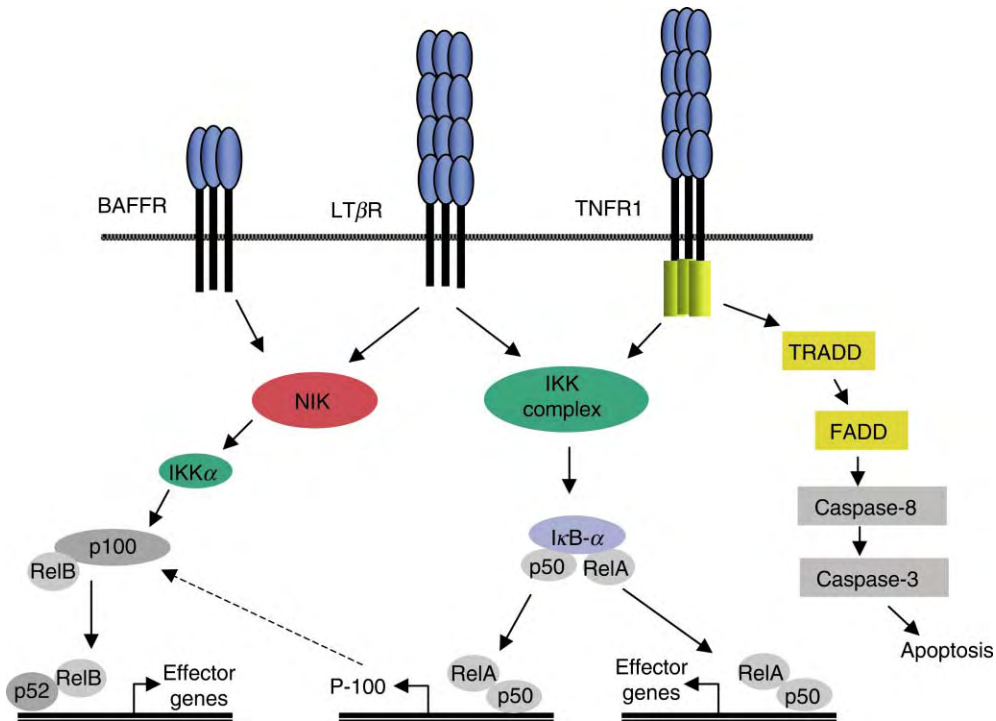
FIGURE 3   Mechanisms of signaling by TNFR. Upon ligation of LTβR, two NFκb pathways are induced. The first leads to induction of the NFκB1 pathway and the activation of IKKβ and RelA, which control expression of inflammatory genes. The second pathway results in the activation of NFκB2 and the processing of p100 to p52 following the activation of NIK and IKKα, leading to the transcription of genes implicated in secondary lymphoid organogenesis and homeostasis. BAFFR also activates the NFκB2 pathway, whereas TNFR1 only activates the NFκB1 pathway. TNFR1 also contains a DD to signal caspase activation and apoptosis.

DNA. NFκB1 is held in the cytosol by an inhibitor protein (IκB) while an inhibitory domain controls the cytosolic localization of NFκB2. The inhibitory proteins are phosphorylated and degraded in response to activation signals coming from diverse sources that all target the inhibitor κB protein kinase complex (IKK). Degradation of the inhibitor allows NFκB to localize to the nucleus and activate transcription. The NFκB1 and B2 pathways induce distinct sets of genes: NFκB1 regulates the expression of proinflammatory chemokines and adhesion molecules, while the NFκB2 pathway controls the expression of distinct chemokines and cytokines involved in lymphoid organogenesis. BAFFR is only able to activate the NFκB2 pathway, suggesting that this pathway is crucial for the expression of survival genes specifically important for B lymphocytes. By contrast, TNFR1 is unable to activate NFκB2, which may account for its strong proinflammatory action.

TRAF2 propagates signals to mitogen-activated protein (MAP) kinases, including JNKs/SAPKs, ERKs, and p38s, to activate the transcription factor AP-1 that plays roles in stress responses and cellular homeostasis. Like TRAF2, TRAF5 cooperates to activate NFκB and AP-1 transcription factors, while TRAF3 negatively regulates NFκB activation and thus may play a role

in promoting cell death. TRAF6 also activates NFκB and AP1.

DEATH RECEPTOR SIGNALING

A number of TNFR, including TNFR1 and Fas, are also termed "death" receptors because they regulate apoptotic cell death. Their cytoplasmic tails contain a region of ~ 80 amino acids that fold into six α-helices, termed the DD. The DD serves as a protein interaction motif to recruit signaling molecules to the inert cytoplasmic domain of TNFRs. As such, the DD of TNFR can self-associate or associate with other DD-containing adaptor proteins such as FADD, TRADD, and RIP.

In the simplest scheme known to activate the cell death machinery, the adaptor FADD is recruited to Fas initiating formation of the death-inducing signaling complex (DISC). Procaspase 8 is recruited to the DISC through a second interaction motif contained in FADD, termed the death effector domain (DED), activating downstream effector caspases (e.g., caspase 3) resulting in the cleavage of critical cellular substrates and the eventual collapse of the cells and death (apoptosis). Similarly, TNFR1 activates apoptosis although it requires the adaptor protein TRADD to facilitate

FADD recruitment, and is also able to recruit another adaptor protein RIP that also plays a role in procaspase 8 activation. The DISC is regulated by two forms of another protein known as FLIP, which can switch the DISC between activating caspases or NFκB, which in turn regulates genes that block apoptosis, such as the inhibitors of apoptosis (IAP) that promote cell survival.

Effector Functions of the TNFR Family

Despite similarities in structure and signaling mechanisms, TNFRs function in diverse, and often opposing, roles in immune physiology. As discussed above, several reasons that explain their functional diversity include differences in ligand binding, tissue and cellular expression, regulation of expression, and association with signaling adaptor proteins.

INFLAMMATION

A role for TNF family members in host defense and inflammation was first appreciated when it was realized that TNF is identical to a factor termed cachectin, a protein known to cause fever and wasting. Inflammatory cells, such as macrophages, induce TNF when specialized innate immune receptors called TLRs recognize nonself patterns on microbial pathogens. Binding of TNF to its receptors, TNFR1 or TNFR2, induces a variety of molecules that are crucial for initiating the acute inflammatory response. For example, TNF signaling induces increased expression of adhesion molecules on endothelial cells and secretion of chemokines to promote inflammatory cell migration to the site of infection. TNF can also activate macrophages and neutrophils to increase their phagocytic capacity and their ability to secrete tissue-degrading enzymes into infected tissues. Although inflammation is crucial to prevent infection, it must also be minimized after the pathogenic challenge subsides in order to prevent uncontrolled damage to fragile host tissues. As such, uncontrolled TNF signaling can cause chronic inflammation and wasting, or acutely released, septic shock. Drugs that block TNF signaling, such as soluble TNF decoy receptors, have been developed to prevent inflammatory diseases such as rheumatoid arthritis and inflammatory bowel disease.

APOPTOSIS

Cell death by apoptosis is one of the major functions of DD-containing TNFR. Fas, for example, mediates cell death for varied purposes. First, Fas, and its ligand FasL, cooperate as major effectors, along with the perforin/granzyme pathway, to trigger killing of pathogen-infected cells by cytotoxic T-lymphoctyes (CTLs) or by natural killer (NK) cells. In a second context, Fas is required to eliminate excessive CTLs and other effector cells. Fas is induced on long-term antigen-activated lymphocytes to elicit their removal and minimize the excessive expansion of T lymphocytes that can occur in response to T-cell activation in the space-restricted lymphoid environment. In fact, two naturally occurring mutations in mice encoding the gene *lymphoproliferative disorder (lpr), lpr,* or *lpr^{cg}* that affect the expression or the biological activity of the Fas receptor, respectively, demonstrate the important role of Fas in lymphocyte homeostasis. Both of these mutant mice develop an autoimmune phenotype characterized by massive lymphadenopathy, splenomegaly, and autoantibody production.

CELL SURVIVAL

BAFFR, BCMA, and TACI are some of the most recently discovered members of the TNFR family that are predominantly expressed on B cells and represent a good example of TNFRs that play a role in cell survival. All three receptors bind the TNF-related ligand BAFF, produced by macrophages and dendritic cells, while BCMA and TACI can also bind the related ligand APRIL. BAFF is required in the maintenance and differentiation of mature B cells in peripheral tissues by supporting B-cell survival and maturation presumably through activation of anti-apoptotic genes. The mechanism for enhanced B-cell survival by BAFF may be partly due to the stimulation of the pro-survival protein bcl-2 and a decrease in the expression of the proapoptotic factors Bak and Blk. A crucial role for BAFFR in BAFF-mediated B-cell survival is evidenced by the fact that neither BCMA- nor TACI-deficient mice have defects in B-cell maturation, whereas A/WySnJ mice, containing a naturally occurring mutation in BAFFR, do show defects in B-cell maturation. Overexpression of BAFF in mice can result in autoimmune symptoms, and similarly patients with systemic lupus erythematosus (SLE), rheumatoid arthritis, and Sjogren's syndrome show elevated levels of BAFF in the blood. Other TNFRs also play roles in the survival and development of various tissues including the TRANCE system in bone, EDAR in ectodermal tissues, and NGFR in neurons.

COSTIMULATION

T lymphocyte activation is also regulated by TNFRs where they act as costimulatory signals to positively or negatively influence the proliferation of antigen-stimulated T cells. OX40, CD27, 4-1BB, and HVEM, expressed on CD4$^+$ and CD8$^+$ T cells, promote T-cell

proliferation and cytokine production in response to their respective ligands expressed on dendritic cells. For example, mice deficient in the OX40–OX40L interaction show impaired T-cell proliferative capabilities and cytokine production in response to viruses such as lymphocytic choriomengiitis virus (LCMV) or influenza virus, and the ability to generate or sustain memory T cells is diminished. Additionally, some evidence indicates that several receptors such as CD30, Fas, or HVEM may play a role in positive and negative selection in thymocyte differentiation. Thus, TNFR family members can contribute to immune tolerance.

On B cells, CD40 is the most prominent costimulatory receptor where it interacts with its ligand, CD40L, on T cells to initiate T-cell-dependent B-cell antibody responses. Mutations in CD40 or its ligand affect a number of immune functions including immunoglobulin class switching; in humans, X-linked hyper-IgM syndrome is due to defects in CD40L structure and function manifesting as an inability to switch immunoglobulin M to IgG classes.

Organogenesis and Development

A number of TNFRs play roles in the organization of tissues. The lymphotoxin system mediated by LTβR, and its ligand LTαβ, play a critical role in lymph node development; mice deficient in either receptor or ligand completely lack lymph nodes and Peyer's patches. RANK and its ligand RANKL mediate osteoclast differentiation and therefore are important for maintaining normal bone homeostasis. Osteoporosis, a condition of bone thinning, can be caused by overexpression of the soluble RANKL decoy receptor, OPG, which is induced by estrogen and prevents normal bone formation. RANK-RANKL signaling also supports the development of mammary gland alveolar buds important for lactation, and consequently an absence of RANK induces accelerated apoptosis of mammary epithelial precursors. EDAR, XEDAR, and TROY, and the ligand EDA, regulate the development of ectodermal tissues including hair follicles. In fact, mice or humans with deficiencies in EDAR or EDA lack primary hair follicles and sweat glands. Finally, NGFR is important for the development of sensory neurons; acting alone it induces apoptosis, but upon association with other nerve growth factor receptors it mediates the differentiation and survival of neurons. NGFR-deficient mice develop cutaneous sensorineuronal defects.

Viral Targeting of TNFR

Since TNFRs and their ligands play a critical role in mediating immune defenses against invading pathogens,

and control death and survival fates for cells, it is not surprising that viruses have evolved mechanisms that directly target the TNF family in order to subvert immune effector mechanisms. In fact, viruses when viewed collectively have targeted virtually every step of the TNFR signaling pathway, from ligand binding to activation of apoptosis. For example, infection by some viruses, including cytomegalovirus (CMV), herpes simplex virus (HSV), adenovirus, and HIV, induce the expression of FasL and/or TRAIL on the host cell resulting in the probable killing of recruited immune effector cells that express the cognate TNFR death receptor. In contrast, adenovirus down-regulates the proapoptotic receptors Fas and TRAILR from the surface of infected cells, thereby preventing the killing of the host cell harboring the virus. Another mechanism of immune evasion employed by poxviruses is the expression of soluble, secreted orthologs of TNFR2 or CD30 that inhibit the interaction of their respective ligands with their cellular receptors. In some cases, viruses and TNFR signaling have been adapted to function cooperatively. For instance, signaling through the LTβR or TNFR1 can prevent the replication of human CMV in fibroblasts without inducing death of the infected cell thereby allowing continued existence of CMV within the host. This molecular example of host–virus coexistence might provide a mechanism for CMV to persist latently in individuals with a strong host immune system, and re-emerge to cause disease in individuals who are immunosuppressed.

See Also the Following Articles

Bax and Bcl2 Cell Death Enhancers and Inhibitors • Cell Death by Apoptosis and Necrosis • Chemokine Receptors • Mitogen-Activated Protein Kinase Family • Nuclear Factor kappaB • Toll-Like Receptors

Glossary

apoptosis An orderly process of programmed cell death whereby cellular machinery undergoes changes leading to death.
autoimmunity An immune response against self-antigens.
cytokine Proteins made by cells that affect the behavior of other cells.
signal transduction Term used to describe the processes inside cells used to respond to changes in their environment.
transcription factor A protein that binds DNA to control the transcription of genes involved in a large number of normal cellular and organismal processes.

Further Reading

Benedict, C. A., Banks, T. A., and Ware, C. F. (2003). Death and survival: Viral regulation of TNF signaling pathways. *Curr. Opin. Immunol.* **15**, 59–65.

Bodmer, J., Schneider, P., and Tschopp, J. (2002). The molecular architecture of the TNF superfamily. *Trends Biochem. Sci.* **27**, 19–26.

Chung, J. Y., Park, Y. C., Ye, H., and Wu, H. (2002). All TRAFs are not created equal: Common and distinct molecular mechanisms of TRAF-mediated signal transduction. *J. Cell Sci.* **115**, 679–688.

Dejardin, E., Droin, N. M., Dehase, M., Haas, E., Cao, Y., Makris, C., Li, Z. W., Karin, M., Ware, C. F., and Green, D. R. (2002). The lymphotoxin-β receptor induces different patterns of gene expression via two NFκB pathways. *Immunity* **17**, 1–11.

Locksley, R. M., Killeen, N., and Lenardo, M. J. (2001). The TNF and TNF receptor superfamilies: Integrating mammalian biology. *Cell* **104**, 487–501.

Orlinick, J. R., and Chao, M. V. (1998). TNF-related ligands and their receptors. *Cell. Signal.* **10**, 543–551.

BIOGRAPHY

Karen G. Potter is an NIH Postdoctoral Fellow in the lab of Dr. Carl F. Ware at the La Jolla Institute for Allergy and Immunology. She received her Ph.D. in Cell and Molecular Biology from Duke University in July 2002. Her research interests include the regulation and immunomodulatory functions of the TNFR HVEM.

Carl F. Ware is Head of Molecular Immunology at the La Jolla Institute for Allergy and Immunology, and a Professor of Biology at UCSD. His research program investigates the regulation of immunity by cytokines and viruses. He is an advisor for the National Institutes of Health and is an editor of scientific journals. Dr. Ware received his Ph.D. in Molecular Biology and Biochemistry from the University of California Irvine, and was an NIH Postdoctoral Fellow at the Harvard Medical School.

Two-Dimensional Gel Electrophoresis

Gerhard Schmid, Denis Hochstrasser and Jean-Charles Sanchez
Biomedical Proteomics Research Group, Geneva University Hospital, Geneva, Switzerland

Two-dimensional polyacrylamide gel electrophoresis (2D PAGE) is a high-resolution protein separation technique exploiting two independent physico-chemical properties (charge and size) of the protein components to be separated. During the first dimension, under the influence of an electrical field, charged proteins migrate in a pH gradient until each of them reaches a distinct pH value, which corresponds to its isoelectric point (pI). At its pI, a protein has zero net charge. This separation technique is called isoelectric focusing (IEF). In the second dimension, the protein mixture previously separated by IEF is sorted orthogonally by electrophoresis in the presence of sodium dodecyl sulfate (SDS). Migration of SDS-coated proteins in a sieving polyacrylamide gel results in a separation according to their relative molecular mass (M_r). Nowadays, 2D PAGE is a powerful and widely used tool for the analysis of complex soluble protein mixtures extracted from cells, tissues, or other biological samples.

Some Milestones in the Development of 2D PAGE

HOW IT ALL BEGAN

In 1975, three papers (Klose, O'Farrell, Scheele) introduced two-dimensional polyacrylamide gel electrophoresis (2D PAGE) using denaturing isoelectric focusing (IEF) and SDS-PAGE. At that time, IEF was carried out in cylindrical rod IEF gels cast in glass capillary tubes (1–1.5 mm inner diameter). The gels contained synthetic carrier ampholytes, which formed a continuous pH gradient under the influence of an electric current. After IEF the gel rods had to be removed from their tubes for second dimension separation by SDS-PAGE.

TECHNICAL IMPROVEMENTS

With the introduction of immobilized pH gradients (IPGs) by Bjellqvist *et al.* in 1982, problems associated with carrier ampholyte pH gradients such as gradient instability, low sample loading capacity, and difficulty to achieve reproducibility could be eliminated. In an IPG, a set of acidic and basic buffering groups is covalently incorporated into a polyacrylamide gel at the time it is cast. Precast IPG gel strips supported by a plastic film backing are commercially available in a variety of narrow and broad, linear and nonlinear pH ranges. This represents an important development and allowed 2D PAGE to become the tool of choice for high-resolution protein separation (Figure 1). Besides improvements in the 2D technique, critical developments in other fields greatly contributed to the more widespread use of 2D PAGE. In 1987, Matsudaira showed that classical N-terminal sequencing by Edman degradation can be applied to picomole quantities of proteins electro-transferred onto polyvinylidene difluoride (PVDF) membranes. In 1993, several groups independently showed that mass spectrometry allows the identification of proteins previously separated by 1D or 2D gel electrophoresis. Improvements on the level of the mass spectrometers as well as rapid access to genome sequence databases for various species from across the world nowadays allow the high-throughput identification of hundreds of proteins separated by 2D PAGE. More powerful, less expensive computers and appropriate software made computer-assisted evaluation of the highly complex 2D patterns possible.

2D PAGE IN THE FUTURE

After more than 25 years, one can say that 2D PAGE finally became a routine protein separation method; however, its further development is a continuous process. A major interest, especially for the application of 2D PAGE in the clinical laboratory, lies in the field of miniaturized and automated 2D protein mapping. Recently, a fully automated 2D electrophoresis system (a2DE^{TM}) has been presented by NextGen Sciences. A microchip format that is based on 2D PAGE (digital ProteomeChip^{TM} from Protein Forest Inc.) allows protein separation, staining, and digital imaging in as little as 20 min, with a detection sensitivity below 1 pg of protein.
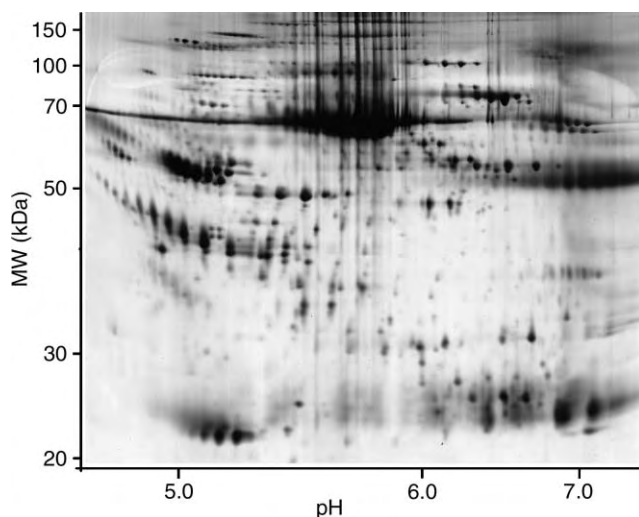
FIGURE 1 High-resolution protein separation of human plasma by 2D PAGE, silver staining.

The Current Protocol

The method described by O'Farrell back in 1975 already contained the magic bullets of 2D PAGE, i.e., urea, dithiothreitol (DTT), a nonionic or zwitterionic detergent for the first and SDS for the second dimension. The procedure of 2D PAGE using IPG gel strips in the first dimension is largely based on the work of Görg and co-workers.

SAMPLE SOLUBILIZATION

An ideal procedure results in the complete solubilization, disaggregation, denaturation, and reduction of all the proteins in the sample. However, dealing with several thousand different proteins with different physico-chemical properties makes complete solubilization practically impossible. In general, cells or tissues are disrupted by various techniques such as grinding in a liquid nitrogen cooled mortar, sonication, shearing-based methods, or homogenization. Proteins are then solubilized in a buffer containing 9 M urea, 4% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 50–100 mM DTT. Various minor modifications from this basic, IEF compatible solubilization buffer have been described over the years, trying to improve the solubilization of certain subclasses of proteins. DNA or RNA molecules, which can interfere with 2D PAGE, are removed by incubation with nucleases. Centrifugation is used to separate the solubilized proteins from nonsolubilized material. Prefractionation of complex samples is a good choice when only a subset of the proteins in a tissue or a cell is of interest or when abundant proteins dominate the sample as in the case of albumin in plasma.

FIRST DIMENSION: IEF

IEF separates proteins according to their isoelectric point. Proteins diffusing away from their pI immediately gain charge and migrate back. This focusing effect allows proteins to be separated on the basis of very small charge differences. Simplified handling and improved performance made the commercially available precast IPG gel strips (from 7 to 24 cm in length) the method of choice for most applications. IPG strips with a pH range from 3 to 10 will display a wide range of proteins, whereas narrower pH ranges allow to zoom into a particular pH range. With a set of narrow range, overlapping pH gradients, a wide pH range can be covered at increased resolution. There are several ways to load the sample on the IPG gel strip. Efficient entry of the solubilized proteins, especially hydrophobic ones, into the gel is always a critical point. IPG dry strips need to be rehydrated before use in IEF. Traditionally, samples are applied on a discrete point via sample cups (Figure 2A). Cups are placed at either the acidic or basic pH extreme of the strip. Under these conditions, most of the proteins in the sample will be charged and immediately start migrating under the influence of an electric current, thereby minimizing sample loss through precipitation. Alternatively, the sample can be applied over the whole IPG gel surface during strip rehydration. This method is of special interest for more dilute samples since larger volumes can be loaded. For strip rehydration, a solution containing 8 M urea, 2% CHAPS, 10 mM DTT, 1% carrier ampholytes, traces of Bromophenol Blue (tracking dye) shows excellent results for most samples. As for the solubilization solution, various modifications were described. A general IEF protocol starts with the sample entry phase at low initial voltage (200 V). Voltage is then increased in a stepwise manner up to the desired final focusing voltage (up to 8000 V). Optimal focusing time depends on the nature of the sample. In general, overfocusing is not deleterious below a total of 100 kVh. State-of-the-art IEF systems allow simultaneous migration of up to 12 strips under temperature-controlled conditions.

SECOND DIMENSION: SDS-PAGE

Prior to the second dimension, the IPG gel strips are equilibrated in a solution containing SDS. This anionic detergent binds to the majority of proteins in a constant mass ratio (1.4 g SDS/g protein), such that the net charge per mass unit becomes approximately constant. This allows protein separation in a polyacrylamide gel according to their relative molecular mass. For an ideal transfer of the proteins separated in the IPG strip onto the second dimension gel, the equilibration step should result in the complete resolubilization of all proteins in the IPG gel. Again, the enormous chemical diversity
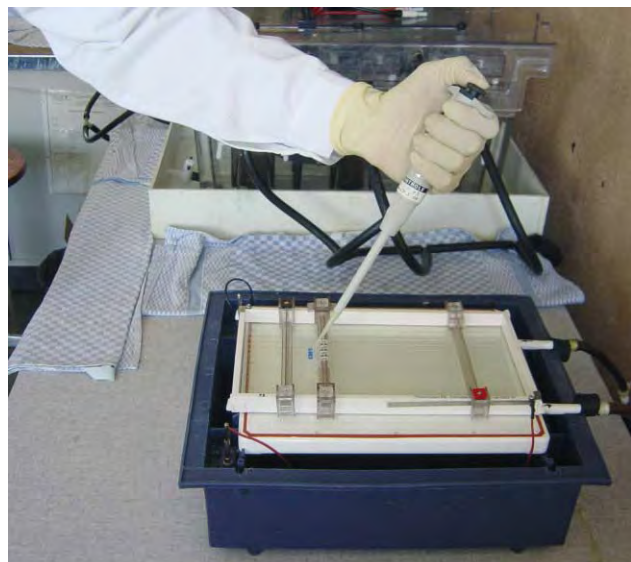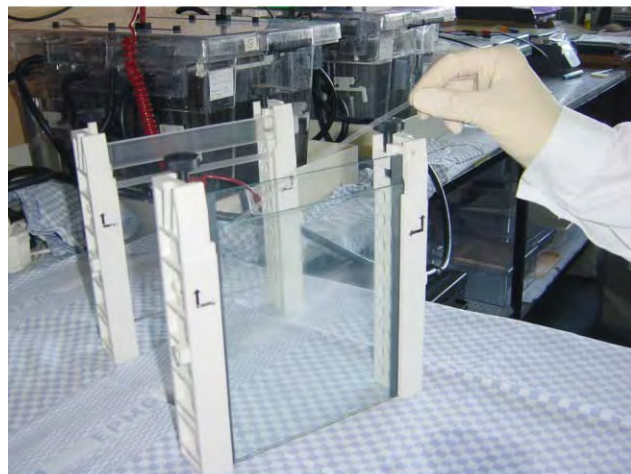
FIGURE 2 (A) Sample application for the first dimension of 2D PAGE – samples are loaded in cups at the cathodic end of the IPG strip. (B) Strip transfer to the second dimension of 2D PAGE – an equilibrated IPG strip is placed on the top of a large format polyacrylamide gel.

present within most samples makes it impossible to use ideal conditions for all its components. Composition of the equilibration solution and especially duration of the equilibration steps represent a compromise between accessing poorly soluble proteins with a tendency to adsorb to the gel matrix and diluting highly soluble proteins out of the gel strip. Equilibration in a buffer containing 6 M urea, 50 mM Tris–HCl buffer (pH 8.4), 2% w/v SDS, 30% w/v glycerol proved to be a good compromise. For protein reduction, IPG strips are first incubated in equilibration buffer containing 2% w/v DTT for 12 min. In a second step, proteins are alkylated by incubation in equilibration buffer containing 2.5% w/v iodoacetamide (IAA) and traces of bromophenol blue (tracking dye) for 5 min. In vertical SDS-PAGE, the equilibrated strip is placed on the surface of the second dimension gel and embedded in molten agarose (Figure 2B). Most commonly, the tris-glycine buffer system described by Laemmli is used. State-of-the-art 2D gel electrophoresis systems allow the simultaneous migration of up to 24 large format gels (20 by 24 cm) under temperature-controlled conditions.

VISUALIZATION

Most of the staining methods used for conventional 1D SDS-PAGE can be applied to 2D gels. Because of its very high sensitivity (down to 1 ng/spot), silver staining is probably most commonly used. However, the use of glutaraldehyde in the silver-staining procedure should be avoided for subsequent protein identification by mass spectrometry. Recently, MS-compatible, fluorescent stains with sensitivities comparable to silver became available. At the time being, application of this stains in large-scale and high-throughput 2D PAGE is hardly affordable. In addition, special imaging systems are required, whereas a simple desktop scanner can be sufficient for "visible" stains. This still leaves plenty of room for other MS-compatible, but less sensitive stains such as coomassie blue and zinc negative staining. Proteins labeled *in vivo* by the use of radioactive amino acids (metabolic labeling) can be visualized by exposing the dried 2D gel to an X-ray film or a storage phosphor screen for use with a PhosphorImager system. As for 1D SDS-PAGE, proteins from 2D gels can be electro-transferred on membranes, which are subsequently probed with antibodies. This so-called Western blot technique allows highly sensitive and specific visualization of a protein of interest.

Applications of 2D PAGE

The combination of a high-resolution protein separation tool (2D PAGE) with a highly sensitive protein identification tool (MS) offers an enormous potential. In 1994, on the occasion of the first Siena 2D Electrophoresis meeting, the expression "proteome" was proposed by Marc Wilkins for the description of the proteins expressed by a genome of a species, an organ, or a cell at a particular moment under particular conditions. Since then, the field of proteomics, i.e., the analysis of proteomes, is booming. The global approach taken with proteomics allows the massively parallel analysis of expressed proteins, including their isoforms originating from posttranslational modifications (PTMs). Global gene expression can also be studied at the level of mRNA. However, it needs to be emphasized that there is not a good correlation between mRNA abundance and protein amount in a cell at a given time and that PTMs

cannot be predicted from gene sequences. In general, applications of 2D PAGE can be divided into descriptive protein mapping without a need for protein quantification and differential analysis, where protein quantification is essential.

Protein Mapping and 2D PAGE Protein Databases

The classical approach, 2D PAGE coupled with MS, allows the establishment of 2D reference maps. One aims at identifying as many proteins as possible. A typical workflow is depicted in Figure 3. With the number of completely sequenced genomes increasing daily, protein identification became possible for a growing number of species. Over the last years, many 2D reference maps have been established using different tissues, body fluids, or cell lines as sample (for an overview see http://ch.expasy.org/ch2d/2d-index.html). Due to the enormous complexity of cells from higher eukaryotes, characterization of their complete proteome in a single step is still impossible. Isolation of subcellular fractions such as organelles, macromolecular structures, and multiprotein

complexes represents a possibility to reduce the complexity of the entire cell.

Differential Analysis

Digitized 2D maps can serve more than just descriptive purposes. With specialized software packages, quantitative image analysis of the complex 2D patterns can be performed. This allows one to compare protein expression levels between different conditions for thousands of proteins in parallel. In general, one aims at identifying only the proteins showing differential expression in the samples under comparison. Since the early 1990s this approach has been used to tackle various biological questions. Comparison of proteomes from tissues or cells representing the healthy and the diseased state, respectively, allows the detection of putative molecular markers of the disease. Such a global approach is particularly useful in polygenic diseases such as cancer and cardiovascular diseases. Analysis of cell differentiation and monitoring therapy are other fields of application of differential analysis.
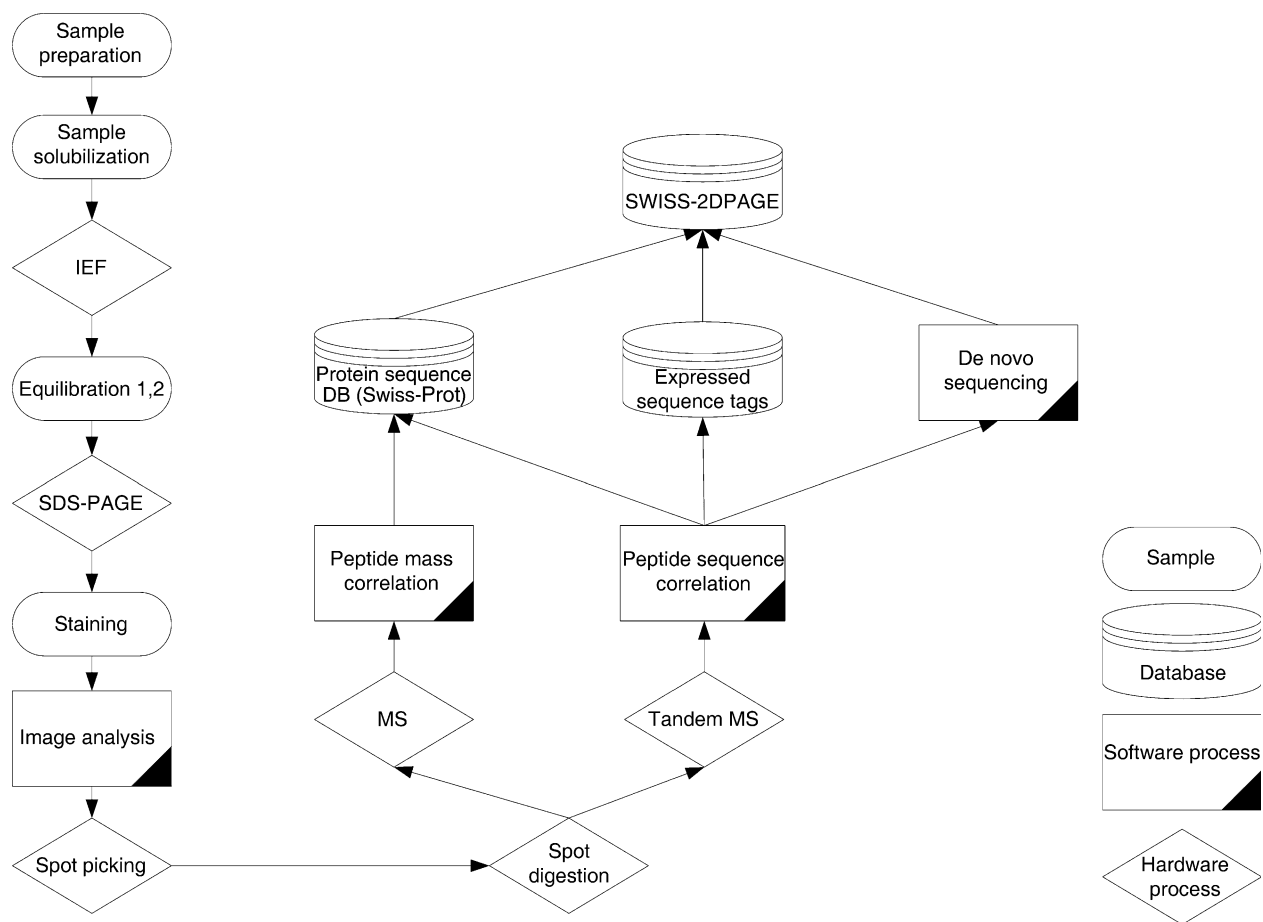
FIGURE 3   Typical proteomic workflow using 2D PAGE as protein separation tool.

Limitations and Perspectives of 2D PAGE

In spite of all benefits, 2D PAGE is not without limitations. On the one hand, there are intrinsic limitations of 2D PAGE and, on the other hand, there are general limitations encountered when dealing with complex protein mixtures.

Intrinsic Limitations of 2D PAGE

Hydrophobic Proteins

Very hydrophobic proteins such as integral membrane proteins and nuclear proteins were shown to be under-represented in 2D gels. They have a tendency to stick to IPG gel matrix, and re-solubilization for the second dimension becomes a problem. Incorporation of thiourea, sulfobetaine surfactants, and organic solvents in the solubilization buffer showed some improvements. However, very hydrophobic proteins will probably remain a major limitation of IEF and therefore also of 2D PAGE.

Basic Proteins

Horizontal streaking on the 2D gel is usually observed in the region with pH > 7. This phenomenon is particularly prominent when basic narrow range pH gradients are used. The streaking is due to the disappearance of the reducing agent from the basic part of the strip. Oxidation of the protein thiol groups results in the formation of inter- and intra-chain disulfide bridges. Recently, oxidation of the protein thiol groups to mixed disulfides using hydroxyethyl disulfide (DeStreak^{TM}) has been introduced, showing massively improved focusing for the very basic proteins.

Automation and Miniaturization

For many years, the low degree in automation and miniaturization of 2D PAGE has hindered this labor-intense technique from becoming routinely used in clinical laboratories and in industry, where high throughput counts and sample amount is limited. Promising developments toward both automation and miniaturization can be observed and their availability seems to be very close.

General Limitations

A biological sample usually represents a complex protein mixture containing thousands of proteins with different physico-chemical properties and widely varying expression levels. In human cells the dynamic range exceeds 6 orders of magnitude and in human plasma it reaches more than 12 orders of magnitude. In contrast, 2D gels and silver staining offer a maximum of 4 orders of magnitude. Consequently, application of unfractionated samples allows one to see only the tip of the iceberg. In order to access the low abundance proteins, samples need to be prefractionated, using traditional biochemical separation techniques, for example.

Perspectives of 2D PAGE

For many years, 2D PAGE has been the core technology for protein separation in proteomics research. This dominant role is now being challenged by various non-gel-based approaches such as protein chips, serial liquid chromatography, or capillary electrophoresis. They can be very demanding in terms of data management and bioinformatic resources, which might be limited for most standard academic research institutions. In fact, these liquid-based techniques should be considered as complementary approaches to 2D PAGE, depending on the biological question one wants to answer. If one aims at identifying as many proteins as possible in a given sample, 2D PAGE coupled to MS no longer represents the method of choice. However, if one aims at differential analysis of protein expression, 2D gels remain very powerful. This is mainly due to the possibility to look at PTMs offered by 2D PAGE. PTMs such as phosphorylations, glycosylations, or proteolytic processing can be very important for the function of a protein.

See Also the Following Articles

Glycation • Protein Data Resources • Proteinase-Activated Receptors

Glossary

immobilized pH gradient (IPG)   In an IPG gradient, a set of acidic and basic buffering groups is covalently incorporated into a polyacrylamide gel at the time it is cast. Precast IPG gel strips supported by a plastic film backing are commercially available in a variety of narrow and broad pH ranges.

isoelectric point (pI)   The pH value at which the net electric charge of an elementary entity, here a protein, is zero. At a pH value below the pI of a protein, it is charged positively, and at a pH value above the pI it is charged negatively.

proteome   Ensemble of proteins expressed by a genome of a species, an organ, or a cell at a particular moment under particular conditions.

proteomics   Qualitative and quantitative comparison of proteomes under different conditions to further unravel biological processes.

Further Reading

Chambers, G., Lawrie, L., Cash, P., and Murray, G. I. (2000). Proteomics: A new approach to the study of disease. *J. Pathol.* **192**, 280–288.

Görg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W. (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **21**, 1037–1053.

Jung, E., Heller, M., Sanchez, J.-C., and Hochstrasser, D. F. (2000). Proteomics meets cell biology: The establishment of subcellular proteomes. *Electrophoresis* **21**, 3369–3377.

Rabilloud, T. (ed.) (2000). *Proteome Research: Two-Dimensional Gel Electrophoresis and Identification Methods*. Springer, Berlin, Heidelberg.

Rabilloud, T. (2002). Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains. *Proteomics* **2**, 3–10.

Swiss-2DPAGE on the Expasy Molecular Biology Server – http://ch.expasy.org/ch2d/.

Wilkins, M. R., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (eds.) (1997). *Proteome Research: New Frontiers in Functional Genomics*. Springer, Berlin, Heidelberg.

BIOGRAPHY

Gerhard Schmid is a Ph.D. student at the Biomedical Proteomics Research Group (BPRG) of the Clinical Pathology Department, Geneva University Hospital.

Denis Hochstrasser is the Director of the Clinical Pathology Department. At the academic level, he is a full Professor both to the Department of Pathology of the Medicine faculty and to the School of Pharmacy of the Science faculty. His innovations in the methodology of 2D gel electrophoresis have contributed decisively to the technique's becoming one of the main protein separation methods used in proteomics.

Jean-Charles Sanchez is the Head of the BPRG since 1995. Working in the field of proteomics since 1989, he contributed to the development of 2D gel electrophoresis and its application in biomedical research. He obtained Ph.D. in Biochemistry at the University of Buckingham (UK) in the field of proteomics and diabetes.

Two-Hybrid Protein–Protein Interactions

Ilya Serebriiskii and Erica A. Golemis
Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA

The two-hybrid system is an artificially constructed genetic system intended to facilitate the detection and assessment of protein–protein interactions. In the two-hybrid system a host organism, typically yeast or bacteria, is so engineered as to contain three components. These are a first protein fused to a DNA-binding domain of known specificity (hybrid 1); a second protein fused to a transcriptional activation domain (hybrid 2), that can interact with the first protein, constituting a functional, albeit composite, transcription factor; and one or more reporter genes transcribed based on the binding of the composite transcription factor. Many permutations of the two-hybrid paradigm have been developed, and two-hybrid systems have become a mainstay of proteomic investigations.

The Yeast Two-Hybrid System

The purpose of the two-hybrid system is to provide a genetic system that is capable of scoring the degree of physical interaction between two proteins of interest. To this end, a simple organism such as yeast (*S. cerevisiae*) or bacteria (*E. coli*) is engineered such that the interaction of two proteins generates a scorable signal, that can be used either to gauge the interaction affinity of two defined proteins, or to identify interacting partners for one defined protein by screening an expression library.

The first working example of a two-hybrid system was described in 1989, by Fields and Song. A schematic showing the system paradigm is shown in Figure 1. There are three minimal components for a functioning two-hybrid system. The first component is an expression construct, which synthesizes a first protein of interest as a fusion with a DNA-binding domain (DBD) that binds a short DNA sequence of defined specificity (hybrid 1: frequently termed the "bait"). The second component is an expression construct that is used either to express a second defined protein, or random clones from a cDNA or genomic library, as a fusion with a transcriptional activation domain (AD) (hybrid 2: frequently termed the "prey"). The third component is a reporter cassette, which consists of the DNA-binding site for the first

hybrid protein in the context of a minimal promoter, upstream of the coding sequence for an easily scored reporter gene.

For the yeast two-hybrid system to work, there are two preconditions. First, the bait must not be independently functional as a transcriptional activator, such that yeast containing only the bait and the reporter produce no reporter signal. Second, the bait and the prey must be capable of localizing to the cell nucleus; hence, proteins containing signal sequences that efficiently target them to the plasma membrane, or other non-nuclear compartments, are not generally acceptable as baits. However, for proteins that meet these preconditions, the interaction between the bait and the prey brings the transcriptional activation domain encoded in the prey to the promoter sequence bound by the bait, and activates the transcription of the reporter gene.

A number of groups have built systems that exploit this two-hybrid paradigm in yeast, with several of these systems in common use. The most common DBDs used in these yeast two-hybrid systems are derived from the yeast transcription factor Gal4p, or from the bacterial repressor protein LexA. A variety of different AD sequences have been used, and include the activation domain from Gal4p; the activation domain from the viral protein VP16; or "artificial" activation sequences encoded from selected fragments of the *E. coli* genome.

Two categories of reporter genes have been used. One category is colorimetric, exemplified by the bacterial *lacZ* gene. Activation of a *lacZ* reporter causes yeast colonies to turn blue when plated on medium containing an appropriate substrate, such as X-gal, or can be assessed quantitatively over a dynamic range in excess of 1000-fold in liquid assays using a spectrophotometer or luminometer. The second category of reporter provides an active selection for yeast growth based on the strength of interaction between two proteins. These reporters include genes encoding critical components of yeast metabolic pathways for amino acid synthesis, such that expression of the gene is required for yeast to grow on medium lacking the relevant amino acid (e.g., show
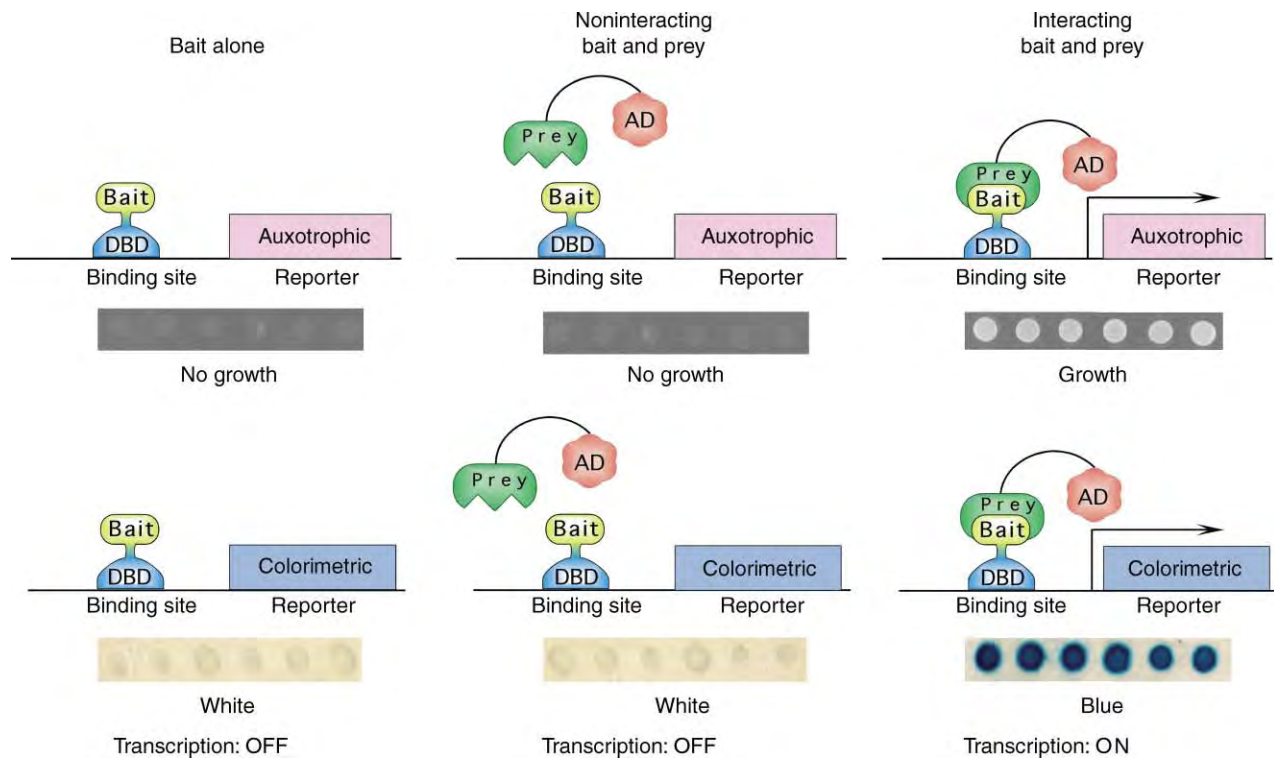
FIGURE 1 Schematic of yeast two-hybrid components, and profile of reporter activation. Details of system components are described in the text. Three situations are shown: (1) yeast containing only a bait and two reporters, left; (2) yeast containing bait, prey that does not interact with bait, and reporters, center; and (3) interacting bait and prey, and reporters, right. For each of these three situations, with a typical auxotrophic and colorimetric reporter, six independent colonies of yeast are dotted on plates with selective medium: expected results (growth/no growth, or color/no color) are shown as panel inserts below each schematic.

an auxotrophic requirement). *HIS3*, required for growth on medium lacking histidine, and *LEU2*, required for growth on medium lacking leucine, have been most commonly used. In general, two hybrid systems now generally incorporate one colorimetric and one auxotrophic selection, both responsive to the same bait. The use of two different reporter genes greatly minimizes the background of false positive and false negative interactions in the system, and is essential for library screening applications, where in excess of 10^6 clones must commonly be analyzed.

Because of the ease of manipulation provided by a genetic system, a yeast two-hybrid system provides a facile means of exploring the interaction of two known partner proteins. This can be done in a predetermined manner, by inserting targeted mutations or deletions into the AD-fused partner, or based on blind selection, for example, by using PCR mutagenesis or DNA sonication to develop a randomized library derived from the AD partner. Each mutant derivative can be readily screened for interaction phenotype. However, the most important use of the yeast two-hybrid system throughout the 1990s has clearly been for the purpose of screening libraries to identify novel interacting partners for proteins of interest. In contrast to other means of identifying interacting proteins, such as biochemical copurification,

two-hybrid screening is extremely inexpensive, requiring only engineered strains of yeast and standard microbiological media. It is rapid, providing the potential of completing a screen from start to finish in less than a month. It is effective at yielding returns for a significant percentage of bait proteins (estimated at $>50\%$). Finally, following completion of a successful library screen, plasmids encoding the novel interacting preys can easily be isolated from yeast, avoiding the need for protein sequencing or mass spectrophotometric analysis required by biochemical methodologies. It is not an exaggeration to note that recent broad adaptation of two hybrid screening methodologies across the scientific community was a major contributing source of elucidation for key cell signaling pathways.

Advanced Applications

Following the validation of the basic two-hybrid system as a useful and reliable tool for study of protein–protein interactions, a number of investigators have explored the potential of the system to study interactions between proteins and partner proteins under additional selective constraints. A first step in this direction was the demonstration that it was possible to identify preys

that interacted with baits contingent upon the modification of the bait by a coexpressed kinase. A second point of interest was the feasibility of identifying proteins that interacted not with a single protein, but with a complex formed by a bait and one or more simultaneously expressed "co-baits." A number of investigators have used a two-hybrid approach to identify proteins that interact as constituents of ternary or even quarternary complexes.

Another application of interest has been to use the two-hybrid system to identify and characterize interactions between proteins and non-protein ligands. To date, two-hybrid derivative systems have been used to score interactions between proteins and small molecules/drugs, and proteins and RNAs. In each case, the derivative system has been shown to function at the level of library screening. While screening for RNAs has become relatively broadly adapted as a research tool, screening for drugs that regulate protein–protein interactions remains at an earlier stage of validation. Detailed description of the derivatization of the two-hybrid system required for the screening of non-protein ligands is beyond the scope of this review.

The two-hybrid system has been used by a number of investigators to identify peptide "aptamers" capable of binding and regulating the functions of specific bait proteins. Peptides isolated based on two-hybrid selection have been found capable of regulating diverse functions of their cognate baits, and may provide the basis for design of peptidomimetics or small molecule derivatives of therapeutic value. This is currently an area of active investigation.

Finally, a recent effort has been to refine the two-hybrid system so as to make it a more useful tool for evaluating the selectivity of protein interactions. In these approaches, the two-hybrid system is "parallelized," using two discrete sets of baits and reporters within a single strain of yeast. In one example, bait 1 (a fusion to LexA) directs the expression of the *lacZ* and *LEU2* reporters, while bait 2 (a fusion to cI) directs the expression of *gusA* (colorimetric) and *LYS2* (auxotrophic) reporters. In a situation where a prey initially interacts with both baits 1 and 2, turning on all four reporters, mutagenesis of the prey can be coupled with selective screening for activation of the *lacZ* and *LEU2*, versus the *gusA* and *LYS2*, reporters, providing a convenient means of identifying altered specificity mutants that can be used to deconvolute signal transduction pathways.

Alternative Two-Hybrid Systems

Although the yeast two-hybrid system has found many useful applications, it has not been appropriate in all cases. Toward the goal of identifying small molecules or mutations that disrupt, rather than permit, protein–protein interactions, a "reverse" two-hybrid system has been developed. In this system, activation of a reporter gene is toxic, providing selection pressure for clones that have lost protein–protein interactions. For proteins that strongly activate transcription in yeast, and hence cannot be used to create baits, one solution has been to move the two-hybrid components to an RNA polymerase III-based reporter system, eliminating transcriptional activity. Other investigators have generated two-hybrid-like systems in bacteria, where eukaryotic transcriptional activation sequences are generally ineffective. For proteins that are membrane associated, or cannot efficiently be nuclear localized, several methods for assessing protein interaction at membranes or in the cytoplasm have been developed. These include systems in which interaction of the bait and prey reconstitutes an intracellular signaling pathway (e.g., the SOS system), or generate an assayable enzymatic by-product, as in the ubiquitin-based split-protein sensor system (USPS). In one recent and promising development, the USPS system has been further modified such that interactions of a bait and prey at the cell membrane triggers the cleavage and release of a transcription factor that migrates to the nucleus and activates standard yeast two-hybrid colorimetric and auxotrophic reporters. This approach combines advantages of membrane-based interaction detection with the ability to use well-validated screening modalities.

Role in Proteomics

Two hybrid protein–protein interaction systems are an important element of strategies to analyze the complete protein–protein interactions of proteomes. To date, extensive protein interaction maps have been constructed for viruses and yeast, and large-scale screening efforts are underway in *Caenorhabditis elegans* (*C. elegans*), *Drosophila*, and humans. These efforts provide a useful complement to simultaneous efforts to analyze protein interactions by mass spectrometry: whereas mass spectrometry is of value in identifying proteins which bind with high affinity to an assembled complex, albeit potentially weakly to any single interactive partner, the yeast two-hybrid system excels at identifying proteins that interact with high affinity with single defined partners. Interactions being detected by these means are being integrated into datasets including transcriptional expression profiles, genetic interaction data, and protein localization data for the same proteins, providing a fundamental tool for the emerging field of systems biology. It is likely that continuing application of the two-hybrid protein–protein interaction will remain a source of scientific insights for many years to come.

SEE ALSO THE FOLLOWING ARTICLES

LexA Regulatory System • Yeast GAL1–GAL10 System

GLOSSARY

bait A term used to describe a DNA-binding domain–protein "X" fusion, used as a probe in the yeast two-hybrid system (e.g., to screen a library).

interaction map A schematic showing a network of protein–protein interactions involving one or more proteins of interest.

interactome A recently introduced term, proposed to describe the complete network of protein–protein interactions occurring within an organism.

prey A term used to describe a transcriptional activation domain–protein "Y" fusion, which interacts with a bait.

two-hybrid system A system in which the interaction of two "hybrid" proteins (created so that (1) one is a fusion between a DNA-binding domain and protein "X," and the second is a fusion between a transcriptional-activation domain and protein "Y"; and (2) X and Y normally interact) in a host organism such as yeast or bacteria causes the activation of one or more scorable reporter genes.

FURTHER READING

Bernstein, D. S., Buter, N., Stumpf, C., and Wickens, M. (2002). Analyzing mRNA–protein complexes using a yeast three-hybrid system. *Methods* **26**, 123–141.

Fashena, S. J., Serebriiskii, I. G., and Golemis, E. A. (2000). The continued evolution of hybrid screening approaches in yeast: How to outwit different baits with different preys. *Gene* **250**, 1–14.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein–protein interaction. *Nature* **340**, 245–246.

Schwikowski, B., Ideker, T., and Uetz, P. (2002). Visualization and integration of protein–protein interactions. In *Protein Interactions* (E. A. Golemis, ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Serebriiskii, I. G., Khazak, V., and Golemis, E. A. (2001). Redefinition of the yeast two-hybrid system in dialogue with changing priorities in biological research. *BioTechniques* **30**, 634–655.

Stagljar, I., and Fields, S. (2002). Analysis of membrane protein interactions using yeast-based technologies. *Trends Biochem. Sci.* **27**, 559–563.

Walhout, A. J., Reboul, J., Shtanko, O., Bertin, N., Vaglio, P., Ge, H., Lee, H., Doucette-Stamm, L., Gunsalus, K. C., Schetter, A. J., Morton, D. G., Kemphues, K. J., Reinke, V., Kim, S. K., Piano, F., and Vidal, M. (2002). Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* Germline. *Curr. Biol.* **12**, 1952–1958.

Zhu, H., and Snyder, M. (2002). "Omic" approaches for unraveling signaling networks. *Curr. Opin. Cell Biol.* **14**, 173–179.

BIOGRAPHY

Ilya Serebriiskii is a Staff Scientist in the Division of Basic Sciences at the Fox Chase Cancer Center in Philadelphia, PA. His research interests include bacterial and yeast genetics and proteomics. He has created and evaluated numerous advanced protein interaction screening tools based on two-hybrid paradigms. He holds a Ph.D. in Biology from VNIIGenetika, Moscow, Russia. He received his postdoctoral training with Prof. Gerard Leblon at The Institute of Genetics and Microbiology, University of Paris-Sud, France, and then with Dr. Erica Golemis at Fox Chase Cancer Center.

Erica Golemis is a Member of the Division of Basic Sciences at Fox Chase Cancer Center. Her interests include the study of protein interactions, and the intersection of cell attachment and cell cycle control signaling in mammalian cancer development. She performed research with Dr. Nancy Hopkins leading to a Ph. D. in Biology from the Massachusetts Institute of Technology in Cambridge, MA. Subsequent postdoctoral work with Dr. Roger Brent at Massachusetts General Hospital in Boston, MA, led to the development of an original LexA-based two-hybrid system. She has worked extensively in the area of yeast two-hybrid system reagent development.

Tyrosine Sulfation

Denis Corbeil and Wieland B. Huttner
Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

The O-sulfation of tyrosine residues of membrane and secretory proteins that transit through the secretory pathway of eukaryotic cells is a ubiquitous posttranslational modification conserved in all multicellular organisms. Tyrosine sulfation is catalyzed by tyrosylprotein sulfotransferase (TPST) isoenzymes, which are integral membrane proteins of the trans-Golgi network. Tyrosine sulfation has been shown to be important for protein–protein interactions occurring in diverse biological processes, ranging from the receptor binding of regulatory peptides to the interaction of viral envelope proteins with the cell surface.

Tyrosine-Sulfated Proteins

OCCURRENCE

Sulfation is one of the most abundant posttranslational modifications of tyrosine, since up to 1% of the tyrosine residues of total protein in an organism can be sulfated. Tyrosine is the only amino acid residue in proteins known to undergo sulfation. Following the first description of a sulfated-tyrosine residue in a peptide derived from fibrinogen by Bettelheim in 1954, it has been known since 1982 that tyrosine-sulfated proteins occur in all animals, from lower invertebrates up to humans. Tyrosine-sulfated proteins also exist in the plant kingdom, for example, in the green alga volvox, one of the earliest truly multicellular organisms, or in higher plants such as rice. While occurring throughout metazoan evolution, tyrosine-sulfated proteins appear to be absent in unicellular eukaryotes and prokaryotes, implicating this post translational modification in some aspect of multicellularity.

In a given animal, tyrosine-sulfated proteins have been observed in all tissues examined. Each tissue appears to contain a characteristic set of tyrosine-sulfated proteins, suggesting that proteins with tissue-specific expression are major targets for tyrosine sulfation. In cell culture, tyrosine sulfation of proteins has been detected in all primary cultures and cell lines investigated, including various secretory cells, epithelial cells, fibroblasts, neuronal cells, and cells of the immune system.

Tyrosine-sulfated proteins can be identified by various methods, including labeling using radioactive sulfate followed by tyrosine sulfate analysis of a given protein. In line with the intracellular localization of tyrosylprotein sulfotransferase (TPST) in the trans-Golgi network, all known tyrosine-sulfated proteins are either secretory or plasma membrane proteins. Reviews with comprehensive lists of tyrosine-sulfated proteins have been published, and several of these proteins have been shown to play important biological roles.

STRUCTURAL DETERMINANTS OF TYROSINE SULFATION

The recognition, by TPST, of the tyrosine residue to be sulfated in a secretory protein or the extracellular domain of a membrane protein requires the presence of certain structural features. These have been deduced from the comparison of identified tyrosine sulfation sites and the *in vitro* tyrosine sulfation of synthetic peptides. Although no strict consensus sequence for tyrosine sulfation exists, all sequences exhibit the presence of acidic amino acid residues in the vicinity, i.e., positions -5 (N-terminal) to $+5$ (C-terminal), of the sulfated tyrosine residue. A particularly critical position appears to be amino-terminal (-1) to the tyrosine. Turn-inducing amino acids (P, G) are also frequently present. The few examples of tyrosine sulfation sites lacking proline and glycine are located near the N- or C-terminus of the protein and/or contain several of the three other amino acid residues (D, S, N) with significant turn-conformational potential. Finally, another characteristic feature common to most identified tyrosine sulfation sites is the absence of cysteine residues or potential N-glycosylation sites (NXS or NXT) in the vicinity (positions -7 to $+7$). In either case, the presence of a disulfide-bridge or N-linked oligosaccharides is likely to prevent sulfation of nearby tyrosine residues due to steric hindrance.

Tyrosylprotein Sulfotransferase (EC 2.8.2.20)

THE TYROSINE SULFATION REACTION

The sulfate transfer reaction to tyrosine residues is catalyzed by TPST, first described in 1983, and used as a
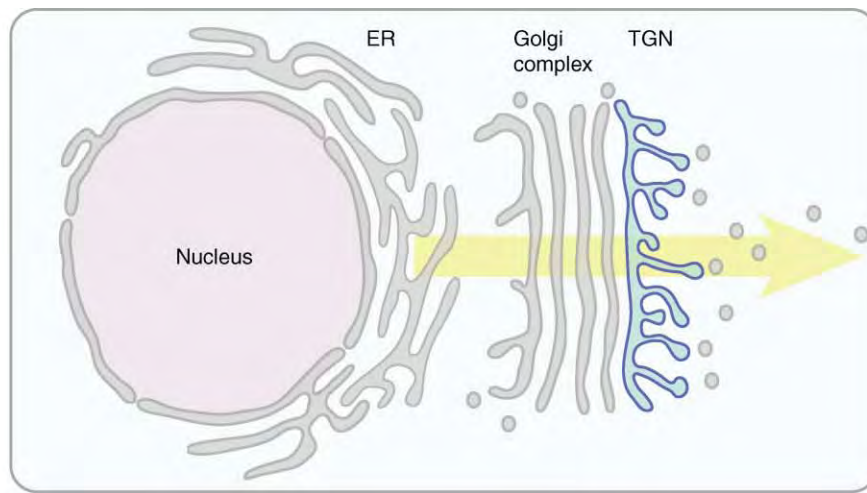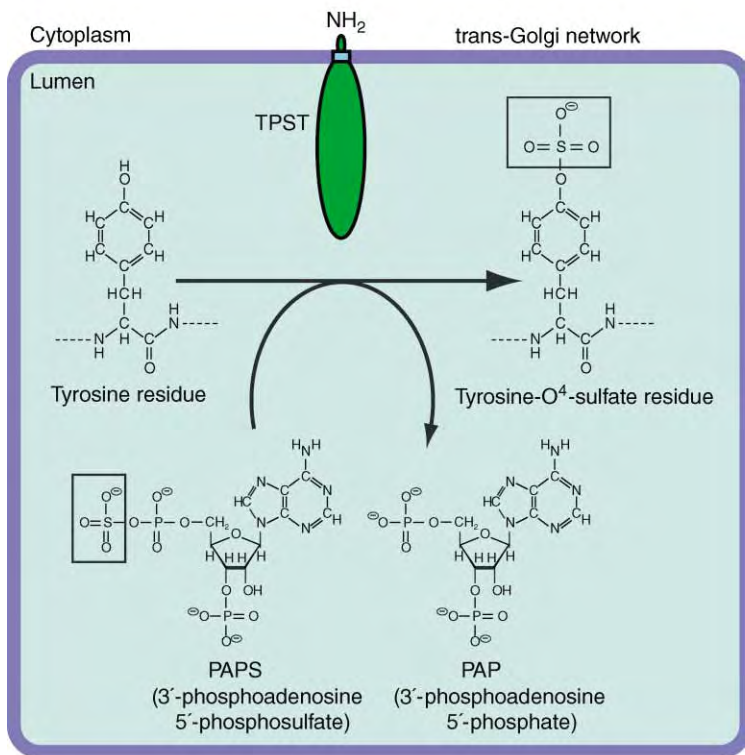
FIGURE 1   (A) Tyrosine sulfation occurs in the trans-Golgi network (TGN). (B) TPST, a type II transmembrane protein, catalyzes the transfer of the sulfate group (boxes) from the co-substrate PAPS to tyrosine residues of secretory proteins and ectodomains of membrane proteins passing through the lumen of the trans-Golgi network. (C) The three principal types of tyrosine sulfation-dependent protein–protein interactions.

sulfate donor, the cosubstrate 3′-phosphoadenosine 5′-phosphosulfate (PAPS) (Figure 1B). The transfer is thought to occur in an ordered reaction mechanism that involves the following sequential steps: (1) cosubstrate PAPS binding; (2) substrate binding; (3) sulfate transfer; (4) tyrosine-sulfated product release and, (5) PAP release. TPST activity can be detected by incubation of appropriate membrane preparations with [^{35}S]PAPS, using either endogenous or exogenous protein substrate.

Biologically, tyrosine sulfation appears to be an irreversible event *in vivo* due to the lack of a sulfatase capable of catalyzing the desulfation of tyrosine-sulfated proteins under physiological conditions.

PROPERTIES OF TPST

TPST is an integral membrane protein residing in the trans-Golgi network (Figure 1A). The TPST protein can

be identified in membrane preparations by MSC (*m*odification after *s*ubstrate *c*ross-linking) labeling, which is based on the cross-linking of a substrate peptide to TPST followed by intramolecular [^{35}S]sulfate transfer from the cosubstrate PAPS. Various nonionic detergents can be used to solubilize TPST from membranes.

TPST purified from bovine adrenal medulla is a 50–54 kDa sialoglycoprotein with an apparent S-value of 6. Its pH optimum is between 6.0 and 6.5, which is in line with the slightly acidic pH of the trans-Golgi network. The catalytic activity of the enzyme, towards endogenously as well as exogenously added proteins, is stimulated by divalent cations (Mg^{2+}, Mn^{2+}), and fluoride ions are required for maximal activity. The K_m for the cosubstrate PAPS is 1.4 µM. Like most sulfotransferases, TPST is inhibited by PAP. *In vitro*, the activity of the enzyme is also inhibited by certain lipids such as sphingosine, but the physiological relevance of this observation is unknown. Synthetic inhibitors of TPST (with IC_{50} values of 30–40 µM) have been generated by a combinatorial target-guided ligand assembly technique.

The K_m for most peptides (8–13 amino acid residues) with a single tyrosine sulfation site is in the range of 10–100 µM. It is likely that similar values hold true for individual tyrosine sulfation sites in proteins. Interestingly, K_m values for synthetic substrates containing multiple tyrosine sulfation sites are considerably lower. Given that several proteins exist with multiple adjacent sulfation sites, e.g., cionin, heparin cofactor, and preprocholecystokinin, the increased affinity of TPST for such substrates might be of physiological significance, e.g., by promoting stoichiometric sulfation. A remarkable example of the sequential sulfation of four tyrosine residues within a short stretch of amino acids is the N-terminal domain of the CC-chemokine receptor 5 (CCR5), which is a coreceptor for HIV.

MOLECULAR CLONING AND MEMBRANE TOPOLOGY OF TPST

Distinct TPST isoenzymes, anticipated from the differential substrate specificities of various TPST preparations, exist in the human and mouse genome. Each contains two TPSTs, TPST-1 (see Swiss-Prot Database, Accession No. 060507) and TPST-2 (see Swiss-Prot Database, Accession No. 060704), which were first characterized at the molecular level in 1998. The human *TPST-1* (7q11) and *TPST-2* (22q12.1) genes encode for 370- and 377-amino acid proteins, respectively, that share an overall 65% identity. The murine *tpst* genes are located on chromosome 5, and the corresponding proteins show ~95% identity to their human counterpart. Both TPST isoenzymes are predicted to have a type

II membrane topology, with a very short NH_2-terminal cytoplasmic domain (8 residues) and the bulk of the polypeptide, which is responsible for the catalytic activity, being located in the Golgi lumen (see Figures 1A and 1B). Although overall, TPST isoenzymes display a low degree of amino acid identity to other cytosolic and Golgi-associated sulfotransferases, most of the residues involved in PAPS binding, as deduced from comparison with the crystal structure of estrogen sulfotransferase, are conserved. The luminal domain of either isoenzyme also contains two potential N-glycosylation sites, consistent with presence of N-linked glycans in the TPST protein.

In agreement with the occurrence of tyrosine sulfation in all metazoan species, cDNAs that predict proteins obviously related to mammalian TPSTs are found in various vertebrates and invertebrates, including fish (GenBank Accession No. BI846282), frog (GenBank Accession No. BG815875), chicken (GenBank Accession Nos. BU142429, BU217098), fly (GenBank Accession No. AY124548), and worm.

TPST-1 VERSUS TPST-2

TPST-1 and TPST-2 transcripts are found in all tissues examined, e.g., brain, heart, skeletal muscle, gut, kidney, liver, lung, and leukocytes. Although ubiquitously expressed, the tissue distribution of TPST-1 and TPST-2 are not identical, recombinant TPST-1 and TPST-2 show similar, but not identical, activities toward certain small peptide substrates, consistent with some functional redundancy as well as a certain degree of differential substrate specificity.

TPST-1- AND TPST-2-DEFICIENT MICE

TPST-1- and TPST-2-deficient mice, generated by targeted disruption of the *tpst-1* and *tpst-2* genes, also point to functional redundancy between the two isoenzymes, as either line of knock-out mice is viable (K.L. Moore, personal communication). Although TPST-1$^{-/-}$ animals appear normal, their body weight is reduced about 5% and an increase in postimplantation fetal death is observed, suggesting that unidentified proteins involved in regulation of body weight and reproductive physiology require tyrosine sulfation for optimal function. TPST-2 deficient mice show a transient delay in growth during postnatal development. In addition, the males, but not females, appear to be infertile. Together, these observations are consistent with the idea that TPST-1 and TPST-2 have distinct, but partially overlapping, physiological roles.

Physiological Function and Medical Relevance

For most tyrosine-sulfated proteins, the physiological function of this posttranslational modification is presently unknown. With regard to the cases in which the biological role of tyrosine sulfation of a particular protein has been elucidated, the common denominator has emerged that tyrosine sulfation promotes extracellular protein–protein interactions. In line with the identification of numerous secretory and plasma membrane proteins that are tyrosine-sulfated, paradigmatic examples exist showing that tyrosine sulfation promotes the interaction between (1) a secretory and a plasma membrane protein, (2) two secretory proteins, or (3) two plasma membrane proteins (Figure 1C).

The regulatory peptide cholecystokinin is a classical example where tyrosine sulfation of a secretory protein dramatically promotes its interaction with a plasma membrane protein, i.e., its cell surface receptor. Thus, sulfated cholecystokinin is 260 times more potent than its unsulfated form. Of the several examples of tyrosine sulfation promoting the interaction between two secretory proteins, the case of the binding of the tyrosine-sulfated blood coagulation factor VIII to von-Willebrand-factor is particularly intriguing, as it also documents the medical relevance of this posttranslational modification. Humans with a mutation in the critical tyrosine residue of factor VIII that is sulfated and involved in its binding to von-Willebrand-factor are afflicted with hemophilia A.

An example of tyrosine sulfation promoting the interaction between two plasma membrane proteins is the important role of this posttranslational modification for the high-affinity binding of leukocyte-associated P-selectin glycoprotein ligand (PSGL)-1 to P-selectin on activated endothelial cells. This crucial interaction initiates adhesion of leukocytes to the vascular wall during inflammation. Tyrosine sulfation also occurs in seven-transmembrane-segment chemokine receptors, e.g., CCR5. Under physiological conditions, these plasma membrane proteins play a central role in chemokine signalling pathway through G proteins. Remarkably, human and simian immunodeficiency viruses use CCR5 as a co-receptor, together with CD4, to mediate their attachment to the host cell membrane. Specifically, sulfation of tyrosine residues in the CCR5 N-terminal domain has been shown to be critical for the interaction of this protein with HIV envelope glycoprotein gp120, leading to HIV infection. Thus, the design of tyrosine-sulfated peptide competitors – mimicking HIV gp120-binding sites – could turn out to be the basis for new therapeutic compounds that will block HIV cellular entry. These examples highlight the medical relevance of protein tyrosine sulfation.

SEE ALSO THE FOLLOWING ARTICLES

Golgi Complex • Oligosaccharide Chains: Free, *N*-Linked, *O*-Linked • Secretory Pathway

GLOSSARY

PAPS 3′-phosphoadenosine 5′-phosphosulfate, sulfate donor in the sulfate transfer reaction. PAPS has been known to be the activated form of sulfate and acts as cosubstrate for the sulfation of a wide variety of substances, including proteins.

trans-Golgi network The last station of the Golgi complex. This site is a major branching point of vesicular transport and the origin of two principal pathways of protein secretion: the regulated and constitutive pathways.

FURTHER READING

Bettelheim, F. R. (1954). Tyrosine-O-sulfate in a peptide from fibrinogen. *J. Am. Chem. Soc.* **76**, 2838–2839.

Huttner, W. B. (1982). Sulphation of tyrosine residues – a widespread modification of proteins. *Nature (London)* **299**, 273–276.

Huttner, W. B. (1984). Determination and occurrence of tyrosine O-sulfate in proteins. *Meth. Enzymol.* **107**, 200–223.

Huttner, W. B., and Baeuerle, P. A. (1988). Protein sulfation on tyrosine. *Mod. Cell Biol.* **6**, 97–140.

Huttner, W. B., Niehrs, C., and Vannier, C. (1991). Bind or bleed. *Curr. Biol.* **1**, 309–310.

Kehoe, J. W., and Bertozzi, C. R. (2000). Tyrosine sulfation: A modulator of extracellular protein–protein interactions. *Chem. Biol.* **7**, R57–R61.

Kehoe, J. W., Maly, D. J., Verdugo, D. E., Armstrong, J. I., Cook, B. N., Ouyang, Y. B., Moore, K. L., Ellman, J. A., and Bertozzi, C. R. (2002). Tyrosylprotein sulfotransferase inhibitors generated by combinatorial target-guided ligand assembly. *Bioorg. Med. Chem. Lett.* **12**, 329–332.

Moore, K. L. (2003). The biology and enzymology of protein tyrosine O-sulfation. *J. Biol. Chem.* **278**, 24243–24246.

Niehrs, C., Beisswanger, R., and Huttner, W. B. (1994). Protein tyrosine sulfation, 1993 – an update. *Chem. Biol. Interact.* **92**, 257–271.

Ouyang, Y.-B., Crawley, J. T. B., Aston, C. E., and Moore, K. L. (2002). Reduced body weight and increased postimplantation fetal death in tyrosylprotein sulfotransferase-1-deficient mice. *J. Biol. Chem.* **277**, 23781–23787.

BIOGRAPHY

Denis Corbeil is a Group Leader at the University of Dresden. His research interests are in the cell biology of stem cells, with a focus on prominin/CD133. He holds a Ph.D. from the University of Montreal and received postdoctoral training in the laboratory of W.B. Huttner, where he participated in the molecular cloning of TPST.

Wieland B. Huttner is a Professor of Neurobiology and Director at MPI-CBG in Dresden. His group made seminal contributions on protein tyrosine sulfation, including the identification, characterization, purification, and cloning of TPST. He holds an M.D. from the University of Hamburg, received postdoctoral training with Nobel Laureate Paul Greengard at Yale University, and has been pursuing research on neurosecretory vesicle biogenesis and neurogenesis in the mammalian central nervous system.

Ubiquitin System

Aaron Ciechanover and Michael H. Glickman

Technion – Israel Institute of Technology, Haifa, Israel

In the ubiquitin system, a target substrate is modified by ubiquitin or a ubiquitin-like protein. In most cases, proteins are modified by multiple moieties of ubiquitin, generating a polyubiquitin chain. This modification leads to their degradation by the 26S proteasome complex. Modification by a single moiety of ubiquitin can target proteins for degradation in the lysosome/vacuole. Modification by ubiquitin-like proteins serves nonproteolytic functions. Conjugation of ubiquitin is carried out by a modular cascade of enzymes, specific to each substrate. Ubiquitination of cellular proteins is a highly complex, temporally controlled, and tightly regulated process that targets in a specific manner thousands of cellular proteins and plays major roles in a variety of basic pathways, such as cell division, differentiation, and quality control. Not surprisingly, aberrations in the system underlie the pathogenesis of many diseases, certain malignancies, and neurodegenerative disorders. Mechanism-based drugs are currently being developed.

Mechanisms of Ubiquitination and Degradation

UBIQUITINATION

Degradation of a protein via the ubiquitin–proteasome pathway involves two discrete and successive steps: (1) tagging of the substrate by covalent attachment of multiple ubiquitin molecules, and (2) degradation of the tagged protein by the 26S proteasome complex with release of free and reusable ubiquitin. This last process is mediated by ubiquitin-recycling isopeptidases (deubiquitinating enzymes). Conjugation of ubiquitin, a highly evolutionarily conserved 76 residue polypeptide, to the protein substrate proceeds via a three-step cascade mechanism. Initially, the ubiquitin-activating enzyme, E1, activates ubiquitin in an ATP-requiring reaction to generate a high-energy thiol ester intermediate, E1-S~ubiquitin. One of several E2 enzymes (ubiquitin-carrier proteins or ubiquitin-conjugating enzymes (UBCs)) transfers the activated ubiquitin from E1, via an additional high-energy thiol ester intermediate, E2-S ~ ubiquitin, to the substrate that is specifically bound to a member of the ubiquitin-protein ligase family, E3. There are a number of different classes of E3 enzymes. For the homologous to the E6-AP C-terminus (HECT) domain E3s, the ubiquitin is transferred once again from the E2 enzyme to an active site Cys residue on the E3, to generate a third high-energy thiol ester intermediate, ubiquitin-S ~ E3, prior to its transfer to the ligase-bound substrate. RING finger-containing E3s catalyze direct transfer of the activated ubiquitin moiety to the E3-bound substrate.

E3s catalyze the last step in the conjugation process: covalent attachment of ubiquitin to the substrate. The ubiquitin molecule is generally transferred to an ε-NH_2 group of an internal lysine residue in the substrate to generate a covalent isopeptide bond. In some cases however, ubiquitin is conjugated to the N-terminal amino group of the substrate. By successively adding activated ubiquitin moieties to internal lysine residues on the previously conjugated ubiquitin molecule, a polyubiquitin chain is synthesized. The chain is recognized by the downstream 26S proteasome complex. Thus, E3s play a key role in the ubiquitin-mediated proteolytic cascade since they serve as the specific recognition factors of the system. In certain cases the first ubiquitin moiety is conjugated to the substrate by one E3, while chain elongation is catalyzed by a different ligase often termed E4. Modification by a single moiety of ubiquitin is catalyzed via an identical mechanism and set of enzymes. The specific enzymes that catalyze modification by ubiquitin-like proteins are somewhat different, though they utilize a similar mechanism.

DEGRADATION

Degradation of polyubiquitinated substrates is carried out by a large protease complex, the 26S proteasome that does not recognize nonmodified substrates. In one established case, that of the polyamine synthesizing enzyme ornithine decarboxylase (ODC), the proteasome recognizes and degrades the substrate without prior ubiquitination. The proteasome is a multicatalytic protease that degrades polyubiquitinated proteins to short peptides. It is composed of two subcomplexes: a 20S core particle (CP) that carries the catalytic activity, and a regulatory 19S regulatory particle (RP). The 20S CP is a barrel-shaped structure composed of four

stacked rings, two identical outer α-rings and two identical inner β-rings. The eukaryotic α- and β-rings are composed each of seven distinct subunits, giving the 20S complex the general structure of $\alpha_{1-7}\beta_{1-7}\beta_{1-7}\alpha_{1-7}$. The catalytic sites are localized to some of the β-subunits. Each extremity of the 20S barrel can be capped by a 19S RP. One important function of the 19S RP is to recognize ubiquitinated proteins and other potential substrates of the proteasome. A ubiquitin-binding subunit of the 19S RP has indeed been identified; however, its importance and mode of action have not been discerned. A second function of the 19S RP is to open an orifice in the α-ring that will allow entry of the substrate into the proteolytic chamber. Also, since a folded protein would not be able to fit through the narrow proteasomal channel, it is assumed that the 19S particle unfolds substrates and inserts them into the 20S CP. Both the channel opening function and the unfolding of the substrate require metabolic energy, and indeed, the 19S RP contains six different ATPase subunits. Following degradation of the substrate, short peptides derived from the substrate

are released, as well as reusable ubiquitin. Proteasomal degradation is not always complete. In some cases, the proteasome, rather than completely destroying its target, processes the ubiquitinated substrate precisely, releasing a truncated product. In the case of the NF-κB transcriptional regulator, an active subunit (p50 or p52) is thus released from a longer inactive precursor (p105 or p100). For a general scheme of the ubiquitin system and its multiple functions, see Figure 1.

SUBSTRATE RECOGNITION

A major unresolved question is how does the system achieve its high specificity and selectivity. Why are certain proteins extremely stable in the cell, while others are extremely short-lived? Why are certain proteins degraded only at a particular time point during the cell cycle or only following specific extracellular stimuli, yet they are stable under most other conditions? It appears that specificity of the ubiquitin system is determined by two distinct and unrelated groups of proteins: E3s and ancillary proteins. First, within
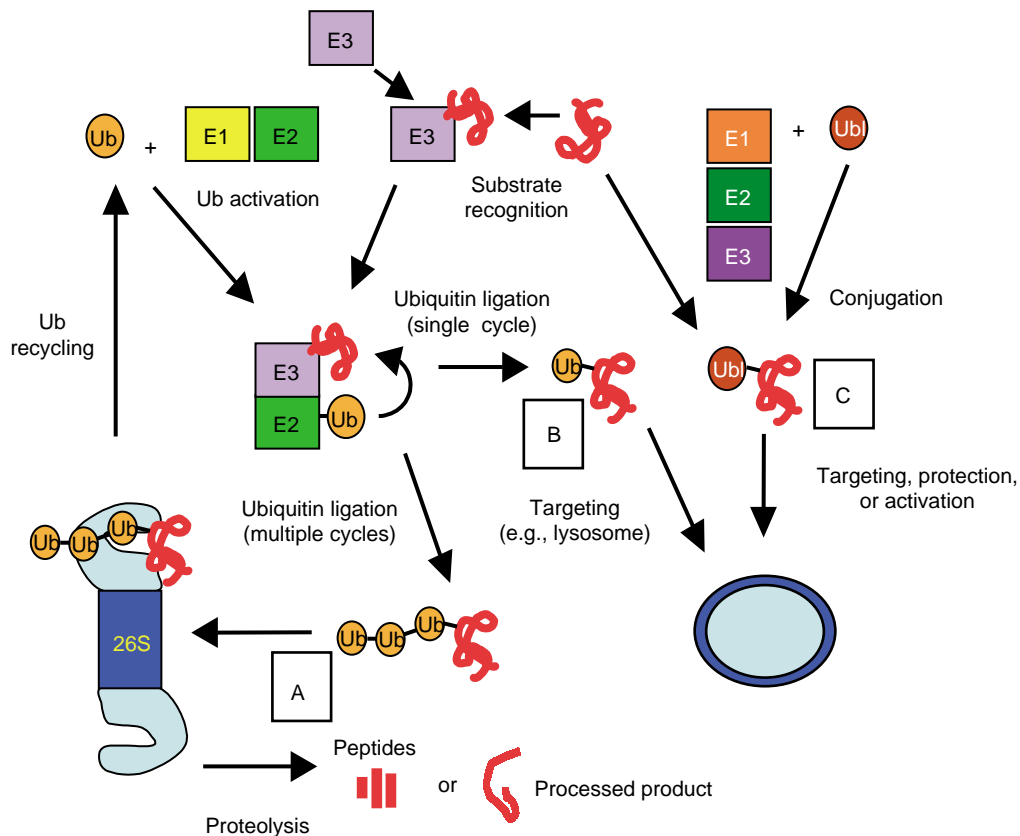
FIGURE 1 The ubiquitin system for protein modification. Ubiquitin (Ub) is activated by the three enzymes E1, E2, and E3, and the active ubiquitin moiety is transferred to the target substrate (red). Substrates can be either modified by a polyubiquitin chain (A) and targeted to the 26S proteasome complex for degradation (or processing), or modified by a single ubiquitin moiety, monoubiquitinated (B) and targeted to other organelles such as the lysosome (wherein it is degraded). Ubiquitin-like proteins (Ubl) are activated by E1, E2, and E3, which are similar but not identical to the enzymes that activate ubiquitin. Modification by Ubl (C) serves a variety of nonproteolytic functions such as routing of proteins to subcellular organelles (e.g., nucleus), protecting them from ubiquitination, or activating them.

the ubiquitin system, substrates must be specifically recognized by an appropriate E3 as a prerequisite to their ubiquitination. In most cases however, substrates are not recognized in a constitutive manner, and in some cases they are not recognized directly by the E3. In some instances, the E3 must "be switched on" by undergoing posttranslational modification in order to yield an active form that recognizes the substrate. In many other cases, it is the substrate that must undergo a certain change that renders it susceptible for recognition. The stability of additional proteins depends on association with ancillary proteins such as molecular chaperones that act as recognition elements in *trans* and serve as a link to the appropriate ligase. Others, such as certain transcription factors, have to dissociate from the specific DNA sequence to which they bind in order to be recognized by the system. Stability of yet other proteins depends on oligomerization. Thus, in addition to the E3s themselves, modifying enzymes (such as kinases), ancillary proteins, or DNA sequences to which substrates bind also play an important role in the recognition process.

Functions and Substrates of the Ubiquitin System

Ubiquitin-mediated proteolysis of a variety of cellular proteins plays an important role in many basic cellular processes. Among these are regulation of cell cycle and division, differentiation and development, involvement in the cellular response to stress and extracellular effectors, morphogenesis of neuronal networks, modulation of cell surface receptors, ion channels and the secretory pathway, DNA repair, transcriptional regulation, transcriptional silencing, long-term memory, circadian rhythms, regulation of the immune and inflammatory responses, and biogenesis of organelles. The list of cellular proteins that are targeted by ubiquitin is growing rapidly. Among them are cell-cycle regulators such as cyclins, cyclin-dependent kinase inhibitors, and proteins involved in sister chromatid separation, tumor suppressors, transcriptional activators, and their inhibitors. Cell surface receptors and endoplasmic reticulum (ER) proteins are also targeted by the system. Finally, mutated and denatured/misfolded proteins are recognized specifically and removed efficiently. In this capacity, the system is a key player in the cellular quality control and defense mechanisms. The products of the proteasome can play an important role in the immune response. In the case of degradation of foreign proteins – such as those of viral origin – the resulting short peptides are presented by MHC class I molecules to the cytotoxic T cell that lyse the presenting cell.

Regulation of the Ubiquitin System

The ubiquitin–proteasome pathway can be regulated at the level of ubiquitination or at the level of proteasome activity. Since conjugation and proteasomal degradation is required for multitude of cellular functions, regulation must be delicately and specifically tuned. In a few cases, general rather than specific, components of the pathway can be modulated by physiological signals. For example, up-regulation of the pathway is observed during massive degradation of skeletal muscle proteins that occurs under normal fasting, but also under pathological conditions such as cancer cachexia, severe sepsis, metabolic acidosis, or following denervation. In most cases however, regulation is specific and the target substrates are recognized by specific ligases that bind to defined motifs. The targeting motif can be a single amino acid residue (e.g., the N-terminal residue) or a sequence (the "destruction" box in cyclins) or a domain (such as a hydrophobic patch) that is not normally exposed. In other cases the motif is a posttranslational modification such as phosphorylation that is generated in response to cell needs or external signals. Phosphorylation can occur either on the substrate or on the ligase.

