



allosteric activator



PRODUCT FORMATION

1

The Machines and Assemblies of Life

The mechanics of life—indeed life itself—depends on processes operating in that crucial level in the structural hierarchy that lies between individual macromolecules and cells or organelles: the realm of macromolecular assemblies. This book describes the assemblies that function in particular biological contexts. Overall we consider one class—macromolecular machines, viewed as assemblies that have an overtly mechanical aspect to their biological activities and, like macroscopic machines, do mechanical work—alongside other kinds of assembly with different properties. The latter assemblies nevertheless exhibit similar complexity in their structures and interactions and in the sophistication with which their synthesis is regulated. Chapter 1 sets the scene, surveying key concepts relating to the interactions that define macromolecular structures, govern their folding, and orchestrate their assembly. We start at the near-atomic scale, discussing interactions between molecules, and work up to the cellular level. En route, we dwell on the symmetries that underlie the architecture of many assemblies, and on the dynamic properties essential to their functioning: assemblies are far from static structures! Recognizing the fundamental need for communication within and between assembly systems, we examine the essentials of information storage, retrieval, and transmission, and, finally, consider assemblies as they operate in the crowded interiors of cells, and their place in the timeline of evolution.

1.1 EXPRESSION OF THE GENETIC BLUEPRINT

The ‘Central Dogma’ of molecular biology formulated by Francis Crick in 1958 described a one-way flow of sequence information between the three major classes of biopolymers (DNA, RNA, and protein), later paraphrased by Marshall Nirenberg as “DNA makes RNA makes protein.” *Transcription* and *translation* are the processes by which cells read out and then express the information stored in genes, which are segments of chromosomal DNA. Transcription transfers the information encoded in a gene into an RNA nucleotide sequence—a messenger RNA (mRNA)—that is complementary to one of the DNA strands. In translation, the RNA nucleotide sequence is rendered into a sequence of amino acids, a polypeptide chain. So what makes DNA—whose replication is essential for the maintenance and transmission of genetic information and involves the unwinding of the double helix and the synthesis of two complementary strands (**Figure 1.1**)? In fact, DNA and RNA on their own make very little. We now know that all the steps of information storage, copying, and transfer are accomplished by elaborate macromolecular machines. These are mostly multi-subunit protein complexes but some also have RNA components. None has a DNA component.

Transcription is mediated by *RNA polymerases* and steered by other proteins called *transcription factors*. In bacteria and archaea, the processes of transcription and translation are not spatially segregated within the cell, and the mRNA is used directly. In *eukaryotes*, transcription takes place in the nucleus and the mRNA is exported into the cytoplasm for translation; first, however, the primary transcript (pre-mRNA) is modified by protein ‘capping’ factors and in many cases is processed further by splicing. In splicing, RNA segments derived from coding regions of DNA (*exons*) are fused together and the intervening noncoding segments (*introns*) are excised (**Figure 1.2**). Splicing is performed by RNA/protein complexes called *spliceosomes*. Combining exons in different ways (alternative splicing) expands the repertoire of mRNAs derived from a single gene. The mature mRNA then associates with a *ribosome*, a large and highly conserved RNA/protein complex, which synthesizes the corresponding polypeptide chain. For the protein to become functional, the polypeptide chain

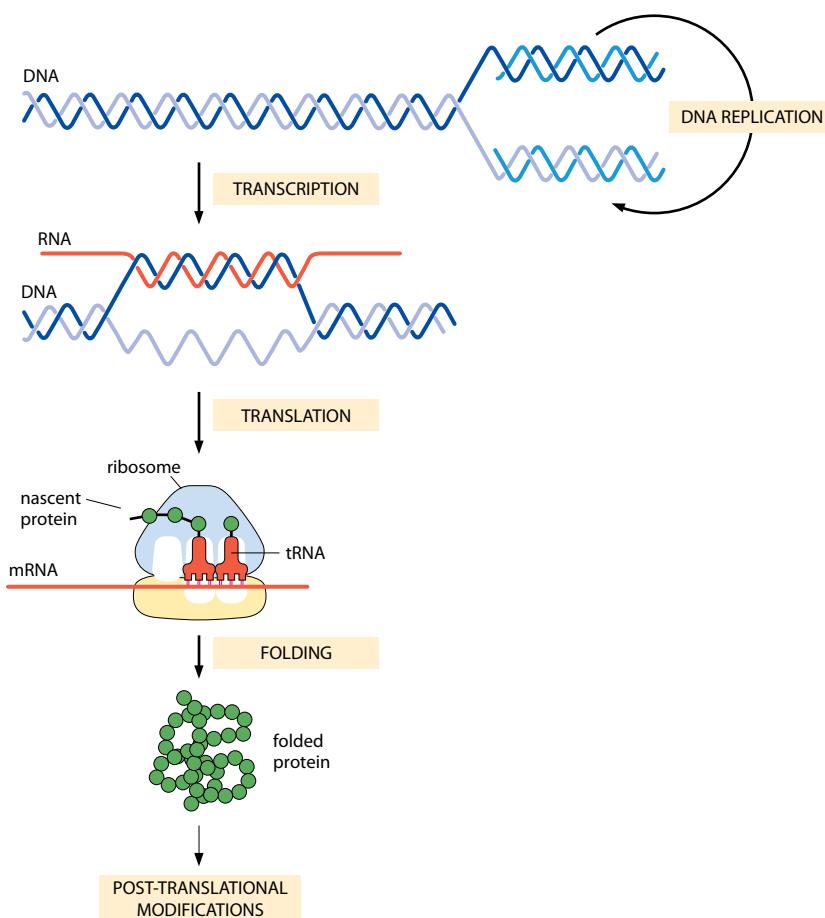


Figure 1.1 The flow of genetic information from chromosomal DNA to a natively folded protein.

must fold into a specific three-dimensional structure, a spontaneous process that in many cases is facilitated by other proteins, called *molecular chaperones*.

The flow of information is not perfect and not always in one direction

DNA replication is carried out by another protein machine, the *replicase*, which consists of *DNA polymerases* and associated proteins. The fidelity of replication is high (error rate $\sim 10^{-9}$ per base pair) but it is not a perfect process: this would be thermodynamically impossible; moreover, the ability to evolve—essential to life—depends on the occasional error. The error rate for mRNA synthesis or translation is much higher, $\sim 10^{-4}$, which is permissible because mRNA and proteins are not information storage molecules and, if faulty, can be replaced.

It has turned out that there are exceptions to this unidirectional flow of information from DNA to RNA to protein. One important exception is **reverse transcription**, the transfer of information from single-stranded RNA to double-stranded DNA. This is the mechanism employed by retroviruses (hence the term *retro-virus*) to replicate their RNA genomes, and by eukaryotic cells in synthesizing **telomeres** to protect the ends of their chromosomal DNA. Many other viruses also have RNA, not DNA, as their genetic material and they reproduce it by direct **RNA-dependent RNA replication**. The RNA polymerase machines responsible are generally encoded by viral genes, unlike many other functions for which the virus relies on resources provided by the host cell.

RNAs that act in a regulatory way are also turning out to be of major importance in eukaryotes. A multicellular organism consists of cells of many different kinds, and yet the chromosomal DNA is exactly the same in each cell. For a given organism, the differences in size, shape, and function of its various cells reflect the differing combinations of genes that are expressed in them. A simple eukaryote, the nematode worm *Caenorhabditis elegans*, and

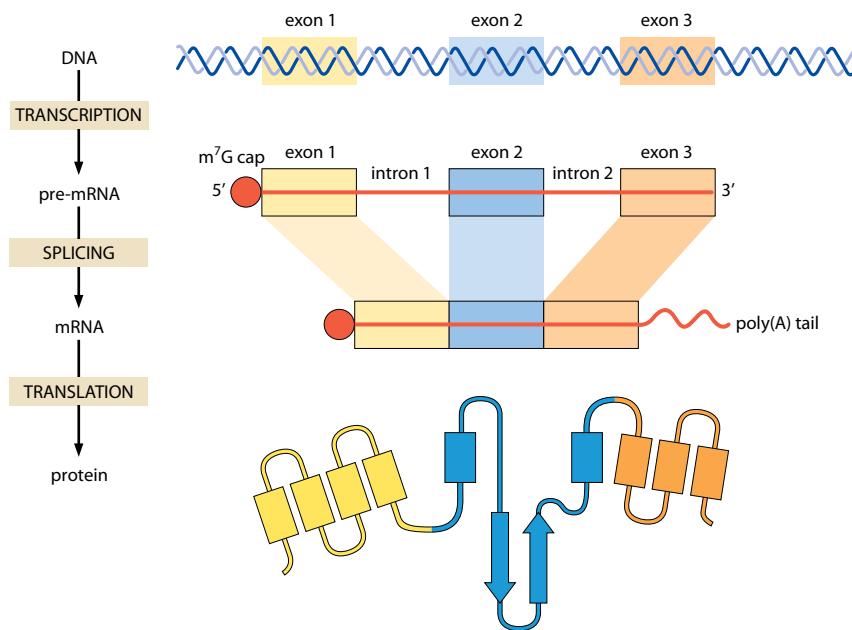


Figure 1.2 Transcription and translation of a eukaryotic gene with three exons.

a more complex one, *Homo sapiens*, both have ~20,000 genes but they differ vastly in cell number (~ 10^3 for the nematode, in contrast with ~ 10^{14} for humans) and cell differentiation. In human DNA, less than 2% of its 3 billion base pairs encode proteins, while much of the remaining noncoding DNA is transcribed into RNAs that have crucial roles in regulating gene expression and the translation of mRNAs. With the nematode, the fraction of protein-coding DNA is much higher, at ~33%. Most probably, it is differences in the complexity of this regulatory RNA world that underlie the huge differences in anatomical and physiological complexity between the worm and the human.

Nor should the effects of chemical modification of DNA and proteins be overlooked. In vertebrates, chemical modification of genomic DNA can affect gene expression without changing the underlying nucleotide sequence. DNA methylation, for example, can lock genes into a silent state. If heritable, effects of this kind are termed *epigenetic* and can have important roles in controlling expression of the genetic blueprint, such as X-chromosome inactivation in females or the preferred expression of maternally or paternally derived genes.

Most eukaryotic proteins undergo *post-translational* changes to their amino acid sequences that modify or modulate their activities. These include selective cleavage by proteases, the self-excision of peptide segments (inteins), and the covalent attachment of polypeptide ‘markers,’ as in *ubiquitylation*. In both prokaryotes and eukaryotes, an extensive range of post-translational modifications expands the set of amino acid side chains far beyond the 20 specified by DNA codons, without requiring genetic changes. Reversible phosphorylation is of particular importance in regulating biological activity in eukaryotes.

A glaring exception to the classical Central Dogma is given by *prions*, the agents of transmissible neuropathic diseases (spongiform encephalopathies) in mammals or of metabolic regulation in fungi. A prion is an infectious aggregate of protein in a fibrillar non-native conformation (termed *amyloid*) that serves as a template for recruiting the native protein into the aggregate, changing its conformation as it does so. In the replication of a prion, there is no participation of DNA or RNA other than in synthesis of the original proteins; here, protein makes protein. Prions differ from other amyloids in their infectivity, whereby the prion is somehow transferred into another previously uninfected cell or organism.

1.2 WEAK FORCES AND MOLECULAR INTERACTIONS

Covalent bonds are required for the synthesis of individual macromolecules; phosphodiester bonds and peptide bonds are of particular importance in nucleic acids and proteins respectively. In a few situations they also contribute to protein folding, stability, and

assembly. Examples include the disulfide bridges found in many secreted proteins, the isopeptide cross-links formed between pairs of Lys and Asn residues in certain viruses and in some bacterial *pili*, and the isopeptide bond formation between Lys and Gln residues catalyzed by transglutaminase enzymes in maintaining the epidermis, the protective barrier of skin. With these few exceptions, it is noncovalent interactions (**weak forces**) that govern protein folding and oligomerization, the assembly of higher-order structures, the structures of nucleic acids and membranes, and every aspect of molecular recognition.

The importance of these noncovalent interactions cannot be overstated. They are of four kinds: electrostatics, hydrogen bonds, van der Waals interactions, and hydrophobic interactions. All must operate in the context of water as medium. Individually, all are much weaker ($\sim 8\text{--}38 \text{ kJ/mol}$) than a typical covalent bond ($\sim 350\text{--}450 \text{ kJ/mol}$). However, in macromolecular assemblies, their effect is amplified by their being deployed in combinations, forming complementary **interaction patches** whose organization imposes the specificity of an interaction.

All weak forces other than hydrophobic interactions are electrostatic in origin

Electrostatic interactions can be attractive or repulsive. With proteins, they include the attraction between a negatively charged residue (Asp or Glu) and a positively charged residue (Arg or Lys) in what is often called a ‘salt bridge,’ after the paired ions, Na^+ and Cl^- , in common salt. The electrostatic potential is a long-range effect, proportional to $1/r$ (**Figure 1.3A**). The force exerted between the charges (which is the derivative of the potential) falls off as $1/r^2$. Fully calibrated, the potential energy per mole for point charges q_1 and q_2 is given by

$$U = q_1 q_2 N_A / 4\pi \epsilon_0 \epsilon r$$

where N_A is Avogadro’s number ($6.02 \times 10^{23}/\text{mol}$); ϵ_0 is the permittivity of free space, $8.85 \times 10^{12} \text{ C}^2/(\text{m}^2 \text{ N})$; ϵ is the dielectric constant of the medium; $1 \text{ J} = 1 \text{ N} \times \text{m}$; 1 elementary charge (that is, + on a proton or – on an electron) = 1.60×10^{-19} coulomb (C). In solution, this interaction is weakened by the high dielectric constant (80) of water and the solvation of charged groups.

Lys and Arg side chains are not fully interchangeable. The $\text{N}^6\text{-NH}_3^+$ group of Lys is essentially a point positive charge, whereas the delocalization of the positive charge over the guanidinium group of Arg enables it to impose directionality in its interaction with a delocalized negative charge. One example is the interaction with the negatively charged carboxyl group (**Figure 1.4A**); another is that with a phosphate ($-\text{OPO}_3^{2-}$) group, which explains why Arg is so often found in binding sites, especially those for nucleotide substrates. The surface charges on a protein molecule generate an electrostatic field, and this influences the orientation of potential interaction partners, as exemplified by the field surrounding the enzyme superoxide dismutase (**Figure 1.4B**).

Another type of electrostatic interaction is the **hydrogen bond** (H-bond), formed when the hydrogen atom attached to an electronegative atom approaches another electronegative atom and there is a partial sharing of the proton. In crystal structures of proteins, where hydrogen atom positions are not usually determined, an H-bond is assumed if two electronegative atoms (the donor D and acceptor A) are closer than 3.3 \AA and the enclosed angle is $>90^\circ$ (**Figure 1.5A**). In proteins, most H-bonds are of the type $\text{N-H}\cdots\text{O}$ or $\text{O-H}\cdots\text{O}$, and

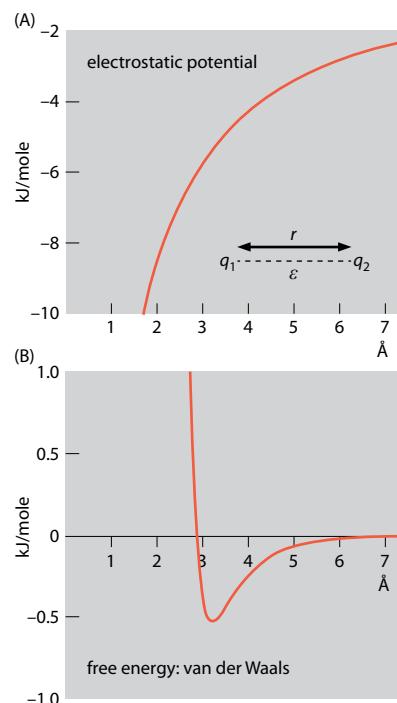


Figure 1.3 Distance dependence of electrostatic and van der Waals interactions. (A) Electrostatics. The potential energy per mole for point charges q_1 and q_2 , distance r apart. This curve shows the field associated with an attractive interaction between oppositely charged residues in water. At $r = 3 \text{ \AA}$, $U = -5.7 \text{ kJ/mole}$. In a protein interior, where $\epsilon \approx 4$, $\Delta U = -115 \text{ kJ/mol}$ for this interaction. (B) A van der Waals energy curve. Energy is plotted as a function of separation of atoms (r), showing the net sum between repulsive energy ($1/r^{12}$ dependence) and the attractive energy ($1/r^6$ dependence) with a minimum at 3.2 \AA , corresponding to an energy $E_m = -0.54 \text{ kJ/mol}$. (Adapted from G.E. Schulz and R.H. Schirmer, Principles of Protein Structure. New York: Springer-Verlag, 1979.)

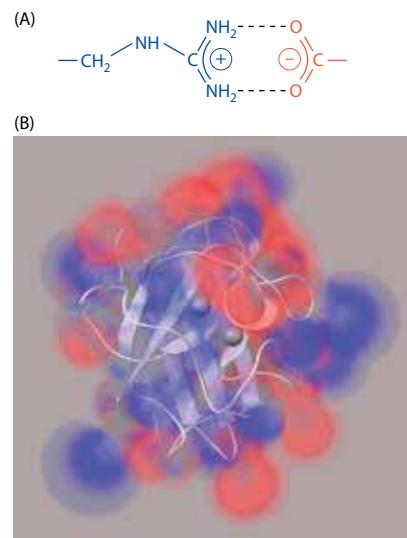


Figure 1.4 Directionality in electrostatic interactions. (A) Bidentate interaction between delocalized charges over the guanidinium group of arginine (blue) and a carboxyl group (red) restricts the relative orientation of the two charges. (B) The enzyme Cu-Zn superoxide dismutase (SOD) functions as a homodimer which, through its electrostatic field, attracts the negatively charged superoxide, O_2^- , a reactive oxygen species (see Section 15.2), to the positively charged metal ions in its catalytic sites, where it is converted to the less toxic hydrogen peroxide. A monomer of bovine SOD is shown, with its backbone depicted as a white ribbon diagram and the Cu^{2+} and Zn^{2+} in the catalytic site as spheres. The surrounding electrostatic field of the protein subunit is displayed as a volume-rendered cloud with positive potential in blue and negative potential in red. (PDB 2SOD) (B, courtesy of A. Olson.)

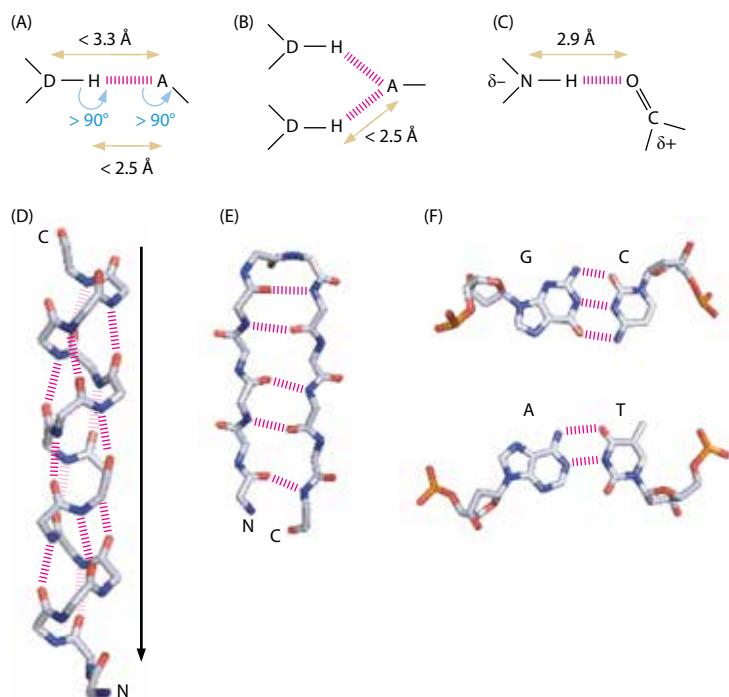


Figure 1.5 H-bond interactions. The bonds are marked by ladders of red dashes. In (A), D is the donor atom and A is the acceptor atom. The standard geometrical criteria for identifying an H-bond in a protein are indicated. (B) Bifurcated H-bond with two sharing donors. (C) The H-bond between peptide groups, showing the partial electric charges ($\delta-$ and $\delta+$) on the atoms. (D) The main chain conformation of an α helix in which the peptide NH of residue $i + 4$ forms an H-bond with the C=O of residue i . Alignment of the peptide dipoles generates an overall dipole for the helix (arrow) with a small positive charge at the N-terminal end. Side chains are not shown. (E) The main chain conformation of a two-stranded anti-parallel β sheet (that is, a β hairpin) with H-bonds between the strands. (F) Watson-Crick H-bonds between G base-paired with C, and A base-paired with T.

less often N-H \cdots N (Figure 1.5C), and almost all (~98%) potential donors and acceptors are found with partners. Sometimes one donor has two acceptors or *vice versa*, generating a bifurcated H-bond (Figure 1.5B). H-bonds have energies of -12 to -38 kJ/mol, depending on the local environment, and they are directional in character. The bond energy is highest when the two atoms and the proton lie on a straight line; a 30° departure from linearity reduces it by about 30%. Two crucial roles (among many) of H-bonding lie in the interactions between main-chain atoms that stabilize the secondary structure of proteins, namely the α helices, β sheets, and turns (Figure 1.5D-E), and in the base pairing of nucleotides in duplex DNA (Figure 1.5F).

Charged (and uncharged) groups also engage in other electrostatic-based interactions. Even nonpolar groups (meaning, here, a group lacking a permanent dipole) develop transient dipoles that interact with similarly fluctuating dipoles in neighboring groups. These interactions, known collectively as **van der Waals interactions**, are short-range and represent a balance between mutual repulsion of the respective electron clouds as atoms come into close contact, for which the energy has a $1/r^{12}$ dependence, and attraction from dipole/dipole interactions whose energy falls off as $1/r^6$. The equilibrium distance is ~ 3.2 Å (see Figure 1.3B). The net forces are individually weak: for example, a methylene group in a crystalline hydrocarbon has a van der Waals energy of ~ -8.4 kJ/mol. However, their influence is enhanced by multiplicity, as in the core of a folded protein, the central region of a *lipid bilayer*, or the base stacking that stabilizes duplex DNA and RNA.

Hydrophobic interactions drive the folding and assembly of macromolecules

The H₂O molecule is unusual. It has a permanent dipole moment and can participate in four H-bonds, two as donor and two as acceptor. Thus water molecules have fluctuating structures based on a tetrahedral H-bonded arrangement: typically, each H₂O participates in three or four H-bonds at any one time. The dual donor/acceptor property means that water molecules often engage in H-bonded networks linking different parts of a macromolecule (Figure 1.6). In bulk water, the free energy change when an H-bond formed is small because it is accompanied by rupture of the H-bonds that each previously made with other water molecules. Some macromolecules have ‘buried’ water molecules that form an integral part of their structure separate from the bulk water. A buried water/protein H-bond contributes ~ -2.5 kJ/mol to the stability of a protein.

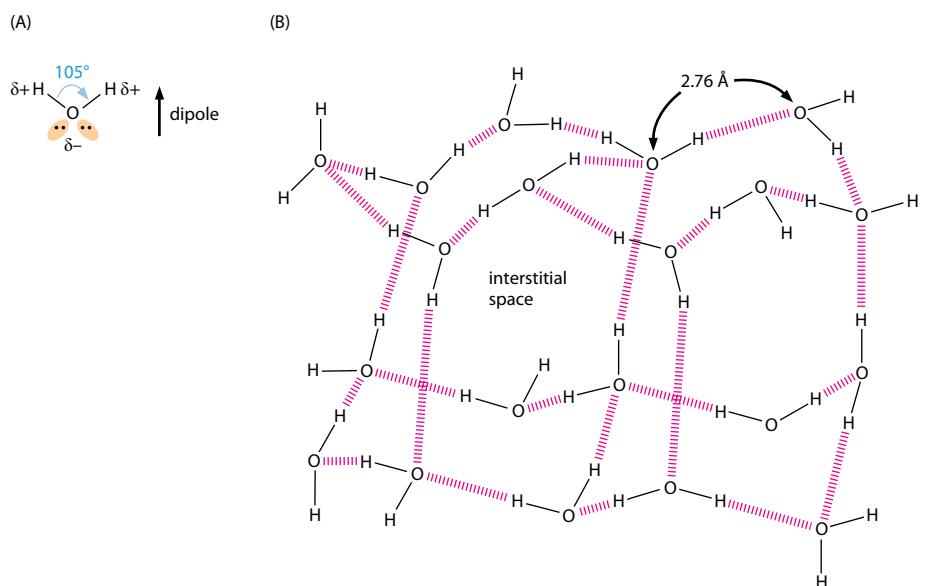


Figure 1.6 Water as a polar solvent.