Ubiquitin-Like Proteins

Both enzymes and substrates of the ubiquitin system have been found to be modified by ubiquitin-like proteins. Modification by ubiquitin-like proteins occurs only once. In the case of enzymes, modification affects their activity. For example, the modification by the ubiquitin-like protein NEDD8 enhances the activity of a certain class of E3 ligases. Conjugation of NEDD8 to one of the components of the ligase complex (Cullin) increases its affinity to other components of the conjugation machinery. In the case of substrates, modification can affect their availability to the ubiquitination/degradation machinery and consequently cellular stability. For example, in the case of IκBα, the inhibitor of the transcriptional regulator NF-κB, modification by SUMO-1 was shown to protect the substrate from ubiquitination. In a completely different case, SUMOylation of RanGAP1 targets the protein to its final subcellular destination in the nuclear pore complex.

Ubiquitination and Pathogenesis of Human Diseases

DISEASES

While inactivation of a major enzyme such as E1 is obviously lethal, mutations or acquired changes in

enzymes or in recognition motifs in substrates that do not affect vital pathways or that affect the involved process only partially, may result in a broad array of diseases. The pathological states associated with the ubiquitin system can be classified into two groups: those that result from (1) loss of function – mutation in a ubiquitin system enzyme or target substrate that result in stabilization of certain proteins, and (2) gain of function – abnormal or accelerated degradation of the protein target.

Alterations in ubiquitination and deubiquitination reactions have been directly implicated in the etiology of many malignancies. In general, cancers can result from "stabilization" of oncoproteins, or "destabilization" of tumor suppressor gene products. Some of the natural substrates for degradation by the proteasome are oncoproteins that, if not properly removed from the cell, can promote cancer. For instance, ubiquitin targets N-myc, c-myc, c-fos, c-jun, Src, and the adenovirus E1A proteins. Destabilization of tumor suppressor proteins such as p53 and p27 has also been implicated in the pathogenesis of malignancies.

In one fascinating case, that of uterine cervical carcinoma, the level of the tumor suppressor protein p53 is extremely low. Most of these malignancies are caused by high-risk strains of the human papillomavirus (HPV). Detailed studies have shown that the suppressor is targeted for ubiquitin-mediated degradation by the virally encoded oncoprotein E6. Degradation is mediated by the native HECT domain E3 enzyme E6-AP, where E6 serves as an ancillary protein that allows recognition of p53 in *trans*. E6-AP will not recognize p53 in the absence of E6. E6 associates with both the ubiquitin-ligase and the target substrate and brings them to the necessary proximity that is assumed to allow catalysis of conjugation to occur. Removal of the suppressor by the oncoprotein is probably an important mechanism used by the virus to transform cells.

Accumulation of ubiquitin conjugates and/or inclusion bodies associated with ubiquitin, proteasome, and certain disease-characteristic proteins have been reported in a broad array of chronic neurodegenerative diseases, such as the neurofibrillary tangles of Alzheimer's disease (AD), brainstem Lewy bodies (LBs) – the neuropathological hallmark in Parkinson's disease (PD), and nuclear inclusions in CAG repeat expansion (poly-glutamine expansion) disorders such as occurring Huntington's disease. However, in all these cases, a direct pathogenetic linkage to aberrations in the ubiquitin system has not been established. One factor that complicates the establishment of such linkage is the realization that many of these diseases, such as Alzheimer's and Parkinson's, are not defined clinical entities, but rather syndromes with different etiologies. Accumulation of ubiquitin conjugates in Lewy inclusion bodies in many of these cases may be secondary, and reflects unsuccessful attempts by the ubiquitin and proteasomal machineries to remove damaged/abnormal proteins. While the initial hypothesis was that inclusion bodies are generated because of the inherent tendency of the abnormal proteins to associate with one another and aggregate, it is now thought that the process may be more complex and involves active cellular machineries, including inhibition of the ubiquitin system by the aggregated proteins. This aggregation of brain proteins into defined lesions is emerging as a common, but poorly understood mechanistic theme in many sporadic and hereditary neurodegenerative disorders.

The case of Parkinson's disease highlights the complexity of the involvement of the ubiquitin system in the pathogenesis of neurodegeneration. Aberrations in several proteins such as mutations in α-synuclein, an important neuronal protein, or in the deubiquitinating enzyme UCH-L1, have been described that link the ubiquitin system to the pathogenesis of the disease. One important player in the pathogenesis of Parkinson's disease is Parkin which is a RING-finger E3. Mutations in the gene appear to be responsible for the pathogenesis of autosomal recessive juvenile parkinsonism (AR-JP), one of the most common familial forms of Parkinson's disease. Parkin ubiquitinates and promotes the degradation of several substrates. It is possible that aberration in the degradation of one of these substrates that leads to its accumulation is neurotoxic and underlies the pathogenesis of AR-JP.

The cystic fibrosis gene encodes the CF transmembrane regulator (CFTR) that is a chloride channel. Only a small fraction of the protein matures to the cell surface, whereas most of it is degraded from the ER by the ubiquitin system prior to its maturation. One frequent mutation in the channel is ΔF508. The mutation leads to an autosomal recessive inherited multisystem disorder characterized by chronic obstruction of airways and severe maldigestion due to exocrine pancreatic dysfunction. Despite normal ion channel function, CFTR$^{\Delta F508}$ does not reach the cell surface at all, and is retained in the ER from which it is degraded. It is possible that the rapid and efficient degradation results in complete lack of cell-surface expression of the ΔF508 protein, and therefore contributes to the pathogenesis of the disease.

DRUG DEVELOPMENT

Because of the central role the ubiquitin system plays in such a broad array of basic cellular processes, development of drugs that modulate the activity of the system may be difficult. Inhibition of enzymes common to the entire pathway, such as the proteasome, may affect many processes nonspecifically, although a narrow window between beneficial effects and toxicity can be identified

for a short-term treatment. Recent experimental evidence strongly suggests that such inhibitors may indeed be beneficial in certain pathologies, such as in multiple myeloma, a malignancy that affects the bone marrow.

A completely different approach to drug development can be the development of small molecules that interfere with the activity of specific E3s. For example, peptides that bind specifically to HPV-E6 can prevent its association with p53 and interfere with its targeting by E6-AP. Such peptides were able to induce p53 in HPV-transformed cells with subsequent reversal of certain malignant characteristics and induction of apoptosis.

SEE ALSO THE FOLLOWING ARTICLES

26S Proteasome, Structure and Function • Ornithine Cycle • Proteasomes, Overview • Protein Degradation • SUMO Modification • Ubiquitin-Like Proteins

GLOSSARY

proteolysis/degradation Hydrolysis of a protein which is a hetero-polymer of amino acids to short peptides, which contain a few amino acids, and/or to single amino acids.

ubiquitination Covalent modification of a protein by the small protein ubiquitin.

FURTHER READING

Glickman, M. H., and Ciechanover, A. (2002). The ubiquitin–proteasome proteolytic pathway: Destruction for the sake of construction. *Physiol. Rev.* **82**, 373–428.

Goldberg, A. L., Elledge, S. J., and Wade, J. (2001). The cellular chamber of doom. *Sci. Am.* January, 68–73.

Hilt, W., and Wolf, D. H. (eds) (2000). *Proteasomes: The World of Regulatory Proteolysis.* Eurekah.com/LANDES BIOSCIENCE Publishing Company, Georgetown, Texas, USA.

Weissman, A. M. (2001). Themes and variations on ubiquitylation. *Nat. Rev. Cell Mol. Biol.* **2**, 169–179.

BIOGRAPHY

Aaron Ciechanover is a Distinguished Professor in the Cancer and Vascular Research Center in the Faculty of Medicine at the Technion, Haifa, Israel. His principal research interests are regulation of transcriptional factors by the ubiquitin system. He holds an M.D. from "Hadassah" Medical School of the Hebrew University in Jerusalem, and a Ph.D. in biochemistry from the Technion. While a graduate student of Dr. Avram Hershko, the two discovered the principles and machinery of ubiquitination as a degradation signal.

Michael Glickman is an Associate Professor in the Department of Biology at the Technion. His principal research interest is mechanisms of proteolysis by the proteasome, and proteasome composition. He holds a Ph.D. in chemistry from the University of California at Berkeley. As a research fellow with Dr. Dan Finley at Harvard Medical School, he uncovered the basic structure of the regulatory complex of the proteasome.

Ubiquitin-Like Proteins

Edward T. H. Yeh
The University of Texas, Houston, Texas, USA

Ubiquitin is a small polypeptide that covalently attaches to other proteins, which marks them for degradation by the proteasome. Ubiquitin regulates many important cellular processes, such as signal transduction, cell-cycle progression, and the transformation of normal to malignant cells. Several ubiquitin-like proteins, such as sentrin (also called SUMO) and NEDD8/Rub1 have also been discovered. Similar to ubiquitin, these ubiquitin-like proteins also modify other proteins. However, they do not trigger proteasomal degradation. Sentrin, in a process called sentrinization, usually produces a change of the target protein's cellular localization and function. NEDD8, with cullin as its substrate, is a key component of the ubiquitin ligase complex. Thus, NEDD8, in a process called neddylation, indirectly regulates the ubiquitin/proteasome system. Sentrinization and neddylation require unique enzymes – E1, E2, and E3 – that are distinct from those involved in the ubiquitin pathway. Furthermore, both sentrinization and neddylation can be regulated by proteases that specifically remove sentrin or NEDD8 from their cellular targets.

Sentrin

There are three sentrins. Sentrin-1 is a 101-amino-acid protein containing a ubiquitin homology domain (residues 22–97) that is 48% homologous to human ubiquitin (see Figure 1). Sentrin-2 is a 95-amino-acid polypeptide that is 66% similar to sentrin-1 in its ubiquitin homology domain. Sentrin-3 is a 103-amino-acid polypeptide that is 97% identical to sentrin-2 in its ubiquitin homology domain. All sentrins have distinct N-terminal amino-acid sequences and C-terminal extensions. In addition, the Gly–Gly residues required for sentrinization are conserved in all sentrins (Figure 1).

ACTIVATION, CONJUGATION, AND LIGATION OF SENTRINS

The amino acids following the conserved Gly–Gly residues in all sentrins are cleaved by C-terminal hydrolases. The resultant sentrin monomer is then transferred to a specific E1 complex (Figure 2). The activating enzyme (E1) for sentrin is composed of two subunits, AOS1 and UBA2. The human AOS1 protein is 56% similar to the N-terminal half of the ubiquitin E1. The human UBA2 protein, which is homologous to the C-terminal half of the ubiquitin E1, contains the active-site cysteine residue required for the formation of a thiol ester linkage with sentrins. The activated sentrin molecule attached to E1 is then transferred to a carrier protein, E2.

In contrast to the large number of E2s involved in ubiquitination, sentrin utilizes only UBC9 as the carrier protein. This exclusive usage of UBC9 is probably explained by the existence of a 5-residue insertion that forms an exposed tight β-hairpin and a 2-residue insertion that forms a bulge in a loop close to the active site of UBC9. The surface of UBC9 involved in sentrin-1 binding is positively charged, whereas the corresponding surface in sentrin-1 is negatively charged. Moreover, the face of UBC9 remote from the catalytic site possesses a distinct electrostatic potential that may account for the propensity of UBC9 to interact with other proteins.

There are at least three different E3 involved in sentrinization. RanBP2, a protein that has been localized in the nuclear pore complex, serves as the E3 for Sp100 and histone deacetylase 4. RanBP2 also binds to UBC9, suggesting that sentrinization could take place at the nuclear pore complex. Another E3, PIAS (protein inhibitor of activated STAT), promotes sentrinization of the androgen receptor, c-JUN, p53, and Sp3. Finally, the human polycomb group protein, Pc2, functions as an E3 for a generalized transcription repressor, CtBP.

PROTEINS MODIFIED BY SENTRIN

The list of sentrinized proteins has been expanding rapidly. Three examples that have been characterized are discussed here, viz., PML, RanGAP1, and IκBα.

PML, a RING finger protein with tumor suppressor activity, has been implicated in the pathogenesis of acute promyelocytic leukemia, which arises following a reciprocal chromosomal translocation that fuses the PML gene with the retinoic acid receptor α (RARα) gene. PML can be modified by all sentrins, but not by ubiquitin or NEDD8. Two forms of PML-RARα fusion proteins have been found to be expressed in acute

```
Sentrin-1(1-50)  MSD----QEAKPST  EDLGDKKEGE  -YIKLKVIGQD  SSEIHFKVKM  TTHLKKLKES
Sentrin-2(1-46)  MAD-----E-KPK-  E--GVKTENN  DHINLKVAGQD  GSVVQFKIKR  HTPLSKLMKA
Sentrin-3(1-45)  MSE-----E-KPK-  E--GVKTEN-  DHINLKVAGQD  GSVVQFKIKR  HTSLSKLMKA
NEDD8(1-29)                                  MLIKVKTLT  GKEIEIDIEP  TDKVERIKER
Ubiquitin(1-29)                              MQIFVKTLT  GKTITLEVEP  SDTIENVKAK

Sentrin-1(51-101)  YCQRQGVPMN  SLRFLFEGQR  IADNHTPKEL  GMEEEDVIEV  YQEQTGGHSTV
Sentrin-2(47-95)   YCERQGLSMR  QIRFRFDGQP  INETDTPAQL  EMEDEDTIDV  FQQQTGGVY
Sentrin-3(46-103)  YCERQGLSMR  QIRFRFDGQP  INETDTPAQL  RMEDEDTIDV  FQQQTGGVPESSLAGHSF
NEDD8(30-81)       VEEKEGIPPQ  QQRLIYSGKQ  MNDEKTAADY  KILGGSVLHL  VLALRGGGGLRQ
Ubiquitin(30-76)   IQDKEGIPPD  QQRLIFAGKQ  LEDGRTLSDY  NIQKESTLHL  VLRLRGG
```

FIGURE 1 Alignment of sentrin-1, sentrin-2, sentrin-3, NEDD8, and ubiquitin. The conserved C-terminal glycine–glycine residues were printed in bold.

promyelocytic leukemia. Remarkably, neither can be sentrinized *in vivo*. Extensive mutational analysis of PML has shown that Lys65 in the RING finger domain, Lys160 in the B1 Box, and Lys490 in the nuclear localization signal constitute three major sentrinization sites. The PML mutant with Lys to Arg substitutions in all three sites is expressed normally but cannot be sentrinized. It also cannot recruit Daxx to the nuclear body. The lack of sentrinization of PML eliminates PML's transcriptional coactivator activity.

RanGAP1, a major regulator of the Ras-like GTPase Ran, which plays a critical role in the bidirectional transport of proteins and ribonucleoproteins across the nuclear pore complex was the first protein shown to be modified by sentrin-1. The 90 kDa sentrinized form of RanGAP1 associates with RanBP2 of the nuclear pore complex, whereas the 70 kDa unmodified form of RanGAP1 is exclusively cytoplasmic. Thus, sentrinization is critical for RanGAP1 to translocate to the nuclear pore. Interestingly, RanGAP1 cannot be modified by sentrin-2/3.

Although most sentrinized proteins are localized in the nucleus or associated with the nuclear envelope, a cytosolic protein, IκBα, can also be modified by sentrin-1.
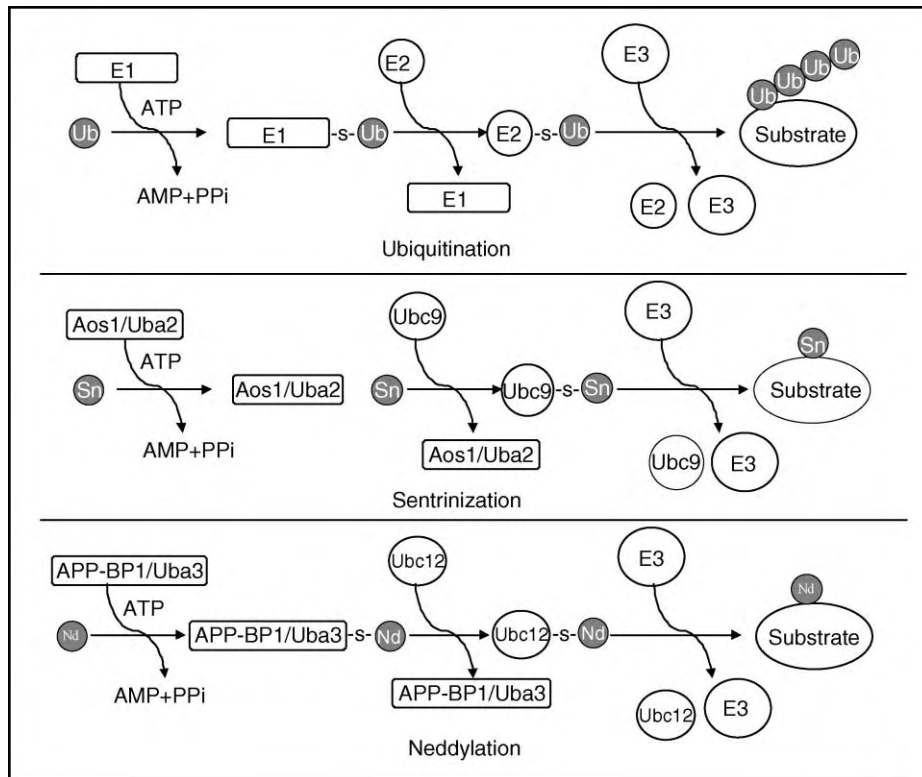
FIGURE 2 Schematic diagram of ubiquitination, sentrinization, and neddylation. E1: activating enzyme; E2: conjugating enzyme; E3: ligases. The E1 for the sentrinization pathway consists of Aos1 and Uba2, whereas the E1 for the neddylation pathway consists of APP-BP1 and Uba3. The E2 for the sentrinization pathway is Ubc9, whereas the E2 for the neddylation pathway is Ubc12. There are multiple E2s for the ubiquitination pathway. Ub: ubiquitin; Sn: sentrin: Nd: NEDD8.

IκBα is a cytosolic inhibitor of NFκB, a transcription factor involved in the induction of a large number of proteins involved in inflammation. IκBα is phosphorylated on serine residues 32 and 36 following tumor necrosis factor stimulation. The phosphorylated form of IκBα is then polyubiquitinated on lysine residues 21 and 22 and degraded by the proteasome. Unlike ubiquitination, sentrinization of IκBα on lysine residue 21 is inhibited by prior phosphorylation. Furthermore, sentrinized IκBα cannot be ubiquitinated and is resistant to proteasomal degradation. Thus, sentrinization can compete with ubiquitination for the same Lys residue on target proteins, effectively constitutes an antiubiquitin process.

By comparing the amino-acid sequences surrounding the acceptor lysine residues, a sentrinization consensus sequence can be formulated. It appears that the acceptor Lys residue is preceded by a hydrophobic amino acid (Ile or Leu) and followed by a polar amino acid, such as Gln, Thr, and a charged amino acid, Glu. The general consensus formula is (I/L)K(Q/T)E.

DE-SENTRINIZATION

In yeast, there are two sentrin-specific proteases (SENPs) that deconjugate sentrin from its substrate. Ulp1 possesses both isopeptidase and C-terminal hydrolase activity and is essential for the transition from G2 to the M phase of the cell cycle, the same phenotype observed for the Ubc9 mutation in yeast. These observations suggest that sentrinization and de-sentrinization are required for cell-cycle progression in yeast. Ulp2, on the other hand, is not essential for cell viability. The Ulp2 mutant exhibits a pleiotropic phenotype that includes temperature-sensitive growth, abnormal cell morphology, and decreased plasmid and chromosome stability. The mutant is also hypersensitive to DNA-damaging agents, such as hydroxyurea. Ulp1 is localized in the nuclear envelope, whereas Ulp2 is in the nucleus.

In humans, there are three well-characterized SENPs. All human SENPs have conserved C-terminal region, even though they vary in size. The similarity between yeast Ulp1, Ulp2, and human SENPs is confined primarily to the C-terminal region of $\sim$200 amino acids, within which a $\sim$90-residue segment forms a core structure common to cysteine proteases. It appears that the conserved C termini of the SENPs contain the catalytic domain, whereas the N termini regulate their cellular localization and substrate specificity. SENP1 is localized in the nucleus, excluding the nucleolus; SENP2 is localized in the nuclear envelope and also forms nuclear speckles; and SENP3 is localized predominantly in the nucleolus. Both SENP1 and SENP2 can deconjugate sentrinized PML but not RanGAP1. In addition, SENP1 and SENP2 can deconjugate many sentrinized

nuclear proteins. However, SENP3 appears to have more restricted substrate specificity.

In summary, the sentrin system possesses its own E1, E2, E3, and deconjugating enzymes, which serve to regulate an increasingly recognized cellular processes.

NEDD8/Rub1

NEDD8 is a small protein consisting of 81 amino acids that is 80% homologous to ubiquitin (see Figure 1). The yeast and plant homologues of NEDD8 have been termed Rub1 (*r*elated to *ub*iquitin *1*) and are processed and conjugated to a small number of cellular proteins. The overall structures of NEDD8 and Rub1 are quite similar to those of ubiquitin. The major difference occurs in two surface regions with a different electrostatic surface potential. Thus, unlike sentrin-1, NEDD8 is structurally more closely related to ubiquitin.

ACTIVATION AND CONJUGATION OF NEDD8 AND RUB1

UCH-L3, one of the ubiquitin C-terminal hydrolases, also binds to NEDD8 and functions as a NEDD8 C-terminal hydrolase. Following processing of its C terminus, NEDD8 is transferred to its E1.

The E1 for human NEDD8 is composed of two subunits, APP-BP1 and UBA3. The human APP-BP1 protein is 56% similar to the N-terminal half of ubiquitin's E1. The human UBA3 protein, which is homologous to the C-terminal half of ubiquitin's E1, contains the active-site cysteine residue required for the formation of thiol ester linkage with NEDD8. A similar E1 has also been identified in yeast and plants. The counterpart of APP-BP1 has been named AXR1 in plant. The importance of the Rub1 conjugation system is highlighted by the requirement of the AXR1 protein for the normal response to the plant hormone auxin in *Arabidopsis thaliana*. The specificity of NEDD8's E1 is due, in part, to the preferential binding of UBA3 to NEDD8, but not to ubiquitin or sentrin-1. The conjugating enzyme for NEDD8 or Rub1 is UBC12. A dominant-negative form of UBC12 could abolish NEDD-8 conjugation *in vivo* and inhibit cell growth.

RUB1/NEDD8 MODIFICATION OF CDC53/CULLINS

The major substrate for Rub1 in yeast is Cdc53 (also known as yeast cullin), a 94 kDa protein required for the G_1 to S phase of the cell-cycle progression. Cdc53 is a common subunit of the SCF complex, a ubiquitin ligase (E3) composed of *S*kp1, *C*dc53/cullin, and an *F* box

protein. There are many F-box proteins, which play a critical role in conferring substrate specificity to the SCF complex. For example, the F-box protein, Cdc4, forms a SCF^{Cdc4} complex with Skip1 and Cdc53/cullin that binds to phosphorylated Sic1 and catalyzes its ubiquitination.

In mammalian cells, NEDD8 modifies a limited number of human cellular proteins that primarily localized in the nucleus. Interestingly, all of the known NEDD8 targets in mammalian cells are cullins. Human Cul-1 is a major component of the SCF complex that is involved in the degradation of IκBα, β-catenin, and p27. The neddylation of Cul-1 dissociates CAND1, an inhibitor of Cul-1–SKP1 complex formation. The suppression of CAND1 therefore increases the level of the CuL-1–SKP1 complex. Thus, the neddylation of Cul-1 enhances the formation of the SCF complex, which in turn stimulates protein polyubiquitination. Human Cul-2 binds to the von Hippel-Lindau gene product through elongin B and elongin C to form a complex (Cul-2–VBC complex) that appears to have ubiquitin ligase activity. The VBC complex itself promotes, but is not essential for, NEDD8 conjugation to Cullin-2. Human Cul-3 was initially identified as a salicylate suppressible protein with unknown mechanism of activation, and was recently shown to be involved in the ubiquitination of cyclin E to control the S phase in mammalian cells. Human Cul-4A associates with UV-damaged DNA-binding protein and may play a role in DNA repair. Thus, the cullins may constitute a family of ubiquitin E3 ligases that have diverse biologic functions, with the neddylation of cullins playing a crucial role in ligase activity.

REGULATION OF NEDD8-CONJUGATES

There are several proteins that regulate NEDD8 conjugates. USP21 is a ubiquitin-deconjugating enzyme that is also capable of removing NEDD8 from NEDD8 conjugates. The overexpression of USP21 profoundly inhibits growth of U2OS cells. DEN1 is a specific protease for NEDD8 conjugates, as it has no activity against ubiquitin or sentrin conjugates. DEN1 can de-neddylate NEDD8 conjugated Cul-1. Interestingly, DEN1 is highly homologous to the catalytic domain of SENP. The COP9 signalosome (CSN) is an evolutionarily conserved multiprotein complex composed of eight subunits. It was first identified as an essential component in the repression of light-regulated development in *Arabidopsis*. CSN interacts with the SCF-type E3 ubiquitin ligases containing Cul1, which suggests that CSN and SCF-type E3 ubiquitin ligases are closely related. Indeed, CSN promotes the cleavage of NEDD8 from neddylated Cul1. The Jab1/MPN domain metalloenzyme motif in Csn5 subunit is responsible for the de-neddylation activity of CSN. Thus, CSN binds to SCF-type E3 ubiquitin ligases and probably regulates their activity through the de-neddylation of conjugated Cul-1. NUB1, an interferon-inducible protein, also regulates the neddylation system. The overexpression of NUB1 leads to the disappearance of most neddylated proteins. NUB1 appears to act by recruiting NEDD8 and its conjugates to the proteasome for degradation, further highlighting the close functional tie between neddylation and ubiquitination.

SEE ALSO THE FOLLOWING ARTICLES

Proteasomes, Overview • SUMO Modification • Ubiquitin System

GLOSSARY

neddylation The process of modifying a substrate with NEDD8.
proteasome A cellular organelle specialized in degrading ubiquitin-conjugated proteins.
sentrinization The process of modifying a substrate with sentrin/SUMO.
ubiquitination The process of modifying a substrate with ubiquitin.

FURTHER READING

Hay, R. T. (2001). Protein modification by SUMO. *Trends Biochem. Sci.* **26**, 332–333.
Melchior, F. (2000). SUMO – nonclassical ubiquitin. *Annu. Rev. Cell Dev. Biol.* **16**, 591–626.
Yeh, E. T. H., Gong, L., and Kamitani, T. (2000). Ubiquitin-like proteins: New wines in new bottles. *Gene* **248**, 1–14.

BIOGRAPHY

Edward T.H. Yeh is Chairman of the Department of Cardiology at the University of Texas M.D. Anderson Cancer Center and Director of the Research Center for Cardiovascular Diseases at the Institute of Molecular Medicine at the University of Texas-Houston Health Science Center. He holds an M.D. from the University of California, Davis and received his postdoctoral training at Harvard University. His laboratory was among the first to show that sentrin and NEDD8 can be covalently conjugated to other proteins.

UmuC, D Lesion Bypass DNA Polymerase V

Zvi Livneh

Weizmann Institute of Science, Rehovot, Israel

DNA polymerase V from the bacterium *Escherichia coli* (*E. coli*) is specialized to perform DNA synthesis across DNA lesions. The latter are sites in DNA that were chemically damaged by radiation or chemicals, and which usually block DNA replication. The reaction catalyzed by pol V is called translesion synthesis (TLS), translesion replication (TLR), or lesion bypass. It is usually mutagenic, since the coding properties of damaged nucleotides in DNA are often different from those of the original undamaged nucleotides. Pol V is a heterotrimer composed of one molecule of the UmuC protein (48 kDa), and two molecules of the UmuD′ protein (12 kDa each). UmuD′ is a proteolytic cleavage product of a larger protein, termed UmuD (15 kDa). The UmuD and UmuC proteins are encoded by an operon that is induced by DNA-damaging agents, under regulation of the SOS response. Pol V is a low-fidelity DNA polymerase, which does not have an exonuclease proofreading activity, and exhibits low processivity. Its lesion-bypass activity requires two additional proteins, RecA, which is also the main recombinase in *E. coli*, and single-strand DNA-binding protein (SSB), a protein essential for replication. In addition, pol V is stimulated by the processivity proteins of pol III, namely, the *β*-subunit DNA sliding clamp, and the *γ*-complex clamp loader. By virtue of its mutagenic lesion-bypass activity, pol V is responsible for the formation of mutations by a variety of DNA-damaging agents that cause replication blocks, such as UV radiation, methymethanesulphonate (MMS), and more. Pol V homologues were conserved in evolution from *E. coli* to humans, and comprise the Y family of DNA polymerases.

Pol V and the SOS Response

The SOS response is a global stress response in *E. coli*, induced by DNA damage that causes interruption of DNA replication. It involves the induction of 30–40 genes, which collectively act to increase cell survival under adverse environmental conditions. SOS genes comprise a regulon, whose members are negatively regulated at the transcriptional level by the LexA repressor. Induction involves autocleavage of free LexA repressor, promoted by an activated form of the RecA.

This cleavage shifts the binding equilibrium of LexA, causing it to dissociate from its target genes, thereby causing induction. RecA is usually activated by forming a complex of multiple RecA molecules bound to a stretch of single-stranded DNA (ssDNA), which is formed as a result of exposure to damaging agents. This RecA nucleoprotein filament has multiple roles in cellular responses to DNA damage. In addition to the induction of the SOS regulon, it promotes homologous recombination repair, and it is essential for lesion bypass by pol V. The *umuD* and *umuC* genes, encoding the two subunits of pol V, are arranged in an operon that is part of the SOS regulon (namely, its transcription is regulated by LexA and RecA), and are therefore inducible by DNA-damaging agents such as UV radiation or MMS. Mutations that inactivate *umuC* or *umuD* cause a drastic decrease in UV- or MMS-mutagenesis, with a minor effect on cell survival. This demonstrated the surprising phenomenon that mutagenesis caused by DNA-damaging agents is an inducible process, which requires specific gene products. In other words, mutagenesis is not merely a by-product of replication, and the presence of lesions in DNA does not cause mutations without the activity of specific inducible proteins.

The Discovery of Pol V

It took nearly two decades from the time the *umuC* and *umuD* genes were cloned, until the discovery that they encode a specialized lesion-bypass DNA polymerase. This long period of time is attributed to the great difficulty in obtaining UmuC in soluble and active form, the lack of any amino acid homology to classical DNA polymerases, and the complexity of the *in vitro* lesion-bypass assay system. In addition, genetic experiments had indicated that the *polC (dnaE)*, encoding the *α* subunit of pol III, was required for lesion bypass *in vivo*. This led to the incorrect model that the UmuC and UmuD′ proteins were accessory proteins, which enabled pol III to bypass DNA lesions. Finally, UmuCD′-dependent lesion bypass systems were reconstituted

from purified components, and led to the finding that UmuC has DNA polymerase activity, and together with UmuD' it forms pol V. This discovery was made simultaneously and independently in the laboratories of Myron Goodman (University of Southern California, Los Angeles) and Zvi Livneh (Weizmann Institute of Science, Rehovot, Israel).

Model for the Mechanism of Lesion Bypass by Pol V

Pol V is unable to bypass lesions on its own. It requires two additional proteins, RecA and SSB, known to be involved in other DNA transactions such as replication and recombination. In addition, pol V is stimulated by the processivity proteins, which are also required for DNA replication. The RecA protein is required as a nucleoprotein complex with ssDNA, in which multiple RecA monomers form a protective helical sleeve around the DNA (Figure 1, stage 1). The RecA filament prevents nonproductive binding of pol V to ssDNA regions

1. Formation of a RecA-ssDNA nucleoprotein filament

UmuC **Pol V**
RecA **UmuD'**

2. Targeting of pol V to the primer terminus

SSB

3. Loading of the β-subunit DNA sliding clamp

β -subunit

4. Lesion bypass

M
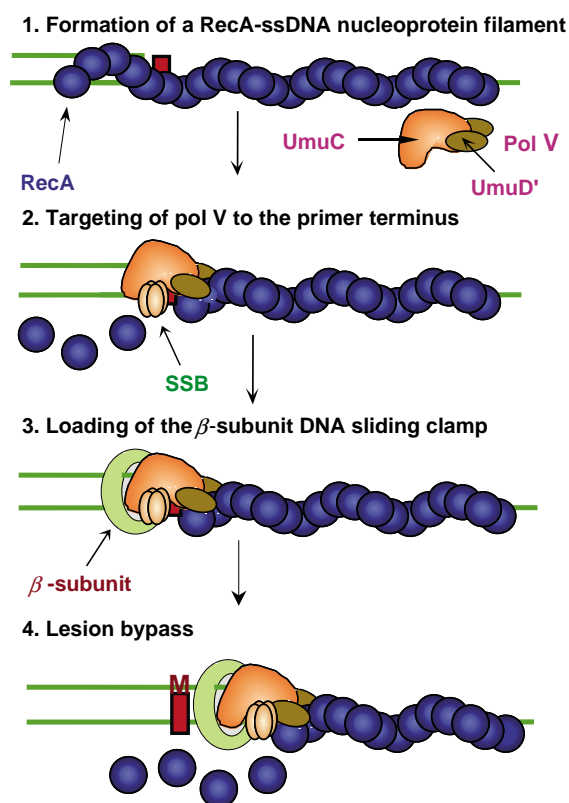
FIGURE 1 Outline of the mechanism of lesion bypass by pol V. The red square represents a blocking lesion. The "M" represents an incorrect insertion opposite the lesion. Reproduced from Livneh, Z. (2001). DNA damage control by novel DNA polymerases: Translesion replication and mutagenesis. *J. Biol. Chem.* **276**, 25639–25642, with permission of the American Society for Biochemistry and Molecular Biology.

remote to the primer terminus, and it targets pol V to the primer-template region (Figure 1, stage 2). Once targeted to the primer-terminus site, the RecA filament needs to locally dissociate, to allow productive binding of pol V. This requires, in addition to pol V, also the SSB protein. The pol V initiation complex is stabilized by RecA–UmuD' interactions, UmuC–DNA interactions, and perhaps interactions with SSB. Once this pol V complex has formed an initiation complex, DNA synthesis can occur. However, to get full pol V holoenzyme activity, the processivity proteins are needed, including the β-subunit DNA sliding clamp, and the γ-complex clamp loader (Figure 1, stages 3, 4). These proteins are essential for highly processive DNA replication by DNA polymerase III; however, they also increase the processivity of DNA polymerases II, IV, and V, although to a much lesser extent than pol III. DNA synthesis by pol V holoenzyme causes dissociation of RecA monomers, as the polymerase progresses. Pol V is a major lesion-bypass machine, which can bypass a wide variety of DNA lesions. This includes abasic sites, and the two major lesions caused by UV light, a cyclobutyl thymine–thymine dimer, and a thymine–thymine 6–4 adduct. Most remarkably, pol V is able to bypass a stretch of 12 methylene residues inserted into the DNA. This artificial lesion is an extreme form of DNA damage, since it bears no resemblance to any of the regular DNA constitutents. Despite this remarkable bypass ability, pol V is poor in bypassing some DNA lesions, most notably benzo[a]pyrene–guanine adducts. The ability of pol V to bypass DNA lesions correlates with its low fidelity. However, low fidelity by itself is not sufficient to ensure effective lesion bypass, since there are enzymes with fidelity similar to pol V, which nevertheless are not lesion-bypass enzymes (e.g., mammalian DNA polymerase β). The crystal structure of pol V was not determined yet. However, based on several crystal structures of other homologous lesion-bypass DNA polymerases, a spacious and flexible active site is likely to be responsible, at least in part, for the ability to bypass DNA lesions.

The Fidelity of Pol V

The fidelity of DNA synthesis by pol V undamaged DNA templates is 10^{-3}–10^{-4} errors/nucleotide replicated. This is 100–1000-fold lower than the fidelity of the replicative pol III holoenzyme. DNA sequence analysis of pol V errors showed that pol V produces all types of errors, including frameshifts and base substitutions. However, it has a propensity to produce transversion mutations, namely, changing a pyrimidine into a purine, or vice versa. Pol V forms transversions at a frequency up to 300-fold higher than pol III holoenzyme (which tends to form transition mutations; Table I). It should be

TABLE I

Mutations Produced during *In Vitro* DNA Synthesis by Pol V and Pol III Holoenzyme

Mutation	Mismatch[a]	Mutation frequency $\times 10^{-5}$/gene		Pol V/pol III
		Pol III	Pol V	
Transition				
T → C	T:G	<1.4	368.3	>263
Transversion				
A → C	A:G	2.8	161.1	58
A → T	A:A	<1.4	414.4	>296
C → A	C:T	<1.4	92.1	>66
T → A	T:T	2.8	230.2	82
T → G	T:C	<1.4	69.1	>49

[a] The mutations were formed in the *cro* repressor gene of phage λ during gap-filling DNA synthesis.

Reproduced from Maor-Shoshani, A., Reuven, N. B., Tomer, G., and Liuneb, Z. (2000). Highly mutagenic replication by DNA polymerase V (UmuC) provides a mechanistic basis for SOS untargeted mutagenesis. *Proc. Natl Acad. Sci. USA* **97**, 565–570, with permission of the National Academy of Sciences, U.S.A.

noted that base–base mismatches produced by DNA polymerases during synthesis on undamaged DNA are potential substrates for the mismatch repair system. This system primarily corrects mismatches that are precursors for transition mutations, namely, the types of mismatches that the replicative pol III holoenzyme tends to make. This significantly reduces mutations that can occur during DNA replication, and contributes to the high replication fidelity. The repair by the mismatch repair system of mismatches that are precursors of transversions, such as pol V tends to make, is 20-fold less effective than transition mismatches. This implies that pol V generates mismatches that are largely immune to mismatch repair and yield mutations. This provides the mechanistic basis for the phenomenon of untargeted SOS mutagenesis, also termed SOS mutator activity. In essence, this phenomenon involves the formation of mutations in undamaged regions of DNA in cells in which the SOS response was induced. This process is *umuDC*- and *recA*- dependent, and primarily yields transversion mutations. Based on the fidelity of pol V, untargeted mutagenesis can be explained by the activity of pol V under SOS conditions in undamaged regions of the *E. coli* chromosome. It should be noted that there is a second branch of untargeted mutagenesis that depends on pol IV rather than pol V.

Regulation of Pol V

The activity of pol V is tightly regulated at several levels, to ensure that this mutator polymerase acts only under situations, and at sites, when it is really needed. As described above, the *umuDC* operon is regulated at the transcriptional level by the LexA repressor and the RecA activator like other SOS genes. Within the timed expression of SOS genes, the *umuDC* operon is induced late, nearly an hour after the onset of induction. When SOS genes become repressed again, as the SOS response is shut down, the *umuDC* operon is repressed early. This late induction, early shutdown expression pattern provides an extra layer of regulation on pol V. It limits the activity of pol V to a stage after error-free repair mechanisms had a chance to repair DNA, and even at that late stage in SOS, the activity of pol V is limited in time, as it is shut off early. Additional regulation is at the posttranslational level. The UmuD$_2$ protein is cleaved to yield a shorter protein termed UmuD$_2'$, which is the active protein in pol V. The proteolytic processing occurs via the self-cleavage of the UmuD$_2$ protein, promoted by activated RecA, in a mechanism that is similar to the cleavage of the LexA repressor. Finally, there is regulation by turnover, as UmuD is degraded by the Lon and ClpXP proteases. The steady-state level of UmuC in noninduced cells is low and undetected by currently available antibodies. Upon SOS induction, UmuC accumulates to a level of ~200 molecules/cell. UmuD is present in higher amounts, starting with 200 molecules in the noninduced cell, and going up to 2400 molecules/cell after induction. Finally, pol V is regulated at the activity level, in its requirement for a RecA nucleoprotein filament. Such filaments are formed in DNA only when replication is arrested (e.g., at a lesion) and a persistent single-stranded region is generated.

In vivo Function of Pol V

The major phenotype of *E. coli* cells lacking pol V is a strong decrease in mutations caused by DNA-damaging agents such as UV radiation or MMS. Survival is decreased too, but to a lower degree compared to defects in other DNA repair pathways, such as excision repair. Based on these phenotypes, there are two views on the *in vivo* role of pol V. One is that pol V repairs replication gaps and this leads to increased survival. According to this model, the mutations formed by pol V are a by-product of the gap-filling activity, and can be viewed as the price paid for this gap repair (repair at the expense of mutations). The other view portrays the generation of mutations as the main role of pol V. According to this view, when a population of *E. coli* cells is under hostile environmental conditions, e.g., lack of nutrients, or presence of DNA-damaging agents such as UV light, it increases its mutations rate via pol V, to facilitate adaptation and increase fitness. Even one successful mutation in a cell (e.g., conferring higher resistance to sunlight) may have a great impact on the

entire population, since in time these cells, which have a selective advantage, may proliferate to a state where they take over the population. Consistent with such an idea is the finding that when mutator strains are grown for many generations along with nonmutator strains, the mutator strain takes over. Similarly, *umuDC*-deficient mutants suffer a severe reduction in fitness when competing against cells with normal pol V. Since mutations form randomly in relation to their functional outcome, a mutator acting long enough will start accumulating deleterious mutations, and will lose its adaptive advantage. The fact that the activity of pol V is transient (namely, as long as the SOS response is on) helps to preserve successful mutations from being neutralized by the formation of other mutations. It was also suggested that pol V initially evolved as a generic bypass DNA polymerse, which functioned to increase survival by overcoming replication blocks, but as more sophisticated DNA repair mechanisms evolved, the importance of the bypass function decreased, and what kept pol V during evolution was its mutator effect. Of course, the various hypotheses are not mutually exclusive, and the *in vivo* role of pol V may be both in survival and mutagenesis.

Bacterial Homologues of Pol V

E. coli contains a homologue of pol V, termed pol IV, which is the product of the SOS gene *dinB*. Like pol V, pol IV is an SOS-inducible, low-fidelity, and low-processivity DNA polymerase, which is able to bypass some DNA lesions. As in pol V, the processivity and lesion-bypass ability of pol IV are increased by the β-subunit clamp, and the γ-subunit clamp loader. However, unlike pol V, pol IV is a single subunit enzyme, and its bypass activity requires neither RecA nor SSB. On undamaged DNA, pol IV tends to form frameshifts, primarily minus-one deletions. The *in vivo* role of pol IV is not as clear as that of pol V. Pol IV is required for untargeted mutagenesis of phage λ. This mutagenic pathway is observed in nonirradiated phage λ when it infects a UV-irradiated *E. coli* host. In addition, pol IV is required for the bypass of a benzo[a]pyrene–G adducts. It is also required for stationary phase mutations.

Homologues of pol V exist in many, but not all bacteria. Pol V homologues were also found on native bacterial conjugative plasmids, such as R46. These are large plasmids, in the range of 100 kbp, with broad host-range specificity, which often carry multiple antibiotics-resistance markers. In the case of the R46 plasmid, the UmuD, UmuD′, and UmuC homologues are MucA, MucA′ and MucB, respectively. MucA′ and MucB were shown to form a lesion-bypass DNA polymerase, pol RI, with properties similar to those of pol V. Pol RI,

like pol V, required both RecA and SSB for bypassing an abasic site. MucA and MucB are present on plasmid pKM101, a natural deletion derivative of plasmid R46. Plasmid pKM101 was introduced into the *Salmonella* strains that are used in the Ames test for mutagens, in order to improve its mutagenic sensitivity.

The emergence of pathogenic bacteria that are resistant to multiple types of antibiotics is associated in part of the cases with the presence of native conjugative plasmids. The fact that native plasmids, which have a limited genome size, and rely heavily on host factors, carry mutator lesion-bypass DNA polymerase genes indicates that the latter have a special role in the life cycle of such plasmid. One possibility is that the mutator activity of the polymerases is required when the plasmids enter a new foreign bacterial host, and need rapid adaptation to the new intracellular environment.

Eukaryotic Homologues of Pol V

Homologues of pol V were found to be conserved in evolution, and comprise the Y family of DNA polymerases. The yeast *S. cerevisiae* contains two homologues of UmuC, pol η (product of the *RAD30* gene), and REV1, a G-template specific DNA polymerase. Human cells contain no less than four homologues of pol V: pol η, pol ι, pol κ, and REV1. The DNA polymerases of the Y family were discovered in 1999 by several investigators, simultaneously and independently. The *S. cerevisiae* pol η was reported first, discovered by Satya Prakash and Louis Prakash (University of Texas Medical Branch, Galveston). Both the *S. cerevisiae* and the human pol η have the remarkable property that they replicate across an unmodified DNA TT sequence, with similar efficiency and specificity to that of a TT cyclobutyl dimer. In the absence of pol η in the cell, another polymerase takes over, but performs the reaction with a lower fidelity, leading to increased UV mutagenesis. This is the situation in the human hereditary disease Xeroderma Pigmentosum Variant (XP-V). Unlike other forms of XP, which are defective in components of the error-free nucleotide excision repair, XP-V patients are deficient in DNA pol η (the XP-V gene product). XP-V patients lacking this enzyme show extreme sensitivity to sunlight and a high susceptibility to skin cancer. This led to the unexpected conclusion that lesion bypass may be functionally nonmutagenic, and that lesion-bypass DNA polymerases may function to protect mammals from at least certain types of cancer.

S. cervisiae has an additional lesion-bypass polymerase, pol ζ, encoded by the REV3 and REV7 genes. It is homologous to the replicative pol δ, and not to the Y-family DNA polymerases. Historically, pol ζ was the first lesion-bypass DNA polymerase, discovered in 1996

by David Hinkle and Christopher Lawrence (Rochester University), followed soon thereafter by the discovery, in the same year, and by the same investigators, that REV1 has nucleotidyl transferase activity. Human cells also contain additional polymerases, which can bypass lesions *in vitro*. This includes pol μ and pol λ, and homologues of the yeast *REV3* and *REV7* genes, encoding a putative pol ζ. Overall this suggests that lesion bypass has an important role in the response to DNA damage in mammalian cells.

SEE ALSO THE FOLLOWING ARTICLES

DNA Base Excision Repair • DNA Damage: Alkylation • DNA Polymerase I, Bacterial • DNA Polymerase II, Bacterial • DNA Polymerase III, Bacterial • Translesion DNA Polymerases, Eukaryotic

GLOSSARY

base substitution mutation A mutation in which a particular DNA base is changed to another base (e.g., A → G).

benzo[*a*]pyrene–guanine adduct A DNA adduct formed between benzo[*a*]pyrene, a major tobacco smoke carcinogen, and a guanine base. Benzo[*a*]pyrene requires metabolic activation before it can react with DNA.

excision repair The major mechanism of DNA repair. It involves excision of the damaged site from DNA, followed by gapfilling DNA synthesis, based on the complementary intact strand and ligation.

frameshift mutation A mutation that changes the reading frame of a gene. Frameshifts involve either deletion or insertion of one or two nucleotides. Larger deletions or insertions may also cause frameshift mutations, when the number of deleted or added nucleotides is not a multiple integer of three.

operon A segment of the genome with two or more genes which are cotranscribed yielding a multicstronic mRNA.

processivity of DNA polymerase The number of nucleotides polymerized by a DNA polymerase per single binding event to the primer-template. Processivity of DNA polymerases varies from 1 (distributive DNA polymerase) up to 10 000 (highly processive DNA polymerase). Usually replicative polymerases have very high processivity, whereas DNA repair polymerases have lower processivity.

regulon A set of genes, residing at different locations in the chromosome, that are regulated by a common regulatory pathway. The genes regulated by the SOS response in *E. coli* comprise a regulon.

transition mutation A mutation in which a pyrimidine is changed to another pyrimidine, or a purine is changed to another purine. Changes of A → G or C → T are examples of transition mutations.

transversion mutation A mutation in which a purine changes to a pyrimidine, or a pyrimidine changes into a purine. Changes of A → C or T → A are examples of transversion mutations.

FURTHER READING

Baynton, K., and Fuchs, R. P. (2000). Lesions in DNA: Hurdles for polymerases. *Trends Biochem. Sci.* **25**, 74–79.

Goodman, M. F. (2000). Coping with replication 'train wrecks' in *Escherichia coli* using Pol V, Pol II, and RecA proteins. *Trends Biochem. Sci.* **25**, 189–195.

Livneh, Z. (2001). DNA damage control by novel DNA polymerases: Translesion replication and mutagenesis. *J. Biol. Chem.* **276**, 25639–25642.

Radman, M. (1999). Enzymes of evolutionary change. *Nature* **401**, 866–869.

Sutton, M. D., Smith, B. T., Godoy, V. G., and Walker, G. C. (2000). The SOS response: Recent insights into umuDC-dependent mutagenesis and DNA damage tolerance. *Annu. Rev. Genet.* **34**, 479–497.

BIOGRAPHY

Zvi Livneh received his Ph.D. from the Weizmann Institute of Science in Israel, and was a postdoctoral fellow at Stanford University School of Medicine in California. He is a Professor of biochemistry and molecular medicine, and Head of the Department of Biological Chemistry at the Weizmann Institute of Science in Israel. His main research interests are molecular mechanisms of DNA repair and mutagenesis in bacteria and in mammals, and the role of DNA repair in cancer risk.

Uncoupling Proteins

Daniel Ricquier and Frédéric Bouillaud
Centre National de la Recherche Scientifique, Paris, France

Uncoupling proteins (UCPs) are mitochondrial transporters present in the inner membrane of mitochondria. They belong to the family of anion mitochondrial carriers including adenine nucleotide transporters, phosphate carrier and other transporters. The term uncoupling protein was originally given to UCP1 which is uniquely present in mitochondria of brown adipocytes, the thermogenic cells devoted to maintenance of body temperature in mammals. In these cells, UCP1 acts as a proton carrier creating a shunt between complexes of respiratory chain and the ATP-synthase. Purine nucleotide inhibit UCP1 whereas fatty acids activate it. Activation of UCP1 stimulate respiration and the uncoupling process results in a futile cycle and dissipation of oxidation energy into heat. UCP2 is ubiquitous and highly expressed in lymphoid system and macrophages. UCP3 is mainly expressed in skeletal muscles. In comparison to the established uncoupling and thermogenic activities of UCP1, UCP2, and UCP3 rather appear to be involved in the limitation of free radicals levels in cells than in physiological uncoupling and thermogenesis. The mechanism of the protonophoric activity of the UCPs is still controversial since it has been proposed that the UCPs transport fatty acid anions and catalyse proton transport by fatty acid cycling. Quinones and superoxide ions may also activate the UCPs.

The Mitochondria and the Coupling of Respiration to ADP Phosphorylation

Cellular respiration, the reactions of the citric acid cycle, fatty acid oxidation, and several steps of urea synthesis and gluconeogenesis take place in specialized cellular organites, the mitochondria.

Mitochondria

In addition to oxidative phosphorylation and metabolic pathways, mitochondria are involved in thermogenesis, radical production, calcium homeostasis, apoptosis, and protein synthesis. Mitochondria contain two compartments bounded by inner and outer membranes. The outer membrane is permeable to many small metabolites whereas the permeability of the inner membrane is controlled in order to maintain the high electrochemical gradient created by the mitochondrial respiratory chain which is necessary for energy conservation and ATP synthesis in mitochondria. The inner membrane transports anion substrates such as ADP, ATP, phosphate, oxoglutarate, citrate, glutamate, and malate.

Coupling of Respiration to ATP Synthesis

It has long been known that respiration and mitochondrial ATP synthesis are coupled. The observation that respiration rate increased when mitochondria synthesized more ATP led to the concept of respiratory control by ADP phosphorylation. In fact, there is a link between mitochondrial ATP synthesis and cellular ATP demand by a feed-back mechanism. In agreement with Mitchell's theory, it was demonstrated that the mitochondrial electrochemical proton gradient, generated as electrons are passed down the respiratory chain, is the primary source for cellular ATP synthesis. In this way, several complexes of the respiratory chain pump protons outside of the inner membrane during reoxidation of coenzymes and generate a proton gradient which is consumed by the reactions of ATP synthesis. Proton leak represents another mechanism consuming the mitochondrial proton gradient. Mitchell's theory predicted that any proton leak non coupled to ATP synthesis would represent an uncoupling of respiration and result in thermogenesis. An excellent example of such an uncoupling of respiration to ADP phosphorylation is represented by the mitochondrial UCP of brown adipocytes (UCP1) which dissipates energy of substrate oxidation as heat.

Brown Adipose Tissue and UCP1: History of a True Respiration Uncoupling

Maintenance of body temperature in a cold environment or at birth requires thermogenesis. Another situation

requiring thermogenesis is arousal from hibernation. The two regulatory thermogenic processes are shivering and metabolic thermogenesis also referred to as non-shivering thermogenesis. The largest part of nonshivering thermogenesis in small mammals is achieved in the brown adipose tissue (BAT).

BROWN ADIPOSE TISSUE (BAT)

BAT is found in all small mammals and in the newborn of larger mammals, such as humans. BAT is located in specific body areas near large blood vessels and consists of brown adipocytes which are distinct from white adipocytes. The brown adipocytes are characterized by numerous mitochondria containing a highly developed inner membrane (Figure 1). Activation of thermogenesis in BAT occurs in newborns, rodents exposed to cold, and in animals emerging from hibernation. It is commanded by the central nervous system and the orthosympathetic fibers innervating each brown adipocyte. The norepinephrine released by these fibers binds to adrenergic receptors on the surface of the brown adipocytes. The later steps of the activation of thermogenesis in brown

adipocytes are production of cyclic AMP, activation of lipolysis, and oxidation of fatty acids by the numerous mitochondria. Released fatty acids stimulate brown adipocyte respiration and heat production.

RESPIRATION UNCOUPLING IS THE THERMOGENIC MECHANISM IN BROWN ADIPOCYTES

In most cell types such as muscular fibers, ATP utilization stimulates ADP phosphorylation and respiration (Figure 2). Original observations made in 1967 revealed that the high respiratory rate in brown adipocyte mitochondria was not controlled by mitochondrial ADP phosphorylation suggesting that energy from substrate oxidation was dissipated into heat instead of being converted in ATP (Figure 2). In line with Mitchell's chemiosmotic theory, if proton leakage is activated by a signal such as free fatty acids in the case of the mitochondria of brown adipocytes, the mitochondrial membrane potential is decreased and concomitantly, respiration is activated. Since, in addition, the activated proton leakage is not linked to ADP phosphorylation, respiration is uncoupled from ATP synthesis and oxidation energy is dissipated in the form of heat.

FIGURE 1   Histology of brown adipocyte. (A) The cytosol of brown adipocytes is characterized by numerous mitochondria and lipid droplets. (B) Magnification of a brown adipocyte mitochondrion showing parallel cristae and UCP1 detected using antibodies (black dots). (Figure kindly provided by Dr Saverio Cinti, University of Ancona.)
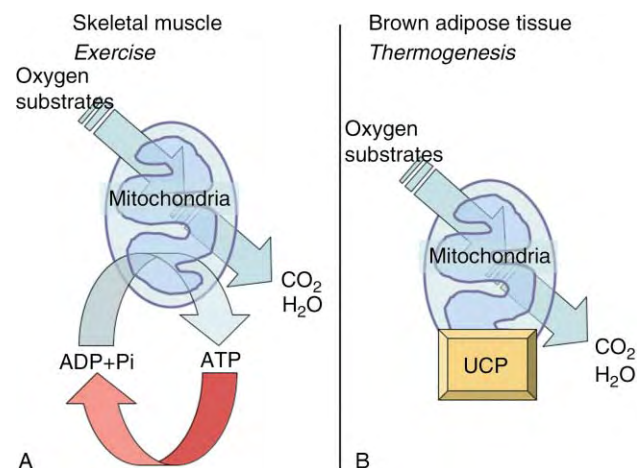
FIGURE 2   Comparison of skeletal muscle and brown adipose tissue mitochondria. (A) Mitochondria are cellular organelles converting the redox energy liberated during oxidation of organic substrates into ATP, a molecule containing energy under a form readily usable by most enzymes working to maintain cell structure and integrity or to perform work such as mechanical work in muscle. The cellular need for ATP controls oxidation rate by mitochondria. (B) In brown adipocyte, a specific uncoupling protein referred to as UCP (precisely UCP1) is present in a large amount in the inner membrane. UCP1 is able to fool mitochondria which accelerates their oxidation rate whereas no ATP is produced. UCP1 removes the control of energy expenditure by ATP demand and therefore oxidation rate and energy expenditure increase under the form of heat. This process is used in newborns, arousal from hibernation, and small mammals exposed to cold when muscle thermogenesis would not be sufficient.

The UCP of Brown Adipocyte Mitochondria

The protein responsible for proton leakage within the inner mitochondrial membrane and for respiratory uncoupling was identified in 1976–77, was purified, and its cDNA was cloned. It has an apparent molecular weight of 33 000 and was originally called UCP. This protein is known as UCP1 since the discovery of UCP2 in 1997. UCP1 is abundant in the inner membrane of the mitochondria of brown adipocytes and is specific to these cells. It short-circuits the proton circuit between complexes of the respiratory chain and the mitochondrial ATP-synthase which is the main consumer of the proton gradient (Figure 3). When there is no need for thermogenesis, purine nucleotides bound to UCP1 inhibit its activity. Upon activation of thermogenesis by the sympathetic nervous system, fatty acids overcome the inhibitory effect of nucleotides and activate UCP1. Heterologous expression of UCP1 in yeasts or mammalian cell lines induces respiration uncoupling. The protonophoric activity of UCP1 has been reconstituted in liposomes where it can be inhibited by nucleotides and activated by free fatty acids. The mechanism of action of UCP1 is still controversial: some scientists believe that it is a true proton transporter, whilst others assert that it
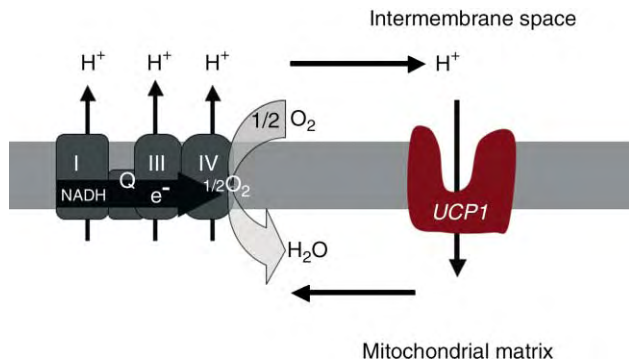
FIGURE 3 UCP1-induced proton cycling. Q (coenzyme Q) refers to complex II which is the succinate dehydrogenase. This enzyme directly reduces coenzyme Q but has no proton pumping activity contrary to complexes I, III and IV. UCP is inserted in the mitochondrial inner membrane where, also present, is a multienzymatic complex called the respiratory chain made of complexes I to IV. The respiratory chain reoxidize reduced coenzymes and electrons are driven to oxygen. This oxido-reduction step liberates energy which is used to generate an electrochemical gradient of protons across the inner membrane. This gradient is normally consumed by the ATP-synthase which phosphorylates ADP. UCP1 transports protons passively and makes possible a futile cycle of protons across the inner membrane leading to increased energy expenditure. This schema illustrates the situation encountered in brown adipocytes of mammals where a large amount of respiratory chains as well as a large amount of UCP1 are present. Activation of the futile cycling increases considerably energy expenditure and thus heat production in these thermogenic cells. In other cells where homologues of the UCP1 are expressed at much lower level, this pathway would represent a minor contributor to energy expenditure, but might be of importance to avoid oxidative damage.
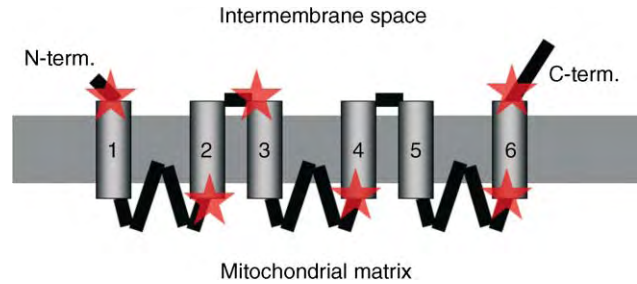
FIGURE 4 Model of insertion of UCP in the membrane. The determination of structure of the UCPs remains difficult because crystallization of such membrane proteins has not been successful yet. To modelize the organization of such proteins, indirect methods such as prediction of secondary structure by computers can be used. Experimentally, restricted domains of the protein in the membrane (indicated by stars) were explored using specific antibodies. The UCPs are made of six transmembraneous segments (probably alpha helices), and three hydrophilic loops stemming from the matricial side of the membrane. An internal repetition in the structural arrangement of the protein suggests that the protein (made of 300 amino acid residues) derived from the ancient triplication and divergence of a 100 amino acid domain (made of 2 helices and 1 loop).

returns anionic fatty acids to the intermembrane space, after they have crossed the membrane in protonated form. The observation that $Ucp1-/-$ mice were unable to maintain body temperature in the cold proved that the UCP1-induced uncoupling of respiration was responsible for cold-activated nonshivering thermogenesis of rodents. In agreement with this observation, ectopic expression of UCP1 in skeletal muscles of transgenic mice promotes substrate oxidation in muscles and resistance to obesity and type 2 diabetes. UCP1 functions as a dimer and each monomer is made of six transmembrane fragments (Figure 4). Moreover, UCP1, the ADP/ATP translocator and other mitochondrial anion carriers derive from the same ancestral gene and probably share a similar structure. They all have a triplicated structure (Figure 4).

The Novel UCPs

Given their specific role in thermogenesis, it has always seemed logical that brown adipocytes be equipped with an original mechanism, partial uncoupling of respiration, brought into play by a specific protein, UCP1, which induces proton leakage. In fact, it is known that mitochondrial respiration is always accompanied by heat production as it is imperfectly coupled to ADP phosphorylation in all types of cells. To explain this incomplete coupling of respiration and the energy loss mechanism, some authors have invoked slippage of the respiratory chains, while others referred to proton leaks. It has been estimated that proton leaks from the inner membrane of mitochondria of hepatocytes

and myocytes could explain about 20% of the body's basal metabolism.

UCP2: A HOMOLOGUE OF THE BROWN FAT UCP PRESENT IN VARIOUS TISSUES

A clone corresponding to a protein with 59% sequence identity with the UCP of brown fat was isolated. This new UCP, called UCP2, was also identified in humans. In contrast with UCP1, the mRNA of UCP2 is present in almost all tissues and many cell types, such as spleen, thymus, digestive tract, lung, brain, adipose tissues, skeletal muscle, heart, adipocytes, myocytes, and macrophages (Figure 5). Expressed in yeasts, murine UCP2 decreased the potential of the mitochondrial membrane, raised the respiration rate and reduced sensitivity to uncouplers: UCP2 therefore was able to uncouple respiration and appeared to be a second mitochondrial UCP. The UCP2 gene is located on chromosome 7 of the mouse and chromosome 11 of humans, near a region linked to hyperinsulism and obesity. The expression of UCP2 mRNA was measured in obesity-prone and obesity-resistant mice given a lipid-rich diet: the obesity-resistant mice overexpressed UCP2 mRNA in their adipose tissue. These findings, together with the function of the protein and the chromosomal location of its gene, led to propose a role for UCP2 in diet-induced thermogenesis. However, this hypothesis was not validated in Ucp2$-/-$ mice.

UCP3: ANOTHER UCP HOMOLOGUE PREDOMINANTLY PRESENT IN SKELETAL MUSCLE

cDNAs corresponding to UCP3, a protein homologous to UCP1 and UCP2, were cloned soon after the discovery of UCP2. The amino acid sequence of UCP3 is 72% identical to that of UCP2 and 57% identical to that of UCP1. UCP3 mRNA is predominantly expressed in the skeletal muscles of humans, mice, and rats (Figure 5). Human and murine UCP2 and UCP3 genes are juxtaposed on the same chromosome.
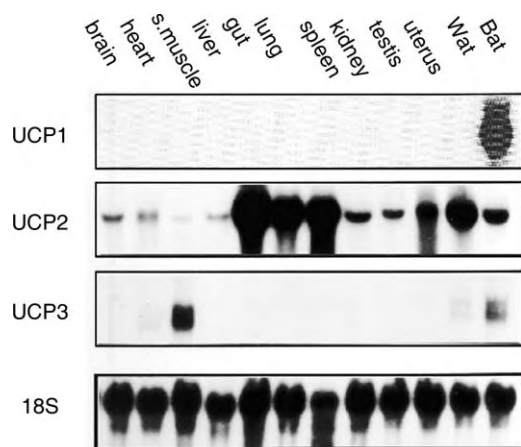
PLANT AND BIRD UCPS

Functional studies of isolated plant mitochondria suggested that plants contain mitochondrial UCPs. Following this observation, a cDNA from a new UCP was isolated from a potato cDNA library. The sequence of the corresponding protein, called stUCP, is 44% identical to that of UCP1 and 47% identical to that of UCP2. The mRNA of stUCP is present in most plant organs. The most surprising result was that stUCP mRNA was markedly induced in the leaves of plants exposed to a temperature of 4°C. Cold, therefore, induced stUCP as it induces UCP1 in the brown adipose tissue of animals. However, plant and animal UCPs may have different functions. More recently, a chicken UCP termed avUCP was characterized. This gene is only expressed in skeletal muscles of birds and ducks and is induced upon exposure to the cold. Therefore avUCP seems to be involved in regulatory thermogenesis. Other proteins referred to as UCP4 and BMCP1/UCP5 were characterized. However their exact relationship to UCP1, UCP2, and UCP3 will require further work.

ROLE AND FUNCTION OF UCPS OTHER THAN UCP1: A ROLE IN CONTROLLING THE LEVEL OF REACTIVE OXYGEN SPECIES

Numerous questions remain unanswered, notably concerning the exact catalytic activity of each UCP, and the nature of the endogenous ligands of UCP2 and UCP3. These proteins may simply translocate protons. As said above for UCP1, it has been proposed that all the UCPs are active as fatty acid cycler through the membrane. According to this later hypothesis, protonated fatty acids cross the membrane and release a proton on the matricial side; then, the UCP facilitates the translocation of the anionic form of fatty acids. High loads of UCP2 or UCP3 can uncouple respiration in yeast or mammalian cell, but it is not certain whether the low physiological levels of these proteins may be sufficient to induce a net uncoupling of respiration *in vivo*. The divergent observations of the regulation of the activities of UCP2 or UCP3 by fatty acids or nucleotides will require further studies. Interestingly, it has been proposed that UCP2 and UCP3 could export fatty acid anion under conditions of elevated fatty acid oxidation.

FIGURE 5 Tissular distribution of UCP1, UCP2, and UCP3 RNAs. The data correspond to mouse tissues. A very similar picture was obtained when analyzing human tissues. 18S refers to ribosomal RNA.
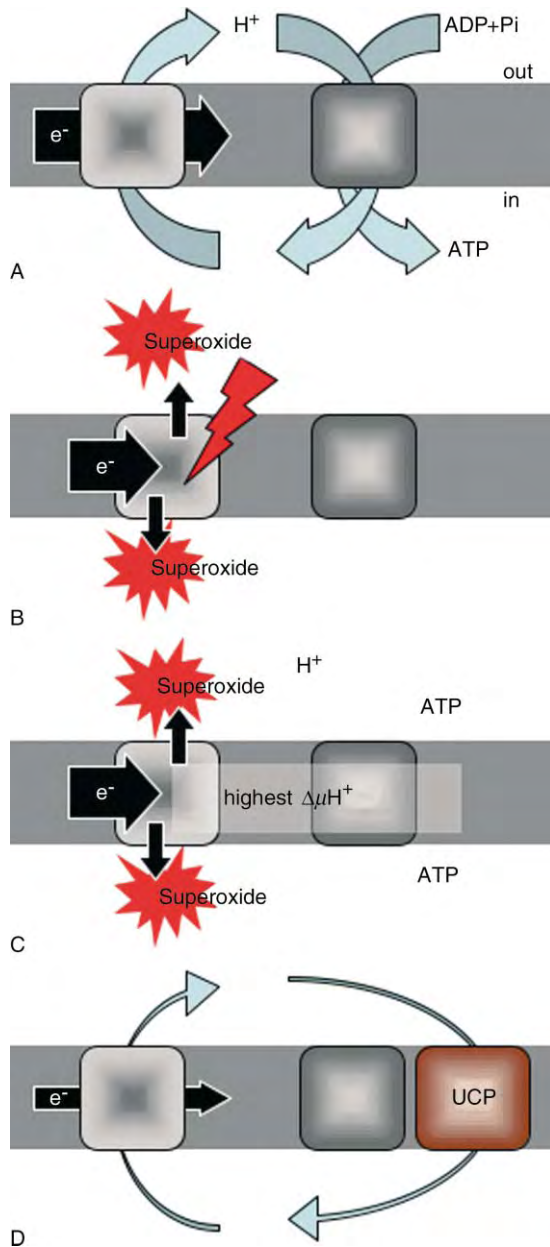
FIGURE 6 Role of UCP and superoxide generation by mitochondria. (A) mitochondria phosphorylating ATP: a proton gradient generator (the respiratory chain, left square) uses the energy liberated by the movement of electrons to pump protons through the membrane (grey zone) towards the intermembrane space. A proton gradient consumer (the ATP-synthase, right square) phosphorylates ADP coming from outside into ATP, that is exported afterwards. Therefore, a proton cycle across the inner membrane couples oxidation to ATP synthesis (and *vice versa*). (B) and (C) when oxidation is impaired (B) or slows down (C), an excess of electrons aiming to travel across the respiratory complexes cannot reach oxygen (to form water). Single electrons react with molecular oxygen and form superoxide anion ($O2^{.-}$). It occurs either when the respiratory chain is poisoned (red arrow in panel B) or under physiological circumstances when the use of ATP by the cell is weak, ATP level rises, and ADP level decreases (panel C). In that case, the membrane potential rises to its highest value which impedes proton pumping, slows down reoxidation, and therefore increases superoxide production. (D) in tissues other than the thermogenic brown adipose tissue, the UCPs may induce a controlled leak of protons across the

Metabolic Activity of UCP2 and UCP3

Conflicting data regarding the association of UCP2 or UCP3 genetic polymorphisms to body mass index, susceptibility to obesity, resting metabolic rate, metabolic efficiency, fat oxidation, insulin resistance, susceptibility to gain fat with age were obtained. However, UCP2 appears to act as a negative regulator of insulin secretion. Moreover, mice overexpressing a large amount of human UCP3 in skeletal muscle weigh less, have a decreased amount of adipose tissue, and an increased resting oxygen consumption.

Reactive Oxygen Species

Guided by the fact that UCP2 is expressed at high level in the immune system, a role for UCP2 in immunity was searched for. In fact, Ucp2 $-/-$ mice are more resistant to infection by a parasite. This phenotype was explained by the fact that disruption of the Ucp2 gene provoked an elevation of ROS level that facilitated the killing of the pathogens. Mice lacking the Ucp3 gene also produced more oxygen radicals in myocytes. Therefore, it appears that UCP2 and UCP3 contribute to limit the level of reactive oxygen species in cells. Such an activity may be explained by a mild uncoupling activity of the proteins since a small decrease of the mitochondrial membrane potential opposes production of superoxide ion in mitochondria (Figure 6). The observation that UCP2 protects against atherosclerosis highlights its ability to counteract radical synthesis.

UCP1, UCP2, and UCP3, Conclusions and Perspectives

UCP1 has a well-demonstrated uncoupling activity and an essential role in maintenance of body temperature in small rodents exposed to the cold, but the exact biochemical and physiological roles of UCP2 and UCP3 remain to be further identified. Certain data support a true activity of these UCPs in respiration uncoupling and substrate oxidation, other data do not agree with a role for these UCPs in controlling adiposity. Several reports support a role for UCP2 and UCP3 in decreasing the level of reactive oxygen species in cells.

inner membrane that allows reoxidation to occur and maintain the membrane potential below the threshold inducing superoxide generation. The mechanism of proton transport (proton channeling or fatty acid cycling), as well as the identity of regulators of UCPs inside the cells are still controversial. Fatty acids would be cofactors of the proton transport whereas nucleotides are inhibitors. The membrane potential itself is a variable influencing heavily UCP1 activity. It has been proposed that superoxide anion and quinones directly activate proton transport by UCPs.

Whilst UCP1 is present at a high level, an open question is that of the exact amount of UCP2 and UCP3 in cells since whatever is the exact intrinsic ability of these proteins to uncouple respiration from ATP synthesis, a certain amount of the proteins will be required to achieve a physiological uncoupling. Finally, the possible contribution of a UCP to the hypermetabolic syndrome (known as Luft's disease) previously described in patients remains to be investigated.

SEE ALSO THE FOLLOWING ARTICLES

ATP Synthesis in Plant Mitochondria: Substrates, Inhibitors, Uncouplers • Luft's Disease • Mitochondrial Membranes, Structural Organization • Quinones • Respiratory Chain and ATP Synthase

GLOSSARY

mitochondria These organelles are the sites of respiration and oxidative phosphorylation in all animal and higher plant tissues as well as in protozoa, fungi, and aerobically grown yeasts. They are ~ 2–$30\ \mu$m long and 0.5–$10\ \mu$m wide.

respiration coupling/uncoupling In coupled mitochondria, respiration rate determines ADP phosphorylation rate and ATP synthesis determines respiration rate. When uncoupling occurs (due to chemical uncouplers or UCP), the respiration rate increases sharply since the control by ADP phosphorylation does not limit respiration. In such a situation, oxidation energy is dissipated as heat.

FURTHER READING

Boss, O., Hagen, T., and Lowell, B. B. (2000). Uncoupling proteins 2 and 3. Potential regulators of mitochondrial energy metabolism. *Diabetes* **49**, 143–156.

Cannon, B., and Nedergaard, J. (1985). The biochemistry of an inefficient tissue: Brown adipose tissue. *Essays Biochem.* **20**, 110–164.

Echtay, K. S., Winkler, E., and Klingenberg, M. (2000). Coenzyme Q is an obligatory cofactor for uncoupling protein function. *Nature* **408**, 609–613.

Enerbäck, S., Jacobsson, A., Simpson, E. M., Guerra, C., Yamashita, E. H., Harper, M. E., and Kozak, L. P. (1997). Mice lacking mitochondrial uncoupling protein are cold-sensitive but not obese. *Nature* **387**, 90–94.

Garlid, K. D., and Jaburek, M. (1998). The mechanism of proton transport mediated by mitochondrial uncoupling proteins. *FEBS Lett.* **438**, 10–14.

Himms-Hagen, J., and Ricquier, D. (1998). Brown adipose tissue. In *Handbook of Obesity* (G. Bray, C. Bouchard and W. P. T. James, eds.) pp. 415–441. Marcel Dekker, New York.

Ledesma, M., Garcià de Lacoba, M., and Rial, E. (2002). The mitochondrial uncoupling proteins. *Genome Biol.* **3** (on line November 29).

Nicholls, D. G., and Locke, R. M. (1984). Thermogenic mechanisms in brown fat. *Physiol. Rev.* **64**, 1–64.

Ricquier, D., and Bouillaud, F. (2000). The uncoupling protein homologues: UCP1, UCP2, UCP3, StUCP & AtUCP. *Biochem. J.* **345**(Pt 2), 161–719.

Skulachev, V. P. (1988). Uncoupling: New approaches to an old problem of bioenergetics. *Biochim. Biophys. Acta* **1363**, 100–124.

Stuart, J. A., Cadenas, S., Jekabsons, M. B., Roussel, D., and Brand, M. D. (2001). Mitochondrial proton leak and the uncoupling protein 1 homologues. *Biochim. Biophys. Acta* **1504**, 144–158.

BIOGRAPHY

Daniel Ricquier is Professor at Neckes-Enfants Malades Faculty of Medicine and Frédéric Bouillaud is Main Investigator at the Institut National de la Santé et de la Recherche Médicale. They work at CNRS Unit 9078 at Faculté de Médecine and Institut de Recherches Necker-Enfants Malades in Paris. They hold Ph.Ds from Pierre and Marie Curie University in Paris. Their laboratory first cloned UCP1, UCP2, BMCP1/UCP5, and avUCP. They also collaborated to the characterization of the first plant UCP. With their collaborators, they are authors of original reports dealing with biochemistry, physiology, and genetics of the UCP family.

Unfolded Protein Responses

David Ron
New York University School of Medicine, New York, USA

The unfolded protein response (UPR) is a transcriptional and translational response to the accumulation of unfolded proteins in the endoplasmic reticulum (ER). The UPR is mediated by highly conserved signaling pathways that are activated by imbalance between the load of unfolded (or malfolded) ER client proteins and the capacity of the organelle to process this load. Collectively these pathways restore equilibrium to the protein-folding environment in the organelle by increasing the expression of genes that enhance nearly all aspects of ER function and by transiently repressing the biosynthesis of new client proteins. Interest in the UPR has been stimulated by the realization that postsynthetic protein processing constitutes an important step in gene expression and that protein malfolding plays an important role in human disease.

ER Function and ER Stress

THE ER, A PROTEIN-PROCESSING MACHINE

Proteins destined for secretion and membrane insertion are translocated across the ER membrane in an unfolded state. In the ER lumen they undergo chaperone-assisted folding, a variety of organelle-specific post-translational covalent modifications and often chaperone-assisted assembly into oligomeric structures. Once properly folded and assembled, most ER client proteins are packaged into vesicles that are transported to more distal sites in the secretory pathway. Proteins that fail to attain their proper folded and oligomeric conformation are retained in the ER by continued binding to chaperones and are ultimately translocated from the organelle to the cytoplasm for proteasomal degradation, in a process known as ER-associated protein degradation (ERAD).

To fulfill these various functions the ER is endowed with a unique complement of proteins. These include enzymes for post-translational covalent modifications (e.g., N-linked glycosylation, disulfide bond formation), chaperones that assist in folding and assembly steps, translocation channels and transporters involved in transmembrane traffic, and various components involved in the postassembly steps of client protein egress from the organelle. The build-up of these components defines the capacity of the organelle to handle its client proteins. In the late 1980s Sambrook, Gething, and their colleagues discovered that manipulations that interfere with the function of various aspects of the ER client protein-handling machinery, and thereby perturb protein folding in the ER, selectively up-regulate the expression of genes that encode components of that machinery. The extent of this transcriptional response was not fully appreciated at the time, however its selectivity was noted, hence its name: the ER unfolded protein response (or UPR).

It was imagined that the cell possessed means to monitor the load of client proteins presented to its ER and responded to an increase in such load by up-regulating the capacity of its ER to process unfolded client proteins. An additional clue to the workings of this response was provided by the seminal observation that forced overexpression of an ER chaperone, BiP (also known as GRP78) markedly suppressed the activity of the UPR. Thus, at least one target gene of the UPR (BiP) is able to exert negative feedback on the entire response. BiP overexpression does not restore function to a challenged ER; this requires the coordinate expression of numerous UPR target genes. However, BiP, which is a member of the highly conserved HSP70/DnaJ family of chaperones, promiscuously recognizes hydrophobic stretches of amino acids. These are normally incorporated into the cores of properly folded polypeptides and assembled oligomeric complexes, but remain exposed on the surface of unfolded, malfolded or unassembled proteins. The ability of BiP to suppress the UPR suggested that it might be doing so by masking a stress signal generated by many different unfolded and malfolded proteins. This phenomenon was assigned the heuristic term ER stress and its level reflects the balance between client protein load and the capacity of the organelle to process that load.

PHYSIOLOGICAL AND PATHOLOGICAL ER STRESS

To the extent that cell types vary in the load of secreted proteins that they are called upon to produce,

they are subject to widely different levels of physiological ER stress. This explains enhanced activity of the UPR in various professional secretory cells, such as pancreatic cells of vertebrates or intestinal cells of the nematode, *Caenorhabditis elegans*. ER stress also occurs when a mutation in an abundantly expressed ER client protein renders that protein especially difficult to fold. An example is provided by various degenerative diseases affecting myelinated neurons in which mutations in a component of the myelin sheath (an ER client protein) cause the protein to malfold and induce high levels of ER stress which, over time, destroys the myelin-producing cell. It is important to emphasize that most mutations that impede folding of ER client proteins do not cause measurable ER stress. However, because they diminish expression of the properly folded protein, such mutations may deprive the organism of the latter's beneficial actions. This genetic mechanism underlies such serious human diseases as cystic fibrosis, familial hypercholesterolemia, and hemophilia. Nonetheless, the level of expression of the mutant protein is not enough to globally challenge ER function in the cell that produces it and the phenotypic expression of the mutation reflects the lack of an important protein, rather than the production of a toxic one.

The above constitute client protein-driven ER stress, however ER stress may also initiate from impaired function of the organelle, which occurs in cells deprived of energy sources or oxygen, or in cells exposed to certain ER-specific toxins. We do not understand in detail how ER stress contributes to further organelle dysfunction and ultimately cell death. However a framework for thinking about this has recently emerged with the realization that unfolded and malfolded proteins present reactive interfaces that have not been vetted by evolution and may thus disrupt the cellular machinery by interacting promiscuously and illegitimately with essential cellular components. According to this theory, proteotoxicity is normally held in check by the chaperones that bind such potentially toxic protein interfaces and let go only once the latter have been buried in the hydrophobic cores of the properly folded client protein. ER stress (by definition) challenges the capacity of the chaperones and may permit illegitimate protein interfaces to emerge. The ability of chaperone overexpression to suppress ER stress signaling suggests that the need to prevent proteotoxicity is an important driving force in evolution of the UPR. The unifying feature of ER stress need not be the presence of toxic moieties on the surface of every unfolded or malfolded protein. The common feature of diseases of protein folding might instead be the exhaustion of a protective chaperone reserve, which normally suppresses the potential proteotoxicity of certain (possibly normal) folding intermediates of ER client proteins.

The Yeast Unfolded Protein Response

IRE1, A PROTOTYPE OF TRANSMEMBRANE STRESS SIGNALING

Early studies on the UPR were carried out in the yeast, *Saccharomyces cerevisiae*, an organism that lends itself well to forward genetic screens. To screen for mutations affecting the UPR, the regulatory region of yeast BiP gene (KAR2) was fused to a reporter and mutant yeast with suppressed activity of this reporter were sought. The first gene thus identified was inositol requiring 1 (IRE1), so-called because its loss of function had been previously noted to result in inositol auxotrophy. IRE1 encodes a transmembrane ER resident protein with an N-terminal domain residing in the ER lumen and a C-terminal domain that is exposed on the cytoplasmic side. The membrane topology and subcellular localization of IRE1 immediately suggest a mechanism for transmitting information on the state of the ER (topologically equivalent to the extracellular space) to the cell's interior. The lumenal domain somehow senses ER stress, conveying the signal across the ER membrane to the cytoplasmic domain, which broadcasts it to the nucleus, turning "on" UPR target gene expression (Figure 1).

IRE1's C-terminal, cytoplasmic, effector-domain, is a protein kinase and undergoes autophosphorylation when activated by ER stress. This suggested that IRE1 might function like other transmembrane receptors that are also protein kinases and convey their signal by phosphorylating downstream targets, often through a kinase relay. However, other than IRE1 itself no substrates for the aforementioned kinase have been identified to date.

UNCONVENTIONAL SPLICING OF HAC1 mRNA

The clues to understanding propagation of the UPR signal, downstream of IRE1 again came from yeast genetics. Mutations in two additional genes were noted to block the yeast UPR. One of these, HAC1, encodes a transcription factor, which binds to and activates the promoters of the primary target genes of the yeast UPR, the second, more mysteriously turned out to be RLG1, which encodes for a multifunctional enzyme involved in tRNA splicing. ER stress promotes accumulation of HAC1 protein and this is blocked by mutations in IRE1 and RLG1.
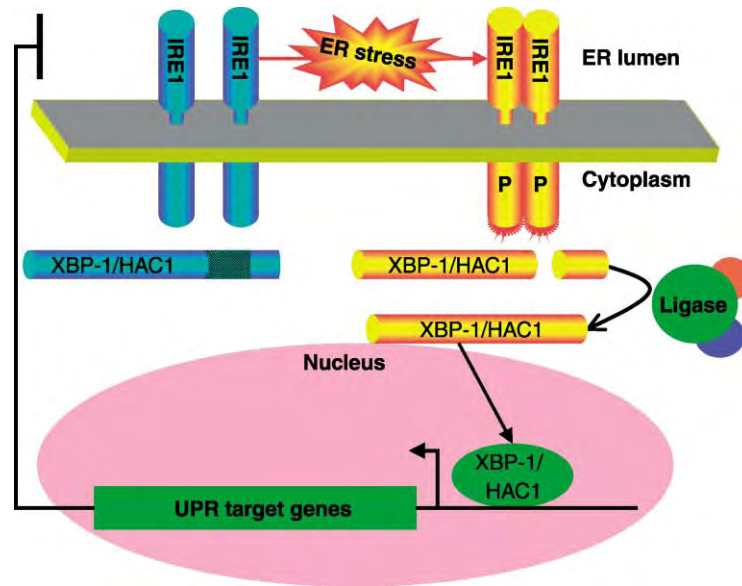
FIGURE 1 IRE1 and unconventional splicing of its target mRNA signal the unfolded protein response. ER stress leads to oligomerization and trans-autophosphorylation of IRE1 (indicated by the "P" on its cytoplasmic effector domain). Phosphorylation unmasks the effector function of IRE1 (cartooned by the red bristles on the cytoplasmic domain), which consists of the endoribonucleolytic processing of its target mRNA, XBP-1 in metazoans, and HAC1 in yeast. The two ends of the cleaved mRNA are joined together by a ligase. The unconventionally spliced XBP-1/HAC1 mRNA is more efficiently translated than the unprocessed mRNA and encodes a protein that is more stable and more effective at trans-activation of UPR target genes in the nucleus.

The underlying mechanisms, which turned out to be full of surprise, were revealed in a series of brilliant experiments carried out in the Walter and Mori laboratories.

The regulated step in HAC1 expression entails the removal of a 242 base internal segment at the 3′ end of the mature mRNA. The enzyme which precisely cuts the mRNA at the 5′ and 3′ ends of this segment turned out to be IRE1 itself, whose highly sequence-specific endoribonucleolytic activity appears to be subordinate to its kinase activity (though the molecular details remain to be worked out). The two ends of the severed HAC1 mRNA are held together by extensive base pairing and are subsequently religated by the protein product of RLG1. IRE1 is the master regulator of this unconventional mRNA splicing reaction, since its kinase and endoribonuclease are activated by ER stress. The removal of the aforementioned internal segment of the HAC1 mRNA de-represses HAC1 translation and the protein encoded by the processed mRNA is more stable and transactivates its target genes more potently than that encoded by the unprocessed mRNA.

Mutations in IRE1 and HAC1 similarly abolish most signaling in the yeast UPR. This was convincingly demonstrated by expression profiling which showed that the vast majority of the genes induced by exposing yeast to toxins that promote ER stress were no longer activated in IRE1 or HAC1 mutant strains. Thus the yeast UPR consists of a linear pathway in which HAC1

is an essential target of IRE1 and together the two proteins are required for the activation of most known UPR target genes.

Diversification in the Metazoan UPR

CONSERVATION OF THE IRE1 PATHWAY

Systematic genomic sequencing of the metazoan *C. elegans* and random cDNA sequencing from several mammalian species, predicted proteins highly related to yeast IRE1 in higher eukaryotes. Their common membrane topology, localization to the ER, and conserved kinase and endoribonuclease domains suggested that they were indeed IRE1 homologues. However, HAC1 homologues could not be identified by simple sequence comparison to the yeast. Eventually, however, a combination of forward genetic screens in *C. elegans* and creative guesswork led to the identification of XBP-1 as a functional homologue of HAC1 in higher eukaryotes. It too encodes a transcription factor whose expression is tightly regulated by IRE1 through the processing of its mRNA.

Less conserved, however, is the role of the IRE1 pathway in the yeast and metazoan UPR. Whereas in the former, IRE1 and HAC1 are absolutely required for activating nearly all UPR target genes, in mammals, IRE1 and XBP-1 are dispensable for activation of

all but a very small set of UPR targets. This is all the more remarkable when one considers that most UPR target genes are conserved between yeast and mammals, it is just that the latter have found alternative means to couple their expression to ER stress. *C. elegans* occupies an intermediate position between yeast and mammals, in that most of its identifiable UPR targets are partially dependent on IRE1 and XBP-1, but can to some extent also be activated independently of that pathway.

Despite this redundancy in the mammalian UPR, both IRE1 and XBP-1 are essential for embryonic development (mammals have two IRE1 genes, a nonessential beta isoform restricted in its expression to the intestinal and bronchial epithelium and a broadly expressed alpha isoform which is essential). The reason(s) for the embryonic lethality of IRE1α and XBP-1 null animals are not fully understood, but it is hypothesized that in mammals the pathway may have diverted to specify the capacity for especially high levels of protein secretion. This hypothesis, which is clearly in need of further experimental support, suggests that specification of a secretory cell fate activates a conserved stress pathway that drives a developmental process, namely acquisition of an apparatus required by professional secretory cells for high-capacity protein secretion. This idea is further supported by the observation that many UPR target genes function far downstream in the secretory pathway rendering it unlikely that such genes merely relieve the stressed ER of its load of unfolded client proteins.

ATF6 AND INTRAMEMBRANE PROTEOLYSIS IN THE METAZOAN UPR

The identification of ATF6 by Mori and colleagues confirmed the predicted redundancy in the metazoan UPR. Like IRE1, ATF6 is also an ER-localized transmembrane protein. However, its cytoplasmic effector domain consists of a transcription factor that activates UPR target genes directly. In its membrane-bound form, ATF6 is inert, as it cannot reach the nucleus. Activation involves regulated intramembrane proteolysis, which liberates the transcription factor part of ATF6 from the ER membrane under conditions of ER stress (Figure 2). The next surprise came when Brown, Goldstein, Prywes, and colleagues identified the proteases involved in this highly regulated event. These turned out to be the same proteases that process and activate the membrane-bound transcription factor, SREBP, which regulates genes involved in sterol and fatty acid biosynthesis and assimilation.

There are no mammalian gene knockouts reported for ATF6 (of which at least two isoforms exist).

However, unlike cells lacking IRE1 or XBP-1, cells lacking the proteases required for ATF6 processing are severely impaired in UPR target gene expression. It seems likely therefore that mammalian ATF6 has picked up some of the role performed by IRE1 and XBP-1/HAC1 in simpler organisms. The similarities between ATF6 and the SREBPs highlight another remarkable feature of the UPR. IRE1 and HAC1 mutant yeast are unable to synthesize adequate amounts of the membrane phospholipid precursor, inositol. Though the details remain to be worked out, insufficiency of membrane lipid components also signals through the yeast UPR and enzymes involved in phospholipid metabolism are targets of IRE1 and HAC1. In metazoans this aspect of the UPR apparently has been split off and relegated to dedicated transcription factors, the SREBPs that are activated by insufficiency of lipid components. However the ancient link between the client protein-based UPR and membrane lipid insufficiency signaling has remained in the form of shared machinery for proteolytic activation of ATF6 and SREBP.

Translational Control in the UPR

PERK COUPLES ER STRESS TO EIF2α PHOSPHORYLATION AND TRANSLATIONAL REPRESSION

In addition to activated gene expression ER stress results in a dramatic reduction in protein synthesis. Translational repression is an active process limiting the influx of client proteins into the stressed ER and thus serves as a counterpart to the gene expression program, which increases the organelle's capacity to process client proteins. Brostrom and colleagues and Kaufman and colleagues noted, early on, that translational repression by ER stress is associated with phosphorylation of the α-subunit of translation initiation factor 2 (eIF2α) on serine 51. This phosphorylation site is conserved in all eukaryotes and serves to regulate translation initiation in diverse stressful conditions. Trimeric eIF2, in complex with guanosine triphosphate (GTP), recruits the aminoacylated initiator methionyl–tRNA to the small ribosomal subunit, allowing translation initiation. Recognition of an AUG initiation codon on the mRNA leads to hydrolysis of GTP to GDP and dissociation of the ribosome–eIF2 complex. To participate in another round of translation initiation, the GDP bound to eIF2 must be exchanged to GTP. The enzyme catalyzing this exchange reaction (eIF2B) is inhibited by phosphorylated eIF2.

Distinct eIF2α kinases were known to be activated by distinct stress signals; PKR by double-stranded RNA
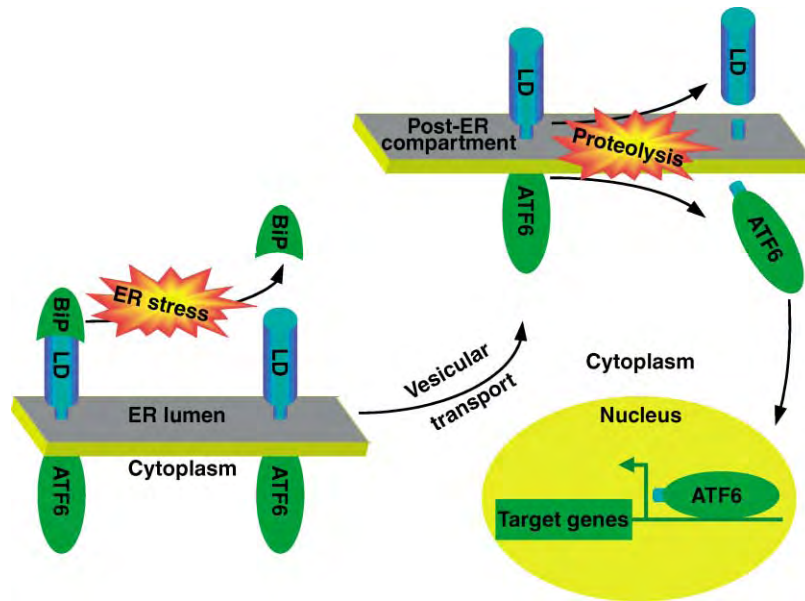
FIGURE 2 ATF6 activation by regulated intramembrane proteolysis. Membrane tethering inactivates the transcription factor ATF6, whose effector domain is prevented from accessing its target genes in the nucleus. BiP binding to the lumenal domain of ATF6 retains the protein in the ER compartment. However, under conditions of ER stress, BiP dissociates from ATF6, which then migrates (presumably by vesicular transport) to a post-ER compartment that contains the proteases that liberate the effector domain to signal in the UPR.

in viral infection and GCN2 by uncharged tRNAs in amino acid starvation. Our group noted the existence of a predicted transmembrane protein in the *C. elegans* genome, with an N-terminal "extracellular" domain similar to IRE1 and an intracellular domain similar to PKR and GCN2, known eIF2α kinases. This protein, which is conserved in metazoans (and absent from yeast) and to which we gave the name PERK (PKR-like ER Kinase), couples ER stress to eIF2α phosphorylation and is essential to translational regulation by ER stress. Cells lacking PERK are hypersensitive to ER stress and in mammals, PERK activity is especially important for preservation of cells engaged in high levels of secretion. Thus humans with PERK mutations and mice lacking the gene, develop diabetes mellitus and skeletal defects attributed to dysfunction of insulin- and collagen-secreting cells in the endocrine pancreas and bone, respectively.

GENE EXPRESSION, eIF2α PHOSPHORYLATION AND INTEGRATION OF SIGNALS IN THE UPR, AND OTHER STRESS PATHWAYS

In addition to limiting the load of client protein on the stressed ER (by inhibiting translation of most mRNA), eIF2α phosphorylation paradoxically activates the translation of the mRNA encoding the transcription factor ATF4 (and likely that of other regulatory proteins which remain to be discovered, Figure 3). PERK thus plays an important role in up-regulating target gene expression in the UPR. Expression profiling of ER stressed wild-type and PERK $-/-$ cells reveals a broad role for PERK in UPR-target gene expression, with mild impairment in most target genes and severe impairment in a smaller group. Among the UPR target genes that are strongly impaired in PERK $-/-$ cells are genes involved in amino acid transport and metabolism and a gene, *GADD34*, which encodes a phosphatase that dephosphorylates eIF2α and terminates signaling by PERK. The aforementioned strong PERK target genes play an important role in maintaining translation, by providing building blocks for secreted protein synthesis and by promoting eIF2α dephosphorylation. Thus, PERK and eIF2α phosphorylation appear to have a dual role in the UPR. Early in the response they protect the stressed cells from ER overload whereas later they promote conditions required for synthesis of secreted proteins.

A considerable overlap exists between genes activated by ER stress and genes activated by other stressful conditions that promote eIF2α phosphorylation. Thus, eIF2α phosphorylation integrates signaling in several stress pathways. The extent of this integrated stress response is not fully known, but it appears to control important aspects of cellular and organismal metabolism.
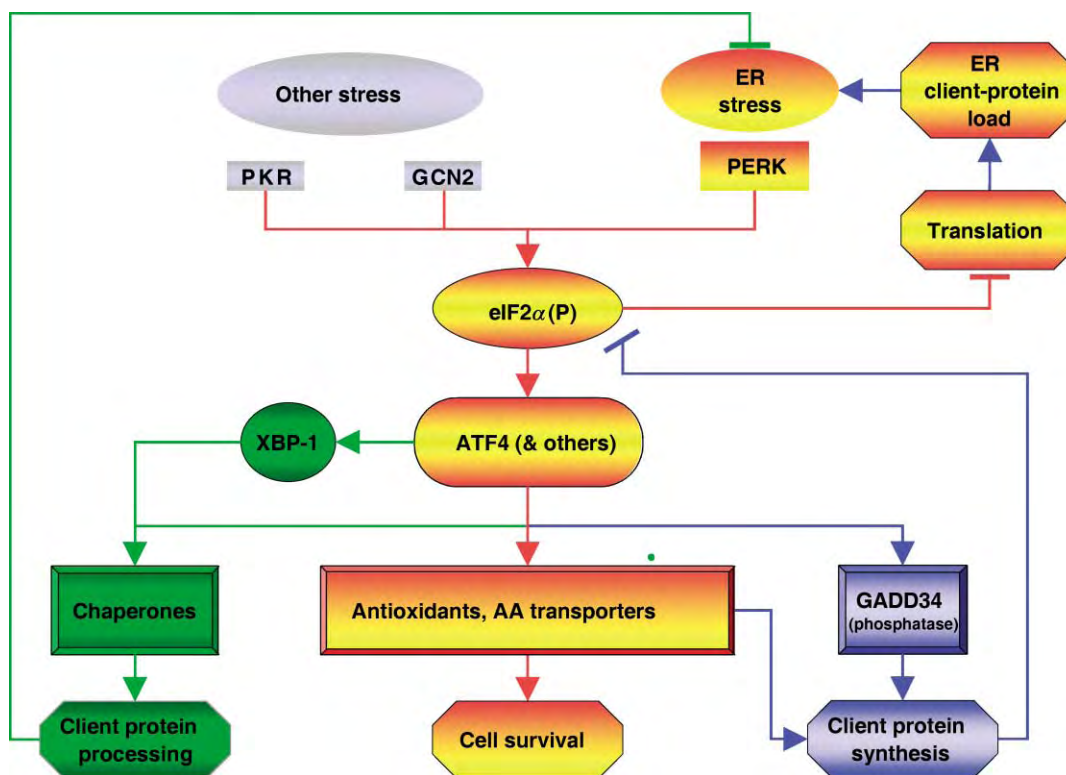
FIGURE 3 The eIF2α phosphorylation-dependent integrated stress response. ER stress and other stresses activate eIF2α kinases. Phosphorylated eIF2 inhibits translation initiation, alleviating ER client protein load, while at the same time activating the transcription factor ATF4 (and other, yet to be identified effectors). These activate a gene expression program with several components: The eIF2α phosphatase GADD34 promotes translational recovery and resumed client protein synthesis. Genes involved in antioxidant responses and amino acid import and assimilation, promote cell survival and client protein synthesis. PERK signaling also functions synergistically with other arms of the UPR (e.g., by transcriptional activation of the IRE1 target gene XBP-1) to enhance ER client protein processing.

A Common Mechanism for Sensing ER Stress

The identification of three different stress sensors in the UPR has allowed a comparison between their mode of activation. As mentioned earlier, overexpression of BiP and other ER chaperones had been noted to repress signaling in the UPR. In the equilibrated ER, all three stress transducers, IRE1, ATF6, and PERK, are found in a simple complex with BiP, which binds on the lumenal side to their stress-sensing domains. These complexes are rapidly dissociated under conditions of ER stress. Furthermore, dissociation precedes activation of the stress transducers. BiP dissociation promotes oligomerization of IRE1 and PERK, which in turns leads to transautophosphorylation and activation of downstream signaling. It seems likely that BiP-dissociation under conditions of ER stress unmasks a homo-oligomerization domain in the similarly structured lumenal domains of IRE1 and PERK.

The chain of events in the case of ATF6 activation is more complex, but conceptually similar. BiP binding retains ATF6 in the ER, segregated from the proteases that release its transcription factor portion, as these are localized to a post-ER compartment. BiP dissociation liberates ATF6 to migrate to the protease-containing compartment, a migration that presumably takes place by vesicular transport. Thus, dispensable molecules of BiP, which are not engaged in chaperoning ER client proteins actively repress signaling by all three known transducers of the UPR. It would thus appear that the recognition of unfolded and malfolded proteins is relegated to professional chaperones and the stress transducers passively monitor the degree to which the latter are engaged by their clients. This arrangement of the upstream steps of the UPR mirrors that of the cytoplasmic heat shock response in which unfolded proteins in the cytoplasm activate genes implicated in folding and degrading cytoplasmic proteins.

SEE ALSO THE FOLLOWING ARTICLES

Chaperones for Metalloproteins • Chaperones, Molecular • Chaperonins • Endoplasmic Reticulum-Associated Protein Degradation • Regulated Intramembrane Proteolysis (Rip)

GLOSSARY

chaperone A protein that assists folding and assembly of other proteins and prevents their aggregation in the unfolded state.

endoplasmic reticulum A membrane-bound organelle, present in all nucleated cells, which serves as the entry point for secreted proteins.

proteotoxicity A hypothetical property of unfolded and malfolded proteins, whereby peptide domains that are illegitimately exposed on the surface perturb cellular function.

FURTHER READING

Cox, J. S., and Walter, P. (1996). A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. *Cell* **87**, 391–404.

Harding, H., Zhang, Y., and Ron, D. (1999). Translation and protein folding are coupled by an endoplasmic reticulum resident kinase. *Nature* **397**, 271–274.

Harding, H. P., Calfon, M., Urano, F., Novoa, I., and Ron, D. (2002). Transcriptional and translational control in the mammalian unfolded protein response. *Annu. Rev. Cell Dev. Biol.* **18**, 575–599.

Kaufman, R. J. (1999). Stress signaling from the lumen of the endoplasmic reticulum: Coordination of gene transcriptional and translational controls. *Genes Dev.* **13**, 1211–1233.

Sidrauski, C., and Walter, P. (1997). The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell* **90**, 1031–1039.

Yoshida, H., Okada, T., Haze, K., Yanagi, H., Yura, T., Negishi, M., and Mori, K. (2000). ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol. Cell Biol.* **20**, 6755–6767.

BIOGRAPHY

David Ron obtained his Medical Doctorate at the Technion in Haifa, Israel. Following residencies in internal medicine and endocrinology at Mount Sinai Hospital and the Massachusetts General Hospital and a research fellowship at Harvard Medical School, he took a faculty position at New York University School of Medicine, where he is currently a Professor of Medicine and Cell Biology. His laboratory studies stress responses in eukaryotic cells, with a special emphasis on signaling between organelles and the nucleus.

Urea Cycle, Inborn Defects of

Marsha K. Fearing and Vivian E. Shih
Harvard Medical School, Boston, Massachusetts, USA

Surplus nitrogen generated by amino acid metabolism cannot be stored in mammals and must be excreted. The urea cycle is the primary pathway for the disposal of ammonium nitrogen by conversion of the toxic ammonia molecule into innocuous urea. Inborn errors of metabolism in the urea cycle can result in severe diseases in humans. An accumulation of ammonia can lead to severe morbidity, including brain damage and death.

Overview

The urea cycle is the major pathway for the disposal of nitrogen in humans (Figure 1). Ninety percent of ingested protein is metabolized to urea and excreted through the urine. Ammonia is derived from a variety of precursor protein sources. Part of the urea cycle resides in mitochondria where ammonia is converted to carbamoyl phosphate by carbamoyl phosphate synthetase (CPS) along with its allosteric activator, N-acetylglutamate. The next step involves carbamoyl phosphate forming citrulline when condensed with ornithine. This process is mediated by the enzyme ornithine transcarbamoylase (OTC). Citrulline exits the mitochondria and condenses with aspartate to produce argininosuccinate. This compound is then cleaved to arginine and fumarate by argininosuccinate lyase. Arginine is hydrolyzed by arginase, thus releasing urea and regenerating ornithine. Ornithine is then shuttled into mitochondria by its own transporter. Within the urea cycle ornithine is a conserved moiety, it is neither formed nor lost. One atom of waste nitrogen from ammonia and one from aspartate forms a molecule of urea.

Liver is the only organ that contains the complete urea cycle and is the major site of ureagenesis. Intestinal epithelial cells have CPS and OTC for citrulline synthesis and are the source of the circulating citrulline in plasma that serves as a precursor for synthesis of arginine, a semi-essential amino acid, in brain and kidney. These organs have negligible arginase activity, hence they do not produce urea.

Clinical Diseases

Urea Cycle Defects

Gene mutations resulting in urea cycle defects have been described for each of the five urea cycle enzymes. The individual disease states are N-acetylglutamate synthetase (NAGS) deficiency, carbamoyl phosphate synthetase (CPS) deficiency, ornithine transcarbamoylase (OTC) deficiency, argininosuccinate synthetase (AS) deficiency or citrullinemia, argininosuccinate lyase (AL) deficiency or argininosuccinic aciduria, and arginase (A1) deficiency.

The clinical features of the urea cycle disorders are similar. Depending on the enzyme involved and the residual enzyme activity, metabolic blocks in the urea cycle will cause a resultant elevation of the plasma ammonia level. Hyperammonemia is associated with many of the clinical features, which include respiratory alkalosis, seizures, acute encephalopathy, coma, and even death. Spasticity, resembling cerebral palsy can be a late complication. The most severe presentation is coma in the neonatal period, but symptoms can occur later in childhood as well. Metabolic stress, intercurrent illness, or fever can result in endogenous protein catabolism and precipitate a hyperammonemic crisis. In milder cases, abnormal neuropsychiatric behavior can be the only manifestation. Short, friable hair (trichorrhexis nodosa) and liver fibrosis are unique to argininosuccinate lyase deficiency.

The exact pathophysiology of hyperammonemia is not known and the etiology of brain edema is still unclear. Accumulation of ammonia and astrocytic glutamine are likely to be contributory to the brain damage.

Except for the X-linked ornithine transcarbamoylase deficiency, the enzyme defects of the urea cycle disorders are inherited as autosomal recessive traits.

The gene for OTC is located on chromosome Xp2.1. Thus, males with OTC deficiency are usually more severely affected with neonatal onset of hyperammonemia, whereas the clinical presentation in females with OTC deficiency can be quite variable, depending on random X-chromosome inactivation (Lyon hypothesis).
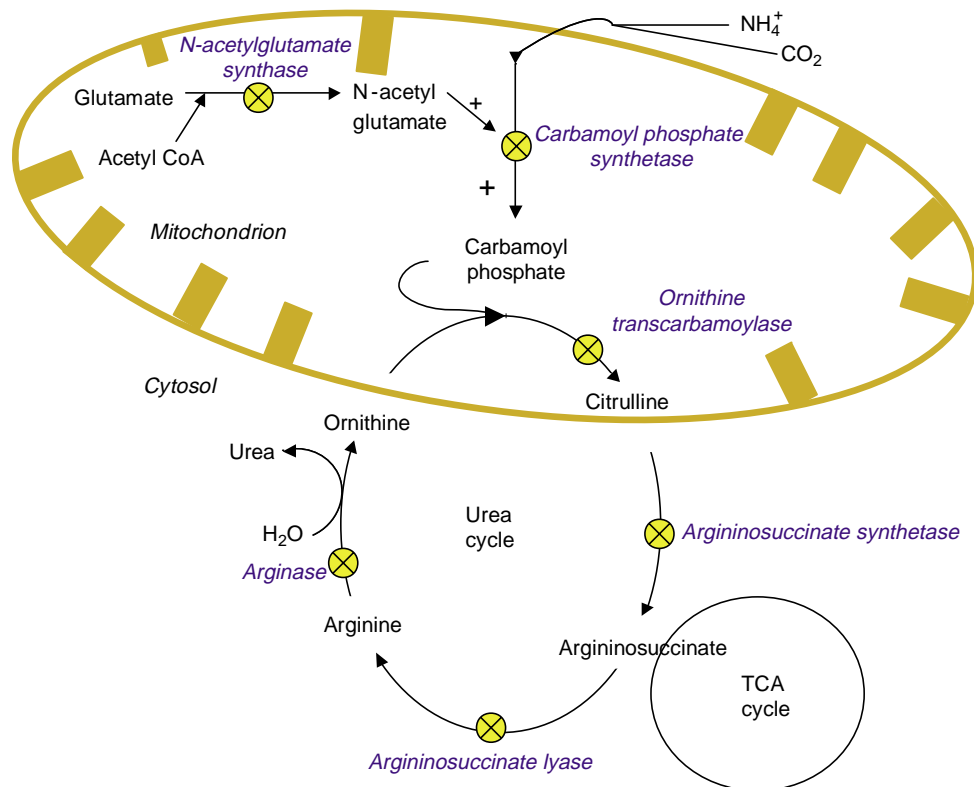
FIGURE 1 The urea cycle. (1) N-acetylglutamate synthetase, (2) carbamoyl phosphate synthetase, (3) ornithine transcarbamylase, (4) argininosuccinate synthetase, (5) argininosuccinate lyase, (6) arginase, (7) mitochondrial ornithine transporter, and (8) ornithine aminotransferase. NAG = N-acetylglutamate, an allosteric activator of carbamoyl phosphate synthetase.

OTHER HYPERAMMONEMIA STATES DUE TO AMINO ACID TRANSPORTER DEFECTS

Amino acid transporter defects are disorders of compartmentation. The urea cycle enzymes are available but the substrates are not transported to the proper site, resulting in a functional enzyme deficiency.

Triple H Syndrome (Hyperornithinemia, Hyperammonemia, Homocitrullinuria)

The HHH syndrome is a defect in ornithine mitochondrial transport, which prevents the recycling of ornithine as the substrate for OTC. The end result is a functional OTC deficiency and reduced urea synthesis. These children present with episodic hyperammonemia, ataxia, vomiting, irritability, lethargy, and coma. They often come to attention due to mental retardation and developmental delay, with occasional seizure activity. The diagnostic markers are a combination of high plasma ornithine, high blood ammonia, increased urine orotic acid, and high urine homocitrulline, the origin of which is unclear. The treatment is similar to that for OTC deficiency.

Lysinuric Protein Intolerance

Lysinuric protein intolerance (LPI), also known as hyperdibasic aminoaciduria, is an autosomal recessive disease with an extremely high prevalence in the Finnish population. The gene mutation has been identified on chromosome 14q12. LPI occurs from defects in the cellular basolateral membrane transporter of the dibasic amino acids, in renal tubules and the intestinal system. This results in excessive renal loss of lysine, arginine, and ornithine in the urine and poor gastrointestinal absorption of the same amino acids. Arginine and ornithine deficiency impairs the function of the urea cycle and results in hyperammonemia after protein intake. Symptoms of LPI include failure to thrive, vomiting, and chronic lysine deficiency, which causes growth failure with decreased bone density resulting in, osteoporosis. LPI patients also have recurrent autoimmune deficiency and infections due to low white and red blood cell counts. Other organ involvement includes interstitial lung disease and renal failure.

Citrin Deficiency (Citrullinemia Type II)

Citrullinemia type II was originally reported as an adult-onset citrullinemia, mainly in Japanese patients, with

TABLE I

Diagnostic Characteristics of Urea Cycle Disorders and Hyperammonemic Syndromes due to Amino Acid Transporter Defects

Disorder	Plasma Cit	Plasma Arg	Urine orotic acid	Other diagnostic amino acids	Genetic markers	Enzyme assay	Prenatal diagnosis
N-acetyl-glutamate synthetase deficiency	↓		N		AR; 17q21.31	Liver	DNA testing
Carbamoyl phosphate synthetase deficiency	↓	↓	N		AR; chromosome 2p	Liver	Linkage analysis; direct mutational analysis
Ornithine transcarbamoylase deficiency	↓	↓	↑		X linked; Xp2.1	Liver	Molecular diagnosis
Argininosuccinate synthetase deficiency	↑↑	↓	↑		AR; chromosome 9q34	Liver; fibroblasts; lymphoblasts, chorionic villi	Enzyme assay; direct mutational analysis; Metabolite measurement in amniotic fluid
Argininosuccinate lyase deficiency	↑	↓	↑	↑ Arginino-succinic acid (blood and urine)	AR; chromosome 7q	Liver; fibroblasts, red blood cells	Metabolite measurement in amniotic fluid
Arginase deficiency	N	↑	↑		AR; chromosome 6q23	Liver, red blood cells	Enzyme assay, DNA analysis
Lysinuric protein intolerance	↓	↓	↑	↑ arginine, lysine, ornithine (urine)	AR; chromosome 14q11.2		
Triple H syndrome	↓	↓	↑	↑ ornithine (blood) ↑ homocitrulline (urine)	AR; chromosome 13q14	Fibroblasts, Amniocytes	Enzyme assay in fetal blood, DNA analysis
Citrin deficiency	↑	↓	↑	↑ citrulline (blood and urine)	AR; chromosome 7q21.3	Liver	DNA analysis

a liver-specific argininosuccinate synthetase deficiency. The clinical presentation depends on the degree of hyperammonemia and includes a sudden disturbance of consciousness, restlessness, drowsiness, and coma with cerebral edema. Recent studies have identified a deficiency of the citrin protein, which is a mitochondrial transporter of aspartate and/glutamate, as the underlying cause. This has been confirmed by mutation analysis. Citrin deficiency has now been identified in citrullinemic patients of different ethnic backgrounds with neonatal intrahepatic cholestasis and other forms of severe hepatic dysfunction.

Diagnosis

All disorders have hyperammonemia. Plasma and urine amino acid analysis reveal the specific diagnostic patterns (Table I). The amino acids proximal to the enzyme block are increased and those distal to the block are decreased. When there is a blockage in the urea cycle beyond the step of CPS, the carbamoyl phosphate synthesized cannot be used for urea synthesis and is shunted to the biosynthetic pathway of pyrimidine to form orotic acid. Thus, orotic acid is increased in all urea cycle disorders except for NAGS deficiency and CPS deficiency. Diagnosis can be confirmed by enzyme assay in red blood cells (arginase deficiency) and cultured fibroblasts (AS deficiency, AL deficiency). Measurement of CPS and OTC activities can only be done in liver biopsy. Enzymatic details can be seen in Table II. Alternatively, molecular DNA mutation analysis can be useful for diagnosis as well as for future genetic counseling.

For all of the urea cycle defects, except NAGS, prenatal diagnosis is available. AS and AL deficiencies are noted by finding increased citrulline and increased ASA in amniotic fluid. Amniocytes do not have all of the urea cycle enzymes; only AS and AL are expressed. Fetal liver biopsy is required for prenatal diagnosis of CPS and OTC deficiencies, and fetal blood is needed for prenatal diagnosis of arginase deficiency. For prenatal diagnosis of HHH syndrome, the ornithine incorporation assay can be performed in cultured amniocytes. In cases of known mutations in a family, DNA analysis is the preferred procedure.

Treatment

LONG-TERM MANAGEMENT

Reduction in protein intake while maintaining nutritional needs to ensure growth and development remains the mainstay of therapy. The goal of therapy is to reduce the flow of nitrogen for disposal through the urea cycle. This can be achieved by restriction of dietary protein intake by providing an alternative pathway for nitrogen disposal.

Alternative pathway therapy for nitrogen excretion diverts ammonia clearance through compounds other than urea. Sodium benzoate is one of the drugs used for this purpose. When administered it conjugates with the amino acid glycine to form hippurate, which has high renal clearance. One mol of nitrogen is excreted per one mole of hippurate. Following the same principle but with greater efficiency, sodium phenylacetate conjugates with the amino acid glutamine to form phenylacetylglutamine, which eliminates two moles of nitrogen.

To maintain normal growth it is important to provide enough essential amino acids by supplementing the diet. In patients with a urea cycle defect, arginine synthesis is reduced and arginine becomes a semi-essential amino acid. Patients can be supplemented with citrulline, the precursor of arginine, to restore total body arginine pools or with arginine directly in patients with AS and AL deficiency.

In addition to medical management, liver transplantation has been used with some success for treatment of these disorders.

TABLE II

Urea Cycle Enzyme Characteristics and Locations

Enzyme	Cellular location	Relative activity in liver	pH optimum	Tissue distribution
NAGS	Mitochondrial matrix	0.01	8.5	L, I, (K)
CPS1	Mitochondrial matrix	3.1	6.8–7.6	L, I, (K)
OTC	Mitochondrial matrix	73	7.7	L, I, (K)
AS	Cytosol	1	8.7	L, K, F, (B)
AL	Cytosol	2.4	7.5	L, K, F, (B)
A1	Cytosol	962	9.5	L, RBC, (K), (B)

B = brain; I = small intestine; K = kidney; L = liver; RBC = red blood cells.

ACUTE MANAGEMENT

Hyperammonemic crisis is often precipitated by stressful situations such as fever, trauma, infection, surgery, and dietary indiscretion. The goal of treatment is to lower the blood ammonia levels to prevent brain edema. The most effective treatment is hemodialysis or peritoneal dialysis. Nitrogen scavenger medications, such as sodium phenylacetate and sodium benzoate, and arginine are also given intravenously. Protein intake is halted and an intravenous fluid with a high dextrose (10%) load is provided immediately to minimize catabolism and therefore the source of ammonia. Intercurrent infection should be treated appropriately.

Future Horizons

Future treatment modalities may include total or partial organ transplant, stem cell transplant, and enzyme replacement. Early diagnosis utilizing newborn screening and prenatal diagnosis will allow early treatment and improve the outcome.

SEE ALSO THE FOLLOWING ARTICLES

Amino Acid Metabolism • Ornithine Cycle

GLOSSARY

autosomal recessive   A mode of genetic inheritance that describes a trait or disorder requiring the presence of two copies of a gene mutation on one of the 22 pairs of autosomes (nonsex chromosomes) at a particular locus in order to express observable phenotype.

enzyme   Biological reactants that are protein catalysts, mediating reactions without themselves being changed in the overall process. They selectively channel substrates into pathways.

essential amino acids   The amino acids that are synthesized by nonmammalian pathways and must be obtained in diet. They include histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine. Arginine is considered a "semi-essential" amino acid because mammals synthesize arginine, but cleave most of it to form urea and require greater amounts than can be produced in mammalian biosynthesis.

Lyon hypothesis   The X chromosome is randomly inactivated in early development in embryonic cells. This results in fixed inactivation in the female's descendant cells. The deactivated chromosome forms the Barr body.

non-essential amino acids   Amino acids that can be produced by mammals from common intermediates and biosynthesis and are not strictly diet dependent. These include alanine, asparagine, aspartate, cysteine, glutamate, glutamine, glycine, proline, serine, and tyrosine.

X-linked disorder   A mode of genetic inheritance that describes a gene mutation on the X chromosome that causes the phenotype to be expressed in males who are hemizygous for the gene mutation and in females who are homozygous for the gene mutation (meaning they must have a copy of the gene mutation on each of their two X chromosomes). Carrier females do not usually express the phenotype, although differences in X-chromosome inactivation can lead to varying degrees of clinical expression.

FURTHER READING

Berry, G., and Steiner, R. (2001). Long-term management of patients with urea cycle disorders. *J. Pediatrics* **138**(Suppl. 1), S56–S61.

Feillet, F., and Leonard, J. V. (1998). Alternative pathway therapy for urea cycle disorders. *J. Inher. Metab. Dis.* **21**(Suppl. 1), 101–111.

Fernandes, J., Saudubray, J. M., and Van den Berghe, G. (eds.) (2000). *Inborn Metabolic Diseases: Diagnosis and Treatment*, 3rd edition. Springer, Berlin.

Morizono, H., Caldovic, L., Shi, D., and Tuchmon, D. (2004). Mammalian N-acetylglutamate synthase. *Mol. Genet. Metab.* **81**, 4–11.

Nyhan, W. L., and Ozand, P. T. (eds.) (1998). *Atlas of Metabolic Diseases*. Chapman and Hall Medical, London.

Steiner, R., and Cederbaum, S. (2001). Laboratory evaluation of urea cycle disorders. *J. Pediatrics* **138**(Suppl. 1), S21–S29.

Summar, M., and Tuchman, M. (2001). Proceedings of a consensus conference for the management of patients with urea cycle disorders. *J. Pediatrics* **138**(Suppl. 1), S6–S10.

BIOGRAPHY

Marsha K. Fearing is a Clinical Fellow at the Harvard Medical School Genetics Training Program, specializing in Biochemical Genetics. She is a member of the Scholars in Clinical Science Program at Harvard Medical School. Her area of research is translational and public health research in metabolism. Dr. Fearing is a member of Alpha Omega Alpha, the American Academy of Pediatrics, and a candidate fellow of the American College of Medical Genetics.

Vivian E. Shih is Professor of Neurology at Harvard Medical School, Associate Neurologist and Pediatrician at Massachusetts General Hospital, and Director of the Clinical Neurochemistry and Amino Acid Disorders Laboratory at Massachusetts General Hospital. Her area of research is inborn errors of amino acid and organic acid metabolism.

Vacuoles

Christopher J. Stefan and Scott D. Emr
University of California, San Diego, La Jolla, California, USA

Vacuoles and their mammalian counterparts, lysosomes, are membrane-bound cytoplasmic organelles that contain an assortment of soluble acid-dependent hydrolases and a set of highly glycosylated integral membrane proteins. Most notably, this organelle is an important site for the degradation of cellular lipids, membrane-associated proteins, and cytoplasmic proteins. In addition to its degradative functions, the vacuole/lysosome plays an important role in pH and ion homeostasis and is a site for the generation/storage of nutrients, amino acids, antigens, and additional signaling factors.

Identification/Discovery of Vacuoles and Vacuolar Constituents

Vacuoles/lysosomes were first visualized by independent cell morphological studies undertaken in the 1950s by the Palade and Novikoff laboratories. These studies revealed that lysosomes appear as organelles with 0.5 μm diameter and a heterogeneous morphology, containing electron-dense cores and, in some cases, membrane vesicles. In the yeast *Saccharomyces cerevisiae*, vacuoles are larger with respect to cell size, but display similar characteristics in terms of electron density and intralumenal vesicles, when observed under certain conditions (Figures 1A and 1B).

Biochemical studies performed by de Duve and colleagues initially identified lysosomes as organelles of highly buoyant centrifugal properties that contained a number of hydrolytic activities. The hydrolases within this fraction demonstrated attenuated activity dependent on the integrity of their surrounding membrane. Specifically, hydrolase activity was retained within the limiting vacuole/lysosome membrane, restricted from the cell lysate unless the limiting membrane was disrupted. Since vacuoles/lysosomes are morphologically heterogeneous, biochemical studies have produced the current definition of vacuoles/lysosomes as membrane-bound acidic organelles that contain mature acid-dependent hydrolases and certain integral membrane-glycoproteins. Vacuoles lack other proteins that distinguish them from endosomes (compartments that deliver material to vacuoles), such as biosynthetic sorting receptors and recycling cell-surface receptors.

As already mentioned, vacuoles/lysosomes are membrane compartments that serve as a major degradative compartment in eukaryotic cells. Both endogenous and exogenous molecules can be delivered to vacuoles through biosynthetic and endocytic pathways. In addition, vacuoles can degrade proteins transported from the cytosol. The degradative function of these organelles is carried out by numerous acid-dependent hydrolases (e.g., proteases, lipases, glycosidases) contained within its lumen (Figure 2). The outer, limiting membrane of vacuoles also contains a set of highly glycosylated, vacuolar/lysosomal-associated membrane proteins. A subset exhibits known hydrolase activities (Figure 2), while the functions of many are still unclear. Additional vacuolar membrane proteins mediate transport of ions, amino acids, and other solutes across the vacuolar membrane and maintain an acidic lumenal pH in the range of 4.5–6.5 (Figure 2).

Vacuole Function in Yeast

The vacuole of the yeast *S. cerevisiae* plays an important role in protein degradation, pH regulation, ion homeostasis, and nutrient storage. The yeast vacuole contains a large number of hydrolases (proteases, lipases, glycosidases; see Figure 2) with various substrate specificities. Thus, it is a major site for degradation of cellular proteins, lipids, and even whole organelles.

A remarkable feature of vacuoles is their acidification. In yeast, a specific vacuolar proton pump ATPase (V-ATPase) mediates this process. Biochemical and genetic approaches have identified 14 subunits of the V-ATPase (encoded by the *VMA* genes for *v*acuolar *m*embrane *A*TPase). The V-ATPase subunit proteins form two complexes, the peripheral V_1 subcomplex, responsible for ATP hydrolysis, and the V_0 subcomplex, responsible for proton translocation across the membrane (Figure 2). Cells lacking a functional V-ATPase are viable, but require an intact endocytic pathway to the vacuole. Moreover, they fail to grow on media with neutral pH, indicating that vacuole acidification is an
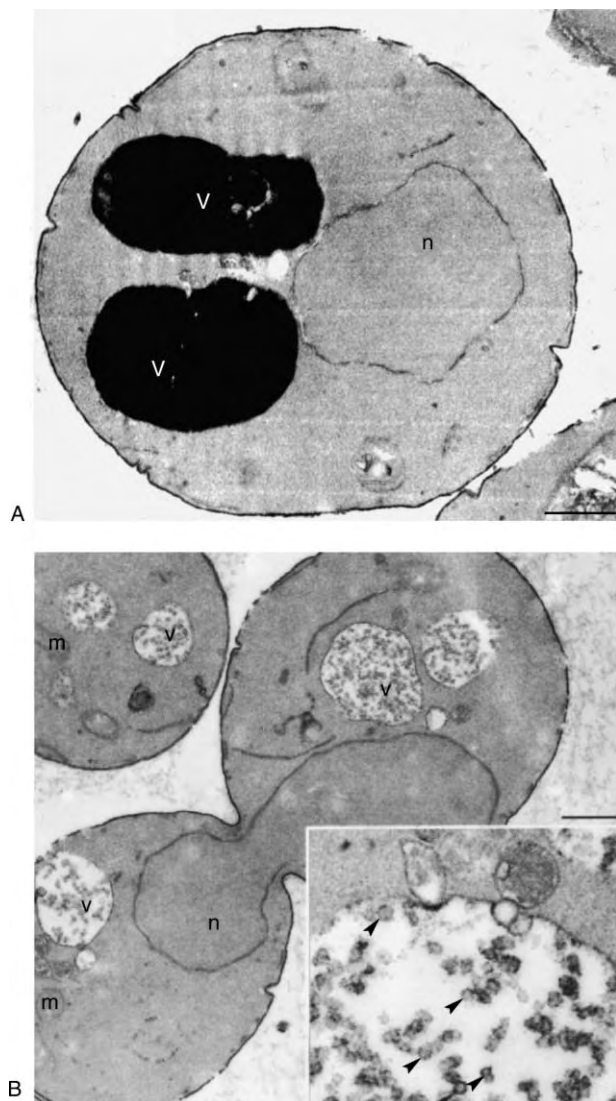
FIGURE 1   Ultrastructural analysis of vacuole morphology in wild type (A) and hydrolase deficient (B) yeast cells reveals electron dense cores and intralumenal membrane vesicles, respectively. (A) Wild type cells were grown to mid-log phase, fixed, and visualized by electron microscopy. Vacuoles appear as dark, electron dense organelles due to the accumulation of cellular proteins, lipids, and other molecules. (B) In cells lacking vacuolar ATPase activity (*vma4Δ* mutant cells), vacuoles appear translucent by electron microscopy and contain intralumenal vesicles. Inset: High magnification image of vacuolar vesicles (arrowheads) (v, vacuoles; n, nucleus; m, mitochondria; bars = 0.1 μm). (Reprinted with permission from Wurmser, A. E., and Emr, S. D. (1998). Phosphoinositide signaling and turnover: PtdIns(3)P, a regulator of membrane traffic, is transported to the vacuole and degraded by a process that requires lumenal vacuolar hydrolase activities. *The EMBO Journal* **17**, 4930–4942, copyright 1998 Oxford University Press.)

essential feature. Consistent with this, the electrochemical gradient generated by the V-ATPase is necessary to drive several other transport systems described in this article (Figure 2).

In addition to its hydrolytic functions, the vacuole serves as a storage compartment for several ions and
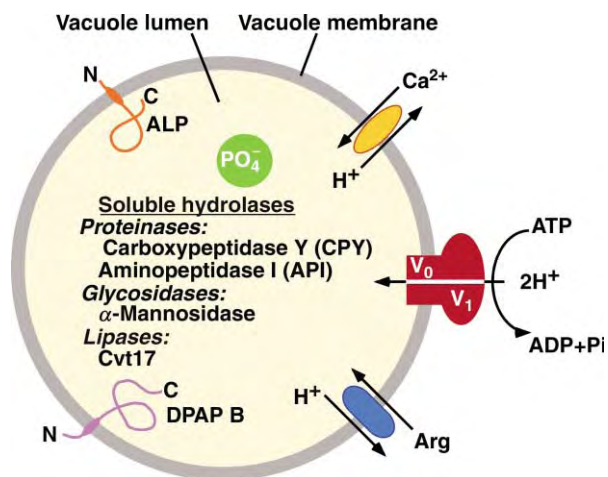
FIGURE 2   Schematic overview of vacuolar components. The vacuole lumen possesses several hydrolase activities and accumulates ions, such as phosphate (PO_4^-) and polyphosphates. The vacuolar membrane also contains known hydrolases, such as ALP, and proteases, such as dipeptidyl-aminopeptidase B (DPAP B). The vacuolar ATPase (V_0/V_1) functions as the primary proton pump that generates an electrochemical gradient used to drive other vacuole membrane transport systems, such as arginine (Arg) and divalent cation (Ca^{2+}) anti-transporters.

nutrients (Figure 2). For example, the Vcx1 protein acts as a high capacity H^+/Ca^{2+} exchanger *in vitro* and has been implicated in the regulation of vacuolar calcium flux *in vivo* (Figure 2). The vacuole also stores both phosphate and polyphosphate (Figure 2). Since hydrolysis of phosphate leads to the release of protons, a role for polyphosphate in pH homeostasis has been suggested. Polyphosphate is proposed to function as a cation counter ion as well. Some amino acids are also stored inside the vacuole at high concentrations. For example, arginine accumulates in vacuoles at concentrations nearly an order of magnitude greater than in the surrounding media (Figure 2). Consistent with this, several H^+/amino acid anti-porter systems have been identified in yeast. Pools of amino acids stored in the vacuole may be utilized during nitrogen starvation.

Vacuole Biogenesis and Transport Pathways to the Vacuole in Yeast

To carry out the numerous functions that vacuoles perform, several transport pathways have evolved that deliver cellular components to vacuoles. Model genetic systems such as the yeast *S. cerevisiae* have been instrumental in identifying both proteins and lipids that constitute components of vacuoles and the machinery involved in transport to vacuoles. Importantly, many of these factors and transport pathways are conserved in mammalian cells.

At least eight pathways have been identified in vacuole biogenesis and protein transport to vacuoles in

yeast: endocytosis from the cell-surface, two distinct biosynthetic pathways, the carboxypeptidase Y (CPY) and alkaline phosphatase (ALP) from the Golgi, multivesicular body (MVB) sorting, the cytoplasm to vacuole pathway (Cvt), macro-autophagy, micro-autophagy, and vacuole inheritance during cell division (Figure 3A). The endocytic pathway is essential for regulating levels of cell-surface proteins and intersects with the CPY pathway at a late endosomal compartment. Components of the CPY-sorting pathway are directly involved in vacuolar biogenesis, morphology, and function, as discussed here in greater detail. ALP travels from the Golgi to the vacuole along a pathway that is independent of the CPY and endocytic pathways. Proteins such as aminopeptidase I (API) are delivered from the cytoplasm to the vacuole through pathways involving products of the *CVT* and *APG* (autophagy) genes. In micro-autophagy, the vacuole membrane invaginates to engulf cytoplasmic material, including entire organelles such as peroxisomes. Vacuolar segregation into dividing daughter cells of budding yeast requires *VAC* (vacuolar inheritance) genes, some of which also participate in the CPY pathway. Previous studies have also identified several distinct factors involved in vacuolar inheritance.

THE CPY PATHWAY

CPY is the prototype of a subset of proteins that traffics from the Golgi to the vacuole via an endosomal intermediate. This pathway depends on the function of over 60 *VPS* gene products, mutations in which result in the mis-sorting of CPY to the secretory pathway, abnormal vacuole morphology, and in some cases, abnormal endosome morphology. In a manner analogous to receptor-mediated sorting in mammalian cells, the CPY receptor Vps10 binds CPY at the Golgi via a targeting signal. Sorting of this receptor–ligand complex into vesicles bound to endosomes requires proteins such as clathrin and accessory factors such as the AP-1 clathrin adaptor complex and the Gga proteins. Next, the class D Vps proteins are involved in CPY vesicle targeting and fusion with endosomes, such as Vps21 (Rab5 homologue), Vps9 (Rab GEF), Vac1 (EEA1), Vps45 (Sec1 homologue), Pep12 (t-SNARE), and Sec18 (NSF). At the endosome, Vps10 dissociates from CPY and recycles to the Golgi in a process requiring the retromer complex (consisting of Vps29, Vps26, Vps35, Vps5, and Vps17). Mutants defective in retromer function mis-localize Vps10 to the vacuole membrane and secrete CPY, as Vps10 becomes limiting for subsequent sorting reactions at the Golgi.

Another set of Vps proteins (the class E Vps proteins) is necessary for efficient sorting at the endosome. Class E Vps mutant cells accumulate abnormal/aberrant endosomes containing both biosynthetic cargo such as CPY and endocytosed proteins. The class E proteins are also involved in the formation of multivesicular bodies (MVBs) at late endosomes. The abnormal/exaggerated endosomes observed in these mutant cells form in part due to impaired MVB vesicle budding into the lumen of the endosome.

Finally, fusion of late endosomes/MVBs with the vacuole requires another set of proteins including Ypt7 (Rab7 homologue) and SNARE proteins (Vti1, Ykt6, Nyv1, Vam7, and Vam3). The class C Vps protein complex (also termed HOPS) consisting of Vps18, Vps11, Vps16, and Vps33 (Sec1 homologue), Vps41, and Vps39 (Rab GEF) is required for this final fusion step as well (Figure 3B). The Vps34 PI 3-kinase also contributes to this fusion step by the generation of PI(3)P that recruits effector proteins such as the PX domain-containing protein Vam7. Mutants with defects in these gene products accumulate numerous endosomal intermediates and MVBs that fail to fuse with the vacuole. Accordingly, mutants defective in components of the vacuolar fusion machinery display highly fragmented vacuoles and can sometimes even lack vacuoles entirely.

THE ALP PATHWAY

ALP is an integral membrane protein that travels from the Golgi to the vacuole independently of endosomal compartments that transport CPY and endocytic cargoes (Figure 3A). A specific adaptor complex, termed AP-3, mediates sorting of ALP into vesicles at the Golgi. However, following formation, fusion of ALP cargo vesicles with the vacuole is dependent on the class C Vps complex, Ypt7, Vam7, and Vam3. Thus, the ALP and CPY pathways converge upon common docking/fusion machinery at the vacuole (Figure 3B).

CYTOPLASM TO VACUOLE TRANSPORT AND MACRO-AUTOPHAGY

Autophagy is a trafficking pathway to vacuoles regulated by changes in nutrient availability. In macro-autophagy, induced during starvation, cytoplasmic material is first sequestered in double-membrane vesicles called autophagosomes and then subsequently delivered to the vacuole (Figure 3A). While autophagy is induced, the Cvt pathway constitutively packages the hydrolases aminopeptidase I (API) and α-mannosidase into autophagosomes for delivery to vacuoles. Autophagosomes are targeted to, and fuse with, the vacuole by the same machinery that mediates endosome-vacuole fusion (Figure 3B). Inside the vacuole, the lipase Cvt17 is responsible for autophagosome turnover (Figure 2).
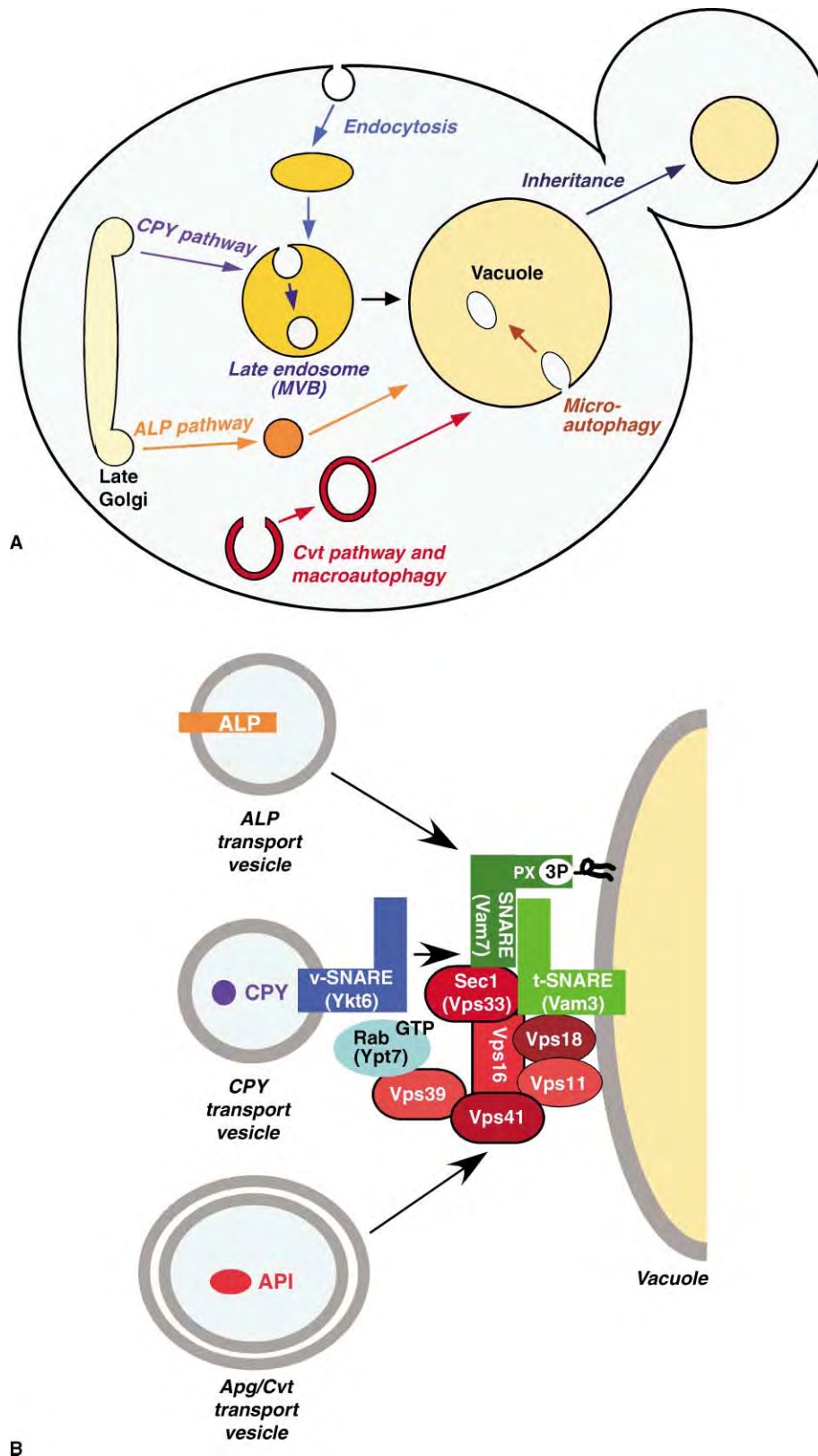
FIGURE 3 (A) Model representation of the transport pathways to the vacuole. Endocytosis of cell-surface proteins is followed by transport to early endosomes and then late endosomes. The CPY pathway sorts proteins from the late Golgi to the late endosome. At the late endosome, select cargo is sequestered into intralumenal vesicles to form multivesicular bodies (MVB) while other proteins, such as sorting receptors, are recycled to the Golgi. The ALP pathway functions as a Golgi to vacuole route that is independent of the late endosome/MVB. In the Cvt pathway, cytoplasmic proteins are enclosed within double membrane structures that subsequently fuse with and deliver their material into the

Vacuole/Lysosome Functions in Higher Eukaryotes

THE CONTRACTILE VACUOLE IN PROTOZOA

Protozoa living in fresh water exist in a hypotonic environment. Water flows across their plasma membrane since their cytosol is hypertonic to the environment. To adapt, many protozoa have an organelle, the contractile vacuole complex (CVC), which collects and expels excess water. Previous work shows that CVCs are composed of a central vacuole with radial arm-like extensions. The extensions are often divided into separate bundles of tubules and contain proton-translocating V-ATPase enzymes that provide an electrochemical gradient necessary for fluid collection. The membrane of the central compartment lacks V-ATPases, but can expand into a reservoir for fluid storage, and is capable of fusing with the plasma membrane. It is then that the central compartment undergoes contraction, resulting in fluid expulsion.

LYSOSOME-RELATED ORGANELLES

The mammalian vacuole-like lysosome also is a site for delivery of cellular materials targeted for degradation, as it is the terminal compartment of the endocytic, phagocytic, and biosynthetic pathways in mammalian cells. However, the idea that vacuoles/lysosomes are simply degradative sites has evolved, in part by the identification of lysosome-like compartments that perform additional cellular functions such as lysosomes that secrete their contents after fusion with the plasma membrane. Many properties of vacuoles/lysosomes are shared with this group of cell type-specific lysosome-related organelles, which include melanosomes (pigment granules), lytic granules, platelet-dense granules, basophil granules, nuetrophil granules, and dendrocyte MHC class II compartments (involved in antigen presentation). However, in addition to lysosomal proteins, these organelles contain cell type-specific components that are responsible for their specialized functions.

PATHOGEN-CONTAINING VACUOLES

The uptake of foreign objects by macrophages often provides an important line of immune defense. To some intracellular pathogens (such as certain protozoa) however, phagocytosis represents an opportunity to gain protected access within a host cell. Other types of pathogens, including *Salmonella* and *Shigella*, actively invade host cells by delivering effectors into the host cell that leads to their direct uptake into intracellular vacuoles. Regardless of the mode of entry, the resulting intracellular vacuole that contains the parasite undergoes a maturation process, involving numerous membrane-trafficking events, as well as transmembrane transport of nutrients. Maturation of the vacuole yields a unique intracellular environment where the parasites are not only provided with essential nutrients, but are also protected from destruction by the host.

DISEASES ASSOCIATED WITH LYSOSOMAL DISORDERS

Several human genetic disorders are associated with defects in vacuolar/lysosomal function and transport, such as I-Cell, Tay-Sachs', Pompe's, Galactosialdosis, and Gaucher's disease. I-Cell disease is manifested by the inappropriate targeting/transport of multiple lysosomal enzymes. Cells of these patients become highly vacuolated and contain numerous dense inclusion bodies. Patients with I-Cell disease display severe clinical defects including skeletal and neurological defects, delayed growth and psychomotor development, and often death before age five. Abnormalities in lysosome-related organelles have also been observed in human genetic diseases such as the Chediak-Higashi and Hermansky-Pudlak syndromes. The similarity of genes affected in these lysosomal diseases to genes involved in vacuolar transport in yeast further demonstrates the importance of understanding the molecular machinery involved in the biogenesis and function of vacuoles, lysosomes, and lysosome-related organelles. Further studies on vacuole biogenesis will likely shed light on additional aspects of vesicular transport in the endosomal, lysosomal, and secretory systems.

SEE ALSO THE FOLLOWING ARTICLES

Endocytosis • V-ATPases

GLOSSARY

active transport Use of energy to transport a substance, often across membranes from an area where it is in lower concentration to an

vacuole. Macro-autophagy is similar to the Cvt pathway but is induced primarily under starvation conditions. In micro-autophagy, the vacuole membrane invaginates to mediate the uptake of cytoplasmic material. (B) Schematic diagram of proteins that function in the final docking/fusion step of various cargo vesicles to the vacuole. SNARE proteins are represented by shaded rectangles (Vam3 and Vam7, green; Ykt6, blue). Components of the class C Vps complex (Vps11, 16, 18, 33, 39, and 41) that mediate SNARE pairing are shown in red. Tethering is also mediated by the Rab-like GTPase, Ypt7.

area where it is in higher concentration. Membrane proteins called transporters perform active transport.

endocytosis   Ingestion of particulate matter or fluid by phagocytosis or pinocytosis; i.e., bringing material into a cell by invagination of its surface membrane and then pinching off the invaginated portion to form an endosome or vacuole.

hydrolases   Enzymes that act as catalysts in the cleavage of covalent bonds with accompanying addition of water. Lipases are hydrolases that break down fatty acids and other lipids. Proteases specifically digest peptide bonds.

vacuole   A large, membrane-bound cytoplasmic organelle that functions in ingestion, digestion, excretion, and storage of water, sugars, proteins, lipids, ions, and other materials. Vacuoles and vacuole-related organelles are found in fungal, plant, protozoan, and mammalian cells. Most plant cells have a single vacuole that takes up much of the cell and helps maintain the shape/turgor of the cell.

vesicular transport   Trafficking of proteins and lipids from one cellular compartment to another by means of membrane-enclosed intermediates such as endosomes or organelle fragments.

FURTHER READING

Allen, R. D. (2000). The contractile vacuole and its membrane dynamics. *Bioessays* **22**, 1035–1042.

Bowers, W. E. (1998). Christian de Duve and the discovery of lysosomes and peroxisomes. *Trends Cell Biol.* **8**, 330–333.

Bryant, N. J., and Stevens, T. H. (1998). Vacuole biogenesis in *Saccharomyces cerevisiae*: Protein transport pathways to the yeast vacuole. *Microbiol. Mol. Biol. Rev.* **62**, 230–247.

Dell'Angelica, E. C., Mullins, C., Caplan, S., and Bonifacino, J. S. (2000). Lysosome-related organelles. *Faseb J.* **14**, 1265–1278.

Jones, E. W. (2002). Vacuolar proteases and proteolytic artifacts in *Saccharomyces cerevisiae*. *Methods Enzymol.* **351**, 127–150.

Knodler, L. A., and Steele-Mortimer, O. (2003). Taking possession: Biogenesis of the salmonella-containing vacuole. *Traffic* **4**, 587–599.

Klionsky, D. J., Herman, P. K., and Emr, S. D. (1990). The fungal vacuole: Composition, function, and biogenesis. *Microbiol. Rev.* **54**, 266–292.

Kornfeld, S., and Mellman, I. (1989). The biogenesis of lysosomes. *Annu. Rev. Cell Biol.* **5**, 483–525.

Mach, L. (2002). Biosynthesis of lysosomal proteinases in health and disease. *Biol. Chem.* **383**, 751–756.

Mullins, C., and Bonifacino, J. S. (2001). The molecular machinery for lysosome biogenesis. *Bioessays* **23**, 333–343.

BIOGRAPHY

Scott D. Emr is a Professor of Cellular and Molecular Medicine at the University of California, San Diego and an investigator of the Howard Hughes Medical Institute. His lab focuses on defining components of the core transport machinery as well as the regulatory apparatus that direct protein and membrane sorting to and from intracellular organelles, such as vacuoles.

Christopher J. Stefan is a Research Associate of the Howard Hughes Medical Institute.

Vascular Endothelial Growth Factor Receptors

Kenneth A. Thomas
Merck Research Laboratories, West Point, Pennsylvania, USA

Vascular endothelial growth factor receptors (VEGFRs) are a set of three homologous transmembrane receptor tyrosine kinases that bind vascular endothelial growth factors (VEGFs). VEGFRs are expressed primarily by endothelial cells lining the lumen of vascular and lymphatic vessels. VEGF-mediated activation of these receptors can induce endothelial cell migration and mitosis promoting the growth of blood vessels, denoted angiogenesis, and of lymphatic vessels, or lymphangiogenesis.

VEGFR Genes

EVOLUTION

The three VEGFRs are denoted VEGFR-1, VEGFR-2, and VEGFR-3, or Flt-1, KDR/Flk-1 and Flt-4, respectively. Each VEGFR gene encodes a protein composed of seven extracellular immunoglobulin (Ig)-like domains, a short transmembrane-spanning polypeptide and an intracellular portion containing a tyrosine kinase. The VEGFRs are most closely related to the hematopoietic receptor tyrosine kinases c-kit, c-fms and Flt3 and to the platelet-derived growth factor receptor (PDGFR)-α and -β, each of which contain five extracellular Ig-like domains and an intracellular tyrosine kinase. The VEGFR-1, -2, and -3 genes are clustered with the hematopoietic receptors and PDGFRs on chromosomes 13, 4 and 5, respectively, consistent with divergence from a common ancestral receptor tyrosine kinase gene.

STRUCTURE

Each of the three VEGFRs is encoded by 30 exons (NCBI genomic database, http://www.ncbi.nlm.nih. gov/genome/guide/human). The corresponding exon-coding regions, in each gene, are of similar size although some of the introns that separate them vary substantially in length resulting in total genomic DNAs ranging from $\sim$190 kb for VEGFR-1 to 47 kb and 46 kb for VEGFR-2 and -3, respectively. In all three genes, exon 1 encodes the secretory leader sequence, exons 2–15 span the extracellular region, exon 16 corresponds to the transmembrane-spanning polypeptide and exons 17–30 encode the cytoplasmic region including the tyrosine kinase containing a kinase–insert domain encoded by exon 21.

Gene Expression

Cellular differentiation status and responses to extracellular signals influence the expression of each of the VEGFR genes. Multiple transcription factor DNA-binding site consensus sequences are present within approximately the first 1 kb 5′ of the translational start site. Transcription factor binding sites within the first intron might also either augment or inhibit transcription.

VEGFR-1

Subsets of vascular endothelial cells, monocytes, dendritic cell precursors, and some types of smooth muscle cells express VEGFR-1. The basal promoter and 5′ sequences that control VEGFR-1 endothelial cell-selective expression contain Ets, Sp1, Egr-1, and CRE transcription factor binding site consensus sequences, several of which have been shown to be functional. In addition, a single hypoxic response element is located within the first 1 kb 5′ of the transcriptional start site, consistent with the observation that the transcription of VEGFR-1 is increased by hypoxia. Transcription is initiated downstream of a TATA box basal transcription factor binding site to generate a 7.5–8 kb mRNA.

VEGFR-2

VEGFR-2 is expressed primarily by vascular endothelial cells although a few other cell types can also express this receptor. Selective endothelial cell expression has been mapped to a region within the first 150 bp 5′ of the transcriptional start site that contains multiple Sp1, Ap-2, and NF-κB sites. *In vivo* expression of the corresponding murine VEGFR-2 gene has been studied using transgenic mice. Blood vessel targeted gene expression in developing mouse embryos is

controlled by sequences not only in the region 5' of the translational start site but also within the first intron. The human gene, which does not contain a TATA box, is transcribed as a 7 kb mRNA.

VEGFR-3

Several consensus transcription factor sequences have also been recognized in the VEGFR-3 gene 5' of the translational initiation site. A 1.6 kb 5' region has been reported to drive lymphatic expression. Full-length VEGFR-3 mRNA is transcribed as a 5.8 kb mRNA.

Protein Structure

EXTRACELLULAR DOMAINS

The primary translation products of the VEGFR-1, -2, and -3 genes are 1338, 1356, and 1363 amino acid residue proteins, respectively. Each $\sim$150 kDa protein is glycosylated at multiple sites on the extracellular Ig-like domains to generate mature proteins of 185–230 kDa. In addition, an alternatively spliced soluble 687 amino acid form of VEGFR-1, denoted sVEGFR-1 or sFlt-1, consists of the 6 N-terminal Ig-like domains. This 75 kDa protein is converted to 110 kDa by glycosylation. Partial proteolysis of VEGFR-3 generates an N-terminal 70 kDa polypeptide disulfide linked to the remaining 120–125 kDa transmembrane protein. A crystal structure of the VEGFR-1 Ig-like domain 2, composed of 5- and 3-stranded β-sheets, in complex with the dimeric VEGF-A ligand is shown in Figure 1.

INTRACELLULAR DOMAINS

The structure of the VEGFR-2 tyrosine kinase, shown in Figure 2, contains N- and C-terminal domains consisting primarily of β-strands and α-helices, respectively. The catalytic site is located at the interface of these two domains denoted by the ADP modeled into the structure. It is flanked by the disordered activation-binding loop, the glycine-rich nucleotide-binding loop, the catalytic loop and that can participate in enzyme activation and substrate binding. The VEGFRs each contain a 65–70 amino acid kinase–insert domain within the N-terminal region of the C-terminal domain that is deleted in the crystallized VEGFR-2 kinase. A C-terminally truncated form of VEGFR-3, missing 65 amino residues, is also generated by alternative splicing.

Ligand Binding

VEGF STRUCTURE AND RECEPTOR BINDING

The VEGFs are a family of five homologous dimeric glycoproteins. In addition, several VEGF homologues are expressed by the orf and pseudocowpox viruses, collectively denoted VEGF-E. HIV tat can also function as a VEGF. Each ligand binds with pM affinity either 1 or
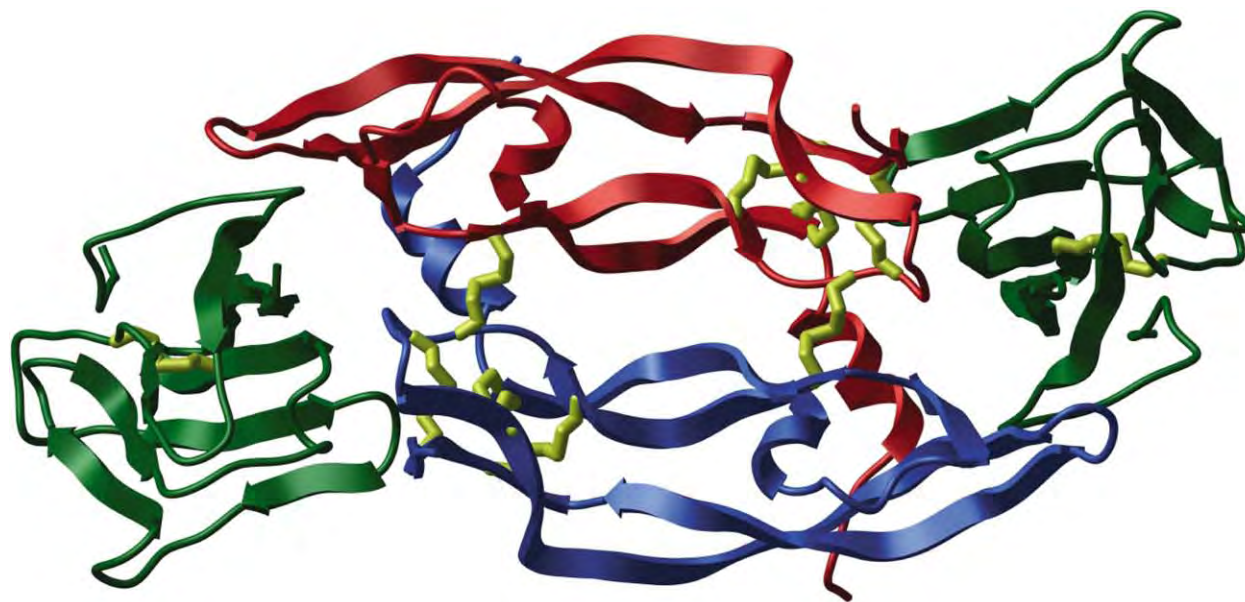
FIGURE 1   Structure of the complex between a VEGF-A dimer and domain 2 of VEGFR-1. The anti-parallel VEGF-A subunits are red and blue and the VEGFR-1 domains are green with disulfide bonds shown in yellow. Secondary structure is illustrated as arrows pointing toward the C-terminal ends. (Reproduced from Harada, S., and Thomas, K. A. (2002). Vascular endothelial growth factors. In *Principles of Bone Biology* (J. P. Bilezikian, L. G. Raisz and G. A. Rodan, eds.) 2nd edition, Vol 2, pp. 883–902. Academic Press, San Diego, with permission.)
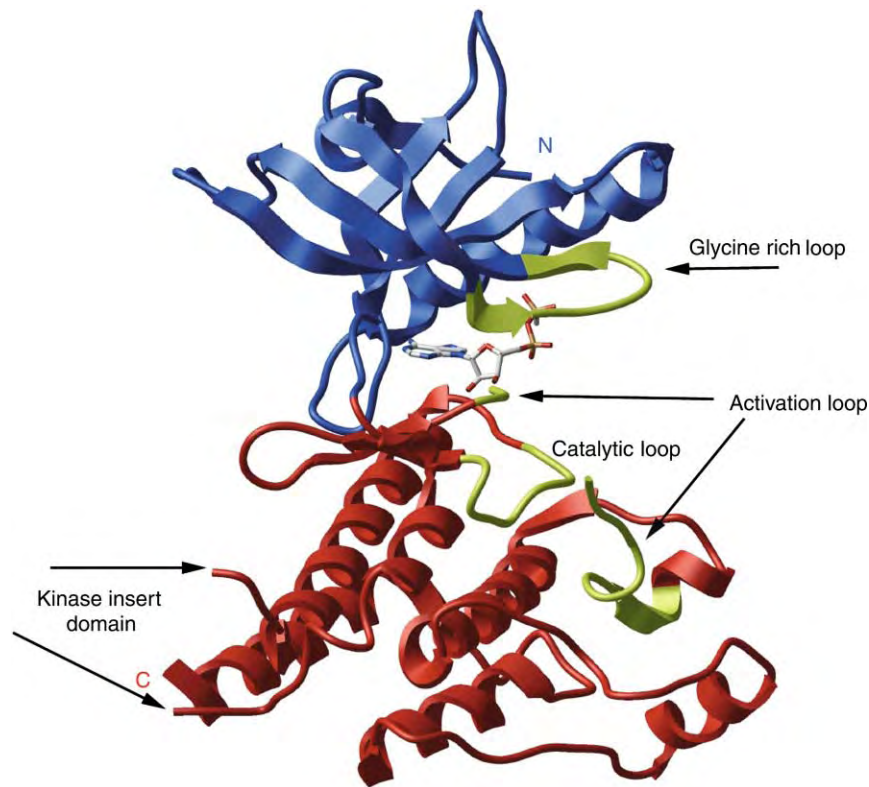
FIGURE 2 Structure of the VEGFR-2 tyrosine kinase. The N- and C-terminal lobes are blue and red, respectively. The active site region between these two domains is denoted by the location of ADP modeled into it. The glycine-rich, catalytic, and the ends of the disordered activation loops are shown in yellow. The ends of the largely deleted kinase-insert domain are also shown and labeled. (Reproduced from Harada, S., and Thomas, K. A. (2002). Vascular endothelial growth factors. In *Principles of Bone Biology* (J. P. Bilezikian, L. G. Raisz and G. A. Rodan, eds.) 2nd edition, Vol 2, pp. 883–902. Academic Press, San Diego, with permission.)

2 of the 3 receptors as shown in Figure 3. The core receptor-binding region of each subunit consists of ∼120 amino acid residues. VEGF subunits, each composed of four β-strands and two short helices, dimerize in an anti-parallel arrangement as illustrated by VEGF-A in Figure 1. The loops between β-strands of both VEGF-A subunits interact with the second and third Ig-like domains of each of two receptor subunits through primarily hydrophobic interactions. Therefore, VEGF binding can promote receptor dimerization, which is additionally stabilized by interactions between the fourth Ig-like domains in VEGFR-1, as illustrated in Figure 4. VEGF-C and -D bind with highest affinity to VEGFR-3, mainly expressed on lymphatic endothelial cells and on the tips of some growing capillaries. Partial proteolysis, which removes sequences N- and C-terminal of the core-receptor-binding region, can increase the affinity of VEGF-C and -D for VEGFR-2 as indicated by the dashed lines in Figure 3.

CORECEPTORS

VEGF-A, VEGF-B, and PlGF are subject to alternative splicing that either incorporates or deletes C-terminal polycationic regions. These positively charged sequences can promote binding to negatively charged soluble heparin and heparin sulfate proteoglycans and to the membrane-anchored proteins neuropilin-1 and -2. Although these receptors do not appear to induce VEGF signal transduction, they might indirectly promote VEGF activity by partitioning the ligands to
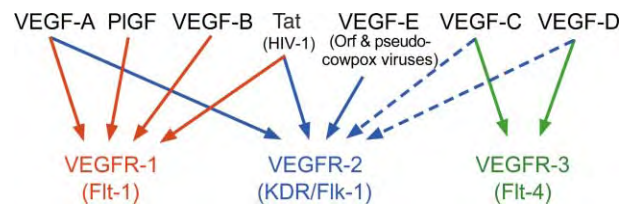
FIGURE 3 Receptor-ligand binding selectivity. VEGFs are listed in the top row with color-coded arrows pointing toward the high affinity VEGFRs that they bind. Tat and VEGF-E function as viral VEGFs. Dashed arrows from VEGF-C and -D to VEGFR-2 indicate high-affinity binding following full proteolytic processing to remove N- and C-terminal polypeptides. (Modified from Harada and Harada, S., and Thomas, K. A. (2002). Vascular endothelial growth factors. In *Principles of Bone Biology* (J. P. Bilezikian, L. G. Raisz and G. A. Rodan, eds.) 2nd edition, Vol 2, pp. 883–902. Academic Press, San Diego, with permission.)
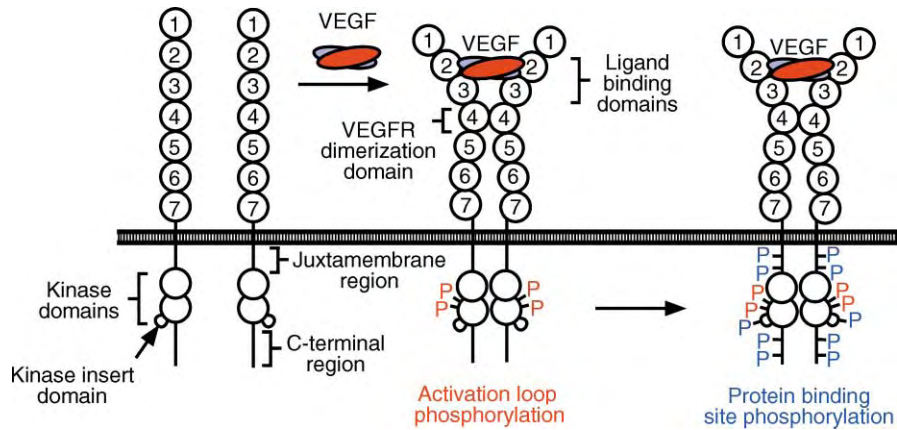
FIGURE 4   The structure of VEGFRs, containing seven extracellular Ig-like domains, a juxtamembrane region, the two domain tyrosine kinase containing a kinase-insert domain and a C-terminal tail, is shown on the left. Binding of dimeric VEGF to Ig domains 2 and 3 induce receptor dimerization, which can be aided by interactions between receptor domains 4. This promotes phosphorylation of two kinase activation loop tyrosines, denoted by the red "P" labels, and enzymatic activation in VEGFR-2 and probably in VEGFR-1 and -3. Additional tyrosines in the juxtamembrane region, kinase-insert loop and C-terminal tail are phosphorylated, as denoted by blue "P" labels, and bind signaling proteins that trigger several VEGF functions.

cellular surfaces, thereby increasing their local concentration in the vicinity of the high-affinity receptors, and by functioning as co-receptors that present the ligands to the high-affinity VEGFRs.

Signal Transduction

RECEPTOR ACTIVATION

Intracellular signaling is initiated by VEGF-induced receptor dimerization that brings the intracellular tyrosine kinases into proximity where they are thought to phosphorylate each other as shown in Figure 4. The initial phosphorylation of tyrosine residues 1054 and 1059 on the tyrosine kinase activation loop of VEGFR-2 decreases the K_m of the enzyme for ATP and for peptide substrates without altering K_{cat}, the effective maximal rate of enzyme activity. This increased affinity for substrates could reflect the movement of the phosphorylated activation loop to provide better substrate access to their binding sites. Equivalent activation loop tyrosines exist on VEGFR-1 and -3 so that they might function in a similar manner.

RECRUITMENT OF SIGNALING PROTEINS

Additional tyrosines in the juxtamembrane region, the kinase–insert domain and the C-terminal tail can be phosphorylated, as shown schematically in Figure 4. Some of these phosphorylated tyrosine residues have been shown to serve as recognition sequences for binding by one or more signal transduction proteins. These include adapter proteins such as Grb2, Grap, Nck, Crk, Sck, Shc, and Vrap, which can bridge the receptors to

other phosphorylated and non-phosphorylated proteins, and several enzymes such as phospholipase Cγ (PLCγ), phosphatidylinositol 3-kinase (PI3K), and the tyrosine phosphatase SHP-2. VEGF-induced phosphorylation is rapid, reaching maximal levels by 2-5 min, followed within 30 min by receptor internalization.

Some of these phosphorylated tyrosine residues have been linked with specific functions. For example, phosphorylation of VEGFR-2 Tyr$_{1175}$ in the C-terminal tail is required for the efficient phosphorylation of PLCγ and mitogenic activity. The longer alternatively spliced form of VEGFR-3 contains tyrosines within the unique C-terminal 65 amino acid residues. At least one of these tyrosines, Tyr$_{1337}$, is required for transforming activity *in vitro* and upon phosphorylation can bind the adapter protein Shc.

SIGNAL TRANSDUCTION CASCADES

The phosphorylation of VEGFR-2-associated PLCγ activates its enzymatic activity catalyzing the hydrolysis of membrane-anchored phosphatidylinositol 4,5-bisphosphate (PIP$_2$) to generate inositol 1,4,5-trisphosphate (IP$_3$) and diacylglycerol (DAG). Water soluble IP$_3$ activates an endoplasmic reticulum Ca^{2+} channel that releases Ca^{2+}, activating several enzymes including endothelial cell nitric oxide synthase (eNOS) and promoting the translocation of specific protein kinase C (PKC) isoforms to the membrane where it is activated by binding membrane-associated DAG. PKC appears to be able to activate the mitogen-activated protein kinase kinase (MAPKK) mitogenic pathway possibly involving Raf kinase. In addition, nitric oxide can activate protein kinase G (PKG) that subsequently activates Raf. MAPKK activates mitogen activated

kinase (MAPK), which can enter the nucleus to modulate transcription. Additional pathways, mediated through PI3K and the anti-apoptotic kinase Akt, promote cell adhesion, migration, and survival.

Biologic Activities

DEVELOPMENT

The developmental activities of VEGFRs have been revealed by mouse gene knockouts, which are each embryonically lethal. Blood vessels exist in VEGFR-1 knockout mice but are disorganized, which appears to be a consequence of modified cell fate leading to increased numbers of endothelial cell progenitor hemangioblasts that altered vessel pattern formation. Surprisingly, knockout of the VEGFR-1 gene segment encoding the tyrosine kinase but retention of the extracellular and membrane-spanning regions leads to normal development of blood vessels and survival. This result is consistent with the possibility that one of the functions of VEGFR-1, which binds VEGF-A with $\sim$ 10-fold higher affinity than VEGFR-2, might be to sequester low levels of VEGF, thereby "buffering" the VEGFR-2 receptor from inappropriate activation by low levels of VEGF. The VEGFR-1 cytoplasmic region might inhibit the VEGFR-2 mitogenic, but not migratory, activity by a mechanism involving VEGFR-1 juxtamembrane Tyr_{794}. The VEGFR-2 knockout is virtually devoid of vascular endothelial cells, consistent with its crucial role as a mediator of endothelial cell mitosis and survival. VEGFR-3 gene knockout mice exhibit vasculogenesis and angiogenesis in early developing embryos but large vessels appear abnormal with defective lumens leading to fluid accumulation in the pericardial cavity and cardiovascular failure. Therefore, at least during early embryogenesis before lymphatic vessels develop from post-capillary venules, VEGFR-3 plays a critical role in vascular development.

ANGIOGENESIS

In adults, VEGFR-2 is the primary angiogenic receptor family member. It can mediate the migration and mitosis of vascular endothelial cells culminating in neovascular growth. Ligands such as VEGF-A and the viral VEGFs that bind VEGFR-2 are sufficient to drive angiogenesis and the growth of some hematopoietic progenitor cells. Antibodies to VEGFR-2 and enzyme inhibitors of the VEGFR-2 tyrosine kinase inhibit angiogenesis and the resulting growth of tumors in mice.

Soluble VEGFR-1 (sVEGFR-1), which is expressed by cultured endothelial cells and has been detected in vivo, is the only identified naturally occurring specific VEGFR inhibitor. It retains the ligand-binding site and so recognizes the same VEGFs as full-length VEGFR-1. The soluble receptor can not only homodimerize but also heterodimerize with the extracellular regions of VEGFR-1 and VEGFR-2. Therefore, it can sequester ligands and perhaps inhibit activation of full-length membrane-spanning VEGFR-1 and VEGFR-2 by the formation of dominant negative heterodimers that contain a single tyrosine kinase so are incapable of kinase activation by trans-phosphorylation. Transfection experiments show that expression of sVEGFR-1 by tumor cells that express transfected sVEGFR-1 exhibit severely inhibited growth in vivo but not in vitro, consistent with an antiangiogenic mechanism.

LYMPHANGIOGENESIS

VEGFR-3 is expressed by lymphatic endothelial cells and also vascular endothelial cells at the tips of growing capillary shoots. The VEGFR-3 ligand VEGF-C is mitogenic for lymphatic endothelial cells in culture. Transgenic mice expressing elevated VEGF-C or -D in skin induce the growth of dermal lymphatic vessels but not blood vessels. However, inhibitory anti-VEGFR-3 antibodies can inhibit tumor angiogenesis. Naturally occurring human missense mutations with inactive tyrosine kinases are linked to lymphoedema, a genetic disease in which fluid accumulates in tissues as a consequence of deficient lymphatic function. Therefore, in adults VEGFR-3 appears to be active on lymphatic endothelial cells and at least some endothelial cells in actively growing capillaries.

NONMITOGENIC FUNCTIONS

In adults, activation of VEGFR-2 induces vascular permeability that can facilitate angiogenesis and promote edema. Although VEGFR-1 does not directly promote permeability, it might play a permissive role. Activation of VEGFR-1 by its selective ligands PlGF and VEGF-B does not appear to directly induce endothelial cell mitosis or angiogenesis under most conditions. However, it can induce vascular endothelial cell expression of specific proteins such as tissue factor, matrix metalloproteases, urokinase, plasminogen activator inhibitor-1, hepatocyte growth factor and pigment epithelium-derived factor. In addition, activation of monocyte and macrophage VEGFR-1 can drive expression of tissue factor, monocyte chemoattractant protein-1 and tumor necrosis factor-α. It can also induce the production of endothelial cell nitric oxide, and the migration of monocytes and some endothelial cells. Although the VEGFR-1 ligand PlGF is not a potent angiogenic agent, it does appear to promote the development of collateral blood vessels, a process involving the conversion of smaller to larger vessels.

Therefore, on the basis of the knockout and cell expression data, VEGFR-1 seems to modulate the expression of several genes associated with differentiated functions and the activity of VEGFR-2.

Summary

The homologous VEGFRs are critical mediators of the growth and function of blood and lymphatic vessels. These receptors and their ligands are intimately involved not only in development but also in the maintenance of normal differentiated functions. Their inappropriate activation or inhibition can have pathologic consequences so they also are potential targets for therapeutic intervention.