(A) The dipole of a water molecule. The oxygen sp^3 orbitals are shown as yellow ellipses containing the lone pair of electrons that act as acceptors in H-bonds. (B) A possible arrangement of H-bonds (red dash ladders) for a network of water molecules in solution.

However, electrostatic-based forces do not explain why nonpolar amino acids tend to congregate in the interior of proteins, leaving most charged residues on the surface and exposed to the solvent. Buried charged residues, when they occur, are almost always neutralized by pairing with complementary charges. Uncharged polar groups are found both on the surface and in the interior (Figure 1.7); in the latter, they usually engage in H-bonding. Thus, the interior of a folded protein is a hydrophobic milieu where nonpolar groups engage in van der Waals interactions with other nonpolar groups. This partitioning is due to **hydrophobic effects**.

The presence of a nonpolar group in water is energetically unfavorable on account of its effect on the local organization of water molecules, which no longer assume fluctuating tetrahedral structures but form a clathrate (cage-like) structure around the nonpolar group. For example, it takes a minimum of 20 water molecules to make a clathrate around CH_4 , and their partial ordering reduces the overall number of degrees of freedom. The corresponding decrease in **entropy** (an energy-related measure of disorder) is relieved if the nonpolar group associates with another nonpolar group, releasing water molecules from the interface. This more than compensates for the loss of configurational entropy from the association of the two nonpolar molecules.

Hydrophobicity may be measured in terms of the free energy change when a nonpolar group is transferred from water into a nonpolar solvent. Some examples are given in Table 1.1. Hydrophobic interactions are nondirectional and they are relatively long-range, making them early contributors to protein folding and assembly. The lack of directionality means that a single hydrophobic group cannot give rise to specificity, but the way in which hydrophobic residues are distributed over an *interaction patch* does confer this property (see Figure 1.7).

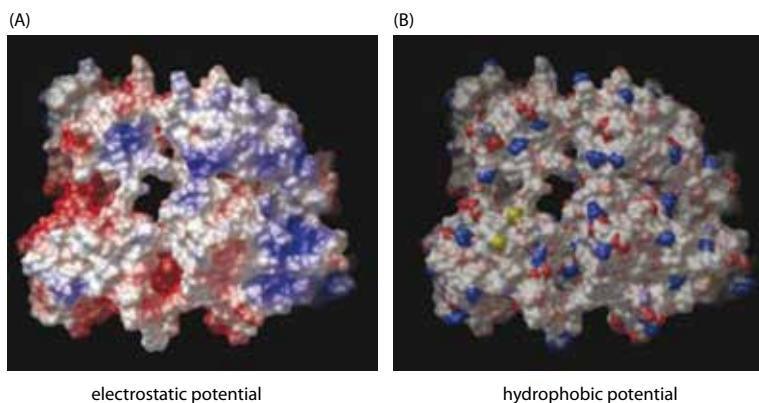


Figure 1.7 Mapping distributions of charge and hydrophobicity over a protein surface.

The images were rendered with PMV (Python Molecular Viewer). (A) The electrostatic potential (positive, blue; negative, red) of a tyrosine kinase was computed with APBS (Adaptive Poisson–Bolzmann Solver). (PDB 1GCF) (B) Hydrophobic potential at the molecular surface shown in a coloring scheme in which all carbon atoms (hydrophobic) are white, while polar atoms are desaturated red and blue, and charged atoms are saturated red and blue. This gives a mixed-color solvent-excluded surface with charged and polar atoms distributed among the white (hydrophobic) carbon atoms. The small yellow patches are sulfur atoms in Cys residues. (From S.I. O'Donoghue et al., *Nat. Methods Suppl.* 7:S42–S55, 2010. With permission from Macmillan Publishers Ltd.)

Table 1.1 Free energy changes on the transfer of hydrophobic groups from water to n-octanol.

Group (relative to H atom)	Free energy change (kJ/mol)
-CH ₃	-2.9
-CH ₂ CH ₃	-5.7
-(CH ₂) ₂ CH ₃	-8.6
-CH ₂ -phenyl	-15.0
-OH	6.6

(Adapted from A. Fersht, Structure and Mechanism in Protein Science. New York: WH Freeman, 1999.)

Intrinsic membrane proteins represent a special case in that they have transmembrane domains consisting mainly of amino acids with hydrophobic side chains that are embedded in the hydrophobic phase of a lipid bilayer. The main chain atoms of these segments of polypeptide chain compensate for their polar nature by H-bonding, most commonly in the form of α helices and occasionally in β sheets.

The energy balance in folding and assembly has both enthalpic and entropic contributions

The energetic exchanges in transitions involving biological (and other) macromolecules are related by the equation

$$\Delta G = \Delta H - T\Delta S$$

where ΔG is the change in Gibbs free energy, ΔH is the change in **enthalpy** (the heat given up or taken in by the system), ΔS is the change in entropy, and T is the absolute temperature (in kelvins). In general, if a system is not already at a free energy minimum, it will spontaneously shift toward a lower (more negative) free energy state. Note, however, that ΔG tells nothing about kinetics, in other words the rate at which a new equilibrium will be reached. In folding and assembly, the enthalpic term (ΔH) with contributions from charge/charge interactions, H-bonds, and van der Waals interactions is generally negative (favorable). The entropic term ($-T\Delta S$) may be positive or negative, depending on whether the loss of configurational entropy on entering a more ordered state exceeds the gain in entropy from liberating water molecules. When ΔH is close to zero, a positive entropy contribution (making $-T\Delta S$ negative) is decisive. This is the basis of entropy-driven processes, which include the assembly of *microtubules* (Section 14.5), *tobacco mosaic virus* (Section 8.2), or *collagen* (Section 11.7) fibrils. Ordering processes of this kind have the counter-intuitive property of being promoted by a rise in temperature, which increases the value of $-T\Delta S$: (normally, raising the temperature promotes kinetic disorder, leading to dissociation).

Size and topography matter for interaction patches

The concept of **solvent-accessible surface** was introduced by B-K. Lee and Frederic Richards as a way of quantifying how much surface is buried when proteins associate. It defines the accessible surface area (ASA) over which a protein and solvent make contact, as marked out by computationally rolling a spherical solvent molecule over the van der Waals surface of the protein (Figure 1.8A). The probe radius is usually set to 1.4 Å for water. Later, the molecular surface area (MSA) of a protein was defined as the area covered by an inward-facing probe rolled over the van der Waals surface (Figure 1.8B). ASA areas tend to be somewhat larger (Table 1.2). The buried surface area (BSA) when macromolecules X and Y form a complex is given by

$$\text{BSA} = \text{ASA}(X) + \text{ASA}(Y) - \text{ASA}(XY)$$

A similar equation applies when MSAs are used instead. The BSA is roughly twice the area of an interaction patch at the X/Y interface because approximately the same amount of exposed surface is lost from both X and Y (it is exactly the same amount if XY is a symmetric

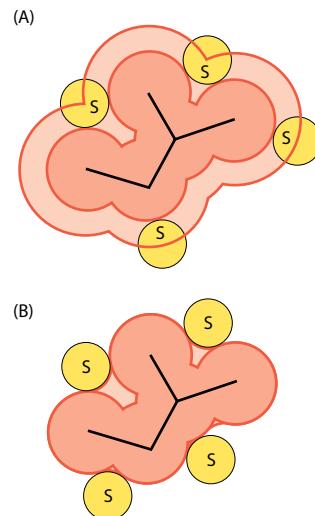


Figure 1.8 Accessible surface area (ASA) and molecular surface area (MSA). In both cases, a solvent probe, S, is rolled computationally over the van der Waals surface of the molecule. In (A), its center defines the ASA. In (B), its point of contact with the van der Waals surface defines the MSA. (Adapted from I. Tunon, E. Silla and J.L. Pascual-Ahuir, *Protein Eng.* 5:715–716, 1992. With permission from Oxford University Press.)

Table 1.2 Solvent-accessible surface areas (ASAs) and molecular surface areas (MSAs) for the Cdk2/cyclin A heterodimer.

	Cdk2 (X)	Cyclin A (Y)	Cdk2/cyclin A (XY)	BSA (X + Y - XY)
ASA (\AA^2)	14,396	12,300	23,346	3350
MSA (\AA^2)	11,733	10,308	19,158	2884

ASAs were calculated with Areaimol from the CCP4 suite of programs. MSAs were calculated with Aesop. (Courtesy of Martin Noble.)

dimer). In another convention used by some investigators (but not here), the areas defined above are halved and the redefined BSAs are similar in size to the interaction patch.

Atoms in the core regions of proteins are inaccessible to solvent. Crystal structures reveal that they are packed as tightly as atoms in an organic solid, such as a small-molecule crystal. Amino acid residues in the core tend to pack in a precise jigsaw fit, although some proteins contain cavities large enough to accommodate one or more water molecules (see above). Surface residues are less tightly packed and some proteins have surface-exposed regions that are flexible or even completely disordered.

The surface area buried between protein subunits can be used as a guideline to discriminate between crystal packing and a physiologically meaningful interaction. Crystal contacts are usually smaller and less complementary. An analysis of crystalline proteins known to form monomers or dimers in solution gave a success rate of ~85% when 1712 \AA^2 was taken as a cut-off value for ΔASA . The Protein Interfaces Surfaces and Assemblies (PISA) program predicts quaternary structures on the basis of calculated interactions and thermodynamic parameters and has an estimated success rate of ~90%. Paired interaction patches also have complementary topographies. A convenient measure of **complementarity**, S_c , averages the following quantity over all points on both surfaces:

$$S_c = \exp(-wd^2) \cos\alpha$$

where d is the distance between x_A , a point on surface A, and x_B the nearest point on surface B; α is the angle between the unit vector normal to surface A at x_A and the unit vector normal to surface B at x_B , and w is a free parameter (Figure 1.9). The higher the value of S_c , the better the fit. A perfect fit has the maximum value $S_c = 1$. Values of 0.65–0.75 are typical for genuine interactions. Favorable shape complementarity is helpful when assessing interfaces, real and potential, but although it is conducive to the formation of hydrophobic and van der Waals interactions, it does not necessarily point to a physical interaction.

As noted above, complementary interaction patches are quite large. Although they include some hydrophobic sites, they also involve charge pairing and H-bonds. An interaction patch of 1000 \AA^2 or so cannot be entirely hydrophobic, because the molecule bearing it would be liable to nonspecific aggregation and therefore insoluble. Two examples of complementary pairs of interaction patches are shown in Figure 1.10. However, interactions within an assembly are more complex in that most subunits have binding interfaces with more than one other subunit: for example, subunits in viral capsids make contact with between four and seven neighbors (see Chapter 8).

A certain minimum strength of interaction is required for specificity

Interactions are gauged in terms of their **affinity** or its inverse, the **dissociation constant** (K_d), defined by

$$K_d = [A][B]/[AB]$$

where $[A]$ is the molar concentration of component A, and so on. Affinities are related to free energies of association by

$$K_d = \exp(-\Delta G/RT)$$

or, rearranging,

$$\Delta G = -RT \ln K_d$$

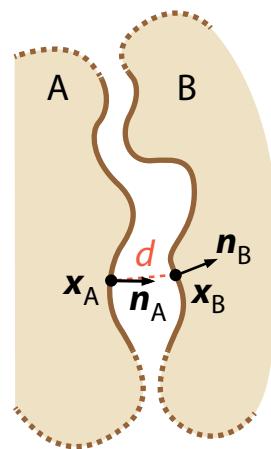
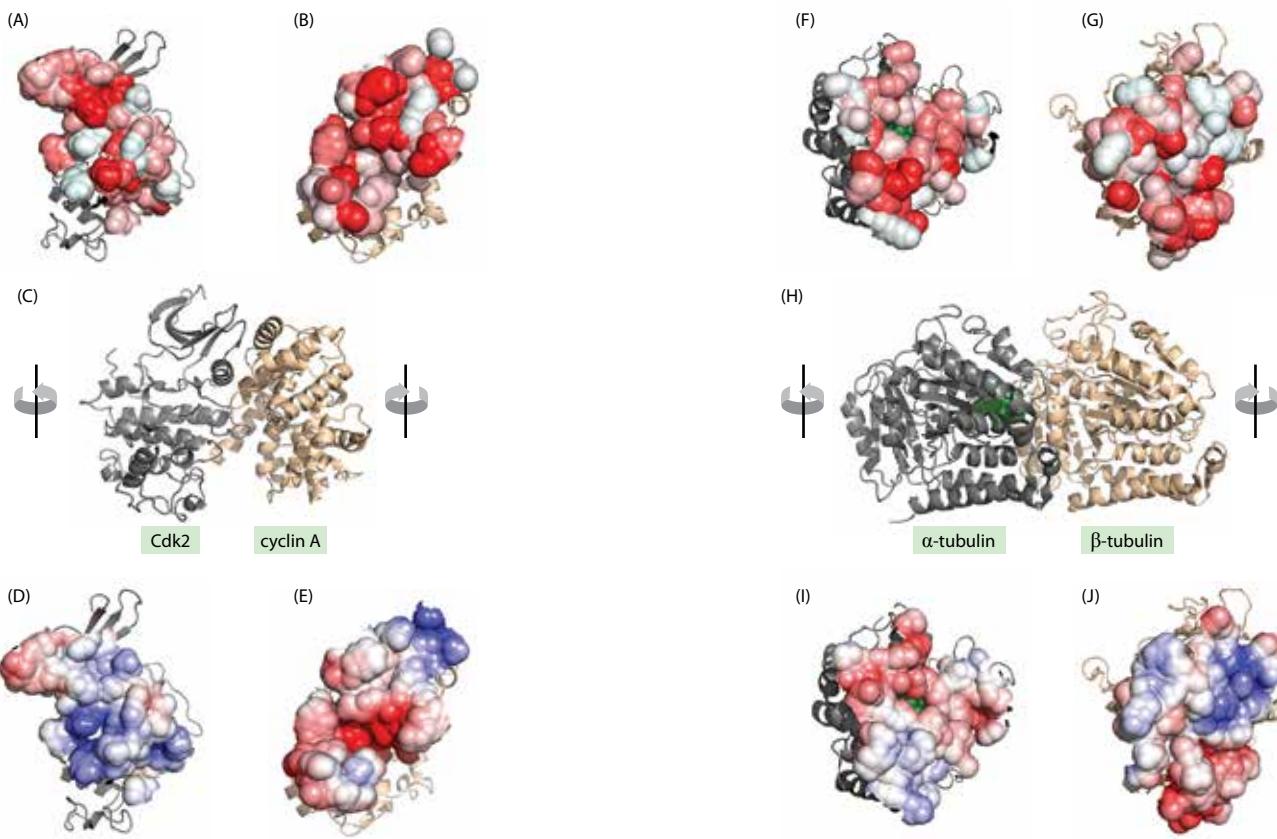


Figure 1.9 Shape complementarity of paired molecular surfaces. The measure S_c is defined in the text. Bold symbols denote vectors. The solid lines mark the regions on molecules A and B that form the interface under consideration. For point x_A , x_B is the nearest point on molecule B. The scalar product $n_A \cdot n_B$ is the cosine of the angle α between the two unit vectors n_A and n_B , and cannot exceed 1. Similarly, the exponential decay according to the distance d between x_A and x_B cannot exceed 1 and only achieves this value if the two surfaces make direct contact ($d = 0$). (Adapted from M.C. Lawrence, CCP4 Newsletter 39.)



R is the gas constant. If solutions of [A] and [B], both at 2 μM (a typical concentration for *in vitro* experiments with 100 kDa proteins), are mixed, and they have a K_d of 1 μM , then at equilibrium there are equal concentrations of A and B, that is 1 μM , free and as AB. This does not state a strong preference for complex formation. If, in contrast, the K_d is 1 nM, then about 98% of both A and B are as AB; in other words, association is strongly favored. Given the large areas and chemical complexity of their surfaces, there are an enormous number of ways in which two macromolecules can come together, and many of them represent attractive interactions. Whether or not an association is meaningful depends on the K_d . A micromolar K_d is considered borderline, but a nanomolar K_d is taken to indicate a stable complex. In antigen–antibody binding, for example, the K_d is typically in the nanomolar to picomolar range (Section 17.4).

Cooperativity enhances stability in multi-subunit complexes

In assemblies, most subunits interact with more than one neighbor. The net ΔASA for one subunit, considered as an interaction between it and the rest of the complex, can be taken as the sum of all the ΔASAs involved. For this reason, complexes tend to be more stable than the individually weak interactions would suggest. The combined effect of several interactions may be described in terms of their **avidity**. For simplicity, let us consider a two-component system (C, D) in which there is more than one interface between C and D, each with its own $K_d^{(i)}$, $i = 1, 2, \dots$. In principle, their combined effect is given by the product of these dissociation constants:

$$K_d^{\text{tot}} = \prod K_d^{(i)}$$

However, it can also be considered in terms of a single ‘effective’ interaction whose dissociation constant, K_d^{eff} , represents the avidity of the reaction. Engagement at one interface may affect the readiness of binding to take place at another and this gives rise to the concept of **cooperativity**, described by parameter α :

$$\alpha = K_d^{\text{tot}}/K_d^{\text{eff}}$$

If $\alpha = 1$, the interaction is non-cooperative; if $\alpha > 1$, there is positive cooperativity (synergy); and if $\alpha < 1$, there is negative cooperativity (interference). Cooperativity is also important

Figure 1.10 Interaction patches at inter-subunit interfaces of two heterodimeric proteins. (A–E) Cdk2/cyclin A. (PDB 1FIN) A side-view ribbon diagram is shown in (C). (A, B) and (D, E) show the paired interaction patches as surface renderings color-coded to convey different properties. Their areas are given in Table 1.2. To present the interaction patches *en face*, the subunits were rotated 90° as shown. Other visible parts are shown as ribbons. (A, B) are coded for hydrophobicity: stronger red means more hydrophobic. Complementarity is expressed in terms of the extent to which the red regions observe mirror symmetry across the vertical axis. (D, E) are coded for electrostatic potential: blue, positive; red, negative; white, uncharged. Complementarity is expressed in terms of red regions being reflected in blue regions across the vertical axis. Greater, although not total, complementarity is observed in terms of hydrophobicity. As the buried surface areas are much larger than the minimum required for a meaningful interaction (see main text), such departures from complementarity can be accommodated. (F–J) α – β Tubulin. (PDB 1JFF) The presentation is as in (A–E). The buried surface area at the interface is 3780 \AA^2 , and about half that for each interaction patch. The bound nucleotide (GTP) is green. The hydrophobicity complementarity is greater than in (A–E) but still not perfect.

in a different context in which an oligomeric receptor molecule has multiple binding sites for a ligand and binding of the first ligand can affect the affinity for a second ligand and so on, as discussed in Section 1.7.

1.3 PROTEIN FOLDING AND STABILITY

Most proteins must adopt a particular three-dimensional structure that allows them to interact specifically with other molecules and to fulfill their biological functions. Moreover, only correctly folded proteins have long-term stability in cellular environments. Misfolded proteins are normally degraded by intracellular proteases, but those that escape this quality-control mechanism can form insoluble aggregates, for example the fibrillar assemblies (*amyloids*) associated with neurodegenerative diseases (Section 6.5).

How a given protein reaches its folded state after synthesis on the ribosome is an important issue often referred to as the ‘protein folding problem.’ Christian Anfinsen’s classic experiment (Figure 1.11) showed that the linear sequence of amino acids in a polypeptide chain contains all the information needed to specify its three-dimensional structure. Moreover, despite the immense conformational space that a polypeptide chain can sample, it must reach its native fold on a biologically relevant timescale. The magnitude of this problem is illustrated by the paradox first stated by Cyrus Levinthal: if, with a protein of 100 aa, each residue can assume only three different configurations, the number of possible conformations would be $\sim 10^{49}$. Even if the time required for switching from one conformation to another is very small, say 10^{-11} seconds, then a random search of conformational space would take 10^{29} years (the estimated age of the universe is only $\sim 14 \times 10^9$ years)! This clearly rules out any random-search process. In reality, proteins fold on a timescale of seconds or less.

Protein folding follows pathways populated with intermediates

A common feature of most hypotheses aimed at resolving the Levinthal paradox is some sort of folding pathway, with the formation of structured elements occurring early in the process. These elements form in parallel, not sequentially, and subsequent sampling of conformational space involves fewer, larger, components; both considerations accelerate the folding process. The ‘hydrophobic collapse’ model assumes that folding starts with a rapid condensation of the polypeptide chain into a dynamic, flexible form with the beginnings of

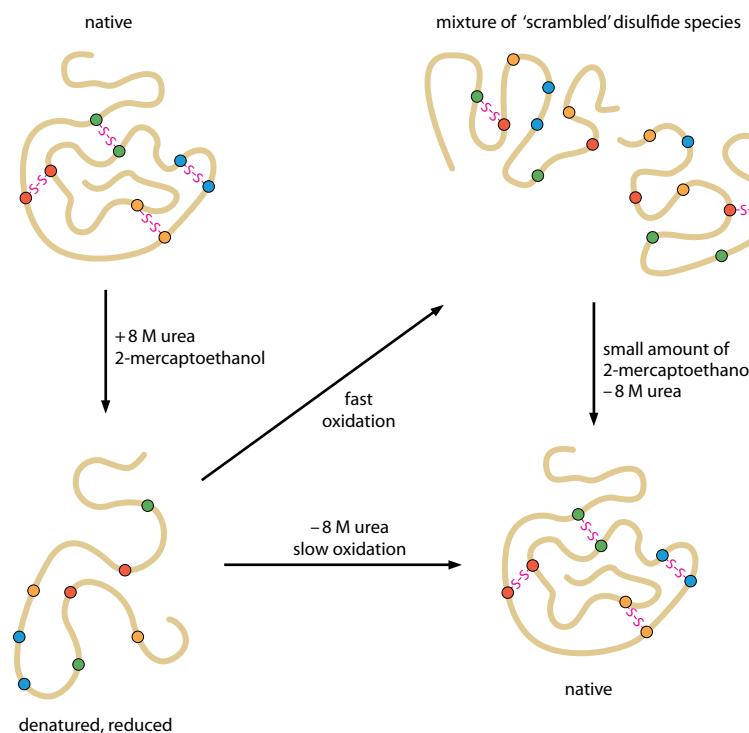


Figure 1.11 The Anfinsen experiment demonstrating protein unfolding and spontaneous refolding. Ribonuclease, an enzyme of 124 aa with four disulfide bridges, is exposed to 8 M urea to denature the protein, and 2-mercaptoethanol to cleave the disulfide bridges. This renders the enzyme inactive. When the cysteines are allowed to reoxidize rapidly in the absence of 2-mercaptoethanol under denaturing conditions, the protein adopts a mixture of conformations with randomly formed disulfide bridges and remains inactive. However, if the urea is removed slowly, disulfide exchange occurs and the enzyme refolds into a conformationally homogeneous fully active species. The four disulfide bridges are correctly formed out of the 105 ($7 \times 5 \times 3 \times 1$) possible pairings of cysteines.

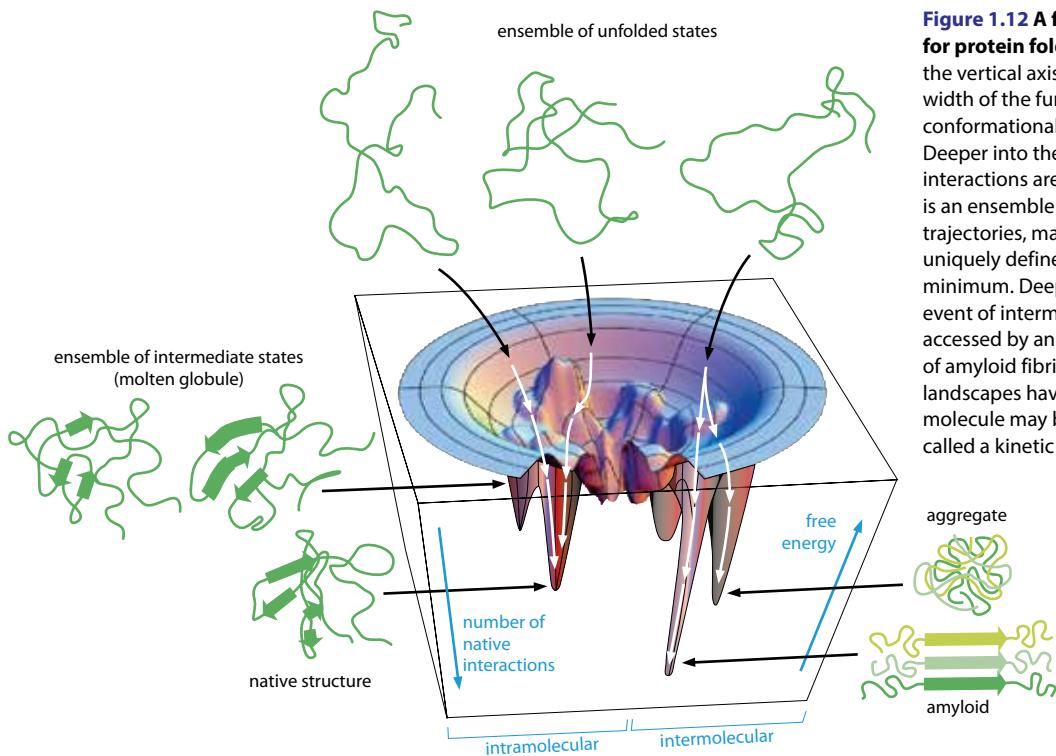


Figure 1.12 A funnel-shaped energy landscape for protein folding. In this schematic diagram, the vertical axis represents free energy, and the width of the funnel at any level represents the conformational freedom of the polypeptide chain. Deeper into the funnel, more and more native interactions are formed. At the top of the funnel is an ensemble of starting structures; their folding trajectories, marked with arrows, converge to the uniquely defined native structure at the global minimum. Deeper minima are possible in the event of intermolecular interactions. The deepest, accessed by an entirely different pathway, is that of amyloid fibrils (Section 6.5). In general, the landscapes have some ruggedness, and the folding molecule may become trapped in a local minimum, called a kinetic trap.

a hydrophobic core, a so-called **molten globule**. Whether hydrophobic collapse or secondary structure formation occurs first is still an open question.