SEE ALSO THE FOLLOWING ARTICLES

Mitogen-Activated Protein Kinase Family • Phosphatidylinositol Bisphosphate and Trisphosphate • Phospholipase C • Protein Kinase C Family

GLOSSARY

angiogenesis Growth of new blood vessels from existing vessels.

lymphangiogenesis Growth of new lymphatic vessels.

vascular endothelial growth factor (VEGF) An extracellular soluble dimeric protein that binds and activates one or more VEGFRs.

vasculogenesis Embryonic *de novo* organization of blood vessels.

VEGFR Full-length membrane-spanning VEGF receptor containing an intracellular tyrosine kinase.

FURTHER READING

Claesson-Welsh, L. (2003). Signal transduction by vascular endothelial growth factor receptors. *Biochem. Soc. Trans.* **31**, 20–24.

Ferrara, N., Gerber, H.-P., and LeCouter, J. (2003). The biology of VEGF and its receptors. *Nat. Med.* **9**, 669–676.

Harada, S., and Thomas, K. A. (2002). Vascular endothelial growth factors. In *Principles of Bone Biology* (J. P. Bilezikian, L. G. Raisz and G. A. Rodan, eds.) 2nd edition, Vol 2, pp. 883–902. Academic Press, San Diego.

Larrivee, B., and Karsan, A. (2000). Signaling pathways induced by vascular endothelial growth factor (Review). *Int. J. Mol. Med.* **5**, 447–456.

McTigue, M. A., Wickersham, J. A., Pinko, C., Showalter, R. E., Parast, C. V., Tempczyk-Russell, A., Gehring, M. R., Mroczkowski, B., Kan, C.-C., Villafranca, J. E., and Appelt, K. (1999). Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2: A key enzyme in angiogenesis. *Structure* **7**, 319–330.

Robinson, C. J., and Stringer, S. E. (2001). The splice variants of vascular endothelial growth factor (VEGF) and their receptors. *J. Cell Sci.* **114**, 853–865.

Weismann, C., Fuh, G., Christinger, H. W., Eigenbrot, C., Wells, J. A., and de Vos, A. M. (1997). Crystal structure at 1.7Å resolution of VEGF in complex with domain 2 of the Flt-1 receptor. *Cell* **91**, 695–704.

Zachary, I., and Cliki, G. (2001). Signaling transduction mechanisms mediating actions of the vascular endothelial growth factor family. *Cardiovasc. Res.* **49**, 568–581.

BIOGRAPHY

Kenneth Thomas is Director of Growth Factor Research at the Merck Research Laboratories in West Point, PA. His general research interest is growth control in normal and pathologic conditions and he has led research efforts focused on the discovery and characterization of Fibroblast Growth Factors, VEGFs and their receptors. He holds a Ph.D. in biochemistry from Duke University and received postdoctoral training at the National Institutes of Health and Washington University School of Medicine in St. Louis. He is a member of the American Society for Biochemistry and Molecular Biology and the Association for the Advancement of Science.

Vasopressin/Oxytocin Receptor Family

Michael J. Brownstein
National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland, USA

Large neurons at the base of the brain (the hypothalamus) of humans and other mammals send their axons through the median eminence to terminate in the posterior lobe (neurohypophysis) of the pituitary. Each of these "magnocellular" nerve cells typically synthesizes one of two neurohypophyseal hormones – vasopressin (VP) or oxytocin (OT). In fact, all vertebrates higher than jawless fish make vasopressin- and oxytocin-like hormones. Jawless fish and invertebrates, on the other hand, appear to produce only one VP/OT-like peptide per species.

The Structures of "Pressins" and "Tocins"

For the last billion years, the peptides in VP/OT family have retained certain structural features. With the exception of the hydrins – partially processed relatives of vasotocin, which are physiologically active in some amphibia – they all have nine amino acids and a C-terminal amide. Cysteine (C) is invariably present in positions 1 and 6, and the two cysteine residues form a cystine bridge, creating conformationally constrained, cyclic peptides. Except for seritocin (serine5, isoleucine8-oxytocin), which was reported in a single species, *Bufo regularis*, all family members have asparagines (N) in position 5, and proline (P) and glycine (G) in positions 7 and 9, respectively. VP-like peptides ("pressins") are characterized by the presence of arginine (R) or lysine (K) in position 8, while OT-like molecules ("tocins") have leucine (L), isoleucine (I), glutamine (Q), valine (V), or threonine (T) there. All of the invertebrate peptides except for the locust diuretic hormone, which has not been detected in other insect species, have an R instead of Q, N, or S in position 4. Thus, the primordial peptide that gave rise to the molecules we know today should have looked rather similar to one of the conopressins or to vasotocin, the likely ancestor of all of the vertebrate hormones. (The molecules detected in *Hydra* with antivasopressin antibodies share the C-terminal PRG-amide motif with VP, but are otherwise unrelated.)

In addition to the peptides shown in Table I and Figure 1, it is worth noting that the major metabolite of VP (pGlu4, Cystine6-AVP) mobilizes calcium in some cells in the central nervous system, and has been suggested to have a unique, but as yet unidentified, receptor of its own.

Peptide Biosynthesis

The peptides in the VP/OT family are synthesized as parts of precursor proteins. The vasopressin precursor has the following structural features:

signal peptide-CYFNQNCPRG*G*-K*R*neurophysin-copeptin

The signal peptide is removed cotranslationally in the endoplasmic reticulum, liberating the N terminus of the peptide. Then, after the precursor is packaged into vesicles by the Golgi apparatus, it is cleaved at the asterisked sites by a Lys-Arg calcium-dependent endoprotease. The remaining basic amino acids are removed by a carboxypeptidase B-like enzyme, leaving a glycine (*G*) on the carboxy terminus. This glycine serves as the donor of the amide group, which is generated by the sequential actions of a peptidyl-glycine monooxygenase and a peptidyl-hydroxyglycine lyase. Neurophysin is thought to serve as an internal chaperone, promoting proper folding and cystine-bridge formation. Copeptin, a glycosylated species, is part of the vasopressin, but not the oxytocin precursor. Its function is unknown. The genes encoding the VP and OT precursors are adjacent to one another on the same chromosome. In mammals, they are oriented tail-to-tail, and share regulatory elements.

The Functions of Vasopressin and Oxytocin

VP/OT-like peptides play different roles in the various species where they are found. They are very commonly involved in water homeostasis and/or

TABLE I

Vasopressin, Oxytocin, and Related Peptides

Oxytocin-like peptides (Tocins)		
Oxytocin (OT)	C YIQ*NCP* LG-amide	Mammals, Pacific ratfish
Mesotocin (MT)	CYIQ*NCP*IG-amide	Nonmammalian tetrapods, marsupials, lungfish
[Phe2]Mesotocin (FMT)	CFIQNCPIG-amide	Australian lungfish
Seritocin (ST)	C YIQ*SCP* IG-amide	Dryness-resistant African toad
Isotocin (IT)	C YIS*NCP* IG-amide	Bony fish
Aspargtocin (AspT)	C YIN*NCP* LG-amide	Spiny dogfish
Asvatocin (AsvT)	C YIN*NCP*VG-amide	Spotted dogfish
Glumitocin (GT)	C YIS*NCP*QG-amide	Rays
Phasvatocin (PhT)	C YFN*NCP*VG-amide	Spotted dogfish
Valitocin (ValT)	C YIQ*NCP*VG-amide	Spiny dogfish
Annetocin (AnT)	C FVR*NCP*TG-amide	Earthworm
Cephalotocin (CT)	C YFR*NCP* IG-amide	Octopus
Vasopressin-like peptides (Pressins)		
A-Vasopressin (AVP)	C YFQ*NCP*RG-amide	Most mammals
L-Vasopressin (LVP)	C YFQ*NCP*KG-amide	Pig, marsupials
Phenypressin (PP)	C YFQ*NCP*RG-amide	Marsupials
Vasotocin (VT)	C YIQ*NCP*RG-amide	Non-mammalian vertebrates
Hydrin 2 (H2)	CYIQ*NCP* RGG	Frogs, toads
A-conopressin (ACP)	C IIR*NCP* RG-amide	Snail
L-conopressin (LCP)	C FIR*NCP* KG-amide	Snail, sea hare, leech
Locust diuretic hormone (LDH)	C LIT*NCP* RG-amide	Locust (but not fruitfly)
Putative Ancestor	C(F/Y)I(Q/R)*NCP*(R/K)G-amide	

Abbreviations: C (cysteine), F (phenylalanine), G (glycine), I (isoleucine), K (lysine), N (asparagine), P (proline), Q (glutamine), R (arginine), S (serine), T (threonine).

reproductive function. In humans vasopressin is an antidiuretic hormone. It is released from the posterior pituitary into the blood stream when the osmolarity (solute content) of the blood increases, and it acts on V2 receptors in the kidney to cause water retention. V2 receptors are Gs-coupled, and increase intracellular cyclic AMP. This, in turn, causes preformed aquaporin (AQP) 2 channels to be inserted into the luminal surfaces of cells in the collecting duct of the kidney. Water enters these channels, traverses the cells, exits through AQP 3 and 4 channels on their interstitial faces, and enters the bloodstream. Mutations in the genes encoding vasopressin, the V2 receptor, or AQP2 cause diabetes insipidus – an inability to concentrate urine, resulting in high urine volumes and a need to consume large amounts of fluid each day. Since the V2 receptor gene is on the X chromosome, female carriers pass X-linked nephrogenic diabetes insipidus along to their sons, but not their daughters.

Nonosmotic as well as osmotic stimuli can trigger vasopressin secretion. Osmoreceptors on cells in the anterior hypothalamus, and baroreceptors in the left ventricle of the heart, aortic arch, and carotid sinus monitor osmolarity and blood pressure, respectively, and participate in controlling the firing of VP-producing cells.

In addition to V2 receptors, vasopressin acts on V1a and V1b receptors. The former are found on blood vessels and are responsible for VP's pressor (blood pressure increasing) activity. The physiological importance of this is moot, but it is clear that certain vascular beds are especially sensitive to VP, including those in the skin and uterus. In fact, it has been suggested that V1a receptor antagonists might be useful for treating Raynaud's disease (excessive constriction of digital arteries) and dysmenorrhea (menstrual cramps which may be caused by dilation of vessels in the uterus).

V1b receptors are found on corticotrophs (adrenocorticotropic hormone or ACTH-producing cells) in the anterior pituitary. In addition to being made by magnocellular neurons, VP is also synthesized by small cells in the paraventricular nucleus of the hypothalamus. These same cells also make corticotropin releasing hormone (CRH), and their axons terminate on portal vessels in the external zone of the median eminence. VP and CRH are released into these vessels and transported to the anterior pituitary, where they act in concert to stimulate ACTH secretion. ACTH, in turn, causes the adrenal to release glucocorticoids. Thus, VP is important in mediating the body's response to stress.
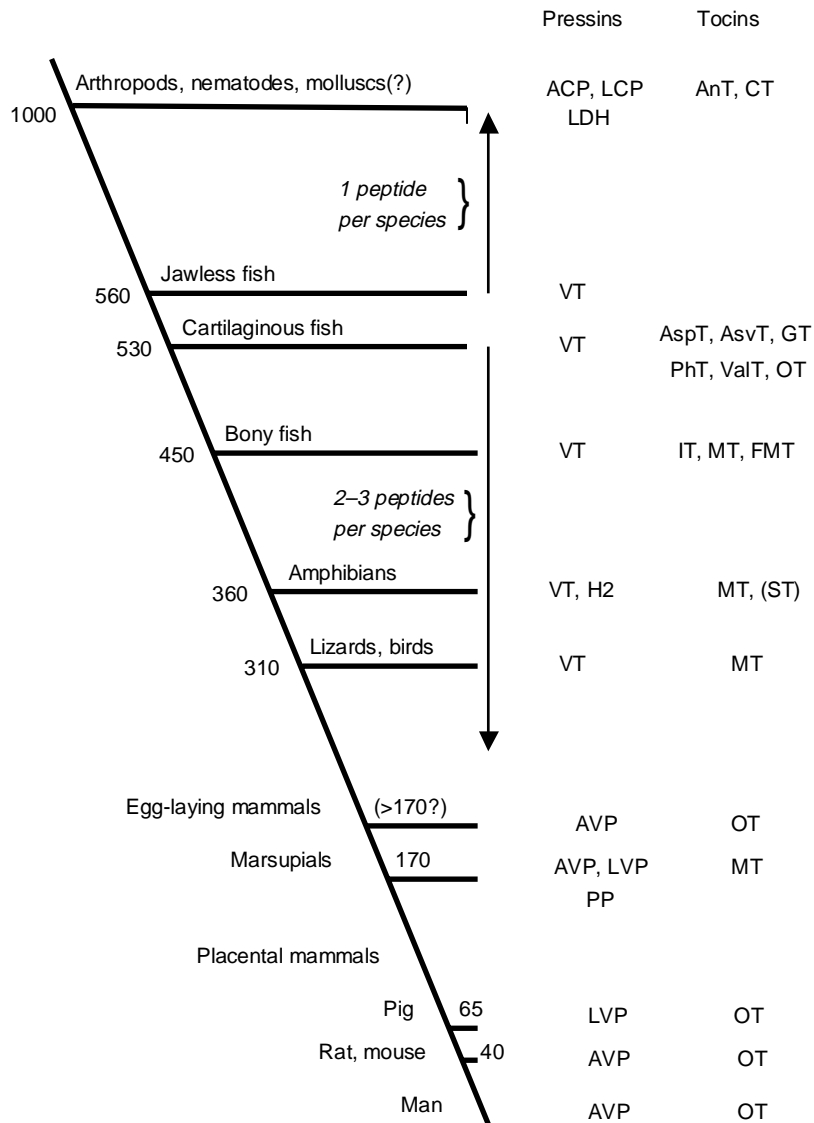
FIGURE 1 Evolution of vasopressin-related (pressins) and oxytocin-related (tocins) peptides. The timescale (millions of years ago) was taken from Hedges (2002). Note that invertebrates and jawless fish make a single pressin or tocin per species, and that vertebrates make two or, in the case of marsupials, three such peptides – one or two pressins and one tocin. It appears that vasotocin was the ancestor of the vertebrate hormones, and it continued to be synthesized by all vertebrates except mammals, in which it was replaced by arginine or lysine vasopressin. Tocins began to be produced in fish. It is not clear why cartilaginous fish make so many of these; lungfish make MT or FMT, and all of the remaining bony fish make IT. Based on the structures of the peptide precursors, it has been argued that lungfish are related to tetrapods (e.g., toads) and that bony fish form a separate lineage. Amphibians, lizards, birds, and marsupials make MT; egg-laying and placental mammals differ in the hormones they produce. Marsupials appear to have branched away as suggested by other analyses. The abbreviations used in this figure are defined in Table I.

V1a and V1b, but not V2, receptors are found in the brain. The central effects of vasopressin agonists and antagonists must be mediated by these receptors. These effects vary from species to species, and they are unknown in humans. It appears, however, that VP may mediate behavioral reactions to stress, aggressive and affiliative behavior, juvenile recognition, and parenting in rodents. In addition, V1a antagonists appear to block the vomiting associated with motion sickness in pigs and ferrets. This appears to

be a central effect of the drugs, but could have a peripheral component.

Oxytocin-producing mammals make a single receptor for this peptide. Like the V1a and V1b receptors, it is Gq-coupled, and activates phospholipase C, increasing intracellular calcium. Oxytocin receptors increase dramatically in the pregnant uterus as term approaches. Despite the fact that it has been used for decades to induce uterine contractions, oxytocin is not essential for this process, and oxytocin receptor antagonists have not

```
                        R1
Human V1B        G R D E E L A K V E I G V L A T V L V L A T G G N L A V L L T L G Q L G R K - - R S R M H L F V L H L A L T D L A V A L
Tree frog MT     K R N E E I A K V E V T V L A L I L F L A L A G N I C V L L G I Y I N R H K - - H S R M Y F F M K H L S I A D L V V A V
Human OT         R R N E A L A R V E V A V L C L I L L L A L S G N A C V L L A L R T T R Q K - - H S R L F F F M K H L S I A D L V V A V
Human V1A        V R N E E A K L E I A V L A V T F A V A V L G N S S V L L A L H R T P R K - - T S R M H L F I R H L S L A D L A V A F
White sucker IT  G R N E E V A K M E I T V L S V T F V V A V I G N L S V L L A M H N T K K K - - S S R M H L F I K H L S L A D M V V A F
Flounder AVT     G R N E E V A Q I E I M V L S I T F V V A V I G N V S V L L A M Y N T K K K - - M S R M H L F I K H L S L A D L V V A F
Chicken VT1      E R D E Q L A Q V E I A V L G V I F L T A S V G N F I L I L V L W R R R K K - - L S R M Y V F M L I H L S I A D L V V A F
Human V2         T R D P L L A R A E L A L L S I V F V A V A L S N G L V L A A L A R R G R R G H W A P I H V F I G H L C L A D L A V A L
Snail CP1        G V D E D L L K I E T A V Q A T I L Y M T L F G N G I V L L V L R L R R Q K - - L T R M Q W F I A H L A F A D I F V G F
Snail CP2        G R D E S L A R I E I L L V Q S I I L A L A I I A G N S C V L T A L A R R G K A - - A S R M H L F I F H L S I A D L L V A V
                 ──────────────────────────────────────────────────────          ──────────────────────────────────
                              TM1                                                            TM2

                     Q1      Q2          V            C1    K   Q3                            R2
Human V1B        F Q V L P Q L L W D I T Y R F Q G P D L L C R A V K Y L Q V L S M F A S T Y M L L A M T L D R Y L A V C H P L R S L Q Q
Tree frog MT     F Q V L P Q L I W D I T F R F Y A P D F V C R L V T Y L Q V V G M F A S T Y M L L L M S L D R C L A I C Q P L R S L H R
Human OT         F Q V L P Q L L W D I T F R F Y G P D L L C R L V K Y L Q V V G M F A S T Y L L L M S L D R C L A I C Q P L R S L R R
Human V1A        F Q V L P Q M C W D I T Y R F R G P D W L C R V V K H L Q V F G M F A S A Y M L V V M T A D R Y I A V C H P L K T L Q Q
White sucker IT  F Q V L P Q L C W E I T F R F Y G P D F L C R I V K H L Q V L G M F A S T Y M M V M M T L D R Y I A I C H P L K T L Q Q
Flounder AVT     F Q V L P Q L C W E I T Y R F F G P D F L C R I V K H L Q V T G M F A S T Y M M V M M T L D R Y I A I C H P L K T L Q Q
Chicken VT1      F Q V L P Q L I W D I T D V F I G P D F L C R I I K Y L Q L L G M F A S T Y M I V V M T V D R Y Q A V C Y P M V T F Q K
Human V2         F Q V L P Q L A W K A T D R F R G P D A L C R A V K Y L Q M V G M Y A S S Y M I L A M T L D R H R A I C R P M L A Y R H
Snail CP1        F N I L P Q L I S D V T I V F H G D D F T C R F T K Y F Q V I A M Y A S S Y V L V M A A I D R Y L S I C H P L T S Q T L
Snail CP2        F N I L P Q L I W D I T E R F Y G G D L L C R I K F M Q V Y M Y L S T Y M L V M T A V D R Y R A V C H P L S A F N T
                 ────────────────────                      ─────────────────────────────────────────
                                                                           TM3

                                             Q4                          C2
Human V1B        P G - Q S T Y L L I A A P W L L A A I F S L P Q V F I F S L R E V I Q G S G V L D C W A D F G F P W G P R A Y L T W T T
Tree frog MT     - - - R S D C V Y V L F T W I L S F L L S I P Q T A I F S L T E V N G - - V Y D C R A D F I L P W G P K A Y I T W I T
Human OT         - - - R T D R L A V L A T W L G C L V A D G - - V F D C W A V F I Q P W G P K A Y I T W I T
Human V1A        P A - R R S R L M A A A W V L S F V L S T P Q Y F V F S M I E V N N V T K A R D C W A T F I Q P W G S R A Y V T W M T
White sucker IT  P T - Q R A Y I M I G S T W L C S L L L S T P Q Y F I F S L S E I Q N G S Y V Y D C W G H F I E P W G I R A Y I T W I T
Flounder AVT     P T - Q R S Y I M I V S T W M C S L V F S T P Q Y F I F S L S E V K N G S T V K D C W A H F I E P W G A R A Y I T W I T
Chicken VT1      K R - A L W N I P I C T S W S I S L I S L P Q V F I F S K I E I S P G - - I F E C W A E F I Q P W G P R A Y V T W I L
Human V2         G S G A H W N R P V L V A W A F S L L L S L P Q L F I F A Q R N V E G G S G V T D C W A C F A E P W G R R T Y V T W I A
Snail CP1        S P - K R V H L M I A L A W L I S F L C A L P Q V F I F S L Q A V G P D - - Q Y D C L A T F E P D W G M Q A Y I T W V F
Snail CP2        S T R T P L Y C M I V S A Y V I S G V L S L P Q P I I F K Y R E K S H G S G D Y E C W G R F E P P W T L N L Y I T S F T
                 ──────────────────────────────────────                              ──────────────────────
                                 TM4

Human V1B        L A I F V L P V T M L T A C Y S L I C H E I C K N L K V K T Q A W R V G G G G - - - - - - - - - - - - - - - - - W R T W D
Tree frog MT     L S V Y I I P V L I L S I C Y G L I S Y K I W Q N I R L K T V C E S N - - - - - - - - - - - - - - - - - - - - - - - L R
Human OT         L A V Y I V P V I V L A T C Y G L I S F K I W Q N L R L K T A A A A A - - - - - - - - - - - - - - - - - - - - - - A E
Human V1A        G G I F V A P V V I L G T C Y G F I C Y N I W C N V R G K T A S R Q S K G A E - - - - - - - - - - - - - - - - - Q A G V A
White sucker IT  V G I F L I P V I I L M I C Y G F I C H S I W K N I K C K T M R G T R - - - - - - - - - - - - - - - - - - - - - - - N
Flounder AVT     G G I F L V P V V I L V M C Y G F I C H T I W K N I K Y K K R K T I P G - - - - - - - - - - - - - - - - - - - - - A A
Chicken VT1      V V I F F I P S T L I I T C Q V K I C K I I K R N I Y K K Q N E Y Q - - - - - - - - - - - - - - - - - - - - - - - - G
Human V2         L M V F V A P T L G I A A C Q V L I F R E I H A S L V P G P S E R P G - - - - - - - - - - - - - - - - - - - - - - - G
Snail CP1        V A N Y V I P F L L L A F C Y G R I C H V V W M S V A A K E S A A Y S S M R N G C T E S S R P - - - - - I K M R I S F H
Snail CP2        F A V Y I V P L A I L I F A Y S I C C T I W R K Y K S A E N E R K H M L N G S D S S L G N R N I Y S N H V T H S A L F
                 ──────────────────────────────────
                         TM5

Human V1B        R P S P S T L A A T - - - - - - - - - T R G L P S R - - - - V S S I N T I S R A K I R T V K M T F V I V L A Y I A C W A P
Tree frog MT     L S T S R R A T - - - - - - - - - - - - L S R - - - - V S S V R L I S K A K I R T V K M T F I I V L A Y I V C W T P
Human OT         A P E G A A A G D G - - - - - - - - G R V A L A R - - - - V S S V K L I S K A K I R T V K M T F I I V L A F I V C W T P
Human V1A        F Q K G F L L A - - - - - - - - - - - P C - - - - V S S V K S I S R A K I R T V K M T F V I V L A Y I V C W A P
White sucker IT  T K D G M I G K - - - - - - - - - - - V S - - - - V S S V T I I S R A K L R T V K M T F V I V L A Y I V C W A P
Flounder AVT     S K N G L I G K - - - - - - - - - - - N S - - - - V S S V T I I S R A K L R T V K M T F V I V L A Y I L C W A P
Chicken VT1      T N Q K Q V L P - - - - - - - - - - - S R - - - - A S S V N C L S K A M I K T V K M T I V T V V A Y I L C W S P
Human V2         R R R G R R T G - - - - - - - - - - - S - - - - - P G E G A H V S A A V A K T V R M T L V I V V Y V L C W A P
Snail CP1        R R R D N T N A T L T S L D R H D A S A V T S S D S K K P R G H Q R G V S K S K M K T I K L T L T V V L C Y L F C W A P
Snail CP2        R H R G V I E R R R N L V Q R C R P A P M A A P R - - - - A H S L R G F S R A K L K T V K L T F V I V A Y V V C W S P
                                                                                           ────────────────────────
                                                                                                     TM6

                     Q5                                     C3
Human V1B        F F S V Q M W S V W D K N A P D E D S T N V A F T I S M L L G N L N S C C N P W I Y M G F N S H L
Tree frog MT     F F F V Q M W S V W D P N P P K E - - - A S L F I I A M L L G S L N S C C N P W I Y M L F T G H L
Human OT         F F F V Q M W S V W D A N A P K E - - - A S A F I I A M L L A S L N S C C N P W I Y M L F T G H L
Human V1A        F F I I Q M W S V W D P M S V W T E S E N P T I T I T A L L G S L N S C C N P W I Y M F F S G H L
White sucker IT  F F I V Q M W S V W D E N F S W D D S E N A A V T L S A L L A S L N S C C N P W I Y M L F S G H L
Flounder AVT     F F T V Q M W S V W D E N F Q Y A D S E N T A V T L S A L L A S L N S C C N P W I Y M I F S G H L
Chicken VT1      F F I A Q L W S V W F P S G I T E G - - - S A F T I I M L L G N L N S C T N P W I Y M Y F C G H I
Human V2         F F L V Q L W A A W D P E A P L E G - - - A P F V L L M L L A S L N S C T N P W I Y A S F S S S V
Snail CP1        F F V Q M W S A F D D D S G I E H - - - P V T V I C M L L A S L N S C T N P W I Y L A F S G R T
Snail CP2        F F L S Q L W W L Y D E Q Q E H N - - - - H A V V I M V L L A S L N S C C N P W I Y L A F S G N L
                 ──────────────────                    ──────────────────────────────────────────────────
                                                                           TM7
```

proven useful for treating preterm labor. There is no doubt, however, that the pulsatile release of OT from the pituitary in response to cervical dilation and vaginal stimulation, facilitates the expulsion of the fetus.

Oxytocin is required for milk ejection. Mechanical stimulation of pressure sensitive receptors in the nipple of the breast by the nursing infant results in activation of magnocellular neurons in the hypothalamus and release of pulses of OT into the bloodstream. The hormone causes breast myoepithelial cells to contract, increasing intramammary pressure and forcing milk into the ducts. In the absence of OT, milk cannot be let down, and the infant will starve if it is not provided an alternative source of food.

In addition to its roles in parturition and lactation, oxytocin appears to affect maternal and social behaviors, stimulate lipogenesis to compensate for lipid loss in the milk (via an action on insulin secretion), and possibly participate in regulating salt and water balance. While OT causes natriuresis in rats, it is not clear that this is the case in humans.

Vasopressin and Oxytocin Receptors

As expected from the fact that their ligands are similar, VP and OT receptors are structurally related. They are members of the rhodopsin superfamily, and have seven α-helical membrane-spanning domains connected to one another by intracellular and extracellular loops. The N terminus of each receptor faces the outside of the cell; the C terminus is cytoplasmic. The intracellular loops and C-terminal tail of the receptors interact with G proteins, coupling agonist binding to activation of second messenger systems. More than 40 VP/OT receptors found in species ranging from snails to humans have been cloned and sequenced. The primary sequences of some of these are shown in Figure 2. It is remarkable that from the beginning of its first transmembrane domain (TM1) to the end of its seventh one (TM7), the snail conopressin receptor 2 is 43% identical in amino acid sequence to the human V1a, V1b, and OT receptors and the white sucker fish vasotocin receptor. Unlike the vertebrate proteins, however, the conopressin receptor responds equally well to lysine8- and isoleucine8-conopressin (an OT-like synthetic

analogue of lysine-conopressin). Duplication of a relatively promiscuous receptor of this sort might have permitted trial-and-error evolution of functionally distinct pressors and tocins in vertebrates.

A number of amino acids are conserved among all of the receptors in Figure 2. Some of these residues are found in most G-protein-coupled receptors. Among them, the arginine (R2) in the DRY motif, found just beneath TM3, is thought to dwell in a pocket formed by polar residues in TMs 1, 2, and 7 when the receptor is in its inactive state. Hormone binding dislodges this arginine from its polar pocket, exposing G-protein docking sites on the cytoplasmic loops.

The cysteines in extracellular loops 1 and 2 (C1 and C2, respectively) are also highly conserved among rhodopsin-like receptors. They form a cystine bridge that links these loops, stabilizing the conformation of the receptors. The pair of cysteines (C3 and its neighbor) located 15 aa's below TM7 in the cytoplasmic tails of most VP/OT receptors are likely to be palmitoylated and are thought to anchor their C termini to the plasma membrane. Like other members of the rhodopsin superfamily, VP and OT receptors appear to be glycosylated on their N termini, and regulated by phosphorylation of their intracellular domains.

A number of attempts have been made to model the binding of VP, OT, and vasotocin to their receptors. The models are fundamentally similar in the sense that they all predict that the peptide hormones fill a cleft located in the upper third of the barrel formed by the seven membrane-spanning α-helices. The hydrophobic amino acids that comprise the cyclic portion of the peptides (cysteine1, tyrosine2, isoleucine or phenylalanine3, and cystein6) appear to reside in a hydrophobic pocket formed by aromatic residues on helices 5 and 6 (and adjacent helices). The more polar amino acids (asparagine4, glutamine5), and the amidated C terminus of the hormones must occupy a hydrophilic region formed by residues on helices 2, 3, and 4. More specifically, residues that are conserved in the N-terminal domain and TMs 2, 3, 4, and 6 in most, if not all, of the VP/OT receptors (labeled R1, Q1, Q2, K, Q3, Q4, and Q5 in Figure 2) have been shown to be important for high-affinity binding, even though their predicted interactions with specific amino acids in the peptide hormones vary from model to model. Parsimony dictates that residues conserved among the various

FIGURE 2 Alignment of vasopressin and oxytocin receptors and selected relatives. To save space, the extracellular N termini and intracellular C termini have been removed. They are quite divergent, but it is remarkable that the transmembrane domains (TMs), the first two extracellular loops (linking TMs 2 and 3, and TMs 4 and 5), and portions of intracellular loops 2 and 3 (linking TMs 3 and 4, and TMs 5 and 6), have remained so similar throughout the course of their evolution. Specific residues in these domains are responsible for ligand binding and selectivity (see text), and other motifs are important for signal transduction. The variable intracellular portions of the receptors allow them to interact with specific G proteins.

VP/OT peptides should interact with conserved domains in the receptors, but perhaps things are not this simple.

One amino acid appears to be crucial for peptide agonist selectivity. This residue (marked with a V in Figure 2) is found in the first extracellular loop, and it interacts with the eighth amino acid of the peptide hormones (arginine and leucine in AVP and OT, respectively). The aspartic acid or tyrosine residues found at V in the human V2 and V1a/V1b receptors, respectively, are responsible for their marked preferences for AVP versus OT. The phenylalanine in this position in the OT receptor accounts for its modest preference for OT over AVP.

VP and OT antagonist-binding sites appear to be different from the ones where the peptide hormones bind.

SEE ALSO THE FOLLOWING ARTICLES

Amino Acid Metabolism • Phospholipase C

GLOSSARY

baroreceptor Receptor in the walls of the heart or blood vessels that is stimulated by alterations in pressure.

diuresis Increased urine excretion.

diuretic An agent that increases urine excretion.

hypothalamus The part of the brain that regulates the endocrine and autonomic and autonomic nervous systems, controlling water balance, blood pressure, body temperature, growth, and sexual function.

lipogenesis Formation of body fat.

magnocellular neurons Large neurons in the hypothalamus that manufacture vasopressin or oxytocin.

natriutesis Sodium excretion in the urine.

pressin An agent that increases blood pressure. In the context of this review, "pressins" are vasopressin-like peptides with nine amino acids, having a basic residue (arginine or lysine in the eighth position).

tocin An agent that promotes childbirth by causing uterine contractions; an oxytocin-like peptide lacking arginine or lysine in the eighth position.

FURTHER READING

Acher, R., and Chauvet, J. (1995). The neurohypophsial endocrine regulatory cascade: Precursors, mediators, receptors, and effectors. *Front. Neuroendocrinol.* **16**, 237–289.

Chini, B., and Fanelli, F. (2000). Molecular basis of ligand binding and receptor activation in the oxytocin and vasopressin receptor family. *Exp. Physiol.* **85S**, 59S–66S.

Hedges, B. S. (2002). The origin and evolution of model organisms. *Nature Rev. Genet.* **3**, 838–849.

Ivell, R., and Russell, J. A. (1995). *Oxytocin: Cellular and Molecular Approaches in Medicine and Research*. Kluwer Academic/Plenum, New York.

Zingg, H. H., Bourque, C. W., and Bichet, D. G. (1998). *Vasopressin and Oxytocin: Molecular Cellular, and Clinical Advances*. Kluwer Academic/Plenum, New York.

BIOGRAPHY

Michael J. Brownstein is Chief of the Laboratory of Genetics, National Institute of Mental Health, NIH, Bethesda, MD. He received his Ph.D. and M.D. degrees from the University of Chicago and did postdoctoral training with Julius Axelrod. His research has been in the areas of neurobiology, endocrinology, genetics, and genomics. He and his co-workers are known for their studies of vasopressin and oxytocin biosynthesis, and for cloning the vasopressin and oxytocin receptors.

V-ATPases

Michael Forgac
Tufts University School of Medicine, Boston, Massachusetts, USA

The V-ATPases (or vacuolar (H$^+$)-ATPases) are ATP-driven proton pumps whose primary function is to acidify intracellular compartments in eukaryotic cells, although they have also been identified in the plasma membrane of certain cells. V-ATPases have been shown to play a crucial role in a variety of normal cellular processes as well as a number of human diseases. The structure, mechanism, and regulation of these proton pumps have been the topics of intense study.

V-ATPase Function

FUNCTION OF INTRACELLULAR V-ATPASES

V-ATPases have been identified in many intracellular compartments, including endosomes, lysosomes, Golgi-derived vesicles and secretory vesicles. V-ATPases within endosomal compartments are important for the process of receptor-mediated endocytosis (Figure 1). During receptor-mediated endocytosis, cells take up ligands (such as the cholesterol carrying complex low density lipoprotein, or LDL) from their environment by binding them to receptors on the cell surface and clustering these receptors in specialized regions of the plasma membrane which then invaginate into the cell. Following this internalization, the ligand-receptor complexes are exposed to a low pH within the endosome that causes the internalized ligand to dissociate from its receptor. This dissociation allows the receptor to recycle to the plasma membrane (where it is reutilized) and the ligand to proceed to the lysosome, where it is degraded. The low pH within the endosome is generated by the V-ATPase.

Acidification of endosomes is also important in the formation of carrier vesicles that carry the released ligands from early to late endosomal compartments, and in the delivery of newly synthesized lysosomal enzymes from the Golgi to lysosomes. The latter process involves the binding of these enzymes to mannose-6-phosphate receptors in the trans-Golgi followed by their delivery to an endosomal compartment. Within this compartment, the low pH created by the V-ATPases causes dissociation of the lysosomal enzymes from their receptors, allowing delivery of the enzymes to the lysosome and recycling of the receptors to the trans-Golgi. Finally, endosomal acidification is involved in the entry of certain envelope viruses (such as influenza virus) into cells. These viruses bind to the surface of cells and are internalized by the process of endocytosis. Upon exposure to a low pH, the virus coat fuses with the endosomal membrane, releasing the viral nucleic acid into the cytoplasm of the host cell. Endosomal acidification is therefore essential in the process by which these viruses infect cells.

Lysosomes are the major compartment in which degradation of proteins and other macromolecules occurs in cells. The lysosomal enzymes responsible for this degradation all require an acidic environment to be active. This acidic environment is created by the V-ATPases. Secretory vesicles, such as synaptic vesicles, are also acidic compartments. Synaptic vesicles are located at the synaptic terminal of nerve cells and release neurotransmitters (that chemically trigger the next nerve cell) by fusion with the plasma membrane. Neurotransmitters become concentrated within synaptic vesicles by transport proteins within the synaptic vesicle membrane that utilize either the proton gradient or the membrane potential generated by the V-ATPases to drive uptake of the transmitter.

FUNCTION OF PLASMA MEMBRANE V-ATPASES

Plasma membrane V-ATPases play an important role in a number of normal and disease processes. In alpha-intercalated cells in the kidney, V-ATPases are located in the apical membrane where they pump protons into the urine, thus helping to control the pH of the blood. A genetic defect in this pump leads to a disease called renal tubule acidosis, in which the kidney is unable to secrete sufficient acid. V-ATPases are also present in the plasma membrane of osteoclasts, which are cells that function in degradation of bone. These cells are essential during development to facilitate bone remodeling. Plasma membrane V-ATPases in osteoclasts create an acidic extracellular environment that is necessary for bone degradation to occur. A genetic defect in the V-ATPase in osteoclasts leads to the human disease autosomal
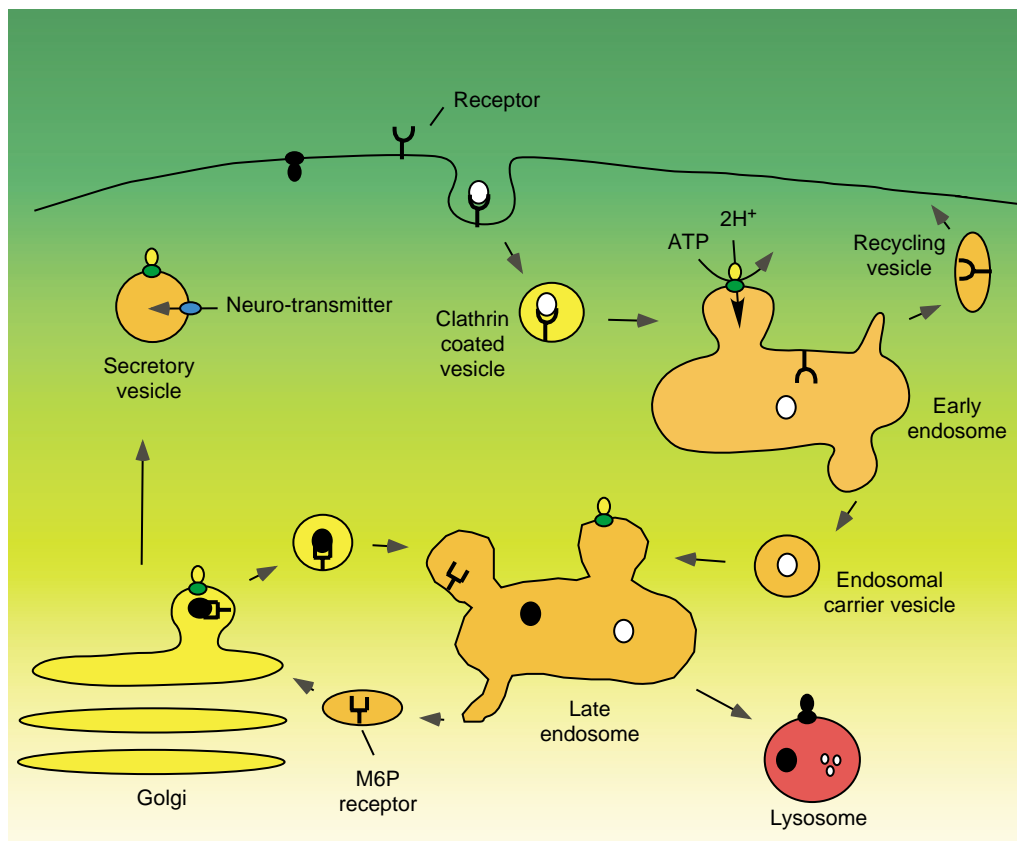
FIGURE 1 Function of intracellular V-ATPases. Acidification of early endosomes by the V-ATPase facilitates receptor recycling following endocytic uptake and formation of endosomal carrier vesicles. Recycling M6P receptors to the trans-Golgi is also low pH dependent. V-ATPase activity is also required to drive neurotransmitter uptake and facilitates protein degradation in lysosomes.

recessive osteopetrosis, in which the inability to degrade bone leads to severe skeletal defects and death.

Plasma membrane V-ATPases in macrophages and neutrophils have been shown to help maintain a neutral internal pH under conditions of severe acid load. In the vas deferens, V-ATPases create a low pH environment necessary for sperm development. V-ATPases in the plasma membrane of tumor cells have also been proposed to function in tumor invasion by providing an acidic extracellular environment necessary for secreted lysosomal enzymes to degrade extracellular matrix. Finally, V-ATPases in intestinal cells in insects create a membrane potential across the apical membrane that is used to drive potassium transport into the gut.

V-ATPase Structure

The V-ATPase is a large complex composed of 13 different subunits. These subunits are arranged into two separate domains termed V_1 and V_0 (Figure 2). The V_1 domain is made up entirely of subunits that are peripheral to the membrane (i.e., not membrane-embedded). This domain has a molecular mass of ~ 640 kDa and contains eight different subunits (subunits A–H) of molecular mass 70–13 kDa (Table I). The V_1 domain is responsible for hydolysis of ATP, which occurs on catalytic sites located on the three copies of subunit A. There are therefore three catalytic nucleotide binding sites per complex. The B subunits (which are also present in three copies per complex) can also bind nucleotides, but these sites are referred to as "non-catalytic" sites, since they do not actually hydrolyze ATP. The function of these sites is not known, but they may play a role in controlling the activity of the V-ATPase. The A and B subunits are arranged in a hexamer, like the segments of an orange, with alternating A and B subunits. ATP is hydrolyzed sequentially at each of the three catalytic sites. The other subunits in the V_1 domain (subunits C–H) function to connect the V_1 domain to the V_0 domain and are discussed here.

The V_0 domain is composed of five different subunits (subunits a, d, c, c', and c'') of molecular mass 100–17 kDa. All of the subunits in the V_0 domain except subunit d are embedded in the membrane, and thus require detergent for solubilization. The V_0 complex has a molecular mass of 260 kDa and is responsible for transport of protons across the membrane.
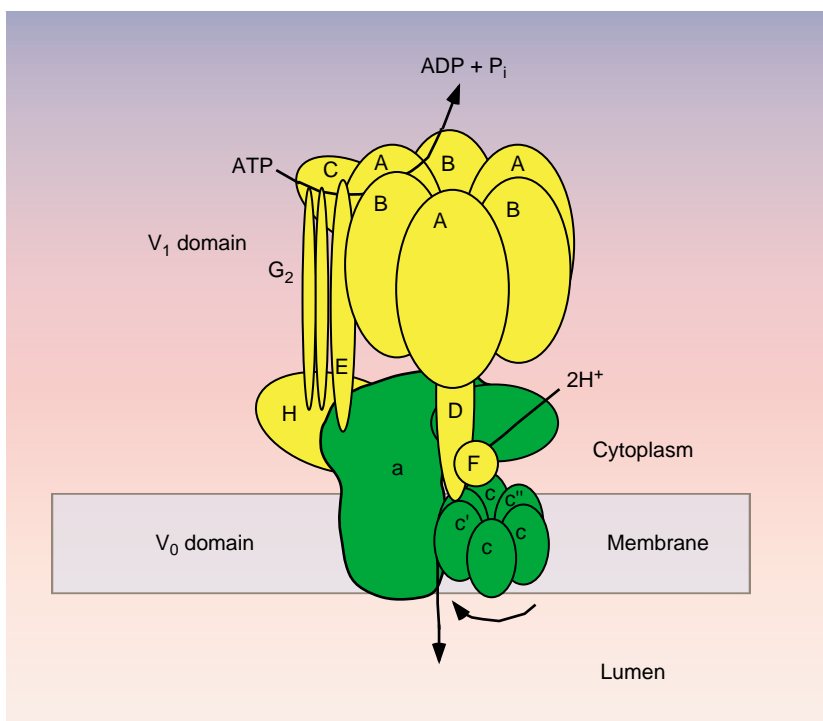
FIGURE 2   Structure of the V-ATPases. The V-ATPases contain two domains. The V_1 domain is responsible for ATP hydrolysis and the V_0 domain carries out proton transport across the membrane. Like the F-ATPases, the V-ATPases operate by a rotary mechanism in which ATP hydrolysis in V_1 drives rotation of a central stalk which is connected to a ring of proteolipid subunits in V_0. It is movement of the proteolipid ring relative to subunit a that drives proton transport (see text).

This proton transport only occurs when the V_1 domain is attached to V_0 and is driven by the hydrolysis of ATP in V_1. Three of the subunits in the V_0 domain are called proteolipid subunits (c, c′, and c″) because they are so

TABLE I

Subunit Composition of the V-ATPase

Domain	Subunit	Gene (yeast)	M_r (kDa)	Function/ location
V_1	A	VMA1	69	Catalytic ATP binding site
	B	VMA2	58	Noncatalytic ATP binding site
	C	VMA5	44	Peripheral stalk
	D	VMA8	29	Central stalk
	E	VMA4	26	Peripheral stalk
	F	VMA7	14	Central stalk
	G	VMA10	13	Peripheral stalk
	H	VMA13	54	Peripheral stalk
V_0	a	VPH1/STV1	100	Proton translocation, targeting
	d	VMA6	40	Cytoplasmic side
	c	VMA3	17	Proton translocation, bafilomycin-binding site
	c′	VMA11	17	Proton translocation
	c″	VMA16	23	Proton translocation

hydrophobic that they can be extracted from the membrane using organic solvent mixtures, such as chloroform:methanol. These subunits are almost completely embedded in the membrane and the polypeptide chain of each one crosses the membrane four times (these are called transmembrane segments). Buried in the middle of one of the transmembrane segments of each of the proteolipid subunits is a single essential glutamic acid residue which is reversibly protonated and deprotonated during proton transport by the V-ATPases. Like the A and B subunits in the V_1 domain, the proteolipid subunits in the V_0 domain form a ring, with four copies of subunit c and one copy each of subunit c′ and c″. The specific V-ATPase inhibitor bafilomycin has been shown to bind to the proteolipid subunits of the V_0 domain.

In addition to the proteolipid subunits, the other V_0 subunit that is embedded in the membrane is the 100-kDa subunit a. This subunit is made up of two domains. The carboxyl-terminal half of the molecule contains nine transmembrane segments while the amino-terminal half is a hydrophilic domain that is present on the cytoplasmic side of the membrane. Like the proteolipid subunits, the subunit a also contains amino acid residues that are essential for proton transport. In particular, there is a positively charged arginine residue near the middle of the seventh transmembrane segment of subunit a that is absolutely required for proton transport

by the V-ATPases. The function of the remaining V_0 subunit d, which is tightly bound to the V_0 domain but is not embedded in the membrane, is not known.

The V_1 and V_0 domains are connected by two stalks. The central stalk is composed of the subunits D and F whereas the peripheral stalk is composed of subunits C, E, G, H, and the soluble domain of subunit a. The function of these stalks is described below.

Mechanism of ATP-Driven Proton Transport by the V-ATPases

The V-ATPases are believed to operate by a rotary mechanism, similar to that demonstrated for the F-ATPases (or ATP synthases), which are enzyme complexes present in mitochondria, chloroplasts, and bacteria that function in the reverse direction (that is in proton-driven ATP synthesis). For the V-ATPases, ATP hydrolysis in the V_1 domain drives rotation of the central stalk (containing subunits D and F), which in turn drives rotation of the ring of proteolipid subunits relative to subunit a in the V_0 domain. Subunit a is held fixed relative to the A_3B_3 hexamer of V_1 by the peripheral stalk (or stator), composed of subunits C, E, G, H, and the soluble domain of subunit a. It is rotation of the ring of proteolipid subunits relative to subunit a that drives active transport of protons from the cytoplasmic to the lumenal side of the membrane. A proton enters from the cytoplasmic side of the membrane via a cytoplasmic access channel in subunit a and protonates a buried carboxyl group on one of the proteolipid subunits. ATP hydrolysis in V_1 forces rotation of the proteolipid ring in the plane of the membrane such that the protonated carboxyl group on the proteolipid subunit reaches a second access channel in subunit a that leads to the lumenal side of the membrane. Interaction between this carboxyl group on the proteolipid subunit and the buried arginine residue of subunit a (which is positively charged) forces the proton off of the proteolipid subunit into the lumenal access channel, where it can be released to the lumenal side of the membrane, thus completing the transport cycle. In this way, the rotary motion driven by hydrolysis of ATP is converted into unidirectional transport of protons across the membrane.

Regulation of V-ATPase Activity *In Vivo*

The activity of V-ATPases in different membranes in the cell is known to be regulated such that the pH of different intracellular compartments and the degree of proton transport across the plasma membrane is carefully controlled, but the mechanisms employed in regulating V-ATPase activity in cells are still being elucidated. One important mechanism of regulation involves reversible dissociation of the V-ATPase complex into its component V_1 and V_0 domains. In yeast, dissociation occurs in response to removal of glucose from the media, probably as a way to preserve cellular energy stores. Dissociation has also been demonstrated to occur in insects and in mammalian cells. A second proposed regulatory mechanism involves the formation of a disulfide bond between two conserved cysteine residues located at the catalytic site on the subunit A.

Differential targeting of V-ATPases to different cellular membranes has also been proposed as a means of controlling proton transport. This has been shown to occur in intercalated cells in the kidney, where exposure to a low pH causes the fusion of intracellular vesicles containing the V-ATPase with the plasma membrane, thus increasing proton transport out of the cell into the renal fluid. Differential targeting of V-ATPases appears to be controlled by different isoforms of the subunit a. Thus the a3 isoform is able to target the V-ATPase to the plasma membrane in osteoclasts whereas the a4 isoform targets the V-ATPase to the intercalated cell plasma membrane. It is mutations in these isoforms that lead to the human diseases osteopetrosis and renal tubule acidosis mentioned earlier. Isoforms have now been identified in many of the V-ATPase subunits in mammalian cells, and these have been shown to be expressed in both tissue- and organelle-specific manner. This has led to the expectation that specific inhibitors can be identified that are selective in their ability to inhibit particular V-ATPase complexes, which may in turn lead to cures for diseases such as osteoporosis.

SEE ALSO THE FOLLOWING ARTICLE

Lipid Rafts

GLOSSARY

access channel An aqueous channel that allows protons to reach buried carboxyl groups in the center of the membrane from one side of the membrane or the other.

osteopetrosis A genetic disease in humans associated with the inability to degrade bone, one cause of which is a defect in the V-ATPase of osteoclasts.

receptor-mediated endocytosis The process by which cells take up specific ligands from their environment (such as low density lipoprotein) via cell surface receptors.

V-ATPase Vacuolar proton translocating ATPase, which carries out active proton transport from the cytoplasmic to the non-cytoplasmic side of the membrane driven by energy released upon hydrolysis of ATP.

FURTHER READING

Arata, Y., Baleja, J. D., and Forgac, M. (2002). Localization of subunits D, E and G in the yeast V-ATPase complex using cysteine-mediated cross-linking to subunit B. *Biochemistry* **41**, 11301–11307.

Bowman, B. J., and Bowman, E. J. (2002). Mutations in subunit c of the V-ATPase confer resistance to bafilomycin and identify a conserved antibiotic binding site. *J. Biol. Chem.* **277**, 3965–3972.

Nishi, T., and Forgac, M. (2002). The vacuolar (H$^+$)-ATPases: Nature's most versatile proton pumps. *Nat. Rev. Mol. Cell Biol.* **3**, 94–103.

Smardon, A. M., Tarsio, M., and Kane, P. M. (2002). The RAVE complex is essential for stable assembly of the yeast V-ATPase. *J. Biol. Chem.* **277**, 13831–13839.

Toyomura, T., Oka, T., Yamaguchi, C., Wada, Y., and Futai, M. (2000). Three subunit a isoforms of mouse vacuolar (H$^+$)-ATPase. Preferential expression of the a3 isoform during osteoclast differentiation. *J. Biol. Chem.* **275**, 8760–8765.

BIOGRAPHY

Michael Forgac is a Professor in the Department of Physiology at Tufts University School of Medicine. His principal research interest is in the structure, mechanism, and regulation of the V-ATPases. He received his B.S. in biology and chemistry from Caltech in 1976 and his Ph.D. in Biochemistry and Molecular Biology from Harvard in 1981. He is author of over eighty publications and a recipient of a MERIT Award from the National Institutes of Health.

Vitamin A (Retinoids)

Joseph L. Napoli
University of California, Berkeley, CA, USA

Vertebrates require vitamin A, all-*trans*-retinol (atROH), for vision, fertility, embryogenesis, growth, and optimum neuro and immune function. atROH generates the visual pigment, 11-*cis*-retinal, and the humoral effector all-*trans*-retinoic acid (atRA). atROH transport and metabolism relies on serum and cellular chaperones (binding proteins) to maximize efficiency and impose specificity. Cleavage of carotenoids in the intestinal mucosa by carotenoid monooxygenase produces all-*trans*-retinal. CRBP(II) binds all-*trans*-retinal and dietary retinol, and allows only reduction and esterification into all-*trans*-retinyl esters (atRE). atRE storage occurs mostly in liver stellate cells. Liver only releases atROH bound with serum retinol binding protein, sRBP. In extra-intestinal tissues, cellular retinol binding-protein (CRBP) binds atROH and allows re-esterification. In the retinal pigment epithelium, the binding-protein RPE65 facilitates concerted atRE hydrolysis and conversion into 11-*cis*-retinol, catalyzed by an isomerohydrolase. Short-chain dehydrogenase/reductases (SDR) convert 11-*cis*-retinol into 11-*cis*-retinal. 11-*Cis*-retinal crosses the interphotoreceptor matrix and enters the rod outer segments, where it binds covalently with opsin to form rhodopsin. Light isomerizes 11-*cis*-retinal into all-*trans*-retinal, changing the conformation of rhodopsin and generating a nerve impulse. The transporter ABCR aids leaching of all-*trans*-retinal from rhodopsin. Other SDR reduce all-*trans*-retinal into atROH. atROH crosses the interphotoreceptor space, rebinds with CRBP in the retinal pigment epithelium, and undergoes re-esterification. atRA biogeneration takes a different path. atROH, from hydrolysis of atRE by retinyl ester hydrolase or from serum, binds with CRBP and undergoes dehydrogenation by microsomal SDR. The all-*trans*-retinal produced undergoes dehydrogenation into atRA, catalyzed by retinal dehydrogenases. CRABP(II) delivers atRA to retinoic acid receptors, whereas the complex CRABP-atRA has higher enzymatic efficiency for atRA degradation than does free atRA. Cytochromes P-450 initiate atRA degradation through C4 and C18 hydroxylation and 5,6-epoxydation. Primary regulators of atROH homeostasis include apo-CRBP, which inhibits esterification of atROH and accelerates hydrolysis of atRE, and atRA, which induces its own catabolism.

Retinoids and their Functions

The term *retinoid* refers to compounds, both naturally occurring and synthetic, with vitamin A activity: vitamin A denotes the organic compound all-*trans*-retinol (atROH). All vertebrates require vitamin A/atROH for vision, fertility, embryogenesis, and growth. atROH does not support the physiological functions attributed to vitamin A; rather it acts as precursor for biosynthesis of retinoids directly responsible for producing "vitamin A activity". These atROH metabolites include 11-*cis*-retinal, the cofactor covalently bound with opsin to form the visual pigment rhodopsin, and all-*trans*-retinoic acid (atRA), the humoral effector of the non-visual, systemic functions attributed to vitamin A. These systemic functions include: controlling the differentiation programs of all epithelia, and of stem cells in the skin, nervous system, bone, immune system, hemopoietic system; acting as a tumor suppressor; and regulating apoptosis. Specific non-visual effects of retinoids include memory formation, immune responsivity, stress adaptation, and cell fate determination. atRA activates three nuclear receptors, RARα, β and γ, and thereby controls the expression of hundreds of genes. atRA, or perhaps other retinoids, may also function through non-genomic mechanisms, but the research is in its infancy.

The Chemistry of atROH Generation, Storage, and Metabolic Activation

Carotenoids with at least one β-ionone ring, especially β-carotene, provide the major retinoid precursors in most diets. Oxidative central cleavage by a plasma-membrane associated, but soluble, carotene 15,15′-monooxygenase (CMO) produces all-*trans*-retinal—a transient intermediate that occurs in very low concentrations outside of the neural retina. Microsomal retinal reductases, RRD, members of the short-chain

dehydrogenase/reductase (SDR) gene family, catalyze reduction of all-*trans*-retinal into atROH, which then undergoes esterification into all-*trans*-retinyl esters (atRE), predominantly palmitate, by microsomal lecithin:retinol acyltransferase (LRAT). Chylomicrons carry atRE into circulation and chylomicron remnants deliver them to liver, where ultimately, most are stored in stellate cells.

In the visual cycle, atRE undergo concerted hydrolysis–isomerization into 11-*cis*-retinol by a microsomal isomerohydrolase (IMH), followed by dehydrogenation into 11-*cis*-retinal (SDR). 11-*Cis*-retinal forms a Schiff's base with a lysine residue in the protein opsin to form rhodopsin. When light strikes the neural retinal, 11-*cis*-retinal in rhodopsin undergoes *cis* to *trans*-isomerization, causing a conformation change that initiates nerve impulses and releases the newly formed all-*trans*-retinal. Reduction of all-*trans*-retinal into atROH (microsomal SDR) and re-esterification (LRAT) into atRE completes the visual cycle.

Activation of atROH into atRA uses the same intermediate used in the visual cycle, all-*trans*-retinal, but relies on a metabolically distinct route. atROH, either from blood or from hydrolysis of retinyl esters by microsomal retinyl ester hydrolase (REH), undergoes reversible dehydrogenation into all-*trans*-retinal, catalyzed primarily by microsomal retinol dehydrogenases, RDH, also members of the SDR gene family. In contrast to the comparatively high concentrations of retinals that allow the visual cycle to function, all-*trans*-retinal concentrations during atRA biosynthesis are kept low by reduction (back reaction of RDH and reduction by microsomal and peroxisomal reductases), and by irreversible dehydrogenation into atRA, catalyzed by soluble, ~ 54 kDa, high V_{max} retinal dehydrogenases (RALDH), members of the aldehyde dehydrogenase (ALDH) gene family. atRA isomers occur *in vivo*, such as 13-*cis*-RA and 9,13-di-*cis*-RA, but their significance and source(s) remain unclear.

These straightforward reactions offer complex opportunities for physiological regulation, owing to compartmentalization, distinct enzymes catalyzing each direction of chemically reversible reactions (*e.g.* dehydrogenation/reduction of retinol/retinal; esterification/hydrolysis of atROH/atRE), cell-distinct expression patterns, and multiple homologs/paralogs that catalyze several reactions. For example, at least four reductases have been identified, which belong to the SDR gene family, Rrd (peroxisomal) and RalR1 outside of the eye, and retSDR and prRDH, in the neural retina. At least three RDH, also SDR, have been identified in the rat: Rodh1, Rodh2 and Rodh3. Four Raldh have been identified in human, rat and mouse: Raldh1, 2, 3, 4 (aka ALDH 1A1, 1A2, 1A6 and 8A1). These constitute a complex enzyme system for absorbing and storing

vitamin A, maintaining atRA homeostasis, and recycling vitamin A in the visual cycle.

Retinoid Binding-Proteins and their Contributions to Retinoid Homeostasis

Processing of dietary retinoids and retinoid precursors and biogeneration of active retinoids relies on serum and cellular chaperones for efficient and specific retinoid use, as demonstrated by studies *in vitro* and the consequences of gene knockouts and/or naturally occurring mutations.

Liver and other tissues synthesize a retinol binding-protein, sRBP, a member of the lipocalin gene family. atROH egress from liver requires sRBP, and sRBP-atROH represents the major form of vitamin A in serum. sRBP circulates as a complex with a tetramer of transthyretin, which protects the ~ 20 kDa sRBP from kidney filtration. The mechanism of atROH delivery from sRBP into cells has not been established. Some data suggest a membrane receptor, other data indicate that cellular retinol binding-protein(s) draw atROH from sRBP through the membrane. A third hypothesis would have a sRBP receptor mainly in eye, the major site of vitamin A consumption. The sRBP null mouse seems phenotypically normal, except for impaired vision after weaning. Feeding a vitamin A-adequate diet for months restores vision. Although the eye relies on sRBP for efficient atROH uptake, atROH obtained from post-prandial lipoprotein delivery can substitute, at least under laboratory conditions. Interestingly, atRA serum levels increase in the sRBP null mouse, indicating that serum delivery of atRA to tissues helps compensate for impaired atROH delivery.

Binding-proteins channel retinoid intermediates through the series of reactions that constitute the visual cycle (Figure 1). Mice null in the retinal pigment protein (RPE) RPE65, for example, cannot produce 11-*cis*-retinoids, consistent with an RPE65 function of binding the hydrophobic atRE (K_d value ~ 20 pM), and accelerating their mobilization and delivery to the next step, acyl hydrolysis and C11 isomerization by the IMH. RPE65 belongs to the same gene family as the carotenoid-metabolizing enzyme CMO (the mouse proteins have only 37% amino acid identity, however), suggesting a gene family devoted to transport/metabolism of hydrophobic substances. The IMH product, 11-*cis*-retinol, undergoes sequestration in the RPE by the 36 kDa cytosolic cellular retinal binding-protein (CRALBP), a member of the gene family that includes the α-tocopherol transfer protein, TTP. CRALBP facilitates dehydrogenation of 11-*cis*-retinol into 11-*cis*-retinal, and drives forward the *trans* to *cis* isomerization.
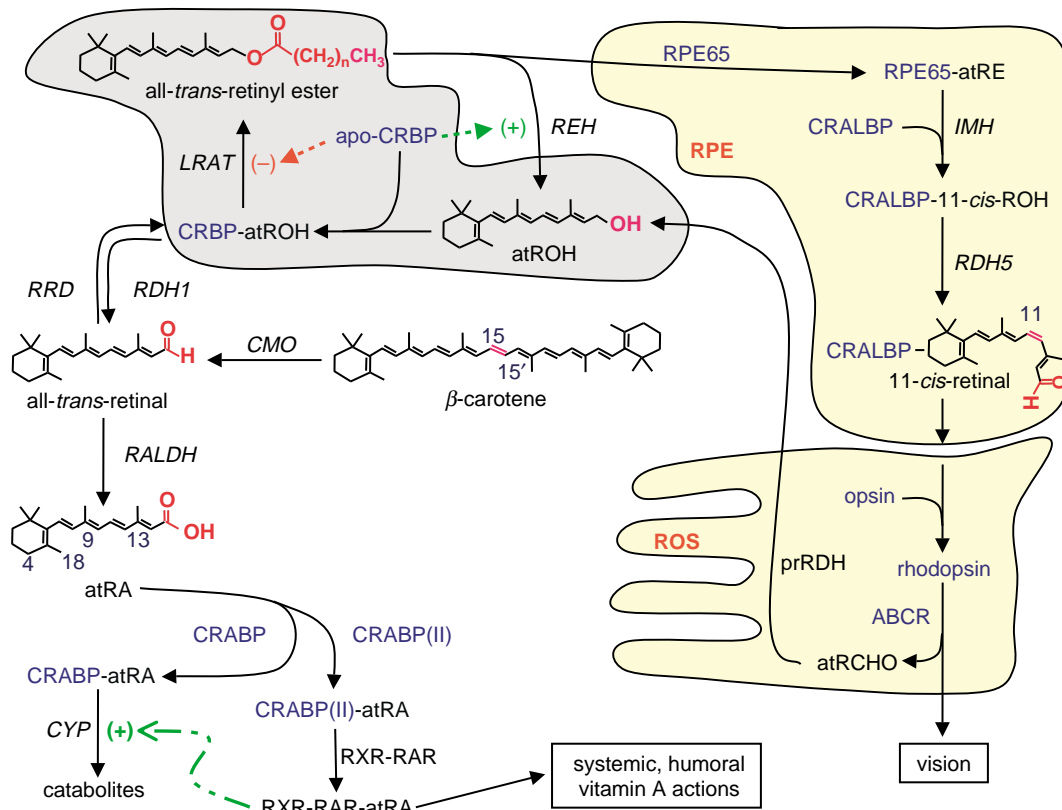
FIGURE 1   Major paths of retinoid metabolism. The white background designates reactions of retinol homeostasis and activation into atRA in the liver and extra-hepatic vitamin A target-tissues. The light gray background depicts reactions that occur both in extra-ocular tissues and in the retinal pigment epithelium (RPE) of the neural retina. The light yellow background designates reactions that occur in the RPE and in the rod outer segments (ROS). Numbers indicate the positions of *trans* to *cis* isomerization (C11), the positions of hydroxylations that occur during catabolism of atRA (C4 and C18), and the positions of atRA isomers of physiological interest (C9, C13). The dotted lines indicate the direct actions of apo-CRBP in regulating atRE hydrolysis and atROH esterification. The dashed-dotted line indicates transcriptional induction of cytochromes P-450 (CYP) by atRA. For simplicity, CRBP(II) is not shown.

Mutations in human CRALBP cause night blindness and photoreceptor degeneration. RDH4 and RDH5, respectively, represent the murine and human 11-*cis*-retinol dehydrogenase, which interacts with and/or accepts substrate from CRALBP. Mutations in RDH5 are associated with the rare, autosomal recessive disease fundus albipunctatus, i.e. night blindness from delayed photopigment regeneration. Lack of total blindness in the case of the RDH5 mutation, indicates that SDR in addition to RDH4/5 contribute to 11-*cis*-retinal biosynthesis. 11-*Cis*-retinal tranverses the interphotoreceptor matrix by an unknown mechanism and covalently binds with opsin to form rhodopsin in the rod outer segment (ROS). After the action of light, an ATP-dependent transporter, ABCR, facilitates release of all-*trans*-retinal from rhodopsin. prRDH then reduces the all-*trans*-retinal into atROH, which then travels back through the interphotoreceptor space to the RPE.

Several additional cellular retinoid binding-proteins occur. The four understood to be the most important, other than those mentioned above with respect to the visual cycle, include CRBP, CRBP(II), CRABP, and CRABP(II) (Table I). All vertebrates express all four, with well-conserved amino acid sequences among orthologs. All four are high-affinity, soluble, and specific for their ligands. CRBP binds atROH, and closely related compounds such as 3,4-didehydro-atROH, and all-*trans*-retinal to a lesser extent, and discriminate against *cis*-retinols and atRA. CRABP binds atRA, metabolites such as 4-OH-atRA, and discriminates against *cis*-RAs and atROH. These ~15 kDa globular proteins belong to the intracellular lipid binding-protein (iLBP) gene family, which includes the various fatty acid binding-proteins.

The cellular retinol/retinal binding-proteins enclose atROH and all-*trans*-retinal with the hydroxyl/aldehyde function, crucial to metabolic activation, sheltered inside. These two proteins apparently confer selective advantage to vertebrates by enhancing efficiency of sequestering, transporting and storing vitamin A, and limiting its catabolism. Vitamin A absorption and biosynthesis in the intestine relies on CRBP(II). CRBP(II)

TABLE I

Major Extra-ocular Retinoid Binding-Proteins

Retinoid binding-protein	Ligand(s)	K_d (nM)	Post-embryonic distribution
CRBP (cellular retinol binding-protein, type I)	atROH	$\sim$0.1	Nearly ubiquitous (low in intestine)
	All-*trans*-retinal	10–50	
CRBP(II) (cellular retinol binding-protein, type II)	atROH	10–50	Intestine, neonatal liver
	All-*trans*-retinal	10–50	
CRABP (cellular retinoic acid binding-protein, type I)	atRA	0.4	Widespread
CRABP(II) (cellular retinoic acid binding-protein, type II)	atRA	2	Limited (inducible?) (skin, uterus, ovary)

null mice pups die within 24 hours after birth, when delivered by dams fed a diet marginal in vitamin A content. CRBP(II) contributes $\sim$1% of the soluble protein to the intestinal enterocyte—an indication of a mass-action function to sequester newly synthesized (from carotenoids) all-*trans*-retinal or newly absorbed dietary atROH. *In vitro*, all-*trans*-retinal bound with CRBP(II) undergoes reduction readily, but neither it nor bound atROH undergoes dehydrogenation. This would limit production of atRA, and other metabolism, from the bolus of all-*trans*-retinal produced during carotenoid metabolism. In the intestine, the esterifying enzyme LRAT recognizes atROH bound with CRBP(II) to produce atRE for incorporation into chylomicrons. Thus, CRBP(II) likely aids atROH uptake and chaperones the products of carotenoid metabolism down the pathway to atRE to enhance efficiency of retinoid recovery from the diet.

Extra-intestinal vitamin A uptake and storage relies on CRBP. CRBP-null mice seem morphologically normal, but eliminate atRE 6-fold faster than wild-type mice, and may sequester/esterify atROH less efficiently. Clearly, *efficient* use of atROH *in vivo* depends on the chaperone. Strikingly, CRBP has a K_d value for atROH far lower than the atROH concentration in tissues (1–30 μM), and CRBP concentrations exceed atROH concentrations, where measured. The law of mass action predicts from these data that non-esterified atROH would occur nearly exclusively in the CRBP-bound state, assuming no alternative high-affinity or very high-capacity acceptors. Alternatives to CRBP exert limited influence *in vivo*, because isolation of CRBP from animal tissues by traditional (time-consuming and containing membranes, lipids and lipid vesicles) biochemical techniques (tissue homogenization, centrifugation and several types of column chromatography) produces largely holo-protein. Evidently, the capacity of membranes and other potential acceptors to sequester atROH does not overwhelm the ability of CRBP to sequester atROH. Like CRBP(II), LRAT can access atROH bound with CRBP to produce atRE. Ultimately, liver stellate cells accumulate most of the atRE.

CRBP seems necessary for retinoid transfer from hepatocytes to stellate cells, because the CRBP null mouse does not accumulate atRE in stellate cells.

atROH sequestering within CRBP, the need to store retinol as esters, and the need for atRA biosynthesis during specific times at specific loci, suggest that transfer of atROH for metabolism might depend on relationships between metabolizing enzymes and CRBP. Like CRBP(II), CRBP allows esterification of bound atROH, but in contrast to CRBP(II), CRBP also allows dehydrogenation of atROH. The CRBP–atROH complex shows Michaelis–Menton relationships with atRE formation by LRAT and atROH dehydrogenation by RDH. This relationship with the microsomal RDH is maintained even with changes in the CRBP/atROH ratio (provided the CRBP concentration exceeds the atROH concentration). This eliminates a "free diffusion" mechanism of transfer from the complex to select enzymes. Specific crosslinking of holo-CRBP with both RDH and LRAT confirms close proximity of CRBP and these two enzymes. Additionally, a single mutation in an exterior residue of CRBP (L35A) reduces the Vm of atROH dehydrogenation by microsomes, but does not alter the K_m, or the K_d of atROH binding to CRBP, consistent with conservation of exterior residues that aid transfer of atROH from CRBP to enzymes. Obviously, atRE and atRA biosynthesis *in vivo* occurs in the absence of CRBP, as indicated by the lack of morphological pathology in the CRBP null mouse, and their ability to sequester esters. This was predicted by the experiments *in vitro*, which showed that neither RDH nor LRAT require presentation of atROH by CRBP. Not surprisingly, the enzymes' active sites recognize their substrates in the absence of CRBP. CRBP operates as a chaperone, which restricts atROH metabolism to select enzymes, and seems required only for *efficient* atROH use *in vivo*. The ability of LRAT and RDH to access retinol from CRBP addresses the issue of how atROH would undergo efficient metabolism in the face of limited diffusion from the binding protein.

RALDH catalyze the irreversible conversion of atRCHO into atRA in the presence of CRBP, and also

can use atRCHO generated *in situ* from CRBP-atROH and RDH, or in cells presented with atROH and transfected with RDH and RALDH. In the rat, RALDH1 and RALDH2 have differing but overlaping expression patterns, and respond differently to changes in atROH status. This suggests a purpose for more than one—precise control over atROH use and atRA generation. The RALDH1 null mouse remains fertile and healthy, but may have decreased ability to produce atRA in the liver. The RALDH2 null mouse dies *in utero* at midgestation, demonstrating its unique contribution to atRA biosynthesis during embryogenesis. The situation may differ in the adult, as testes express RALDH2 strongly, but RALDH1 prevails outside of the testis. The RALDH3 null mouse dies during suckling from an obstruction in the nose. Apparently, RALDH can compensate for each other after critical developmental milestones.

A CRBP(III) has been detected in mouse heart and skeletal muscle, which express little or no CRBP or CRBP(II), but has not been detected in other retinoid target tissues, such as liver, kidney, brain, etc. CRBP(III) seems to bind about equally well with atROH, 9-*cis*-retinol and 13-*cis*-retinol, but with much lower affinity that either CRBP or CRBP(II) ($K_d \sim 80-110$ nM). Humans express yet another CRBP, originally referred to as CRBP(III), but distinct from mouse CRBP(III), and therefore CRBP(IV). CRBP(IV) mRNA is much more abundant in human liver and intestine than CRBP mRNA, but the mouse does not encode a complete CRBP(IV) gene. CRBP(IV) binds atROH with a K_d value of ~ 60 nM, and does not bind *cis*-isomers. The precise functions of CRBP(III) and CRBP(IV) have not been clarified, and unlike CRBP and CRABP, their endogenous ligands have not been established.

The atRA binding-proteins, CRABP and CRABP(II), do not have well-defined functions. Doubly null mice have a ~ 4-fold higher rate of death from unknown causes by 6 weeks after birth than wild-type, but the survivors appear normal, with one exception. The doubly null mouse, as well as the CRABP(II)-only null mouse, show 83% and 45%, respectively, incidence of a small outgrowth anomaly on the post-axial side of digit five, predominantly in the forelimbs. Mice doubly null in CRABP and CRABP(II) do not exhibit enhanced sensitivity to exogenous atRA, suggesting that the binding-proteins do not protect against atRA toxicity. In contrast to CRBP, both CRABP and CRABP(II) allow projection of the β-ionone ring of their ligand. Significantly, the first reactions of atRA degradation occur at these comparatively accessible sites, i.e. C4 and C18. Presenting atRA to microsomes bound with CRABP enhances kinetic efficiency (K_{cat}/K_m) of catabolism 7-fold. There seems to be little doubt that CRABP sequesters atRA: delivering the sequestered atRA for efficient catabolism seems a logical

mechanism to discharge the ligand without releasing it back into the cell. Unfortunately, this insight doesn't reveal the primary purpose for CRABP impounding atRA in the first place, although CRABP tends to express in cells different from those that express CRBP and CRABP(II). CRABP(II), but not CRABP, seems to deliver atRA to RAR, via a transfer that does not require free diffusion. This would complete the chaperoning of atROH on its journey from atRE through atRA biogenesis to nuclear localization.

Many other enzymes, including medium-change alcohol dehydrogenases, and aldo-keto reductases, metabolize retinoids *in vitro*, which seems to confirm the need for evolution of CRBP to protect the sparse and valuable vitamin A from clearance as a "xenobiotic". These enzymes do not access atROH bound with CRBP. Neither the ADHI-null deermouse, a natural mutant, nor mice made null in ADHI, ADHIV, or ADHIII or doubly null in ADHI and ADHIV, present with vitamin A deficiency symptoms, revealing no inability to activate atROH. ADHI-null mice do show decreased ability to convert a very large dose (50 mg kg^{-1}) of atROH into atRA. Such a dose has no natural equivalent: the results indicate only that determination can overwhelm physiological chaperones.

Other Naturally Occurring Retinoids

Discrete loci, such as skin and the chick limb bud, synthesize 3,4-didehydro-atRA. 3,4-Didehydro-atRA binds retinoic acid receptors with affinity similar to that of atRA. The purpose has not been clarified of creating a signaling molecule that functions as atRA in specialized loci that also biosynthesize atRA. Although 9-*cis*-RA was reported as an activated retinoid, it is virtually undetectable *in vivo*: its putative function as a physiological ligand that controls RXR remains uncertain.

Control of Vitamin A Homeostasis

The primary regulators of retinoid homeostasis appear to be apo-CRBP and atRA. apo-CRBP inhibits LRAT and stimulates retinyl ester hydrolysis. Thus, the direction of flux into or out of atRE would reflect the ratio apo-CRBP/holo-CRBP, which would reflect the atROH status of the cell. atRA may serve as a signal to liver to release atROH into the serum. Little has been revealed about humoral regulation of atRA biosynthesis, but estrogen and PGE, increase and decrease, respectively, atRA biosynthesis in cultured cells. The metabolism of atRA limits its activity; conversely, inhibitors of atRA metabolism enhance atRA potency. atRA

induces its own oxidative metabolism into 5,6-epoxy-atRA, 18-hydroxy-atRA, and 4-hydroxy-atRA, through inducing cytochrome P-450 s (CYP). CYP26A1 and CYP2C39 appear to have major functions in the catalysis of atRA catabolism. CYP26A1 null mice die in mid to late gestation with serious morphological defects. Two other CYP, 26B1 and 26C1, also catabolize atRA, but their precise function has not been clarified. Many other CYP reportedly catabolize atRA (e.g. CYP1A1/2, CYP2A6, CYP2C8/9, CYP2E1, CYPP3A4/5), but atRA does not induce these isoforms and most have inefficient kinetics with atRA *in vitro*.

PPARβ induces transcription of CMO, CRBP, CRBP(II), LRAT, and RARβ, suggesting correlation between vitamin A homeostasis and general lipid metabolism, whereas LRAT and CRBP are among the numerous genes induced by atRA. This action of atRA may represent a "housekeeping" function, rather than acute control over retinoid homeostasis.

Several xenobiotics, including ethanol and poly-chlorinated aromatic hydrocarbons reduce atRE stores, possibly through enhancing atRA catabolism by inducing CYP: the polychlorinated hydrocarbons, such as dioxin, function through the AH receptor to decrease atRE stores.

Clinical Uses of Retinoids

Numerous studies have correlated vitamin A insufficiency in laboratory animals with increased incidence of spontaneous and carcinogen-induced cancer. Chemo-preventive trials in humans show some promise for retinoids in actinic keratoses, oral premalignant lesions, laryngeal leukoplakia, and cervical dysplasia. The FDA has approved retinoids for acute promyelocytic leukemia and in non-life-threatening diseases such as cystic acne and psoriasis. Retinoids also provide the active ingredients in agents to treat sun-/age-damaged skin.

The WHO recognizes vitamin A-deficiency as a mortality factor for childhood measles. Two large doses (60,000 REQ each) of a water-soluble vitamin A formulation given to children on each of two days decreases the risk of death from measles 81% in areas of prevalent vitamin A-deficiency.

SEE ALSO THE FOLLOWING ARTICLES

Retinoic Acid Receptors • Steroid/Thyroid Hormone Receptors

GLOSSARY

K_{cat}/K_m A measure of the efficiency of an enzyme for its substrate from dividing the rate of turnover of the substrate by the enzyme by the Michaelis constant.

K_d (equilibrium dissociation constant) A measure of the affinity of a protein for its ligand. Lower numbers indicate higher affinity.

retinoid binding-protein Proteins that bind specific retinoids with high affinity, several with K_d values <10 nM.

short-chain dehydrogenase/reductase (SDR) A gene family consisting of ~50 mammalian members in the range of 25–35 kDa that uses pyridine nucleotide cofactors to dehydrogenate or reduce steroids, retinoids, prostanoids, and intermediates in lipid metabolism.

FURTHER READING

Blomhoff, R., Green, M. H., Green, J. B., Berg, T., and Norum, K. R. (1991). Vitamin A metabolism: new perspectives on absorption, transport, and storage. *Physiological Reviews* **71**, 951–990.

Harrison, E. H. (1998). Lipases and carboxyesterases: possible roles in the hepatic metabolism of retinol. *Annual Reviews of Nutrition* **18**, 259–276.

Maden, M. (2001). Role of retinoic acid in embryonic and post-embryonic development. *Proceedings of the Nutrition Society* **59**, 65–73.

Napoli, J. L. (2000). Retinoic acid: its biosynthesis and metabolism. *Progress in Nucleic Acids Research* **63**, 139–188.

Newcomer, M. E. (1995). Retinoid-binding proteins: structural determinants important for function. *FASEB Journal* **9**, 229–239.

Saari, J. C. (2000). Biochemistry of visual pigment regeneration: the Friedenwald lecture. *Investigative Ophthalmology and Visual Science* **41**, 337–348.

Stephensen, C. B. (2001). Vitamin A, infection, and immune function. *Annual Reviews of Nutrition* **21**, 167–192.

Sun, S. Y., and Lotan, R. (2002). Retinoids and their receptors in cancer development and chemoprevention. *Critical Reviews in Oncology and Hematology* **41**, 41–55.

Wolf, G. (1984). Multiple functions of vitamin A. *Physiological Reviews* **64**, 873–937.

BIOGRAPHY

Joseph L. Napoli is a Professor in the Department of Nutrition Sciences and Toxicology and a biochemist in the experimental station at UC-Berkeley. His main research interests are in the metabolism and functions of retinoids. He received his Ph.D. from the University of Michigan-Ann Arbor in medicinal chemistry and was a post-doctoral fellow in biochemistry at the University of Wisconsin-Madison.

Vitamin B$_{12}$ and B$_{12}$-Proteins

Bernhard Kräutler
University of Innsbruck, Innsbruck, Austria

Vitamin B$_{12}$ (CNCbl), the antipernicious anemia factor, is required for human and animal metabolism and was discovered in the late 1940s. The B$_{12}$-derivatives belong to the tetrapyrrolic natural compounds and are cobalt complexes of the unique and remarkably complex corrin ligand. The B$_{12}$-coenzymes are the cofactors in important organometallic enzymatic reactions and are particularly relevant in the metabolism of anaerobes. Indeed, the microorganisms are the only natural sources of the B$_{12}$-derivatives, whereas most living organisms (except for the higher plants) depend on these cobalt corrinoids. Vitamin B$_{12}$ and its derivatives thus hold an important position in the life sciences and have attracted strong interest from medicine, biology, chemistry, and physics.

B$_{12}$: Structure and Reactivity

The red, cyanide-containing cobalt-complex vitamin B$_{12}$ (cyanocobalamin, CNCbl) is a relatively inert and physiologically rather inactive Co(III)-corrin. The biologically relevant B$_{12}$-derivatives are the light sensitive and chemically more labile organometallic coenzyme forms, coenzyme B$_{12}$ (5′-deoxy-5′-adenosylcobalamin, AdoCbl), and methylcobalamin (MeCbl, see Figure 1).

STRUCTURE OF VITAMIN B$_{12}$-DERIVATIVES

The structures of vitamin B$_{12}$ (CNCbl) and of coenzyme B$_{12}$ (AdoCbl) were established by X-ray crystallographic studies from the laboratory of D. C. Hodgkin. This work helped clarify the nature of the corrin ligand and to discover the organometallic nature of the coenzyme AdoCbl (see Figure 1). CNCbl and other cobalamins, in which the cyanide ligand is replaced by another "upper" β-ligand (see Figure 1), represent the most common of the nucleotide containing (i.e., "complete") B$_{12}$-derivatives. The crystal structures of various (organometallic) Co(III)-corrins were analyzed, including MeCbl, to study the axial bonding at the corrin-bound cobalt center and the inherent "non-planar" nature of the corrin ligand (as opposed to the porphyrin ligand), as well as possible implications of this for B$_{12}$-catalyzed

enzymatic reactions. An interesting structure is that of the oxygen-sensitive Co(II)-corrin cob(II)alamin (B$_{12r}$), the corrinoid moiety resulting from (Co−C)-bond homolysis of AdoCbl during the catalytic cycle of coenzyme B$_{12}$-dependent enzymes.

UV/vis- and circular dichroism (CD)-spectroscopy have been used to study the colored and chiral B$_{12}$-derivatives in solution. Nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry helped to identify the corrinoids from anaerobes (methanogens, sulfur metabolizing, and acetogenic bacteria) and to characterize their solution structures. The Co(II)-forms, in turn, have been investigated by electron spin resonance (ESR) spectroscopy, a technique used increasingly to analyze for paramagnetic intermediates in B$_{12}$-catalyzed enzymatic reactions.

The natural B$_{12}$-derivatives are either "complete" or "incomplete" corrinoids (which lack the nucleotide function, see Figure 1). The natural "complete" corrinoids carry different functional β-axial ligands. In addition they may vary by the constitution of the "nucleotide base," a 5,6-dimethylbenzimidazole (DMB) in the cobalamins (such as CNCbl), but an adenine in pseudovitamin B$_{12}$. The "complete" corrinoids are also unique due to the unusual α-configuration of their (pseudo-)nucleotide appendage. The specific build-up of this function enables the heterocyclic base to bind in an intramolecular fashion to the "lower" α-axial coordination site of the corrin-bound cobalt center. In this way, the nucleotide function steers the organometallic reactivity at the cobalt center of "complete" B$_{12}$-derivatives and is also relevant for recognition and tight binding by the B$_{12}$-binding proteins.

B$_{12}$-DERIVATIVES IN ELECTRON TRANSFER REACTIONS

Oxidation−reduction processes are of key importance for the chemistry and biology of B$_{12}$. Under physiological conditions B$_{12}$-derivatives may exist in three different oxidation states (Co(III), Co(II), or Co(I)), each possessing different coordination properties and reactivities: Axial coordination to the corrin-bound cobalt center depends primarily on the formal oxidation
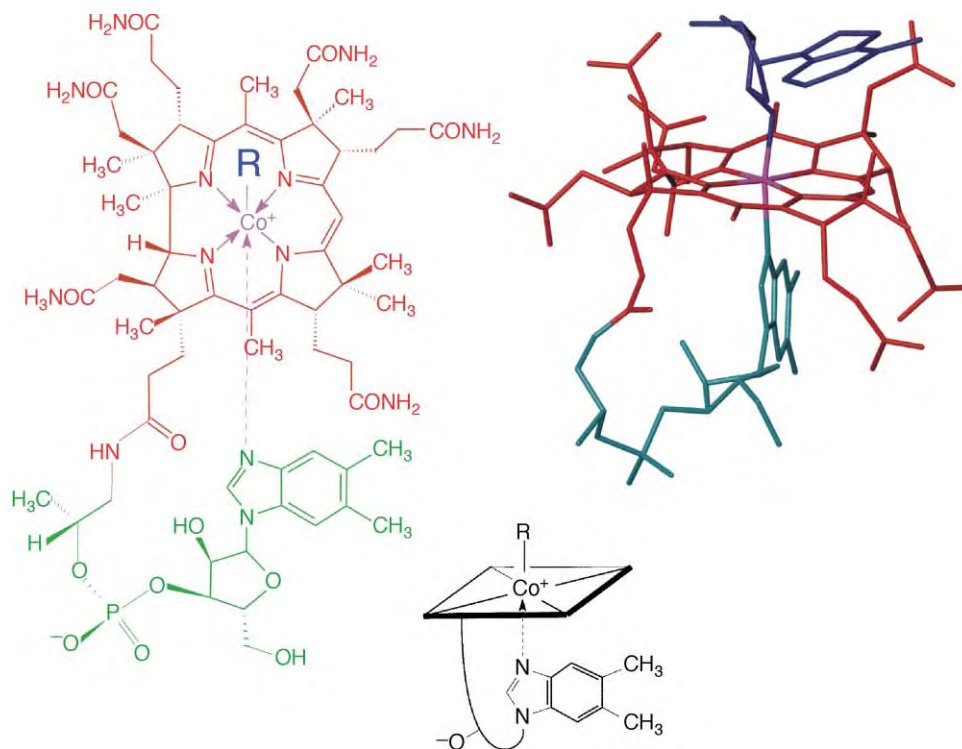
FIGURE 1   Left: structural formulas of vitamin B$_{12}$ (R = CN, CNCbl), of coenzyme B$_{12}$ (R = 5′-deoxy-5′-adenosyl, AdoCbl), methylcobalamin (R = methyl, MeCbl), cob(II)alamin (R = e$^-$, B$_{12r}$). Right/top: structure of coenzyme B$_{12}$ (AdoCbl). Bottom: symbols used for vitamin B$_{12}$ (R = CN: CNCbl) and other cobalamins.

state of the cobalt ion. As a rule, the number of axial ligands decreases in parallel with the cobalt oxidation state: In the thermodynamically predominating forms of cobalt corrins, two axial ligands are bound to the diamagnetic Co(III)-center, one axial ligand is bound to the paramagnetic Co(II)-center and axial ligands are assumed to be absent at the diamagnetic Co(I)-center. Electron transfer reactions involving B$_{12}$-derivatives, therefore, are accompanied by changes of the number of axial ligands and depend upon the nature of axial ligands. Axial coordination of the nucleotide base and of strongly coordinating ligands stabilizes the cobalt center against reduction and the reduction of alkyl-Co(III)-corrins typically occurs at potentials more negative than that of the Co(II)/Co(I)-redox-pair B$_{12r}$/B$_{12s}$.

ORGANOMETALLIC REACTIONS OF B$_{12}$-DERIVATIVES

The reactivity of B$_{12}$-derivatives in organometallic reactions holds the key to much of the biological activity of the B$_{12}$-dependent enzymes: formation and cleavage of the (Co–C)-bond in the B$_{12}$-cofactors are essential steps of the reactions catalyzed by B$_{12}$-dependent enzymes and are of particular interest.

In solution formation and cleavage of the (Co–C)-bond in organometallic B$_{12}$-derivatives were observed to

occur on all oxidation levels of the cobalt center. Two of these reaction modes were also found to be relevant for B$_{12}$-dependent enzymatic reactions:

1. the homolytic mode of formation/cleavage of the organometallic axial bond at the cobalt center (formally a one-electron reduction/oxidation of the metal center, see Figure 2), is of particular importance for the role of AdoCbl as a cofactor AdoCbl is considered a "reversible carrier of an alkyl radical" (or a reversibly functioning "radical source"). The (Co–C)-bond of AdoCbl has been determined to be ~ 30 kcal mol^{-1} strong and to be affected only slightly by the nucleotide (in the "base-on" form). The (reverse) reactions of B$_{12r}$ with alkyl radicals (such as the 5′-deoxy-5′-adenosyl radical) are remarkably fast. Indeed, the radicaloid B$_{12r}$ is a highly efficient "radical trap" and its reactions with radicals can occur with minimal restructuring of the cobalt-corrin moiety.

2. the heterolytic, nucleophile induced (SN$_2$) mode of formation/cleavage of the (Co–C)-bond at the cobalt center (formally a two-electron oxidation/reduction of the metal ion) is accompanied by formation/cleavage of a second axial bond (Figure 2): Heterolytic formation/ cleavage of the (Co–C)-bond is particularly important in enzyme-catalyzed methyl-transfer reactions (Figure 3). This mode is represented by the reaction of
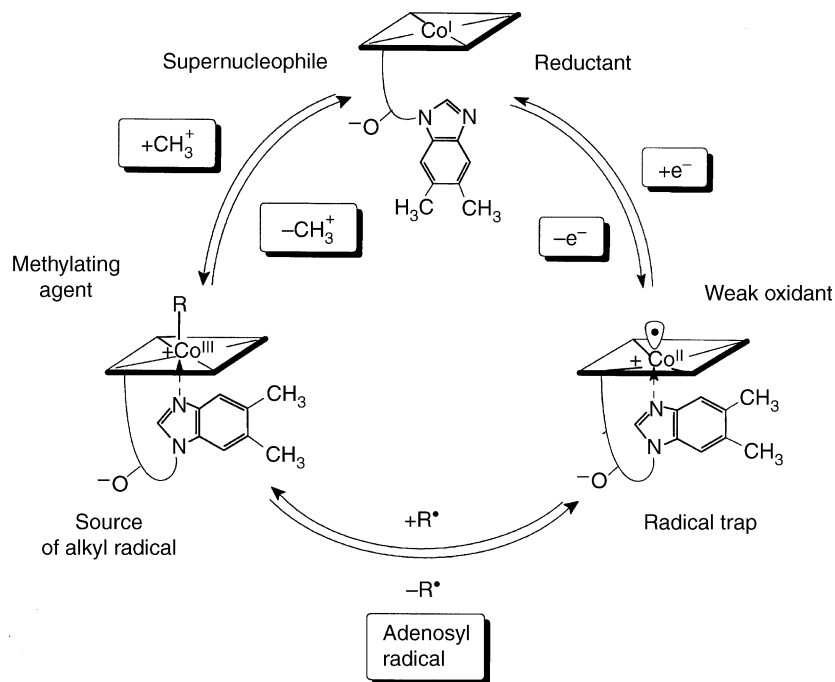
FIGURE 2 Elementary formal reaction steps of "complete" corrinoids characterizing their patterns of reactivity relevant for their cofactor function in B_{12}-dependent enzymes.

("supernucleophilic") Co(I)-corrins with alkylating agents and by the nucleophile-induced demethylation of methyl-Co(III)-corrins. Alkylation at the Co(I)-center usually occurs via "classical" bimolecular nucleophilic substitution (SN_2). The intramolecular coordination of the DMB-base in MeCbl has a notable thermodynamic effect on this type of reaction: e.g., it stabilizes "base-on" MeCbl by ~ 4 kcal mol^{-1}.
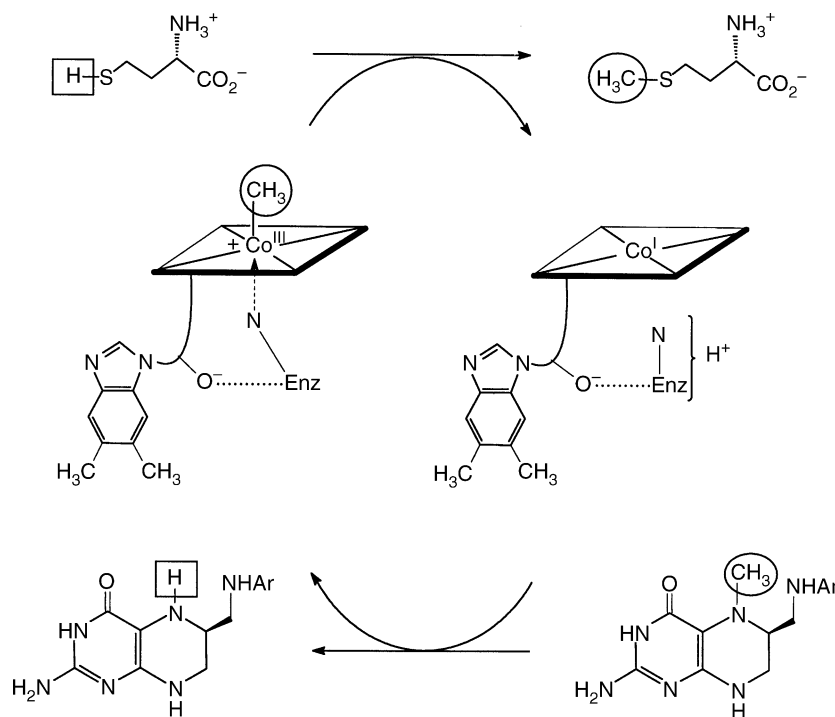
FIGURE 3 Biosynthesis of methionine by methylation of homocysteine is catalyzed by methionine synthase (MetH, Enz signifies the MetH-apoenzyme), where the bound corrinoid shuttles between MeCbl, in a "base-off/His-on" form, and cob(I)alamin.

Organometallic B$_{12}$-derivatives are rather resistant against proteolytic cleavage of the (Co−C)-bond under physiological conditions, crucial for the cofactor role of the B$_{12}$-coenzymes, but they are very sensitive to visible light, which induces the homolysis of the (Co−C)-bond.

B$_{12}$: Biosynthesis and Total Synthesis

The natural corrinoids are made exclusively by microorganisms, which are used for the industrial production of vitamin B$_{12}$ (world output ~ 10 tons per year). A "last" common intermediate in the biosynthesis of the corrins and of other porphyrinoids is the tetra-pyrrole uroporphyrinogen III. The 'corrin pathway' separates off with the sequential incorporation of 'extra' methyl groups into the reduced ligand. A major difference of the two main biosynthetic paths in aerobic and in anaerobic microorganisms concerns the timing of the cobalt insertion on the way to cobalt-containing corrins. The nucleotide portion of the "complete" corrinoids is then built-up in the later phases of the B$_{12}$-biosynthesis, beginning at the stage of cobalt-adenosylated cobyric acid.

The total synthesis of vitamin B$_{12}$ was achieved in the 1970s in the laboratories of Eschenmoser and Woodward. This conquest of the B$_{12}$-structure by abiotic synthesis led to the "incomplete" corrinoid cobyric acid first, to which the nucleotide part was attached in a specific way. The Co(II)-ion was incorporated easily into the corrin ligand, but attempts to remove cobalt from intact natural corrinoids have not been successful.

B$_{12}$-Dependent Enzymatic Reactions

B$_{12}$-Dependent Methyl Transferases

The reactivity of the "supernucleophilic" Co(I)-corrins and of methyl-Co(III)-corrins makes B$_{12}$-derivatives well-suited as cofactors in B$_{12}$-dependent enzymatic methyl group transfer reactions, which are widely relevant in many organisms. Anaerobic acetogenesis, methanogenesis and catabolism of acetic acid to methane and carbon dioxide make use of B$_{12}$-catalyzed enzymatic methyl transfer reactions. Various substrates act as sources of methyl groups, such as methanol, aromatic methyl ethers, methyl amines or N^5-methyltetrahydropterins (such as N^5-methyltetrahydrofolate). Thiols are the methyl group acceptors in methanogenesis and in methionine synthesis. In the anaerobic biosynthesis of acetyl-coenzyme A from one-carbon precursors the methyl group acceptor is suggested to be the nickel center of an Fe/Ni-cluster.

B$_{12}$-dependent methionine synthesis is one of the two essential roles of B$_{12}$ in mammalian metabolism. In methionine synthase of *Escherichia coli* (MetH) B$_{12}$-dependent methyl transfer involves a sequential mechanism, in which tetrahydrofolate and methionine are formed and homocysteine and N^5-methyltetrahydrofolate act as methyl group acceptors and donors, respectively (see Figure 3). MetH is a modular protein, where the B$_{12}$-binding domain is bound to an N^5-methyltetrahydrofolate-binding module, the homo-cysteine-binding module and a reactivating module (that binds S-adenosyl-methionine). During turnover MetH catalyzes two methyl group transfer reactions which occur with an overall retention of configuration (consistent with two inversions). In a formal sense, these methyl-transfer reactions are SN$_2$ reactions with heterolytic cleavage/formation of the (Co−CH$_3$)-bond.

X-ray crystal analysis of the B$_{12}$-binding domain of MetH provided the first insight into the structure of a B$_{12}$-binding protein. This work revealed the cobalt-coordinating DMB-nucleotide tail of the bound cofactor MeCbl to be displaced by a histidine and to be anchored in a pocket of the protein. Accordingly, in MetH the corrinoid cofactor is bound in a "base-off/His-on" constitution. The crucial cobalt-ligating histidine residue is part of a Gly-X-X-His-X-Asp-sequence, which is conserved in a group of B$_{12}$-binding proteins.

Coenzyme B$_{12}$-Dependent Enzymes

About ten coenzyme B$_{12}$-dependent enzymes are now known. These enzymes are four carbon skeleton mutases, diol dehydratase, glycerol dehydratase, ethanolamine ammonia lyase, two amino mutases, and B$_{12}$-dependent ribonucleotide reductase. Of these AdoCbl-dependent enzymes, only methylmalonyl-CoA mutase is indispensable in human metabolism.

The coenzyme B$_{12}$-dependent enzymes perform transformations, that are difficult to achieve by typical "organic" reactions. With the exception of the enzymatic ribonucleotide reduction, the results of the coenzyme B$_{12}$-catalyzed enzymatic reactions formally correspond to isomerizations with vicinal exchange of a hydrogen atom and of a group with heavy atom centers. Homolytic cleavage of the (Co−C)-bond of the protein-bound AdoCbl to a 5′-deoxy-5′-adenosyl radical and cob(II)-alamin (B$_{12r}$) is the entry to H-abstraction reactions induced by the 5′-deoxy-5′-adenosyl radical. The AdoCbl-dependent enzymes rely upon the reactivity of bound organic radicals, which are formed (directly or indirectly) by a H-atom abstraction by the 5′-deoxy-5′-adenosyl radical. The substrate radicals rearrange rapidly to the product radicals in a (pseudo-)intramolecular rearrangement and without noticeable participation of the bound B$_{12r}$ (i.e., the corrinoid is a "spectator").

The major tasks of the enzymes thus concern the enhancement of the critical radical reactions, the reversible generation of the radical intermediates and the protection of the proteins from nonspecific radical chemistry.

Coenzyme B$_{12}$-Dependent Carbon Skeleton Mutases

Methylmalonyl-CoA mutase interconverts *R*-methylmalonyl-CoA and succinyl-CoA. Binding of the substrate triggers the homolysis of the (Co–C)-bond of the bound AdoCbl. The radical carbon skeleton rearrangement reaction then proceeds, as outlined in Figure 4.

In all coenzyme B$_{12}$-dependent carbon skeleton mutases (methylmalonyl-CoA mutase (MMCM), glutamate mutase (GM), methyleneglutarate mutase, isobutyryl-CoA mutase) the B$_{12}$-binding motif (Gly-X-X-His-X-Asp) occurs. Consistent with this, the B$_{12}$-cofactor is bound "base-off/His-on" at the interface between the B$_{12}$-binding and the substrate-activating domains. The B$_{12}$-binding domain in MMCM, the B$_{12}$-binding subunit in GM and even the B$_{12}$-binding domain of MetH exhibit high sequence homology. Such a homology does not extend to the substrate-binding domains of these enzymes. The crystal structures of MMCM (from *Propionibacterium shermanii*) and GM (from *Clostridium cochlearium*) are available. The X-ray
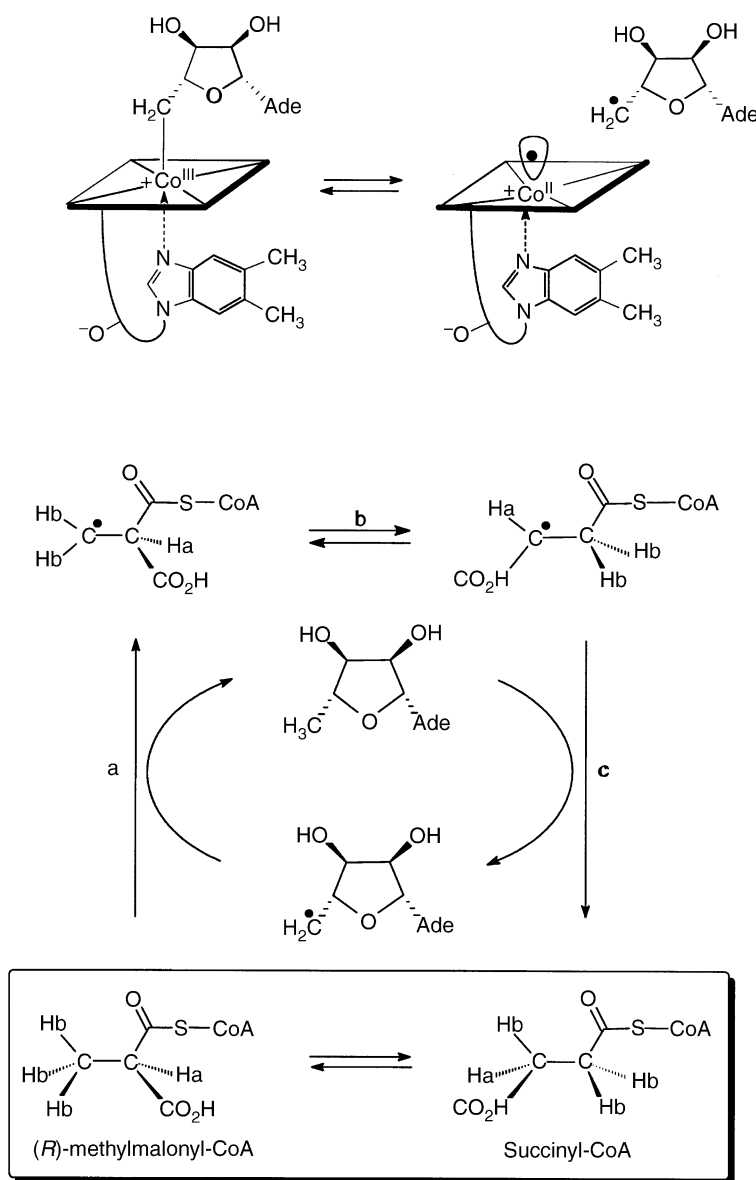
FIGURE 4 Top: coenzyme B$_{12}$ (AdoCbl) is a reversible source of the 5′-deoxy-5′-adenosyl radical and cob(II)alamin (B$_{12r}$). Bottom: methylmalonyl-CoA mutase interconverts (*R*)-methylmalonyl-CoA and succinyl-CoA. This rearrangement is proposed to involve H-atom abstraction (step a), radical rearrangement (step b) and back transfer of H-atom (step c). (Formation of the 5′-deoxy-5′-adenosyl radical from protein bound AdoCbl is substrate triggered.)

analysis of MMCM provided the first crystal structure of a coenzyme B$_{12}$-dependent enzyme. Detailed crystallographic analysis of GM with bound AdoCbl revealed a likely structural basis for the operation of the radical mechanism in AdoCbl-dependent mutases: the ribose part of the 5'-deoxyadenosyl moiety is present in two conformations, related to each other by a pseudo-rotation of the furanose ring. The solution structure of the B$_{12}$-binding subunit of GM from *Clostridium tetanomorphum* (analyzed by heteronuclear NMR) provided a first structure of a cofactor-free B$_{12}$-binding protein and indicated it to be largely preorganized for B$_{12}$-binding.

In contrast to MMCM and GM, the crystal structure of the AdoCbl-dependent diol dehydratase from *Klebsiella oxytoca* showed the corrinoid to be bound "base-on" at the interface of the B$_{12}$-binding and of the substrate-binding subunits.

B$_{12}$-Dependent Ribonucleotide Reductase

In all organisms, ribonucleotide reductases (RNRs) play essential roles in the biosynthesis of DNA by catalyzing the reduction of nucleoside di- or triphosphates to the corresponding 2'-deoxynucleotides. The RNRs use various metal cofactors to initiate the nucleotide reduction by a radical reaction. The reductase from *Lactobacillus leichmanii* (RNR-*Ll*) uses AdoCbl as cofactor and nucleoside triphosphates as substrates, while 2'-deoxynucleoside triphosphates are allosteric effectors. In RNR-*Ll* the (Gly-X-X-His-X-Asp)-"B$_{12}$-binding motif" is absent and ESR-spectroscopy and crystallography showed "base-on" binding of the B$_{12}$-cofactor. In RNR-*Ll* homolysis of the bound AdoCbl generates a protein-centered thiyl-radical, which then induces the radical reactions that lead to the reductive removal of the 2'-OH group of the ribonucleotide.

Other B$_{12}$-Dependent Enzymatic Transformations

Methanogens and acetogens dechlorinate chloro-methanes with reduced metal cofactors (corrinoids, in particular). Environmentally relevant microbiological dehalogenation reactions of chloroethenes were found recently. The anaerobe *Dehalospirillum multivorans* uses tetrachloroethene as the terminal electron acceptor, which is dechlorinated by a B$_{12}$-dependent tetrachloroethene reductive dehalogenase to trichloroethene and (then) *cis*-trichloroethene.

B$_{12}$: Medical Aspects

B$_{12}$-deficiency was found in the last century to be the cause for "pernicious anemia." Recently B$_{12}$-deficiency was recognized as a rather common condition with older persons. In most cases B$_{12}$-deficiency results from impaired uptake from the ingested food. A dose of about 3–4 μg per day of B$_{12}$ are considered necessary for sustained physical well-being. In the human body B$_{12}$ is metabolically active as MeCbl and as AdoCbl.

Three soluble B$_{12}$-binding proteins are known to be involved in the uptake and transport of cobalamins in humans: intrinsic factor (IF), transcobalamin (TC) and haptocorrin (HC). These three genetically related B$_{12}$-binders (apparent binding constants of $> 10^{12} \, l \, mol^{-1}$) ensure that the needed amounts of B$_{12}$ reach the two enzymes, methionine synthase (in the cytosol) and methylmalonyl-CoA mutase (in the mitochondria). IF is synthesized in the stomach and binds cobalamins in the small intestines. The IF-cobalamin complex is absorbed by specific receptors on the brush borders of the epithelial cells in the small intestines. Intracellular B$_{12}$-trafficking depends upon a complex interplay between the B$_{12}$-binders and cellular surface receptors that recognize complexes between B$_{12}$ and the B$_{12}$-binding proteins IF, TC, and HC.

See Also the Following Articles

Giant Mitochondria (Megamitochondria) • Inorganic Biochemistry • Porphyrin Metabolism • Propionyl CoA–Succinyl CoA Pathway

Glossary

heteronuclear NMR Nuclear magnetic resonance (NMR) experiments that correlate nuclei of different elements (e.g., ^{1}H and ^{13}C) to obtain information on their mutual positions in a molecule (or a supramolecular complex).

organic radical Highly reactive organic molecules with unpaired electrons.

pernicious anemia Microscopic abnormalities of blood and bone marrow correlating with B$_{12}$-deficiency.

Further Reading

Banerjee, R. (1999). *Chemistry and Biochemistry of B$_{12}$*. Wiley, New York.

Buckel, W., and Golding, B. T. (1996). Glutamate and 2-methylene-glutarate mutase: From microbial curiosities to paradigms for coenzyme B$_{12}$-dependent enzymes. *Chem. Soc. Rev. (London)* **25**, 329–338.

Eschenmoser, A. (1988). Vitamin B$_{12}$: Experiments concerning the origin of its molecular structure. *Angew. Chem. Int. Ed. Engl.* **27**, 5–40.

Kratky, C., and Gruber, K. (2002). Coenzyme B$_{12}$ dependent glutamate mutase. *Curr. Opin. Chem. Biol.* **6**, 598–603.

Kräutler, B., and Ostermann, S. (2003). Structure, reactions and functions of B$_{12}$ and B$_{12}$-proteins. In *The Porphyrin Handbook*, Vol 11, pp. 227–274.

Kräutler, B., Arigoni, D., and Golding, B. T. (eds.) (1998). *Vitamin B$_{12}$ and B$_{12}$-Proteins*. Wiley-VCH, Weinheim, Germany.

Matthews, R. G. (2001). Cobalamin-dependent methyltransferases. *Acc. Chem. Res.* **34**, 681–689.

Stubbe, J. (2000). Ribonucleotide reductases: The link between an RNA and a DNA World? *Curr. Opin. Struct. Biol.* **10**, 731–736.

Toraya, T. (2003). Radical catalysis in coenzyme B$_{12}$-dependent isomerization (eliminating) reactions. *Chem. Rev.* **103**, 2095–2127.

BIOGRAPHY

Bernhard Kräutler studied chemistry at the ETH (Zürich), where he did a dissertation with Prof. Albert Eschenmoser. He had postdoctoral association with Prof. Allen Bard (Austin, Texas) and Prof. Nick Turro (Columbia University). After return to the ETH he obtained his habilitation in organic chemistry (1985). Since 1991 he is a Professor of Organic Chemistry at the University of Innsbruck. His main research interests include bioorganic chemistry, porphyrinoid compounds, coenzyme B$_{12}$, chlorophyll, fullerenes, structure/function of biomacromolecules, and supramolecular chemistry.

Vitamin C

Robert B. Rucker and Francene Steinberg
University of California, Davis, California, USA

One of the most important redox cofactors in plant and animal systems is ascorbic acid or vitamin C. Although most animals make sufficient ascorbic acid, for some animals, ascorbic acid is a true vitamin because of their inability to carry out synthesis. For example, in humans, ascorbic acid deficiency can result in the nutritional disease scurvy, which causes a range of pathologic symptoms, because of defects in ascorbic-acid-specific enzymatic steps and processes.

In most plants and animals, ascorbic acid is derived as a product from the direct oxidation of glucose, galactose, and mannose (in plants), e.g., glucose or galactose → UDP-D-glucuronic acid → glucuronic acid/glucuronolactone → gulono-1, 4-lactone → ascorbic acid. In animals, a key enzyme in this process is L-gulonolactone oxidase (GLO, EC 1.1.3.8), which catalyzes the terminal step in ascorbic acid synthesis, i.e., gulono-1, 4-lactone is oxidized to ascorbic acid. L-gulonolactone oxidase resides in the kidney of most birds and reptiles, and during the course of evolution was transferred to the liver of mammals. For reasons that are not clear, the ability to express sufficient L-gulonolactone oxidase activity disappeared from the guinea pig, some fruit-eating bats, and most primates, including humans. Accordingly, a dietary source of ascorbic acid is needed in these animals. There is also a possibility that minor alternative pathways exist in plants and mammals for ascorbic acid synthesis.

Chemistry

The chemical designation for ascorbic acid is 2-oxo-L-theo-hexono-4-lactone-2, 3-enediol. Ascorbic acid is a near planar five-member ring with two chiral centers that resolves into the four stereoisomers. The oxidized form of ascorbic acid, dehydroascorbic acid, retains vitamin C activity and can exist as a hydrated hemi-ketal. Crystalline dehydro-L-ascorbic acid can exist as a dimer. The most important chemical property of ascorbic acid is the reversible oxidation of ascorbic acid to semidehydro-L-ascorbic acid and to dehydroascorbic acid (Figure 1).

In addition to facilitating reduction–oxidation reactions, ascorbic acid has the ability to form relatively stable free radical intermediates. Ascorbic acid can act as a free radical scavenger in reactions involving reactive oxidant species (ROS). In this regard, the rate constants for the generation of ascorbate radicals vary considerably, but often dictate rapid radical formation, e.g., 10^5–10^{10} k_{obs}/M^{-1} s^{-1}. Once formed, ascorbate (Asc) radicals decay relatively slowly by a process of disproportionation (2 Asc$^{·-}$ + H^+ → AscH$^-$ + DHA). The ability to form free radical intermediates can significantly delay or prevent free radical-initiated oxidations. Ascorbic acid readily scavenges reactive oxygen and nitrogen species, such as superoxide, hydroperoxyl, peroxynitrite, and nitroxide radicals. Asc[H$^-$] donates a hydrogen atom (H$^·$ or H^+ + e^-) to an oxidizing radical to produce the resonance-stabilized tricarbonyl ascorbate free radical. Asc[H$^·$] has a pK_a of -0.86; thus, it is not protonated. Further, the unpaired electron of Asc [$-$] resides in the π-system that includes the tricarbonyl moiety of ascorbate. It is a weakly oxidizing and reducing radical. Due to its π-character, Asc [$-$] does not react with oxygen to form peroxyl radicals capable of damaging oxidations. It is relatively unreactive with a one-electron reduction potential of only $+282$ MV, and as such is considered a terminal, small-molecule antioxidant.

Ascorbic acid is often associated with the protection of lipid, DNA, and proteins from oxidants. As examples, when peroxyl radicals are generated in plasma, vitamin C is consumed faster than other antioxidants, e.g., uric acid, bilirubins, and vitamin E. Ascorbic acid is 10^3 times more reactive than a polyunsaturated fatty acid in reacting with peroxyl radicals. In contrast, ascorbic acid can be viewed as a pro-oxidant under aerobic condition when metals capable of redox ($Fe^{+2} \leftrightarrow Fe^{+3}$; $Cu^{+1} \leftrightarrow Cu^{+2}$) are also present. Metals, such as iron and copper in their reduced states, are effective Fenton catalysts.

With regard to other chemical properties, the acidity of ascorbic acid is due to the low pK_a of the proton on oxygen-3. In addition, ascorbic acid is not very stable in aqueous media, wherein it can decay within a few hours or even minutes at high pH (>10.0). In contrast, ascorbic acid is relatively stable in blood (a day or more), or if stored at acid pH (<3.0) or below $-20°C$ (often weeks to months).
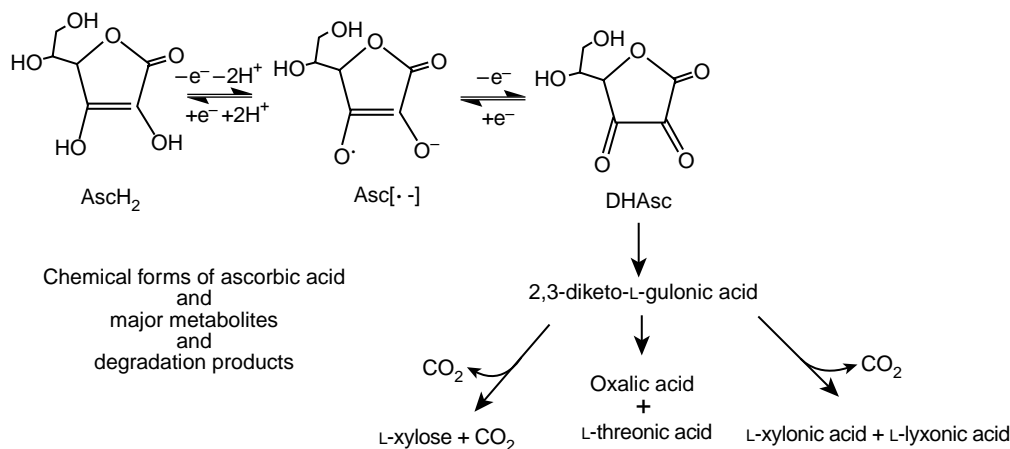
FIGURE 1 The major redox forms of ascorbic acid are shown in addition to products that result from both the chemical and enzymatic degradation of ascorbic acid. In alkaline solutions ascorbic acid is easily modified to 1 carbon (CO_2) and various 5-carbon compounds (e.g., xylonic acid, lyxonic acid) or the 2-carbon compound, oxalic acid and the 4-carbon compound, L-threonic acid. In addition, L-xylose and CO_2 may be formed *in vivo* following decarboxylation at the C-1 position of ascorbic acid followed by reduction.

Nutritional and Biochemical Importance

Ascorbic acid functions in many mono- and dioxygenases to maintain metals in a reduced state. For example, mono- and dioxygenases usually contain copper or iron as redox cofactors, respectively. As an additional characteristic, dioxygenases require α-ketoglutarate and O_2 as cosubstrates in reactions, while mono-oxygenases require only O_2. Important reactions and processes that require ascorbic acid include:

• Norepinephrine synthesis – dopamine-β-hydroxylase, EC 1.14.17.1, is the final and rate-limiting step in the synthesis of norepinephrine and requires copper and ascorbic acid for activity.

• Hormone activation – posttranslational steps involving α-amidations activate many hormones and hormone-releasing factors. Examples of hormone activation include melanotropins, calcitonin, releasing factors for growth hormone, corticotrophin and thyrotropin, pro-ACTH, vasopressin, oxytocin, cholecystokinin, and gastrin. Petidylglycine α-amidating mono-oxygenase (EC 1.14.17.3), the enzyme that carries out α-amidation is found in secretory granules of neuroendocrine cells.

• Carnitine Biosynthesis – ascorbate is a cofactor for two of the hydroxylation steps in the pathway of carnitine biosynthesis, gamma-butyrobetaine hydroxylase and epsilon-N-trimethyllysine hydroxylase. Ascorbate deficiency results in as much as 50% decrease in carnitine in heart and skeletal muscle in animals that require ascorbic acid.

• Prolyl and lysyl hydroxylations – ascorbic acid is an important cofactor in the hydroxylation of proline and lysine amino acid residues. In scurvy, poor wound healing, bruising, and osteopenic abnormalities can result from the under-hydroxylation of collagens. The various types of collagens are distinguished by their relative amounts of hydroxyproline and hydroxylysine. Hydroxylation of proline and lysine is important in collagen assembly and maturation into fibers. Because collagen is the principal connective tissue protein in the body, it is essential to all phases of normal growth, development, and repair. There are also other proteins that are not defined as collagens, which contain hydroxyproline, e.g., the subcomponent C1q of complement, acetylcholine esterase, pulmonary surfactant proteins, and proteins that function as cell surface receptors. In plants, prolyl hydroxylase hydroxylates proline residues in cell wall hydroxyproline-rich glycoproteins required for cell division and expansion. Approximately 50 hydroxyproline-containing proteins have been identified in viruses and bacteria, wherein ascorbic acid may also play a role as cofactor.

Along with glutathione, ascorbic acid is also an important nonenzymatic antioxidant/reductant in the water-soluble compartment of cells. Glutathione is L-(-glutamyl-L-cysteine-glycine (GSH). The antioxidant actions of glutathione and ascorbate are closely linked and involve mechanisms in which decreased glutathione may even stimulate ascorbic acid synthesis in those animals that can produce it. During postnatal development, a rapid change occurs as animals adapt, from a relatively hypoxic to a relatively hyperoxic environment. In this regard, one of glutathione's many functions is to keep ascorbic acid in a reduced form. In adult animals that can make ascorbic acid, a reduction in glutathione levels can lead to a rapid increase in liver dehydroascorbic acid. An adequate ascorbic acid intake is

particularly important in newborns and neonates, who do not yet have the potential to synthesize ascorbate. A deficiency of both ascorbic acid and glutathione can cause severe damage to liver and other organs.

With regard to plants, ascorbic acid reaches a concentration of over 20 mM in chloroplasts and occurs in all cell compartments, including the cell wall. It has a major role in photosynthesis, acting in the Mehler peroxidase reaction with ascorbate peroxidase to regulate the redox state of photosynthetic electron carriers and as a cofactor for violaxanthin de-epoxidase, an enzyme involved in xanthophyll cycle-mediated photoprotection. Ascorbic acid also serves as a cofactor in metabolic pathways important to ethylene, gibberellins, and anthocyanins. There is a rapid response of ascorbate peroxidase expression to (photo)-oxidative stress. Of interest, the ascorbate-deficient vtc 1 mutant of *Arabidopsis thaliana*, is very sensitive to ozone, ultraviolet-B radiation, and other forms of oxidant stress.

Cellular Regulation of Ascorbic Acid

Specific non-overlapping transport proteins mediate the transport of ascorbic acid across biological membranes. Dehydroascorbic acid uptake is via the facilitated-diffusion glucose transporters, (GLUT 1, 3, and 4), but under physiological conditions these transporters do not appear to play a major role in the uptake of dehydroascorbic acid due to the high concentrations of glucose that effectively block influx. L-ascorbic acid enters cells via Na^+-dependent systems. Two isoforms of these transporters have recently been cloned from human and rat sources, which differ in their tissue distribution. Cell accumulation of ascorbic acid occurs, because of cellular dehydroascorbate reduction systems that are capable of rapidly generating ascorbic acid.

In addition to cell regulation, ascorbate reduction systems are also important in reducing the ascorbic acid radical. As indicated, the ascorbic acid radical is relatively stable. An accessory enzymatic system is needed to reduce the potential of transient accumulation of the ascorbate radical. Excess ascorbate radicals may initiate free radical cascade reaction or nonspecific oxidations. In plants, NADH-monodehydroascorbate reductase (EC 1.6.5.4) has evolved to maintain ascorbic acid in its reduced form. NADH-monodehydroascorbate reductase plays a major role in stress-related responses in plants. In animal tissues, glutathione dehydroascorbate reductase (EC 1.8.5.1) serves this purpose.

Digestion, Absorption, and Nutritional Requirements

When ascorbic acid is consumed in the diet, the ileum and jejunum are major sites of ascorbic acid absorption. The bioavailability of vitamin C is dose dependent. In humans, saturation of transport occurs with dosages of 200–400 mg per day. Approximately 70% of a 500 mg dose is absorbed. However, much of the absorbed dose (>50%) is nonmetabolized and excreted in the urine. With a dose of 1250 mg, only 50% of the dose absorbed and most (>85%) of the absorbed dose is excreted. Vitamin C is not protein-bound and is eliminated with an elimination half-life of 10–12 hours. In Western populations plasma, vitamin C concentrations range from 54 to 91 $\mu mol\, l^{-1}$. Tissue concentrations range from μM to mM, i.e., high tissue levels of vitamin C are well tolerated in mammalian systems and lead to the conclusion that ascorbic acid is relatively nontoxic, although there is also some evidence that accelerated metabolism or disposal of ascorbic acid may occur after prolonged supplementation of high doses.

Sources and Detection

A number of approaches are available for ascorbic acid detection. Ascorbic acid absorbs strongly in UV light, the basis for some spectrophotometric assays. The redox properties of ascorbic acid allow for electrochemical detection or detection from interacting with redox-sensitive chromophores and dyes. Enzymatic methods (using ascorbate oxidase) and conventional and isotope ratio mass spectrometry are used for detection and quantitation of ascorbic acid following separation by means of ion exchange, absorption or partition chromatographic approaches. Typical food and tissue values are given in Table I.

Defining Ascorbic Acid Status

The requirements for vitamin C in humans (given as the range for female–male and expressed as the recommended dietary allowance, RDA or adequate uptake, AI) are as follows: infants, 40 mg per day (AI based on what is present in human milk); children (3–13 years), 15–45 mg per day (extrapolated from the RDA for adults); adolescence, 45–65 mg per day; adults (>19 years), 75–90 mg per day; lactating mothers, 115–120 mg per day. In keeping with these estimates, more recent clinical investigations have noted vitamin C deficiency in 2–6% of individuals surveyed (often defined as individuals meeting less than two-thirds of the RDA). With regard to marginal vitamin C status

TABLE I

Vitamin C in Selected Foods and Tissues

Sources of vitamin C	Mg of ascorbic acid per 100 g of wet weight or edible portion
Animal products	
Cow's milk	0.5–2
Human milk	3–6
Beef, pork, veal	2–10
Liver, chicken	15–20
Kidney, chicken	6–8
Heart, chicken	5
Crab muscle, lobster	1–4
Shrimp	2–4
Fruits	
Apple	3–30
Banana	8–16
Blackberry	8–10
Cherry	15–30
Currant, red	20–50
Currant, black	150–200
Grapefruit	30–70
Kiwi fruit	80–90
Lemon, orange	40–50
Melon	9–60
Strawberry	59–70
Pineapple	15–25
Rose hips	250–800
Vegetables	
Beans, various	10–15
Brussels sprouts	100–120
Cabbage	30–70
Carrot	5–10
Cucumber	6–8
Cauliflower	50–70
Eggplant	15–20
Chive	40–50
Kale	70–100
Onion	10–15
Peas	8–12
Potato	4–30
Pumpkin	15
Radishes	25
Spinach	35–40
Condiments	
Chicory	30–40
Coriander (spice)	90
Garlic	15–25
Horseradish	50
Lettuce, various	10–30
Leek	15
Parsley	200–300
Pepper, various	150–200

based on plasma ascorbic acid values, e.g., $<20\ \mu g/dL$, the prevalence ranges from 17 to 24% US. In this regard, smokers are more likely to have marginal vitamin C status compared to nonsmoking adults. Several studies suggest that smokers require over 200 mg vitamin C daily to maintain plasma vitamin C concentrations at a level equivalent to that of nonsmokers.

Summary

Ascorbic acid usually carries out redox reactions by mechanisms dependent upon free-radical processes. Ascorbate metabolism is linked to the metabolism of glutathione. Ascorbic acid is also required in animals that lack or have mutations in the gene for L-gulonolactone oxidase. Ascorbic deficiency results in reduced mono- and dioxygenase activities. The consequences of severe deficiency are profound, since growth, extracellular matrix, and hormonal regulation are impaired. Recent data suggest optimal intakes of ascorbic acid, based on a range of criteria, should be as high as 75–90 mg per day for adult humans.

SEE ALSO THE FOLLOWING ARTICLES

Carnitine and β-Oxidation • Peptide Amidation

GLOSSARY

adequate intake (AI) One of the four terms used to define dietary reference intakes (DRIs), which are reference values that are quantitative estimates of nutrient intakes to be used for planning and assessing diets for healthy people. The AI is used when the RDA for a nutrient is not available. It is an RDI that is based on observed or experimentally determined approximations of nutrient intake by a group of healthy people.

amidation A complex reaction, which in some hormones results in a C-terminal protein modification. A C-terminus consensus sequence is required. The consensus is glycine, followed by two basic amino acids (Arg or Lys). The end product is an amide of glycine ($-NH-CH_2-CO-NH_2$) with loss of the basic amino acid residues.

anthocyanins Naturally occurring compounds that impart color to fruit, vegetables, and plants. They also have antioxidant and insecticidal properties.

carnitine Essential component of fatty acid transport into and out of mitochondria for their eventual oxidation. Carnitine is derived from the amino acid lysine following extensive modification.

complement A term originally used to refer to the heat labile factor in serum that causes immune cytolysis, the lysis of antibody-coated cells. The term now refers to the entire functionally related system comprising at least 20 distinct serum proteins that is the effector immune cytolysis and other biologic functions immune system related functions.

dioxygenase An oxidoreductase that incorporates two atoms of oxygen (from one molecule of O_2) into the (reduced) substrate.

Fenton reaction The formation of OH˙, OH−, and Fe^{3+} from the nonenzymatic reaction of Fe^{2+} with H_2O_2.

gibberellins   Plant hormones that stimulate growth in the stem and leaves, and trigger the germination of the seed

Mehler reaction   Describes the photoreduction of oxygen in chloroplasts by photosystem I in plants yielding O_2^-.

monooxygenase   Oxidoreductases that induce the incorporation of one atom of oxygen from O_2 into the substance being oxidized.

recommended dietary allowance (RDA)   The average daily dietary intake level that is sufficient to meet the nutrient requirement of nearly all (97–98%) healthy individuals in a group.

redox   oxidation–reduction reactionary reaction in which electrons are removed from one molecule or atom and given to another molecule or atom.

transporter   A class of transmembrane protein that allows substances to cross plasma membranes far faster than would be possible by diffusion alone. A major class of transport proteins that expend energy to move substances (called active transport) are the transport ATPases.

violaxanthin   Photopigment involved in photoprotection in plants. When light energy absorbed by plants becomes excessive (relative to the capacity of photosynthesis), the xanthophyll (a chemical classification designation), violaxanthin is reversibly modified by violaxanthin de-epoxidase as a protective function in plants.

FURTHER READING

Frei, B. (ed.) (1994). *Natural Antioxidants in Human Health and Disease*. Academic Press, San Diego.

National Academy Press (2000). *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids*. A report of the Panel on Dietary Antioxidants and Related Compounds, Subcommittees on Upper Reference Levels of Nutrients and Interpretation and Uses of Dietary Reference Intakes, and the Standing Committee on the Scientific Evaluation of Dietary Reference Intakes, Food and Nutrition Board, Institute of Medicine, National Academy Press, Washington, DC.

Packer, L., and Fuchs, J. (eds.) (1997). *Vitamin C in Health and Disease*. Marcel Dekker, New York.

Simopoulos, A. P (Vol. Ed.) (1991). Selected vitamins, minerals, and functional consequences of maternal malnutrition. In *The Series World Review of Nutrition and Dietetics*,Vol. 64, Karger, Basel, New York.

BIOGRAPHY

Robert Rucker is a Professor of Nutrition and Internal Medicine at the University of California, Davis. Dr. Rucker's research interests are focused on extracellular matrix assembly, the role of micronutrients in early growth and development, and the physiological roles of quinone cofactors derived from tyrosine, such as pyrroloquinoline quinone. He holds a Ph.D. from Purdue University. Honors and appointments include serving as a past president of the *American Society for the Nutritional Sciences*, serving as an associate editor for the *American Journal of Clinical Nutrition*, receiving the Borden Research Award from the *American Society for Nutritional Sciences*, and recognition as a Fellow of the AAAS.

Francene Steinberg is an Associate Professor in the Department of Nutrition at UC Davis and Director of the Didactic Program in Dietetics. She received the Ph.D. in nutrition at UC Davis, received additional postdoctoral training in the Department of Endocrinology at the University of Washington. She is also a registered Dietitian. Dr. Steinberg's research area concerns the metabolism of lipids and lipoproteins, with an emphasis on the effects of nutrition on cardiovascular disease risk and the development of atherosclerosis. Areas of particular interest include dietary soy phytoestrogens, structured lipids, and antioxidant vitamins.

Vitamin D Receptor

Diane R. Dowd and Paul N. MacDonald
Case Western Reserve University, Cleveland, Ohio, USA

The vitamin D receptor is a member of the nuclear receptor/steroid hormone receptor superfamily. These receptors function as ligand-activated, transcriptional regulatory proteins. Thus, the vitamin D receptor selectively binds the 1,25-dihydroxyvitamin D_3 [$1,25(OH)_2D_3$] hormone and controls the expression of selected genes in target cells. The main actions of $1,25(OH)_2D_3$, VDR, and the vitamin D endocrine system are to maintain overall calcium and mineral homeostasis. A fundamental physiological role is to maintain adequate absorption of dietary calcium at the level of the small intestine. Dietary deficiency of vitamin D or mutations in the vitamin D receptor result in human conditions of hypocalcemia and secondary skeletal undermineralization. The molecular details involved in vitamin D-regulated gene expression by the VDR in key mineral regulating tissues such as the intestine, kidney, and bone are revealing important insights into the manner in which our bodies maintain the structural integrity of skeletal tissue.

Introduction

Vitamin D_3, or cholecalciferol, was discovered nearly a century ago as a micronutrient that is essential for normal skeletal development and for maintaining bone integrity. $1,25(OH)_2D_3$ is the bioactive, hormonal form of vitamin D. However, vitamin D_3 is more appropriately classified as a hormone since it is produced in the body in response to serum calcium levels and it functions through a mechanism that is analogous to other steroid hormones. It is generated by two sequential hydroxylations of vitamin D_3 in response to hypocalcemia and elevated parathyroid hormone. The predominant role of $1,25(OH)_2D_3$ is to enhance the intestinal absorption of dietary calcium and phosphorus. $1,25(OH)_2D_3$ also acts on mineral-regulating target tissues such as intestine, bone, kidney, and parathyroid glands to maintain normal calcium and mineral homeostasis. $1,25(OH)_2D_3$ also directly affects skeletal bone remodeling by causing osteoblasts to terminally differentiate into osteocytes and deposit calcified matrix. $1,25(OH)_2D_3$ also promotes the differentiation of precursor cells into mature osteoclasts which function to resorb bone and maintain appropriate bone remodeling.

Thus, it is through an integrated series of diverse effects that vitamin D is thought to preserve and maintain the integrity of the bony tissues. To highlight this role, vitamin D_3 deficiency leads to rickets in children and osteomalacia in adults. The vitamin D endocrine system is also involved in a number of other important physiological processes including blood pressure regulation, immune function, mammary gland development, and hair follicle cycling.

The biological effects of $1,25(OH)_2D_3$ are mediated through a soluble receptor protein termed the vitamin D receptor (VDR), a member of the steroid receptor or nuclear receptor (NR) superfamily. The VDR binds $1,25(OH)_2D_3$ with high affinity and high selectivity. In the target cell, the interaction of the $1,25(OH)_2D_3$ hormone with VDR initiates a complex cascade of molecular events culminating in alterations in the rate of transcription of specific genes or gene networks (Figure 1). Central to this mechanism, is the requisite interaction of VDR with retinoid X receptor (RXR) to form a heterodimeric complex that binds to specific DNA sequence elements (VDREs) in vitamin D-responsive target genes and ultimately influences the rate of RNA polymerase II-mediated transcription. An emerging concept in this mechanism is that protein–protein interactions between the VDR–RXR heterodimer, transcriptional comodulatory proteins and the transcription machinery are essential for the mechanism of vitamin D-mediated gene expression. This article discusses the molecular biology of the VDR with a focus on the macromolecular interactions that are required for the transcriptional regulatory activity of the VDR.

Functional Domains of the VDR

THE N-TERMINAL DNA-BINDING DOMAIN (DBD)

Receptors in the nuclear receptor superfamily generally have two transcriptional activation domains, termed activation function (AF) 1 and 2, which are required for the receptor to function as a ligand-activated
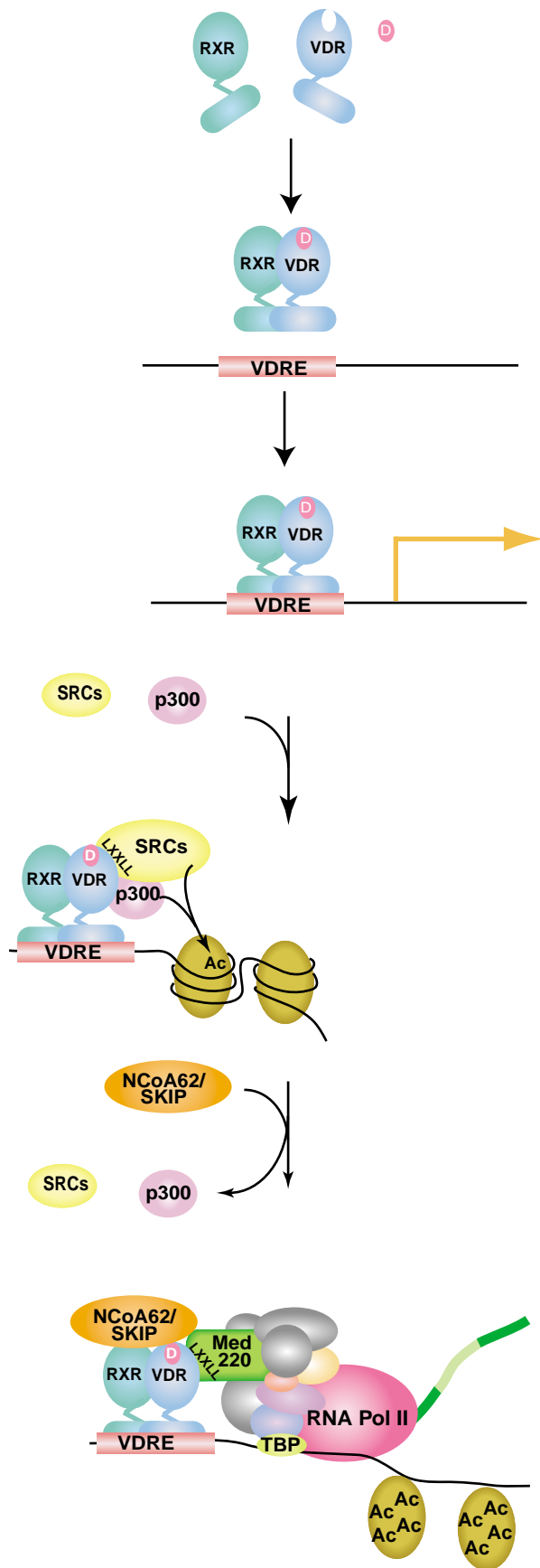
transcription factor. The AF-2 domain exists at the extreme COOH terminus while the AF-1 domain is located in the N-terminal region (or A/B domain) of the receptors (Figure 2). The AF-1 domain is a constitutive (i.e., hormone-independent) activation domain. In the VDR, the N-terminal A/B domain is truncated compared to other nuclear receptors and thus an analogous constitutive AF-1 domain is lacking in the VDR.

In order for VDR to regulate transcription of target genes, it must recognize and bind to DNA in the promoter regions of vitamin D-responsive genes. It does so through a specialized DNA-binding domain (DBD) located near the amino terminus of the receptor (Figure 3). This DBD is required for the VDR to bind selectively and with high affinity to specific DNA sequences termed vitamin D-responsive elements (VDREs). The minimal DBD of the VDR that mediates VDR–DNA interactions resides between amino acid residues 22 and 113 in the human sequence. There are nine cysteine residues within the DBD that are conserved throughout the family members. The first eight of these cysteines (counting from the N terminus) tetrahedrally coordinate two zinc atoms to form two zinc finger DNA-binding motifs. Mutation of the first eight of the nine cysteine residues to serines eliminates VDR binding to both nonspecific and specific DNA sequences and eliminates VDR-mediated transactivation. Indeed, inactivating mutations within the DBD of the human VDR are responsible for the rare inherited disorder termed vitamin D-resistant rickets type II. These patients express a nonfunctional VDR that cannot bind to DNA, and they exhibit classical symptoms of vitamin D deficiency that are not corrected by providing an external source of vitamin D.

LIGAND-BINDING DOMAIN (LBD)

The large carboxyl-terminal ligand-binding domain (LBD) of VDR is organized into 12 α-helices. As the name suggests, the LBD is responsible for high-affinity binding of the $1,25(OH)_2D_3$ ligand, exhibiting equilibrium binding constants of the order of 10^{-10} to

FIGURE 1 Mechanism of action of the vitamin D receptor. VDR binds to its ligand, $1,25(OH)_2D_3$, heterodimerizes with RXR, and binds to a DR-3 type VDRE in the promoter of $1,25(OH)_2D_3$-responsive genes. VDR occupies the proximal or 3′ half-site while RXR occupies the distal or 5′ half-site. VDR/RXR then recruits coactivators such as SRCs and p300/CBP. The HAT activity of these coactivators loosens the chromatin structure, allowing for a more transcriptionally permissive environment. SRCs and p300/CBP dissociate, allowing for binding of NCoA62/SKIP and the Mediator-D multimeric complex. This latter complex is thought to recruit RNA Pol II and the core transcriptional machinery to initiate active transcription of the target gene.

FIGURE 2   Functional domains of the VDR. A/B, amino terminal region; DBD, DNA- binding domain containing two zinc-finger motifs; LBD, ligand-binding domain; AF-2, activation function-2 domain encompassing helix 12.

10^{-11} M. Both the 1-hydroxyl and 25-hydroxyl moieties are crucial for efficient recognition by VDR, and the nonhydroxylated, inactive parent vitamin D_3 compound does not bind appreciably to the VDR.

In addition to hormone binding, the LBD is required for several other aspects of receptor function, particularly in mediating protein–protein interactions. One important protein–protein contact is the heterodimerization of VDR with the retinoid X receptor (RXR). As mentioned above, VDR–RXR heterodimer formation is generally required for high-affinity interaction of the receptor with VDREs and an extensive heterodimerization interface that coalesces around helices 10 and 11 in the LBD of VDR mediate protein–protein contacts with RXRs.

The LBD also mediates association of VDR with comodulatory proteins. Many of these interactions require at the extreme C-terminus of the receptor, where the AF-2 domain is located (Figure 2). The AF-2 domain is highly conserved throughout the hormone receptor superfamily and its main structural feature is that of an amphipathic α-helix. Removing 25 amino acids from the C-terminus of hVDR ($\Delta 403$–427), which contains the AF-2 domain, results in a complete loss of $1,25(OH)_2D_3$/VDR-activated transcription. This loss of function is not due to altered binding of RXR, VDRE, or hormone and the mutant receptor was appropriately targeted to the cell nucleus. Thus, the AF-2 domain of VDR, corresponding to helix 12, plays a central role in $1,25(OH)_2D_3$-activated transcription mediated by VDR.

Molecular Mechanism of Transcriptional Control by VDR

VDR INTERACTION WITH VITAMIN D RESPONSIVE ELEMENTS

VDR modulates transcription by binding to specific VDREs, in the promoter regions of responsive genes. VDREs from a variety target genes have been identified and on this basis a consensus VDRE may be generally described as an imperfect direct repeat of a core hexanucleotide sequence,*G/AGGTG/CA,* with a spacer region of three nucleotides separating each half element (also termed DR-3 for direct repeat with a three nucleotide spacer). The mechanism of VDR binding to VDREs is reflected in the direct repeat nature of the element. For example, purified VDR alone does not bind to a VDRE with high affinity. Rather, RXR is required for high-affinity binding of VDR to VDREs. The asymmetry of the directly repeating motif causes the VDR–RXR heterodimer to bind to the VDRE with a defined polarity, with RXR occupying the 5′ half site and VDR occupying the 3′ half site.

The natural VDREs identified thus far provide only a snapshot of the DNA sequences that mediate the transcriptional effects of the VDR. Variations on the DR-3 motif for VDREs have been identified in the elements that mediate vitamin D responsiveness in the calbindin D_{9k} and calbindin D_{28k} genes. Moreover, several synthetic elements with large spacer regions and inverted arrangements can mediate vitamin D responsiveness under certain conditions. It is likely that the affinity of VDR for these atypical elements may vary from that of the classic DR-3 motif adding yet another level of regulatory complexity to the process of VDR-mediated gene expression.
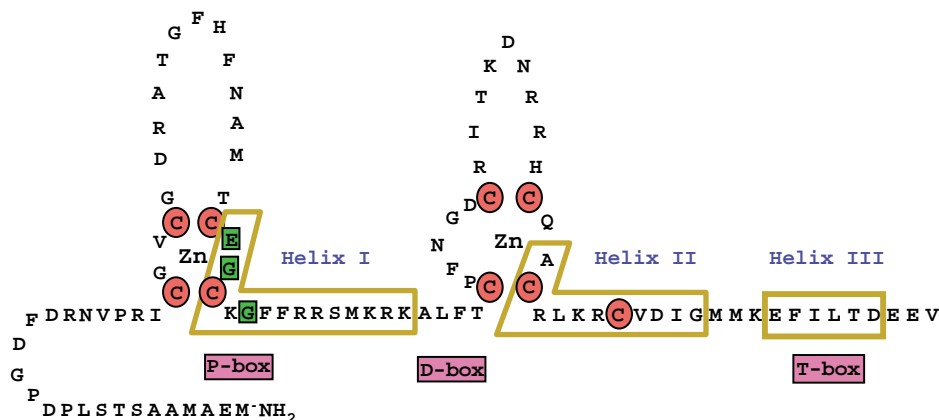
FIGURE 3   Schematic of the DNA-binding domain of the VDR. The cysteines responsible for coordinating the zinc atoms are shown in red.

One important role for the $1,25(OH)_2D_3$ ligand in the transactivation process is to dramatically enhance both the formation of the VDR–RXR heterodimer and the subsequent interaction of the heterodimer with the VDRE (Figure 1). Upon ligand binding, there is a clear $1,25(OH)_2D_3$-dependent decrease in the formation of VDR homodimers and a concomitant increase in VDR heterodimerization with RXR. These ligand-induced changes in VDR–RXR interactions are likely due to altered conformations of the VDR that disrupts weak homodimers of unliganded VDR and promotes liganded VDR heterodimerization with RXR. The interaction of VDR and RXR generates a heterodimeric complex that is highly competent to bind DR-3 like VDREs and subsequently affect the transcriptional process. Thus, the VDR–RXR heterodimer is the primary active complex involved in controlling DR-3 VDRE-driven promoters.

COMMUNICATION BETWEEN VDR AND THE TRANSCRIPTIONAL MACHINERY

The regulation of VDR-mediated transcription involves a complex series of macromolecular interactions occurring in a temporally coordinated fashion (Figure 1). Other transcriptional components that associate with the liganded VDR/RXR heterodimer may be classified into two general categories: general transcription factors and the comodulatory proteins. The interaction of VDR with the first group results in direct contacts with the preinitiation complex (PIC), which may facilitate assembly or recruitment of the PIC and thereby stimulate transcription by RNA Pol II. Transactivator interaction with the PIC may occur via adapter proteins such as the TBP-associated factors (TAFs) or via direct interaction with other core transcription factors. In this regard, TFIIB may be a key target since it is known to interact directly with VDR and augment vitamin D-activated transcription in transient gene expression studies. Indeed, differences in the activities of NH_2-terminal VDR isoforms are attributed to differential interactions with TFIIB thus, supporting an important role for the VDR–TFIIB interaction in determining the overall transactivation potential of the liganded VDR.

In addition to direct contacts with the general transcription machinery, the liganded VDR is also linked to the transcriptional PIC by the NR comodulatory factors. The general functional properties of the NR comodulators are their ability to interact with nuclear receptors and control their transcriptional responsiveness to ligand. NR comodulators are classified either as coactivators or corepressors, and they aid in the induction or repression, respectively, of ligand/receptor-mediated transcription. NR coactivators may function as macromolecular bridges between the receptor and the transcriptional machinery that aid in the assembly or promote the stability of the preinitiation complex. Moreover, several coactivator proteins either possess intrinsic histone acetyltransferase (HAT) activities, or recruit other proteins that possess HAT activity (e.g., CBP/P300). Acetylation of histones near the promoter presumably results in a loosening of chromatin structure and greater accessibility of the promoter for the transcriptional machinery (Figure 1). The best-characterized coactivators are the steroid receptor coactivator (SRC) family of nuclear receptor coactivators that includes three members at present: SRC-1 (NCoA-1), SRC-2 (GRIP-1, TIF2, NCoA-2), and SRC-3 (pCIP, RAC3, ACTR, AIB-1, TRAM-1).

Structural studies of related receptors provide insight into the mechanism of ligand-induced interaction of VDR with SRC coactivators. It is hypothesized that the $1,25(OH)_2D_3$ ligand promotes coactivator interaction by inducing a repositioning of the AF-2 activation helix, or helix H12 (Figure 4). In the unliganded state, the AF-2 domain (helix H12) projects out away from the globular core of the LBD, while in the liganded state the AF-2 domain is folded over onto the LBD globular core domain. One outcome of helix H12 folding is the creation of a platform or protein interaction surface through which coactivator proteins such as SRCs effectively dock with the VDR. The domain in the coactivator proteins, which is responsible for this interaction is termed the nuclear receptor or NR box and is an amphipathic α-helix composed of the consensus core sequence LXXLL. This interaction is essential for transactivation mediated by the VDR.

A large multiprotein complex called Mediator D is another coactivator required for VDR-mediated transcription (Figure 1). Mediator D is also known as vitamin D receptor interacting protein (DRIP), thyroid receptor activating protein (TRAP) complex, and the mammalian mediator complex. The Mediator D complex is composed of at least ten different proteins anchored by Med220, which interacts directly with ligand-activated VDR/RXR heterodimers through one of two LXXLL motifs. DRIP is essential for *in vitro* VDR-mediated transcription from chromatinized templates. However, as the DRIP complex does not contain any SRCs and does not possess HAT activity, it appears to potentiate NR-mediated transcription through distinct mechanisms. In particular, DRIP directly recruits the RNA polymerase II holoenzyme to $1,25(OH)_2D_3$-activated VDR, indicating that DRIP may serve as a bridge between VDR and the core transcriptional machinery. Indeed, chromatin immuno-precipitation studies show that DRIPs enter transcriptional complexes at a later time than SRC coactivators, supporting the nonredundant roles of these coactivator classes in NR-mediated transactivation.

Another protein named NCoA62/SKIP was identified as a VDR and NR coactivator, and as such, it augments
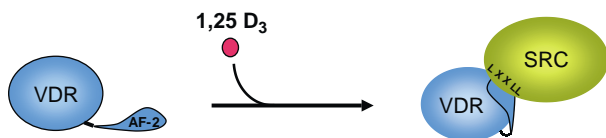
FIGURE 4 The vitamin D receptor undergoes a conformational change upon binding hormone. The VDR binds ligand and this induces a conformational change in the AF-2 domain to trap the ligand in the binding pocket. This change also creates a hydrophobic cleft or surface on VDR that LXXLL motifs in coactivator proteins use for docking to the VDR.

VDR- and other NR-activated transcriptional processes in transient reporter gene assays. On the basis of its primary amino acid sequence and its function, NCoA62/SKIP is a unique protein. One key difference is the lack of LXXLL motifs which are characteristic of a large variety of coactivators including the SRCs, CBP/P300 and the Mediator D coactivator complex. In addition, NCoA62 interacts with VDR in a ligand- and AF2-independent manner. Although NCoA62/SKIP does not interact directly with SRC coactivators, NCoA62/SKIP, VDR, and SRCs form a ligand-dependent ternary complex. NCoA62/SKIP also functions cooperatively with SRC coactivators to augment VDR-activated transcription and it is recruited to vitamin D responsive target genes in a distinct, temporal manner from the SRC coactivators. Presently, the mechanisms involved in this synergistic action of these two coactivator classes are unknown.

Summary

Research spanning the 1980s and 1990s reveals that $1,25(OH)_2D_3$-mediated transcription is more complex than the simple binding of the receptor to DNA and the recruitment of RNA Pol II to initiate transcription. VDR/RXR-activated transcription involves complex interactions that may occur in a spatially distinct and temporally coordinated fashion to increase the rate at which $1,25(OH)_2D_3$-responsive genes are transcribed and at which the resulting RNA transcript is processed. A model for VDR-mediated transcription is proposed which incorporates numerous properties of VDR that were discussed in this section (Figure 1). The initial event in this model is high-affinity binding of the $1,25(OH)_2D_3$ ligand to the VDR. Ligand binding induces VDR/RXR heterodimerization and the heterodimer specifically binds VDREs in the promoter regions of vitamin D responsive genes. The heterodimer then recruits coactivator molecules that acetylate core histones to make the DNA more accessible to the transcriptional machinery. Meanwhile, other coactivators and the PIC are recruited to the site and activated transcription proceeds. Understanding the complex interplay that occurs between these various factors is crucial to unraveling the complexities of activated or repressed transcription mediated by vitamin D and the VDR.

SEE ALSO THE FOLLOWING ARTICLES

Calcium Buffering Proteins: Calbindin • Vitamin D • Zinc Fingers

GLOSSARY

comodulatory proteins Proteins that interact directly with the nuclear receptor and alter the transcriptional regulatory activity of the receptor. Coactivators enhance the activity of the receptor, and corepressors attenuate the transcriptional regulatory activity of the receptor.

1,25-dihydroxyvitamin D_3 ($1,25(OH)_2D_3$) The active form of vitamin D in mammals.

histone acetyltransferase (HAT) Enzymes that add acetyl groups to the positively charged lysine residues on core histones, effectively negating the positive charge. Histone hyperacetylation is correlated with areas of active transcription, while hypoacetylation is correlated with nontranscribed promoters.

preinitiation complex (PIC) The complex of basic transcription factors that are necessary and sufficient for the initiation of transcription to occur. In addition to RNA Polymerase II, the PIC consists of transcription factor IIA (TFIIA), TFIIB, TFIIE, TFIIF, and TATA-binding protein or TBP.

retinoid X receptor (RXR) A member of the nuclear receptor superfamily of transcription factors that serves as a common heterodimeric partner for many of the class II nuclear receptors including retinoic acid receptor, vitamin D receptor, thyroid hormone receptor, and peroxisome proliferator-activating receptor.

vitamin D responsive elements (VDRE) Specific sequences of DNA that the VDR–RXR heterodimer selectively recognize and bind. This binding event is one of the initial steps in the mechanism through which 1,25-dihydroxyvitamin D_3 and the VDR regulate gene transcription.

FURTHER READING

Feldman, D., Glorieux, F. H., and Pike, J. W. (eds.) (1997). *Vitamin D.* Academic Press, San Diego.

Haussler, M. R., Whitfield, G. K., Haussler, C. A., Hsieh, J. C., Thompson, P. D., Selznick, S. H., Dominguez, C. E., and Jurutka, P. W. (1998). The nuclear vitamin D receptor: Biological and molecular regulatory properties revealed. *J. Bone Miner. Res.* **13**, 325–349.

Jones, G., Strugnell, S. A., and DeLuca, H. F. (1998). Current understanding of the molecular actions of vitamin D. *Physiol. Rev.* **78**, 1193–1231.

Jurutka, P. W, Whitfield, G. K., Hsieh, J. C., Thompson, P. D., Haussler, C. A., and Haussler, M. R. (2001). Molecular nature of the vitamin D receptor and its role in regulation of gene expression. *Rev Endocr. Metab. Disord.* **2**, 203–216.

Malloy, P. J., Pike, J. W., and Feldman, D. (1999). The vitamin D receptor and the syndrome of hereditary 1,25-dihydroxyvitamin D-resistant rickets. *Endocr. Rev.* **20**, 156–188.

McKenna, N. J., Lanz, R. B., and O'Malley, B. W. (1999). Nuclear receptor coregulators: Cellular and molecular biology. *Endocr. Rev.* **20**, 321–344.

Norman, A. W., Mizwicki, M. T., and Okamura, W. H. (2003). Ligand structure–function relationships in the vitamin D endocrine system from the perspective of drug development (including cancer treatment). *Recent Results Cancer Res.* **164**, 55–82.

Rachez, C., and Freedman, L. P. (2001). Mediator complexes and transcription. *Curr. Opin. Cell Biol.* **13**, 274–280.

Sutton, A. L., and MacDonald, P. N. (2003). Vitamin D: More than a "bone-a-fide" hormone. *Mol. Endocrinol.* **17**, 777–791.

BIOGRAPHY

Paul MacDonald and Diane Dowd are Faculty Members in the Department of Pharmacology at Case Western Reserve University. They are actively involved in basic research on the molecular, cellular, and physiological roles of the vitamin D receptor in normal and pathological states. Dr. MacDonald holds a Ph.D. in biochemistry from Vanderbilt University and received his postdoctoral training at the University of Arizona. Dr. Dowd received her Ph.D. in biochemistry from Vanderbilt University and her postdoctoral training at the Arizona Cancer Center and University of Arizona.

Vitamin D

Hector F. DeLuca and Margaret Clagett-Dame
University of Wisconsin, Madison, Wisconsin, USA

This article discusses the chemical identity and production of vitamin D. Its endocrine system, the nuclear receptor, and its mechanism of action are treated, along with the degradation of vitamin D hormone and its roles.

Chemical Identity, Natural Occurrence, and Production

The vitamin Ds are secosterols in which the B-ring of the steroid nucleus is replaced by a diene bridge. The two most common forms are vitamin D_2 (ergocalciferol) and vitamin D_3 (cholecalciferol). The structures of these two nutritionally important compounds are illustrated in Figure 1.

Although it is classified as a vitamin because of the nature of the discovery process, vitamin D in actual fact should not be considered a true vitamin for the following reasons. First and foremost, vitamin D utilized by higher organisms is manufactured in the epidermis of skin by photolysis of 7-dehydrocholesterol, a side reaction product of cholesterol synthesis. Secondly, vitamin D is nearly absent from the food supply for the most part. Vitamin D is found in fish liver oils and in egg yolk but is not found in virtually all plant materials, in skeletal meats, seeds, fruits, and vegetables. In fact, very little is found in an expected source, milk. We consider milk an important supply of vitamin D in many countries, primarily because it is fortified by the addition of vitamin D made artificially by man.

The chemical process whereby 7-dehydrocholesterol is converted to vitamin D_3 is well known. The 5,7-diene portion of either ergocalciferol or 7-dehydrocholesterol will absorb 280–310 nm ultraviolet light. This causes an isomerization to form a compound, previtamin D_3, which in itself is biologically inactive and remains in skin. Previtamin D is in equilibrium with vitamin D_3. The plant sterol, ergosterol, is converted to vitamin D_2 by an identical process. For many years, vitamin D_2 was the major synthetic form of vitamin D used in vitamin capsules, animal feeds, and in the fortification of foods. Because of improved synthetic methods, vitamin D_3 has become the major form used for both fortification and nutrition of domestic animals. The photolysis of 7-dehydrocholesterol is quite efficient in skin, and summer sun on hands and face and will in 10 minutes produce the daily requirement of vitamin D (currently believed to be 10–20 µg per day). Winter sun is unable to induce production of vitamin D because of the angle of the sunlight and the filtration of the critical wavelengths that cause vitamin D production. It is generally accepted that skin production of vitamin D_3 is quantitatively the most important source of the vitamin for human health.

Conversion of Vitamin D to its Hormonal Form

The vitamin D_3 that is produced in skin or that is absorbed from the intestine from vitamin pills or fortified foods is biologically inactive as such. The body must process vitamin D to form a final vitamin D hormone that is believed to carry out all of the known functions of vitamin D. Vitamin D_3 is first hydroxylated in the liver to form the circulating form of vitamin D_3, 25-hydroxyvitamin D_3 (25-OH-D_3). Interestingly, this compound is also biologically inactive and must be further modified before function. The second step occurs in the proximal convoluted tubule cells of the kidney where a 1 α-hydroxyl group is added to the molecule to produce the final vitamin D hormone, $1\alpha,25$-dihydroxy-vitamin D_3 ($1,25$-$(OH)_2D_3$). The reaction steps that result in this transformation are shown in Figure 2. The first step in conversion of vitamin D to the circulating form occurs in the hepatocytes of liver but not exclusively so. At least two enzymes are known to make this conversion, a mitochondrial enzyme and a microsomal enzyme, both of which have been cloned. No natural human mutation blocking this initial step has yet been discovered. However, the second step occurring in the proximal convoluted tubule cells of the kidney is a highly regulated one. This 1α-hydroxylase, also known as CYP27B1, has also been cloned and natural human mutants of this enzyme are now well known. These mutations result in a disease called vitamin D-dependency rickets type I in which children present
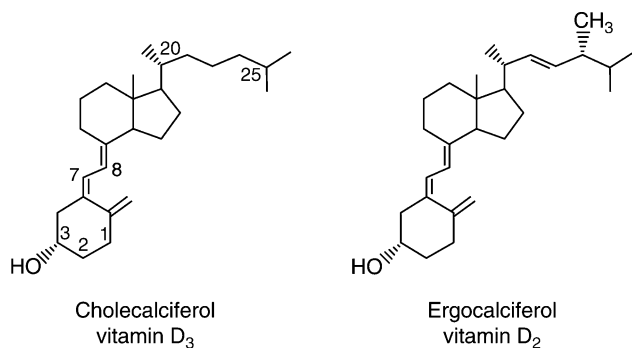
FIGURE 1 The nutritionally important forms of vitamin D. Vitamin D_3 is the form manufactured in skin and vitamin D_2 is the form produced by the irradiation of ergosterol. In mammals, both are equally active and both appear to be metabolized by almost identical routes.

vitamin D-deficiency diseases, i.e., rickets, hypocalcemia, and hypophosphatemia, even though they are naturally irradiated with sunlight or are provided vitamin D as a supplement. These children are cured by the administration of physiologic amounts of the natural vitamin D hormone, 1,25-$(OH)_2D_3$, establishing clearly the two-step activation process of vitamin D for man.

The Role of 1,25-$(OH)_2D_3$ in Calcium, Phosphorus, and Bone Metabolism

The basis for discovery of vitamin D in the first place is the disease rickets in children. This disease is also accompanied by hypocalcemia (low blood calcium) causing convulsive tetany. In adults, the deficiency of vitamin D causes the disease, osteomalacia. The diseases, rickets and osteomalacia, result from a failure to mineralize the organic matrix of newly synthesized bone. Early investigations to understand how vitamin D functions were, of course, directed toward the formation

of mineralized bone. However, it is now abundantly clear that 1,25-$(OH)_2D_3$ does not act directly on the mineralization process in the healing of rickets or osteomalacia but rather acts on intestine, kidney, and bone to provide calcium and phosphorus in the plasma required for the mineralization of skeleton and prevention of the convulsive disease, hypocalcemic tetany. The essence of vitamin D action, therefore, in curing rickets and osteomalacia is the elevation of plasma calcium and phosphorus to supersaturating levels that are required for normal mineralization. The vitamin D hormone acts to provide this calcium by stimulating the enterocyte of the small intestine to absorb calcium as well as phosphate by an independent mechanism both of which require active metabolic energy.

Because of the constant need for calcium to support normal neuromuscular function, calcium must be available in the plasma even when it is not available in the diet. Thus, the vitamin D hormone, i.e., 1,25-$(OH)_2D_3$ acting together with parathyroid hormone will cause the mobilization of calcium from the skeletal fluid compartment into the plasma compartment. Thus, the skeleton serves not only a structural function but also as a source of calcium and phosphorus. A final point is that 1,25-$(OH)_2D_3$ works in concert with the parathyroid hormone to induce the distal nephron of the kidney to reabsorb the last 1% of the filtered load of calcium into the plasma compartment. These then work to maintain plasma calcium and phosphorus at levels that support both bone formation and prevent hypocalcemic tetany. A diagram representing these functions is shown in Figure 3.

The Vitamin D Endocrine System

One of the most elegantly regulated substances in the body is the plasma calcium level. The parathyroid glands continuously monitor serum calcium levels and a calcium receptor therein responds when calcium falls even slightly below the normal level. By a G protein
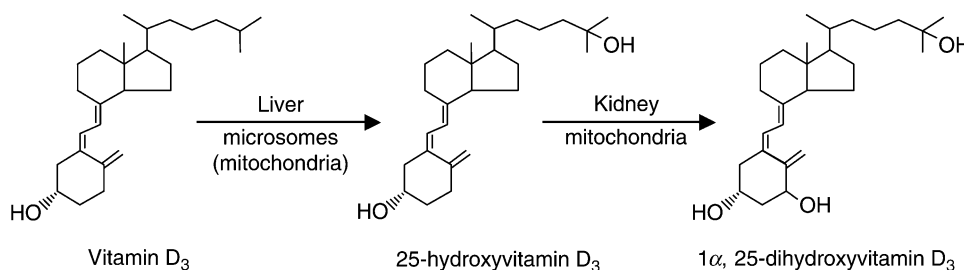
FIGURE 2 The required metabolism of vitamin D for function. Shown here is the metabolism of vitamin D_3 to the major blood form of vitamin D, 25-hydroxyvitamin D3 and its subsequent metabolism in the kidney to form the final vitamin D hormone, 1α,25-dihydroxyvitamin D_3. It is this form of vitamin D_3 that carries out all of the known functions of the vitamin. Vitamin D_2 not shown here is metabolized in an almost identical fashion to its active form, 1α,25-dihydroxyvitamin D_2.
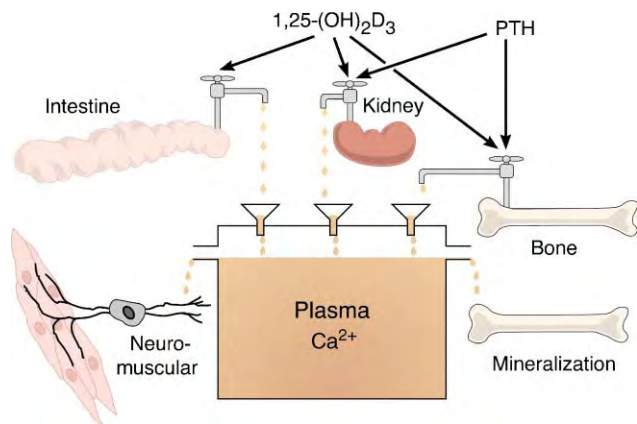
FIGURE 3 A diagrammatic representation of how the vitamin D hormone $1\alpha,25\text{-}(OH)_2D_3$) functions to maintain plasma calcium at supersaturating levels which are required for prevention of hypocalcemic tetany (neuromuscular) and for the prevention of rickets and osteomalacia (mineralization of bone).

mechanism, the calcium receptor stimulates the secretion of parathyroid hormone, a peptide hormone, that is the body's signal requesting calcium. It binds to the entire nephron of the kidney and to the osteoblasts of bone. Among its many functions in the kidney is to stimulate the 1α-hydroxylase (CYP27B1) to produce the vitamin D hormone in substantial quantities. The vitamin D hormone then acts on the bone, kidney, and intestine as described earlier to mobilize calcium as needed to raise blood calcium into the normal range, clearing the set-point of the parathyroids shutting down parathyroid secretion. Overshoot of calcium results in the C-cells of the thyroid secreting calcitonin, a 32 amino acid peptide hormone that blocks calcium mobilization from the skeleton. Calcitonin also has a stimulatory effect on the 1α-hydroxylase to produce the vitamin D hormone in small amounts to support its noncalcemic activities.

The Vitamin D Receptor

$1,25\text{-}(OH)_2D_3$ is a true steroid hormone that acts through a nuclear receptor called the vitamin D receptor (VDR). The human receptor is a 427 amino acid protein that serves as a ligand-dependent transcription factor. It is a member of the nuclear receptor superfamily of receptors. Other members of this superfamily include the retinoic acid, thyroid, estrogen, testosterone, gluco-corticoid, and other steroid hormone receptors. VDR is the smallest of the steroid hormone receptors and largely acts in a fashion quite similar to other steroid hormone receptors. Unlike other nuclear receptors, there is only a single VDR in all tissues. No subtypes are known and it is this receptor through which all vitamin D actions

occur. Natural mutants of the human VDR are known and produce a disease called vitamin D-dependency rickets type II. In this disease, high blood levels of $1,25\text{-}(OH)_2D_3$ occur in the face of severe rickets and severe hypocalcemia and hypophosphatemia. VDR null mutant mice have been produced and are available as experimental tools. The vitamin D-dependency rickets type II patients present a variety of diseases depending upon where the mutation has occurred. If there is a complete absence of a functional receptor, then very brittle vitamin D-resistant rickets occurs and can only be treated by the infusion of calcium and phosphorus into the blood stream to mineralize the skeleton but many of the other functions of vitamin D go unmet. The fact that VDR null mutant mice are normal at birth shows that vitamin D is not an embryonic developmental factor but functions primarily after parturition.

There are reports of nongenomic functions of vitamin D. These have been largely cellular-based observations, which have yet to be shown to occur *in vivo*. In fact, the administration of large amounts of $1,25\text{-}(OH)_2D_3$ to receptor null mutant mice produces no significant phenotype, suggesting that the only mechanism of action of vitamin D is through its receptor under physiologic circumstances.

Molecular Mechanism

Vitamin D-response elements (VDE) are found in the promoter region of target genes. These are hexameric repeats separated by three nonspecified bases. There can be as many as three repeat elements in a responsive element system. The VDR must heterodimerize with another transcription factor, the retinoid X receptor (RXR), in order to bind to the responsive elements. The RXR binds to the 5' repeat and the VDR to the 3' repeat. Figure 4 provides a cartoon to illustrate what is currently known about the mechanism of action of the vitamin D hormone working through its receptor to regulate expression or suppression of target genes. As with other steroid hormones, it is anticipated that the unliganded receptor may be retained in an inactive state by being bound to a corepressor. This has not been specifically shown for the VDR but is known for other class II nuclear receptors. There is increasing evidence that such a corepressor might be identified for the VDR. In this model, the corepressor that interacts directly with the VDR would be released upon binding of the ligand to the receptor probably because of a change in VDR conformation. The receptor would then be free to interact with a coactivator protein that is believed to be involved in chromatin remodeling and to link the complex bound at the hormone response element to the basal transcription machinery. Among the known coactivators believed to interact directly with the VDR
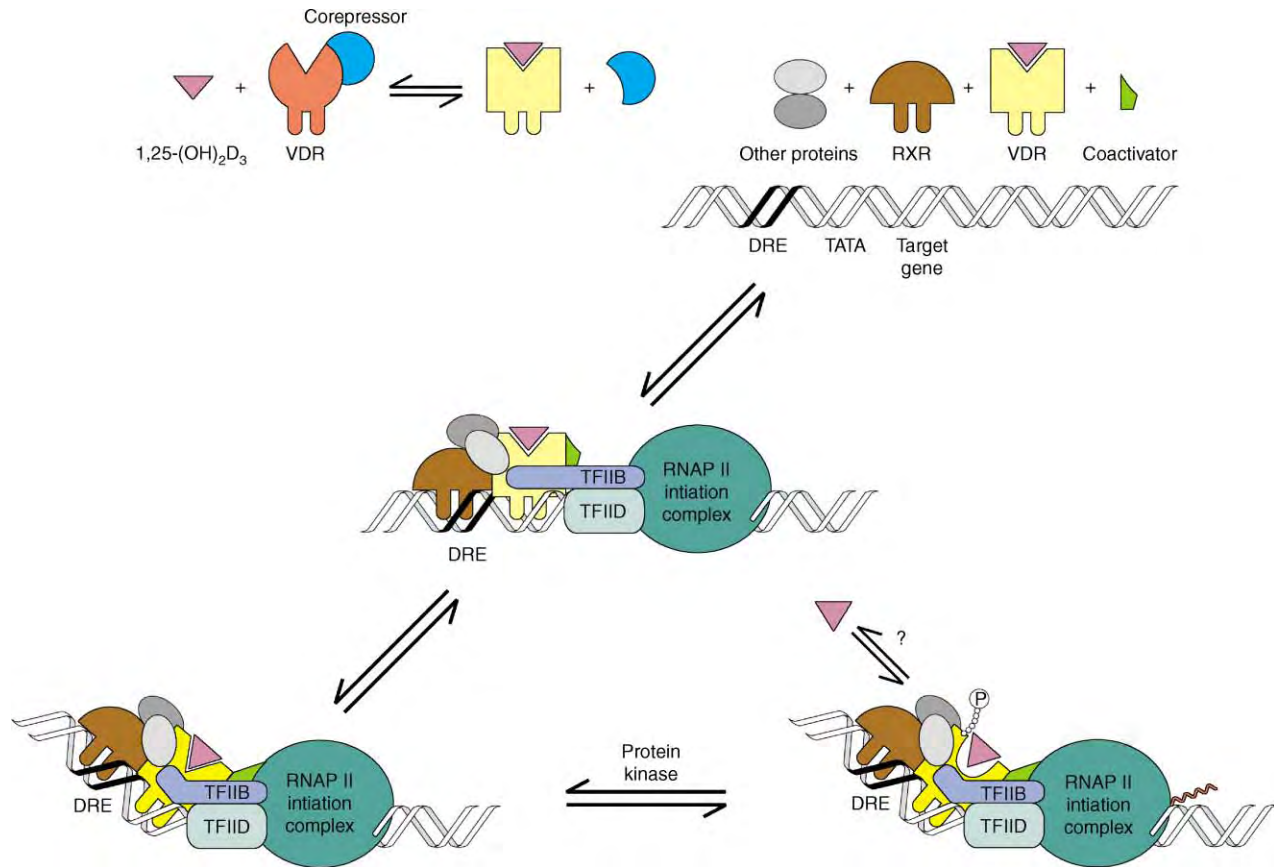
FIGURE 4 A cartoon representing the molecular mechanism of action of the vitamin D hormone ($1,25$-$(OH)_2D_3$) in regulating transcription of target genes in the target tissues. Binding of the ligand, i.e., $1,25$-$(OH)_2D_3$ to the VDR causes a conformational change and a rejection of the corepresssor. It also allows binding of a coactivator and formation of a heterodimer with RXR on the vitamin D responsive elements. Several proteins are required to form the transcription complex, all of which have not yet been identified. There is a bending of the DNA and a phosphorylation on serine 205 but their exact role in either suppression or activation of transcription remains unknown.

are SRC 1, 2, 3, and DRIP 205. Likely other participating coactivators will be discovered. A phosphorylation does occur on serine 205 of the VDR but whether this is of functional importance remains to be determined.