There is now ample evidence that folding proceeds in stages and that there is no single, directed pathway. The process is determined by the protein's **free energy landscape**, which can be represented graphically as a funnel (Figure 1.12). In this view of folding, the non-native protein exists as a large ensemble of different structures. The molecules move spontaneously down free energy gradients, passing through multiple partly folded states until they converge on a unique global minimum. On the way, more and more native interactions are established. In an alternative folding pathway, fibrillar aggregates called **amyloid** may be formed. Most amyloids that have been studied are the result of misfolding to a state that is a lower free energy minimum than the native fold. However, there are some proteins for which amyloid is their native state (see Section 6.5).

For proteins that fold via a two-state transition (denatured/native), the free energy funnel has no significant valleys or barriers (that is, no local minima). Such a scenario is rare and largely confined to small proteins (see below). More commonly, the energy landscape is rugged and the folding protein must navigate through local minima. This ruggedness reflects the temporary formation of folding intermediates that depend on multiple weak interactions not necessarily found in the end state (the native fold). These local minima represent **kinetic traps** that slow the folding process. If a local minimum is deep, the corresponding intermediate is longer-lived and the likelihood of off-pathway events such as aggregation increases. A full characterization of all the non-native species (unfolded, transitional states, and partly folded intermediates) poses a formidable challenge that has yet to be solved.

Larger proteins (longer than ~200 aa) do not usually behave as a single folding unit; many consist of two or more regions whose folds are recognized as they appear in several, even many, other proteins. These regions, called **domains**, vary in size between ~50 and ~250 aa. Some are capable of fully independent folding if separately expressed, or they may be excised proteolytically in folded form from the parent protein: such domains make few, if any, contacts with the other part(s) of the protein. Other domains make many stabilizing interactions with other parts in their parent protein and are essential contributors to its hydrophobic core. They would leave unacceptably large amounts of hydrophobic surface exposed in the absence of the rest of the protein: in this sense, they fold semi-autonomously. Many domains are responsible for particular functions, such as the Rossmann fold that

binds NAD⁺/NADP in oxidoreductases or the SH2 domain that binds phosphorylated Ser residues in other proteins (see **Guide**). Domains may tolerate substantial changes to their amino acid sequences, as indicated by the many instances of domains with substantially different sequences sharing the same fold. However, single amino acid substitutions in crucial locations can affect folding, some observed as temperature- or cold-sensitive mutants (see below). A limited number (probably somewhat above 1000) of domain folds is thought to exist (see **Guide**). Domains are thus fundamental building blocks of protein structure, recurring in different combinations in the vast array of known proteins.

Folded proteins in solution are not rigid; rather, they may be considered as undergoing rapid fluctuations around their native structure. The extent of these fluctuations varies with temperature and from case to case, but surface regions, especially loops and projections, are in general less well ordered than core regions. Multidomain proteins sometimes have disordered segments that act as hinges or extended flexible linkers (see, for example, Section 9.4), and some proteins have sizable surface-exposed regions (in extreme cases, the entire molecule) that remain unfolded in the absence of a suitable interaction partner. These regions, which said to be *natively unfolded*, are covered in greater detail in Sections 1.5 and 6.4.

Protein structures are only marginally stable

It has long been known that the native structures of folded proteins can be lost (denatured) by heating or exposure to extremes of pH or denaturing agents such as urea or guanidinium chloride, or detergents such as sodium dodecylsulfate (SDS). **Denaturation** can be understood as a disruption of the multiple weak forces that combine to stabilize the native folds. Denatured proteins are highly protonated or deprotonated (at extremes of pH) or coated with denaturants, and thus tend to have extended shapes. Indeed, the extended form of SDS-denatured proteins saturated with bound negatively charged dodecylsulfate ions is the basis for their separation by size in SDS polyacrylamide-gel electrophoresis (SDS-PAGE). Unfolding can be partial and, in some cases, reversible on removal of the denaturing agent, but it generally involves a cooperative loss of structure as the weak forces are collectively disrupted. The denatured protein is best regarded as an ensemble of unfolded structures, for which the term **random coil** is sometimes used. However, it may also refer to specific conformations that lack regular secondary structure elements: that is, α helices and β strands.

Denaturation exposes hydrophobic groups that are normally confined within the protein core, and these will tend to interact, causing nonspecific aggregation. Binding of chemical denaturants such as urea keeps the denatured protein in solution, but thermal denaturation or exposure to low pH or organic solvents leaves the exposed hydrophobic groups free to interact and this can lead to precipitation.

The **free energy of stabilization** (ΔG) of folded proteins is small, usually in the range of -21 to -42 kJ/mol at 25°C. This is distributed over many weak interactions but is energetically equivalent to just two or three H-bonds. It has been argued that this property of marginal stability is important in facilitating functionally relevant conformational changes and unfolding prior to degradation and recycling. Others have argued that, provided that a protein achieves a certain threshold of stability, further increases are evolutionarily neutral and are not selected for.

Protein stability correlates with size and other factors such as covalent cross-links

Data from studies of the reversible folding of 63 globular proteins have shown that thermal denaturation depends more on the length of the polypeptide chain than on its amino acid composition or its secondary or tertiary structure. The principal contributions to stability come from interactions in the hydrophobic core, whose volume increases with the cube of the radius (r^3), whereas the surface area increases only with the square (r^2). On this basis, it is possible to assess the entire proteomes of organisms whose genomes have been sequenced and arrive at an estimate of the average stability. For *E. coli* proteins at 37°C, the average ΔG of folding comes to ~30 kJ/mol; for *S. cerevisiae* it is ~37 kJ/mol, and for *C. elegans* it is ~34 kJ/mol.

It is widely accepted that a minimum length of ~50 aa is required to generate a sufficient hydrophobic core to sustain a folded structure, although there are a few exceptions. The

peripheral subunit-binding domain (PSBD, 33 aa in its structured region) of 2-oxo acid dehydrogenase complexes is one of the smallest proteins to have a stable fold (see Figure 9.4B). Its two short parallel α helices, separated by a short 3_{10} helix and a loop, enclose a close-packed hydrophobic core; folding takes place on an ultrafast (microseconds) time-scale with a ΔG of ~ -8 to -17 kJ/mol at 25°C in an apparent two-state transition. Its single robust folding pathway suggests the presence of a single nucleation-competent motif (or *foldon*).

Another factor that affects stability is covalent cross-linking (see Section 1.2), notably disulfide bridges, which are largely confined to extracellular (secreted) proteins such as the protease chymotrypsin. The oxygen-rich extracellular environment is much harsher than the reducing environment found inside cells. A disulfide bridge restricts the conformational space open to an otherwise unfolded polypeptide chain and thus stabilizes folds by limiting the increase in entropy from disorder. Experiments indicate that a disulfide bridge can contribute up to ~ 350 kJ/mol to the free energy of stabilization. For this reason, disulfide-bridged proteases such as subtilisin have found wide application in washing powders, for example, and the biotechnology industry seeks to render many proteins less prone to denaturation by engineering new disulfide bridges into them.

Many cellular proteins denature collectively under thermal stress

Based on the spread of protein sizes in its proteome, the range of protein stabilities in any one organism is predicted to be broad (Figure 1.13). About 650 (15%) of the 4300 proteins of *E. coli* have predicted ΔG values of denaturation of <17 kJ/mol, so that they are only marginally stable at the ambient temperature of 37°C . Temperatures that are substantially higher should lead to catastrophic levels of denaturation, with 1% of *E. coli* proteins predicted to denature at 47°C , rising to 50% at 54°C . These temperatures coincide with the range beyond which *E. coli* cells cannot survive, suggesting that collective denaturation of the less stable proteins is responsible for halting cell function. Other calculations suggest that similar behavior is widespread.

In the account given above, no attention was paid to the fact that many proteins have multiple domains. This should result in lower stabilities as estimated on the basis of domain length rather than total polypeptide chain length. Nor does it recognize that many proteins are stabilized by incorporation into multi-subunit assemblies. These qualifications notwithstanding, changes of just a few degrees can shift thermal equilibria substantially: for example, it is estimated that raising the temperature from 37°C to 41°C destabilizes the average protein by almost 20%. To combat thermal perturbation, cells mobilize their *heat-shock response* systems, up-regulating chaperone proteins that promote the refolding of thermally perturbed proteins (Section 6.3).

In temperature-sensitive mutants, minor changes to the primary structure of a protein that folds correctly and is biologically active at one temperature—say, 30°C —render it inactive and presumably denatured at a slightly higher temperature—say, 37°C . Such mutations have proved invaluable in unraveling many biochemical and assembly pathways. Conversely, cold-sensitive mutants function at the higher but not at the lower temperature, presumably because the mutated protein remains stuck in a kinetic trap until released by additional thermal (enthalpic) energy.

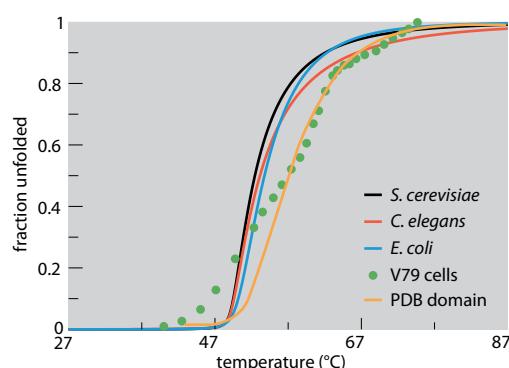


Figure 1.13 Fractions of proteins in the proteomes of *E. coli*, *S. cerevisiae*, and *C. elegans* unfolded as a function of temperature. The blue, black, and red curves are based on the predicted distributions of domain lengths in the respective proteomes of a bacterium, a fungus, and a nematode worm. The fraction of denatured proteins for mammalian V79 cells (green points) was measured by differential scanning calorimetry. The yellow curve is the prediction based on domain-length distribution obtained from structures in the Protein Data Bank. (Adapted from K. Ghosh and K. Dill, *Biophys. J.* 99:3996–4002, 2010. With permission from Elsevier.)

Proteins from thermophilic organisms are not very different from mesophilic homologs

Unlike organisms that live at temperatures between 0°C and 40°C (mesophiles), thermophiles flourish only at higher temperatures (up to 80°C) and hyperthermophiles thrive in even hotter environments (80–120°C), such as hot springs or deep sea vents. Many thermophiles and hyperthermophiles are archaea (Section 1.9), but some bacteria have adapted to life at high temperatures. Nevertheless, regardless of preferred growth temperature, these organisms all have similar proteomes, but proteins from thermophiles are substantially more heat-resistant than their mesophile counterparts. Indeed, the ability of the *Taq* polymerase from the bacterium *Thermus aquaticus* to function at temperatures above those that denature mesophile polymerases and cause mesophile DNAs to melt (strand separation) is the basis of the polymerase chain reaction (PCR).

The structures of proteins from thermophiles, esteemed by crystallographers because of their amenability to crystallization, are generally similar to homologs from mesophiles. The differences reside mostly in their having tighter surface loops and more stabilizing electrostatic interactions (notably, salt bridges) at the protein surface. These changes do not have to increase the free energy of folding very much to be effective. If a mesophile protein at, say 100°C (373 K), denatures in 2 seconds and the thermophile homolog survives 1000 times as long (~30 min), that is potentially long enough for a typical bacterial cell division. The change in free energy of folding to achieve this extra stability, $\Delta\Delta G \approx RT\ln 1000$, or $2 \times 373 \times 2.3 \times 3 = 21$ kJ/mol, is roughly equal to a couple of H-bonds. So it is perhaps not surprising that it is hard to recognize the subtle underlying changes in weak forces distributed over the entire protein molecule.

1.4 SELF-ASSEMBLY AND SYMMETRY

Just as the primary structures of protein subunits specify their secondary structures (α helices and β sheets) and tertiary structures (three-dimensional folds), they also underlie the quaternary structures (oligomeric and polymeric) of the machines and assemblies that permeate biology. In many cases, just one kind of protein is involved (**homomeric** assembly), but in others several different polypeptide chains co-assemble (**heteromeric** assembly), and other components (such as nucleic acids or lipids) may also be incorporated. The exquisite selectivity exercised in the choice of components to be incorporated into any given complex is a striking feature of all assembly processes. Nonetheless, it is possible to codify the various types of quaternary structure that are encountered and to see how some might have been evolutionarily favored.

Many assembly processes can be recreated from purified components—for example, in the reversible denaturation of oligomeric enzymes, whereby refolding of the subunits and reassembly of the oligomers take place upon removal of the denaturant. Another classic example is the reassembly of tobacco mosaic virus from its RNA and coat protein (see Section 8.3). We call this **self-assembly**, and the fact that it is spontaneous indicates that, like protein folding, it is driven by an overall decrease in Gibbs free energy. In other systems, assembly may require other proteins (assembly factors or scaffolding proteins) that are not retained in the final structure. Alternatively, an input of energy (deriving ultimately from the hydrolysis of ATP) may be needed.

Most proteins form symmetrical oligomers with two or more subunits

Bioinformatics, structural and functional genomics, and protein structure databases all concur that the majority of proteins exist as oligomers. Of the complexes found in *E. coli* cells, it is estimated that about half are homomeric. Dimers and tetramers are common, but there are also many larger assemblies. Assemblies with odd numbers of subunits are much rarer than those with even numbers. Inter-subunit interfaces in oligomers are generally more highly conserved across species than is the rest of the protein surface. Given that these interfaces are largely hydrophobic, it is tempting to imagine that dimers arose by mutations that formed a hydrophobic interaction patch on the surface of a monomer. If the interaction were face-to-face (an **isologous** interaction), a symmetrical dimer would result (Figure 1.14). A second face-to-face interaction generated in the same way would give a

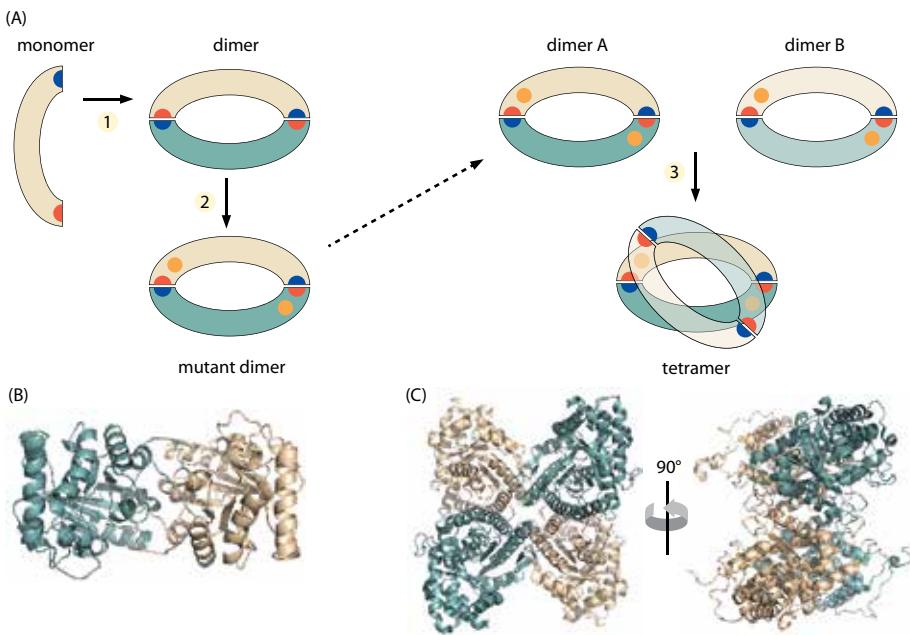


Figure 1.14 Formation of symmetrical homodimers and homotetramers.

(A) In step 1, a monomer (in wheat) with complementary binding patches (red and blue) interacts with an identical monomer (in teal), creating a dimer with a 2-fold axis perpendicular to the plane of the page. In step 2, mutation creates a second complementary binding patch (orange) on each monomer and, in step 3, two such dimers interact across a 2-fold axis, giving a tetramer. (B) A symmetrical homodimer of chicken muscle triose phosphate isomerase. (PDB 8TIM) (C) Two views of the symmetrical homotetramer of rabbit muscle fructose-1,6-bisphosphate aldolase. (PDB 3B8D) In (B) and (C), the subunits are colored as in (A).

tetramer, but the emergence of a third interaction patch on each subunit in the tetramer could not generate a closed octamer with all potential binding surfaces satisfied internally. Interaction surfaces would be left open for head-to-tail interactions of one tetramer with another, with no limit on the number. This would probably be selected against, which may explain the prevalence of dimers and tetramers with 2-fold axes.

Symmetrical interactions of this kind create closed dimers and tetramers and these have several potential advantages. In a dimeric enzyme, two active sites can be generated across the 2-fold axis, with contributions from both subunits to each one. In addition, in a symmetrical homo-oligomer the effect of any mutation is multiplied and this can enhance the sensitivity to evolutionary selection, as in the properties of *cooperativity* and *allostery*, which are almost always associated with multiple binding sites for substrates and/or ligands (Section 1.7).

Symmetry defines a set of larger structures composed of multiple copies of identical subunits

Over and above dimers and tetramers, there are many much larger structures among the assemblies and machines that this book addresses. Some are homomeric; others are heteromeric. If only one type of subunit is used, the simplest arrangements have the subunits packed such that the bonding pattern is the same for each; in other words, they are *equivalent*. (Departures from equivalence are discussed in Section 8.2.) The final structure is then necessarily a symmetrical one. Since proteins are chiral—that is, different from their mirror images—reflections are forbidden and the only symmetry operations permitted are rotations and translations. Thus all structures must belong to one of the line, point, plane, or space groups. Plane groups are relevant to biological membranes and space groups to protein crystals, but here we confine the discussion to finite structures defined by line and point groups.

Line and cyclic point group symmetries generate helices and rings

If protein subunits bind head-to-tail in identical fashion, every subunit will have exactly the same environment except those at the two ends, an arrangement known as **line symmetry** (Figure 1.15A). In general, this gives rise to a helical structure, which may be further stabilized by axial interactions between subunits in successive turns. (A simple linear stack is rare in biology.) Some protein helices are essentially straight, reflecting rigidity, whereas others are more flexible. There is no restriction on the number of subunits per turn, nor does it have to be integral.

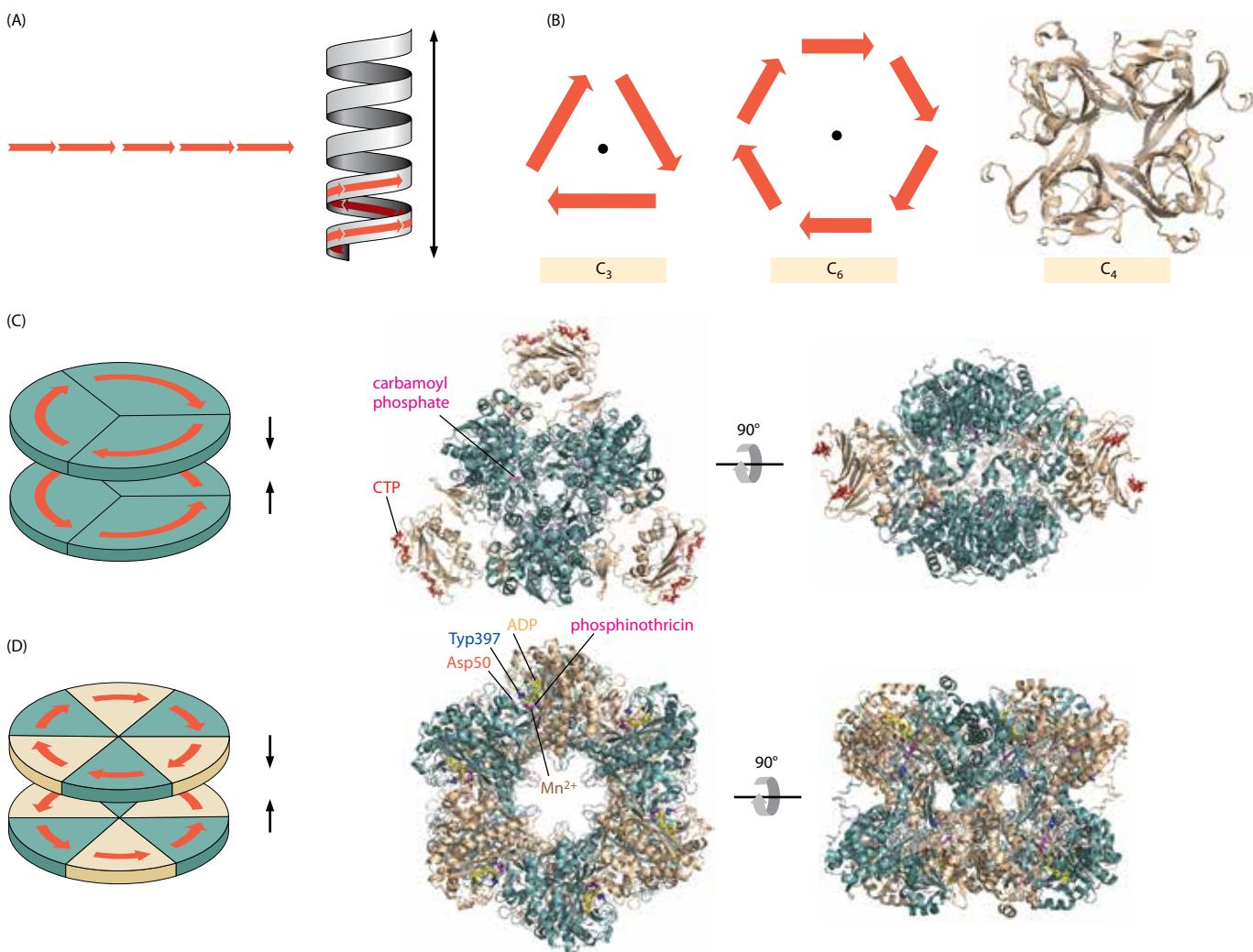


Figure 1.15 Line symmetry and cyclic symmetry. (A) In line symmetry, a succession of asymmetric subunits (red arrows) bind head-to-tail. If they wind into a helical coil, line symmetry is still observed. (B) In cyclic symmetry, the subunits form closed rings. At the right is the crystal structure of *E. coli* aspartate 1-decarboxylase, a C_4 tetramer. (PDB 1PQF) (C) In the dihedral symmetry of *E. coli* aspartate transcarbamoylase, two C_3 rings of catalytic subunits (teal) are arranged face-to-face (black arrows). In the ribbon diagram, they are shown associating with three homodimers of regulatory subunits (wheat), each bound on a 2-fold axis. The enzyme is thus an $\alpha_6\beta_6$ hetero-dodecamer. Carbamoyl phosphate in the active sites of the catalytic subunits is in magenta; CTP, an allosteric effector bound to the regulatory dimers, is in red. (PDB 1ZA2) (D) *Salmonella typhimurium* glutamine synthetase has two face-to-face C_6 rings whose subunits are colored alternately in teal and wheat. The active sites are located between subunits in the C_6 rings. Asp50 in the active site is in red, ADP is in yellow, Mn^{2+} ions are brown, and phosphinothricin, an inhibitor, is in magenta. Tyr397, the site of regulatory adenylylation, is in blue. (PDB 1FPY)

The lengths of helical polymers are generally not fixed but depend on the respective propensities for nucleation and elongation and their resistance to fragmentation by hydrodynamic shear. However, defined lengths result if assembly is directed by a **molecular tape measure**. For example, in helical viruses, the genomic nucleic acid (of defined length) interacts with coat protein subunits to initiate assembly, and subsequently to specify the length of the virion. In other systems, a long fibrous protein interacts with the protein subunits to define the length of a helical filament, as with the tail-tubes of tailed viruses or the *sarcomeres* in muscle fibers (Section 14.4). Many systems employ specialized initiation and termination complexes made of other proteins. For example, microtubules are helical polymers of the cytoskeletal protein tubulin, which can self-assemble *in vitro* under appropriate conditions. *In vivo*, however, their growth is nucleated on heterologous complexes called microtubule organizing centers (MTOCs), whose positioning helps to specify the overall cytoarchitecture (Section 14.6).

In another type of head-to-tail arrangement, symmetrical closed structures are generated in the form of rings. **Cyclic symmetry**, a form of point group symmetry, places no restriction on the number of subunits in a ring, though it is seldom more than six and it is the only form of symmetry that can accommodate odd as well as even numbers of subunits. A typical example is the C_4 tetramer, *E. coli* aspartate 1-decarboxylase (Figure 1.15B). However, cyclic symmetry is often extended to **dihedral symmetry**, in which a 2-fold axis is added at right-angles to the n -fold axis, giving $2n$ subunits in the assembly. Two examples of cyclic symmetry combined with dihedral symmetry are shown in (Figure 1.15C,D). Both are 12-mers, one $\alpha_6\beta_6$ (aspartate transcarbamoylase) and the other α_{12} (glutamine synthetase). Note how the active sites are located between subunits in the C_6 rings of glutamine synthetase.

Cubic symmetry is employed in a variety of oligomeric proteins

The other point group is **cubic symmetry**. This is based on the Platonic solids (Figure 1.16A). On each of the four faces of a tetrahedron, C₃ trimers can be arranged and each face is related to its neighbors by 2-fold axes. Hence, the other name for this arrangement is (23) symmetry; it has a total of 12 subunits. An example of this relatively rare symmetry is dehydroquinase, an enzyme involved in aromatic biosynthesis in bacteria and plants (Figure 1.16B). Similarly, C₃ trimers can be arranged on the eight faces of an octahedron. Viewed from the top, there is a 4-fold axis of rotation; viewed on a face, there is a 3-fold axis of rotation; and there is a 2-fold axis on each edge. This conforms to octahedral or (432) symmetry, with a total of 24 subunits. Arrangement on a cube is formally equivalent, with 4-fold axes on the faces, 3-fold axes at the corners, and 2-fold axes on the edges. Examples of octahedral symmetry are ferritin (the iron storage protein) and the core assembly of the pyruvate dehydrogenase (PDH) complex depicted in Figure 1.16C.

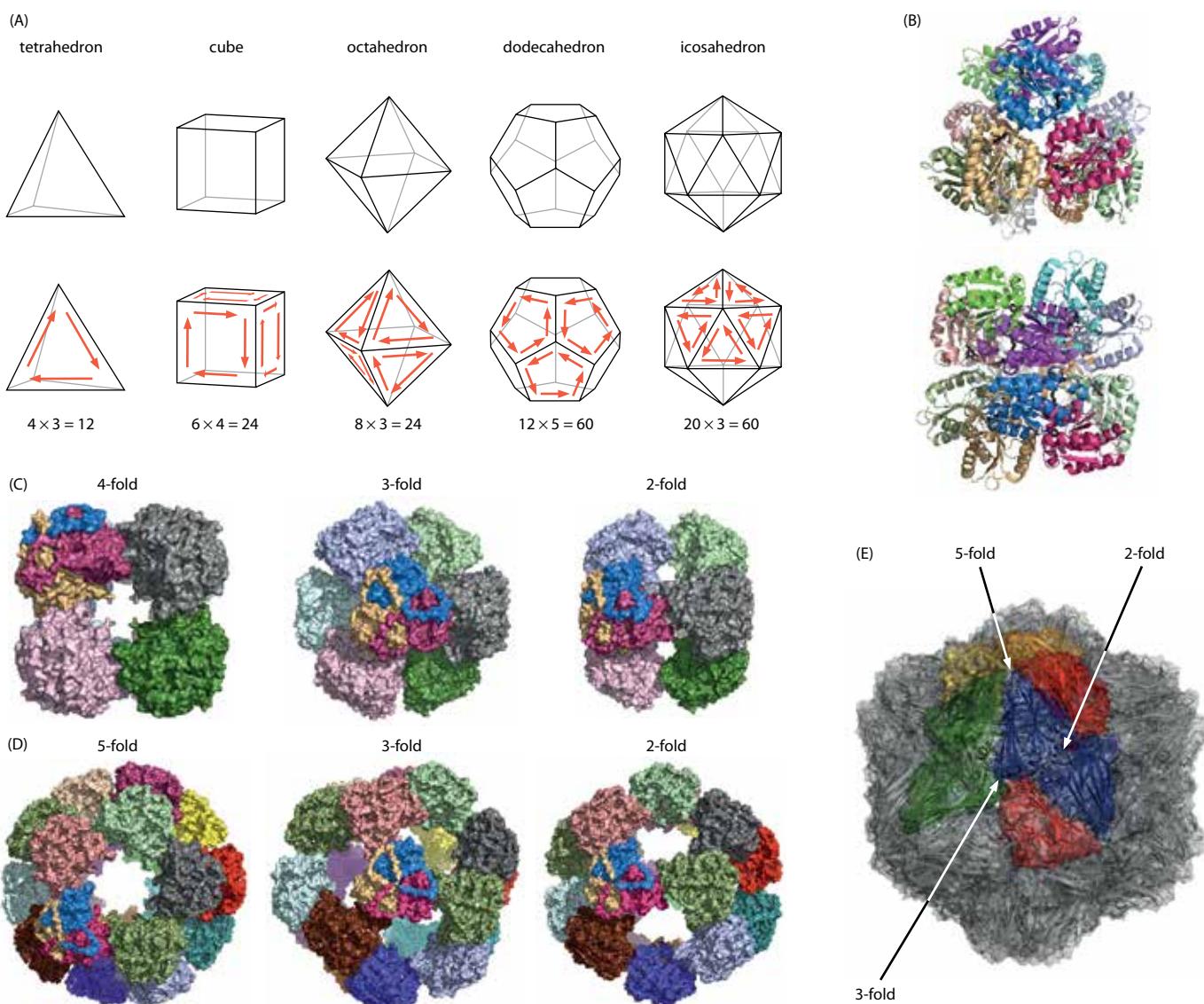


Figure 1.16 Cubic symmetry. (A) The Platonic solids with protein subunits placed as represented by red arrows. (B) The 12 subunits of *Streptomyces coelicolor* dehydroquinase arranged with (23) symmetry. Bottom, view along a 2-fold axis; top, view along a 3-fold axis. Each subunit is colored differently. (PDB 1GTZ) (C) Core assembly of acetyltransferase domains of the E2 subunits of a PDH complex with (432) symmetry, arranged as eight trimers at the vertices of a cube. One trimer has the three domains in blue, gold, and magenta, and the other trimers

are in different colors. (PDB 1EAA) (D) Core assembly of acetyltransferase domains of the E2 subunits of a PDH complex with (532) symmetry, arranged as 20 trimers at the vertices of a dodecahedron. One trimer has the three domains in blue, gold, and magenta. The other trimers are in different colors. (PDB 1B5S) (E) The capsid of satellite tobacco necrosis virus has 60 subunits distributed over an icosahedral lattice. One 5-fold axis (labeled) is surrounded by red, orange, yellow, green, and blue subunits; and similarly at the 3-fold and 2-fold axes. (PDB 2BUK)

In icosahedral or (532) symmetry, C_3 trimers can be placed on the 20 faces of an icosahedron, giving a total of 60 subunits with 12 5-fold axes at the vertices, 20 3-fold axes on the faces, and 30 2-fold axes on the edges. A formally equivalent arrangement is 12 5-fold axes on the faces of a dodecahedron, 20 3-fold axes at the vertices, and 30 2-fold axes on the edges. Examples are the core assembly of another PDH complex (a dodecahedron with 20 trimers) and the capsid of satellite tobacco necrosis virus (an icosahedron): Figure 1.16D,E. This symmetry is widely observed in viral capsids (Chapter 8).

Assembly proceeds along pathways

If a structure is generated from a large number of subunits, it is evident that it cannot arise from multiple simultaneous collisions: intermediate subcomplexes and assembly pathways are required, mirroring the situation with protein folding. Sometimes one can guess at these intermediates from inspection of the final structures, but the subcomplexes may be only transiently stable and consequently difficult to isolate. Moreover, they may not have exactly the same structure as in the final assembly. Tetramers may be dimers of dimers, suggesting a route from monomer to dimer to tetramer. Trimeric sub-assemblies are important intermediates in the assembly of octahedral and icosahedral cores of PDH complexes (different symmetries are observed in homologs in different organisms). With icosahedral virus capsids, dimers, trimers, or pentamers and hexamers can all serve as the building blocks.

The discussion thus far has been in terms of homomeric structures, but the same symmetry arguments apply equally well to heteromeric assemblies with mixed **protomers** (building blocks). For example, the thiocyanate hydrolase of *Thiobacillus thioparus* (Figure 1.17) has α , β , and γ subunits and can best be understood as four $\alpha\beta\gamma$ trimers arranged with dihedral symmetry. In later chapters, we describe functional complexes in which sub-assemblies with differing symmetries are combined, invoking symmetry mismatches, of which numerous examples are found in viruses (Chapter 8), in motile systems (Chapter 14), and in ATP synthase (Chapter 15). They pose a structural conundrum, but there is usually a plausible explanation for the mismatch, having to do with biological function.

Why are there so many large macromolecular assemblies?

Oligomerization makes it possible to have active sites that bridge the interface between subunits, and symmetrical assembly underlies the regulatory properties of cooperativity and allosteric. However, higher-order structures also offer other advantages. In multienzyme complexes, the mutual proximity of different active sites improves the speed and efficiency of catalysis. For carriers or repositories, bigger and bigger ‘containers’ can be constructed as one goes from tetrahedral to octahedral and then to icosahedral symmetry,

In other cases, reasons for oligomerization are less clear. For example, dimeric triose phosphate isomerase and tetrameric fructose 1,6-bisphosphate aldolase (see Figure 1.14) are two of the many proteins that have independent active sites and do not display any cooperative or allosteric effects. Nor is it clear why dehydroquinase (see Figure 1.16B) is found as a 12-mer with tetrahedral symmetry. One long-standing proposal has been that larger proteins are more stable, and stabilization is more economically achieved by aggregating small proteins than by increasing the sizes of individual proteins. Another argument in favor of oligomerization has been that of ‘editing’: thus, if a mistake is made in transcription or translation, the resulting defective subunit can be rejected in the subsequent assembly process. Here, long polypeptide chains are again at a disadvantage compared with shorter ones. There is some evidence for this. Single aldolase subunits are more thermolabile than the tetramer, and a naturally occurring mutation that limits assembly at the dimer stage is accompanied by decreased thermostability. Similarly, free pentamers and hexamers of the capsid protein of bacteriophage HK97 melt 20°C below the denaturation temperature of the assembled capsid (Figure 1.18). In contrast, it has been argued that, in the case of triose phosphate isomerase, a dimer enables fluctuations in the subunit interface to be channeled to the active site and thus to contribute to catalytic activity.

In oligomeric assemblies, individual subunits can make multiple interactions with more than two of their neighbors, except in the cyclic symmetries without accompanying dihedral symmetry. For example, in C_4 symmetry (see Figure 1.15B), the subunits on opposite sides of the ring make little or no contact, and even less as the number of subunits increases. This may account for the relative rarity among soluble enzymes of single rings with cyclic

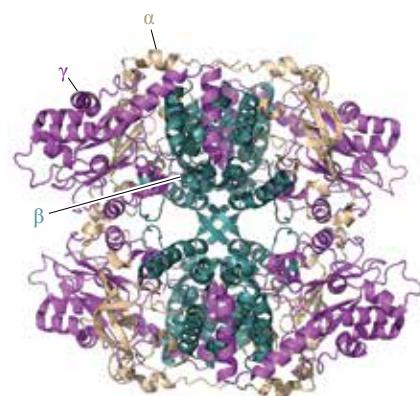


Figure 1.17 A dodecamer with dihedral symmetry composed of three different types of subunits. The thiocyanate hydrolase of *T. thioparus* has α subunits (wheat), β subunits (teal), and γ subunits (magenta). They are labeled on one of the four $\alpha\beta\gamma$ trimers. (PDB 2DD4)

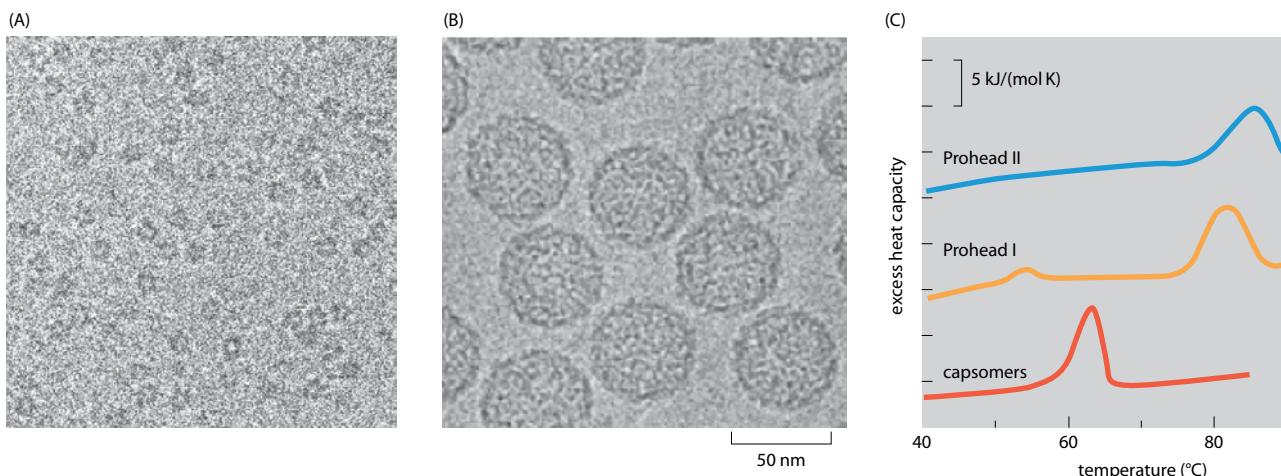


Figure 1.18 Enhancement of thermal stability in a macromolecular complex. The 41 kDa precursor capsid protein of bacteriophage HK97 forms hexameric and pentameric rings (capsomeres) that are the building blocks of the icosahedral procapsid Prohead I. Shown are cryo-electron micrographs of (A) capsomeres and (B) Prohead I. (A reconstruction is shown in Figure 8.41.) In differential scanning calorimetry, the specimen is heated so as to maintain a constant rate of increase in temperature. When it undergoes an endothermic phase transition, additional heat (enthalpy) must be applied to maintain the constant rate. The signal recorded is the excess heat capacity. Here, the main transition is thermal denaturation. At its peak (melting temperature, T_m), equal fractions of the protein are folded and unfolded. (C) The 20°C increase in T_m between the unassembled capsomeres (red curve) and assembled capsomeres (Prohead I, yellow curve) represents a major stabilization, resulting from inter-capsomer interactions. Proteolytic removal of the N-terminal domain (25% of the protein) converts Prohead I to Prohead II with a further stabilization (blue curve). The minor event at 54°C in the Prohead I thermogram is a thermally induced phase transition in which the pentamers switch to a more mature conformation. (A and B, courtesy of N. Cheng and R.L. Duda; C, adapted from P.D. Ross et al. *J. Mol. Biol.* 364:512–525, 2006. With permission from Elsevier.)

symmetry although they are more common in other kinds of complexes such as membrane-associated ones. In some cases, however, a hole in the center can have important functional roles, as in the hexameric AAA⁺ ATPases, in which the holes accommodate macromolecular substrates on which mechanical work is done (see Boxes 3.1 and 7.2). There are clearly numerous ways to create different symmetrical assemblies, even with the same number of subunits (for example, C₄ and D₂ tetramers and the three different 12-mers described above), and the evolutionary advantages will vary from one system to another. Mechanisms that regulate macromolecular assembly are discussed further in Chapter 8.

1.5 MACROMOLECULAR DYNAMICS

Our understanding of the mechanisms of molecular machines is based largely on static crystal and electron microscopy (EM) structures interpreted in the light of biochemical and other data. However, molecular assemblies are almost always dynamic, and they undergo a variety of movements and transitions during their duty cycles or during the course of their assembly and disassembly. For a given machine, insight into its functional properties depends on achieving an understanding of how many states it may populate, the structural transitions involved, their time course, and how they are regulated. Similarly, assemblies that lack machine-like properties but nevertheless have essential roles in defining cellular architecture and cell/cell interactions, also undergo structural changes.

The timescale of transitions in molecular biology spans many orders of magnitude. Those characteristic of various motions undergone by proteins are compared in Figure 1.19. At

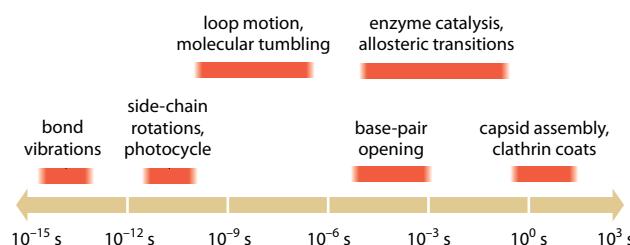


Figure 1.19 Timescales of molecular movements.

one extreme, steps in the photocycles of light-harvesting complexes, involving electron transfers or the *cis-trans* isomerization of *chromophores*, are in the femtosecond (10^{-12} s) range. On the other hand, changes that entail global movements of a macromolecular machine, involving many or all of its components, usually take place in milliseconds or, in some cases, much longer. For example, the assembly of a clathrin-coated pit (Section 10.2) can take 30–60 seconds, and its final pinching off perhaps 1 second. The times characteristic of some of the machine-based processes described in this book are summarized in **Table 1.3**.

Table 1.3 Times characteristic of movements of some macromolecular machines.

Movement	Time (ms)	Book section
Photocycles of light-harvesting complexes	10^{-9}	15.3
One revolution of a bacterial flagellum ^a	0.8–3.2	14.11
Viral DNA packaging motor translocates one base pair ^b	Minimum times: $1.5(\lambda) - 6(\phi29)$	8.5
DNA polymerase of phage T7 adds one base pair to a replicating DNA molecule ^c	3–4	3.2
Passage of a water molecule through an aquaporin channel ^d	8–50	16.4
Turnover number of a typical multienzyme complex (~100/s)	10	9.1–9.3
RNA polymerase adds one nucleotide to an mRNA ^e	10 (Pol II); 15–150 (Pol I)	5.2
Beating cycle of a sperm flagellum	25–33	14.8
One revolution of ATP synthase (F_1F_o ATPase) ^f	30	15.4
Translocation step by a motor protein ^g	0.2–400	14.3, 14.7
Chaperone-assisted protease ClpXP; 0.1 ms for scission; 4–15 sec for a substrate translocation step ^h	$0.1-10^4$	7.2
Ribosome adds one amino acid to a nascent polypeptide chain	50 (bacterial)–500 (eukaryotic)	6.2
A lamellipodium advances by 10 nm ⁱ	300	14.9
Translocon transfers one aa of polypeptide chain across a membrane	10^2-10^3	10.2

^a300 Hz (driven by proton gradient)–1200 Hz (driven by gradient of Na^+).

^bMaximum rates achieved early in packaging by bacteriophages λ and $\phi29$. Rates become slower as packaging proceeds on account of mounting back-pressure. Two base pairs are translocated for each ATP hydrolyzed. (D.N. Fuller et al., *J. Mol. Biol.* 373:1113–1122, 2007.)

^cIn elongation phase, replication having been initiated. Bases are added in separate processes to the two strands of a replicating duplex at approximately equal rates. There are two main steps: polymerization and proof-checking. (K.A. Johnson, *Biochim. Biophys. Acta* 1804:1041–1048, 2010.)

^dTransport of glycerol (through a glycerol-specific channel) is slower than water transport.

^eElongation phase, assuming that transcription has been initiated and is proceeding processively. Transcription also involves other, slower, steps such as promoter recognition and binding, and engagement with transcription factors. (G.L. Hager, J.G. McNally and T. Misteli, *Mol. Cell* 35:741–753, 2009.)

^fThree molecules of ATP synthesized per revolution. Each synthesis involves multiple substeps.

^gFor successive steps in processive motion along a cytoskeletal filament. They correspond to translocation speeds of 0.2–60 $\mu\text{m}/\text{s}$ for myosins (0.2–50 ms per cycle), and 0.02–2 $\mu\text{m}/\text{s}$ for kinesins (4–400 ms per cycle). (L.C. Sweeney and A. Houdusse *Phil. Trans. R. Soc. B* 359:1829–1841, 2004.)

^hAssumes that the protease complex has already bound substrate and initiated translocation or cleavage steps. There may be 50–100 ATP hydrolyzed to achieve a translocation step. (S.R. Barkow et al., *Chem. Biol.* 16:605–612, 2009.)

ⁱ2 $\mu\text{m}/\text{min}$ for leading edge. (J.I. Lim et al., *Exp. Cell Res.* 316:2027–2041, 2010.)

The structural changes involved in these transitions are of several kinds: the addition or detachment of protein subunits; rigid-body movements (rotations and translations) of individual domains or subcomplexes; order-disorder transitions; and refolding (**Figure 1.20**). There is also strong evidence that complexes (and individual proteins) undergo rapid (picoseconds to milliseconds) thermal fluctuations or ‘breathing’ motions whereby they fluctuate about their ground state, transiently exposing surfaces that are normally buried and/or unfolding peripheral motifs.

The experimental methods used to probe the dynamic behavior of macromolecules and complexes may be placed in two classes: ensemble methods and ‘single-molecule’ methods. Ensemble methods measure the net signal from numerous contributors. This property amplifies the signal to a detectable level but raises the problem that when a transition of interest is induced, the contributors do not necessarily proceed in synchrony, so that the signal recorded is a temporally smeared time-average. This problem is overcome in ‘single-molecule’ approaches that interrogate the complex of interest, one copy at a time, but the resulting signals are necessarily very weak. Nevertheless, by combining data from multiple measurements, it is possible to strengthen the signal as well as to assess its stochastic variability. In ‘single-molecule’ experiments, it is often necessary to adjust the experimental conditions so as to slow the transition down from its natural rate *in vivo*, to allow measurements to be made. It may also be necessary to modify the complex chemically to render the moving parts detectable.

Ensemble methods measure the net signal from numerous contributors

In **time-resolved X-ray diffraction**, the specimen is illuminated with an intense synchrotron-derived beam of X-rays. A transition is initiated—for instance, by an electrical stimulus or a flash of light—and diffraction is monitored on a fast read-out detector. Millisecond time resolution is possible with current devices. One may distinguish two kinds of experiment, depending on the specimen and the portion of the diffraction pattern that is monitored. The smaller this portion, the more rapid the read-out that is possible.

Time-resolved fiber diffraction studies have been pursued mainly in the context of muscle fiber activation. The specimen has components that exhibit two forms of helical symmetry so that the signal is confined to two sets of discrete reflections (‘layer-lines’) in the diffraction pattern. One set relates to the actin filaments and the other to the myosin filaments. An example is shown in Figure 14.24C,D. To interpret these data, a mathematical model of the interacting structures is formulated and its parameters are fitted to reproduce the observed changes in diffraction intensity and their time course.

In **SAXS (small-angle X-ray scattering)** experiments, the specimen consists of many randomly oriented molecules that may diffuse rotationally and translationally during the measurement. As a result, the diffraction pattern is rotationally symmetric. In this situation, it may be sufficient to sample the diffraction intensity on a single radial line, although a higher signal-to-noise ratio is obtained if a full two-dimensional pattern is recorded and the data are averaged around each ring. SAXS diffraction patterns have a relatively low information content but can yield valuable information about kinetics, particularly if other information is available about the structures of components. As an example, an analysis of the R→T transition of aspartate transcarbamoylase, which takes place within a tenth of a second, is given in **Figure 1.21**.