Metabolism and Degradation

The vitamin D hormone is degraded specifically by an enzyme it induces. That enzyme is the CYP24 (24-hydroxylase), which carries out a series of reactions on the side chain of the vitamin D hormone and its precursor, 25-OH-D_3. It first causes hydroxylation in the 24R-position followed by oxidation of that same hydroxyl to a ketone. The next step again carried out by the same enzyme is a 23-hydroxylation followed by cleavage and oxidation giving rise to the biliary excretion product called calcitroic acid. Calcitroic acid is biologically inactive. Thus, the vitamin D hormone, an extremely potent biological substance, ensures its limited activity by inducing its own destruction.

Other metabolism of vitamin D is known including the formation of the 26,23-lactone derivative, 26-hydroxylation, but quantitatively the most important reaction regulating the level of vitamin D hormone is the CYP24. The CYP24 is regulated in a negative fashion by the parathyroid hormone causing an instability of the mRNA encoding for that enzyme. Thus, the parathyroid hormone not only stimulates the 1 α-hydroxylase but also stimulates the destruction of the enzyme that destroys the hormone, thereby ensuring that large amounts of the vitamin D hormone are available to carry out its calcium functions.

New and Nonclassical Roles of the Vitamin D Hormone

With the discovery of the VDR came the observation that it is found in tissues not previously appreciated as target tissues of vitamin D action. The discovery of

the VDR in the intestine, kidney tubules, and osteoblasts was not unexpected. However, its presence in the parathyroid gland, the islet cells of the pancreas, the immune cells such as T-cells, macrophages, and keratinocytes made it possible that the vitamin D hormone had important functions besides raising blood calcium and phosphorus and mineralizing the skeleton. It is now clear that an important function of the vitamin D hormone is to suppress the growth of parathyroid glands and to suppress production of the parathyroid hormone. This basic function has been utilized by physicians using $1,25\text{-}(OH)_2D_3$ and its analogues to treat secondary hyperparathyroidism in kidney failure patients.

Another important VDR site are the cells of the immune system and in the promyelocytes which are the precursors of the monocyte. The vitamin D hormone stimulates the formation of the monocyte by causing the production of RANKL by osteoblasts and stromal cells. This results in a coalescence of the monocytes to form giant osteoclasts which are further activated by RANKL. Thus, the formation of the giant osteoclast that causes bone turnover is a vitamin D function. Furthermore, vitamin D also functions to cause the formation of new bone, thus playing an important role in the bone remodeling and modeling systems.

The role of the vitamin D hormone in the keratinocyte has also become of importance therapeutically. The vitamin D hormone added to the keratinocytes in cultures or added topically will cause a cessation of proliferation of the keratinocyte and will cause its differentiation. Thus, $1,25\text{-}(OH)_2D_3$ and analogues have been successfully used in the topical treatment of psoriasis.

Finally, attention must be focused on the immune system where the vitamin D hormone is now known to cause important immuno-modulatory effects. The vitamin D hormone can be used to block certain animal models of autoimmune disease. For example, it can be used to prevent experimental autoimmune encephalomyelitis (EAE), a model of multiple sclerosis in mice. However, in so doing, it does induce an abnormal rise in serum calcium. $1,25\text{-}(OH)_2D_3$ can be used to prevent diabetes in the nonobese diabetic mouse model of type 1 diabetes. It can be used to prevent rheumatoid arthritis, systemic lupus, and inflammatory bowel disease. Thus, the immuno-modulatory role of the vitamin D hormone under normal circumstances is only now becoming realized.

Therapeutic Uses of $1\alpha,25\text{-}(OH)_2D_3$ and its Analogues

As discussed earlier, besides the treatment of vitamin D-dependency rickets and vitamin D-resistant rickets of many types, the vitamin D hormone and its analogues are certainly central actors in the treatment of bone disease secondary to kidney failure. This is primarily by suppression of parathyroid proliferation and by suppression of parathyroid hormone production. $1,25\text{-}(OH)_2D_3$ and its analogue, $1\alpha\text{-}OH\text{-}D_3$ has been successfully used in certain parts of the world to treat postmenopausal and age-related osteoporosis, giving a modest rise in bone mass along with a clear reduction in fracture rate. The danger of hypercalcemia has limited its use in the treatment of this disease. However, newer vitamin D analogues with greater bone anabolic activity than $1,25\text{-}(OH)_2D_3$ are currently under development and may make it possible to safely treat osteoporosis in the future.

An analogue of $1,25\text{-}(OH)_2D_3$ called Dovonex is currently marketed for the topical treatment of the hyperproliferative skin disorder, psoriasis. Improved analogues will likely appear for treatment of psoriasis. There are potential new uses for the vitamin D analogues in the autoimmune diseases as described above or in the treatment of malignant cancer growth. The vitamin D hormone has been found to suppress growth of malignant cells *in vitro* and to cause their differentiation into functional cells. Thus, the use of vitamin D compounds as a potential treatment of metastatic neoplastic disease has been exciting but it has not yet yielded a therapeutic agent. The major problem with the use of $1,25\text{-}(OH)_2D_3$ and many of its analogues for treatment of disease is that its primary purpose is to raise blood calcium and phosphorus which means that is a major side effect that must be dealt with before it can be used to treat noncalcemic diseases. One approach to this is the synthesis of analogues that retain the ability to suppress cancerous growth but do not raise blood calcium. New analogues with these desirable characteristics are currently under basic study, but are not yet available for pharmaceutical use.

See Also the Following Articles

Calcium Buffering Proteins: Calbindin • Cholesterol Synthesis • Cytochrome P-450 • Lipid Rafts • Parathyroid Hormone/Parathyroid Hormone-Related Protein Receptor • Vitamin D Receptor

Glossary

coactivator   Protein(s) that enhance receptor activity in transcription.

convulsive tetany   A disease precipitated by either low blood magnesium or low blood calcium resulting in a severe convulsive state. It is lethal unless immediately treated with calcium and/or magnesium.

corepressor   A protein that silences transcription factors.

7-dehydrocholesterol   The sterol precursor of vitamin D_3.

Dovonex A drug marketed by Leo Pharmaceuticals containing an analogue of 1,25-$(OH)_2D_3$ and used to treat psoriasis s by topical administration.

hepatocytes Cells comprising $\sim 30\%$ of the liver that carry out many of the liver's functions.

keratinocytes Cells of the skin that produce the keratin proteins and contribute to the skin barrier.

macrophages Cells of the immune system that are responsible for destroying foreign materials.

nongenomic Referring to or describing activities that are carried out not involving gene expression.

osteoblasts Cells of the skeleton that carry out synthesis of new bone and that mediate regulation of bone activities.

previtamin D An isomer of vitamin D formed during the irradiation process for producing vitamin D. It is in equilibrium with vitamin D.

proximal convoluted tubule cells Cells lining the proximal convoluted tubule of the kidney.

RANKL *Receptor *a*ctivator of *NF-K*appaβ *li*gand, a signal protein produced by osteoblasts and stromal cells in response to such stimulants as 1,25-$(OH)_2D_3$ and parathyroid hormone.

transcription factor A protein usually found in the nucleus that has a marked effect or role in the transcription of genes.

FURTHER READING

Darwish, H. M., and DeLuca, H. F. (1996). Recent advances in the molecular biology of vitamin D action. In *Progress in Nucleic Acid Research and Molecular Biology* (W. E. Cohn and K. Moldave, eds.) pp. 321–344. Academic Press, San Diego, CA.

DeLuca, H. F. (1974). Vitamin D: The vitamin and the hormone. *Fed. Proc.* **33**, 2211–2219.

DeLuca, H. F. (1997). Historical overview. In *Vitamin D* (D. Feldman, F. H. Glorieux and J. W. Pike, eds.) Vol 1, pp. 3–11. Academic Press, San Diego, CA.

Feldman, D., Glorieux, F. H., and Pike, J. W. (1997). *Vitamin D*. Academic Press, San Diego, CA, 1285pp.

Jones, G., Strugnell, S. A., and DeLuca, H. F. (1998). Current understanding of the molecular actions of vitamin D. *Physiol. Rev.* **78**, 1193–1231.

BIOGRAPHY

Hector F. DeLuca is Steenbock Research Professor and Chairman of the Department of Biochemistry at the University of Wisconsin-Madison. He has been a major contributor to the current understanding of the metabolism and mechanism of action of vitamin D. He has published over 2000 papers and patents. He has trained well over 250 students in vitamin D, vitamin A, bone and calcium fields. He is a member of the National Academy of Sciences, the American Academy of Arts and Sciences, a Fellow of the AAAS and has won many awards for his work on vitamin D.

Margaret Clagett-Dame is a Professor of Biochemistry and Pharmaceutical Sciences in the Department of Biochemistry at the University of Wisconsin-Madison and is best known for her work in the developmental role of the retinoids. She has published some 48 papers and 6 patents.

Vitamin E

Ute C. Obermuller-Jevic
BASF Aktiengesellschaft, Ludwigshafen, Germany

Lester Packer
University of Southern California, Los Angeles, California, USA

Vitamine E is the name for a group of biologically active substances including tocopherols and tocotrienols. Alpha-tocopherol shows the highest biological activity of all vitamin E forms and is the most common form found in the human body. Vitamin E was discovered in 1922 as essential micronutrient for reproduction. In the 1950s it was recognized that vitamin E is the body's major lipid-soluble antioxidant protecting lipoproteins and membranes against oxidative damage. The biological role of vitamin E goes, however, beyond its antioxidant function. Studies on the effects of vitamin E on gene expression have revealed many genes and signal transduction pathways which are influenced by vitamin E. All these actions of vitamin E help to explain the observed beneficial effects of vitamin E in chronic and degenerative diseases. A well-balanced diet rich in vitamin E seems a basic requirement for human health. Intake of supplementary vitamin E may be advisable for individuals who do not get adequate levels from diet, or who are at high risk for, or who are already suffering from, chronic and degenerative diseases.

Molecular Structure of Vitamin E

Vitamin E is the generic name for a group of eight plant-derived, lipid-soluble substances (tocols) including four tocopherols and four tocotrienols (Figure 1). The molecular structure of vitamin E is comprised of a chromanol ring with a side chain located at the C-2 position. Tocopherols have a saturated phytyl side chain and tocotrienols have an unsaturated isoprenoid side chain. The number and position of methyl groups located around the chromanol ring varies among the different tocopherols and tocotrienols, and accounts for the designation as alpha-, beta-, gamma-, or delta-forms. Natural and natural-source alpha-tocopherol occur in *RRR*-configuration only (formerly designated as d-alpha-tocopherol). In contrast, synthetic alpha-tocopherol consists of an equal racemic mixture of eight stereoisomers (*RRR, RSR, RRS, RSS, SRR, SSR, SRS, SSS*) arising from the three chiral centers of the molecule

at positions C2, C4′ and C8′ and designated as *all-rac*-alpha-tocopherol (or dl-alpha-tocopherol).

Biological Activity of Vitamin E

The various forms of vitamin E differ in their biological activities. Alpha-tocopherol is the most common form of vitamin E occurring in human blood and tissues and it has the highest biological activity among all tocopherols and tocotrienols. Moreover, the human body preferentially accumulates the *RRR*- or 2*R*-forms of vitamin E. Thus natural *RRR*-alpha-tocopherol has a higher bioavailability and "biological activity" than synthetic *all rac*-alpha-tocopherol. According to the Food and Nutrition Board (FNB) of the US National Academy of Sciences, which publishes recommendations for dietary intake of vitamin E, only *RRR*-alpha-tocopherol itself, or the *RRR*- and 2*R*-stereoisomers of *all rac*-alpha-tocopherol meet the vitamin E requirements in humans. The FNB suggests a 2-fold higher potency of natural versus synthetic alpha-tocopherol sources, however, this value has been challenged. The biological activity of vitamin E in dietary supplements is usually expressed as international units (IU). As published in the United States Pharmacopeia (USP), 1 IU is defined as the biological activity of 1 mg *all rac*-alpha-tocopheryl acetate. Other biological activities of vitamin E are 1 mg *all rac*-alpha-tocopherol = 1.1 IU; 1 mg *RRR*-alpha-tocopherol = 1.49 IU; 1 mg gamma-tocopherol = 0.15 IU.

Dietary Sources and Recommended Intake of Vitamin E

Vegetable oils and lipid-rich plant products (e.g., nuts, seeds, grains) are the main dietary sources of vitamin E. In Western diets, vitamin E intake mainly derives from fats and oils as found in margarine, mayonnaise, salad dressing, and desserts, and increasingly also from fortified food (e.g., breakfast cereals, milk, fruit juices).
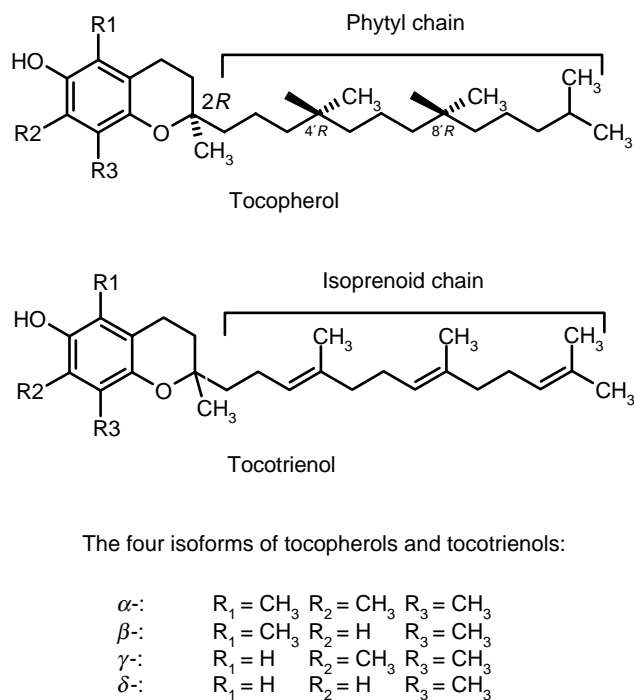
FIGURE 1 The molecular structure of vitamin E. Tocopherols and Tocotrienols exist in four different isoforms (α-, β-, γ-, and δ-forms). Tocopherol is shown in its naturally occurring *RRR*-configuration. (Adapted from Food and Nutrition Board, 2000.)

Noteworthy, the diet of Americans contains large amounts of gamma-tocopherol when compared to populations in other Western countries, which is a result of the high consumption of soybean and corn oil containing more gamma- than alpha-tocopherol. Vitamin E used for food fortification or dietary supplements consists mainly of alpha-tocopherol, either derived from natural sources (i.e., methylated gamma-tocopherol from vegetable oil) or from synthetic production and it is usually esterified to increase stability.

The FNB recommends for adults a daily intake (recommended dietary allowance, or RDA) of 15 mg vitamin E including food, fortified food, and supplements. The recommendations are largely based on *in vitro* studies of M. Horwitt dating back to 1960, which are still considered the most adequate data for defining the physiological status and a health benefit of vitamin E in humans. In these studies, prevention of H_2O_2-induced erythrocyte hemolysis was used as test system. However, except when vitamin E is clearly deficient the erythrocyte hemolysis test is not a useful indicator of vitamin E functional status. Thus, the guidelines of the FNB were critically discussed in the literature and it was strongly agreed upon that other assays and new biomarkers need to be identified to define the physiologic role and beneficial potential of vitamin E in humans.

The recommended daily intake of 15 mg vitamin E is considered rather unlikely to be achieved by North Americans and other Western countries through diet alone. Although universal dietary supplementation has not been recommended, a dose of 200 mg/day vitamin E (alpha-tocopherol) has been suggested to saturate plasma levels and elevate tissue levels which may have possible health effects in the long term.

As an upper limit for supplemental vitamin E intake the FNB has published for adults a dose of 1 g/day alpha-tocopherol (i.e., 1 500 IU *RRR*- or 1 100 IU *all rac*-alpha-tocopherol) showing no apparent side effects, which is considered safe. In human intervention studies, various doses of vitamin E up to 3 600 IU/day have been used. Nevertheless, conclusive evidence from long-term studies regarding biological effects and safety of chronic intake of pharmacologic doses of vitamin E are lacking. In the Alpha-Tocopherol, Beta Carotene (ATBC) Cancer Prevention Study with Finnish smokers consuming 50 IU/day vitamin E for 6 years, an increase in mortality from hemorrhagic stroke was observed, however, other intervention studies did not report such an adverse effect. It was suggested that pharmacologic doses of vitamin E are contraindicated in persons with blood coagulation disorders as vitamin E might exacerbate defects in the blood coagulation system based on its inhibitory effects on platelet aggregation.

Uptake, Distribution, and Metabolism of Vitamin E

Together with dietary fat, alpha-tocopherol and all other forms of vitamin E are absorbed in the digestive tract, incorporated into chylomicrons and transported in the lymphatic system. Part of the absorbed stereoisomers are taken up into extrahepatic tissues by the action of lipoprotein lipase and the remainder is delivered in chylomicron remnants to the liver. Distribution of vitamin E into the circulation is regulated by a cytosolic alpha-tocopherol transfer protein (alpha-TTP) in the liver, which is selective for alpha-tocopherol in its *RRR*- or 2*R*-forms. The other stereoisomers of alpha-tocopherol as well as the other tocopherols and tocotrienols exert much less affinity for alpha-TTP. Alpha-TTP plays an important role in maintaining plasma levels of vitamin E.

In the liver, vitamin E is incorporated into very low density lipoproteins (VLDLs) and released to the systemic blood circulation. Excess amounts of alpha-tocopherol along with the other absorbed forms of tocopherols and tocotrienols are metabolized or eliminated by the biliary tract. VLDLs are converted into low-density lipoproteins (LDLs) and excess surface components including alpha-tocopherol are transferred

to high-density lipoproteins (HDLs). Delivery of alpha-tocopherol to peripheral tissues takes place via binding of LDL to LDL receptors and subsequent cellular uptake. Vitamin E preferably accumulates in adipose tissues. Cytosolic tocopherol-associated proteins (TAP) have been reported showing alpha-tocopherol-specific binding characteristics and also nuclear translocation and transcriptional activation in various mammalian cell types and organs. TAP seems to play an important role in vitamin E-induced gene expression, however, its biological functions are not widely known. A further cytosolic vitamin E regulatory protein, i.e., tocopherol-binding protein (TBP) has been reported whose functions are not fully understood.

The lowest acceptable vitamin E level in plasma is 11.6 μmol L^{-1} (0.5 mg dL^{-1}) and a ratio of vitamin E:cholesterol of 2.25 μmol mmol^{-1}. Serum levels of vitamin E correlate with cholesterol levels and hence do not necessarily correlate with vitamin E intake. Except for non- or poorly responding subjects, serum levels of vitamin E can usually be increased up to threefold by intake of dietary supplements reaching a saturation level.

Biological Importance of Vitamin E in Reproduction and Essentiality

Vitamin E was discovered in 1922 by H. Evans and K. Bishop as essential micronutrient needed to ensure normal reproduction in rats. It was named according to a consecutive alphabetical order preceded by the discovery of vitamins A to D. Later vitamin E was called alpha-tocopherol, according to the greek term "tokos" child-birth, "phero" to bear, and -ol, indicating an alcohol. Rats on a diet low in vitamin E showed reduced fertility and a high rate of fetal resorption. However, when animals were fed a lipophilic fraction from lettuce or, as later shown, wheat germ oil, their fertility was retained and a successful implantation of the fetus was observed. The biological activity of vitamin E is based on this so-called rat resorption-gestation assay. The family of natural vitamin E molecules as well as the stereoisomers of *all rac*-alpha-tocopherol all exhibit to varying degrees vitamin E activity in this bioassay. Unfortunately, this assay of reproductive activity in pregnant rats may bear limited relevance to human health.

Unique Role of Vitamin E as a Lipophilic Antioxidant in Lipoproteins and Cell Membranes

In the 1950s it was recognized under leadership of A. L. Tappel's group that vitamin E is the body's major lipid soluble antioxidant protecting lipoproteins and membranes where it resides against free radical-mediated lipid peroxidation which, if not prevented or interrupted by vitamin E, causes widespread oxidative molecular damage and pathology. All natural isoforms and synthetic stereoisomers of vitamin E exhibit to varying degrees the ability to inhibit lipid peroxidation as a "chain-breaking" antioxidant. Vitamin E primarily destroys peroxyl radicals and thus protects unsaturated fatty acids from oxidation. Additionally, vitamin E scavenges a variety of oxygen-derived free radicals including alkoxyl radicals, superoxide, and other reactive oxygen species (ROS) such as singlet oxygen and ozone, and it reacts with nitrogen species. Vitamin E participates in an *antioxidant network* (a concept advanced by L. Packer's group) and thus vitamin E radicals can be recycled or regenerated back to their native form, e.g., by vitamin C. The antioxidant network is strengthened by bioflavonoids (recycle vitamin C) and by carotenoids (free radical traps sparing vitamin E) (Figure 2).

Effects of Vitamin E on Cell Signaling and Gene Expression

In the early 1990s, inhibition of protein kinase C (PKC) activity by vitamin E was suggested by A. Azzi's group as the crucial factor for inhibition of cell proliferation in smooth muscle cells. PKC activity is an important factor contributing to disorders such as vascular disease, cancer, diabetes, and other age-related degenerative diseases. Vitamin E was found to inhibit PKC activity in many cell types including smooth muscle cells, monocytes, macrophages, neutrophils, fibroblasts, and mesangial cells and the effects were repeatedly confirmed in animal studies. Inhibition of PKC activity by vitamin E occurs indirectly via activation of a phosphatase that cleaves the active, phosphorylated form of PKC, or by modulating diacylglycerol kinase activity. What is novel, is that inhibition of PKC is apparently independent of antioxidant activity and rather is the result of inhibition due to the specific molecular structure of the RRR stereoisomer of alpha tocpopherol. Hence the biological role of vitamin E goes beyond its antioxidant function.

Recent advances in molecular biology and availability of microarray techniques for studying effects of vitamin E on gene expression have revealed novel vitamin E-sensitive genes and signal transduction pathways. Vitamin E regulates at the transcriptional level the expression of several genes including collagen alpha-1 and alpha-TTP in liver, collagenase in skin, adhesion molecules, and chemokines such as VCAM-1 and MCP-1 in endothelial cells, different integrins in erythroleukemia
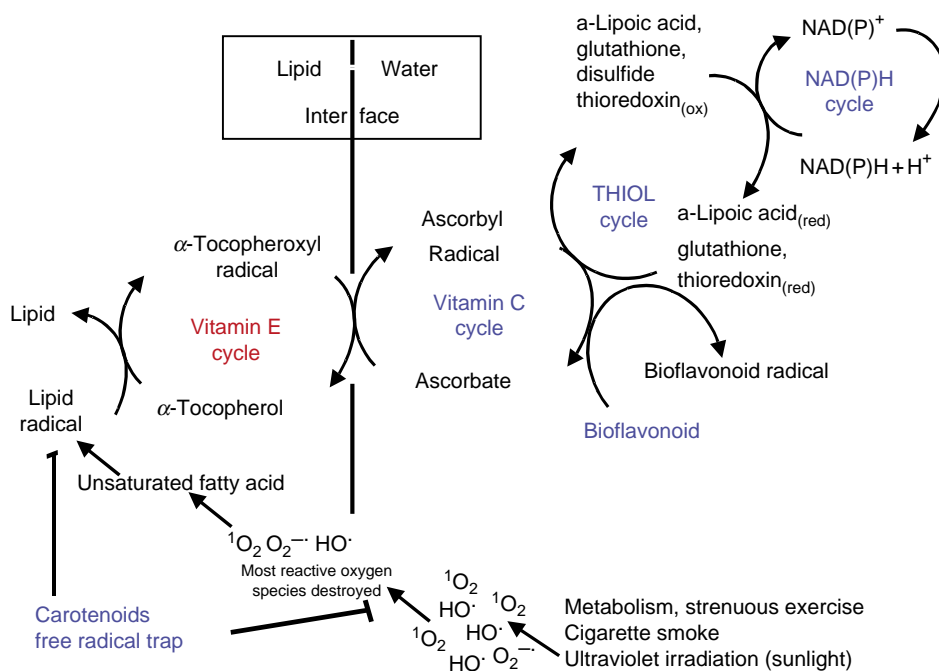
FIGURE 2 The antioxidant network concept.

cells, alpha-tropomyosin in smooth muscle cells, and scavenger receptors class A (SR-A) and CD 36 in macrophages and smooth muscle cells. At the posttranslational level, vitamin E regulates the expression of cyclooxygenase in monocytes leading to a decrease in prostaglandin E_2 levels. As this latter effect was also observed using other vitamin E homologues compatible with antioxidant activity this function of vitamin E may involve redox-signaling. In alpha-TTP knockout mice many genes associated with functions of liver or brain are remarkably modulated by deficiency in vitamin E, whereas the majority of genes over a broad range of expression levels show no change. This indicates that vitamin E regulates tissue-specific gene expression. These newly discovered actions of vitamin E help to explain the observed beneficial effects of vitamin E in chronic and degenerative diseases of aging.

Recognition of Vitamin E Deficiency Diseases and the Beneficial Effects of Vitamin E Supplements in Human Health and Disease Prevention

Clinically manifest vitamin E deficiency in humans is rare. Clear indications of vitamin E deficiency have been found in premature babies and in children suffering from an abnormal ability to absorb vitamin E as in abetalipoproteinemia, chronic cholestatic liver disease, cystic fibrosis, and in short-bowel syndrome.

A mutation in the alpha-TTP gene as found in "familial isolated vitamin E deficiency" results in a failure to deliver alpha-tocopherol to the systemic circulation. In these cases clinical vitamin E deficiency is recognized in babies showing symptoms of retrolental fiblasia and intraventricular hemorrhage, and in children suffering from muscular dystrophy, ataxia, and other disorders. High-dose vitamin E supplementation can reduce or eliminate clinical symptoms in these patients, however, to achieve amelioration of neurological symptoms early diagnosis and early start of treatment are crucial.

More frequently, a chronic sub-optimal supply in vitamin E (i.e., theoretically an intake below RDA levels) occurs population-wide and in all age groups. This may cause impaired defense against oxidative stress and increased susceptibility to oxidative injury and adverse health effects. In most diseases that have been examined with any degree of scrutiny, evidence for "oxidative stress" and oxidative damage has been observed in some stage of disease initiation or progression. Therefore it seems obvious that vitamin E like other antioxidants may prevent or delay disease progression. In numerous epidemiologic studies it has been observed that high intake of vitamin E is associated with a lower risk of age-related and chronic diseases, and experimental studies have suggested substantial

health benefits from vitamin E in disease prevention and therapy. Most large-scale human intervention trials on the prevention of cardiovascular disease, the main cause of morbidity and mortality in the Western world, however have failed to convincingly show that (relatively short-term) supplementation with vitamin E lowers the incidence of cardiovascular events or the rate of mortality from heart disease or stroke. Interestingly, a plethora of small clinical trials report a significant improvement of health status and/or retardation of disease progression by vitamin E supplementation in patients suffering from cardiovascular disease or other chronic and age-related diseases. Vitamin E appears to have important health benefits such as skin health, in reproductive diseases (pre-eclampsia), age-related eye diseases (cataract, AMD), metabolic disorders (Diabetes mellitus), neurodegenerative disorders (Alzheimer's and Parkinson's disease), skin aging (photoaging), and healthy aging.

Conclusion

An adequate dietary intake of vitamin E seems of crucial importance to guarantee throughout life normal function of physiologic processes in the body as well as adequate defense against oxidants generated by endogenous sources or from exposure to environmental stress. Vitamin E plays a unique role in the prevention of chronic and degenerative diseases, and thus in healthy aging. A well-balanced diet rich in vitamin E as part of a healthy lifestyle seems a basic requirement for human health. Moreover, intake of supplementary vitamin E may be advisable for individuals who are unable to sustain adequate levels, or who are at high risk for, or who are already suffering from chronic and degenerative diseases.

SEE ALSO THE FOLLOWING ARTICLES

Lipoproteins, HDL/LDL • Protein Kinase C Family • Vitamin C

GLOSSARY

α-tocopherol-binding protein A protein found mainly in liver, which discriminates between the different forms of vitamin E. It preferentially binds to RRR alpha-tocopherol leading to mainly retention of this vitamin E form in the human body whereas the other forms of vitamin E are mainly metabolized and excreted.

RRR-α-tocopherol The only stereoisomer of alpha-tocopherol occurring in nature. When alpha-tocopherol is made by chemical synthesis, it contains not only the RRR-stereoisomer, but a racemic mixture (called all-rac-alpha-tocopherol) of eight stereoisomers. Synthetic alpha-tocopherol is mainly used in dietary supplements and in fortified foods.

vitamin E The name for a group of eight substances of plant origin including four tocopherols and tocotrienols. Alpha-tocopherol is the most prominent form of vitamin E, which is considered essential in humans and animals. Besides its essentiality, alpha-tocopherol as well as the other forms of vitamin E are the major antioxidants in blood lipids and in cell membranes.

FURTHER READING

Brigelius-Flohe, R., and Traber, M. G. (1999). Vitamin E: Function and metabolism. *FASEB J.* **13**, 1145–1155.

Food and Nutrition Board (2000). *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids*, pp. 186–283. National Academy Press, Washington, DC.

Gordon, N. (2001). Hereditary vitamin-E deficiency. *Dev. Med. Child Neurol.* **43**, 133–135.

Hoppe, P., and Krennrich, G. (2000). Bioavailability and potency of natural-source and all-racemic alpha-tocopherol in the human: A dispute. *Eur. J. Nutr.* **39**, 183–193.

Packer, L., and Obermüller-Jevic, U. (2002). Vitamin E in disease prevention and therapy: Future perspectives. In *The Antioxidant Vitamins C and E* (L. Packer, M. G. Traber, K. Krämer and B. Frei, eds.) AOCS Press, Champaign.

Ricciarelli, R., Zingg, J. M., and Azzi, A. (2002). The 80th anniversary of vitamin E: Beyond its antioxidant properties. *Biol. Chem.* **383**, 457–465.

Rimbach, G., Minihane, A., Majewicz, J., Fischer, A., Pallauf, J., Virgli, F., and Weinberg, P. (2002). Regulation of cell signaling by vitamin E. *Prog. Nutr. Soc.* **61**, 415–425.

BIOGRAPHY

Ute Obermüller-Jevic Ph.D. is a Nutrition Scientist holding a doctoral degree in the field of antioxidant nutrients. Currently she is a scientific affairs manager for BASF Aktiengesellschaft, Ludwigshafen, focusing on the health benefits of vitamins and carotenoids.

Lester Packer received his Ph.D. from Yale University in 1956 and in 1961 joined the University of California at Berkeley where he was Professor of Molecular and Cell Biology and since 2000 is Adjunct Professor in the Department of Molecular Pharmacology and Toxicology, University of Southern California Health Sciences Center. Dr. Packer is a leading researcher on Antioxidants who has published over 800 articles, edited more than 90 books, is a member of numerous professional societies and editorial boards. He was Founder and Honorary President of The Oxygen Club of California, Past President of the Society for Free Radical Research International: he has received numerous awards including three Honorary Doctoral degrees.

Vitamin K: Biochemistry, Metabolism, and Nutritional Aspects

J. W. Suttie
University of Wisconsin-Madison, Wisconsin, USA

The term vitamin K refers to a number of compounds containing a 2-methyl-1,4-naphthoquinone ring with a hydrophobic side chain constituted at the 3-position (Figure 1). The biochemical role of the vitamin was not established until the mid-1970s, when it was shown that it was a substrate for a microsomal enzyme that converted protein bound glutamyl residues to γ-carboxyglutamyl (Gla) residues. This posttranslational modification has been found in relatively few proteins.

Vitamin K-Dependent Proteins

Prothrombin (clotting factor II), the zymogen of the plasma procoagulant thrombin, was the first protein shown to be dependent to vitamin K for its synthesis, and was also the first protein demonstrated to contain γ-carboxylglutamyl (Gla) residues. Plasma clotting factors VII, IX, and X were all initially identified in patients with hereditary bleeding disorders and were subsequently shown to be vitamin K dependent. Until the mid-1970s, these four vitamin K-dependent clotting factors were the only proteins known to require this vitamin for their synthesis. The amino-terminal, Gla domains of these four vitamin K-dependent procoagulants are very homologous, and the 10–13 Gla residues in each are in essentially the same position as in prothrombin.

Following the discovery of Gla, three more Gla-containing plasma proteins with similar homology were discovered. Protein C and protein S are involved in an anticoagulant rather than a procoagulant role in normal hemostasis, and the seventh Gla-containing plasma protein (protein Z) also has an anticoagulant function under some conditions. As these proteins play a critical role in hemostasis, they have been extensively studied, and the cDNA and genomic organization of each of them is well documented.

The discovery of Gla residues in the plasma vitamin K-dependent proteins led to a search for other proteins with this modification, and a 49 residue protein that contained three Gla residues, called osteocalcin (OC), was isolated from bone. It has little structural homology to the vitamin K-dependent plasma proteins. Although it is the second most abundant protein in bone, its function is not clearly defined. A second low-molecular-weight (79 residue) protein, matrix Gla protein (MGP), contains five Gla residues and was also first isolated from bone, but is also synthesized in cartilage and many other soft tissues. The function of these two proteins has not been clearly defined, but it has been shown that osteocalcin gene knockout mice develop more dense bones, and MGP knockout mice die from spontaneous calcification of arteries and cartilage. A limited number of other mammalian proteins have been found to contain Gla residues: Gas 6, a ligand for the tyrosine kinase Ax1, two proline rich integral membrane Gla proteins (PRGP-1, PRGP-2), and two additional Gla proteins (TMG-3 and TMG-4). A large number of Gla-containing toxic venom peptides are secreted by marine Conus snails and are found in some snake venoms. The vitamin K-dependent carboxylase has been cloned from a number of vertebrates, the Conus snail, a tunicate, and *Drosophila*, indicating the ancient evolutionary origin of this posttranslational modification.

Biochemical Role of Vitamin K

Beginning in the early 1960s, studies of prothrombin production in humans and experimental animals led to an understanding of the metabolic role of vitamin K. A biologically inactive form of prothrombin was demonstrated to be present in the plasma of patients treated with oral anticoagulants, and studies with hypoprothrombinemic rats were consistent with the presence of a hepatic precursor protein pool that was rapidly being synthesized and which could be converted to prothrombin by a posttranslational modification. Characterization of the abnormal prothrombin isolated from the plasma of cows fed the anticoagulant dicoumarol led directly to an understanding of the metabolic role of vitamin K. This protein lacked the specific calcium-binding sites present in normal prothrombin. Acidic peptides obtained by proteolytic enzyme digestion of prothrombin were shown to contain Gla

residues, but Gla residues could not be obtained by proteolysis of the abnormal prothrombin.

THE VITAMIN K-DEPENDENT CARBOXYLASE

In 1975, it was demonstrated that crude rat liver microsomal preparation contained an enzymatic activity (the vitamin K-dependent carboxylase) that promoted a vitamin K-dependent incorporation of $H^{14}CO_3^-$ into Gla residues of endogenous precursors of vitamin K-dependent proteins present in these preparations. Detergent-solubilized microsomal preparations retained this activity, and small peptides containing adjacent Glu−Glu sequences were found to be substrates for the enzyme. A general understanding of the properties of this unique enzyme was gained from studies utilizing this crude enzyme preparation, and these data have been adequately reviewed. The vitamin K-dependent carboxylation reaction does not require ATP, and the energy to drive this carboxylation reaction is derived from the oxidation of the reduced, hydronaphthoquinone, form of vitamin K (vitamin KH_2) by O_2 to form vitamin K-2,3-epoxide (Figure 2). The lack of a requirement for biotin and studies of the CO_2/HCO_3^- requirement indicate that carbon dioxide rather than HCO_3^- is the active species in the carboxylation reaction. Active substrates for the enzyme are 2-methyl-1,4-naphthoquinones substituted at the 3-position with a rather hydrophobic group. Although some differences in biological activity can be measured, phylloquinone, the plant form of the vitamin, and the predominant bacterial menaquinones are all effective substrates (Figure 1).

Only a small fraction of the proteins secreted by the liver to the plasma are vitamin K dependent, so an efficient mechanism for recognizing the precursors of the vitamin K-dependent proteins is essential. The primary gene products of the vitamin K-dependent proteins contain a very homologous domain between the amino terminus of the mature protein and the signal sequence. This propeptide region appears to be both a docking or recognition site for the enzyme and a modulator of the activity of the enzyme by decreasing the apparent K_m of the Glu site substrate. The role of vitamin K in the overall reaction catalyzed by the enzyme is to abstract the hydrogen on the γ-carbon of the glutamyl residue to allow attack of CO_2 at this position. A number of studies that utilized substrates tritiated at the γ-carbon of each Glu residue have been used to define the action and stoichiometry involved and have shown that it was an apparent equivalent stoichiometry between vitamin K-2,3-epoxide formation and Gla formation. The mechanism by which epoxide formation is coupled to γ-hydrogen abstraction is key to a complete understanding of the role of vitamin K. The association between epoxide formation, Gla formation, and γ-C−H bond cleavage has been studied, and the reaction efficiency defined as the ratio of Gla residues formed to γ-C−H bonds cleaved has been shown to be independent of Glu substrate concentrations, and to approach unity at high CO_2 concentrations. Identification of an intermediate chemical form of vitamin K which could be sufficiently basic to abstract the γ-hydrogen of the glutamyl residue has been a challenge. A possible candidate is that first proposed by Dowd, who suggested that an initial attack of O_2 at the naphthoquinone carbonyl carbon adjacent to the methyl group results in the formation of a dioxetane ring which generates an alkoxide intermediate. This intermediate is hypothesized to be the strong base that abstracts the γ-methylene hydrogen and leaves a carbanion that can interact with CO_2. This pathway leads to the possibility that a second atom of molecular oxygen can be incorporated into the carbonyl group of the epoxide product as well as the
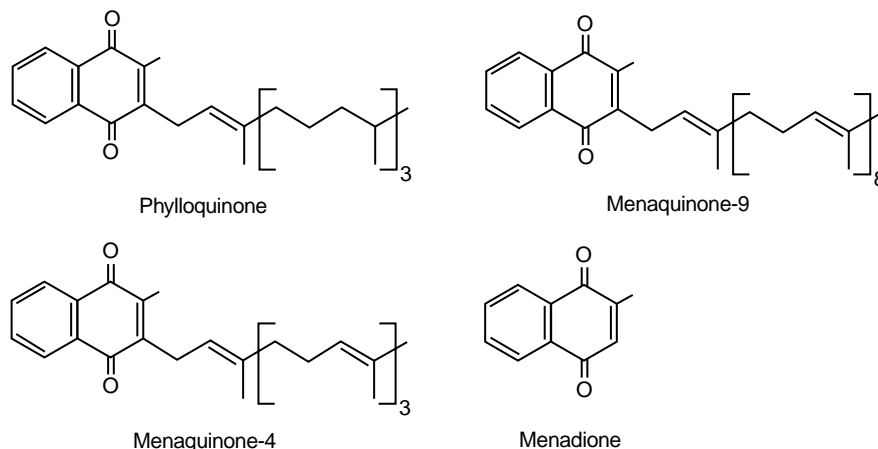
FIGURE 1 Structures of vitamin K active compounds. Phylloquinone (vitamin K_1) synthesized in plants is the main dietary form of vitamin K. Menaquinone-9 is a prominent member of a series of menaquinones (vitamin K_2) produced by intestinal bacteria, and menadione, vitamin K_3, is a synthetic compound that can be converted to menaquinone-4 by animal tissues.
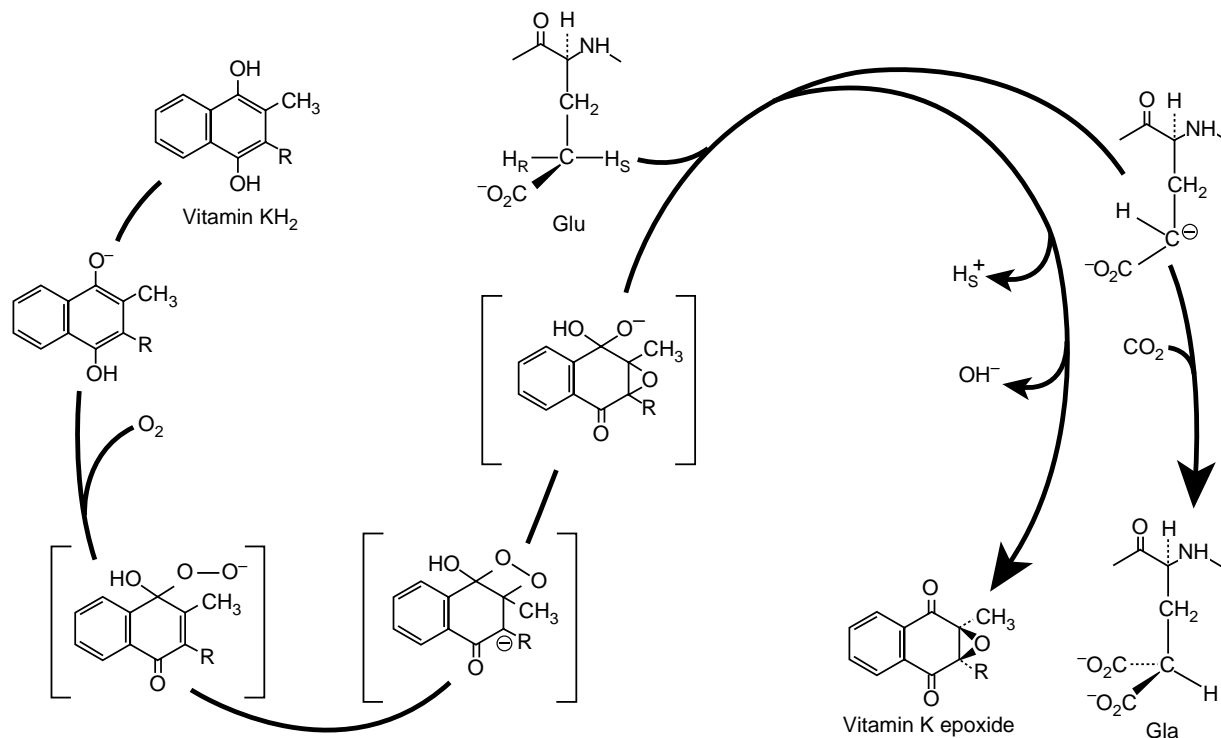
FIGURE 2 The vitamin K-dependent, γ-glutamyl carboxylase. The available data support an interaction of O_2 with vitamin KH_2, the reduced (hydronaphthoquinone) form of vitamin K, to form an oxygenated intermediate which is sufficiently basic to abstract the γ-hydrogen of the glutamyl residue. The products of this reaction are vitamin K-2,3-epoxide and a glutamyl carbanion. Attack of CO_2 on the carbanion leads to the formation of a Gla residue. The bracketed peroxy, dioxetane, and alkoxide intermediates have not been identified in the enzyme catalyzed reaction but are postulated based on model organic reactions, and the available data are consistent with their presence.

epoxide oxygen, and this partial dioxygenase activity has been demonstrated. Although the general scheme shown in Figure 2 is consistent with all of the available data, the mechanism remains a hypothesis at this time.

Progress in purifying this enzyme was slow. Utilization of bound propeptide as an affinity column aided progress in this field, and the enzyme was eventually purified to near homogeneity and cloned. The carboxylase is a unique 758 amino acid residue protein with a sequence suggestive of an integral membrane protein with a number of membrane-spanning domains in the N terminus, and a C-terminal domain located in the lumen of the endoplasmic reticulum. It has been demonstrated that the multiple Glu sites on the substrate for this enzyme are carboxylated processively as they are bound to the enzyme via their propeptide, while the Gla domain undergoes intramolecular movement to reposition each Glu for catalylsis, and that release of the carboxylated substrate is the rate-limiting step in the reaction.

Metabolic Interconversion of Vitamin K

Gla residues are not metabolized but are excreted in the urine. The amount of Gla excreted by an adult human is in the range of 50 μmol per day, so a similar amount must be formed each day. As the dietary requirement of vitamin K is only ∼0.2 μmol per day, and tissue stores are very low, it is clear that the vitamin K 2,3-epoxide generated by the carboxylase can be actively recycled by the pathway shown in Figure 3. The vitamin K-epoxide reductase is capable of reducing both the epoxide to the naphthoquinone, and naphthoquinone to the hydro-naphthoquinone, and this is the enzyme which is blocked by the common oral anticoagulant warfarin. In addition to the epoxide reductase, the quinone and hydronaphthoquinone forms of the vitamin can also be interconverted by a number of NAD(P)H-linked reductases including one that appears to be a micro-somal-bound form of the extensively studied liver DT-diaphorase activity. The epoxide reductase utilizes a sulfhydryl compound as a reductant *in vitro*, but the physiological reductant has not been identified. Efforts to purify and characterize this enzyme have not yet been successful.

A second metabolic pathway utilizing phylloqui-none has been discovered more recently. Chick liver contains a high concentration of menaquinone-4 (MK-4), and a number of extrahepatic tissues such as brain, salivary gland, and pancreas of rats and humans fed phylloquinone also contained a much higher
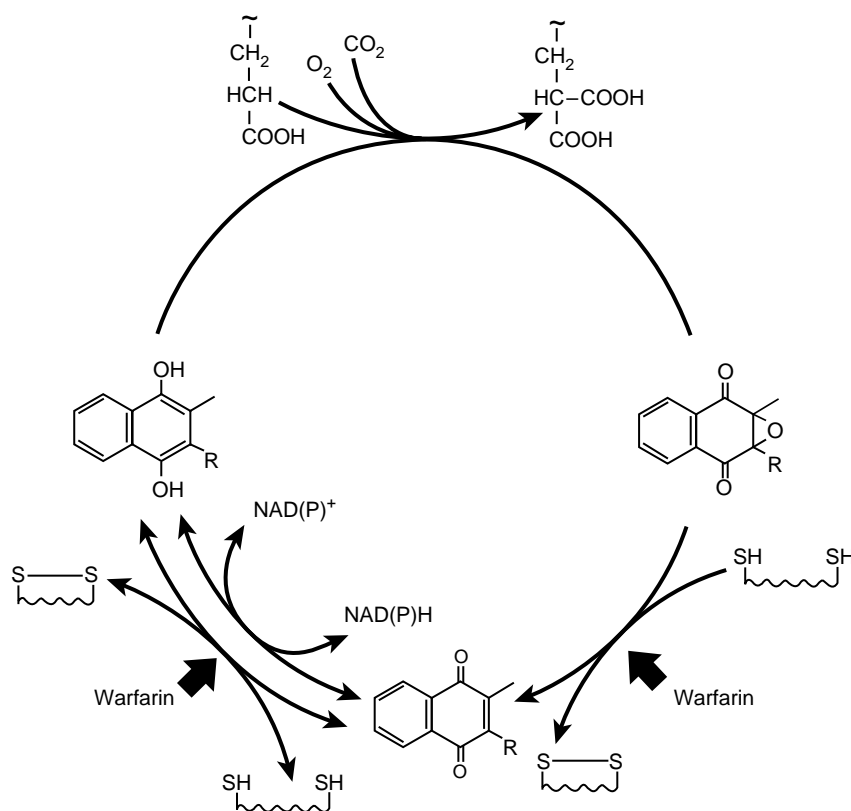
FIGURE 3 Tissue metabolism of vitamin K. Vitamin K epoxide formed in the carboxylation reaction is reduced to the quinone form of the vitamin by a warfarin-sensitive pathway, vitamin K epoxide reductase, that is driven by a reduced dithiol. The naphthoquinone form of the vitamin can be reduced to the hydronaphthoquinone form either by the same warfarin-sensitive dithiol-driven reductase or by one or more of the hepatic NADH or NADPH-lined quinone reductases which are less sensitive to warfarin.

concentration of MK-4 than of phylloquinone. Tissue formation of MK-4 from phylloquinone fed to germ-free rats and the observation that cultured kidney cells can convert phylloquinone to MK-4 have now demonstrated that bacterial action is not involved in the conversion and that numerous cell types have the ability to carry out this transformation. Details of the mechanism of this conversion are lacking, but it seems unlikely that a metabolic pathway leading to MK-4 has evolved unless there is some specific role for this vitamin. This role is unlikely to involve the vitamin K-dependent carboxylase, as phylloquinone and MK-4 have similar activity as a substrate for this enzymatic activity.

Nutritional Aspects of Vitamin K

Reports of uncomplicated adult deficiencies of vitamin K are extremely rare, and most diets contain an adequate amount (~ 100 mg per day) of vitamin K. Low lipid intake or the impaired lipid absorption resulting from the lack of bile salts will adversely affect vitamin K absorption, and depression of the vitamin K-dependent coagulation factors has frequently been reported in various malabsorption syndromes. This classic example of a human vitamin K deficiency is that of early vitamin K deficiency bleeding occurring during the first week of life in healthy appearing neonates. Low placental transfer of phylloquinone, low clotting factor levels, a sterile gut, and the low vitamin K content of breast milk all contribute to the disease. Although the incidence is low, the mortality rate from intracranial bleeding is high, and prevention by oral or parenteral administration of vitamin K immediately following birth is the standard cure. The presence of large amounts of osteocalcin in bone and reports that low vitamin K intake or increased amounts of circulating under-γ-carboxylated osteocalcin are correlated with low bone mineral density or fracture rate have focused significant attention on a possible role for vitamin K in bone health. A clear understanding of the significance of these reports is not yet available.

See Also the Following Articles

Quinones • Vitamin K: Blood Coagulation and Use in Therapy

Glossary

γ-carboxyglutamic acid An amino acid found in a limited number of proteins, formed by a posttranslational modification of Glu residues.

γ-glutamyl carboxylase The enzyme involved in the formation of Gla residues; also called vitamin K-dependent carboxylase.

Further Reading

Berkner, K. L. (2000). The vitamin K-dependent carboxylase. *J. Nutr.* **130,** 1877–1880.

Binkley, N. C., and Suttie, J. W. (1995). Vitamin K nutrition and osteoporosis. *J. Nutr.* **125,** 1812–1821.

Dowd, P., Ham, S.-W., Naganathan, S., and Hershline, R. (1995). The mechanism of action of vitamin K. *Annu. Rev. Nutr.* **15,** 419–440.

Furie, B., Bouchard, B. A., and Furie, B. C. (1999). Vitamin K-dependent biosynthesis of γ-carboxyglutamic acid. *Blood* **93,** 1798–1808.

Presnell, S. R., and Stafford, D. W. (2002). The vitamin K-dependent carboxylase. *Thromb. Haemost.* **87,** 937–946.

Suttie, J. W. (2001). Vitamin K. In *Handbook of Vitamins* (R. B. Rucker, J. W. Suttie, D. B. McCormick and L. J. Machlin, eds.) pp. 115–164. Marcel Dekker, New York.

Biography

John W. Suttie is Professor Emeritus of Biochemistry and Nutritional Sciences at the University of Wisconsin-Madison. He holds a Ph.D. from the University of Wisconsin-Madison. His research efforts have been directed toward an understanding of the biochemical and nutritional roles of vitamin K.

Vitamin K: Blood Coagulation and Use in Therapy

Matthew D. Stone and Gary L. Nelsestuen
University of Minnesota, Twin Cities, Minnesota, USA

Vitamin K is a fat soluble, quinone derivative (Figure 1) that is critical for healthy blood coagulation or hemostasis. It functions as a cofactor for a carboxylase in the post-translational production γ-carboxyglutamic acid (Gla), a unique amino acid that has calcium ion affinity. Henrik Dam first proposed the presence of this vitamin in the late 1920s and early 1930s in Germany through feeding chicks a strict organic solvent extracted diet. It was observed that these chicks died from uncontrolled bleeding. As a result, the vitamin was recognized as lipid soluble and named after the German word *koagulation*. Many proteins involved in blood coagulation contain Gla residues and these residues are essential for function.

Nomenclature

Vitamin K is defined as any molecule that contains the 2-methyl-1,4-naphthoquinone functional group (Figure 1A) and has biological function; i.e., it can reverse physiological effects of vitamin K deficiency. Two common forms are vitamin K_1 or phylloquinone (Figure 1B) and vitamin K_2 or menaquinone (Figure 1C). Vitamin K_1 has three repeating saturated prenyl groups adjoined to an initial unsaturated prenyl group at carbon 3 of the quinone. On the other hand, vitamin K_2 describes a class of compounds that contain multiple repeating unsaturated prenyl groups at carbon 3 of the quinone.

Biochemical Role of Vitamin K in Carboxylation

Vitamin K is a cofactor for a carboxylase that catalyzes the addition of CO_2 onto a peptidyl glutamic acid residue to produce γ-carboxyglutamic acid. Formation of a new carbon–carbon bond through carboxylation is energetically unfavorable. A more familiar method for biological carboxylation uses biotin as a cofactor and is coupled to adenosine triphosphate (ATP) hydrolysis to provide energy for product formation. In the vitamin K-dependent reaction, however, the energy to drive carboxylation is provided by vitamin K oxidation.

The reaction mechanism of the carboxylase is still under investigation. However, it is believed that the redox cycle (Figure 2) starts with the reduced form of vitamin K (KH_2). Molecular oxygen addition results in an epoxide, which is coupled to a concerted condensation of CO_2 and glutamic acid to form Gla. Within the concerted reaction, vitamin K oxidation has been proposed to result in a very strong base, which could abstract one of the γ protons of glutamic acid. The vitamin K epoxide is no longer active and must be reduced back to KH_2 in two stages by at least one reductase with concomitant disulfide oxidation.

The carboxylase is localized to the lumen of the endoplasmic reticulum from where it can modify newly synthesized peptides. Proteins to be carboxylated are recognized by a signal sequence near the amino terminus. The signal sequence is removed at a further processing event. The carboxylase only modifies specific glutamic acid residues.

Use of vitamin K as a biofactor appears to have evolved recently, since the vitamin K-dependent carboxylase has only been discovered in the animal kingdom. Naturally, it is expressed in all vertebrates with circulatory systems. However, it is also found in snails of the Conus genus, and the gene was recently identified in arthropods in *D. melanogaster*. High homology in sequence alignments suggests that the invertebrate carboxylase is an ancestor of the vertebrate form and that vitamin K function arose early in animal evolution.

Function of Vitamin K-Dependent Proteins

BLOOD COAGULATION

The most well-understood biological process involving vitamin K-dependent proteins is blood coagulation. Blood coagulation is a complex pathway mediated by a
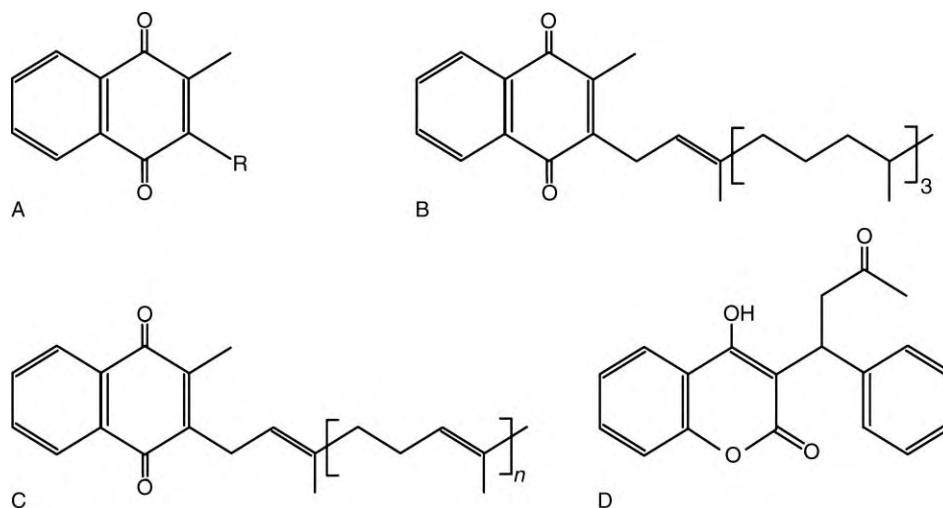
FIGURE 1 Vitamin K derivatives. (A) 2-methyl-1,4-napthoquinone, the functional group of vitamin K. (B) Phyloquinone or vitamin K_1. (C) Menaquinone or vitamin K_2. (D) Warfarin, a vitamin K antagonist.
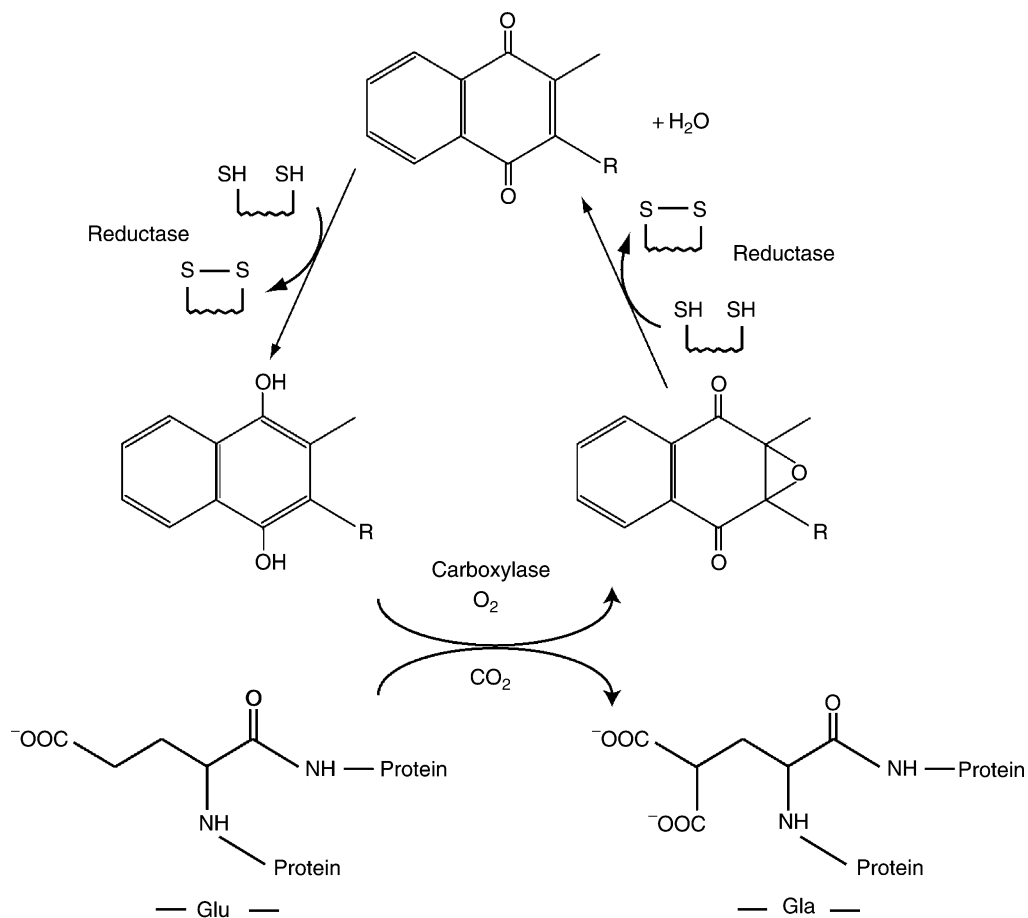
FIGURE 2 The vitamin K cycle. Oxidation of KH_2 to the epoxide provides sufficient energy for carboxylation and may participate in hydrogen abstraction. The cycle is completed by two reduction steps that regenerate KH_2.

series of proteolytic activation steps resulting in signal amplification (Figure 3). Many of the proteolytic reactions of the cascade depend on specific binding to membranes composed of acidic phospholipids such as phosphatidylserine. Binding to membrane aids the blood coagulation reactions by (1) directing the clotting factors to the site of injury and (2) increasing the concentration of clotting factors through localization to a restricted area.

The Procoagulation Pathway

Vitamin K-dependent plasma proteins within the procoagulant cascade include factor VII, factor IX, factor X, and prothrombin. The amino terminal 1–45 amino acids of each protein contain 9–12 Gla residues, which are not found elsewhere in the protein. As a result, this region is referred to as the Gla domain. The Gla residues bind calcium ions and the resulting conformational change shifts the Gla domain from a disordered structure to one that is

able to bind to membranes. The exact membrane-binding site is still controversial.

Factor VII, factor IX, factor X, and prothrombin are zymogens containing a trypsin-like serine protease domain. The corresponding active enzymes are denoted with an "a" after the Roman numeral (e.g., factor VIIa). The exception is thrombin, the activated form of prothrombin (Table I).

Procoagulation is traditionally divided into two pathways, extrinsic and intrinsic, both leading to the production of factor Xa (Figure 3A). The extrinsic pathway is so-called because it relies on the expression of the cell surface receptor, tissue factor (TF), which is not found in plasma.

Blood coagulation is tightly regulated. Two important modes of regulation are (1) the above-mentioned selective membrane binding and (2) allosteric activation through cofactor protein binding. Initiation of the extrinsic pathway results from vascular injury. Injury disrupts cell membranes thereby exposing acidic
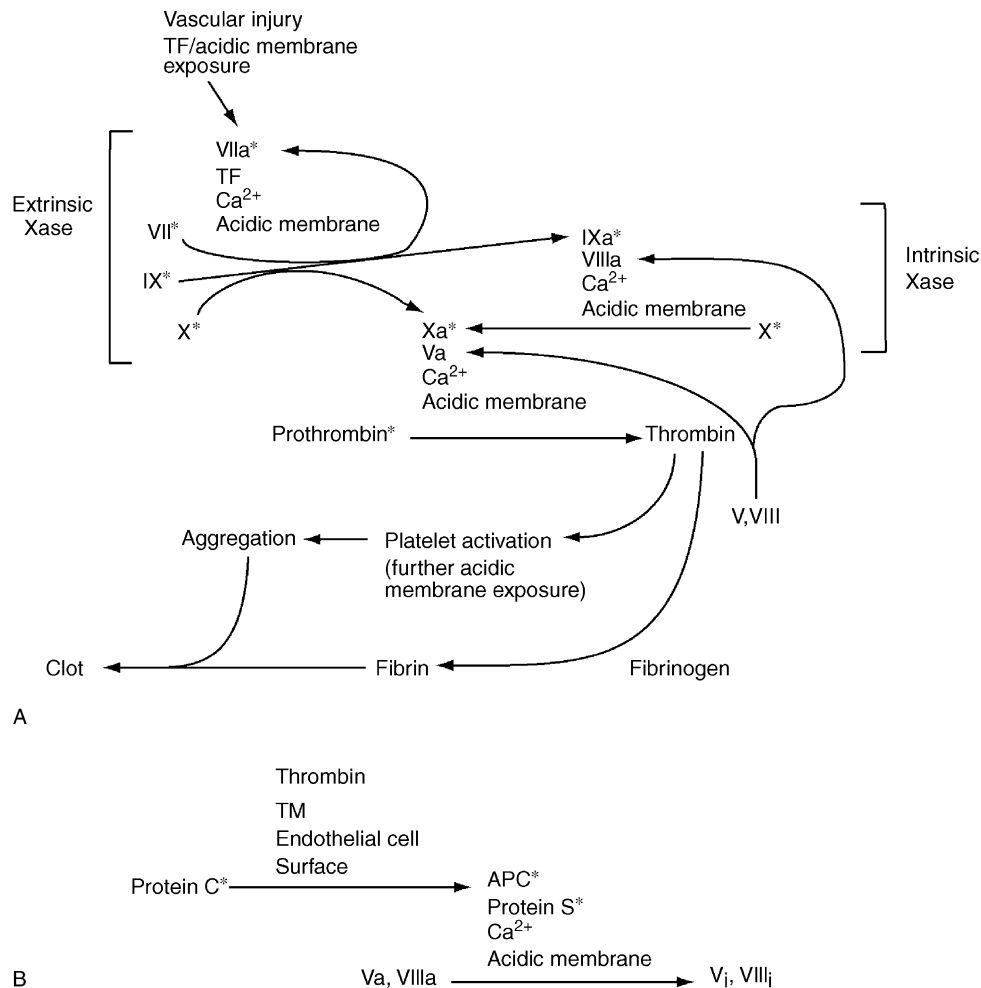
FIGURE 3 The coagulation cascade. (A) The procoagulation cascade is mediated by a series, a proteolytic activation steps with feedback activation for signal amplification. (B) The anticoagulant pathway inactivates factor Va and VIIIa. Enzyme complexes are depicted as vertical text columns with the proteolytic enzyme first, protein cofactors second, and calcium and membrane components last. The asterisk denotes a vitamin K-dependent protein.

TABLE I

List of Coagulation Factors

Factor	Active form	Alternative name	Alternative name of active form	Function
I	Ia	Fibrinogen	Fibrin	Clot forming protein
II[a]	IIa	Prothrombin	Thrombin	Enzyme
III		Membrane surface with tissue factor		Cofactor
IV		Calcium		Cofactor
V	Va	Proaccelerin	Accelerin	Protein cofactor
VII[a]	VIIa[a]	Proconvertin	Convertin	Enzyme
VIII	VIIIa	Anti-hemophilic factor		Protein cofactor
IX[a]	IXa[a]	Christmas factor		Enzyme
X[a]	Xa[a]	Stuart factor		Enzyme

[a] Denotes vitamin K-dependent protein.

phospholipids, such as phosphatidylserine, that are not normally exposed on healthy cell surfaces. Injury also exposes TF, the cofactor for factor VIIa, which is expressed in subendothelial layers. Protease activity of VIIa is greatly enhanced when bound in complex with membrane and TF allowing it to efficiently catalyze three reactions. It activates itself, factor VII to VIIa, factor IX to IXa, and factor X to Xa. Factor Xa similarly produces a small amount of thrombin, which initiates the intrinsic pathway through activation of factor VIII to factor VIIIa. Thrombin also activates factor V to factor Va and it activates platelets. Activation of platelets causes them to expose acidic membrane, thereby expanding the clotting reaction surface and to aggregate, thereby forming a plug at the site of injury.

The intrinsic pathway constitutes the second phase of the coagulation reaction and functions to amplify the clotting stimulus. Factor IXa, in complex with membrane and Factor VIIIa, catalyzes further conversion of factor X to factor Xa. Classic hemophilia is a result of deficiencies in either factor VIII or IX. Buildup of factor Xa and its cofactor, factor Va, leads to efficient large-scale production of thrombin. Finally, thrombin hydrolyzes fibrinogen to fibrin, which self-polymerizes into a macro-molecular protein net that, along with the platelet plug, comprises a clot.

The Anticoagulation Pathway

Of numerous anticoagulation pathways, one is comprised of two vitamin K-dependent plasma proteins, protein C and protein S (Figure 3B). This pathway is initiated when thrombin, in complex with the endothelial cell surface receptor thrombomodulin, catalyzes proteolytic activation of protein C to activated protein C (APC). Protein S, which is not an enzyme, acts as a cofactor for APC. Together these proteins inhibit amplification of the coagulation reaction through proteolytic inactivation of factor Va and factor VIIIa. Protein Z is another non-enzymatic vitamin K-dependent plasma protein involved in anticoagulation. It acts as a cofactor for a specific factor Xa inhibitor protein.

OTHER VITAMIN K-DEPENDENT PROTEINS

Vitamin K has diverse, although less characterized, roles outside of blood coagulation. One vitamin K-dependent protein, the product of the growth arrest specific 6 gene (Gas6), has high homology to protein S and appears to be involved in development. Skeletal tissues express two vitamin K-dependent proteins, bone matrix Gla protein, and osteocalcin. These proteins are rather small, contain only two Gla residues, and have little similarity to the plasma proteins. The bone matrix protein has been implicated in the prevention of vascular calcification.

Small unique peptide toxins called conantokins from snails of the Conus genus have also been identified as vitamin K dependent. These toxins function as neurotransmitter inhibitors that paralyze fish for these predatory snails.

The majority of vitamin K-dependent plasma proteins are expressed in the liver. However, virtually every cell type aside from muscle expresses the vitamin K-dependent carboxylase, implying that vitamin K is vital to many tissues. Some cell receptor proteins have been shown to contain Gla domains. However, the dominant symptom of vitamin K deficiency is a loss of blood coagulation potential.

Nutrition

Vitamin K_1 is found mainly in green leafy vegetables and to a lesser extent in cooking oils such as soybean and canola oil. The equally active vitamin K_2 molecules are synthesized by intestinal bacteria and absorbed with bile salts. However, the contribution of these derivatives toward the total vitamin K requirement has not been accurately determined. Nevertheless, the recommended daily amount of dietary vitamin K is 1 μg per kg of body weight.

Since the requirement is rather low and a typical adult human diet contains an excess of vitamin K, deficiencies are rare. Those that do occur are probably due to a variety of factors such as poor diet, liver disease or failure, low surface area for intestinal absorption, inadequate production of bile salts, or administration of antibiotics that could disturb bacterial vitamin K production.

Vitamin K deficiency leading to hemorrhagic disease in newborns is observed and is probably due to a lack of

established bacteria in the intestines of newborns, and a low vitamin K content in breast milk. As a result, newborns are given a 0.5–1.0 mg vitamin K injection upon birth and baby formula is typically fortified with vitamin K.

Use in Therapy

VITAMIN K ANTAGONISTS AS ANTICOAGULANTS

Several structural analogues of vitamin K, known as 4-hydroxycoumarins, have been synthesized and function as anticoagulants. The most common is warfarin (Figure 1D), also described by the term, coumadin. Warfarin functions as a vitamin K antagonist through inhibition of the epoxide reductase in the vitamin K reaction cycle (Figure 2). Treatment results in under-carboxylation of plasma proteins leading to loss of function and an anticoagulant state. Warfarin is typically used as a long-term anticoagulant and can be administered orally. However, therapy must be monitored, since overdose can lead to uncontrolled bleeding. Additionally, warfarin is an effective rodenticide and is the active ingredient in many pest traps.

RECOMBINANT VITAMIN K-DEPENDENT PROTEINS AS ANTICOAGULANTS

Advances in recombinant DNA technology have allowed the expression and purification of vitamin K-dependent proteins for use in factor replacement therapy. Recombinant proteins are more attractive than plasma concentrates because of the threat of viral contamination in donated blood.

Recombinant APC is useful for treatment of severe sepsis. Sepsis is caused by bacterial infection leading to a massive inflammatory response, which causes aberrant endothelial expression of TF as well as down-regulation in protein C levels. Severe sepsis leads to undesirable activation of the coagulation cascade inducing stroke or aneurysm and is often deadly. APC has proven to be an effective therapy of severe sepsis. Not only does it inhibit coagulation, but it also inhibits inflammation, and has been shown to significantly reduce associated fatalities.

RECOMBINANT VITAMIN K-DEPENDENT PROTEINS AS PROCOAGULANTS

Recombinant factor IX is used in replacement therapy for factor IX deficiency. Extreme cases of factor VIII or IX deficiency arise in hemophiliacs who develop antibodies to either factor VIII or IX that is administered through replacement therapy. In these patients, high dose recombinant factor VIIa can be used to treat bleeding episodes by a mechanism that bypasses the intrinsic pathway.

SEE ALSO THE FOLLOWING ARTICLES

Proteases in Blood Clotting • Quinones • Vitamin K: Biochemistry, Metabolism, and Nutritional Aspects

GLOSSARY

carboxylase An enzyme that catalyzes the addition of a carboxylic acid group onto a substrate.
coagulation The process of blood clot formation to stop bleeding.
cofactor Any agent (protein, metabolite, metal ion, etc.) that aids an enzyme during catalysis.
hemophilia A disease characterized by pronounced bleeding and a prolonged response to a coagulation stimulus.
warfarin or coumadin Orally administered vitamin K antagonists.

FURTHER READING

Booth, S., and Suttie, J. (1998). Dietary intake and adequacy of vitamin K. *J. Nutr.* **128**, 785–788.

Dowd, P., Ham, S., and Geib, S. (1995). The mechanism of action of vitamin K. *Annu. Rev. Nutr.* **15**, 419–440.

Hirsch, J., Ginsberg, J., and Marder, V. (1994). Anticoagulant therapy with coumarin agents. In *Hemostasis and Thrombosis: Basic Principles and Clinical Practice* (R. Colman, J. Hirsch, V. Marder and E. Salzman, eds.) pp. 1567–1583. J. B. Lippincott, Philadelphia.

McCoy, C., and Matthews, S. (2003). Drotrecogin alfa (recombinant human activated protein C) for the treatment of severe sepsis. *Clin. Ther.* **25**, 396–421.

Nelsestuen, G., Shah, A., and Harvey, S. (2000). Vitamin K-dependent proteins. *Vitam. Horm* **58**, 355–389.

Stenflo, J. (1999). Contributions of Gla and EGF-like domains to the function of vitamin K-dependent coagulation factors. *Crit. Rev. Eukayotic Gene Exp.* **9**, 59–88.

Suttie, J. (2001). Vitamin K. In *Handbook of Vitamins* (R. Rucker, J. Suttie, D. McCormick and L. Machlin, eds.) pp. 115–164. Marcel Dekker, New York.

Veldman, A., Hoffman, M., and Ehrenforth, S. (2003). New insights into the coagulation system and implications for new therapeutic options with recombinant factor VIIa. *Curr. Med. Chem.* **10**, 797–811.

BIOGRAPHY

Gary Nelsestuen is currently the Samuel Kirkwood Professor of Biochemistry, Molecular, and Biophysics at the University of Minnesota, Twin Cities. He was one of the investigators who discovered the relationship of vitamin K to Gla. He has also made numerous contributions in understanding the nature of interaction between calcium-binding proteins and membranes. He received his Ph.D. from the University of Minnesota in 1970 and was a postdoctoral fellow at the University of Wisconsin.

Matthew D. Stone currently holds the Thomas Reid Fellowship at the University of Minnesota, Twin Cities in the Department of Biochemistry, Molecular Biology, and Biophysics. He is a new researcher in the vitamin K field and currently studies functional consequences from the binding of vitamin K-dependent plasma proteins with membranes.

Voltage-Dependent K$^+$ Channels

Ramón Latorre and Francisco J. Morera
Centro de Estudios Científicos, Valdivia, Chile

Voltage-dependent K$^+$ channels are the intrinsic membrane proteins that by forming a conduction pathway, or pore, through which K$^+$ selectively diffuses down their electrochemical gradient across the membrane, draw the cell-resting potential toward the K$^+$ equilibrium potential. The voltage imposed across the cell membrane governs pore opening – hence the name voltage-dependent K$^+$ channels. By hyperpolarizing the cell membrane, these channels serve as the appropriate negative feedback to excitatory inputs on a cell when they are active. In general, voltage-dependent K$^+$ channels increase their activity on membrane depolarization and their activation dampens the excitatory events that elevate the cytosolic Ca^{2+} and/or depolarize the cell membrane. Voltage-dependent K$^+$ channels are widely distributed in mammalian tissues and play important physiological roles. Setting the duration of action potentials and the interspike interval during repetitive firing in neurons and in the heart, K$^+$ secretion in epithelia, setting the smooth muscle tone, and determining the resonance frequency in hair cells are some of the roles that voltage-dependent K$^+$ channels play. Voltage-dependent K$^+$ channels are made of subunits that assemble in the bilayer as tetramers to form highly selective K$^+$ channels. The large number of different genes, alternative splicing, auxiliary β-subunits, and metabolic regulation explain the tremendous diversity of voltage-dependent K$^+$ channels.

Voltage-Dependent K$^+$ Channel Structure and Diversity

Voltage-dependent K$^+$ (eag, AKT1, KAT1, hslo, KCNQ, K$_v$) channels are tetrameric assemblies, each subunit consisting of six hydrophobic segments, S1–S6. Hslo is an exception containing seven hydrophobic segments, S0–S6. In 2003, the group of Rod MacKinnon determined the crystal structure of the full-length K$_v$ channel from the bacterium *Aeropyrum pernix* (K$_v$AP) at a resolution of 3.2Å. In the K$_v$AP channel, the linker between the S5 and S6 segment dips back down into membrane and participates in formation of the channel pore and determines ion selectivity (Figures 1(Aa) and 1(Ab)). This region is commonly referred to as the P region or the pore loop. Figures 1(Aa) and 1(Ab) show a comparison of the K$_v$AP pore (blue) with the KscA K$^+$ channel (green). The pore-lining, inner, S6-helices contain a conserved glycine (red dots in Figure 1(Aa)). This residue has been proposed to be a gating hinge that allows the inner helix bends 30° when the pore opens. In Figures 1(Aa) and 1(Ab), the K$_v$AP pore is in the open and the KcsA pore is in closed configuration. All K$^+$ channels have, inside the P region, a consensus amino acid sequence: TXXTVGYGD, dubbed the "signature sequence." The residues TVGYG line the selectivity filter (in some K$^+$ channels the tyrosine (Y) residue is replaced by phenylalanine (F), for example, in the eag channel). In the selectivity filter, carbonyl oxygen atoms are directed toward the pore to coordinate dehydrated K$^+$ ions and hydrophobic chains of valine and tyrosine directed toward the hydrophobic core surrounding the filter stabilize the main chain (Figure 1(Ac)).

Gating in the K$^+$ channels is conferred through the attachment of gating domains to the pore. The basic function of these gating domains is to perform mechanical work on the ion conduction pore to change its conformation between closed and open states. In voltage-dependent channels, a "voltage sensor" domain (S1–S4) is present on each subunit. Thus, a voltage sensor converts energy stored in the membrane electric field into mechanical work. There is evidence that the positive charges contained in S4 are the voltage-sensing elements. Thus, ion channel gating is essentially an electromechanical coupling between a gating unit and a pore unit.

K$_v$AP channels contain a central ion-conduction pore surrounded by the voltage sensor. Each S4 segment forms half of a "voltage-sensor paddle" (Figure 1(B)). Whether or not this is the structure involved in voltage-dependent activation is at present under heavy scrutiny.

Voltage-dependent K$^+$ channels are ubiquitously distributed in different cell and tissues and, therefore, K$^+$ channel diversity is of great importance in determining the variety of electrical responses of cells when subjected to stimuli. The possible mechanisms that originate the immense voltage-dependent K$^+$ channel diversity are: (1) multiple genes; (2) alternative splicing;

(3) formation of heteromultimeric channels; and (4) coexpression with regulatory β-subunits.

Voltage-dependent K⁺ channel diversity seems to have accompanied animal cells through evolution, and some K⁺ channels are the most conserved proteins in eukaryotes. Potassium channels became fundamental to animal cell physiology very early in evolution, and through voltage-dependent K⁺ channel control of membrane potential they couple the cell inner dynamics to the outer environment. A differential expression of voltage-dependent K⁺ channel mRNAs is found to accompany the events of animal development and this also occurs in adult animal tissues. This differential expression of voltage-dependent K⁺ channels is particularly noteworthy in the brain where it has been documented from invertebrate to mammals.

Some idea of the diversity of voltage-dependent K⁺ channel genes is indicated in the dendrogram shown in Figure 2. The dendrogram shows the major classes of voltage-dependent K⁺ channels. The ramification of voltage-dependent K⁺ channels was constructed by sorting the genes according to their similarity of amino acid sequence. In order to compare the amino acid sequences of the different voltage-dependent K⁺ channels, the sequences were aligned using at least 300 amino acids from six hydrophobic segment (S1–S6) channels and their respective linkers. Branch lengths do not represent time, but the branching is expected to preserve evolutionary relationships.

Physiological Function of Voltage-Dependent K⁺ Channels

ETHER-A-GO-GO K⁺ CHANNELS

Mutant fruit flies which shook their legs like go-go dancing when anesthetized with ether were carriers of a mutated gene of a voltage-gated K⁺. This gene was called ether-a-go-go or *eag*. A family of *eag*-like genes was isolated from *Drosophila*, known as aeg, erg, and elk. The human ether-a-go-go related gene (HERG) was cloned by homology from a human hippocampal cDNA library. HERG was strongly expressed in heart and LQT2, a mutation on this gene, impairs the
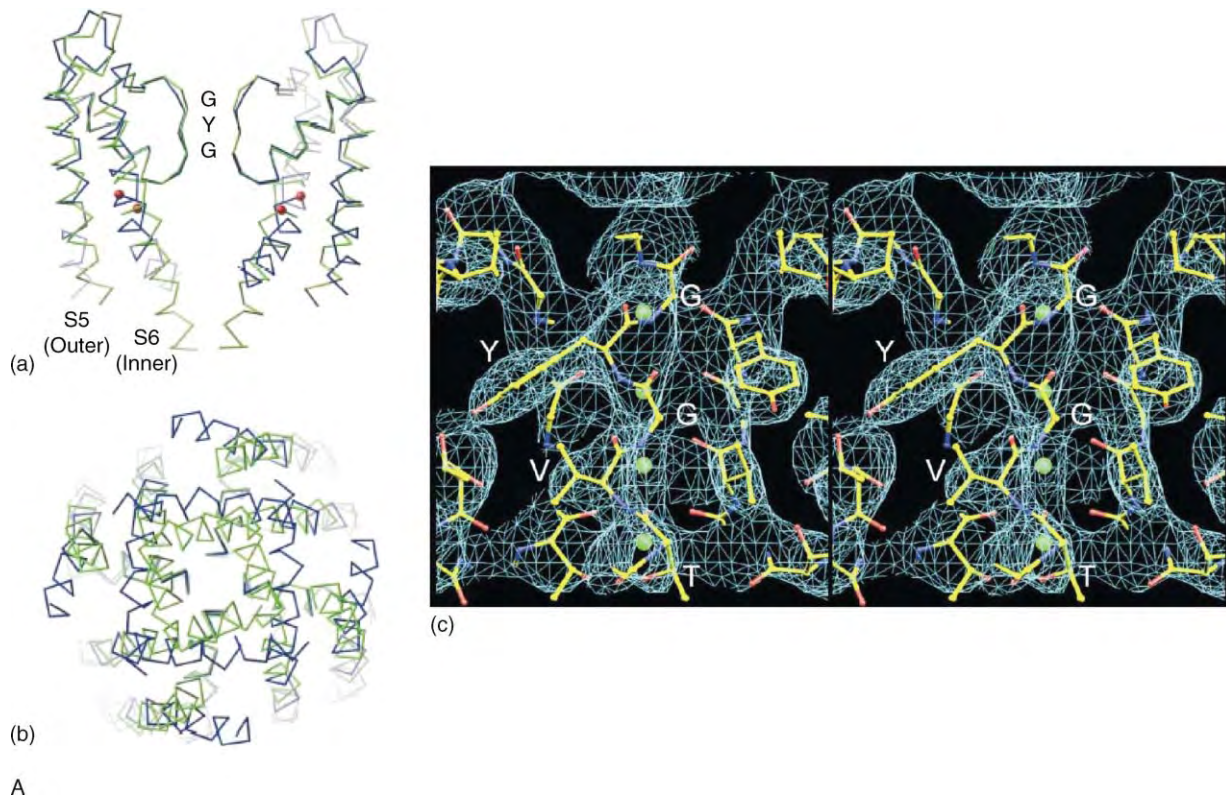
FIGURE 1 (A) The crystal structure of K$_v$ channel from the bacterium *Aeropyrum pernix* (KvAP) (a); view from the side of the K$_v$AP pore (blue) and comparison with the KcsA channel (green) (b); and the α-carbon traces are shown from the intracellular solution (c). (B) Architecture of the K$_v$AP channel. Image of the K$_v$AP channel tetramer viewed from the intracellular side (a) of the membrane and from the side with the intracellular solution (b) below, and a single subunit viewed from the side (c and d). Each subunit has a different color. (Modifed from Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chalt, B. T., and MacKinnon, R. (2003). X-ray structure of a voltage-dependent K⁺ channel. *Nature* **423**, 33–41.)
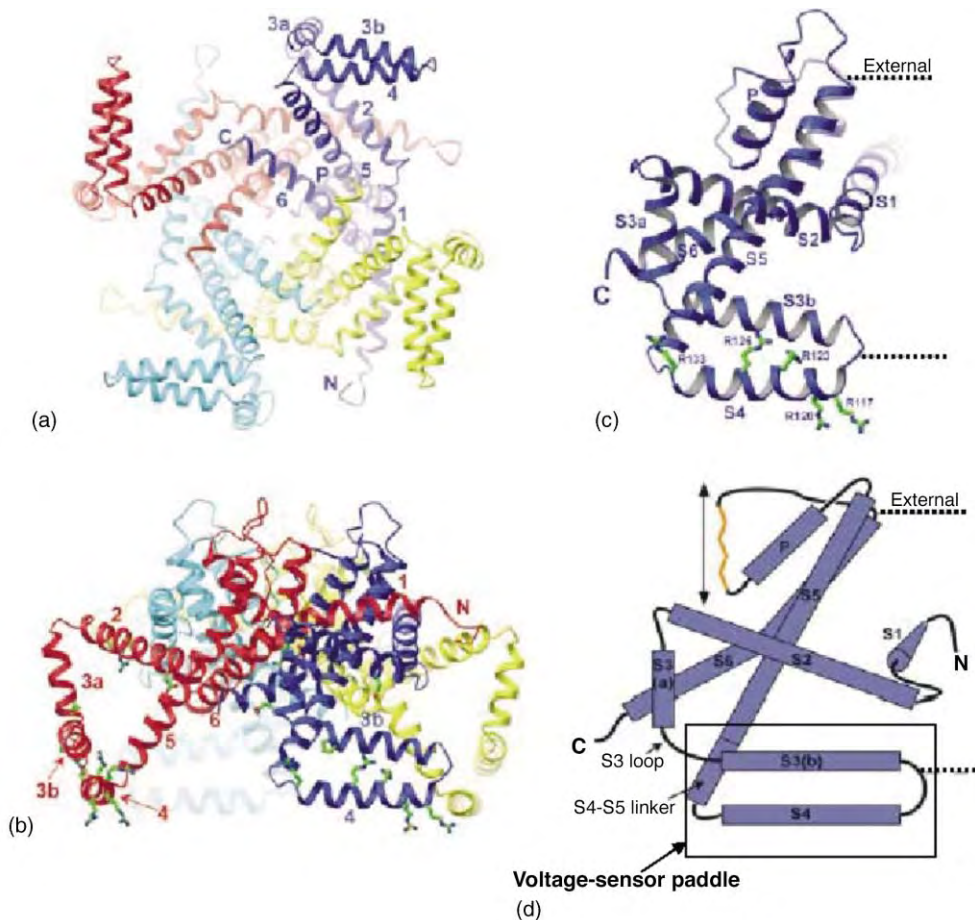
FIGURE 1 *(continued)*

repolarization of the cardiac action potential, causing a form of long QT syndrome.