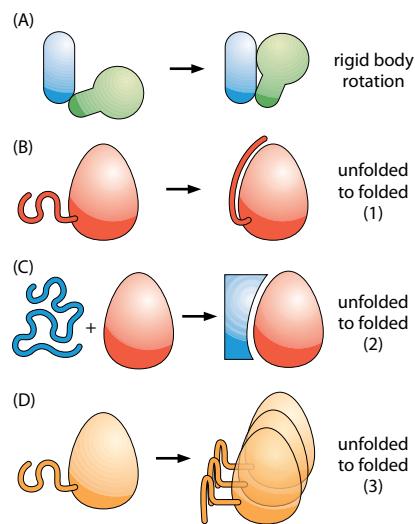
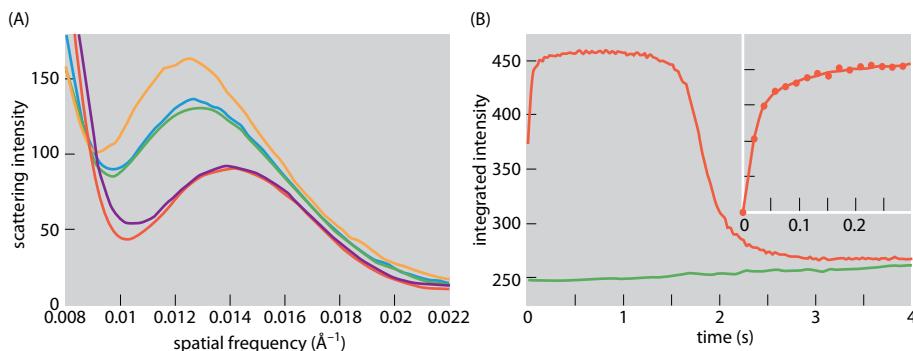


Figure 1.20 Different kinds of conformational change. (A) Rigid-body rotations of domains or subunits. (B–D) Association-induced folding of initially unfolded motifs, subunits, and domains. (B) The motif folds on binding to the folded core of the same protein. (C) An unfolded subunit (blue) folds on binding to another already folded subunit (red). (D) An initially unfolded domain polymerizes into an amyloid fibril decorated with folded globular domains of the same protein.

Figure 1.21 Kinetics of T→R transition of aspartate transcarbamoylase monitored by SAXS. ATCase catalyzes the reaction of carbamoyl phosphate (CP) with L-aspartate (Asp) to form N-carbamoyl-L-aspartate and inorganic phosphate (Pi) (see Figure 1.31). The *E. coli* protein has a catalytic core consisting of a dimer of homotrimers of 34 kDa subunits plus three regulatory trimers of 17 kDa subunits (see Figure 1.15C). The enzyme is regulated by binding nucleotide triphosphates. Crystal structures are known for both the low-activity (no substrates) T state and the high-activity R state. In the T→R transition, the catalytic trimers move apart by 11 Å and rotate by 5°. (A) SAXS patterns recorded 38 ms (blue curve), 380 ms (orange curve), and 3800 ms (purple curve) after mixing 0.75 mM ATCase with 50 mM substrates (CP and L-Asp). The pattern from the enzyme with CP and D-Asp at 3800 ms (red curve) is a T-state control. The green curve represents the sum of $0.33 \times$ (T-state curve) + $0.67 \times$ (R-state curve). Spatial frequency $s = 2\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the X-ray wavelength (1.38 Å). (B) Time courses of the quaternary structure change (red curve) and control (green curve), as monitored by scattering in the frequency band 0.085–0.152 Å⁻¹. Inset: the first 300 ms with a curve (two exponentials) fitted to the data points. (Adapted from J.M. West, J.R. Xia and E.R. Kantrowitz, *J. Mol. Biol.* 384:206–218, 2008. With permission from Elsevier.)

Time-resolved X-ray crystallography. Although crystallography has been the primary source of high-resolution information on static structures, it is less well suited to dynamic studies. Nevertheless, several avenues have been explored. They typically involve inducing a change in a protein crystal of known structure—for instance, by a flash of laser light—then measuring the resulting changes in diffraction intensities and interpreting them in structural terms. To collect enough photons in a fraction of a second, very intense X-ray sources are required, risking radiation damage. It is essential that the unit cell parameters of the crystal remain unchanged in the transition, which limits applicability to very small and localized changes in structure. One conceptually appealing approach uses Laue diffraction in which a broadband source is used instead of the usual monochromatic X-ray beam. Reflecting the formidable technical difficulties of these experiments, there have been only a few applications to proteins, for instance a study of the dissociation of carbon monoxide from hemoglobin.

Nuclear magnetic resonance (NMR) spectroscopy is based on the ability of a nucleus with a spin of $\frac{1}{2}$ (for example, ^1H , ^{13}C , ^{15}N , and ^{31}P) to adopt two different orientations in a magnetic field. The distribution of nuclei between the two states can be changed by subjecting them to a short pulse of radiation with a frequency commensurate with the energy difference between them. Monitoring the magnetic signals in the subsequent decay can yield dynamic information about the orientation and spacing of the nuclei, which provide restraints that can be turned into structural information.

A molecule in solution is free to tumble; however, tumbling slows with increasing size, causing broadening and overlapping of signals. Thus, the determination of protein structures by solution NMR has been restricted to molecules of no more than about 40 kDa. It is therefore of limited applicability to molecular machines but has been invaluable in determining the structures of components that could not be crystallized and in studies of the conformationally flexible parts of larger proteins that are invisible to X-ray crystallography. One of the earliest examples was the detection and identification of the Ala/Pro-rich linkers required for catalytic activity in a pyruvate dehydrogenase multienzyme complex (Section 9.4). Another example is the detection of mobility on a timescale of seconds in the N-terminal motifs that gate the α ring of subunits in the 20S proteasome (Section 7.3).

In time-resolved FRET (fluorescence resonance energy transfer), two fluorophores—a donor and an acceptor—conjugated to different components of a complex shift when the transition is induced (Figure 1.22A). Excitation of the donor is followed by a transfer of energy to the acceptor if they are sufficiently close. The acceptor then emits light at its given wavelength. This emission signal is monitored in time. Each donor/acceptor pair has a characteristic spacing (the Förster radius), the spacing at which FRET efficiency is 50% and falls off thereafter as the sixth power. An example of time-resolved FRET is given in Figure 1.22B, C for an investigation of the weak ($K_d \approx 10 \mu\text{M}$) interaction in which a 210 kDa complex is formed between the complement protein, C3b and a regulating factor (see Section

Figure 1.22 Kinetics of a conformational change from time-resolved FRET. (A) The distance between donor (D) and acceptor (A) fluorophores attached to different sites changes in a conformational transition. The transition may be initiated by rapid mixing with a suitable ligand. (B) The decrease in donor fluorescence and the accompanying increase in acceptor fluorescence report on a shortening of the interfluorophore distance. Fluorescence decay is generally used to monitor the kinetics. (C) Time-resolved fluorescence-decay curves for the binding of C3bAx488 to FH1–4Ax555. C3b is a key component of complement alternative pathway (Section 17.3) and FH1–4 is a four-domain, functionally competent, fragment of Factor H which helps regulate that pathway. The donor dye Ax488 (Alexa Fluor 488) was attached to a unique site in C3b and the acceptor Ax555 was attached to a Cys residue substituted for Asn102 in FH1–4. Donor-only decay is in green and donor/acceptor decay in pink. (B, adapted from D. Klostermeier and D.P. Millar, *Methods* 23:240–254, 2001. With permission from Elsevier; C, from I.C. Pechtl et al., *Protein Sci.* 20:2102–2112, 2011. With permission from John Wiley & Sons.)

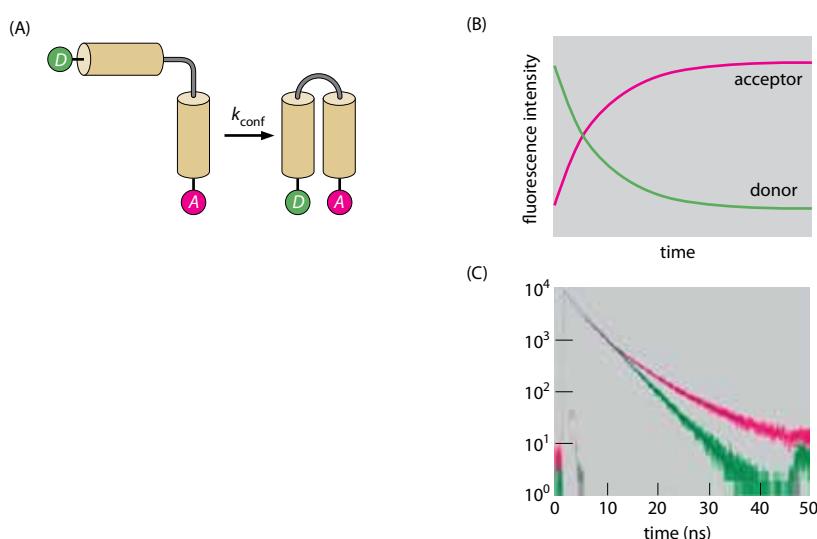
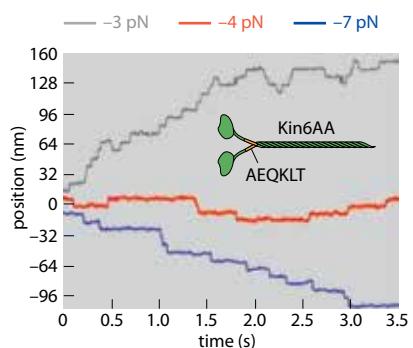


Figure 1.23 Time-resolved observation of kinesin-1 stepping by optical trapping. Kinesin-1 is a dimeric motor protein that moves along microtubules in 8 nm steps, driven by ATP hydrolysis (Section 14.7). Its stepping is highly processive. The steps can be forward (toward the plus end of the microtubule) or backward. These plots record the stepping histories in three experiments with individual molecules of the mutant, Kin6AA, which has a six-residue insertion in its neck-linker region and is more liable than wild-type kinesin-1 to take backward steps. The plots were smoothed by median filtering. Against an applied force of 3 pN most steps are forward, but at 7 pN all steps are backward. The time between successive steps varies stochastically. Data of this kind were used to formulate a five-state model of stepping kinetics. (From B.E. Clancy et al., *Nat. Struct. Mol. Biol.* 18:1020–1027, 2011. With permission from Macmillan Publishers Ltd.)



17.3). Cysteines were introduced at potentially informative positions in the proteins as sites for fluorophore attachment.

'Single-molecule' methods interrogate macromolecules one at a time

In force measurements by **optical trapping**, two reference points on the complex of interest are attached to the experimental apparatus either by antibodies or by covalent attachment to a glass bead, and an optical trap (Figure 14.17) is connected to each. The reaction is induced and the apparatus measures temporal changes in force exerted by the complex on the sensor. An example of using this approach to characterize the discrete steps made by a motor protein moving along microtubules is shown in **Figure 1.23**.

In time-resolved **atomic force microscopy** (AFM), a very fine tip is scanned across the surface of a specimen, mapping its height above a flat substrate. The technique has been refined to the point that dynamic events involving mobile components of a complex anchored on a substrate can be observed at video rates (~0.03 seconds per frame). An example of this approach has been visualization of the 36 nm 'hand over hand' steps taken by the motor protein myosin V (Section 14.3) moving along an actin filament (**Figure 1.24**).

There are several ways in which fluorescent probes can be used to track movements of macromolecular machines. In one, a fluorescent dye (or better, multiple copies of the dye) is conjugated to the particle of interest and its progress into and through a cell is followed at video rate in **live cell imaging** with a camera of sufficient sensitivity and resolution. Thus, movement of the particle of interest—viewed as a dot—can be charted in a biologically relevant context (**Figure 1.25**). In another application, rotational stepping by the F₁F₀ ATPase has been visualized by attaching a fluorescently labeled actin filament to the rotary motor and observing its movements, the filament being long enough for its rotation to be observed by light microscopy (Section 15.4).

Although cryo-EM visualizes individual molecules, it is not possible to record the same moving molecule at successive time points. Rather, in **time-resolved cryo-electron microscopy** (also called four-dimensional EM), dynamics are visualized by comparing images of molecules captured in successive states; in this sense, it is an ensemble method that depends on being able to distinguish between different conformational states. This source of variability has to be disentangled from that arising from the differing orientations that the molecule

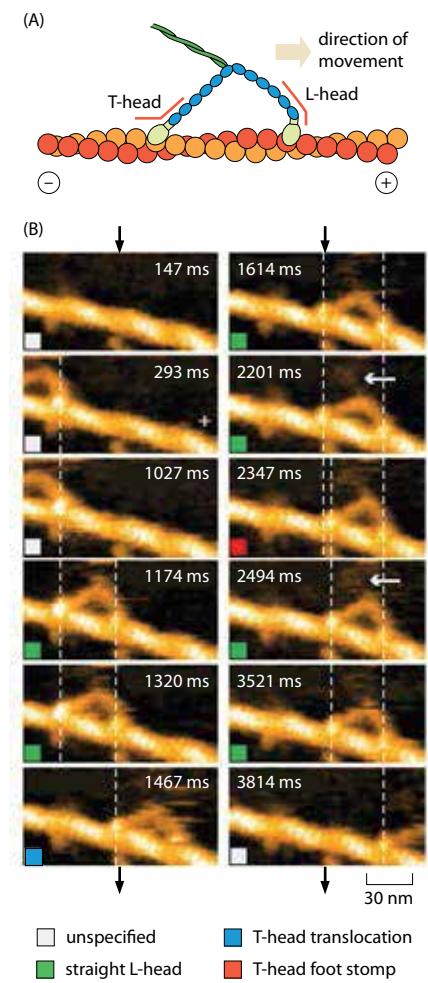


Figure 1.24 Video-rate imaging of myosin V stepping by atomic force microscopy. (A) Diagram of a myosin V construct with two heads and a coiled-coil linker (heavy meromyosin, HMM) moving along an actin filament. L-head, leading head; T-head, trailing head. (B) Successive AFM images taken at 146.7 ms per frame, showing the movement of M-V-HMM in 1 mM ATP. Pixel brightness conveys the height of the sampled surface above the substrate. The white arrow indicates the coiled-coil tail tilted toward the minus end of actin. Vertical dashed lines mark the centers of mass of the motor domains. The color-coded squares denote the states captured in each frame. The 'foot-stomp' is a head movement in which the head detaches and reattaches without translocating. (Adapted from N. Kodera et al., *Nature* 468:72–76, 2010. With permission from Macmillan Publishers Ltd.)

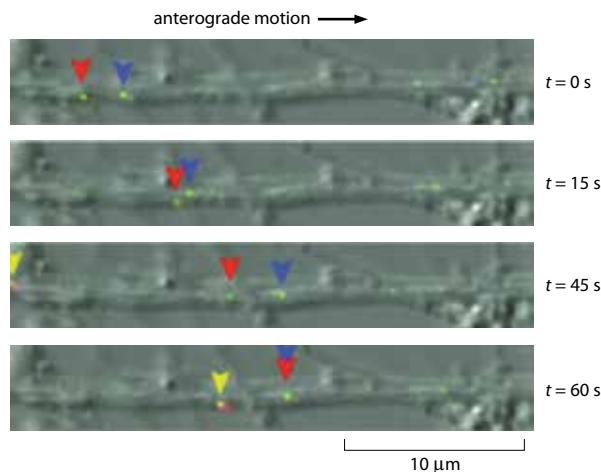


Figure 1.25 Intracellular transport of herpesvirus capsids monitored by time-resolved fluorescence microscopy.

Anterograde transport (that is, outward from the cell body toward the synaptic terminal) of pseudorabies virus (PRV) capsids along a neuronal axon imaged by confocal microscopy. PRV is a porcine virus belonging to the same subfamily as herpes simplex virus. The time-series images are merged overlays of differential interference contrast and fluorescence images with GFP and red fluorescent protein (RFP). Blue, red, and yellow arrowheads track the movement of fluorescent dots—mostly individual capsids, but see below—through the field of view. Red and blue arrowheads highlight yellow dots, which contain both GFP and RFP fused to a small capsid protein of which there are ~900 copies per capsid. The yellow arrowhead marks a predominantly red dot and a predominantly green dot transported together that eventually separate later, not shown. (From M.G. Lyman et al., *J. Virol.* 81:11363–11371, 2007. With permission from American Society for Microbiology.)

presents in projection. For this to be done reliably, the differences between the states must be substantial, and the states manageably few in number. Images are classified by computational techniques. The images in each class are averaged to reduce noise or reconstructed to obtain a three-dimensional structure. Data are collected at successive time points after initiating the transition, and the time courses with which the various conformers wax and wane are determined. Two examples of this approach have been to visualize the maturation of viral capsids (Figure 1.26) and to describe the power-stroke of the motor protein dynein.

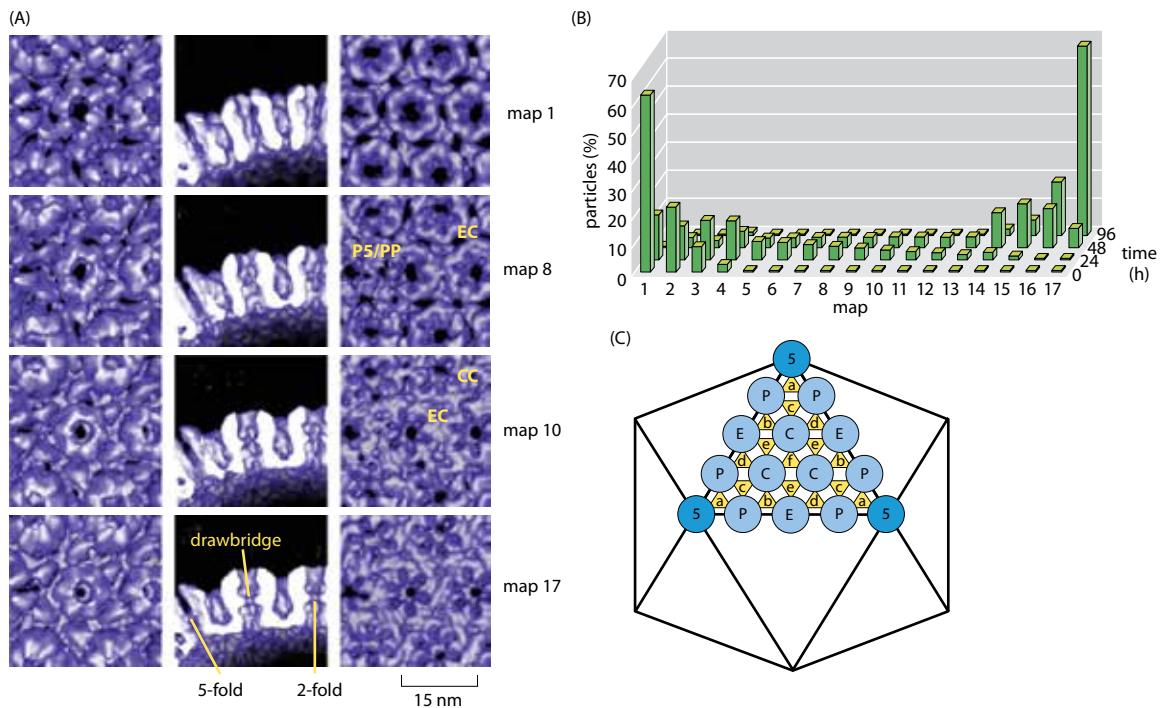


Figure 1.26 Successive states in the maturation of herpes simplex virus capsid captured by time-resolved cryo-electron microscopy.

Micrographs of the maturing capsids were recorded at four time points out to 96 hours. (*In vivo* maturation is much faster but is likely to proceed along the same pathway.) Capsid images were sorted into 17 classes, and a three-dimensional structure was calculated for each class. (A) Details from maps 1 (earliest precursor state), 8, 10, and 17 (final mature state). There is an axial channel through each capsomer, illustrated for a peripentonal

hexon in the cutaway views. The ‘drawbridge’ domain moves into a position in which it almost closes off this channel in the mature state. (B) Kinetics of maturation. The populations of each class were determined for each time point. (C) Mapping of the various capsomers: edge hexons (E), central hexons (C), peripentonal hexon (P), pentons (5), and triplexes (yellow, a–e). In (A), EC is the interface between an edge hexon and a central hexon, and so on. (Adapted from J.B. Heymann et al., *Nat. Struct. Biol.* 10:334–341, 2003. With permission from Macmillan Publishers Ltd.)

Molecular dynamics models the motions of crystal structures in the presence of a force field

To explore the full dynamic behavior of macromolecules, one enters the realm of computational biology. Here, ‘molecular dynamics’ (MD) studies are typically based on numerical simulation of movements taking place, starting from a known crystal structure in the presence of a mathematically defined force field. An all-atom simulation will involve many hundreds, if not thousands, of variables (the coordinates of each atom), each considered as a function of time. The scope of an MD experiment depends on their number as well as the duration of the simulation (now often as long as 1 μ s) as well as the reliability of the force field employed. These are complex calculations. Because most assemblies and machines discussed in this book are at least an order of magnitude larger than individual macromolecules, and their motions are slower—in the millisecond range or slower—they are tackled instead in **coarse-grained** approaches, in which the elements used to describe the dynamic structure are larger than single atoms; examples are amino acids, secondary structure elements, domains, and even protein subunits. Molecular dynamics has been particularly effective in analyzing the transport of solutes through membranes via channels or transporters (see Chapter 16). Indeed, the interior of a cell has been portrayed as being in a state of constant motion.

1.6 CATALYSIS

Most cells grow and replicate in an aqueous environment (about 70% of the weight of a cell is water), at or near neutral pH, and at temperatures no higher than 40°C. These conditions, not least water as solvent, are far from those that a chemist would expect to use in the laboratory. In biology, reactions require catalysis. Enzymes are catalysts that speed up the rate at which chemical reactions reach equilibrium. They cannot alter the free energy change (ΔG) of a reaction and thus its equilibrium position; rather, they accelerate the forward and back reactions equally. The rate enhancements achieved by enzymes are vast—up to 10^{17} -fold in one or two cases—and our understanding of their ability to do this continues to grow but is still incomplete.

All enzymes are proteins, apart from a few important exceptions where RNA acts catalytically, as in the ribosome (Section 6.2). In uncatalyzed chemical reactions, the rate is normally approximately doubled for a 10°C rise in temperature ($Q_{10} \approx 2$). However, enzymes function best at or near the growth temperature of the host organism and cease to function beyond their characteristic denaturation temperatures. An enzyme also usually works best on a given substrate over a particular pH range, centered on its optimum pH. This is usually near-neutral, except for enzymes that function in special environments, such as pepsin, a protease that operates at the low pH (<2) of the stomach, or enzymes of some extremophiles that live in conditions of low or high pH.

Enzymes form highly specific but transient complexes with their substrates

The network of biochemical reactions—the metabolic map—varies from organism to organism but in all cases is highly complex (see Nicolson maps cited in References 1.6). In general, each reaction is catalyzed by a dedicated enzyme that exhibits exquisite specificity for its substrate and can distinguish even between **stereoisomers** (Box 1.1). Biological chemistry is almost always restricted to particular stereoisomers, such as the L-series of amino acids and the D-series of sugars. The ‘wrong’ isomers are normally not recognized by the enzymes concerned; in some important instances, organisms may have taken evolutionary advantage of this. For example, the presence of D-amino acids in some bacterial antibiotics makes them resistant to degradative defense systems. In the pharmaceutical industry, the presence of the ‘wrong’ stereoisomer can be of crucial importance, for example in the disastrous effects of the drug thalidomide (see Box 1.1).

Enzymes follow **saturation kinetics**: as the substrate concentration is increased, the rate of the reaction catalyzed also rises, but it levels off at high substrate concentrations (Figure 1.27). This, and their exacting specificity, led to the concept of an enzyme **active site** and an **enzyme–substrate complex** being formed as a key step in catalysis. The **lock-and-key hypothesis** of Emil Fischer more than a century ago postulated a perfect fit between the

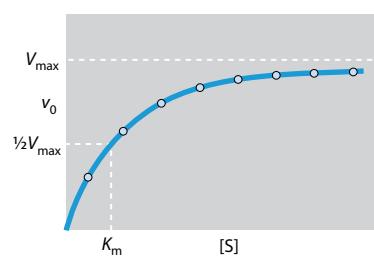


Figure 1.27 Plot of the initial velocity, V_0 , against substrate concentration $[S]$ for an enzyme-catalyzed reaction. The Michaelis constant, K_m , is the substrate concentration at which V_0 is half the maximal value, V_{\max} , achieved at saturating substrate concentration.

Box 1.1 Isomerism

There are three sorts of isomerism important in biology. **Structural isomerism** is the existence of two or more molecules made of the same atoms but with different bond arrangements, for example the sugars glucose and fructose ($C_6H_{12}O_6$). The presence of a double bond in a molecule introduces the possibility of **geometric isomerism**: *cis* and *trans* forms, different molecules with alternative dispositions of the substituents across the double bond. The third sort is **stereoisomerism**, in which a molecule contains a carbon atom with four different substituents, a so-called **chiral** carbon. There is no internal plane of symmetry and the molecule can exist in two mirror-image forms called **enantiomers**.

The original naming of all stereoisomers was based on glyceraldehyde, which contains one chiral carbon (Figure 1.1.1). Enantiomers rotate plane-polarized light in opposite directions; *levo* (–) to the left and *dextro* (+) to the right. All compounds with an absolute configuration of atoms around a chiral carbon atom that can be related to D-(+)-glyceraldehyde are part of the D-series, and conversely for the L-series. The designations D and L do not indicate whether the molecule is *dextro*- or *levo*-rotatory. Another naming convention is the R/S notation, but we use D- and L- in this book.

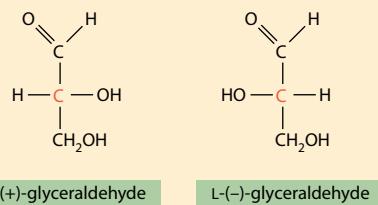


Figure 1.1.1 Glyceraldehyde exists in two enantiomeric forms.

The chiral carbon atom is in red. The two forms are mirror images in an imaginary plane between them.

If a molecule contains more than one chiral carbon it will exhibit **diastereoisomerism**; there will be as many chemically different compounds as there are chiral carbons, each of which exists as a pair of enantiomers. For example, the four-carbon sugars threose and erythrose (two chiral carbons) are diastereoisomers but not enantiomers; in other words, they are not mirror images of one another.

Both exist in D- and L-forms, which are enantiomeric. The forms depicted in Figure 1.1.2 have the same configuration about

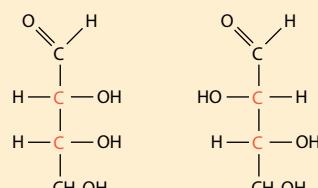


Figure 1.1.2 D-Erythrose and D-threose. The chiral carbon atoms are in red.

the bottom chiral carbon atom as D-glyceraldehyde (see Figure 1.1.1) and are therefore designated D-erythrose and D-threose. Note that D-erythrose and D-threose are *levo*-rotatory, whereas D-glyceraldehyde is *dextro*-rotatory.

Stereoisomerism is of major importance in biology. Proteins almost always consist of only L-amino acids, and carbohydrates of D-sugars. Macromolecules are asymmetric and therefore likely to recognize only one enantiomer of any given compound. Failure to take account of this can be dangerous. Thalidomide was sold as a medication effective against morning sickness for pregnant women. It contains a chiral carbon atom and so exists in D- and L-forms (Figure 1.1.3). The commercial preparation was a **racemic mixture**—it contained equal amounts of the two enantiomers, one of which induced developmental abnormalities and led to a tragic outbreak of birth defects. It is thought that this isomer may be able to insert into the major groove of DNA at GC-rich regions, interfering with gene transcription and thus inhibiting the formation of blood vessels. Unfortunately the enantiomers are able to interconvert *in vivo* and administering only one enantiomer means that both enantiomers will still be present in the serum. More recently, thalidomide has found favor again, in carefully controlled application, for such diseases as AIDS, leprosy, lupus, and tuberculosis.

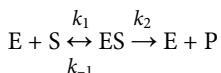


Figure 1.1.3 The structure of thalidomide. The chiral carbon is arrowed.

enzyme and its substrate. Today we speak instead of **molecular complementarity**. This accommodates the possibility of conformational change on the part of the protein when it comes into contact with the substrate, or the selection of a particularly efficacious conformation of the enzyme from an ensemble of related structures.

Enzyme kinetics are governed by a few equations

For the single-substrate reaction



where S is the substrate and P the product, if $[E_0]$ is the total concentration of the enzyme,

the initial rate, v , of the forward reaction is given by

$$v = \frac{k_2[E_0][S]}{[S] + (k_2 + k_{-1})/k_1}$$

and the **Michaelis constant**, K_m , by

$$K_m = (k_2 + k_{-1})/k_1$$

an equation first derived by George Briggs and J.B.S. Haldane in 1925. This followed earlier work of Leonor Michaelis and Maud Menten in 1913, who derived a similar equation but assumed that k_2 was very small compared with k_{-1} . If $k_2 \ll k_{-1}$, $K_m \approx k_{-1}/k_1$, which is the same as K_s , the dissociation constant of the enzyme–substrate complex. The **catalytic rate constant** for the forward reaction, conventionally known as k_{cat} , equals k_2 . It is sometimes referred to as the **turnover number**, because it reflects the number of substrate molecules converted to product per second at the enzyme active site. From the equation above, K_m can be defined as the substrate concentration at which half-maximal velocity ($\frac{1}{2}V_{\max}$) is achieved (see Figure 1.27). At low substrate concentrations, $v = (k_{\text{cat}}/K_m)[E_0][S]$, which implies that k_{cat}/K_m can be regarded as an apparent second-order rate constant. Thus it is a measure of the specificity of the enzyme for its particular substrate—hence its other name of **specificity constant**.

In many instances, $k_2 \ll k_{-1}$ as Michaelis and Menten assumed. In such cases, $K_m \approx k_{-1}/k_1$ (see above) and gives a direct measure of the affinity ($k_1/k_{-1} = 1/K_s$) of the enzyme for its substrate. However, the assumption should be tested for any given enzyme. If $k_2 \gg k_{-1}$, $k_{\text{cat}}/K_m \approx k_1$, the rate constant for the formation of the enzyme–substrate complex. k_1 is typically about $10^8/(s\text{ M})$ for enzymes and their substrates, and a measurement of k_{cat}/K_m approaching this value is a strong indication that Briggs–Haldane rather than Michaelis–Menten kinetics are being followed. With some enzymes, additional intermediates occur on the reaction pathway, and then $K_m < K_s$.

A key feature of enzyme catalysis is the tight binding of the transition state

In a typical chemical reaction, the reactants pass through one or more **transition states** on the way to product, and the energy required to reach the highest point in the profile is the **activation energy** that must be surmounted if the reaction is to go to completion (Figure 1.28A). In the transition state, chemical bonds are being made and broken, whereas the depressions in an energy profile are populated by intermediates in which bonds are fully formed. The rate of the overall reaction is the concentration of the transition state multiplied by the rate constant for its decomposition. In an enzyme-catalyzed reaction, the activation energy is lowered (Figure 1.28B) and the rate of reaction is thereby enhanced.

Following earlier work by J.B.S. Haldane, Linus Pauling suggested that an enzyme speeds up a reaction by stabilizing the transition state, binding it more tightly than the substrate and thereby lowering the activation energy. Specificity, represented by the formation of an ES complex, and catalysis, represented by the transition state, are linked, but the crucial feature is that the enzyme and transition state should adopt complementary structures. The results of protein engineering experiments designed to dissect the contributions made by individual amino acid side chains to the catalytic mechanisms of particular enzymes are consistent with this idea. Moreover, if a molecule resembling the transition state of a chemical reaction is used as an antigen, **catalytic antibodies** can be elicited that are able to catalyze, albeit sometimes only poorly, the reaction represented by the transition state analog.

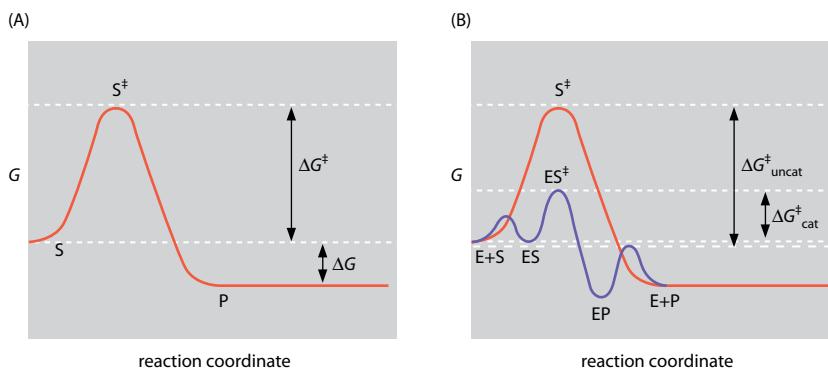


Figure 1.28 Free energy profiles of uncatalyzed and enzyme-catalyzed reactions. G , Gibbs free energy; S , substrate (reactant); E , enzyme; P , product. (A) A simple uncatalyzed chemical reaction passes through the transition state, S^\ddagger . ΔG^\ddagger is the free energy of activation (that is, the difference in free energy between reactant and the transition state). ΔG represents the difference in free energy between reactant and product. (B) In a simple enzyme-catalyzed reaction (blue curve), ES^\ddagger is the transition state and ES and EP represent the enzyme–substrate and enzyme–product complexes. ΔG is unchanged. $\Delta G^\ddagger_{\text{uncat}}$ is the free energy of activation for the uncatalyzed reaction, and $\Delta G^\ddagger_{\text{cat}}$ is the free energy of activation for the enzyme-catalyzed reaction. The difference, $\Delta\Delta G^\ddagger$, is a measure of the efficacy of the catalyst.

Enzyme-catalyzed reactions usually proceed through the direct making or breaking of bonds over distances of up to 2 Å (see below). However, in some electron transfer proteins (Chapter 15), *redox centers* are as much as 15–20 Å apart. In such enzymes, electron transfer proceeds by *quantum mechanical tunneling*; that is, the electron tunnels through a transition state barrier that classically it could not surmount. The rate of electron transfer is rapid, much faster than the millisecond turnover times that are characteristic of most enzymes, and can readily accommodate the picosecond reaction times common in photosynthesis. Some enzymes involved in C–H transfers also appear to operate by quantum mechanical tunneling of the hydrogen, though because of its much higher mass (1800× that of an electron) the tunneling distance is limited to at most 1 Å. In contrast with conventional transition state theory, tunneling is more a function of barrier width than of barrier height (Chapter 15).

Enzymes generate catalytic rate enhancements in multiple ways

An enzyme creates a ‘pre-organized environment’ that promotes catalysis. It selectively binds substrates and orients them advantageously in the active site, creating a microenvironment that facilitates reactions that would be difficult to achieve in an aqueous environment. It positions catalytically important amino acid side chains to facilitate acid–base catalysis and/or other features of an efficient catalytic mechanism. It also promotes a high effective concentration of intramolecular species that overcomes the loss of entropy that is unavoidable in intermolecular reactions. It makes possible covalent catalysis in which relevant groups (for example phosphate) can be transiently lodged on the protein, and it binds **cofactors**. Cofactors can be metal ions, **coenzymes** (such as NAD⁺ or NADP⁺, which shuttle as carriers between oxidoreductases), and **prosthetic groups**, which differ from coenzymes in being bound very tightly or even covalently to the parent protein. The chemical mechanism of the protease chymotrypsin (Figure 1.29) illustrates some of these features.

Enzymes are not static structures. In addition to allowing the enzyme to adopt a structure complementary to the transition state, conformational flexibility may be required to allow substrates to enter and products to leave, and to close off the active site from the aqueous environment. If water molecules are expelled, there will be a gain in entropy that promotes catalysis. Groups essential to catalysis can be moved into position as the reaction proceeds, and high precision (<1 Å) in their placing is usually necessary.

Most enzymes operate on a millisecond timescale or faster. NMR spectroscopy and other techniques have shown that proteins undergo multiple reversible motions away from their average structures. The timescales involved vary from seconds to milliseconds for the rearrangement of subunits or domains down to nanoseconds to picoseconds for local movements of side chains. How coupled ensembles of fluctuating conformations contribute to the rate enhancement of reactions as enzymes progress through their catalytic cycles is not yet settled, but the concept of an energy landscape like that observed in protein folding (see Figure 1.12) is gaining acceptance. Nonetheless, the conclusion remains that the catalytic power of enzymes resides in a lowering of the activation energy and an increase in the **generalized transmission coefficient** that results from dynamic crossing and recrossing of the activation barrier.

Enzymes can be inhibited reversibly and irreversibly

Enzymes can be inhibited in various ways. **Competitive inhibitors** closely resemble the substrate in structure and can bind reversibly at the active site, thereby blocking the reaction. They do not affect the value of V_{max} , because saturating concentrations of substrate will swamp a finite concentration of inhibitor. However, the K_m for the substrate will apparently increase, as the enzyme requires a higher concentration of substrate to reach $\frac{1}{2}V_{max}$. In another type of reversible inhibition, a **transition state analog** (a molecule that resembles the transition state but cannot be acted on by the enzyme) binds more tightly to the active site than does the substrate itself. Transition state analogs have been developed for several enzymes and have helped in the elucidation of catalytic mechanisms and in structure-based drug design.

A third and different form of inhibition, generally irreversible, is **covalent modification** of the enzyme. For example, a reagent may preferentially react with a protein functional group that displays unusual reactivity. If this group is in or near the active site, the catalytic reaction may be impaired. In a related form of inhibition, **active site-directed**

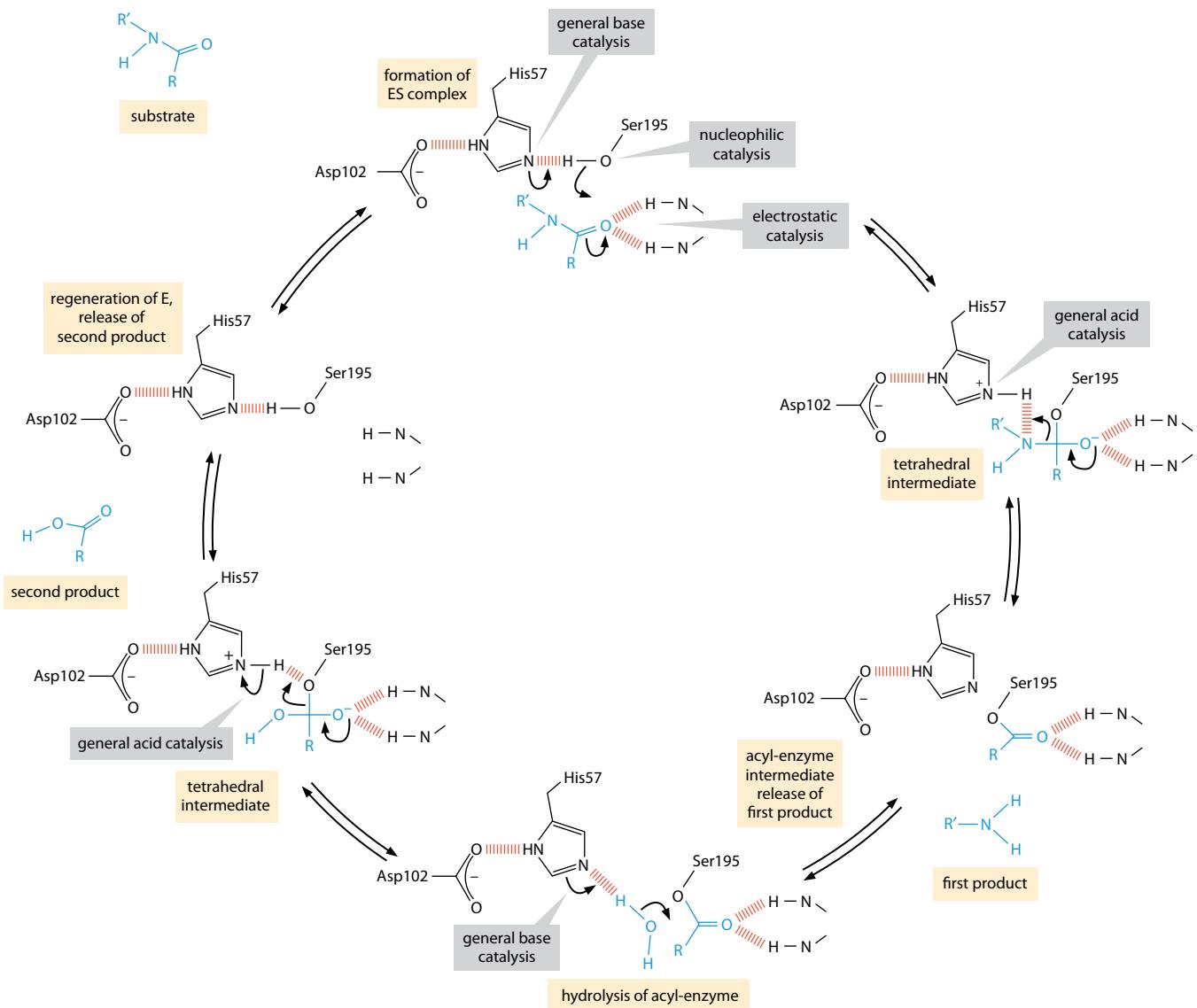


Figure 1.29 Catalytic mechanism of the pancreatic protease

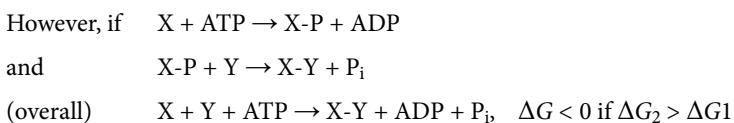
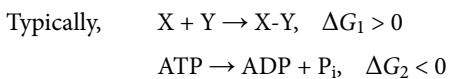
chymotrypsin. Substrate specificity is provided by the bulky hydrophobic side chain of the amino acid (Phe, Tyr, Trp) on the N-terminal side of the scissile peptide bond, which enters a narrow hydrophobic pocket located close to the catalytically important groups of the enzyme. The catalytic reaction can be considered as comprising two halves. In the first, the negative charge on the carboxyl group of Asp102 acts to stabilize a positive charge on the imidazole ring of His57, to which it is H-bonded. This makes it easier for the hydrogen of the hydroxyl group in the side chain of Ser195 to be donated to the uncharged imidazole ring (general base catalysis), thereby promoting nucleophilic attack of Ser195 on the >C=O group of the scissile peptide bond of the bound peptide substrate. The trigonal carbon of the >C=O is held in the correct position for this attack by two fractionally charged >NH groups (Gly193 and Ser195) in a so-called oxyanion hole provided by the peptide backbone of the enzyme (electrostatic catalysis). This generates the tetrahedral intermediate in which the newly formed

>C=O⁻ moves deeper into the oxyanion hole, where it forms H-bonds with the NH groups of Gly193 and Ser195 (stabilization of the transition state). This adduct in turn collapses to the acyl-enzyme intermediate, facilitated by the donation of a proton from the positively charged imidazole ring of His57 to the R'NH₂ leaving group (general acid catalysis), and release of the first product, R'NH₂. In the second half of the reaction, a water molecule interacts with the now uncharged imidazole ring of the Asp102-His57 pair. This helps activate it (general base catalysis) as a nucleophile for attack on the acyl-enzyme, thereby forming another tetrahedral intermediate. Collapse of this tetrahedral intermediate, with the return of a proton from the positively charged imidazole ring (general acid catalysis), regenerates the hydroxyl group of Ser195 and thus the starting enzyme, which is accompanied by release of the second product, RCOOH. The numbering of residues is that of the inactive precursor (zymogen), bovine chymotrypsinogen. (Adapted from A.R. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, W.H. Freeman, 1998.)

inhibitors—molecules that resemble the substrate but also carry a ‘chemical warhead’—are targeted at the active site. Selective reaction of the warhead with a nearby functional group can irreversibly inhibit the enzyme. The warhead might be an alkylating agent, or an electrophile in search of a protein nucleophile, or even a highly reactive species generated *in situ* from a benign substituent by a laser flash (such as a nitrene from an azide).

Coupling of enzyme-catalyzed reactions allows energetically unfavorable reactions to occur

A reaction will proceed in a given direction only if it has a favorable free energy change (a negative value of ΔG). However, many reactions in cells have unfavorable changes in free energy and are brought about by being coupled with another reaction that has a larger, negative value for ΔG —often, for example, the hydrolysis of ATP.



Facilitating the catalysis of coupled reactions is one of the key features of many molecular machines; to give just three diverse examples, the synthesis of DNA by polymerases (Chapter 3), the formation of peptide bonds by the ribosome (Section 6.2), and carboxylation reactions catalyzed by biotin-dependent multienzyme complexes (Section 9.4).

1.7 SIGNALING AND REGULATORY MECHANISMS

Biological processes must be coordinated and regulated if cells are to function in an orderly way. Signals are of crucial importance, whether coming from outside the cell or sent between or within compartments inside the cell. The fate of a cell (or indeed an organism) can rest on the speed and efficiency with which signals are transmitted and acted upon. Growth factors and hormones—chemical messengers such as insulin and epinephrine (adrenaline) that circulate in the blood of higher animals—have long been known to transmit signals across the cell membrane, generating intracellular responses. Many molecular machines and assemblies are key players in these and other signaling processes. Their activity can be regulated in several ways, some slow and irreversible, others faster and reversible.

Cells are able to turn on and off the synthesis of a given protein by controlling the transcription of its gene (Chapter 5). However, this process is slow (minutes to hours) and the protein remains active until it decays or is degraded. Another widespread control mechanism is the action of proteases to generate active enzymes from inactive precursors, called **pro-enzymes** or **zymogens**. This is how, for example, the proteases responsible for initiating apoptosis (Section 13.6) or for activating complement in an immune response (Section 17.3) are generated. Again, though, the relevant proteins remain active until degraded. In contrast, a protein can undergo a conformational change on the millisecond timescale in response to binding a ligand or to post-translational modification. This mechanism underlies many rapid signaling processes and can be envisaged as having conferred great evolutionary advantage.

Ligand-induced conformational change and cooperativity are widespread methods of controlling biological activity

A general feature of signal transduction is that of **ligand-induced conformational change** (Figure 1.30). The ligand can be small—an ion or small molecule (a *chemical messenger*)—or large, such as another protein. Binding of the ligand to a protein induces a conformational

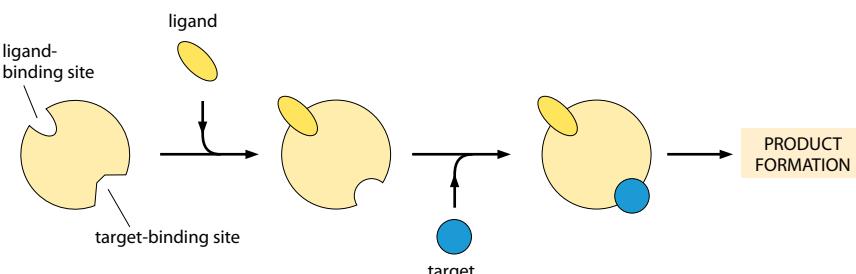


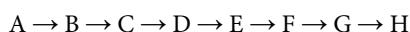
Figure 1.30 Ligand-induced conformational change. Binding of ligand at one site induces a change at the remote active site on the same protein molecule, thereby enabling it to bind the target and process it.

change that enables the protein to recognize its target molecule and act on it to initiate the desired biological activity. Activity can thus be turned on or off in a fraction of a second. The effect of a ligand-induced conformational change might be simple and direct, such as stimulating an enzyme to bind its substrate and initiate a catalytic reaction, but it can be more elaborate. For example, the binding of lactose to the *lac* repressor induces a conformational change that causes the repressor to dissociate from the *lac* operator sequence in *E. coli* DNA, allowing the transcription of genes in the *lac* operon; equally, the binding of insulin to its transmembrane receptor (Section 16.2) or of growth factors to G-protein-coupled receptors (Section 12.2), sets in train intracellular *protein kinase* cascades that amplify and transmit signals to the nucleus or other targets (Chapter 12). Many more examples recur throughout this book.

Protein oligomers offer the possibility of a ligand-induced conformational change in one subunit being transmitted to a neighboring subunit. If the oligomer has two conformational states, it might then be locked into one or the other state by ligand binding. These possibilities are manifested in cooperativity (Section 1.2); that is, one event can make another similar event either more or less likely to follow. Cooperativity is generally a property of oligomers and commonly of symmetrical oligomers. It is a widespread method of transmitting signals through macromolecular complexes, as in the oxygenation of the hemoglobin $\alpha_2\beta_2$ tetramer whereby the binding of one O_2 molecule makes it easier for a second to bind, and so on up to a maximum of four.

Allosteric proteins are regulated by a special form of cooperativity

Consider a metabolic pathway of the kind



In the biosynthesis of pyrimidine nucleotides, the first ($A \rightarrow B$) of seven steps is the formation of *N*-carbamoylaspartate from aspartate and carbamoyl phosphate, catalyzed by aspartate transcarbamoylase (ATCase) (Figure 1.31). ATCase is reversibly inhibited by CTP, the end product (H) of the pathway. Thus, if the concentration of CTP in the cell rises, it down-regulates its own biosynthesis (and conversely, if the concentration of CTP falls). This process is termed **feedback inhibition**. CTP has a structure totally different from either of the substrates for ATCase and is therefore unlikely to act as a competitive inhibitor. ATCase is a complicated oligomeric enzyme (see Figure 1.15C). Substrates are bound and acted upon at a site in a catalytic subunit, whereas CTP binds to a site in a regulatory subunit yet is able to exert its inhibitory effect on a catalytic site more than 50 Å away. A time course of activation is shown in Figure 1.21. In contrast, the presence of ATP stimulates the activity of ATCase, and a high concentration of purine nucleotides in the cell therefore acts as a signal to ATCase to increase the production of the complementary pyrimidine nucleotides for nucleic acid synthesis. ATP, like CTP, is quite different in structure from aspartate and carbamoyl phosphate and binds at a different site, also remote from the catalytic site.