PLANT K⁺ CHANNELS: KAT1 AND AKT1

In plants, voltage-dependent inward rectifiers mediate potassium uptake. KAT1 and AKT1 were the first inward rectifier K⁺ channels cloned from the plant *Arabidopsis thaliana*. The KAT1 channel is expressed in guard cells and the vascular system of the stem. AKT1 is present in guard cells and in the root cortex and it plays an important role in K⁺ uptake from the soil.

CALCIUM-ACTIVATED K⁺ CHANNEL

A change in the intracellular calcium concentration is a ubiquitous signaling mechanism that is frequently coupled with changes in membrane potential. These signals are integrated by the activation of a heterogeneous family of potassium channels, Ca^{2+}-activated potassium channels (K_{Ca}), whose open probability is increased by the elevation of cytosolic calcium to cause membrane hyperpolarization. They are found in a wide range of different tissues and in virtually all multicellular organisms. In nerve cells, they play a key role in controlling the action potential waveform and in regulating cell excitability. In secretory epithelia they are also involved in K⁺ secretion. K_{Ca} channels contribute to the repolarization and after-hyperpolarization of vertebrate neurons. Since this article is related to a voltage-dependent K⁺ channel only, in this section the importance of one K_{Ca} channel family code by a single gene called *slowpoke* (*slo*) will be discussed. Expression of *Slo* in mammalian cells or *Xenopus* oocytes induces the appearance of a highly K⁺-selective, Ca^{2+}-activated and voltage-dependent channel also known as MaxiK or BK (big K) because of its large single-channel conductance (>200 pS in 100 mM symmetrical K⁺). At difference of its other voltage-dependent K⁺ channel congeners, the Slo protein has an extra hydrophobic segment at the NH2 terminus denominated S0. Thus, the NH2 terminus of the protein is placed in the extracellular side of the membrane. The intracellular COOH-terminal comprises about two-thirds of the protein containing four hydrophobic segments, several alternative splicing sites, and a stretch of negatively

	Localization	**Function**
eag	Heat	Cardiac action potential repolarization
AKT1	Root in plants	Potassium uptake from soil
KAT1	Guard cell in plants	Stomatal opening
hslo	Ubiquitous	Control of smooth muscle tone, neurosecretion, hearing
KCNQ5	Brain, skeletal muscle	Neuronal M-current
$K_v1.1$	Neurons, SA nodal cell	Repolarization of the action potential
$K_v1.2$	Smooth muscle	Contractile tone regulation
$K_v1.3$	Leukocytes, lymphocytes, osteoclasts, heart	Inflammatory response?, cell volume regulation
$K_v1.6$	Smooth muscle	Contractile tone regulation
$K_v1.5$	Smooth muscle, Schwann cells	Contractile tone regulation, potassium buffering
$K_v1.7$	Heart, lung, skeletal muscle, kidney	
$K_v1.4$	Smooth muscle, mossy fiber	Contractile tone regulation
$K_v3.1$	Auditory brain stem neurons, lymphocytes	High-frequency firing, lymphocyte activation?
$K_v4.1$	Pulmonary arterial smooth muscle	Pulmonary circulation
$K_v4.2$	Myocardium	Action potential prolongation in heart
$K_v4.3$	Brain, heart	Setting frequency of neuronal firing and heart pacing
$K_v6.1$	Brain, heart	Modulate activity of $K_v2.1$
$K_v6.2$	Heart	Cardiac action potential repolarization
$K_v2.1$	Smooth muscle, pulmonary arteries, brain	Contractile tone regulation, pulmonary circulation
$K_v2.2$	Brain, smooth muscle	
$K_v8.1$	Brain	Inhibits activity of K_v2
$K_v9.1$	Brain	Modulate activity of $K_v2.1$ and $K_v2.2$
$K_v9.2$	Brain	Modulate activity of $K_v2.1$ and $K_v2.2$
$K_v9.3$	Smooth muscle	Modulate activity of $K_v2.1$ and $K_v2.2$
$K_v5.1$	Heat	Modulate activity of $K_v2.1$
$K_v11.1$	Lung, liver, kidney, testis, colon, spleen	Heterotetramer with $K_v2.1$

FIGURE 2 The major classes of voltage-dependent K^+ channels, their tissue localization and function. Comparison of the amino acid sequences of the different α-subunits for voltage-dependent K^+ channels. The dendrogram was constructed using ClustalW and TreeView. The more similar the sequences for a pair of genes, the shorter the path connecting them.

charged amino acids (aspartates) denominated the "Ca^{2+} bowl." The Ca^{2+} bowl has emerged as a good candidate for the high-affinity Ca^{2+}-binding site.

Slo channels control a large variety of physiological processes, including smooth muscle tone, neurosecretion, and hearing. Despite being coded by a single gene, the diversity of Slo channels is great. Regulatory β-subunits, splicing, and metabolic regulation create this diversity fundamental to the adequate function of many tissues. Except in heart myocytes, Slo channels are almost ubiquitously expressed among mammalian tissues, and they play a variety of roles. In smooth muscle, local Ca^{2+} increases, called Ca^{2+} sparks, generate spontaneous transient outward currents, which are generated by Slo channels. This hyperpolarizes the membrane causing muscle relaxation. In chromaffin cells, Slo channels are important for rapid termination of the action potential. In the cochlea of turtles, frogs and chicks, gradients of spliced Slo channels, and β-subunits through an interplay with voltage-dependent Ca^{2+} channels, they are able to generate the different oscillatory responses in the hair cells.

KCNQ CHANNELS

The first member of the KCNQ family to be cloned, *KCNQ1*, was isolated by virtue of the fact that mutations in this gene give rise to the most common form of long QT syndrome, LQT1. Exploiting their homology to KCNQ1, other members of this family were subsequently cloned. KCNQ1 encodes the major subunit of the cardiac I$_{Ks}$ channel, which is involved in repolarization of the ventricular action potential. KCNQ1 mRNA is expressed strongly in the heart, with lower levels in pancreas, kidney, lung, placenta, and ear.

It has been suggested that KCNQ2 and KCNQ3 is a heteromer, an idea which is consistent with their overlapping tissue distribution and the fact that antibodies directed against KCNQ2 are able to coimmunoprecipitate KCNQ3, and vice versa. The heteromeric KCNQ2/KCNQ3 channel corresponds to neurone's M-channel. They are widely distributed throughout the brain, being expressed at high levels in hippocampus, chordate nucleus, and amygdala. Mutations in *KCNQ2* and *KCNQ3* rise to a form of idiopathic epilepsy known as benign familial neonatal convulsions.

K$_v$ CHANNELS

The vertebrate fast voltage-gated K$^+$ delayed rectifier channels fall into a variety of families (K$_v$1–K$_v$11) according to their amino acid sequence similarity (Figure 2). A channel from a fruit fly *Drosophila* was the first K$^+$ channel to be cloned (denominated in vertebrates as K$_v$1.1, the first number denoting subfamily and the second, order of discovery). The identification started with mutant flies called shakers, which shake their legs while under ether anesthesia. These flies lack a fast transient K$^+$ current in presynaptic terminals, so repolarization of the action potential is delayed and the evoked release of neurotransmitter from a single action potential becomes enormous.

K$_v$ channels are found in a wide range of different tissues and they have many different functions. Some ideas of the diversity of roles of K$_v$ channels are shown in Figure 2. For example, K$_v$3.1 channels are found in some neurons that are specialized to fire very short action potentials at high rates, such as those of the auditory system; K$_v$1.3 channels are found in leukocytes playing a role in inflammatory response; K$_v$1.2 channels are found in smooth muscle participating in contractil tone regulation; and K$_v$4.2 channels are expressed in myocardium and their function is action potential prolongation in heart.

Defects in K$_v$ channels function therefore may have profound physiological effects. Linkage studies have identified the gene responsible for episodic ataxia type-1. This gene encoded for K$_v$1.1 channel, which is expressed in the synaptic terminals and dendrites in brain neurones.

SEE ALSO THE FOLLOWING ARTICLES

Ion Channel Protein Superfamily ● Voltage-Sensitive Ca^{2+} Channels

GLOSSARY

delayed rectifier A K$^+$ channel that changes the membrane conductance with a delay after a depolarizing voltage step.

gating The opening or closing of a channel in response to some stimulus.

inward rectifier A K$^+$ channel that opens when the membrane is hyperpolarized.

long QT syndrome An inherited cardiac arrhythmia.

voltage sensor A channel structure able to detect changes in membrane potential. In voltage-dependent ion channels the voltage-sensing elements are located in the S4 segment.

FURTHER READING

Hille, B. (2001). *Ions Channels of Excitable Membranes*. Sinauer Associates, Sunderland, Massachusetts.

Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chalt, B. T., and MacKinnon, R. (2003). X-ray structure of a voltage-dependent K$^+$ channel. *Nature* **423**, 33–41.

Lainé, M., Papazian, D. M., and Roux, B. (2004). Critical assessment of a proposed model of Shaker. *FEBS Lett.* **564**, 257–263.

Orio, P., Rojas, P., Ferreira, G., and Latorre, R. (2002). New disguises for and old channel: MaxiK channel β-subunits. *News Physiol. Sci.* **17**, 156–161.

BIOGRAPHY

Ramon Latorre is Professor and Chair of the Laboratory of Biophysics and Molecular Physiology, CECS. He holds a Ph.D. from the University of Chile and received his postdoctoral training at the Biophysics Laboratory in the National Institutes of Health, Maryland. He is a Foreign Associate of the National Academy of the United States of America. The main scientific interest of Dr. Latorre is the molecular workings of ion channels.

Francisco Morera holds a degree in biochemistry from the Catholic University, Chile and is at present a Ph.D. in the Austral University, Valdivia, Chile.

Voltage-Sensitive Ca^{2+} Channels

Harald Reuter
University of Bern, Bern, Switzerland

Ion channels are integral, pore-forming transmembrane proteins that serve to pass ions across biological membranes. The channels differ in their ion selectivity, i.e., each type allows certain ion species to permeate through the open channel pore. Conformational changes of the proteins leading to openings or closings of the channels are called "gating." It is achieved either by binding of ligands to special sites on the channels (ligand-gated ion channels), or by changes in membrane potential (voltage-gated ion channels). Calcium (Ca^{2+}) channels are voltage-gated and have a high selectivity for Ca^{2+} ions. Opening of these channels by membrane depolarization leads to an influx of Ca^{2+} ions into the cell. The resulting increase in the intracellular Ca^{2+} concentration initiates essential physiological functions, such as cardiac contraction, secretion of hormones and neurotransmitters, gene transcription, or repetitive electrical activity of neurons and the heart. Most of these functions result from binding of Ca^{2+} to specific intracellular proteins (e.g., calmodulin, troponin C) which subsequently mediate the Ca^{2+}-dependent cellular responses. The diversity of voltage-gated Ca^{2+} channels in surface membranes of different cells is important for initiation of specific signaling cascades within the cells. The various types of Ca^{2+} channels are encoded by different genes and show tissue-specific expression. They are targets for toxins and for drugs used for therapeutic purposes.

Molecular Properties

STRUCTURES

Voltage-gated Ca^{2+} channels are oligomeric protein complexes composed of the voltage-sensitive, pore-forming α_1-subunits, together with auxiliary disulfide-linked $\alpha_2\delta$-subunits and intracellular β-subunits (Figure 1). An additional transmembrane γ-subunit is associated with α_1S in skeletal muscles. The diversity of Ca^{2+} channel types not only results from different genes coding for the different channel subunits (10 for α_1, 4 for $\alpha_2\delta$, and 4 for β), but also from multiple splice variants of the subunits. The α_1-subunits consist of four homologous domains (I–IV), each being composed of six transmembrane α-helical segments (S1–S6). The N and C termini are located intracellularly, and the domains are linked together by intracellular loops.

In the folded structures of α_1-subunits, hairpin loops between transmembrane segments S5 and S6 constitute the channel pore. Segments S4 contain highly conserved, positively charged amino acids that act as voltage-sensitive sensors during gating of the channels. The β-subunits are involved in the regulation of the channels' gating properties and, like $\alpha_2\delta$-subunits, play a role in the functional expression of the channels in cell surface membranes.

SELECTIVITY

Ca^{2+} channels are selective pathways for the movement of Ca^{2+} ions down their electrochemical gradient into the cell. The pore-forming structures of the channel contain four negatively charged glutamate residues, each being in a homologous position on the four domains. The resulting ring of negative charges facing the pore coordinates binding of an entering Ca^{2+} ion. A second Ca^{2+} ion that enters the channel reduces the binding affinity of the first one by electrostatic repulsion, thus kicking it off from its binding site and letting it move down its electrochemical gradient. Thus, the selectivity of the channel for Ca^{2+} is critically dependent on the correctly placed negative charges of glutamate in the pore.

CALMODULIN BINDING

Two forms of autoregulation, Ca^{2+}-dependent inactivation and facilitation of Ca$_v$1.2 channels (Table I), involve calmodulin (CaM). The subunit α1C of Ca$_v$1.2 contains a consensus CaM-binding isoleucine-glutamine (IQ)-motif in its cytoplasmic C-terminal tail. The IQ-motif consists of 12 amino acids. Extensive mutations within this sequence revealed structural determinants for the two opposing forms of channel regulation. Two short stretches of aminoacids upstream of the IQ-motif seem to be sites for constitutive, Ca^{2+}-independent binding of apoCaM. The IQ-motif binds CaM only in the presence of Ca^{2+}. It represents the regulatory site for Ca^{2+}-dependent inactivation, while facilitation may require additional activation of Ca/CaMkinaseII.
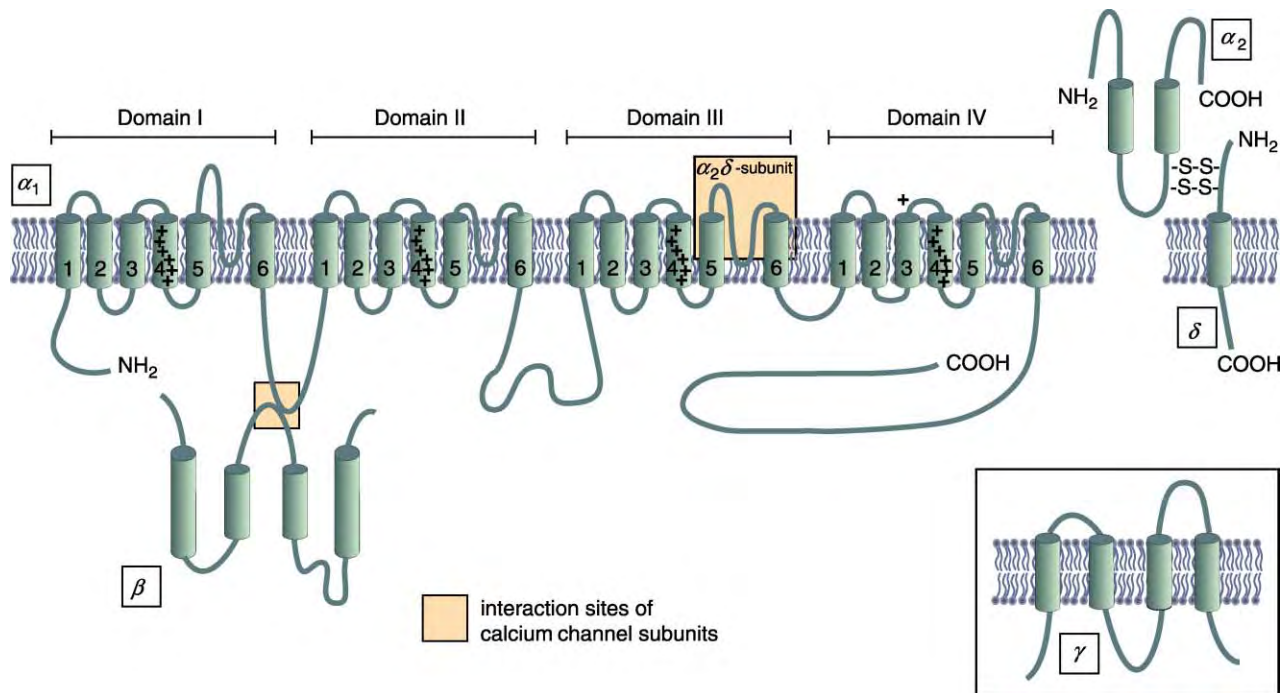
FIGURE 1   Oligomeric subunit complex of voltage-gated Ca^{2+} channels. The subunit composition consists of the pore-forming α1-subunit with its four homologous domains, the intracellular β-subunit, and the disulfide-linked α2-subunit. The sites of interaction between the subunits are indicated by the squares. An extra γ-subunit is present in skeletal muscle Ca$_v$1.1 complex (inset). Kindly prepared by Dr. R. D. Zuehlke, Bern.

PHOSPHORYLATION

Ca^{2+} channel function can be modulated by second messenger-activated protein kinases (PKs) such as cyclicAMP-dependent PKA, diacylglycerol-dependent PKC, or Ca/calmodulin-dependent kinases. Although the exact molecular mechanisms of these modulatory pathways are not yet fully understood, they do depend on phosphorylation by kinases of specific aminoacid residues of the Ca^{2+} channel protein. Multiple intracellular sites of phosphorylation have been identified on $\alpha2_1$- and β-subunits. However, which phosphorylation sites are necessary and sufficient for different forms of modulation of Ca^{2+} channel function by the various protein kinases has not been clearly resolved.

G PROTEINS

Ca^{2+} channels of the Ca$_v$2 family (Table I) that are involved in neurotransmitter and hormone secretion are inhibited by G proteins. There is excellent evidence that the rapid inhibition of the channels results from direct binding of the $\beta\gamma$-subunits of G proteins to an identified segment of the intracellular loop between domains I and II (Figure 1), but other sequences in the C and N termini of the $\alpha2_1$-subunits may also be important. Rapid inhibition of the channels by this mechanism does not require formation of soluble intracellular messengers.

The inhibition results from a membrane-delimited pathway leading to the release of the $\beta\gamma$-subunit of the G-protein complex.

Ca^{2+}-Currents

PHYSIOLOGY

Depolarization of cell membranes during excitation leads to spontaneous openings (activation) and closings (inactivation) of Ca^{2+} channels. At a given voltage, stochastic openings of multiple channels produce a considerable flow of Ca^{2+} ions into a cell and can thus be measured as an electrical current. The currents are recorded by electrophysiological methods, at a level of resolution where even the current through an individual channel can be measured (patch-clamp method). This allows detailed quantitative analyses of the kinetic properties (gating) and selectivities of the channels. Currents through different types of Ca^{2+} channels have traditionally been classified as L-, P/Q-, N-, R-, and T-type (Table I). This classification resulted from functional and pharmacological analyses before the channels were cloned and sequenced. A general overview of the distribution of the channels in different tissues is summarized in Table I. The specific expression of different Ca^{2+} channel types in different tissues is of

TABLE I

Ca²⁺ Channels

Channel types[a]	Pore-forming α-subunits[b]	Tissue distribution	Pharmacology (blockers)
HVA			
Ca$_v$1.1 (L)	α1S	Skeletal muscle	DHPs
Ca$_v$1.2	α1C	Heart, smooth muscle, CNS	
Ca$_v$1.3	α1D	CNS, kidney, cochlea, glands	
CA$_v$1.4	α1F	Retina	
Ca$_v$2.1 (P/Q)	α1A	CNS	ω-Agatoxin
Ca$_v$2.2 (N)	α1B	CNS	ω-Conotox.GVIA
Ca$_v$2.3 (R)	α1E	CNS, heart, glands	
LVA			
Ca$_v$3.1 (T)	α1G	CNS, sinus node	Mibefradil
Ca$_v$3.2	α1H	CNS, heart, kidney, liver	
Ca$_v$3.3	α1I	CNS	

    [a] Channel types: HVA = high-voltage-activated channels; LVA = low-voltage-activated channels; letters in brackets refer to the nomenclature of currents carried through the different channel types.
    [b] Pore-forming α_1-subunits encoded by identified genes have multiple splice variants (not shown); CNS = central nervous system; DHPs = Dihydropyridines.

great physiological and pharmacological importance. The localization of individual channels in association with other cellular constituents governs specific signaling pathways. For example, P/Q-,and N-type channels are directly linked to the SNARE (Soluble *N*-ethylmaleimide sensitive fusion protein Attachment protein REceptor) protein complex in neural and endocrine cells that is essential for secretion of neurotransmitters and hormones. L-type channels that are predominant in cardiac cells, are in close proximity to intracellular Ca²⁺ stores. When Ca²⁺ ions move through the channels into these cells during excitation, they release more Ca²⁺ from the stores which eventually leads to contraction of the heart. Thus, changes in the intracellular Ca²⁺-concentration are coding for important information of diverse cellular functions. Autoregulatory mechanisms, by which the channels inactivate in voltage- and Ca²⁺-dependent manners, prevent Ca²⁺ overload of the cells. Ca²⁺-dependent inactivation has been studied most extensively in L-type Ca²⁺ channels, but may also apply to other channel types. During the flow of Ca²⁺ ions through the pore, they bind to constitutively tethered apoCaM at the cytoplasmic C terminus of the channel. Occupancy of CaM by Ca²⁺ triggers binding of its C-terminal lobe to the nearby IQ-motif and accelerates inactivation of the channel, possibly by disinhibiting voltage-dependent inactivation (Figure 2).
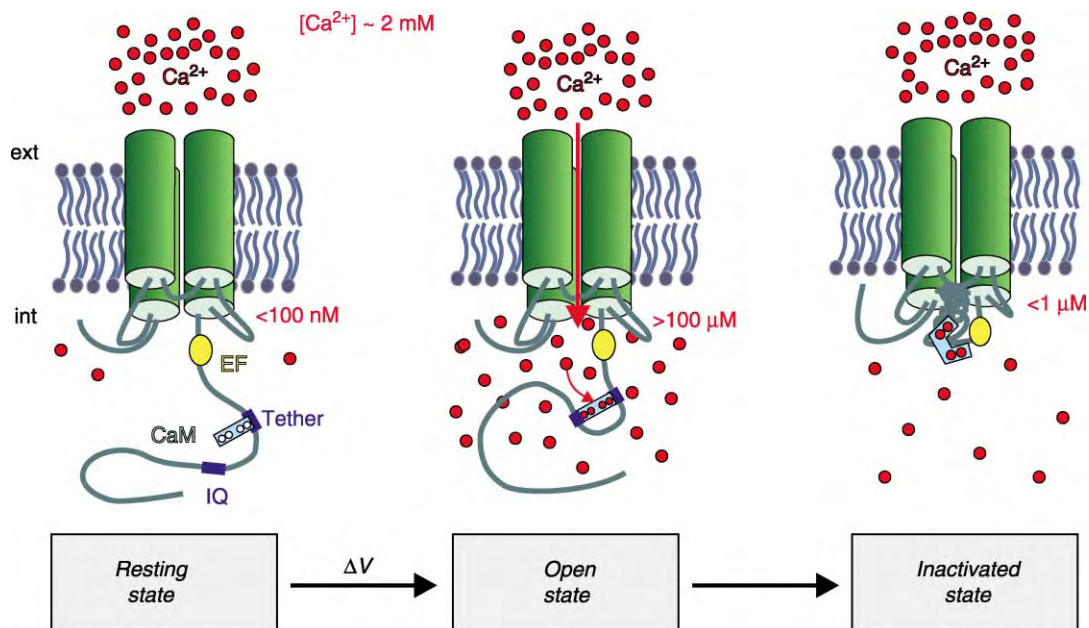
FIGURE 2   Model of voltage- and Ca²⁺-dependent gating of a Ca²⁺ channel. The model shows a schematic folding of the α1-subunit in Figure 1. At membrane resting potential the pore is closed (resting state); upon depolarization (ΔV) the pore opens and Ca²⁺ ions move across the membrane down their electrochemical gradient (open state); when the intracellular Ca²⁺-concentration increases, Ca²⁺ binds to tethered calmodulin (CaM) which then attaches itself to the regulatory IQ-motif and promotes closure of the pore (inactivated state). Kindly prepared by Dr. R. D. Zuehlke, Bern.

MODULATION

Modulation of Ca^{2+} channel activity by neurotransmitters and hormones has become a major field of research. The up-regulation of cardiac Ca^{2+} current (Ca$_v$1.2) by epinephrine was the first example of modulation of an ion channel by a neurotransmitter. It has been shown to be the result of activation of the cyclicAMP pathway via β_1-adrenoceptors and subsequent phosphorylation by protein kinase A of aminoacid (serine) residues at intracellular sites of the channel. This causes enhanced probability of openings of the channels, resulting in largely increased Ca^{2+} currents. Other examples of modulation mainly concern N- and P/Q-type Ca^{2+} channels (Ca$_v$2.1 and Ca$_v$2.2; Table I). They are inhibited by neurotransmitters, such as norepinephrine. Release of G protein $\beta\gamma$-subunits by neurotransmitters shifts channel activation into a reluctant state, thus causing a reduction of Ca^{2+} currents during depolarization. These are just two examples of modulation of Ca^{2+} channel activities. There are several others that are not included under the scope of this article.

PHARMACOLOGY

Ca^{2+} channels are important targets for toxins and drugs. Although a full discussion of their effects is beyond the scope of this article, a few of them are listed in Table I. Dihydropyridines (e.g., nifedipine, nimodipine) are relatively specific inhibitors of the Ca$_v$1 channel family, mainly Ca$_v$1.2 type. Other drugs acting on these channels are the phenylalkyamines verapamil and the benzothiazepine diltiazem. They are grouped together as "Ca-channel blockers," and are therapeutically used for the treatment of cardiovascular diseases, notably hypertension and angina pectoris. The toxins have no, or very limited, therapeutic use. In some cases of otherwise untreatable pain, the N-type channel blocker ω-conotoxin GVIA has been successfully applied. Because of unwanted side effects, the T-type channel blocker mibefradil has been withdrawn from therapeutic use, although it was quite effective in the treatment of essential hypertension.

SEE ALSO THE FOLLOWING ARTICLES

Calcium Oscillations • Calcium Transport in Mitochondria • Calcium/Calmodulin-Dependent Protein Kinase II • Intracellular Calcium Channels: NAADP$^+$-Modulated • Ion Channel Protein Superfamily • Store-Operated Membrane Channels: Calcium

GLOSSARY

apoCaM Ca^{2+}-free calmodulin.

electrical activity Results from ion flow through channels in membranes.

electrochemical gradient Driving force for ions due to differences in concentrations and electrical potentials across membranes.

G proteins Small intracellur proteins involved in signaling pathways.

oligomeric proteins Composition of several protein subunits.

protein kinases Enzymes that catalyze phosphate transfer from ATP to proteins; protein kinases are activated by distinct second messengers.

second messengers Intracellular molecules formed after receptor activation and involved in signaling (e.g. cyclic-AMP, diacylglycerol).

signaling cascade Information transfer within a cell.

FURTHER READING

Catterall, W. A. (2000). Structure and regulation of voltage-gated Ca^{2+} channels. *Annu. Rev. Cell Dev. Biol.* **16**, 521–555.

Hille, B. (2001). *Ionic Channels of Excitable Membranes.* Sinauer, Sunderland, MA.

McDonald, T. F., Pelzer, S., Trautwein, W., and Pelzer, D. J. (1994). Regulation and modulation of calcium channels in cardiac, skeletal, and smooth muscle cells. *Physiol. Rev.* **74**, 365–507.

Perez-Reyes, E. (2003). Molecular physiology of low-voltage-activated T-type calcium channels. *Physiol. Rev.* **83**, 117–161.

Reuter, H. (1996). Diversity and function of presynaptic calcium channels in the brain. *Curr. Opin. Neurobiol.* **6**, 331–337.

Saimi, Y., and Kung, C. (2002). Calmodulin as an ion channel subunit. *Annu. Rev. Physiol.* **64**, 289–311.

BIOGRAPHY

Harald Reuter is Professor Emeritus and former Director of the Institute of Pharmacology at the University of Bern, Switzerland. He received his M.D. at the University of Mainz, Germany. Dr. Reuter first described Ca^{2+} currents and the sodium/calcium-exchange transport in the heart, and discovered regulation of Ca^{2+} channels by catecholamines and calmodulin. He is further interested in synaptic function in the brain. He has received many awards for his discoveries and has been elected a Foreign Associate of the National Academy of Sciences, USA.

Voltage-Sensitive Na$^+$ Channels

William J. Brammar
University of Leicester, Leicester, UK

Voltage-sensitive sodium channels are the transmembrane protein complexes that mediate the increased permeability to Na$^+$ ions during the initial rising phase of the action potential in most excitable cells. The channels are closed at resting membrane potentials, but open in response to membrane depolarization to become selectively permeable to Na$^+$ ions. The open or activated state is transitory: channels are inactivated and Na$^+$ permeability returns to the baseline level within about a millisecond, allowing the repolarization phase of the action potential that is controlled by K$^+$ channels.

A family of genes encoding multiple sodium channel isoforms has been revealed by molecular cloning. Mutations in genes encoding Na$^+$ channel subunits give rise to genetic diseases affecting the function of skeletal muscles, the heart or the brain. Sodium channel proteins include target-sites for the action of a number of neurotoxins, insecticides, and local anesthetics.

Isolation and Purification of Sodium Channel Proteins

The first biochemical identification of sodium channel proteins relied on specific covalent labeling of the proteins with a photoreactive derivative of the north African scorpion toxin. Denaturing polyacrylamide gel electrophoresis of proteins derived from rat brain synaptosomes revealed two labeled polypeptides of 260 and 36 kDa that were subsequently designated as the α- and β1-subunits of the sodium channel. Purification of the sodium channel solubilized from rat brain revealed three polypeptides, α of 260 kDa, β1 of 36 kDa and β2 of 33 kDa, in a 1:1:1 stoichiometry. The β1-subunit is non-covalently associated with the α-subunit, whereas the β2-subunit is covalently attached via disulfide bonds.

The sodium channels purified from rat skeletal muscle sarcolemma or transverse tubules consist of a 260 kDa subunit and one or two subunits of 38 kDa. In contrast, those purified from eel electroplax and chicken heart contain only the 260 kDa component. Reconstitution experiments, in which the purified proteins were incorporated into phospholipid vesicles to generate fully functional, neurotoxin-sensitive sodium channels, proved the validity of the protein purification methods based on neurotoxin-binding.

Primary Structures of Sodium Channel Proteins

THE α-SUBUNITS OF SODIUM CHANNELS

Sequences encoding a complete Na$^+$ channel protein were first cloned as cDNAs derived from the mRNA of the electroplax of the electric eel. The deduced sequence of 1820 amino acids contains four internally homologous domains, each having six putative membrane-spanning α-helical segments (Figure 1). In each homology unit, five segments (S1, S2, S3, S5, and S6) are predicted to be hydrophobic, membrane-spanning helices. The interspersed S4 transmembrane segments contain positively charged amino acids located at every third position, forming an α-helix with a spiral of positive charges.

The suggestion that the four S4 segments act as voltage-sensors for the gating of the Na$^+$ channel has been confirmed by mutational alteration of the positively charged residues. The absence of an N-terminal signal sequence suggests that the N terminus is cytoplasmic.

The structural motif of four internally homologous domains, first promulgated for the sodium channel of the eel electroplax, is also evident in the voltage-gated calcium and potassium channels and in other channels within the large superfamily. Discernable sequence similarity ($\sim 32-37\%$ identity) between the four homologous domains of the voltage-gated Na$^+$ channels and those of the voltage-gated Ca^{2+} channels reflects the common ancestry of these two channel types.

Sequences encoding mammalian sodium channel α-subunits have been isolated using segments of the eel electroplax Na$^+$ channel cDNA as hybridization probes and by screening expression libraries with antibodies prepared against purified components. DNA sequences encoding regions of brain Na$^+$ channels have been used in low stringency hybridizations to
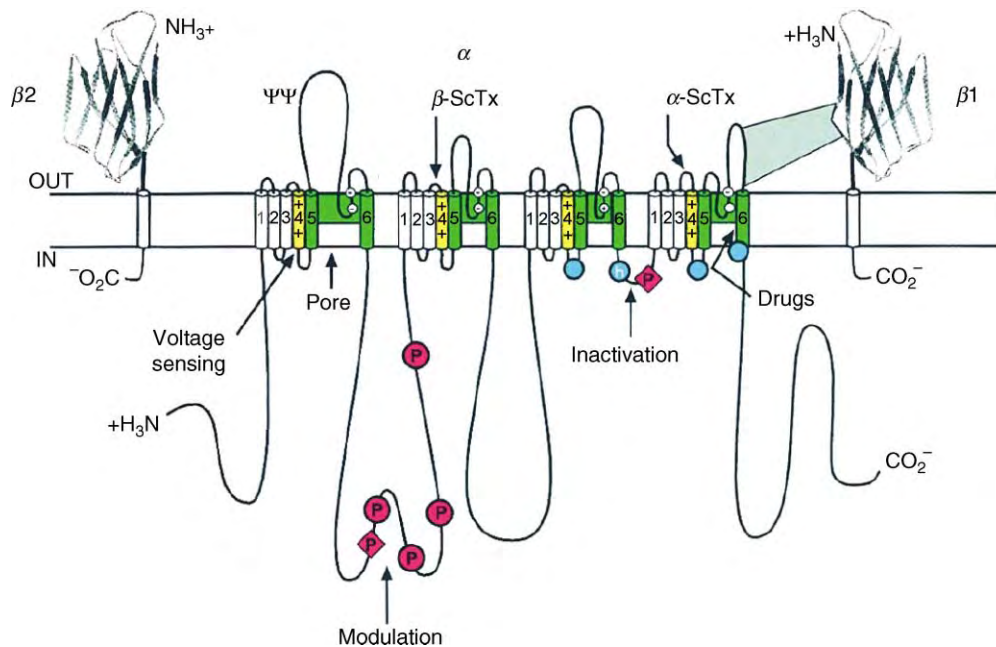
FIGURE 1   Primary structures of the subunits of the voltage-gated sodium channel. The α, β1, and β2 polypeptide chains are shown to illustrate their relationship to each other and to the cell membrane. The cylinders represent putative α-helices. The lengths of the intracellular and extracellular loops are roughly proportional to the number of amino acid residues in the brain sodium channel subtypes. The extracellular domains of the two β subunits are drawn as immunoglobulin-like folds. Ψ shows the site of probable N-linked glycosylation: P in red circles, sites of phosphorylation by (circles) and PKC (diamonds); the green regions are pore-lining segments; white circles represent the outer and inner rings of amino acid residues that form the ion selectivity filter and the tetrodotoxin-binding site: yellow shows the S4 voltage sensors; h in a blue circle represents the inactivation particle and blue circles the sites forming the receptor for the inactivation gate. Binding sites for α and β scorpion toxins (α-ScTx and β-ScTx) are shown by the downward arrows. (Reproduced from Catterall, W. A. (2000). From ionic currents to molecular mechanism: The structure and function of voltage-gated sodium channels. *Neuron* **26**, 13–25, with permission from Elsevier).

isolate cDNAs encoding α-subunits from other tissues such as skeletal muscle and heart. Although the different channel α-subunits vary slightly in length and sequence, they are clearly homologous and share fundamentally similar structures.

THE β-SUBUNITS OF SODIUM CHANNELS

Both the β1 and β2 subunits of the voltage-sensitive Na$^+$ channel are predicted to contain a single transmembrane region, located near the intracellular C terminus, and an extracellular domain containing a single "immunoglobulin-like fold" (Figure 1). The latter consists of two superimposed β-sheets held together by hydrophobic interactions and is characteristic of the family of cell adhesion molecules that mediate cell–cell interactions. In such molecules, immunoglobin-like folds bind extracellular proteins and influence interactions with the extracellular matrix and with other cells to affect cell shape and mobility. Rat brain Na$^+$ channels and the β2-subunit have been shown to bind to the extracellular matrix proteins tenascin-C and tenascin-R, leading to the suggestion that such interactions may influence the formation of specialized areas of high Na$^+$ channel density such as nodes of Ranvier.

Heterologous expression of cDNAs encoding Na$^+$ channel α and β subunits has shown that the β1 and β2 subunits modulate channel-gating behavior. Expression of the unaccompanied α-subunit of brain or skeletal muscle Na$^+$ channel in *Xenopus* oocytes produces channels that activate and inactivate more slowly than the native channels. Coexpression of β1 and β2 subunits with the α-subunit yields channels that gate normally, and the accelerating effect of the β1-subunit is mediated by the immunoglobulin-like fold in the extracellular domain.

The Molecular Basis of Sodium Channel Function

THE OUTER PORE AND SELECTIVITY FILTER

The guanidinium-containing compounds tetrodotoxin and saxitoxin, selective and reversible blockers of Na$^+$ channels at nanomolar concentrations, have been influential in allowing identification of the outer pore and selectivity filter of the channel. These toxins are membrane-impermeant and act only from the extracellular side of the membrane to plug the selectivity filter in the outer pore of the channel. Mutational studies have identified a pair of amino acids, mostly negatively charged, in analogous positions in all four domains of

the rat brain Na$^+$ channel α-subunit, that are important for binding tetrodotoxin and saxitoxin (Figure 1). These four pairs of amino acids were postulated to form outer and inner rings that acted as both the receptor-site for tetrodotoxin and saxitoxin and the selectivity filter in the outer pore of the channel. The concept is strongly supported by the finding that replacement of the four amino acid residues in the inner ring by four glutamic acid residues, their counterparts in voltage-gated Ca^{2+} channels, made the Na$^+$ channels calcium-selective.

Cardiac Na$^+$ channels bind tetrodotoxin with an affinity that is 200-fold lower than that of the brain or skeletal muscle channels, because of a replacement of an aromatic amino acid by a cysteine residue in the outer pore region of domain I. Cadmium is a high-affinity blocker of cardiac Na$^+$ channels because of its interaction with this same cysteine residue. Electrophysiological measurements of the voltage-dependence of the cadmium block of cardiac Na$^+$ channels place the target cysteine residue, and thus the selectivity filter, about 20% of the way through the membrane electric field.

THE INNER PORE

Local anesthetics, such as lidocaine and procaine, are generally lipid-soluble tertiary amine compounds that inhibit propagated action potentials by blocking Na$^+$ channels. A quaternary amine analogue of lidocaine, QX-314, which is positively charged at all pHs and not lipid-soluble, blocks Na$^+$ channels only when applied to the inside of the membrane. Blockage requires prior opening of the channels by depolarization. Mutagenesis studies demonstrated that the binding-site for local anesthetics lies in the IVS6 transmembrane segment (Phe1764 and Tyr1771 in rat brain type IIA channels).

Similar amino acid residues in the S6 transmembrane segment of certain voltage-gated K$^+$ channels act as the targets for tetraethylammonium-based blockers. These amino acids are located in an aqueous cavity within the inner pore region of the K$^+$ channel, by extrapolation from a bacterial K$^+$ channel whose quaternary structure has been determined.

THE MECHANISM OF VOLTAGE-DEPENDENT ACTIVATION

The voltage-dependent activation of sodium channels involves the net outward movement across the membrane electric field of $\sim$12 electronic charges in the channel protein. The S4 segments of the α-subunits, with positively charged amino acids at every third position, were predicted to act as the voltage-sensor and to undergo outward movement on depolarization of the membrane, initiating a conformational change associated with pore-opening.

Mutagenesis studies showed that neutralization of the positive charges in the S4 domains reduced the steepness of the potential-dependence of activation, equivalent to reducing the gating charge. Physical movement of the S4 segments has been demonstrated by substituting cysteine for the basic amino acids, then assessing the availability of the individual cysteine-SH groups to chemical modification by membrane-impermeant sulfhydryl reagents, before and after depolarization. Three successive basic amino acids in the S4 segment of domain IV (Arg1448, Arg1451, and R1454) move from being inaccessible within the membrane to being available for reaction from outside the membrane.

Similar outward movements of positively charged S4 segments have been demonstrated for voltage-sensitive K$^+$ channels. In this case it has been shown by mutagenesis that the positive charges in S4 are paired with negatively charged residues in either S2 or S3 segments, stabilizing the S4 segments in the membrane.

THE BASIS OF INACTIVATION

Native sodium channels spontaneously inactivate within milliseconds of opening. The sensitivity of this rapid inactivation to limited cytoplasmic applications of proteases lead to the proposal of the 'ball and chain' hypothesis of inactivation, in which an intracellular inactivation gate ('the ball'), tethered by a flexible 'chain', was able to interact with the intracellular mouth of the pore and block the channel. The inactivation gate has been localized to a short intracellular loop connecting regions III and IV. Antipeptide antibodies directed against this sequence, cutting the loop in this region and mutagenesis of a hydrophobic triad of Ile, Phe, and Met in the region all prevent fast inactivation. Voltage-dependent movement of this segment has been detected by the depolarization-induced loss of accessibility to cytoplasmically applied reagents.

The receptor for the inactivation gate has been delineated by scanning mutagenesis experiments. Multiple amino acid residues at the intracellular end of transmembrane segment IVS6 and in intracellular loops IIIS4-S5 and IVS4-S5 contribute to the receptor for the inactivation the gate, allowing its binding to block the pore of the channel (Figure 1).

Modulation of Channel Activity

Sodium channels in neurons and skeletal muscle are susceptible to modulation by several protein kinases. Cyclic AMP-dependent protein kinase A (PKA) phosphorylates sodium channel α-subunits in brain synaptosomes and intact neurons on four sites in the intracellular loop between domains I and II. The consequence of phosphorylation by PKA is a reduction

in sodium conductance, with little effect on the voltage-dependence of activation and inactivation. Activation of adenylate cyclase via dopamine acting at D1-like receptors reduces action potential generation and peak sodium currents in hippocampal neurons. The presence of a protein kinase-anchoring protein, AKAP-15, which brings PKA close to the sodium channel, is essential for this regulation. Cocaine, which increases dopaminergic neurotransmission by inhibiting dopamine-uptake, also reduces sodium currents in nucleus accumbens neurons.

Protein kinase C (PKC) also acts on sodium channel α-subunits, slowing channel inactivation and reducing peak sodium currents. Phosphorylation of a site in the inactivation gate is responsible for the slowing of channel inactivation. The reduction in peak current is a consequence of phosphorylation at sites in the intracellular loop between domains I and II (see Figure 1).

The effects of PKA are enhanced by PKC-dependent phosphorylation and by steady membrane depolarization, so that the information from three distinct signaling pathways are integrated via the sodium channels in the hippocampus and CNS.

The Genetics of Sodium Channels

THE GENE–PROTEIN FAMILY OF SODIUM CHANNELS

A large family of sodium channel genes and proteins, containing at least ten members, has been recognized through molecular cloning and heterologous expression to make characterizable protein products. The nomenclature for sodium channels has been informal and inconsistent, but a standardized system has recently been proposed by a consortium of leading researchers (see Table I).

The nine isoforms of the sodium channel α-subunit that have been characterized are greater than 50% identical in the transmembrane and extracellular domains. The continuously variable pattern of sequences is consistent with their constituting a single family. The Na$_V$1.1, Na$_V$1.2, Na$_V$1.3, and Na$_V$1.7 α-subunits are the most closely related. These four channels are all highly tetrodotoxin-sensitive and expressed in the central and peripheral nervous system. The corresponding genes are all located on human chromosome 2q23–24. The Na$_V$1.5, Na$_V$1.8, and Na$_V$1.9 channel α-subunits, all forming tetrodotoxin-resistant sodium channels, are produced in heart and neurons of the dorsal root ganglia from genes located on chromosome 3p21–24. The skeletal muscle sodium channel isoform, Na$_V$1.4, formerly called type μ1, is at least 84% identical to the group of channels encoded on chromosome 2, but has a more distant phylogenetic relationship as determined by parsimony comparison.

Comparison across mammalian species shows that the chromosome segments containing sodium channel α-subunit genes are paralogous segments containing many sets of related genes, generated by large-scale duplication events during early vertebrate evolution.

Several sodium channel α-subunits have been recognized by molecular cloning and DNA sequencing, but have not yet been characterized by functional expression (Table I, Na$_X$). These sequences have more than 80% identity to each other and are phylogenetically closely related to the group of sodium channel subunits encoded on human chromosome 2q23–24, where the human SCN6A gene is located.

The genes encoding mammalian sodium channel α-subunits are complex, with more than 20 exons. Splice variants of several sodium channel proteins have been identified and more are likely to be found.

GENETIC DEFECTS AFFECTING VOLTAGE-GATED SODIUM CHANNELS

Mutations in genes encoding sodium channel subunits are associated with several inherited human diseases characterized by hypersensitivity, including several monogenic epilepsy syndromes. Mutations in the human SCN1A gene have been associated with familial generalized epilepsy with febrile seizures plus (GEFS+), an autosomal dominant syndrome characterized by febrile seizures (FS) and a variety of afebrile generalized seizure types. *De novo* mutations in SCN1A are a major cause of severe myoclonic epilepsy of infancy (SMEI or Dravet syndrome). Benign familial neonatal-infantile seizures have been ascribed to mutations in the SCN2A gene.

Mutations in the SCN4A gene encoding the skeletal muscle sodium channel α-subunit are the cause of the hereditary disorders of sarcolemmal excitability, hyperkalemic periodic paralysis (HYPP or hyperPP) and paramyotonia congenita (PMC). Mutations have been shown to cause reduced rates of channel inactivation, altered voltage-sensitivity or slowed coupling between activation and inactivation. HYPP is a muscle disorder that is prevalent in American quarter horses, where selective breeding for muscular definition and hypertrophy has favored spread of the mutant SCN4A gene in the gene-pool.

One of the causes of long QT syndrome, an inherited cardiac arrhythmia, is mutation in the SCN5A gene, leading to non-inactivating sodium currents that prolong the plateau of the cardiac action potential. The result is an elongated interval between the QRS complex and the T wave in the electrocardiogram. The mutations usually affect amino acid residues in the inactivation gate or its receptor region, reducing the rate of inactivation.

TABLE I

Genes Encoding Mammalian Sodium Channel α-Subunits

Gene	Location	Channel type	Tissue specificity	Associated disease
SCN1A	2q24	Na$_v$1.1	CNS, PNS	Epilepsy
SCN2A	2q23–24	Na$_v$1.2	CNS	Neonatal-infantile seizures
SCN3A	2q24	Na$_v$1.3	CNS	
SCN4A	17q23–25	Na$_v$1.4	Sk.m.	Myotonia (HYPP, PMC)
SCN5A	3p21	Na$_v$1.5	Heart, Uninnervated sk.m.	Long QT syndrome
SCN8A	12q13	Na$_v$1.6	CNS,PNS	Cerebellar ataxia[a]
SCN9A	2q24	Na$_v$1.7	Schwann cells	
SCN10A	3p22–24	Na$_v$1.8	DRG	
SCN11A	3p21–24	Na$_v$1.9	PNS	
SCN6A[b]	2q21–23	Na$_x$	Heart, uterus, sk.m.	

[a] This condition is associated with mutations in the mouse SCN8A gene.
[b] This gene was originally designated SCN6A in human and SCN7A in mouse. The two genes represent homologues and the SCN6A symbol will probably be deleted. CNS = central nervous system; PNS = peripheral nervous system; sk.m. = skeletal muscle; DRG = dorsal root ganglion neurons.

Mutations in the SCN5A gene can also contribute to idiopathic ventricular fibrillation (IVF), a common cause of sudden cardiac death. A variant SCN5A gene segregating with IVF in an affected family contained two missense mutations, affecting amino acid residues in the extracellular loops between transmembrane segments S1 and S2 of domain III and between S3 and S4 of domain IV. The latter extracellular loop is involved in the coupling between channel activation and fast inactivation. The mutations affect the voltage dependence of steady-state inactivation of the sodium channels, shifting it more than 10 mV towards more positive potentials and accelerating rates of recovery from inactivation. Frameshift mutations and splice-site mutations in SCN5A have also been shown to be associated with IVF in affected families.

Mutations of the mouse *Scn8a* gene of the mouse have been shown to cause motor endplate disease, a lethal skeletal muscle atrophy. Less severe phenotypes involving 'jolting', a rhythmic tremor of the head and body and unsteady gait, are associated with missense mutations in the *Scn8a* gene.

A mutation affecting the sodium channel β1-subunit has been shown to be the cause of inherited febrile seizures. The mutation alters a cysteine residue involved in stabilizing the immunoglobulin-like fold in the extracellular domain of the β1-subunit that is involved in modulating sodium channel gating.

Toxins and Channel Modifiers

The voltage-gated sodium channels are subject to modification by a wide range of toxins produced by animals, for use in predation or defence, or by plants as protection against herbivores. Six distinguishable toxin-binding sites have been identified on sodium channels, as summarized in Table II. The groups of toxins have various effects on sodium channels, from complete block to prolongation of activation.

The guanidinium-containing compounds, tetrodotoxin, isolated from the internal organs of the puffer fish and its relatives in the family *Tetraodontidae*, and saxitoxin, from marine dinoflagellates, act as selective, reversible blockers at nanomolar concentrations. These membrane impermeant toxins bind at the extracellular opening of the channel pore and cause prolonged block of nerve and muscle action potentials. Sodium channels in heart and sensory neurons of dorsal root ganglia are relatively insensitive to tetrodotoxin and saxitoxin. The peptide μ-conotoxins, isolated from the venom of the fish-hunting cone snails of the *Conus* genus, also block muscle action potentials by binding to the extracellular opening of the pore of the skeletal muscle sodium channel. The venom, which is injected via a harpoon-like disposable tooth, rapidly results in paralysis and death of the prey.

A group of lipid-soluble alkaloids, including batrachotoxin, aconitine, veratridine, and grayanotoxin I, cause hyperexcitability and cardiac arrhythmias by promoting prolonged opening of sodium channels and long-lasting membrane depolarization. Batrachotoxin, secreted by the skin of Columbian frogs of the genus *Phyllobates*, is used by indigenous natives to make poison arrows for hunting. Brevetoxin, a fused cyclic ether produced by the marine dinoflagellate, *Ptychodiscus brevis*, the organism responsible for 'red tides', acts like the alkaloids, but via a different binding site on the sodium channel.

The peptide toxins of the scorpion α-toxin family, produced by north African genera *Androctinus*, *Buthus*, and *Leirus* and the American genus *Centruroides*, act to

TABLE II

Toxin-Binding Sites on Voltage-Gated Sodium Channels

Site	Toxin	Source	Chemical type	Effect
1	Tetrodotoxin	*Tetraodontidae* ("puffer fish")	Heterocyclic, Guanidinium	Channel block, via binding site at the extracellular opening of the pore of the channel. Result in blockade of nerve and muscle action potentials
	Saxitoxin	Marine dinoflagellates		
	μ-Conotoxins	Fish-hunting cone snails of the *Conus* genus	Small peptides (∼17–22 aa)	
2	Veratridine	*Veratrum lilies*	Alkaloid	Slow inactivation, shift activation to more negative potentials and reduce selectivity. Result in hyperexcitability due to long-lasting membrane depolarization
	Batrachotoxin	*Phyllobates aurotaenia* (Columbian frog)		
	Aconitine	*Aconitum napellus* (monk's hood)		
	Grayanotoxins	Ericaceous plants	Diterpenoid	
3	Scorpion α-toxins	*Leiurus quinquestriatus*	Peptides, 64 aa	Slowing of inactivation and stabilization of the open state of the channel. Macroscopic effects include hyperexcitability and repetitive firing in motor units, accelerated respiration, convulsion, spastic paralysis, and eventual respiratory failure
	Sea anemone toxins	*Anemonia sulcata*	Peptides, 27–46 aa	
4	Scorpion β-toxins	*Centruroides* species	Peptides, 66 aa	Shift voltage-dependent activation in the positive direction. Toxin plus conditioning pre-pulse sift activation in the negative direction
				Hyperexcitability results, with heavy perspiration and tremor
5	Brevetoxins Ciguatoxins	*Ptychodiscus brevis* (marine dinoflagellate)	Cyclic polyethers	Inhibit inactivation and shift voltage-dependent activation to more negative potentials
6	Pyrethrins Pyrethroids	Chrysanthemums Synthetic	Organic esters	Slow activation and inactivation, prolong channel open-time and promote membrane depolarization, leading to lethal paralysis in insects

Data based on *Trends in Neurosciences, Neurotoxins Supplement*, 1996.

prolong the action potential duration in muscle and nerves by inhibiting sodium channel inactivation. The nonhomologous sea anemone toxins act at the same site and produce similar effects.

A second group of peptide-based scorpion toxins, the β-toxins, slow both activation and inactivation of sodium channels in skeletal muscle. The consequence of this modulation is that channels close very slowly at rest. The resulting inward Na$^+$ currents support repetitive firing in response to minimal stimulation, producing hyperexcitability, with heavy perspiration and muscle tremor.

The pyrethrins, natural insecticides produced by flowers of the genus *Chrysanthemum*, and their synthetic analogues the pyrethroids, are organic esters that prolong sodium channel activation and inhibit inactivation in nerve axons. The consequence of this effect is repetitive firing and membrane depolarization, leading to lethal paralysis in insects. The synthetic insecticide dichlorodiphenyltrichloroethane (DDT) acts in a similar fashion.

Several local anesthetics, such as lidocaine and procaine, are lipid-soluble tertiary amines that block sodium channels in nerve axons to inhibit propagation of action potentials. Various sodium channel blockers, including local anesthetics and alkaloids, are being used clinically as neuroprotective agents in ischemic stroke patients and as antiarrhythmic agents in heart disease.

SEE ALSO THE FOLLOWING ARTICLES

Ion Channel Protein Superfamily • Voltage-Dependent K$^+$ Channels • Voltage-Sensitive Ca^{2+} Channels

GLOSSARY

action potential   An electrical impulse arising from local changes in membrane permeability to Na$^+$ and K$^+$ ions that travels along nerve axons.

coexpression   Expression of two or more coding sequences, usually following transfection, in the same host cell.

depolarization Change in membrane potential from the normal resting potential of -70 mV towards more positive potentials.

domain A compact, independently folded region of a polypeptide chain.

gating The process of opening or closing an ion channel.

heterologous expression Expression of a cloned coding sequence in a host cell that would not normally express it.

modulation Change in ion conductivity brought about by external influences, such as hormone-mediated intracellular signalling events.

pore The ion-permeable pathway created by the ion channel protein that provides the route through the cell membrane.

selectivity filter The region of an ion channel that imposes the selectivity for one ion over another.

FURTHER READING

Aidley, D. J., and Stanfield, P. R. (1996). *Ion Channels: Molecules in Action*. Cambridge University Press, Cambridge, UK.

Catterall, W. A. (2000). From ionic currents to molecular mechanisms: The structure and function of voltage-gated sodium channels. *Neuron* **26**, 13–25.

Goldin, A. L., Barchi, R. L., Caldwell, J. H., Hofmann, F., Howe, J. R., Hunter, J. C., Kallen, R. G., Mandel, G., Meisler, M. H., Netter, Y. B., Noda, M., Tamkun, M. M., Waxman, S. G., Wood, J. N., and Catterall, W. A. (2000). Nomenclature of voltage-gated sodium channels. *Neuron* **28**, 365–368.

Hille, B. (2001). *Ion Channels of Excitable Membranes*. 3rd edition, Sinauer Associates, Sunderland, Massachusetts.

BIOGRAPHY

William J. Brammar is a Professor of Biochemistry and Pro-Vice-Chancellor (Research) at the University of Leicester. He has a long-term interest in the regulation of expression of both prokaryotic and eukaryotic genes. He holds a Ph.D. from University College, London and was a postdoctoral fellow at Stanford University. He has written texts on molecular cloning and ion channels and is a member of the European Molecular Biology Organization.

Von Hippel-Lindau (VHL) Protein

Ronald C. Conaway and Joan Weliky Conaway
Stowers Institute for Medical Research, Kansas City, Missouri, USA

The Von Hippel-Lindau (VHL) protein is a tumor suppressor. Mutations in the VHL protein can give rise to tumors of multiple organ systems, including the central nervous system, the endocrine system, and the kidney. The VHL protein functions as a subunit of a multiprotein ubiquitin ligase that negatively regulates expression of a large collection of hypoxia-inducible genes controlled by hypoxia-inducible transcription factors (HIFs). The VHL ubiquitin ligase prevents inappropriate expression of these hypoxia-inducible genes when cells are grown in a plentiful supply of oxygen by targeting HIFs for rapid ubiquitylation and degradation by the proteasome.

Clinical Consequences of VHL Mutations

Mutations in the *VHL* gene are found in a variety of human diseases, including VHL disease, sporadic clear cell renal carcinoma, sporadic hemangioblastoma, and congenital polycythemia. VHL disease is an autosomal dominant familial cancer syndrome. In VHL kindreds, inactivation of both copies of the *VHL* gene can give rise to a variety of highly vascularized tumors, including clear cell renal carcinoma, cerebellar hemangioblastoma and hemangioma, retinal angioma, and pheochromocytoma. *VHL* mutations are also responsible for the majority of cases of sporadic clear cell renal carcinoma and for sporadic hemangioblastoma. In addition, *VHL* mutations are believed to give rise to some forms of polycythemia, a disorder characterized by elevated levels of expression of the hypoxia-inducible *erythropoietin* gene.

Function of the VHL Protein

The VHL protein is present in eukaryotes from worms to mammals, where it performs an evolutionarily conserved function as a subunit of a multiprotein ubiquitin ligase that negatively regulates expression of a large number of hypoxia-inducible genes controlled by hypoxia-inducible transcription factors (HIFs). The expression of these hypoxia-inducible genes is repressed in normal cells grown in a plentiful supply of oxygen (normoxic conditions), but is strongly induced in cells starved for oxygen (hypoxic conditions). The VHL ubiquitin ligase prevents the inappropriate expression of these hypoxia-inducible genes when cells are grown in normoxic conditions by targeting HIFs for rapid ubiquitylation and degradation by the proteasome. When cells are starved for oxygen, ubiquitylation of HIFs by the VHL ubiquitin ligase is blocked. Under these hypoxic conditions, cellular HIF levels rise, and HIFs enter the nucleus, bind to hypoxia response elements (HREs) in the promoters of hypoxia-inducible genes, and activate their transcription and expression.

The cellular levels of HIFs correlate well with the levels of expression of hypoxia-inducible genes in both normoxic and hypoxic conditions (Figure 1). In normal cells growing under normoxic conditions, HIFs are present in cells at low levels, and expression of hypoxic-inducible genes is barely detectable. Under hypoxic conditions, HIFs are present at relatively high levels, and expression of hypoxia-inducible genes is robust. Under both normoxic and hypoxic conditions in cells lacking a functional VHL protein, HIFs are present in cells at high levels, and maximal expression of hypoxia-inducible genes is observed. Inappropriate overexpression of hypoxia-inducible genes in cells lacking a functional VHL protein may contribute to the pathology of VHL-associated diseases. Among these are the genes encoding vascular endothelial growth factor (VEGF), transforming growth factor α (TGFα), and erythropoietin (EPO). VEGF is a secreted protein that appears to play a major role in the vascularization of VHL tumors by promoting ingrowth into tumors of blood vessels needed to carry oxygen and nutrients to the tumor cells. TGFα is also a secreted protein that appears to function as an autocrine growth factor for some types of VHL tumor cells and may be responsible for promoting their uncontrolled proliferation. An elevated EPO level in the serum of polycythemic patients is a hallmark of the disease.

STRUCTURE OF THE VHL UBIQUITIN LIGASE

Purification of the VHL ubiquitin ligase revealed that it is composed of multiple proteins. In addition to the VHL
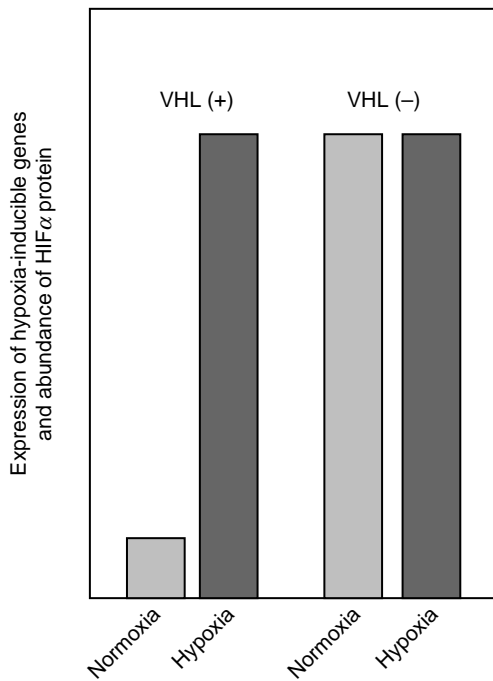
FIGURE 1   Regulation of expression of hypoxia-inducible genes and HIFα protein levels in normoxia and hypoxia. VHL (+), normal cells; VHL (−), cells lacking a functional VHL protein.

protein, the VHL ubiquitin ligase includes the elongin B and C proteins, cullin protein family member Cul2, and RING finger protein Rbx1 (Figure 2). In the VHL ubiquitin ligase, the VHL protein functions to recruit HIFs for ubiquitylation by interacting directly with them. The elongin B and C proteins bind stably to each other to form an adaptor that links the VHL protein to a Cul2-Rbx1 heterodimeric module. The Cul2-Rbx1 module functions to recruit the E2 ubiquitin conjugating enzyme Ubc5, which is responsible for ubiquitylation of HIFs. The VHL ubiquitin ligase is the founding member of a family of elongin BC-based ubiquitin ligases that
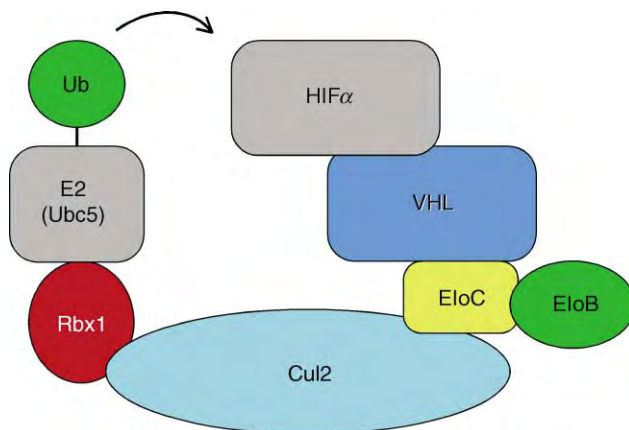
FIGURE 2   Architecture of the VHL ubiquitin ligase. Ub, ubiquitin; E2 (Ubc5), E2 ubiquitin conjugating enzyme Ubc5; EloB, elongin B; EloC, elongin C.

include Cul2-Rbx1 or Cul5-Rbx1 modules. In the VHL ubiquitin ligase, the VHL protein binds directly to elongin BC through a conserved elongin BC-binding site motif (BC-box) with consensus sequence [(A,P,S,T)LXXXCXXX(A,I,L,V)]. In other elongin BC-based ubiquitin ligases, the VHL protein is replaced by one of many known BC-box proteins, including members of the family of more than thirty suppressors of cytokine signaling (SOCS) proteins, which play important roles in a variety of signal transduction pathways.

REGULATION OF HIF UBIQUITYLATION BY THE VHL UBIQUITIN LIGASE

To date, three members of the HIF family of transcription factors have been identified in mammalian cells and shown to be regulated by the VHL ubiquitin ligase. HIFs are heterodimeric DNA-binding transcription factors that bind to HREs in the promoters of hypoxia-inducible genes and activate their expression. In mammalian cells, HIF heterodimers are composed of one of three related HIFα proteins and a common subunit, the aryl hydrocarbon receptor nuclear translocator protein (ARNT). The HIFα and ARNT proteins are basic helix-loop-helix (bHLH) transcription factors that all contain an additional conserved domain, referred to as the PAS domain (PER-ARNT-SIM domain), of unknown function. The HIFα subunits, but not the ARNT subunits, of HIFs are targets of the VHL ubiquitin ligase. Whereas the ARNT subunit of HIFs is present in normoxic and hypoxic cells at similarly high levels, the HIFα subunits are continuously synthesized but rapidly ubiquitylated by the VHL ubiquitin ligase and degraded by the proteasome in normal cells grown under normoxic conditions. In hypoxic cells, ubiquitylation of HIFα subunits is attenuated. Under these conditions, cellular HIFα levels rise, and HIFα subunits heterodimerize with ARNT, enter the nucleus, and activate expression of hypoxia-inducible genes.

HIFα ubiquitylation by the VHL ubiquitin ligase is regulated by oxygen. Elegant studies have revealed that binding of HIFα subunits to the VHL protein and their recruitment to the VHL ubiquitin ligase depends on hydroxylation of a specific HIFα proline in a HIFα region referred to as the oxygen-dependent degradation domain (ODD). The HIFα ODD is an approximately 20 amino acid HIFα region that is essential for interaction of HIFα with the VHL protein and ubiquitylation of HIFα by the VHL ubiquitin ligase. The ODD is a portable domain capable of promoting oxygen-dependent degradation of a variety of ODD-containing fusion proteins. Binding of HIFα to the VHL protein depends on hydroxylation of a proline in the ODD by an iron- and oxoglutarate-dependent dioxygenase of the EGL-9 protein family. The EGL-9 proline hydroxylase uses O_2 as the oxygen donor for hydroxylation of HIFα

subunits in a reaction with kinetics proportional to the concentration of cellular O_2. In cells grown in hypoxic conditions, HIFα hydroxylation is significantly reduced, and the nonhydroxylated HIFα subunits are no longer substrates for ubiquitylation by the VHL ubiquitin ligase. Thus, oxygen-dependent hydroxylation of HIFα serves as a convenient molecular switch for regulation of expression of hypoxia-inducible genes.

SEE ALSO THE FOLLOWING ARTICLES

Transforming Growth Factor-β Receptor Superfamily • Vascular Endothelial Growth Factor Receptors

GLOSSARY

dioxygenase   Enzyme that catalyzes the insertion of both oxygens of molecular oxygen, O_2, into substrate molecules.

erythropoietin   An acidic glycoprotein hormone that regulates red cell production by promoting erythroid differentiation and the initiation of hemoglobin synthesis.

hemangioblastoma   A highly vascularized tumor composed of capillaries and stromal cells.

hemangioma   A benign lesion originating from blood vessels.

pheochromocytoma   A tumor of the adrenal gland.

retinal angioma   A benign lesion found in the retina that originates from blood vessels.

vascular endothelial growth factor   A growth factor for vascular endothelial cells that promotes angiogenesis.

FURTHER READING

Conaway, R. C., and Conaway, J. W. (2002). The Von Hippel-Lindau tumor suppressor complex and regulation of hypoxia-inducible transcription. *Adv. Cancer Res.* **85**, 1–12.

Min, J. H., Yang, H., Ivan, M., Gertler, F., Kaelin, W. G., Jr., and Pavletich, N. P. (2002). Structure of an HIF-1alpha-pVHL complex: Hydroxyproline recognition in signaling. *Science* **296**, 1886–1889.

Pugh, C. W., and Ratcliffe, P. J. (2003). The Von Hippel-Lindau tumor suppressor, hypoxia-inducible factor-1 (HIF-1) degradation, and cancer pathogenesis. *Semin. Cancer Biol.* **13**, 83–89.

Semenza, G. L. (2001). Hypoxia-inducible factor 1: Oxygen homeostasis and disease pathophysiology. *Trends Mol. Med.* **8**, 345–350.

Stebbins, C. E., Kaelin, W. G., and Pavletich, N. P. (1999). Structure of the VHL-elongin C-elongin B complex: Implications for VHL tumor suppressor function. *Science* **284**, 455–461.

BIOGRAPHY

Ronald C. Conaway and Joan Weliky Conaway are Senior Scientists at the Stowers Institute for Medical Research. Research in their laboratory is aimed at contributing to an understanding of the molecular mechanisms governing the synthesis of eukaryotic messenger RNA. They hold Ph.D.s from the Stanford University School of Medicine.
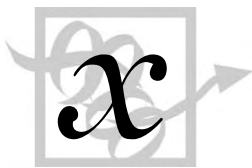
XPV DNA Polymerase and Ultraviolet Damage Bypass

Alan R. Lehmann
University of Sussex, Brighton, UK

DNA damage blocks the progress of the replication fork. In order to bypass the damage, cells have evolved special DNA polymerases, which are uniquely able to replicate damaged DNA. DNA polymerase η (eta) is the enzyme that carries out this bypass function for the major form of UV damage in DNA. A deficiency in this polymerase leads to a severe human condition, the variant form of xeroderma pigmentosum.

XP Variants

All cells sustain many kinds of DNA damage. This can be generated either endogenously from hydrolysis and oxidation reactions or exogenously from exposure to carcinogens. One of the most prevalent DNA damaging agents is UV radiation from the sun, which generates photoproducts in the DNA of cells in exposed areas of the skin. The most important photodimer, namely the cyclobutane pyrimidine dimer (CPD) and the pyrimidine 6-4 pyrimidone photodimer (6-4 PP) are removed from cellular DNA by the process of nucleotide excision repair (NER). The importance of this process can be seen from the features of patients with xeroderma pigmentosum (XP). XP patients show a wide variety of skin abnormalities on sun-exposed areas including freckling, hypo- and hyperpigmentation and ultimately multiple skin cancers. The incidence of skin cancer in XP individuals has been estimated to be 1000-fold higher than that in unaffected individuals. In the majority of XP patients this hypersensitivity to the effects of sunlight results from a genetic deficiency in NER. In ~20% of XP individuals however, NER is normal. These so-called XP variants (XPV) are deficient in their ability to produce intact daughter DNA strands during DNA replication after exposure of the cells to UV-irradiation. In other words, they are defective in their ability to replicate DNA damaged by ultraviolet light.

Polη and the Y-Family of DNA Polymerases

Although the cellular deficiency in XPV was identified in 1975, the molecular basis for the defect was not discovered until 1999. The gene defective in XPV cells was found to encode a new DNA polymerase designated DNA polymerase η or polη. Polη is a protein of 713 amino acids (aa) with no sequence similarity to previously discovered DNA polymerases, but it does share sequence similarity with several other DNA polymerases that were discovered at about the same time. This new group of polymerases was designated the "Y-family." They show sequence conservation over a region of ~400 aa, which contains the catalytic domain, and usually forms the N-terminal part of the protein. Y-family polymerases have a C-terminal extension, which is not conserved between family members.

CATALYTIC ACTIVITY

DNA polymerases that carry out the replication of undamaged DNA are very efficient and accurate, but they are blocked by most types of DNA damage including both major UV photoproducts. In contrast, polη is a rather poor polymerase on undamaged DNA. It dissociates after incorporating a few nucleotides and it is rather error-prone. However polη, in contrast to all other eukaryotic polymerases, is able to synthesize past a T-T CPD with the same efficiency as past undamaged T's. Furthermore in most instances the "correct" bases, namely two A's are inserted opposite the T's of the CPD. The reason for these properties became clear when the catalytic domains of several members of the Y-family (including yeast polη) were crystallized and their three-dimensional structures solved. Despite the lack of sequence similarity between classical and Y-family polymerases, the three-dimensional structures did have the three subdomains, designated finger, thumb and palm, found in all other DNA polymerases, with the Y-family having an extra subdomain designated
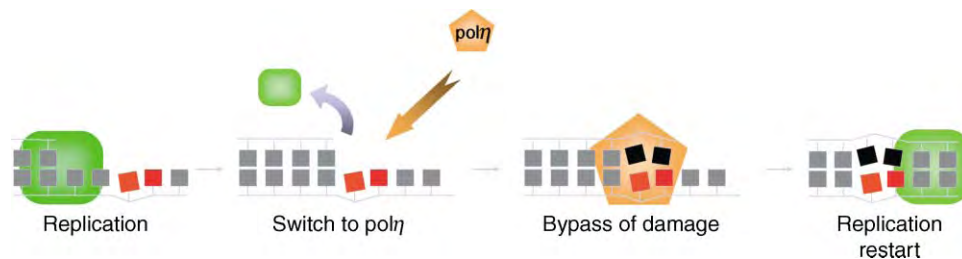
FIGURE 1   Translesion synthesis by polη. Replicative DNA synthesis is blocked by a CPD (indicated by red blocks). This leads to a switch to polη (step 1), which replicates past the CPD (step 2) and then dissociates, so that normal replication can restart (step 3).

pad or little finger. However the classical polymerases form a closed tight structure, which only permits normal bases into the active site. In contrast, polη has a much more open structure with two important consequences. Firstly, it has a relatively low stringency, so that incorrect bases can be incorporated rather easily. Secondly the open structure is able to accommodate the CPD in the template strand and the structure ensures that in general the correct bases are inserted opposite the bases in the CPD. Thus polη is tailor-made for replicating past CPDs. It can insert the appropriate bases opposite the damage and extend from these bases, but then it will soon dissociate from the template, so that the more accurate replicative polymerases can take over again. The ability to bypass DNA damage is termed translesion synthesis (TLS) and is shown schematically in Figure 1.

Apart from the CPD, polη is able to replicate past some other types of damage *in vitro*, but in most cases, it is much less efficient with other types of damaged templates than with undamaged or CPD-containing DNA. It is not yet known if it is involved in bypassing any of these other types of damage *in vivo*.

The importance of polη is seen in the severe clinical characteristics of XPV patients and in the very high UV-induced mutation frequency in XPV cells. The UV-mutability of XPV cells is several-fold higher than that of normal cells and the frequency of all types of mutations is increased. This is evidence that in the absence of polη, something else is able to do its job, but the replacement is less efficient and makes more mistakes.

LOCALIZATION

The catalytic domain of polη is contained within the first 430 aa. What then is the function of the C-terminal 280 aa? DNA replication takes place during the S phase of the cell cycle in "replication factories" inside the cell nucleus. These factories can be visualized either by immuno-fluorescence with an antibody to a replication protein, or by tagging proteins involved in DNA replication with the green fluorescent protein (GFP).

The factories appear as bright nuclear foci when viewed by fluorescence microscopy. Polη also localizes in replication factories during DNA replication, so that it is always on hand in case the replication machinery encounters DNA damage (Figure 2). It is the C-terminal 120 aa of polη (aa 595–713) that are responsible for this localization into replication factories. This C-terminal domain contains a nuclear localization sequence that is required for transporting polη into the nucleus, as well as a C_2H_2 zinc finger and a sequence found in proteins that bind to the replication accessory protein, PCNA. Both of these motifs are needed to localize polη in replication factories. The function of aa 430–595 in polη has not yet been determined but it is doubtless involved in interactions with other proteins. The domain structure of polη is summarized in Figure 3.
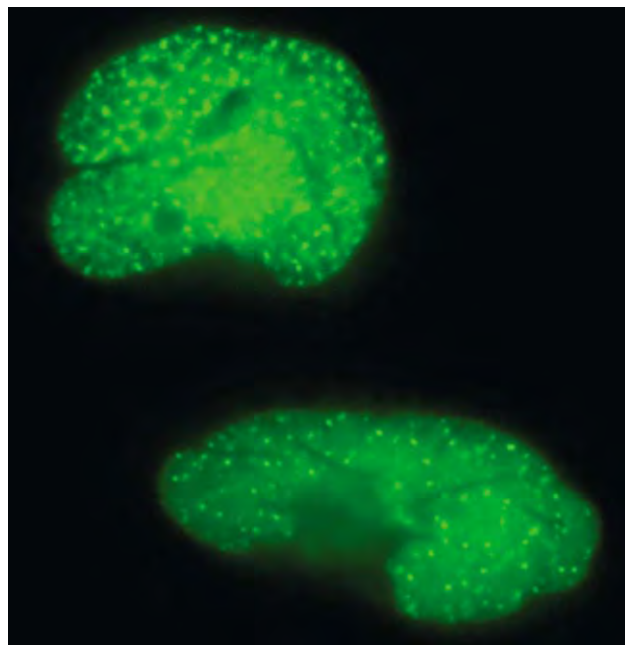
FIGURE 2   Localization of polη in replication foci. Polη, tagged with green fluorescent protein, localizes into nuclear foci during DNA replication.
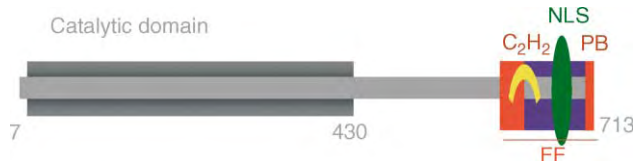
FIGURE 3 Domain structure of polη. C_2H_2, zinc finger; NLS, nuclear localization signal; PB, PCNA-binding motif; FF, foci formation.

Mutations in XPV Patients

Mutations have been identified in the polη gene in some 35 patients. The majority of the mutations result in truncations within the catalytic subunit and would not be expected to result in any functional polη. This suggests that polη is not an essential protein, since its absence is compatible with life. Several mis-sense mutations within the catalytic domain have also been identified and they are predicted either to inhibit binding to DNA or to interfere with the 3D structure of the protein. A small number of mutations do not affect the catalytic domain, but truncate the protein close to the C terminus. These mutant proteins would be expected to be fully active, but to lack the localization domain in the C terminus, which is required for transporting the polymerase into the nucleus. There is no clear relationship between the site or type of mutation and the severity of the clinical features, even though the severity of features differs widely between patients.

Replication Past Other Types of UV Damage

Although the CPD is the major UV photoproduct, the 6-4 PP is also formed in substantial quantities ($\sim 1/3$ the frequency of the CPD). No single polymerase has been found that is able to carry out TLS past a 6-4 PP, which produces a much greater distortion in the DNA than a CPD. However two polymerases acting in concert are able to accomplish this. Insertion of the nucleotides opposite the 6-4 PP can be carried out either by polη or its paralogue, polι, but neither of these can extend further from the incorporated nucleotides. However once nucleotides have been inserted opposite the damaged bases, they can be extended by DNA

polymerase ζ. So far, evidence for this dual polymerase mechanism has only been obtained *in vitro*. It is likely that this mechanism is also used *in vivo*, but as yet it is not known which polymerase carries out the insertion step.

SEE ALSO THE FOLLOWING ARTICLES

DNA Damage: Alkylation • DNA Oxidation • DNA Polymerase α, Eukaryotic • DNA Polymerase β, Eukaryotic • DNA Polymerase δ, Eukaryotic • DNA Polymerase ε, Eukaryotic • Nucleotide Excision Repair in Eukaryotes • Translesion DNA Polymerases, Eukaryotic • Zinc Fingers

GLOSSARY

DNA polymerase Enzymes which are able to synthesize new DNA molecules using a DNA template.

green fluorescent protein A highly stable intensely fluorescent protein, isolated from the jellyfish *Aequoria victoria*. It is widely used to "tag" other proteins to render them visible by fluorescence microscopy.

nuclear localization signal A sequence of amino acids in a protein that results in its localization in the nucleus.

nucleotide excision repair The enzymatic pathway by which many types of DNA damage are first recognized, then cut out of the DNA, and the excised region is replaced with new DNA.

zinc finger A sequence of amino acids which forms a finger-like structure and binds a zinc atom.

FURTHER READING

Goodman, M. F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* **71**, 17–50.

Lehmann, A. R. (2002). Replication of damaged DNA in mammalian cells: New solutions to an old problem. *Mutat. Res.* **509**, 23–34.

Prakash, S., and Prakash, L. (2002). Translesion DNA synthesis in eukaryotes: A one- or two-polymerase affair. *Genes Dev.* **16**, 1872–1883.

BIOGRAPHY

Alan Lehmann is a Professor of Molecular Genetics and Chairman of the Genome Damage and Stability Centre at the University of Sussex, Brighton, UK. His research interests are in DNA repair and its relation to human health and disease, in particular the disorders xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. He identified the cellular defects in Cockayne syndrome, trichothiodystrophy and the variant form of xeroderma pigmentosum. Translesion synthesis polymerases are one of his current major interests. He obtained his Ph.D. from the Institute of Cancer Research, University of London, UK.

X-Ray Determination of 3-D Structure in Proteins

Martha L. Ludwig
University of Michigan, Ann Arbor, Michigan, USA

X-ray crystallography can decipher the three-dimensional arrangements of atoms in biological macromolecules. The primary data for such structure analyses are the intensities of diffracted X-rays (reflections). Determination of the positions and displacement parameters of the atoms requires assignment of the relative phases of these reflections, a task that can be accomplished by any of several strategies, including multiple isomorphous replacement, multiwavelength anomalous diffraction, and molecular replacement. Initial models are subsequently refined to optimize agreement with the diffraction data. Using X-ray crystallography, it has been possible to determine the structures of very large particles such as spherical viruses and of machines such as ATPase and the ribosome at near-atomic resolution. Because crystallographic studies of proteins and of their molecular complexes provide fundamental information about proteins and enzymes, it is important to consider both the scope and the limitations of X-ray structure determination.

Stages in an X-Ray Structure Analysis

CRYSTALLIZATION

Crystals of proteins and other biological macromolecules are remarkable. They form well-ordered three-dimensional arrays but nevertheless contain a significant fraction of mobile solvent that fills the interstices between the irregularly shaped molecules. Obtaining suitable crystals is the largest hurdle in the structure analysis of biomacromolecules. Most important among the criteria for suitability is that crystals must diffract X-rays to a resolution (Figure 1) that allows construction of a reliable atomic model. The requirements for resolution are somewhat flexible and depend on the problem. Most enzymologists would consider 3.0Å to be the lowest acceptable resolution for studies relating structure to mechanism. The crystal itself also must be of sufficient size to allow accurate measurement of scattered X-rays (enough photons for an adequate signal:noise ratio). This criterion has become less stringent as intense beams available at synchrotrons allow the use of crystals as small as 10 μm in thickness.

The parameters that affect nucleation and growth of crystals have been the subject of both theoretical and empirical studies. Finding appropriate buffers, precipitants, additives, and temperatures for crystallization has increasingly been made systematic, and sparse sampling of these variables has been incorporated in the design of several commercial crystallization kits. Robots can be deployed to mix protein samples with crystallization reagents and to disperse samples into sitting drops of sizes that vary from 5–10 μl to 50 nl or less. Mutations of individual residues and expression of domains, truncated sequences, or sequences with deletions often yield proteins that will crystallize when the full-length or wild-type counterparts resist all attempts at crystallization.

DIFFRACTION MEASUREMENTS

The intensities of diffracted X-ray beams (reflections) are the fundamental measurements in a crystallographic experiment. Thanks to the very bright X-ray sources at synchrotrons and the advent of area detectors (imaging plates or charge-coupled devices) that allow many reflections to be recorded simultaneously, data collection is no longer the rate-limiting step in a structure analysis. At modern synchrotron facilities, complete data sets for routine problems can be obtained on a timescale of minutes.

X-irradiation leads to loss of crystalline order and resolution by rupturing chemical bonds and by secondary reactions of the protein with radical species such as OH· that are generated by irradiation of the solvent. Immobilization of solvent by flash-cooling of crystals to about 100 K minimizes the radical reactions. Development of apparatus and reagents for cryoprotection and low-temperature data collection has helped to revolutionize structural biology. Cryotechniques are particularly important for measurements of anomalous diffraction effects (see below) because crystal lifetimes
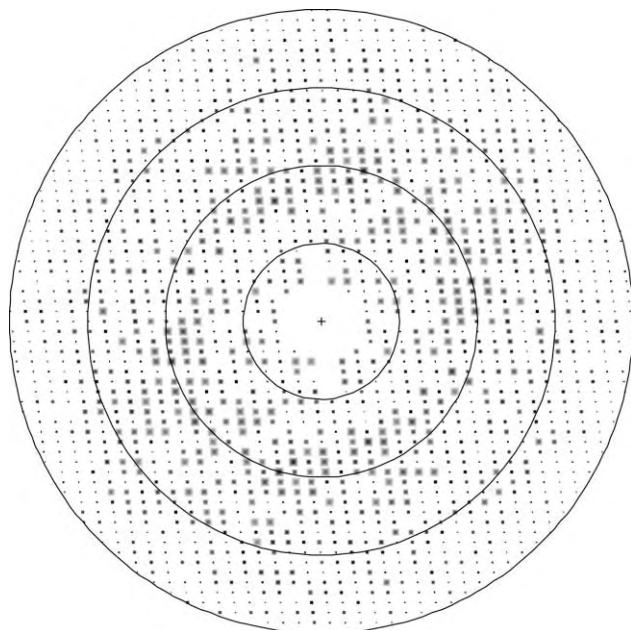
FIGURE 1 The definition of resolution in crystallography. This figure shows the relative intensities of diffracted X-rays (reflections) in a two-dimensional planar section through three-dimensional diffraction space. Each reflection occurs at a particular angle (2θ) determined by Bragg's law, $\lambda/d = 2\sin\theta$, where d is the distance between reflecting planes; reflections farther from the center have larger scattering angles and smaller d spacings. The superimposed circles are drawn with radii corresponding to resolutions of 7.2, 3.6, 2.4, and 1.8Å. All of the unique reflections within a sphere swept out by the 1.8Å circle will be used to calculate the electron density at a resolution of 1.8Å, whereas only the smaller number of reflections within the second sphere will be used for a 3.6Å electron density map. As the limiting d spacing becomes smaller, more data are incorporated in calculation of the electron density, and the resolution is said to be higher.

are extended sufficiently to permit all of the data required for structure determination to be collected from one crystal.