Pyrimidine biosynthesis is just one of many examples of this kind of metabolic control. The fact that the inhibitory or stimulatory molecules are structurally different from the substrates in such enzymes led Jacques Monod, Jeffries Wyman, and Jean-Pierre Changeux to introduce the term **allostery** (Greek *allos*, other, and *stereos*, shape). ATCase is an extreme example of the principle of separate but communicating binding sites for substrates, allosteric inhibitors, and allosteric activators in that its catalytic and regulatory sites are, unusually, in different subunits (Figure 1.32).

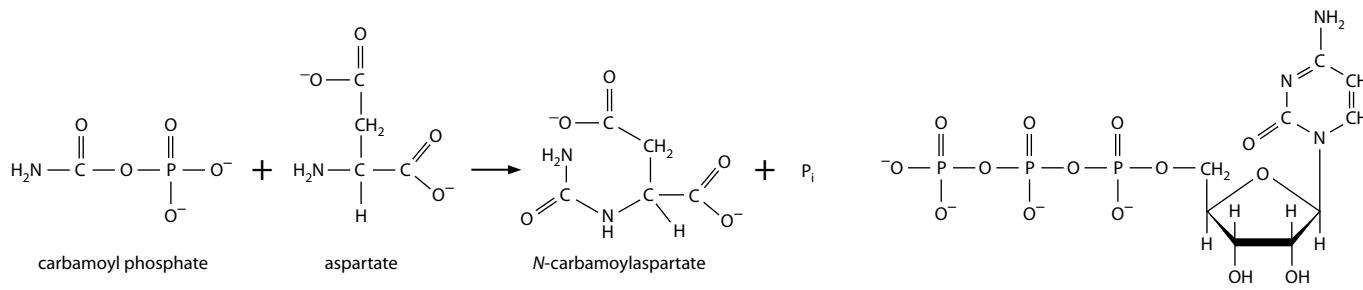


Figure 1.31 The reaction catalyzed by aspartate transcarbamoylase and the structure of the allosteric inhibitor, CTP.

cytidine 5'-triphosphate (CTP)

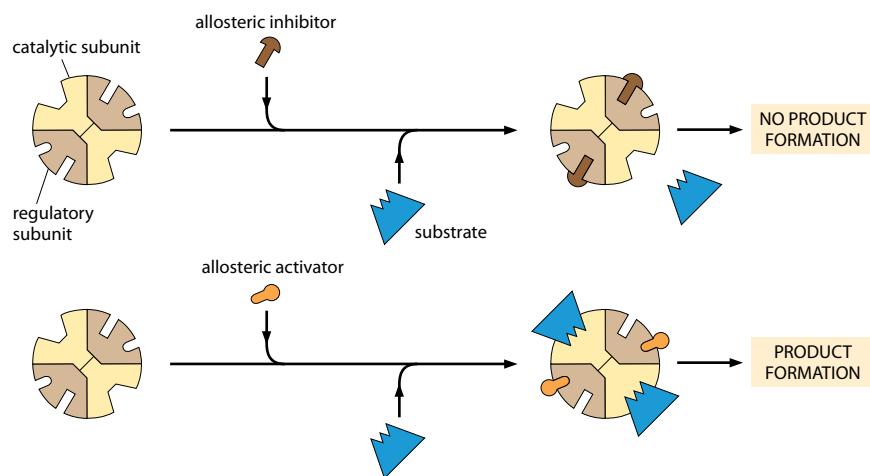


Figure 1.32 Allosteric regulation of a multi-subunit enzyme. This model enzyme is a hetero-oligomer of regulatory and catalytic subunits. Binding of an effector molecule (inhibitor or activator) to the regulatory subunits is transmitted to, and controls the activity of, the catalytic subunits.

Allosteric enzymes do not follow Michaelis–Menten kinetics

A plot of velocity against substrate concentration for an allosteric enzyme is an S-shaped (sigmoid) curve (Figure 1.33A), not the hyperbola typically observed for unregulated enzymes. This is the result of positive cooperativity. A sigmoidal $v/[S]$ plot means that a small change in substrate concentration produces a bigger change in catalytic rate than it does with a hyperbolic Michaelis–Menten $v/[S]$ plot. Moreover, the effect can be further amplified by **allosteric inhibitors** or **activators** (Figure 1.33B). Interactions involving identical molecules (such as the cooperative binding of substrate to an enzyme or of O_2 to hemoglobin) are described as **homotropic**, and those involving allosteric effectors (inhibitors and activators) as **heterotropic**.

Allostery is mediated by protein/protein interactions and conformational changes

Most allosteric proteins are small oligomers, and the Monod–Wyman–Changeux (MWC) theory explained cooperativity in allosteric enzymes and hemoglobin on the basis of the following assumptions (Figure 1.34A):

- The quaternary structure is symmetrical (or pseudo-symmetrical, as in $\alpha_2\beta_2$ hemoglobin).
- The oligomers are in dynamic equilibrium between two conformations, a ‘tense’ (T) state and a ‘relaxed’ (R) state.
- In the absence of any substrate or signaling molecule, at equilibrium the concentration of the T state is higher than that of the R state.
- The T state has a lower affinity for ligands, and hence lower catalytic activity, than the R state.
- Binding of a substrate or allosteric activator shifts the equilibrium in the direction of the R state, whereas an allosteric inhibitor shifts it toward the T state.

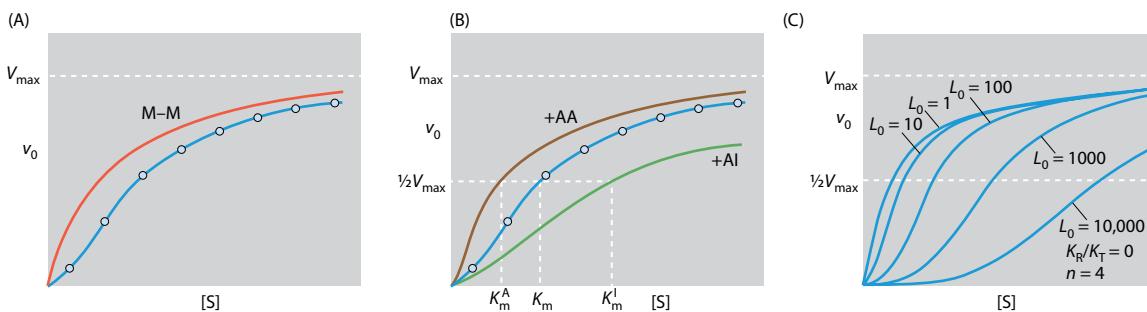
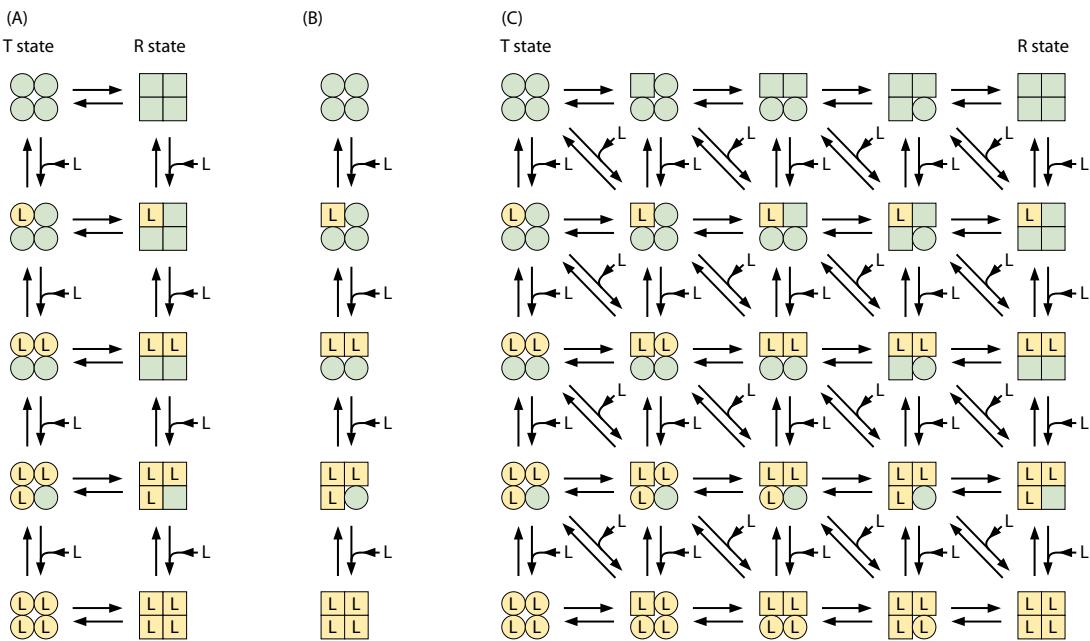


Figure 1.33 Kinetics of allosteric enzymes. (A) Plot of the initial velocity, V_0 , against substrate concentration $[S]$ for an allosteric enzyme (blue) and an enzyme obeying Michaelis–Menten (M–M) kinetics (red). (B) Plot of V_0 against $[S]$ for an allosteric enzyme in the presence of an allosteric activator (AA) or an allosteric inhibitor (AI). The K_m for substrate

S decreases in the presence of the activator and increases in the presence of the inhibitor. This is typical for the class of allosteric enzymes in which the value of V_{max} remains unchanged. (C) Effect of differing values of L_0 for an allosteric tetramer ($n = 4$) with $K_R/K_T = 0$.



The sigmoid binding curve of an allosteric protein can then be calculated from a few parameters: a constant $L_0 = [T]/[R]$, the ratio of the T and R states in the absence of any ligand binding, and K_R and K_T , the dissociation constants for substrate binding to the R and T states, respectively. In comparison with the Michaelis–Menten equation, the $v/[S]$ relationship becomes

$$v = \frac{V([S]/K_R)(1 + [S]/K_R)^{n-1}}{L_0(1 + [S]/K_T)^n + (1 + [S]/K_R)^n}$$

For most allosteric enzymes K_R/K_T is in the range 0.01–0.04 (meaning that, to a good first approximation, the $[S]/K_T$ term can be neglected) and L_0 ranges from ~60 to 10^4 (see Figure 1.33C).

If an allosteric activator displaces the R/T equilibrium so that the R state predominates, cooperativity in substrate binding is no longer observed and the enzyme follows Michaelis–Menten kinetics. This has been verified experimentally for several enzymes and is beautifully clear for ATCase: if separated from the regulatory dimers, the catalytic trimers exhibit straightforward Michaelis–Menten kinetics.

An alternative model for allostery was proposed by Daniel Koshland, George Nemethy, and David Filmer (KNF) in which, in the absence of substrate, the oligomer exists in a single state constrained by subunit/subunit interactions (Figure 1.34B). Symmetry is not conserved, and the conformation of each subunit changes when it binds substrate. If, as a result, this makes it easier for a neighboring subunit to switch to the active conformation, there is positive cooperativity; if it is made more difficult, negative cooperativity is observed. The KNF model can readily be extended to incorporate heterotropic interactions by regulatory activator or inhibitor molecules. Unlike the MWC theory, the KNF theory is able to explain negative cooperativity and is also compatible with ‘half-of-the-sites reactivity’, in which only half of the active sites in the oligomer are active at any one time.

If the subunit/subunit constraints are sufficiently tight, once one subunit is in the R state, all others must take up the R state, and the KNF and MWC theories coalesce. Indeed, Manfred Eigen pointed out that they can both be accommodated in a general overall framework (Figure 1.34C) in which the MWC model populates only the extreme states, whereas the KNF model populates the states found on the diagonal. There are many examples of positive cooperativity; negative cooperativity, although less common, is observed in machines such as the GroEL/GroES chaperone (Section 6.3), the E1 components of pyruvate dehydrogenase multienzyme complexes (Section 9.4), and ATP synthase (Section 15.4). However, allosteric complexes often show elements of both MWC and KNF theory.

Figure 1.34 Models of allostery. The models are illustrated for a tetrameric protein. L, a ligand (substrate or allosteric modulator); circles represent T-state subunits (low affinity for L) and squares represent R-state subunits (high affinity for L). (A) The MWC model. The conformational changes in the protein subunits are concerted; no tetramer contains both R- and T-state subunits at any one time. (B) The KNF model. Conformational changes are not concerted and each subunit can be in either the T-state or the R-state. In both models, increasing the concentration of L causes more subunits to switch into the R-state. (C) The generalized model of allostery in which the MWC model is the two outer columns and the KNF model is the diagonal from top left to bottom right.

Reversible covalent modification controls the activities of some proteins

Chemical modifications afford another way to rapidly regulate the activity of a target protein. In eukaryotes, post-translational modifications of proteins are many and varied, but phosphorylation (catalyzed by kinases) and dephosphorylation (catalyzed by phosphatases) are the most common (Chapter 12). It is important that the modification and its reversal are carried out by different enzymes, so that a signal used to turn on the targeted activity differs from the one that turns it off. The sites of phosphorylation are the hydroxyl groups of serine, threonine, or tyrosine residues. The hydroxyl group, although polar, carries no charge at neutral pH (the pK_a of the OH group in serine or threonine is around pH 14; that of the OH in tyrosine is about pH 10.5). The phosphorylated side chain, however, is negatively charged, which can have the effect of promoting a conformational change in the modified protein that modulates its activity.

The widespread importance of phosphorylation and dephosphorylation as a signaling and control mechanism is reflected in the fact that 518 putative protein kinase genes have been identified in the human genome; of these, 244 map to disease loci. Some 150 different phosphatases, >100 of which are phosphotyrosine-specific, catalyze dephosphorylation. As many as one-third of human intracellular proteins may be subject to regulatory phosphorylation.

Homeostasis is an important aspect of response to environmental change

The concept of the **rate-limiting step** has been a key feature for understanding how a sequence of reactions in a metabolic pathway is regulated. However, **metabolic control analysis** has demonstrated, counter-intuitively perhaps, that each enzyme contributes appreciably to the overall flux control (Box 1.2). In response to a signal, an increase or decrease in the flux through the pathway is accompanied by smaller changes to the concentrations of individual metabolites. It appears that maintenance of stable levels of metabolite concentrations (**homeostasis**, from the Greek *homoios*, same; *stasis*, standing still) is a key feature of metabolic regulation, as indeed it has long been noted to be in the hormonal control of blood glucose or pH. Homeostasis also appears to be a key feature in the maintenance of high concentrations of macromolecules in cells experiencing changes in their

Box 1.2 Metabolic control analysis

Metabolic control analysis posits that each step in a metabolic pathway contributes in some degree to the control of flux through the pathway. Each enzyme is assigned a *flux control coefficient*, and the sum of the control coefficients for all the enzymes in the pathway must come to 1.0. The higher its flux control coefficient, the more significant will be the part played by a given enzyme in the control process. Likewise, for each metabolite there is a concentration control coefficient derived from the effects that each enzyme in the pathway has on its concentration. The 'connectivity theorem' states that the overall sum of the concentration control coefficients should be 0 for each metabolite.

According to theory, a change in the activity of a single enzyme is likely to produce a larger change in the concentrations of metabolites in a pathway than in the flux through the pathway. Moreover, changing the activity of enzymes catalyzing reactions that are at or near equilibrium is unlikely to generate large effects on either metabolite concentrations or metabolic flux. Raising the flux through a metabolic pathway is rarely achieved by controlling the activity of a single enzyme, which would require its flux control efficient to be close to 1.0.

This explains why the overexpression of genes encoding what were thought to be rate-limiting enzymes has been observed to have little or no effect on the flux through the relevant metabolic pathway *in vivo*.

The usual outcome of a signal in a metabolic pathway is an increase or decrease in the flux through the pathway, but this is accompanied by smaller changes in the concentrations of individual metabolites. It appears that maintenance of stable levels of metabolite concentrations is a major factor in metabolic regulation. In turn, it has been argued that feedback inhibition of allosteric enzymes catalyzing supposedly rate-limiting reactions has more to do with metabolite homeostasis than with flux changes. Metabolic control analysis predicts that if the activities of all the enzymes in a pathway are modified by the same numerical factor, the metabolic flux will be modified identically without change to the concentrations of metabolic intermediates. The effect of the feedback inhibition in a pathway of the kind described above will be to lower the flux control coefficient of the target enzyme, which therefore means that control must pass to subsequent steps.

environments (Section 1.8). We are still only in the early stages of a quantitative understanding of the complexity of signaling and the regulation of interconnected systems, but homeostasis is increasingly viewed as an important property.

1.8 MACROMOLECULAR CROWDING

Up to this point, we have described the properties of macromolecular assemblies mainly in terms of their structures and those of their components as determined by X-ray crystallography and other techniques, and their functional attributes as established by biochemical and genetic approaches. In general, the specimens studied are in dilute solution. Most *in vitro* experiments with proteins are carried out at concentrations of <1 mg/ml; for example, enzyme kinetic studies usually use only a few µg per ml; at the upper end, proteins crystallize from solutions at 5–40 mg/ml. *In vivo* the situation is quite different: the interior of a cell is a very crowded environment (Figure 1.35). The total concentration of protein inside an *E. coli* cell is 200–300 mg/ml and that of RNA is 75–150 mg/ml, giving a total macromolecular concentration of ~300–400 mg/ml. Eukaryotic cells are similarly congested, and this congestion has major consequences for biological activity within a cell.

In crowded situations, the concentrations of macromolecules can be so high that they account for a significant fraction of the total volume, which is thus unavailable to other macromolecules. This is termed **macromolecular crowding** (or sometimes, the **excluded volume effect**). The concentrations given above imply that 25–30% of the volume in living cells is excluded. For comparison, the densest possible packing of uniform spheres (cubic or hexagonal close packing) occupies ~74% of total volume, as shown by Carl Friedrich Gauss in 1831. With random packing, this number falls to ~64%. In protein crystals, the fractional volume occupied by proteins is typically ~55%, but values as low as ~25% are observed. It appears that lattice contacts in protein crystals do not alter the overall conformations of the molecules, although they may cause some local perturbations. On this basis, we assume that the effects of molecular crowding are not due to any major perturbation of protein structure.

Advances in mass spectrometry have made it possible to examine the concentrations of individual proteins in cells. For example, in the pathogenic spirochete *Leptospira interrogans*, with a total protein concentration of ~250 mg/ml, more than 2200 proteins were identified; their abundances ranged from 40 or fewer to 40,000 copies per cell. (A cellular quota of 40,000 copies of a 25 kDa protein corresponds to a concentration of ~20 µg/ml.) Proteins



Figure 1.35 Dense packing of macromolecules and organelles in a eukaryotic cell. An electron tomogram is shown for a 400 nm-thick section of a cultured rat cell. The specimen was prepared by high-pressure freezing, freeze substitution, and plastic embedding. A tilt series of 80 micrographs was collected and used to calculate the tomogram, which was then 'segmented' by an observer who manually traced out the various subcellular components in each slice. The segmented tomogram is viewed *en face* from one side. The plough-shaped feature is the Golgi apparatus with seven cisternae. Color coding: yellow, endoplasmic reticulum; blue, membrane-bound ribosomes; orange, free ribosomes; bright green, microtubules; bright blue, dense core vesicles; white, clathrin-negative vesicles; bright red, clathrin-positive compartments and vesicles; purple, clathrin-negative compartments and vesicles; dark green, mitochondria. (From B.J. Marsh, *Biochim. Biophys. Acta* 1744:273–292, 2005. With permission from Elsevier.)

involved in protein synthesis and quality control and proteins needed for motility were the most abundant. Other studies indicate that similar protein concentrations and abundances are widespread in prokaryotes. In response to changes in conditions, protein compositions can change. For example, when the spirochete is exposed to an antibiotic, it responds by expressing the genes for a small number of normally unrepresented proteins (of unknown function), but its total protein concentration remains constant. This homeostatic behavior resembles the response of cells in keeping concentrations of individual metabolites constant even when the flux through a metabolic pathway changes (Section 1.7).

Molecular crowding affects reaction rates, protein folding, assembly, and stability

The effective concentration of a macromolecule in any reaction (a property defined as its **thermodynamic activity**) is higher than its actual physical concentration. The **activity coefficient** is defined as the ratio of the two concentrations. At low concentrations, as in most experiments *in vitro*, the activity coefficient has a value very close to 1, but at high concentrations (crowding), it is well above 1. This has several important kinetic and thermodynamic consequences, among them effects on diffusion rates, reaction rates, and assembly (association or dissociation).

Studies of bacterial DNA replication were among the first *in vitro* experiments to be carried out under crowding conditions, when it was discovered in Arthur Kornberg's laboratory that it was necessary to add a synthetic polymer, polyethylene glycol (PEG), to the reaction mixture for DNA synthesis to occur. For reasons described below, this was probably due to PEG-induced molecular crowding, promoting the assembly of a particular complex at the origin of replication. It is important to bear such considerations in mind when attempting to relate the properties of a macromolecular system studied *in vitro* to the situation *in vivo*.

Crowding promotes the binding of macromolecules to one another, but specificity in these interactions depends on their structures. When macromolecules associate, there is a reduction in excluded volume and the association constants for macromolecular complexes are increased in comparison with their values in dilute solution. The degree of these changes depends on the sizes, shapes, and binding affinities of the macromolecules.

If another molecule is put into a solution of a globular protein at 300 mg/ml, its activity coefficient increases appreciably for molecules above 1 kDa and steeply for those above 10 kDa. For example, it has been estimated that for a 40 kDa protein in a monomer–dimer equilibrium, the association constant at the overall protein concentration found inside an *E. coli* cell would be 8–40 times higher than in dilute solution. If it were a monomer–tetramer equilibrium, the association constant would increase by a factor of 10^3 – 10^5 . Thus, it appears that crowding effects are essentially confined to macromolecules, and small molecules are largely unaffected.

In intracellular protein folding, crowding favors the condensation of newly synthesized polypeptide chains into molten globule-like conformations (Section 1.3); in parallel, however, crowding also enhances the tendency of partly folded chains to come together in aberrant aggregates. Macromolecular machines called *chaperones* safeguard nascent polypeptide chains by sequestering them during the folding process (Section 6.3).

The increase in activity coefficients at high concentrations of reactants is not linear with respect to concentration. For example, the activity coefficient of hemoglobin at 200 mg/ml is ~10, and at 300 mg/ml it is ~100. The high concentration of hemoglobin in the red blood cell (~350 mg/ml) underlies the disastrous pathology of sickle-cell disease (**Box 1.3**). In contrast, the crystallin proteins of the eye lens, essential for transparency, do not aggregate to the point of scattering light, which would impair transparency, despite total protein concentrations of >500 mg/ml. Crystallins remain unreplaceable throughout life, and their extraordinary stability and durability have been attributed to their high concentration.

Macromolecular crowding affects diffusion rates

Diffusion of macromolecules in cells generally follows Fick's laws (**Box 1.4**). The high concentration of protein might be expected to reduce mobility for all sorts of molecules. However, data obtained by single-particle tracking (SPT) using selectively labeled proteins or by fluorescence recovery after photobleaching (FRAP—the return of fluorescent protein

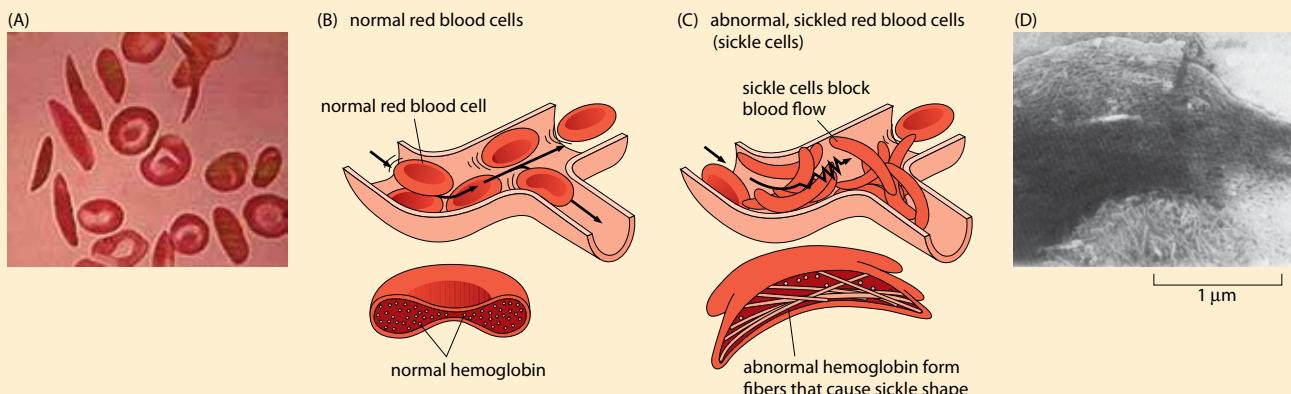
Box 1.3 Sickle-cell anemia and aberrant polymerization of a mutant hemoglobin


Figure 1.3.1 Capillary blocking by red blood cells distorted by fibers of a mutant hemoglobin. (A) Scanning electron micrograph of normal erythrocytes (flattened concave disks) and sickled (elongated crescent) erythrocytes. (B,C) Schematic diagrams of (B) normal erythrocytes which are pliable and pass easily along narrow capillaries, and (C) sickled erythrocytes, which are stiffer and block capillaries. (D)

Electron micrograph of a lysed sickled erythrocyte, negatively stained, showing fibers spilling out of the cell. (A, courtesy of A. Cavanagh and the EM Unit, Royal Free Hospital School of Medicine/Wellcome Trust Photo Library; B and C, adapted, courtesy of NIH Gov. NHLBI Diseases; D, from R. Josephs, H.S. Jarosch and S.J. Edelstein, *J. Mol. Biol.* 102:409–426, 1976. With permission from Elsevier.)