DE NOVO STRUCTURE DETERMINATION

The result of a structure analysis is a 3-D electron density map, which is computed from the X-ray scattering by Fourier transformation. This calculation requires not only the measured amplitudes (square roots of the intensities) of the scattered X-rays but also the assignment of phases to all of the reflected waves. These phases are not determined by the X-ray experiments. Two strategies, multiple isomorphous replacement (MIR) and multiwavelength anomalous diffraction (MAD), are commonly employed to assign phases for structure analysis of a new or unknown protein fold. In both strategies, contributions of added heavy atoms or of incorporated selenomethionine are used to calculate the phases of the protein by triangulation. The first step in both methods is determination of the positions of the heavy atoms. MIR requires measurements from native crystals and from a series of derivative crystals modified

by addition of heavy atoms. These derivatives are obtained by immersing preformed crystals in holding solutions containing heavy atom complexes, or less frequently, by co-crystallization. Typical reagents are mercurials or platinum complexes. The method assumes that X-ray scattering by the protein atoms is exactly the same in the derivative as it is in the native crystals so that all differences in the diffraction pattern can be attributed to the added heavy atoms. Perturbation of the protein induced by addition of heavy atoms (loss of isomorphism) is the most common source of errors in MIR.

Multiwavelength anomalous diffraction exploits scattering effects that occur at wavelengths near the X-ray absorption edges of elements such as Se, Br, and lanthanides. At the peak of the absorption edge of these elements, the imaginary anomalous component (f'') is large, producing differences between intensities of h, k, l and $-h, -k, -l$ reflections. The real anomalous component (f') is large (and negative) at the inflection point of an absorption edge, but is small at wavelengths remote from the edge. Effects of the real (dispersive) component appear as wavelength-dependent differences in intensities. Both f' and f'' vary with wavelength, and both components are used to determine phases. MAD experiments thus require X-rays that can be tuned to appropriate energies and are performed at synchrotron X-ray sources. Although the changes in intensities that arise from anomalous scattering effects are usually smaller than for MIR, MAD has the advantage that the requisite data are obtained from one crystal. Effects of nonisomorphism are thus circumvented. With precise measurements of the relevant intensity differences and algorithms based on direct methods, it is currently possible to locate ~ 100 Se atoms per asymmetric unit. Structures as large as 300 kDa have been determined from anomalous scattering by selenomethionine. Experience demonstrates that measurements of anomalous diffraction at the peak wavelength of lanthanides or even Se (single-wavelength anomalous diffraction, or SAD) can be sufficient to determine a structure.

In the determination of the phases that are used to calculate an initial experimental electron density map, it can be advantageous to combine data from isomorphous replacement and anomalous scattering. Requiring bulk solvent regions to have uniform electron densities (solvent flattening) is a useful restraint that improves the accuracy of phases. Computational suites such as CCP4, CNS, SHARP, or SOLVE/RESOLVE are designed to exploit all these and other sources of phase information.

FITTING ELECTRON DENSITY MAPS

Atomic models are built into electron density maps using graphics displays that allow interactive and menu-driven manipulations of the models. Modern software for model fitting incorporates a number of features that

simplify and speed this process, including local optimization of models within the density and sampling of likely side chain rotamers.

REFINEMENT

Initial atomic models that are constructed by interactive fitting of experimental electron density maps are usually incomplete and inaccurate. They serve as the starting point for refinement, a process in which the agreement between the observed and computed structure factors is improved by adjustment of the atomic positional and thermal parameters. Modified parameters are used in turn to compute phases and new maps that permit refitting of the model and addition of missing parts. This cyclic procedure is repeated until the fitting appears to have converged. It is a fortunate consequence of the physical and mathematical relationship between a crystal structure and its diffraction pattern that structures can be corrected or completed if most of the scattering has been accounted for in an atomic model. Missing and/or incorrectly placed atoms are detected in difference maps that subtract the computed from the observed structure. There are two fundamentally different approaches to optimization of molecular models: the first is straightforward minimization by algorithms such as conjugate gradient and the second is sampling of conformations, usually with molecular dynamics simulations. Molecular dynamics calculations that begin with atom velocities corresponding to elevated temperatures (simulated annealing) are a very efficient way to improve initial models. Standard minimizations are performed after the annealing cycles.

Minimizations are large, iterative least-squares computations whose convergence is hampered by small ratios of data to adjustable parameters. There are at least four parameters to be assigned to each atom: its position, x, y, z, and a temperature factor (or displacement parameter) that reports excursions of the atom about its equilibrium position. Even at 2.0Å resolution, a typical data:parameter ratio for this simplest set of parameters is about 2.5. To compensate for limitations in the data, most refinements are conducted with harmonic restraints that maintain standard polypeptide geometries and in effect increase the number of observations. Data:parameter ratios and robustness can also be increased by replacing Cartesian with torsional variables or by imposing strict constraints to maintain the identity of chains that are repeated in the asymmetric unit.

Various target functions can be chosen for minimization in least-squares refinement. The conventional target is the discrepancy between the observed amplitudes of reflections and the amplitudes calculated from the model. The corresponding reliability index, or R-factor, is defined as $\Sigma \|F_{\mathrm{obs}}| - |F_{\mathrm{calc}}\| / \Sigma |F_{\mathrm{obs}}|$, where the sums are taken over all reflections. It thus resembles an average fractional error. Alternative targets substitute intensities for amplitudes or utilize correlation functions. R_{free}, an agreement index calculated with a subset of reflections that is never sampled in refinement, is an important monitor of refinement.

STRUCTURE DETERMINATION BY MOLECULAR REPLACEMENT

When the protein of interest is related to a known structure, coordinates for the known homologue may be employed to solve the unknown. The initial computation is usually carried out in two stages, referred to as rotation and translation searches. The model is oriented and then positioned in the unit cell of the unknown crystal by correlating observed intensities or structure factor amplitudes with those calculated from the oriented model. It is then subjected to refinement. There is no hard-and-fast rule about the level of sequence similarity that augurs success in molecular replacement (MRP), or about the fraction of the structure to be used in searches. A standard practice has been to truncate side chains that are not identical in the two proteins. For multidomain proteins it may be desirable to use individual domains, rather than the intact protein, as search models.

The principal hazard in MRP is bias from the model structure. Although incorrect features should appear at lower-than-average electron densities, they do not disappear completely. One effective strategy to minimize bias is the computation of "omit" maps, in which local regions of the model that need adjustment or verification are not included in the phase calculations.

Similarity of structures is exploited in another form of molecular replacement in which the electron densities corresponding to structural repeats are averaged to generate modified maps and modified phases. This approach assumes that the multiple copies of chains or subunits found in an asymmetric unit (the smallest motif from which the crystal can be generated by translation and rotation) adopt identical conformations. These copies are related by coordinate transformations that are local and do not obey crystallographic symmetry; hence they are referred to as non-crystallographic symmetry operations. Implementation of this mode of molecular replacement was the key to the pioneering structure analyses of spherical viruses.

Studies of Molecular Complexes and Conformational Changes

The classic comparisons of deoxy- with methemoglobin and the difference Fourier analyses that revealed azide

bound to myoglobin presaged the central role of crystallography in the descriptions of conformation changes and molecular interactions. Subsequent structures have firmly established the plasticity of proteins and have demonstrated that a surprising variety of conformational states can be accessed by a particular polypeptide.

LIGAND BINDING IN SITU

Myoglobin–azide and lyzozyme–tetrasaccharide complexes were the prototypes for experiments in which structures of molecular complexes have been determined from preformed crystals after immersion in holding solutions containing ligands. This approach is feasible because the extensive solvent-filled channels within the crystals allow small ligands or substrates to diffuse to their binding sites. Density corresponding to the ligands is found in maps computed with amplitudes measured from the complex and phases from the ligand-free protein. Subsequent refinement often reveals local changes in the protein, elicited by interactions with the ligands. If the binding sites are not saturated, the electron densities will include images of both the ligand-free and ligand-bound species, and it may be necessary to refine the occupancies of the ligands and to include alternate conformations for parts of the protein.

LARGE CONFORMATION CHANGES

In contrast, major conformation changes such as those accompanying oxygenation of hemoglobin cannot be accommodated by the original crystal lattice and either are suppressed by competing crystal-packing forces or disrupt packing interactions and disorder the crystals. Descriptions of these larger structural changes therefore require analysis and comparison of different (non-isomorphous) crystals obtained under conditions that favor one or another conformation. Determination of the alternate conformation entails solution of a new structure by *de novo* or molecular replacement methods. Dramatic conformation changes are not uncommon and can include displacements of loops or flaps, swinging or pivoting of domains, and even remodeling of secondary structures. Numerous examples have been gathered in compendia of molecular motions.

STUDIES OF REACTION INTERMEDIATES: TRAPPING METHODS

Structures related to intermediates in enzyme-catalyzed reactions can be obtained from complexes with unreactive substrate analogues or transition-state analogues. More elegant experiments are also feasible, as many enzymes are active catalysts in the crystalline state.

With sufficient knowledge of the reaction mechanism, it is possible to design experiments in which true intermediates accumulate in the crystal, or to trap structures of these intermediates at appropriate times by flash-cooling. In studies of isocitrate dehydrogenase, Stoddard and co-workers exploited mutations to favor accumulation of intermediates either prior to or following hydride transfer. The analysis of several intermediates in catalysis by P450-cam was accomplished by controlling addition of reactants. Data were collected from the ferrous enzyme–substrate complex after chemical reduction, and crystals were then allowed to react with oxygen to form the subsequent dioxygen intermediate, which was in turn reduced by X-irradiation to yield a putative activated-oxygen intermediate.

TIME-RESOLVED CRYSTALLOGRAPHY

The most challenging experiments are those that aim to determine structure as a function of time. They require a narrow time window for initiating reactions and very rapid data collection, usually by the Laue method, which employs broadband radiation to capture most of the reflections in a single image of the diffraction pattern. The mixtures of structures that are present vary with time and must be sorted out. Time-resolved structural studies of the photodissociation of CO from crystals of myoglobin have revealed non-heme binding sites for CO and established the nature of relaxations in the heme and globin that follow photolysis.

Accuracy of X-Ray Structures and Metrics of Reliability

THE IMPORTANCE OF RESOLUTION

The exact definition of resolution in crystallography is illustrated in Figure 1, which depicts a plane from the 3-D lattice in diffraction (reciprocal) space. Resolution is probably the most important parameter in any assessment of a structure determination. The dramatic effects of resolution on the appearance of electron density maps have been illustrated in several texts and reviews. The number of reflections included in a structure determination increases as $(1/d_{min})^3$. Resolution thus controls the data:parameter ratios that are critical in refinement and is a primary determinant of positional accuracy. The choice of a resolution limit for a structure analysis is dictated by the completeness of the outermost shell of data and the agreement between measurements of symmetry-related reflections. These metrics are usually displayed in tables describing the structure determination.

UNCERTAINTIES IN COORDINATES

One would like to obtain estimates of the positional uncertainties for each atom and the derived uncertainties of bond lengths, bond angles, and torsions. The diagonal elements of the inverse matrix that is calculated in least-squares refinement provide these uncertainties in mathematically rigorous fashion. However, for proteins the size of the least-squares matrices (at least $3N$ by $3N$, where N is the number of atoms in the asymmetric unit) have generally precluded this computation. Approximations that derive global estimates of average coordinate errors from R-factors versus resolution are partly flawed because they assume the same temperature factors for all atoms. One compromise has been to invert large diagonal blocks taken from the full matrix. Evaluation of the estimated standard deviations arising from experimental errors is complicated by the inclusion of restraints in most refinements. The average errors in bond lengths are often dominated by the deviations chosen for restraints.

Refinements of a few structures have included inversion of the full matrix. Analysis of concanavalin A at 0.94Å resolution provided details of the geometry at the manganese- and calcium-binding sites with accurate estimates of the deviations of bond lengths and angles in the metal cluster, information that is critical for understanding the properties of the metal center. Comparison of rigorous with approximate error estimates for this and other ultra-high-resolution structures demonstrates, as might have been expected, that positional uncertainties depend strongly on the values of the temperature factor (displacement parameter). This dependence is especially pronounced in unrestrained refinements.

Solvent molecules (usually water) that occupy defined sites are included in archived coordinate lists (.pdb files). They are important for structure and function, and they contribute significantly to X-ray scattering even though their residence times may be short. Waters in buried or active sites may be very well-defined, but in general water oxygens are the least well-positioned atoms and the most subject to error. Some publications report refinement of both occupancy and isotropic B values for solvents, but because these parameters are highly correlated, others choose to allow only B to vary in refinement. It is customary to verify that waters are placed at canonical distances from neighbors and make reasonable interactions with protein atoms or other solvents.

RELIABILITY INDICES (R-FACTORS)

R-factors are widely cited criteria of the accuracy and reliability of a structure and are used to judge the progress and convergence of refinements. Readers of structure reports need to be aware of their shortcomings. The actual values will depend on the refinement algorithm, the choices of restraints or constraints, the omission or inclusion of weak reflections, and on the resolution (minimum d-spacing). The conventional R-factor, $\Sigma\|F_{obs}| - |F_{calc}\|/\Sigma|F_{obs}|$, can be decreased artificially by overfitting, e.g., by including many solvents with small scattering contributions or by using anisotropic temperature factors at too low a resolution.

The introduction of R_{free} in the early 1990s provided a better index of reliability and accuracy. R_{free} is calculated from a subset of data, randomly chosen with respect to intensity and resolution, that is never used in refinement. Acceptable refinement strategies such as inclusion of additional solvents should produce decreases not only in the conventional R but also in the R_{free} value.

OTHER QUALITY INDICES: STRUCTURE VALIDATION

Crystal structures are full of details, and models are built interactively and subjectively into the electron densities. It is useful to have objective methods to flag possible errors and regions that deserve further adjustment. Very occasionally, incorrect folds have been fit to electron densities, and wrong connections between pieces of the polypeptide have been introduced in maps at relatively low resolution. More common errors are mistaken registration of sequences and misoriented peptide planes. It is important to examine Ramachandran plots of Φ, Ψ angles for outliers with unusual (high-energy) backbone conformations; side-chain conformations should resemble one of the common rotamers. Comparisons of model density with experimental or omit maps identify residues that may be incorrectly modeled because they are mobile or adopt multiple conformations. The Protein Data Bank routinely subjects submitted coordinates to examination by PROCHECK, which analyzes the stereochemistry and other properties, flags possible errors, and assigns a quality index.

STRUCTURES AT VERY HIGH RESOLUTION

The library of structures at resolutions beyond 0.9Å is small but is growing steadily. Data:parameter ratios for these analyses allow unrestrained refinement and anisotropic modeling of the temperature factors, which requires specification of nine parameters per atom rather than four. In these structures it is possible to see what had to be surmised or inferred
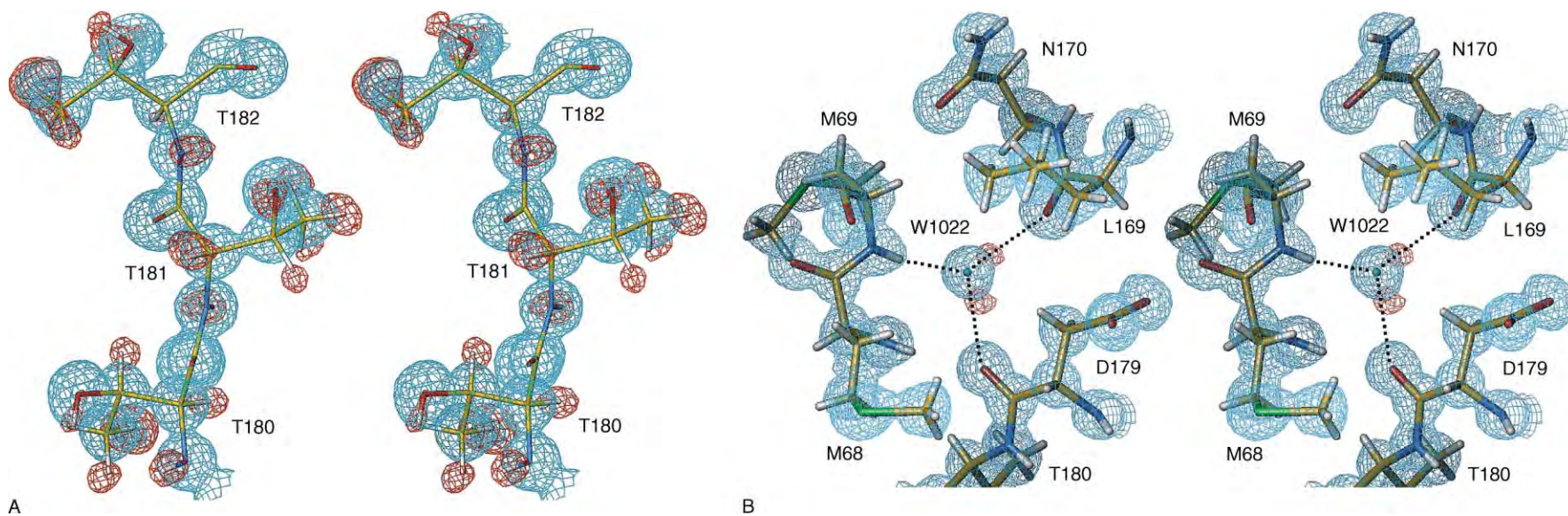
FIGURE 2 Stereoviews of electron density from the crystal structure determination of TEM-1 β-lactamase at 0.85Å resolution. At this resolution, densities corresponding to individual atoms are apparent (cyan), and difference densities (red) identify the positions of hydrogen atoms. The data:parameter ratio was 6:1 in refinement with anisotropic temperature factors; alternate conformations were included for 169 residues; R_{free} is 0.112. Reproduced from Minasov, G., Wang, X., and Shoichet, B. K. (2002). An ultrahigh resolution structure of TEM-1 beta-lactamase suggests a role for Glu 1 66 as the general base in acylation. *J. Am. Chem. Soc.* 124, 5333–5340.

from lower resolution structures—many of the hydrogen atoms, alternate conformations, and distinctions between oxygen and nitrogen atoms. Direct observation of hydrogen bonds is especially valuable for enzymologists, as is resolving ambiguities about the orientations of Asn, Gln, and His. As more high-resolution structures are completed, it should be possible to document true deviations of geometries from the canonical values embedded in restraint libraries. The example shown in Figure 2 illustrates the clear definition of densities corresponding to individual atoms and the assignments of hydrogens from difference maps.

Displaying and Comparing Structures

The computing power of current desktop machines allows the non-crystallographer to display and analyze structures that have been deposited in the Protein Data Bank. Particularly useful features of available programs are algorithms that align structures for comparisons of conformations, facile analysis of noncovalent interactions, routines for mutation and model building, and the capability to generate illustrations in a variety of styles.

SEE ALSO THE FOLLOWING ARTICLES

Imaging Methods • Protein Data Resources • Secondary Structure in Protein Analysis

GLOSSARY

asymmetric unit The smallest motif from which the crystal can be generated by translation and rotation operations. The unit cell that is repeated by translation to form the crystal may contain a number of asymmetric units.

isomorphism When addition of a ligand or heavy atom does not alter the scattering contribution (structure factor) of the protein atoms, the derivative and native crystals are said to be isomorphous. Cell dimensions are expected to be unchanged.

non-crystallographic symmetry (NCS) When the asymmetric unit includes more than one copy of a polypeptide, the multiple copies are related by local coordinate transformations that do not obey the symmetries that define the space group.

structure factor A vector **F** with phase $\alpha_{h,k,l}$ whose length |F| is the amplitude of the reflection h, k, l. The structure factor is the resultant of summing over scattering contributions from all the atoms in the asymmetric unit and can be calculated from the atom parameters. Fourier transformation of the structure factors yields the electron density.

temperature factor (displacement parameter) A measure of the motion of an atom about its equilibrium position. Temperature factors may be isotropic (B-values) or anisotropic. In the general anisotropic case, six parameters are required to describe the displacements.

FURTHER READING

Baldwin, J., and Chothia, C. (1979). Hemoglobin: the structural changes related to ligand binding and its allosteric mechanism. *J. Mol. Biol.* **129**, 183–191.

Brünger, A. T., Adams, P. D., and Rice, L. M. (1999). Annealing in crystallography: A powerful optimization tool. *Prog. Biophys. Mol. Biol.* **72**, 135–155.

Carter, C. W., Jr., and Sweet, R. M. (eds.) (1997). *Macromolecular Crystallography, Parts A and B. Methods in Enzymology*, Vols 276 and 277. Academic Press, San Diego.

Cruickshank, D. W. (1999). Remarks about protein structure precision. *Acta Crystallogr.* **D55**, 583–601.

Drenth, J. (1999). *Principles of Protein Crystallography*. Springer-Verlag, New York.

Garman, E. F., and Schneider, T. R. (1997). Macromolecular cryocrystallography. *J. Appl. Cryst.* **30**, 211–237.

Guex, N., and Peitsch, M. C. (1997). SWISS MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723.

Kleywegt, G. J. (2000). Validation of protein crystal structures. *Acta Crystallogr.* **D56**, 249–265.

Moffat, K. (2001). Time-resolved biochemical crystallography: A mechanistic perspective. *Chem. Rev.* **101**, 1569–1581.

Rossmann, M. G., and Arnold, E. (eds.) (2001). *Crystallography of Biological Macromolecules, International Tables for Crystallography*, Vol F. Kluwer Academic Publishers, Dordrecht.

Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000). The catalytic pathway of cytochrome P450cam at atomic resolution. *Science* **287**, 1615–1622.

Stoddard, B. L. (2001). Trapping reaction intermediates in macromolecular crystals for structural analyses. *Methods* **24**, 125–138.

Stryer, L., Kendrew, J. C., and Watson, H. C. (1964). The mode of attachment of the azide ion to sperm whole methemoglobin. *J. Mol. Biol.* **8**, 96–104.

Terwilliger, T. C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr.* **D55**, 849–861.

BIOGRAPHY

Martha L. Ludwig is Professor of Biological Chemistry and Research Biophysicist at the University of Michigan in Ann Arbor. She is an X-ray crystallographer whose primary research interest is the structure and function of enzymes that require metal- and vitamin-based cofactors. She is a Fellow of the American Association for the Advancement of Science and a member of the National Academy of Sciences.
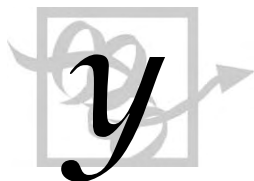
Yeast GAL1–GAL10 System

Dennis Lohr and Ralph Bash
Arizona State University, Tempe, Arizona, USA

The GAL1 and GAL10 genes in the budding yeast *S. cerevisiae* encode enzymes that help convert galactose to a glycolytic intermediate, thus allowing it to be used as a carbon source for cell growth. Genetic characterization of the GAL genes and dissection of their regulatory mechanisms by Douglas and Hawthorne in the 1960s provided a firm and crucial foundation for the molecular characterization of GAL1–10 regulation that began in the 1980s. Expression of GAL1 and GAL10 is strikingly carbon source-dependent. Transcription occurs at extremely high levels in galactose-grown cells but is undetectable in cells grown in other carbon sources. This clear-cut and efficient regulation plus the ability to couple genetic and biochemical studies in the analysis of single-copy genes have made GAL1–10 a very attractive model for eukaryotic gene regulation studies and a paradigm in which many general eukaryotic regulatory themes were first uncovered. The GAL1 promoter is also widely used to drive expression of heterologous genes, in various applications. The basic molecular aspects of GAL1–10 gene regulation are outlined below (genes referred to as GAL1, proteins as Gal1p).

The Elements Contributing to GAL1–10 Regulation

GAL1 and 10 share a common ~ 600 bp (base pair) intergenic region from which they are divergently transcribed (Figure 1). They are considered to be coexpressed and coregulated, mainly at the transcriptional level. Their regulation combines inputs from DNA sequence (promoter) elements, protein factors (gene-specific and general), and chromosome structure.

DNA SEQUENCE ELEMENTS

The most important gene-specific promoter elements for GAL1–10 expression are the upstream activation sequence (UAS_G) elements. These ~ 17 bp motifs are necessary and sufficient for galactose-dependent gene expression because; they provide the binding sites for Gal4p, the specific activator of GAL gene transcription.

GAL1–10 share four UAS_G asymmetrically located between the two genes (Figure 1). Each gene has its own TATA sequence for binding the general transcription factor TBP (TATA binding protein).

THE GENE-SPECIFIC REGULATORY FACTORS

Gal4p

As the required activator for GAL gene expression, Gal4p is a key player in GAL1–10 expression. In the presence of galactose, Gal4p activates transcription through a domain located near its carboxy terminus, residues 767–881, while bound to (UAS_G) DNA via a domain located near its amino terminus. It binds as a homodimer to individual UAS_G elements.

Gal80p

In carbon sources other than galactose, the negative regulator Gal80p binds specifically to the C-terminal activation domain of Gal4p such that it masks Gal4p activation activity and thus prevents GAL1–10 transcription. Gal80p binds Gal4p quite strongly ($K_d \sim 5$ nM).

Gal3p

In the presence of galactose, the Gal80p-mediated inhibition of Gal4p is relaxed by a Gal3p-dependent process, thus freeing the Gal4p activation domain to activate GAL1–10 transcription. Gal3p behaves like a signal transducer and appears to be located solely in the cytoplasm, which should facilitate its interaction with galactose. Gal3p can bind to Gal80p but how this (cytoplasmic) interaction might influence the nuclear Gal80p–Gal4p complex is unclear at this time. Interestingly, the product of the GAL1 gene contains a Gal3p-like activity that apparently makes it capable of fulfilling a similar function to Gal3p.
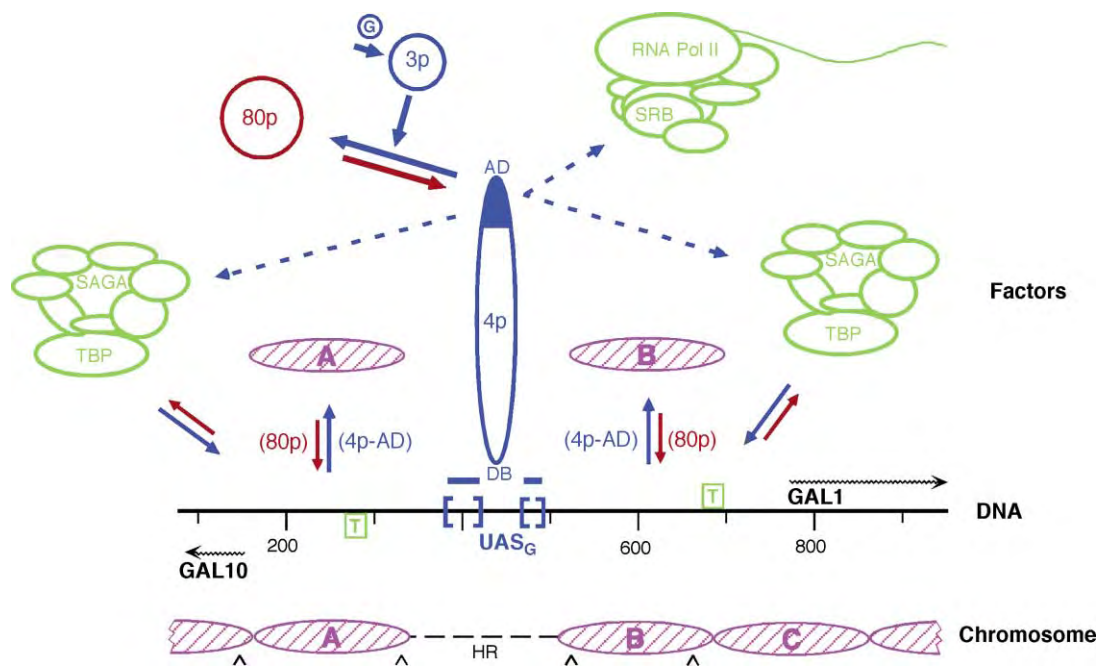
FIGURE 1   An outline of *GAL1–10* regulation. The DNA sequence organization of the yeast *GAL1–10* promoter region is shown in the center of the diagram ("*DNA*"), with UAS_G (blue brackets) and TATA boxes (green, boxed "T") located to scale along the sequence (gray line). The numbers refer to base pairs from an EcoRI site in the *GAL10* gene. The transcription start sites are located at the origins of the two wavy arrows below (*GAL10*) or above (*GAL1*) the line. The chromosomal structure of this region in the inactive state of gene expression is shown at the bottom ("Chromosome") with nucleosomes A, B and C (pink) located to scale on the sequence and the nonnucleosomal, highly accessible region in the chromosome indicated by a dashed line ("HR"). The locations of intrinsic, sequence specific DNA bends are shown by " ∧ ". In the topmost portion of the Figure ("Factors"), various factors and complexes known or thought to be involved in regulating and implementing *GAL1–10* expression are shown. The *GAL*-specific factors ("3p", Gal3p; "4p", Gal4p) are shown in blue and the reactions that occur in, or are characteristic of, the active (galactose-induced, "G") state are designated by solid blue arrows. Dotted blue arrows indicate the suggested recruitment of complexes by Gal4p during gene activation, RNA polymerase II ("RNA pol II"), including SRB (*S*uppressor of *R*NA Polymerase *B* subunit) proteins and the SAGA (*Spt-Ada-Gen5-A*cetyltransferase) complex. The activation domain ("AD") of Gal4p is solid blue and the Gal4p DNA binding domain is identified ("DB"). The negative *GAL*-specific factor, Gal80p ("80p"), is shown in red and the reactions that occur in, or are characteristic of, the inactive state of expression are designated by red arrows. The association/dissociation of TBP with the *GAL1* and *GAL10* TATA and the disruption/redeposition equilibria of promoter nucleosomes A and B (pink) are also designated by solid blue/red arrows. The involvement of Gal4p and Gal80p in the promoter nucleosome disruption/redeposition equilibrium is indicated by their appearance in parentheses next to the appropriate arrow. Nucleosome C, which also undergoes disruption/redeposition, is not shown in this upper portion of the figure.

OTHER FACTORS

Global processes, such as the general glucose repression system, also participate in GAL1–10 regulation and elements of the RNA polymerase II transcription apparatus function in GAL1–10 transcription. Several other genes show galactose-sensitive and Gal4p-dependent expression, suggesting that their gene products may be part of the regulatory network: GAL6 (possible alternative regulator); GCY1 (oxidoreductase); PCL10 (cyclin); FUR4 (uracil permease); RPA; and YP53 (protein metabolism). Their precise roles are unknown.

CHROMOSOMAL ORGANIZATION

The promoter region of GAL1–10 has a distinctive chromosomal organization that is an aspect of gene control. In all carbon sources, the UAS_G elements reside within a sizable (at least 150 bp), highly accessible nucleosome-free region ("HR," Figure 1). Thus, the UAS_G are always available to Gal4p; neither nucleosome removal nor competition with nucleosomes is required to expose these elements for Gal4p binding. The strategy of locating major promoter elements in a large, constitutively accessible region is not common on eukaryotic promoters and the features that maintain it have yet to be defined. In nongalactose carbon sources, nucleosomes bracket this highly accessible HR region, covering the TATA and transcription start sites ("A–C," Figure 1). The DNA near the termini of nucleosomes A and B contains intrinsic, sequence-dependent DNA bends (Figure 1) that may contribute to the preferential location of these promoter nucleosomes.

The (Three) States of GAL1–10 Gene Expression

THE ACTIVE (GALACTOSE-INDUCED) STATE

In galactose, GAL1–10 are transcribed at very high levels, indicating that the activation mechanisms for their expression are quite robust. Under these conditions, Gal4p is bound, via its N-terminal DNA-binding domain, to the UAS$_G$ elements that lie between GAL1 and GAL10 while promoting transcription of the genes through its C-terminal activation domain. In this state, binding of TBP to the GAL1 and GAL10 TATA will be greatly aided by the apparent removal (disruption) of the promoter nucleosomes A–C, a process mediated directly or indirectly (through other factors) by the Gal4p activation domain. Transcription activation by Gal4p also involves the direct or indirect recruitment of cellular complexes (Figure 1). The UAS$_G$ are >200 bp from the GAL1–10 transcription start sites suggesting that the three-dimensional structure of Gal4p and promoter chromatin architecture, both unknown, may also impact on the activation process. This variety of Gal4p activation depends upon the galactose-induced release of Gal80p inhibition that is mediated through Gal3p (and Gal1p?).

THE INACTIVE STATES (REPRESSED OR POISED)

There are two distinct types of inactive states, repressed (in glucose) or poised for expression (in nonfermentable carbon sources like glycerol/lactate). In both inactive states, Gal80p binds to the Gal4p activation domain, masking its activity, and nucleosomes A–C are present on the GAL1–10 promoter region, covering the TATA and transcription start sites.

The presence of glucose results in an additional negative feature, Gal4p absence from the UAS$_G$. This absence prevents GAL1–10 from being (rapidly) inducible in glucose. The UAS$_G$ are still present in an open chromosomal region and should thus be available to Gal4p. Gal4p absence from the UAS$_G$ is thought to result mainly from a decrease in Gal4p levels due to decreased expression of the *GAL4* gene, imposed by the global catabolite repression apparatus. The decrease in expression is fairly modest but Gal4p–UAS$_G$ binding should be very sensitive to Gal4p levels due to its highly cooperative nature (multiple UAS$_G$/two Gal4p per UAS$_G$).

In carbon sources that are neither repressing nor inducing, such as glycerol lactate, the GAL1–10 genes are not expressed at all but they can be very rapidly (within minutes) induced to full expression by galactose.

Rapid inducibility is due to the presence of Gal4p on the UAS$_G$ in these carbon sources, strongly protecting these elements as in galactose-grown cells, and to the (low-level) presence of the signal transducer Gal3p. Elevated Gal4p levels (GAL4 is most actively expressed in these carbon sources) plus constitutive accessibility of the UAS$_G$ elements probably account for the UAS$_G$ occupation by Gal4p in this state. Thus, although inactive, GAL1–10 are poised for expression; only Gal80p inhibition of Gal4p (and the presence of nucleosomes A–C) prevents transcription. A poised state would probably have been a major advantage to wild yeast growing on poor carbon sources, by allowing them to utilize galactose even if it were only transiently available.

Key Themes in GAL1–10 Regulation

GAL4P: COMPLETELY SEPARABLE DNA BINDING AND TRANSCRIPTION ACTIVATION FUNCTIONS

The ability of Gal4p to bind strongly to the UAS$_G$ in the poised inactive state demonstrates that activator binding and transcription activation are separable aspects of GAL1–10 expression. As shown unambiguously by the Ptashne lab, this reflects the independence of the DNA binding and transcription activation functions of Gal4p. (This feature found an important general application in two-hybrid analysis.) These two functions of Gal4p are also differentially controlled: DNA binding by Gal4p levels, transcription activation by Gal80p.

HOW GAL4P ACTIVATES TRANSCRIPTION

As shown by Ptashne and co-workers, Gal4p can function throughout the eukaryotic kingdom, in microbes, animals, and plants. Thus, it must utilize basic and conserved mechanisms of transcriptional activation and target universal components of the transcription apparatus. TBP is a likely recruitment target of the Gal4p activation domain and recent work has suggested that the spt-ada-Gcn5-acetyltransferase (SAGA) complex mediates this recruitment. Gal4p has also been suggested to recruit RNA polymerase. Disruption of the promoter nucleosomes (A–C, Figure 1) in galactose is another explicit function of the Gal4p activation domain; this event is not simply an indirect consequence of transcription (as shown by the Majors lab). Exposing the DNA in these nucleosome-covered promoter regions is necessary to provide access for factors that initiate the transcription

process like TBP, which cannot bind to DNA that is nucleosome-covered. Surprisingly, the activation functions of Gal4p do not appear to depend on specific amino acids or protein structural motifs and sequence variations in the activation domain are exceptionally well tolerated as shown the labs of Johnston, Ptashne, and others.

THE CENTRAL ROLE OF THE NEGATIVE REGULATOR GAL80P

In many respects, Gal80p may be viewed as the key regulator of GAL1–10 expression. It directly inhibits the activator Gal4p in nongalactose carbon sources, it responds to the Gal3p-dependent signal in galactose, thus mediating the activation response, and it even appears to temper GAL1–10 expression levels in galactose. The latter function may explain the surprising fact that expression of this negative regulator is significantly *increased* in galactose (up five- to tenfold), in a Gal4p/UAS_G-dependent transcription process. Gal80p also mediates, directly or indirectly, redeposition of the disrupted promoter nucleosomes A–C when conditions (activation signals, cellular energy levels) are unfavorable, as shown in the Lohr lab, and its effects on GAL1–10 expression in galactose, in particular, may reflect this activity.

THE IMPORTANCE OF PROTEIN–PROTEIN CONTACTS IN GAL1–10 REGULATION

GAL1–10 regulation is implemented mainly through protein–protein interactions: Gal4p with Gal80p, Gal80p with Gal3p, Gal4p with transcription factors (TBP/SAGA/RNA pol II), and Gal4p/Gal80p with nucleosomes. The latter two may also involve other factors. The prominent role of protein–protein contacts might account for the importance of regulator (Gal4p, Gal80p, Gal3p) stoichiometries for proper regulation, which in turn are reflected in the levels and carbon source variations in regulatory gene expression.

THE ROLE OF CHROMOSOME STRUCTURE IN GAL1–10 REGULATION

Chromatin structure is now seen as a major facet of eukaryotic gene regulation; studies of GAL1–10 chromatin as early as the mid 1980s in the Lohr lab provided indications of this feature. The advantages of maintaining the major promoter elements in an accessible chromosomal region, particularly for the poised state of expression, have been discussed previously. As shown by the Grunstein lab, the promoter nucleosomes A–C play regulatory roles in both the inactive and active states of GAL1–10 expression. These nucleosomes help inhibit transcription in nongalactose carbon sources; nucleosome depletion allows TATA-driven (Gal4p-independent) GAL1 expression, even in glucose. The N-terminal tails of histone H4 are needed for full levels of galactose-induced GAL1 expression and therefore play a positive role in this process; this role involves the nucleosome B region. In contrast, the removal of H3 histone tails results in an elevated level of induced GAL1 expression, in an effect that depends on the UAS_G region; therefore the H3 tails must normally function in a process that attenuates induced expression. Occupancy of the TATA regions on the GAL1–10 promoter probably involves a competition between TBP and the promoter region nucleosomes A and B/C (Figure 1). Galactose-inducing conditions favor TBP binding and thus transcription, as the nucleosomes are disrupted and the promoter region is exposed by a Gal4p-dependent process(es); under inactivating conditions, nucleosome binding to the regions is favored, and transcription is disfavored, by a process(es) that is dependent on Gal80p (Figure 1). Therefore, nucleosome presence on the promoter region (and thus transcriptional activity) is dependent on the state of a disruption/redeposition equilibrium, which is controlled by the gene-specific regulators Gal4p and Gal80p. The specific roles of the histone tails mentioned above may be linked to this regulator-dependent promoter region disruption/redeposition equilbrium. Only the promoter nucleosomes A–C show this behavior, not the GAL1 coding region nucleosomes, which is consistent with the suggested regulatory role of the promoter nucleosomes and their link to Gal4p/Gal80p action.

SEE ALSO THE FOLLOWING ARTICLES

Chromatin: Physical Organization • Chromatin Remodeling • RNA Polymerase II and Basal Transcription Factors in Eukaryotes • RNA Polymerase II Elongation Control in Eukaryotes • RNA Polymerase II Structure in Eukaryotes • RNA Polymerase Reaction in Bacteria • RNA Polymerase Structure, Bacterial • T7 RNA Polymerase

GLOSSARY

activator   Proteins that enable or enhance transcription of genes.
gene-specific factors   Proteins whose function is associated with expression of a specific gene(s).
nucleosome   The complex of histone proteins wrapping up 147 bp of DNA that is the basic structural unit of eukaryotic chromosomes.
promoter   DNA sequences that mediate gene expression.

FURTHER READING

Bash, R., and Lohr, D. (2001). Yeast chromatin structure and regulation of GAL gene expression. *Prog. Nucl. Acids Res. Mol. Biol.* **65**, 197–259.

Johnston, M. (1987). A model fungal gene regulatory mechanism: The GAL genes of *Saccharomyces cerevisiae*. *Microbiol. Rev.* **51**, 458–476.

Johnston, M., and Carlson, M. (1992). Regulation of carbon and phosphate utilization. In *The Molecular and Cellular Biology of the Yeast Saccharomyces* Vol. II, pp. 193–281. Cold Springs Harbor Laboratory Press, Cold Springs Harbor, New York.

Lohr, D. (1997). Nucleosome transactions on the promoters of the yeast *GAL* and PHO genes. *J. Biol. Chem.* **272**, 26795–26798.

Lohr, D., Venkov, P., and Zlatanova, J. (1995). Transcriptional regulation in the yeast GAL gene family: A complex genetic network. *FASEB J.* **9**, 777–787.

Reece, R., and Platt, A. (1997). Signalling activation and repression of RNA polymerase II transcription in yeast. *BioEssays* **19**, 1001–1010.

BIOGRAPHY

Dennis Lohr is a Professor in the Department of Chemistry and Biochemistry at Arizona State University, with principal interests in the relationships between chromatin structure and gene expression in eukaryotes. He holds a Ph.D. from the University of North Carolina (Chapel Hill) and was a postdoctoral associate in the Biochemistry and Biophysics Department at Oregon State University, in the lab of K. VanHolde. At Oregon State, he carried out most of the initial characterizations of yeast chromatin structure and in his own lab carried out the first detailed structural analyses of a single copy eukaryotic promoter, the GAL1–10 promoter, and its accompanying gene, GAL1.

Ralph Bash is a postdoctoral Research Associate at the Arizona Biodesign Institute and the Department of Physics and Astronomy at Arizona State University with interests in chromatin structure, chromatin function and molecular imaging. He holds a Masters from San Jose State University and a Ph.D. from Arizona State University, where he studied intrinsic DNA bending on GAL genes.

Zinc Fingers

Mark Isalan

European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

Zinc fingers are small, compact protein subunits, with structured peptide chains folding around chelated zinc ions. Functionally, these subunits carry out a wide variety of tasks within cells by providing stable structural scaffolds and driving critical binding interactions, especially between proteins, DNA, and RNA. Zinc fingers are particularly important in gene regulation, where many proteins employ them to bind DNA in a sequence-specific manner, so as to activate or inhibit particular genes. Although many families of zinc fingers exist, each with a characteristic fold or structure, perhaps the best studied is the "classical" ("Cys_2–His_2" or "C2H2") type. Classical zinc fingers are extremely versatile, as can be seen by their abundance in nature: over 700 proteins contain such domains in the human genome alone, making them the second most common human protein motif. Zinc fingers may be seen as convenient molecular building blocks, often used in clusters, whose modular nature allows them to be conveniently duplicated, altered, and shuffled by evolution. In a biotechnology context, zinc fingers may be re-engineered to have novel binding properties with numerous applications in both research and medicine.

Zinc Finger Structures and Families

CLASSICAL ZINC FINGERS

History

The zinc finger motif was discovered in 1985 by Miller, McLachlan, and Klug in a protein responsible for controlling gene expression: the *Xenopus* transcription factor IIIA (TFIIIA). Originally, nine novel consecutive repeats were found, each ~30 amino acids in length. Since every repeat had a conserved pair of cysteines and histidines at defined positions, Klug proposed that these residues could chelate a zinc ion, stabilizing an independent structural domain, the "zinc finger." Notably, the term "finger" was derived from the finger-like appearance of the sequences when drawn schematically around a zinc ion, as in Figure 1(A). The repeats found in TFIIIA are now used as a definition of the classical zinc finger, where individual fingers conform to the archetypal consensus sequence shown in Figure 2.

Secondary Structure and Hydrophobic Core

Structural studies revealed that the classical Cys_2–His_2 finger is a simple and compact unit based on an antiparallel β-hairpin packed against an α-helix (Figure 1(B)). In addition to chelating zinc, classical fingers are stabilized at the core by a hydrophobic cluster of residues, converging from three fixed positions (see Figures 1 and 2). Moreover, this hydrophobic triad is highly conserved, often consisting of tyrosine, phenylalanine, and leucine. By contrast, zinc fingers may withstand extensive amino acid substitutions at other positions, without significantly altering the overall structure.

The Role of Zinc

Small protein domains are inherently unstable and, even though the $\beta\beta\alpha$ fold found in zinc fingers can form around hydrophobic cores alone, the stability of the fold is dramatically improved by zinc chelation. Such stability cannot be obtained using the disulfide bridge – a common extracellular alternative – because of the highly reducing environment inside cells. Consequently, intracellular proteins have solved the problem of stabilizing small domains by utilizing metal-binding sites. Zinc is ideally suited for this purpose as it has a single oxidation state, a fixed and distinguishable ionic radius, and can accommodate both nitrogen and sulfur ligands.

Evolution

Zinc fingers may have evolved originally from peptides which maintained a simple $\beta\beta\alpha$ structure (similar to that shown in Figure 1(B)) independently of zinc. In any case, the $\beta\beta\alpha$ fold is one of the smallest possible units of protein tertiary structure, and may therefore be one of the oldest. Zinc fingers are very abundant in nature and the huge variety found in various organisms, from yeast to human, may in part be due to their modular nature.
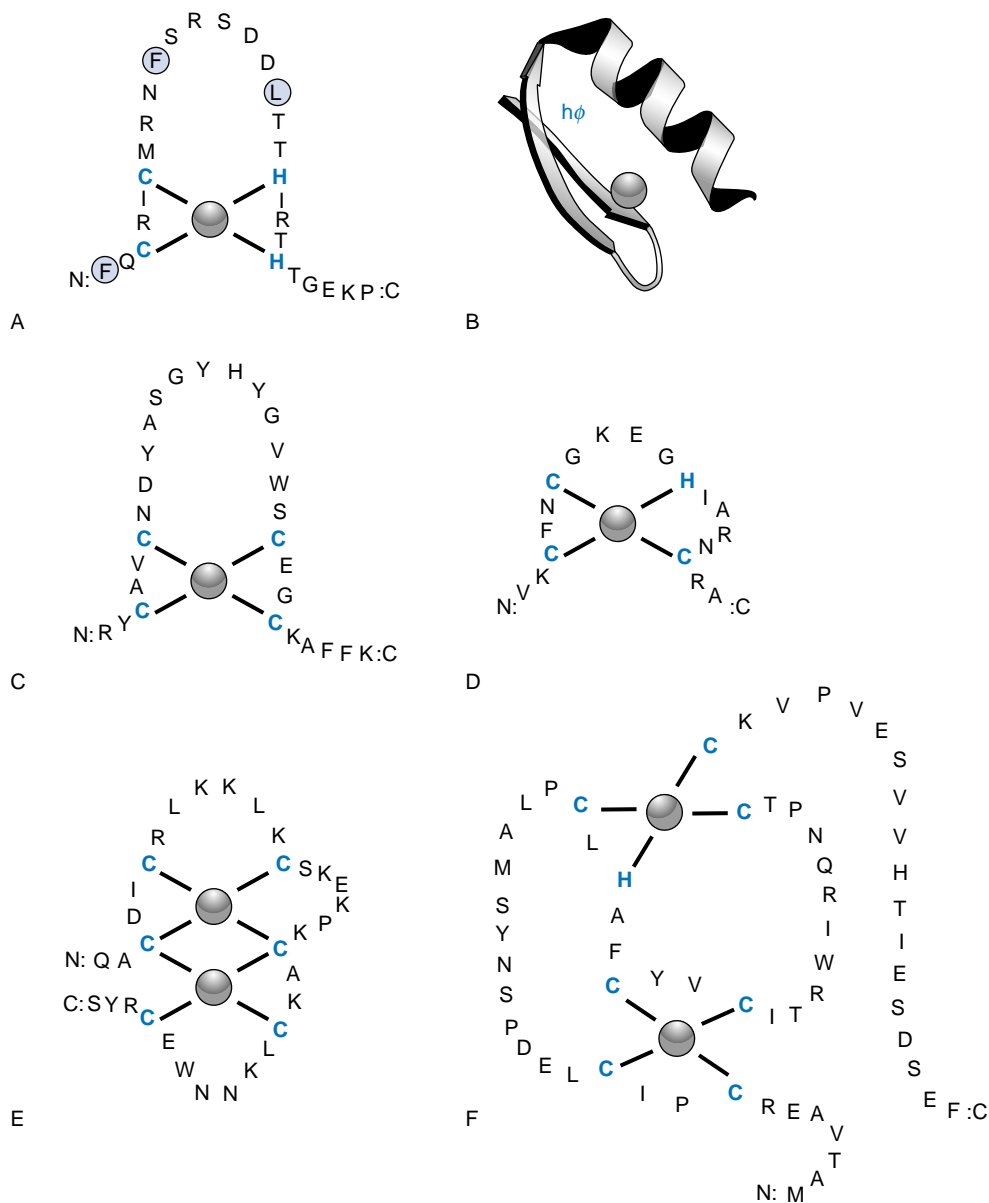
FIGURE 1 Examples of zinc finger families and folds. (A) A classical Cys_2–His_2 zinc finger is drawn schematically as a string of amino acids (1-letter abbreviations), with cysteine (C) and histidine (H) pairs chelating the central zinc ion. Three circled residues (F, F, L) come together in the structure to form a hydrophobic core. The finger shown is Finger 2 from the mouse transcription factor Zif268. (B) Representation of the actual $\beta\beta\alpha$ structure of a classical Cys_2–His_2 zinc finger. Note the position of the hydrophobic core (hϕ). (C) A Cys_4 steroid-receptor zinc finger from the estrogen receptor. Note that this protein actually contains two contiguous Cys_4 zinc finger repeats, forming a single structural unit. (D) A Cys_3–His zinc finger from a viral Nucleocapsid protein. (E) A yeast GAL4 domain. (F) The "ring" finger fold.

In this regard, not only are zinc fingers independently folding structural building blocks, but they are often encoded by single exons. Indeed, exon shuffling and gene duplication may account for why zinc fingers have spread so widely during evolution.

Genomic Abundance

The recent initial drafts of the human genome have revealed the extent to which classical zinc fingers are used by cells: there are 706 genes containing one or more classical zinc fingers in humans, totaling several thousand fingers, and making them the second most common human protein family. It is remarkable that only immunoglobulin domains are more abundant in humans, and indeed only slightly more so, being found in 765 genes. By contrast, the *Drosophila melanogaster* fly has 357 classical zinc finger genes, actually making these the most common protein family in this organism. Despite the fact that zinc finger usage varies from
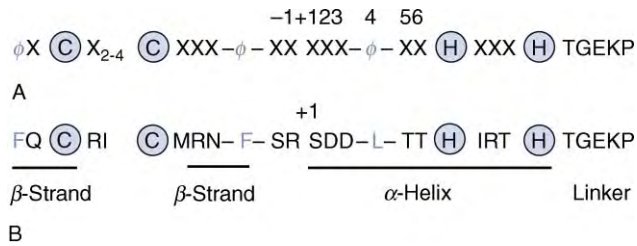
FIGURE 2 Zinc finger amino acid sequences. (A) Consensus sequence of a classical Cys_2-His_2 zinc finger, where X is any amino acid and ϕ represents hydrophobic residues (typically Y, F, and L). Zinc-chelating cysteine and histidine residues are circled. α-helical positions (numbered relative to position $+1$) are indicated above the sequence. (B) The sequence of Finger 2 of Zif268, showing conformity with the consensus sequence.

organism to organism, these motifs are clearly a widespread feature in all eukaryotic cells.

NON-CLASSICAL ZINC FINGERS

Although zinc-binding proteins have been known for many years, the usual definition of zinc fingers maintains that they are distinct from other zinc-binding protein subregions by virtue of being autonomously folding, functional mini-domains. Following the discovery of the classical Cys_2-His_2 zinc finger, many other classes of "zinc finger" have emerged. Although these form a diverse collection, often with radically different folding topologies, these proteins are often referred to in the literature as zinc fingers. Therefore some well-studied examples are outlined here below.

Cys₄ Zinc Fingers

One of the best-characterized nonclassical zinc fingers is the Cys_4 type, including the steroid-hormone receptor family. The estrogen receptor (Figure 1(C)) was the first protein of this type to be elucidated structurally and the gene contains two Cys_4 zinc finger repeats, forming a single structural unit that is quite unlike the independent classical zinc fingers of TFIIIA. Functionally, the steroid receptors form dimers, that bind hormones like estrogen, and then mediate activation of gene expression through sequence-specific DNA-binding interactions in the promoter regions of target genes. Further examples of proteins containing Cys_4 finger motifs include other nuclear hormone receptors, such as the glucocorticoid and retinoic acid receptors. Many Cys_4 proteins function as transcription factors, such as the elongation factor TFIIS and the adenoviral E1A transactivator.

Cys₃−His Zinc Fingers

The Cys_3-His zinc finger is another variant defined by its characteristic combination of cysteine and histidine residues. For example, the nucleocapsid protein, found in certain viruses such as HIV, contains Cys_3-His fingers, with various nucleic acid binding properties that facilitate hybridization and the viral life-cycle (Figure 1(D)). Other examples of Cys_3-His proteins include MetRS and the first finger in the LIM domain.

Multi-Zinc Proteins

Moving beyond the simpler zinc finger domains, which fold around a single zinc ion, there exist larger more complex folds, each employing two or more zinc ions. Examples include the yeast GAL4 domain (Figure 1(E)) and the "ring" finger (Figure 1(F)). Note that although these larger domains are not quite as prevalent as the single-zinc modules, they nonetheless form quite abundant families. For example, the "ring" finger family ranks as the 15th most abundant domain in the human genome, with 210 members, while another fold, the "PHD" finger, ranks 34th with 84 members.

Zinc Finger Functions

INTERACTIONS WITH DNA AND RNA

Since their discovery, zinc fingers have been associated with binding nucleic acids, namely, DNA and RNA. Indeed, the archetypal zinc finger protein, TFIIIA, is bi-functional, having distinct modes of binding which allow it to bind both to the DNA internal control region of 5S RNA genes and to the transcribed RNA gene product. In this sense, zinc fingers are master gene regulators, with the potential to influence both transcription and translation.

Structure of Zinc Finger Bound to DNA

The majority of our knowledge of zinc finger functions has focused on their ability to bind specific DNA sequences. The first structure of zinc fingers bound to DNA, that of the DNA-binding domain of the mouse transcription factor Zif268, was solved by Carl Pabo and colleagues in 1991. The structure revealed that each of the three classical zinc fingers occupies the major groove, wrapping around the DNA for almost one turn of the double helix (Figure 3A). Base specific contacts occur primarily from consecutive turns of the zinc finger α-helices to the guanine-rich strand of the DNA.

DNA Recognition Helix

Essentially, classical zinc fingers use the α-helix as a "reading head" to recognize specific DNA sequences. Contacts from amino acids consist primarily of electrostatic interactions and hydrogen bonds, although hydrophobic interactions also play a role. Conventionally, the amino acids in the α-helix are numbered
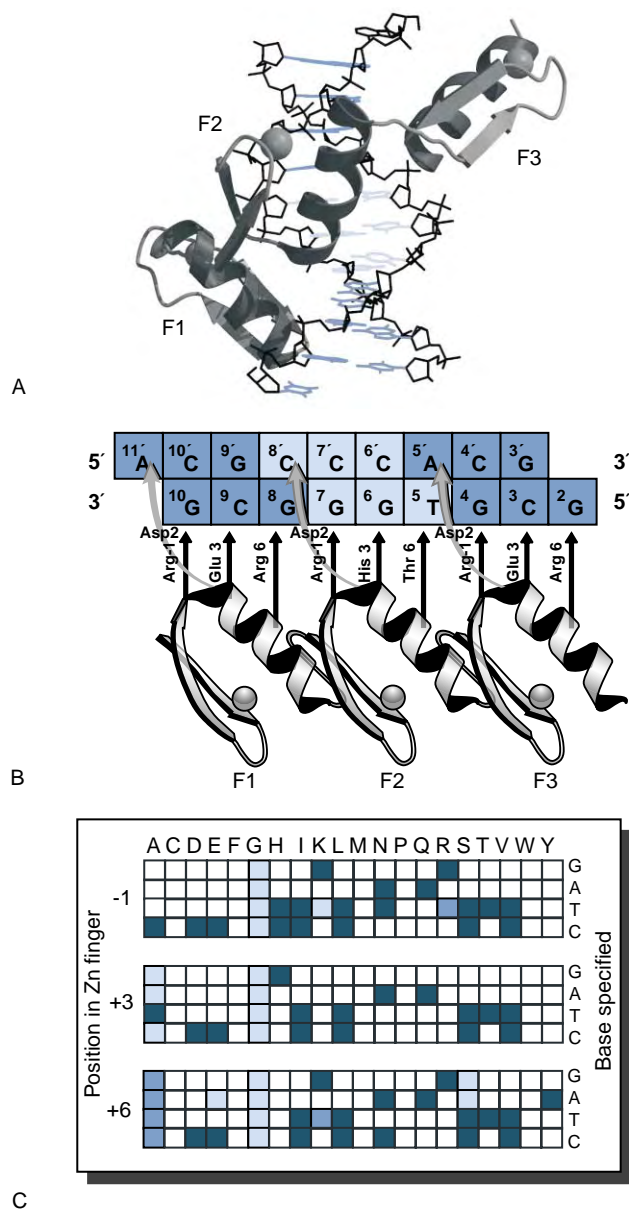
FIGURE 3 Interactions between zinc fingers and DNA. (A) Side view of Zif268 binding DNA (Pavletich and Pabo (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1Å. *Science* **252**, 809–817. The three zinc fingers (F1–F3) bind the major groove of the DNA. (B) Schematic diagram of a simple model of recognition between the three zinc fingers of Zif268 and the triplet subsites of an optimized DNA-binding site. Arrows indicate contacts by recognition residues with bases in each DNA strand. Note that most contacts are to the G-rich strand. Since the N-terminal finger contacts the 3′ end of the DNA and the C-terminal finger the 5′ end, binding is said to be antiparallel. (C) Zinc finger DNA-recognition code. The grid summarizes common associations between amino acid residues (A–Y) and the bases they face in a classical zinc finger DNA-binding site. Darker shading of boxes indicates stronger associations. Note that contacts from overlapping C-terminal zinc fingers (especially from position 2) can significantly alter the coding specificity shown here for position +6.

relative to a defined first position, Position +1 (see Figures 2 and 3(B)). Since positions −1, +3, and +6 are all to be found on the external face of the helix, these can form a binding surface that is especially useful in DNA recognition because the spacing of the residues matches the register of the base pairs. It should be noted, however, that contacts from other positions, especially position +2, are also important in DNA recognition and that not all zinc finger–DNA interactions are as regular as the canonical structure of Zif268.

DNA Recognition "Code"

The most striking feature of the canonical Zif268 structure is the presence of regularly spaced one-to-one contacts between amino acids and DNA bases, from the same fixed helical positions in each finger (Figure 3B). By studying mutations at these positions, it has emerged that classical zinc fingers are a special case in terms of protein–DNA interactions. Unlike for most DNA-binding proteins, these positions do seem to fit a residue-base code-like pattern. Positions −1, +2, +3, and +6 in the α-helix are particularly important (Figure 3B), and some major residue-base associations are outlined in Figure 3C.

As a note of caution, it is important to mention that the code is not absolute, and is certainly an over-simplification of the interplay between different residues in the protein and between factors in the DNA structure, such as base stacking and unwinding. Nonetheless, thinking about zinc finger DNA interactions in this way can give a valuable approximation into the possible binding modes of a given zinc finger protein.

RNA Recognition

Unlike the double helical scaffold of DNA, RNA has much more complex folds, with every RNA presenting a unique binding surface. Nonetheless, zinc fingers have the versatility to accommodate such structures, with contacts from all over the finger framework. Like DNA recognition, electrostatic interactions and hydrogen bonding play a major role, with amino acids contacting either RNA bases or the phosphodiester backbone. Note that in RNA recognition there is a greater potential for hydrophobic interactions, with aromatic residues such as Tyr, Phe, and Trp stacking around extruding bases.

NON-NUCLEIC ACID INTERACTIONS

Protein–Protein Interactions

Zinc fingers are not limited to controlling gene expression, but also play an important cellular role in mediating specific protein–protein interactions.

Whereas nucleic acid recognition usually requires clusters of fingers, cytoplasmic proteins often contain only one or two fingers. This is possibly because a single protein-binding finger has sufficient affinity to bind effectively, given the more variable binding surface presented by other proteins.

Lipid Binding

Recently, the known binding repertoire of zinc fingers has been extended to include lipid recognition. This is exemplified by the FYVE zinc finger family, that is found in a wide range of eukaryotic cells. These nonclassical fingers have a complex structure consisting of multiple loops, two twin-stranded β-sheets and an α-helix stabilized by a pair of zinc ions. Functionally, the domains bind to phosphatidylinositol 3-phosphate (PI3P) in cellular membranes, commonly found on endosomes, and mediate further recruitment of a range of proteins.

ZINC FINGERS IN BIOTECHNOLOGY

Classical zinc fingers are the ideal scaffold for protein engineering. They are small and versatile, withstanding extensive mutations without destroying their basic structure. Moreover, individual finger modules may be linked in tandem repeats to allow recognition of extended target surfaces such as DNA grooves. Novel DNA-binding zinc fingers have been routinely engineered for nearly a decade and are beginning to demonstrate their utility in biotechnical applications. By fusing with suitable effector domains, effective transcription activators and repressors can be synthesized that recognize specific DNA sequences, typically 9–18 bp long. These constructs can up-regulate or down-regulate specific genes in a cell and have been used in a wide variety of contexts, including as antiviral agents and angiogenesis-promoting factors.

SEE ALSO THE FOLLOWING ARTICLES

Aminopeptidases • Calpain • DNA Polymerase ε, Eukaryotic • DNA Sequence Recognition by Proteins • Inorganic Biochemistry • Nucleotide Excision Repair, Bacterial: The UvrABCD System • Opioid Receptors • Phosphatidylinositol Bisphosphate and Trisphosphate • Phosphatidylinositol-3-Phosphate • Retinoic Acid Receptors

GLOSSARY

classical finger A zinc finger of the TFIIIA type, where a zinc ion is coordinated by two cysteines and two histidines ("Cys_2–His_2" or "C2H2").

protein motif A small structural element of a protein with a characteristic fold or shape.

TFIIIA A transcription factor protein from *Xenopus laevis*, in which zinc fingers were originally discovered.

transcription factor A protein involved in gene regulation that helps to activate or repress the expression of a particular gene.

zinc finger A small structural protein motif consisting of peptide chain folded around a central chelated (coordinated) zinc ion.

FURTHER READING

Beerli, R. R., and Barbas, C. F., III (2002). Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol.* **20**, 135–141.

Berg, J. M., and Shi, Y. (1996). The galvanization of biology: A growing appreciation for the roles of zinc. *Science* **271**, 1081–1085.

Choo, Y., and Isalan, M. (2000). Advances in zinc finger engineering. *Curr. Opin. Struct. Biol.* **10**, 411–416.

Klug, A., and Schwabe, J. W. (1995). Protein motifs 5. Zinc fingers. *FASEB J.* **9**, 597–604.

Lander, E. S., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

Mackay, J. P., and Crossley, M. (1998). Zinc fingers are sticking together. *Trends Biochem. Sci.* **23**, 1–4.

Pabo, C. O., Peisach, E., and Grant, R. A. (2001). Design and selection of novel Cys_2His_2 zinc finger proteins. *Annu. Rev. Biochem.* **70**, 313–340.

BIOGRAPHY

Mark Isalan is a fellow of the Wellcome Trust, UK, and is currently at the European Molecular Biology Laboratory (EMBL), Heidelberg. His principal research interests include combinatorial and design approaches to protein engineering, gene regulation, and gene networks. He holds a Ph.D. from the University of Cambridge, where he trained at the MRC Laboratory of Molecular Biology, developing methods to engineer novel DNA-binding zinc fingers.

Editors-in-Chief

William J. Lennarz

State University of New York at Stony Brook, Stony Brook,
New York, USA
Section: Lipids, Carbohydrates, Membranes and Membrane Proteins

WILLIAM J. LENNARZ received his B.S. in Chemistry from Pennsylvania State University and a Ph.D. in Organic Chemistry from the University of Illinois. Subsequently he carried out postdoctoral work at Harvard with Konrad Bloch on fatty acid biosynthesis. In 1962 he was appointed Assistant Professor at Johns Hopkins in the Department of Physiological Chemistry. After promotion to Associate Professor in 1967, and full Professor in 1971, he remained at Hopkins until 1983. At that time, he was appointed Robert A. Welch Professor and Chair of the Department of Biochemistry and Molecular Biology at the University of Texas Cancer Center, M.D. Anderson Hospital. In 1989 he became a Leading Professor and Chair of the Department of Biochemistry and Cell Biology at SUNY at Stony Brook. In 1990 he founded and became Director of the Institute for Cell and Developmental Biology at Stony Brook.

Dr. Lennarz has served on many national and international committees. He has served as President of the Biochemistry Chairman's Organization, President of the American Society for Biochemistry and Molecular Biology and President of the Society for Glycobiology. He was a member of the Executive Committee of the International Union of Biochemistry and Molecular Biology for almost a decade.

He has presented special lectures at the University of Notre Dame, the NIH, the University of West Virginia, Johns Hopkins University, Florida State University, the University of California at San Diego, the University of Arkansas, Indiana University and the Medical College of Virginia.

He is a member of the National Academy of Sciences. The focus of his early work was on lipids and bacterial cell surfaces. More recent efforts have been in the structure, biosynthesis and function of cell surface glycoproteins. The biosynthesis studies initially were carried out in liver and oviduct, but these efforts now are focused in yeast. The functional studies have concentrated on the role of cell surface glycoproteins in fertilization and early development in the sea urchin and, more recently, the frog. For over 30 years Dr. Lennarz' research has been supported by federal sources, primarily the National Institutes of Health. Recently he was appointed Distinguished Professor and Chair of his department.

M. Daniel Lane

The Johns Hopkins University, School of Medicine, Baltimore,
Maryland, USA
Section: Metabolism, Vitamins and Hormones

M. DANIEL LANE received B.S. and M.S. degrees in 1951 and 1953 from Iowa State University and a Ph.D. in 1956 from the University of Illinois. He was a Senior Postdoctoral Fellow with Professor Feodor Lynen at the Max-Planck Institute Fur Zellchemie in Munich. Following faculty positions at Virginia Polytechnic Institute and New York University School of Medicine, he joined the faculty at the Johns Hopkins University School of Medicine in 1969 and served as DeLamar Professor and Director of the Department of Biological Chemistry from 1978 to 1997. He is presently Distinguished Service Professor at Johns Hopkins. In 2002 he received an honorary degree, Doctor of Humane Letters, from Iowa State University.

Dr. Lane was elected to membership in the National Academy of Sciences (in 1987) and was elected as a Fellow of the American Academy of Arts and Sciences (in 1982) and of the American Society of Nutritional Sciences (in 1996). He received the Mead Johnson Award from the American Society for Nutritional Sciences in 1966 for his research on biotin-dependent enzymes and in 1981, the William C. Rose Award from the American Society for Biochemistry and Molecular Biology for his work on the insulin receptor. In 1990–1991 Lane served as President of the American Society of Biochemistry and Molecular Biology. He has presented many named lectureships (including the

Feodor Lynen Lecture in Germany in 1999) and served on numerous editorial boards including the Journal of Biological Chemistry and the Annual Reviews of Biochemistry. Currently he is Associate Editor for Biochemical and Biophysical Research Communications.

Dr. Lane has published 280 research papers in major scientific journals. His early work focused on various enzymatic CO_2 fixation reactions, notably the mechanisms by which the B-vitamin, biotin, functions in enzymes to catalyze carboxylation. Dr. Lane's work on the regulation of acetyl-CoA carboxylase, the key regulatory enzyme of fatty acid synthesis, led him to his present interests which are to understand the basic mechanisms of lipogenesis, adipogenesis and the consequence of aberrations in these processes, most notably obesity. Research currently underway in his laboratory focuses on: (1) the genes that signal stem cell "commitment" to the adipocyte lineage and subsequent differentiation into adipocytes, and (2) the mechanisms by which the region of the brain, known as the hypothalamus, monitors and controls the drive to eat.

Associate Editors

Ernesto Carafoli

Università degli Studi di Padova, Padova, Italy
Section: Bioenergetics

ERNESTO CARAFOLI earned his M.D. degree at the University of Modena in Italy in 1957. After postdoctoral studies in the Laboratory of Albert L. Lehninger at Johns Hopkins University in the mid 1960s he returned to his home institution in Italy where he worked until 1973, when he was appointed Professor of Biochemistry at the Swiss Federal Institute of Technology (ETH) in Zurich. He returned to Italy in 1998 as a Professor of Biochemistry at the University of Padova, where he now also directs the newly founded Venetian Institute of Molecular Medicine (VIMM).

Dr. Carafoli became interested in calcium as a signaling agent during his post-doctoral days at Johns Hopkins. When he arrived there his main interests were in mitochondrial bioenergetics and it was thus natural for him to expand them to the newly discovered area of mitochondrial calcium transport. He was involved in most of the early discoveries in the field, and he continued to work on mitochondria and calcium after his return to Italy and until he moved to the ETH. There his interests still remained focused on calcium, but the emphasis shifted to the proteins that transport it across membranes and to those that process its signal. His favorite object of study became the calcium pumps, especially that of the plasma membrane, an enzyme which is essential to the regulation of calcium homeostasis and thus to the well being of cells. His contributions on the enzyme, especially after he purified it in 1979, have helped establishing most of its properties and have clarified important problems of mechanism, regulation and structure.

Dr. Carafoli has authored or co-authored about 450 peer-reviewed articles and reviews, and has edited or co-edited about 20 books. He has served on the Editorial or Advisory Boards of several periodicals and has organized about 30 International Workshops and Symposia. He has been featured as a plenary or honorary lecturer at numerous events ranging from specialized Workshops to International Symposia and Congresses. Dr. Carafoli's honors and awards include several international prizes and medals, memberships in several Academies, and three honorary degrees.

Don W. Cleveland

University of California, San Diego, La Jolla, CA, USA
Section: Cell Architecture and Function

DON W. CLEVELAND has been a longstanding contributor to the elucidation of regulation of assembly of mitotic spindles and chromosome movement and how errors in these contribute to the chromosome loss characteristic of human tumors. He discovered the tubulin gene families encoding the major subunits of microtubules and the first mammalian example of control of gene expression through regulated RNA instability. He identified components required for microtubule nucleation and anchoring during spindle assembly. He identified the first human centromeric protein (CENP-B). He then discovered CENP-E, the centromere-associated, microtubule-motor that he showed to be essential for chromosome attachment and for activation and silencing of the mitotic checkpoint, the cell cycle control mechanism that prevents errors of chromosome segregation in mitosis.

Dr. Cleveland has also been a leading force in dissecting the disease mechanism for major human neurodegenerative disorders. He initially purified and characterized tau, the microtubule-associated protein that assembles aberrantly in human dementias including Alzheimer's disease and Pick's disease. He established that the extreme asymmetry of neurons acquired during development is achieved with a deformable array of interlinked neurofilaments, microtubules and actin. He showed that disorganization of neurofilament arrays caused selective death of motor neurons in mice and humans. He also demonstrated that neuronal death could also arise by a toxicity of mutant superoxide dismutase unrelated to its normal activity, thereby uncovering the mechanism underlying the major genetic form of amyotrophic lateral sclerosis. He showed that this toxicity could be

sharply ameliorated by lowering the content of neurofilaments.

Dr. Cleveland is currently Head, Laboratory for Cell Biology in the Ludwig Institute for Cancer Research and Professor of Medicine, Neurosciences and Cellular and Molecular Medicine at the University of California at San Diego. He is also the Editor of the Journal of Cell Biology and Current Opinion in Cell Biology.

Jack E. Dixon

University of California, San Diego School of Medicine,
La Jolla, CA, USA
Section: Protein/Enzyme Structure, Function, and Degradation

JACK E. DIXON earned his Ph.D. in Chemistry at the University of California, Santa Barbara in 1971 and did his postdoctoral training in Biochemistry at the University of California, San Diego.

Dr. Dixon is a pioneer and leader in the structure and function of the protein tyrosine phosphatases (PTPases). He demonstrated that the unique catalytic mechanism of the PTPases proceeds via a novel cysteine-phosphate intermediate. He discovered the first dual-specificity phosphatase, which led to the identification of the cell cycle protein, $p80^{cdc25}$, as a phosphatase. He also showed that the bacteria responsible for the plague or "black death" harbor the most active PTPase ever described. He and his colleagues went on to demonstrate that this PTPase gene product is essential for the pathogenesis of the bacteria. Dr. Dixon and his colleagues determined X-ray structures for both tyrosine and dual specificity phosphatases. Dr. Dixon also found that sequences outside of the PTPase catalytic domain could function to direct the subcellular localization of the PTPases and to restrict their substrate specificity. This is now a widely acknowledged regulatory paradigm for the PTPases. Recently, his laboratory demonstrated that the tumor suppressor gene, PTEN, which shares sequence identity with the PTPases, catalyzes the dephosphorylation of a lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate (PIP3). This represents the first example of a PTPase dephosphorylating a lipid second messenger. PIP3 activates the protein kinase, AKT, which plays a critical role in controlling the balance between apoptosis and cell survival. The loss of the PTEN gene elevates PIP3 levels leading to constitutive activation by AKT and oncogenesis. Recently, Dr. Dixon in collaboration with Nikola Pavletich determined the X-ray structure of PTEN. Their structure–function studies explain the PIP3 substrate specificity of PTEN and also provide a rationale for many of the mutations seen in human cancers. Earlier in his career, Dr. Dixon adopted the tools of molecular biology as they became available in the 1970s, and his laboratory was among the first to use synthetic oligonucleotides to isolate and extensively characterize cDNAs encoding peptide hormones.

Dr. Dixon is Professor of Pharmacology, Cellular and Molecular Medicine and Chemistry and Biochemistry and Dean of Scientific Affairs at the University of California, San Diego. He is a member of the National Academy of Sciences, the Institute of Medicine and the American Academy of Arts and Sciences. Dr. Dixon was the recipient of the 2003 William C. Rose Award from the American Society for Biochemistry and Molecular Biology.

John H. Exton

Howard Hughes Medical Institute, Vanderbilt University School of
Medicine, Nashvillie, TN, USA
Section: Signaling

JOHN H. EXTON was born and educated in New Zealand where he received his medical training and a Ph.D. in Biochemistry from the University of Otago in 1963. He did postdoctoral work at Vanderbilt University under Charles R. Park and Earl W. Sutherland, and became an Investigator of the Howard Hughes Medical Institute in 1968 and Professor of Physiology in 1970. He is presently Professor of Molecular Physiology and Biophysics, Professor of Pharmacology and a Hughes Investigator at Vanderbilt.

Dr. Exton's research initially focused on the changes in carbohydrate metabolism in liver during diabetes and treatment with various hormones using the perfused rat liver as the experimental system. His work concentrated on gluconeogenesis and identified the enzymatic reactions that were under control by insulin, epinephrine (adrenaline), glucagon and glucocorticoids, and demonstrated the importance of cyclic AMP in the regulation of these reactions. The role played by the supply of substrates, especially of alanine, was also shown.

Dr. Exton then turned his attention to the action of epinephrine (adrenaline) and demonstrated that many of its actions were not mediated by cyclic AMP but by calcium ions. This led to study of the breakdown of inositol phospholipids by phospholipase C that underlay the increase in calcium. Later this resulted in the discovery of G_q, a novel G protein that activated phospholipase C. Further studies demonstrated that agonists caused the breakdown of another phospholipid (phosphatidylcholine) by another phospholipase (phospholipase D). Current work is focused on the physiological role of phospholipase D.

Dr. Exton has authored over 350 scientific articles and is presently an Associate Editor of the Journal of Biological Chemistry. He has served on many scientific review groups and as a reviewer for many journals. He has won numerous awards, most notably the Lilly

Award of the American Diabetes Association, Fellow of the American Association for the Advancement of Science and election to membership in the National Academy of Sciences.

Paul Modrich

Duke University Medical Center, Durham, NC, USA
Section: Molecular Biology

PAUL MODRICH is an Investigator of the Howard Hughes Medical Institute and James B. Duke Professor of Biochemistry at Duke University Medical Center. He received his undergraduate degree from M.I.T. and his Ph.D. in Biochemistry from Stanford University. His current research addresses the mechanisms of DNA repair. He has served on the editorial boards of the Journal of Biological Chemistry, Biochemistry, Proceedings of the National Academy of Sciences, and DNA Repair. His honors include election to National Academy of Sciences and the Institute of Medicine, the Pfizer Award in Enzyme Chemistry, the General Motors Mott Prize in Cancer Research, and the Pasarow Foundation Award in Cancer Research.

Preface

Biological Chemistry is defined as the chemistry of the compounds and processes that constitute living organisms. The ultimate goal, of course, is to understand and define biology at a mechanistic level. This was aptly stated in an historical treatise on the founding of the *Journal of Biological Chemistry*, where John Edsall quoted a statement in a letter from J. L. Loeb (in Berkeley), "The future of biology lies with those who attack its problems from a chemical point of view." What was an emerging field in 1900 with its origins in physiology, nutrition and chemistry has broadened and expanded to include numerous other fields including mechanistic enzymology, molecular biology, structural biology, cell biology, genomics, proteomics, bioinformatics, metabolomics and others, that were not defined as discrete fields at that time.

Modern biochemistry (biological chemistry) began with the accidental discovery by Eduard Buchner in 1897 that a cell-free yeast extract could carry out fermentation of glucose to alcohol and CO_2 *in the absence of intact cells*. He named the dissolved substance responsible for this process zymase, the substance(s) we now refer to as enzymes. Importantly, Buchner recognized the significance of his discovery. This ended the dogma of the time, perpetuated by Pasteur, the concept of *vitalism*; i.e., that fermentation (and presumably other complex biological phenomena) required the action of intact cells. Thus, serendipity and a prepared mind ushered in a new era of discovery. Now it became possible to dissect complex physiological processes and to study them with preparations free of the constraints of intact cells. Once a metabolic pathway/process was established, it became possible to purify the enzymes, cofactors and substrates involved, to reconstitute the process with purified components and to characterize the components chemically. What followed was an information explosion in the field of biochemistry and progression through a series of trends, each "in vogue" in its time. The identification of the dietary essentials, the hunt for the vitamins/cofactors, the hormones, identification of metabolic pathways and the enzymes involved, oxidative phosphorylation, protein synthesis, molecular biology—each developed as a primary focus.

The need to associate chemistry with function came early and was evident in the naming of departments and journals. Over time names changed from Agricultural Chemistry to Physiological Chemistry to Biochemistry to Biological Chemistry. An example is the Department of Biochemistry at the University of Wisconsin, which began in 1883 as the Department of Agricultural Chemistry.

Where are we headed? We have reached the point where the borders of these areas have become blurred. What constitutes cell biology, molecular biology/genetics, developmental biology, physiology, immunology—ultimately reduces to chemistry. To understand these processes we must know what the molecules are and understand how they interact, i.e. the basic chemistry. That is what this encyclopedia is about.

The breadth of content of this encyclopedia aims to cover major topics of modern biochemistry, each authored by an expert in the area. We feel that the coverage is broad and we have been inclusive in choice of topics. The encyclopedia is a reference work encompassing four volumes containing over 500 articles with more than 750 authors or coauthors. Each article/topic covers an important area of the field which reflects the point of view of the authors. Together the articles cover virtually every aspect of biology for which we have "mechanistic" information. For those who wish to probe more deeply into a topic, references to further readings are included at the end of each article. The editorial board that made decisions on coverage consists of seven members, each an expert representing a major area in the field of biochemistry. A dedicated effort was made to provide coverage that is as complete as possible. The content is presented at a level that we hope will be interpretable to interested individuals with some background in chemistry and biology. It is intended for such individuals rather than specialists with extensive scientific backgrounds in specific areas. It is aimed at the generalist as opposed to the specialist.

Finally, we would like to single out Gail Rice and Dr. Noelle Gracy for their enormous contribution in putting this encyclopedia together. They, in fact, were a driving force that brought this major work to completion.

Notes on the Subject Index

Abbreviations used in subentries without explanation:

CoA	coenzyme A	NADPH	nicotinamide-adenine dinucleotide phosphate
DAG	diacylglycerol		
ELISA	enzyme-linked immunosorbent assay	PFK-2/	6-phosphofructo-2-kinase/fructose-
ERK	extracellular-signal regulated kinase	FBPase-2	2,6-bisphosphatase
GlcNAC	N-Acetylglucosamine	PI3K	phosphatidylinositol 3-kinase
HPLC	high-pressure liquid chromatography	PIP_2	phosphatidylinositol 4,5-bisphosphate
IP_3	inositol 1,4,5-triphosphate	PIP_3	phosphatidylinositol-3,4,5-triphosphate
MAP	mitogen-activated protein	PPAR	peroxisome proliferator-activated receptor
MMP	matrix metalloproteinase	RPLC	reversed-phase high-performance liquid chromatography
mtDNA	mitochondrial DNA		

Volume 1

Volume 2

Volume 3

Volume 4

Lipids, Carbohydrates, Membranes and Membrane Proteins

Metabolism, Vitamins and Hormones

Cell Architecture and Function

Protein/Enzyme Structure Function and Degradation

Allosteric Regulation; Vol.1 – Pages 68-73, Barry S. Cooperman

Aminopeptidases; Vol.1 - Pages 96-98, Ralph A. Bradshaw

Amyloid; Vol.1 - Pages 99-104, Ronald Wetzel

Aspartic Proteases; Vol.1 – Pages 123-127, Ben M. Dunn

B12-Containing Enzymes; Vol.1 - Pages 145-151, Vahe Bandarian and Rowena G. Matthews

Biotin; Vol.1 - Pages 174-178, Steven W. Polyak and Anne Chapman-Smith

Biotinylation of Proteins; Vol.1 - Pages 179-181, Ronald A. Kohanski

Calpain; Vol.1 – Pages 300-306, Hiroyuki Sorimachi and Yasuko Ono

Chemiluminescence and Bioluminescence; Vol.1 – Pages 399-404, Thomas O. Baldwin

Coenzyme A; Vol.1 – Pages 475-477, M. Daniel Lane

Collagenases; Vol.1 – Pages 478-481, Kenn Holmbeck and Henning Birkedal-Hansen

Collagens; Vol.1 – Pages 482-487, Darwin J. Prockop

Cysteine Proteases; Vol.1 – Pages 516-520, David J. Buttle and John S. Mort

Disulfide Bond Formation; Vol.1 – Pages 598-602, Hiram F. Gilbert

Elastin; Vol.2 – Pages 10-12, Judith Ann Foster

Enzyme Inhibitors; Vol.2 – Pages 31-37, Vern L. Schramm

Enzyme Kinetics; Vol.2 – Pages 38-44, Irwin H. Segel

Enzyme Reaction Mechanisms: Stereochemistry; Vol.2 – Pages 45-50, Ming-Daw Tsai, Li Zhao and Brandon J. Lamarche

Flavins; Vol.2 – Pages 118-122, Barrie Entsch and David P. Ballou

Heme Proteins; Vol.2 – Pages 354-361, Johannes Everse

Bioenergetics

Molecular Biology

Signaling

Techniques and Methodology

Affinity Chromatography; Vol.1 - Pages 51-56, Pedro Cuatrecasas and Meir Wilchek

Affinity Tags for Protein Purification; Vol.1 - Pages 57-63, Joseph J. Falke and John A. Corbin

Genome-Wide Analysis of Gene Expression; Vol.2 - Pages 175-180, Karine G. Le Roch and Elizabeth A. Winzeler

HPLC Separation of Peptides; Vol.2 - Pages 398-403, James D. Pearson

Imaging Methods; Vol.2 - Pages 405-410, Gyorgy Szabadkai and Rosario Rizzuto

Inorganic Biochemistry; Vol.2 - Pages 417-420, Robert J. P. Williams

Multiple Sequence Alignment and Phylogenetic Trees; Vol.2 - Pages 770-774, Russell F. Doolittle

Oligosaccharide Analysis by Mass Spectrometry; Vol.3 - Pages 155-160, Andrew J. Hanneman and Vernon N. Reinhold

PCR (Polymerase Chain Reaction); Vol.3 - Pages 208-210, Michael J. Brownstein

Polysialic Acid inMolecular Medicine; Vol.3 - Pages 407-414, Frederic A. Troy, II

Protein Data Resources; Vol.3 - Pages 478-483, Philip E. Bourne

Secondary Structure in Protein Analysis; Vol.4 - Pages 1-6, George D. Rose

Spectrophotometric Assays; Vol.4 - Pages 67-75, Britton Chance

Two-Dimensional Gel Electrophoresis; Vol.4 - Pages 284-289, Gerhard Schmid, Denis Hochstrasser and Jean-Charles Sanchez

X-Ray Determination of 3-D Structure in Proteins; Vol.4 - Pages 422-428, Martha L. Ludwig

ENCYCLOPEDIA OF

Biological Chemistry

Editors

William J. Lennarz

M. Daniel Lane