Hemoglobin (Hb) is an $\alpha_2\beta_2$ tetramer, whose α and β subunits (each 17 kDa) are homologous with similar folds. Hb serves as an oxygen carrier in red blood cells (erythrocytes). In humans, the mature erythrocyte has no nucleus or other organelle; its cytosol is essentially a highly concentrated solution of Hb.

Sickle-cell anemia is a debilitating and often fatal disease, in which the erythrocytes are deformed ('sickled'; see **Figure 1.3.1**). Sickled erythrocytes are fragile, and their rupture causes hemolytic anemia. They are also more rigid than normal erythrocytes and obstruct capillaries, restricting blood flow, causing pain and other symptoms. In 1945, Linus Pauling proposed that sickling is caused by a mutation in the Hb gene. In 1956, Vernon Ingram showed that there was a Glu \rightarrow Val substitution at position 6 in the β subunit of sickle-cell Hb. This was the first instance of a disease being attributed to an amino acid replacement in a protein, a discovery that marked the birth of 'molecular medicine'.

Sickling is due to the deoxy-form of hemoglobin S (HbS) aggregating into filaments whose growth distorts the erythrocyte into an elongated shape (**Figure 1.3.2**). The hydrophobic side chain of Val6 protruding from one HbS tetramer can fit into a hydrophobic pocket on another tetramer. Repetition of this interaction generates paired strands of tetramers in which the two strands are half-staggered and related by a 2-fold screw axis. Seven such pairs twist about a common axis to make 14-stranded fibers with a diameter of ~24 nm and a mean helical pitch of ~270 nm. In wild-type Hb, the charged Glu side chain at position

$\beta 6$ prevents any such interaction, and the surface pocket is absent in oxy-Hb or oxy-HbS because of the conformational change that accompanies oxygenation. However, when the erythrocyte enters a capillary, the Hb becomes deoxygenated and the surface pocket forms. This allows the deoxy-HbS tetramers to fibrillize, causing sickling, and the pathological symptoms ensue. Oxygenation causes the HbS fibers to redissolve. Much effort has gone into seeking treatments that can inhibit fiber formation, thus far without major success.

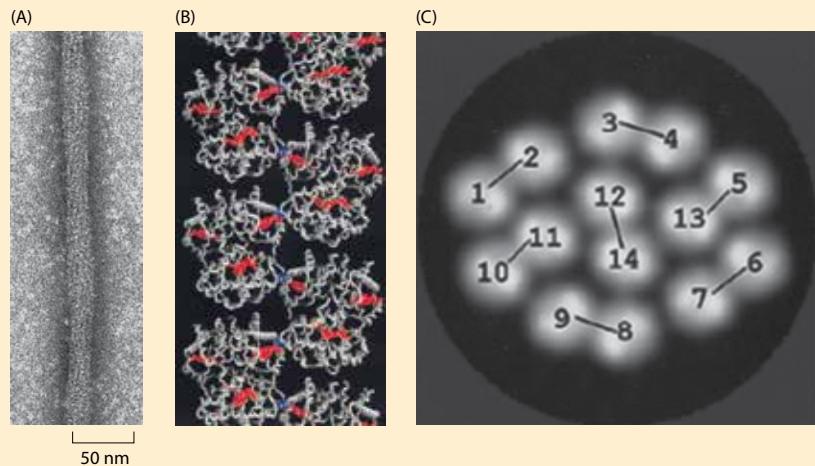


Figure 1.3.2 Molecular packing in HbS fibers. (A) Electron micrograph of a negatively stained fiber. (B) Two-stranded fibers of HbS tetramers are present in the crystal structure. The backbones of the HbS subunits are in white; the heme groups in red, and the mutant β Val6 residues in blue. (PDB 2HBS) (C) Cross section of a fiber model reconstructed from the electron micrographs, showing the 14 strands (numbered 1–14) of HbS tetramers. The strands are paired as shown by the connecting lines. The fiber is elliptical in cross section, ~24 nm in diameter on the major axis. (A and C, from Z. Wang et al., *J. Struct. Biol.* 131:197–209, 2000. With permission from Elsevier; B, from D.J. Harrington, K. Adachi and W.E. Royer, *J. Mol. Biol.* 272:398–407, 1997. With permission from Elsevier.)

Box 1.4 Diffusion: Fick's laws

Molecules in solution are in constant motion, a random process called normal **diffusion**, synonymous with Brownian motion. The primary determinants of diffusion are the size and shape of the molecule and the viscosity of the solvent. If there is a concentration gradient, the molecules will diffuse from regions of higher concentration to lower concentration (in other words, their motion is entropy-driven).

Fick's First Law states that the net flux of diffusing substance, J ($\text{mol}/(\text{m}^2 \text{s})$), in one dimension is given by

$$J = -Dd\phi/dx$$

where D is the diffusion coefficient (in m^2/s), ϕ is the local concentration of the diffusing molecule (in mol/m^3), and x is distance (in m).

In three dimensions, this becomes

$$\mathbf{J} = -D\nabla\phi$$

where \mathbf{J} is the flux vector and ∇ represents the vectorial derivative.

Fick's Second Law gives the change in concentration gradient with time:

$$d\phi/dt = Dd^2\phi/dx^2$$

or, in three dimensions,

$$\partial\phi/\partial t = D\nabla^2\phi$$

Fick's laws were derived in 1855. In 1851, Stokes put forward his analysis of a small sphere moving with velocity v through a viscous medium in laminar flow (smooth flow without turbulence). He derived the equation

$$F = -6\pi\eta rv,$$

where F is the frictional force (viscous drag), η is the viscosity, and r is the radius of the sphere.

In 1905, Einstein analyzed Brownian motion in terms of the kinetic theory of heat and showed that, for a particle under such conditions,

$$D = RT/Nf_0$$

where R is the universal gas constant, T is the absolute temperature, N is Avogadro's number and f_0 is the solvent friction coefficient. This simplifies to

$$D = kT/f_0$$

where k = Boltzmann's constant. f_0 for a small spherical particle is, as Stokes had formulated, $6\pi\eta r$.

Biological macromolecules are not perfect spheres. This causes departures from the Einstein theory, and the value of the frictional coefficient, f , is greater than that of f_0 for the equivalent sphere. The extent to which the **frictional ratio**, f/f_0 , is greater than 1.0 gives an estimate of how well a macromolecule can be modeled instead as, say, a prolate ellipsoid or a cylinder. For *tobacco mosaic virus*, f/f_0 has a value close to 2.0, reflecting its rod-like structure (axial ratio $\sim 17:1$). However, if a macromolecule or assembly has a high value of f/f_0 but electron microscopy or some other method of shape determination indicates that it is roughly spherical, it follows that the macromolecule or assembly carries unusually large amounts of bound water. The 2-oxo acid dehydrogenase multienzyme complexes described in Section 9.4 are a good example of this. For most globular proteins, the amount of bound water is $\sim 0.35 \text{ g/g protein}$, and the value of f/f_0 is typically ~ 1.3 , reflecting both the departure from perfect sphericity and their typical hydration.

At the macroscopic level, the importance of water is becoming apparent in some surprising ways. Given the strong crowding effects of intracellular proteins, small changes in cellular water cause large changes in activity coefficients. In older people, there is a significant loss of cellular and tissue water; a diminution in the occupiable volume available to proteins may contribute to the age-related onset of protein misfolding and aggregation diseases.

On the basis of Fick's laws and Einstein's work, the root mean square displacement (RMSD) $\langle r \rangle$ of a solute particle in time t can be calculated from the equation

$$\langle r \rangle = \sqrt{(6Dt)}$$

For diffusion in one or two dimensions, the factor 6 is replaced by 2 or 4 respectively.

into a volume bleached by a laser flash) show only modest slowing (threefold to fourfold) for small solutes in the eukaryotic cytosol compared with their rates in water. With larger proteins, the effects are more pronounced: hemoglobin (68 kDa) in an erythrocyte diffuses six times more slowly than in dilute solution, and green fluorescent protein (GFP, 27 kDa) recombinantly expressed in the *E. coli* cytoplasm is slowed 11-fold. In a typical FRAP experiment, the bacterial chemotaxis signaling protein CheY (14 kDa; Section 14.12) fused with a fluorescent protein yielded a diffusion coefficient of $2 \mu\text{m}^2/\text{s}$ (Figure 1.36). Other measurements using fluorescent correlation spectroscopy gave a value of $4.6 \mu\text{m}^2/\text{s}$ and the diffusion coefficient of CheY alone has been estimated at $10 \mu\text{m}^2/\text{s}$. Thus, it would take $\sim 50 \text{ ms}$ ($r^2/2D$; Box 1.4) for a CheY molecule to diffuse $1 \mu\text{m}$ from a pole of an *E. coli* cell to the flagellar motors along the sides of the cell (see Figure 14.73). This is consistent with the timescale of the chemotactic response.

Figure 1.36 Time course of protein diffusion in an *E. coli* cell. Simulation *in silico* of the return of fluorescence attributed to CheY tagged with yellow fluorescent protein (CheY-YFP) after photobleaching. The results were fitted to a single exponential curve on the assumption that the labeled molecules were homogeneous and in an isotropic environment, yielding a diffusion coefficient of $2 \mu\text{m}^2/\text{s}$. (From M.A. DePristo et al., *Prog. Biophys. Mol. Biol.* 100:25–32, 2009. With permission from Elsevier.)

In a cell, the central assumption of a continuous fluid of uniform viscosity across the cytoplasm or within an organelle does not normally hold. High concentrations of solutes and the presence of cytoskeletal elements, membranes, or other organelles affect the process, turning it into **anomalous diffusion**. Nevertheless, in many instances, solutes do observe the equations of normal diffusion.

Other factors may come into play; for example, the cytoskeleton restricts diffusion. For some complexes (such as dextrans above 500 kDa) or for linear DNAs, diffusion through the eukaryotic cytosol has been found to be greatly slowed. Within organelles, effects on diffusion may be smaller than theoretical calculations have suggested. Thus in the mitochondrial matrix, the diffusion of GFP is slowed down only about 3-fold, which has been attributed to the clustering of matrix enzymes into membrane-associated complexes (Chapter 15), thereby increasing the space through which other proteins can diffuse. However, in all cells and compartments, any diffusion-limited process stands to be affected; and the bigger the molecule, the larger the effect.

Models of crowded intracellular environments can now be built

It is now possible, conceptually and technically, to perform realistic computational modeling of the densely crowded cellular interior. For example, a simulation has been performed of the *E. coli* cytoplasm, based on 100 macromolecules representing 50 of the most abundant constituents of *E. coli* (45 proteins, the 50S and 30S ribosomal subunits and three typical tRNAs) plus 8 molecules of GFP, at a total concentration of 275 mg/ml. The structures of the chosen molecules were taken from the Protein Data Bank, and their relative abundances were derived from proteome measurements. Together they account for ~85% by weight of the cytoplasmic proteome. Their diffusion was calculated under conditions of Brownian motion, allowing for steric exclusion and electrostatic and short-range hydrophobic interactions. The process was calibrated by monitoring the GFP molecules, for which estimates of the translational diffusion coefficient ($7.7\text{--}9.0 \mu\text{m}^2/\text{s}$) *in vivo* had been obtained by FRAP. In a 15 μs simulation, the GFP molecules were found to move, on average, about six molecular diameters (320 \AA), encountering ~80 different neighbors in doing so. A still image is shown at Figure 1.37. Larger molecules moved more slowly, a ribosome hardly moving at all but still experiencing encounters with many more mobile neighbors.

1.9 CELLULAR COMPARTMENTATION AND EVOLUTION

As long ago as the 1930s, Frederick Gowland Hopkins made clear that it is wrong to perceive the cell as “just a bag of enzymes” but it is only now that the complexities are becoming recognized and, to some extent, amenable to investigation. The cell interior is not a homogeneous medium in which macromolecules are free to diffuse: in eukaryotic cells, much of the cytosol is occupied by organelles and crisscrossed by the filamentous network of the cytoskeleton. This leads to the confinement of many cytosolic macromolecules. Some attach (by electrostatics) to surfaces of organelles or the cytoskeleton, and others are anchored (by hydrophobic interactions) in membranes. Even in a prokaryotic cell, there exist microenvironments that differ significantly from the general cytosol.

Each kind of organelle represents a chemically distinct environment populated by a particular set of macromolecules. There is nevertheless communication and exchange of materials between organelles. Proteins are conveyed in vesicles that bleb off the donor compartment and fuse into the membrane that encloses the acceptor compartment. Some of these vesicles have protein coats, and transport is effected mainly by motor proteins traveling along cytoskeletal filaments.

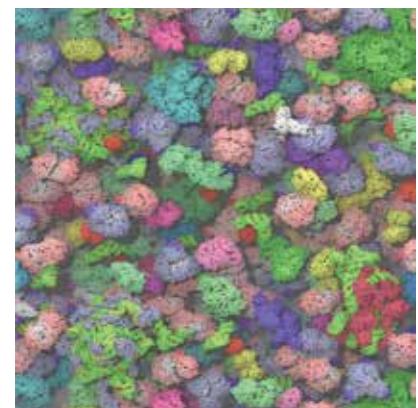
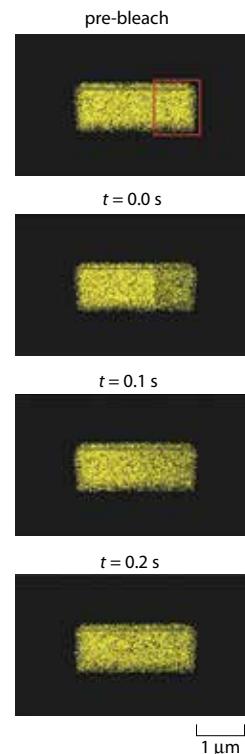


Figure 1.37 A dynamic model of the *E. coli* cytoplasm. Shown is a rendering of the model cytoplasm after a simulation of Brownian motion. RNA is depicted in green and yellow. (From S.R. McGuffee and A.H. Elcock, *PLoS Comp. Biol.* 6:e1000694, 2010.)

All cells belong to one of the three Urkingdoms: Archaea, Bacteria, Eukarya

Until the 1980s, organisms were classified as either **prokaryotes** or **eukaryotes**, a division based on morphological criteria. The defining feature of eukaryotic cells was the presence of a nucleus housing the genetic material; in addition, they possessed other organelles (membrane-enclosed compartments) (Box 1.5) that provide the chemical environments required for specific reactions or separate concurrent metabolic processes. Prokaryotes, which were mostly unicellular and small (1–5 µm), lacked a nucleus; intracellular organelles were found in only a few species such as photosynthetic bacteria. These defining properties remain in effect today except that prokaryotes are now recognized as falling into two kingdoms, bacteria and archaea, as discussed below.

The fundamental similarities at the molecular level shared by all present-day organisms make it likely that life as we know it has emerged on Earth only once. Most probably, all extant organisms are descended from a **progenote**, a hypothetical entity with an organizational level simpler than that of prokaryotic cells (Figure 1.38). Given that the subcellular organization of even the simplest prokaryotes is already quite complex, there must have been simpler intermediates such as self-replicating RNA-like molecules contained in semi-permeable vesicular structures from which the ability to assemble amino acids into proteins subsequently evolved. At the level of the progenote, ‘genes’ and the proteins encoded by them were probably small and the accuracy of transcription and translation was probably low. The progenote is distinct from and precedes the **last universal common ancestor** (LUCA), which lies at the base of phylogenetic trees and is thought to have been a prokaryote-like cell (see below).

The idea of a Tree of Life goes back to Charles Darwin (Figure 1.39), but nowadays phylogenetic trees are based on molecular criteria rather than comparative anatomy or morphology. Protein or gene sequences and, increasingly, whole genomes are used to construct the trees (Figure 1.40). In the 1970s, it started to become apparent that prokaryotes can be partitioned into two fundamentally different life forms. In particular, a comparison of ribosomal RNA sequences indicated the existence of two distinct groups, archaea and bacteria. Carl Woese proposed that all organisms belong to one of three kingdoms, later renamed **Archaea**, **Bacteria**, and **Eukarya** (see Figure 1.38). This reclassification was challenging but is now supported by multiple lines of evidence. Nevertheless, it is still useful in some contexts to retain the term ‘prokaryote’ for non-eukaryotes; in other words, archaea or bacteria.

Bacteria have an open compartment, the nucleoid, and a membrane-delimited compartment, the periplasm

The absence of organelles in most prokaryotes does not mean that their cytosol lacks organization. The genome, usually in the form of a single circular chromosome, is confined to a region, the **nucleoid**, which can be detected by light microscopy using DNA-specific stains. Bacteria belonging to the phylum *Planctomycetes* are an exception in that they have membrane-enclosed compartments; their nucleoid is enclosed by a membrane and is therefore analogous to the eukaryotic nucleus. Moreover, some bacteria possess microcompartments enclosed by polyhedral protein shells that contain specific enzymes; an example in cyanobacteria is the **carboxysome**, which houses the CO₂-fixing enzyme RuBisCo (ribulose-1,5-bisphosphate carboxylase).

In addition to membrane-enclosed compartments, cells have ‘open’ compartments, regions where certain components congregate. The bacterial nucleoid affords one example. (Eukaryotes also have open compartments such as the *nucleolus*—a site for ribosome synthesis within the nucleus.) Open compartments are likely to have some sort of scaffold. In the cytosol, this role is fulfilled in at least some cases by the cytoskeleton. For a long time it was held that prokaryotes do not have a cytoskeleton, but recently genomic sequences, molecular structures, and advanced imaging have shown that homologs or analogs of the major eukaryotic cytoskeletal proteins are found in prokaryotes. Gram-negative bacteria also have a clearly defined compartment distinct from their cytosol, the **periplasm** or **periplasmic space** situated between the plasma membrane (also called the inner membrane) and the outer membrane.

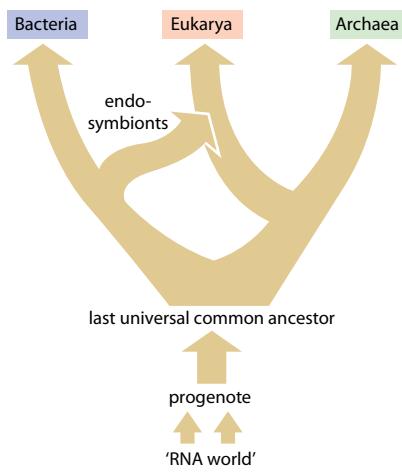


Figure 1.38 A likely scenario for the early evolution of cellular life. After the RNA world, DNA gradually replaced RNA for information storage, and proteins supplanted RNA in most aspects of biocatalysis. The LUCA is held to have been a prokaryote, whereas the progenote represents an earlier and simpler life form. (Adapted from J.P. Gogarten and L. Oelundzenksi, in Encyclopedia of Molecular Biology [T. Creighton, ed.]. Wiley On-line Library.)

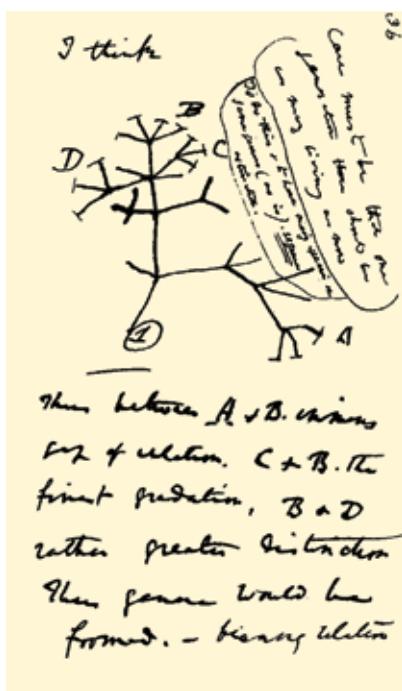


Figure 1.39 Darwin's tree of life. Annotated sketch from his notebook B (1837). Right bubble: “Case must be that in one generation then, there should be as many living as now”. Second bubble: “To do this & to have many species in the same genus requires extinction.” Bottom: “Thus between A & B, immense gap of relation. C & B, the finest gradation. B & D, rather greater distinction. Thus genera would be formed – bearing relation”. (Transcripts courtesy of the American Museum of Natural History.)

Box 1.5 Major compartments in eukaryotic cells

Chloroplasts: Organelles found in plant cells and some algae, enclosed by a double membrane. Their interior (*stroma*) houses a complex system of stacked *thylakoid membranes*, known as *grana*. Chloroplasts, ~5 µm across, are the site of photosynthesis.

Cytosol: Main part of the cytoplasm, excluding the nucleus and other organelles. Typically it constitutes about 50% of the cell volume. It is the major site of metabolic activity.

Endoplasmic reticulum (ER): Convoluted compartment where lipids, secretory proteins, and membrane proteins are synthesized. Part of it, the rough ER, is densely studded with ribosomes on the cytosolic surface. Together with the cisternae of the Golgi apparatus, it occupies 15–20% of the cell volume.

Endosomes: Organelles containing materials newly ingested into the cell (by **endocytosis**) and destined to be fused into lysosomes where their protein cargo is degraded.

Golgi apparatus (Golgi complex): A stacked system of flattened vesicles (cisternae); the part proximal to the ER is the *cis* face, and the part distal to it is the *trans* face. Here, proteins from the ER are modified (for example, by glycosylation) and sorted for secretion from the cell or delivery to other compartments.

Lysosomes: Vesicular structures containing enzymes that degrade foreign materials taken into the cell by endocytosis. These enzymes are fully active only at the acidic pH (~2) within a lysosome.

Mitochondria: Organelles 1–2 µm long by ~0.5 µm wide. The inner mitochondrial membrane has deep infoldings called **cristae**. Mitochondria occupy up to 20% of the cellular volume. In aerobic metabolism, they produce most of the ATP.

Nucleus: Surrounded by the nuclear envelope, a double membrane, the nucleus occupies 5–10% of the cell volume. The outer nuclear membrane is continuous with the ER.

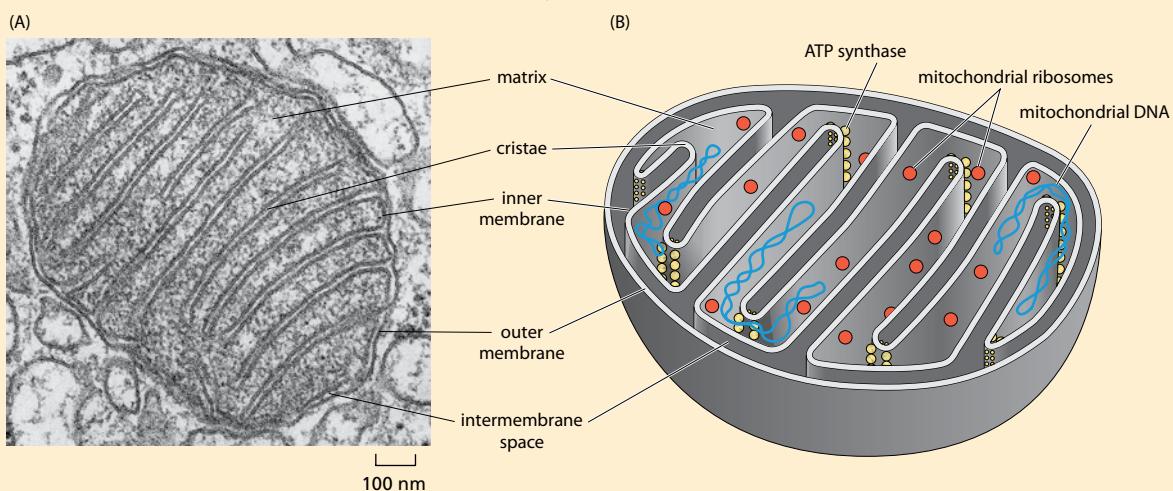
Peroxisomes: These are small vesicles harboring enzymes that carry out oxidative reactions. Some use molecular oxygen to oxidize organic substrates, and this generates toxic hydrogen peroxide (H_2O_2), which is decomposed into H_2O and O_2 by catalase.

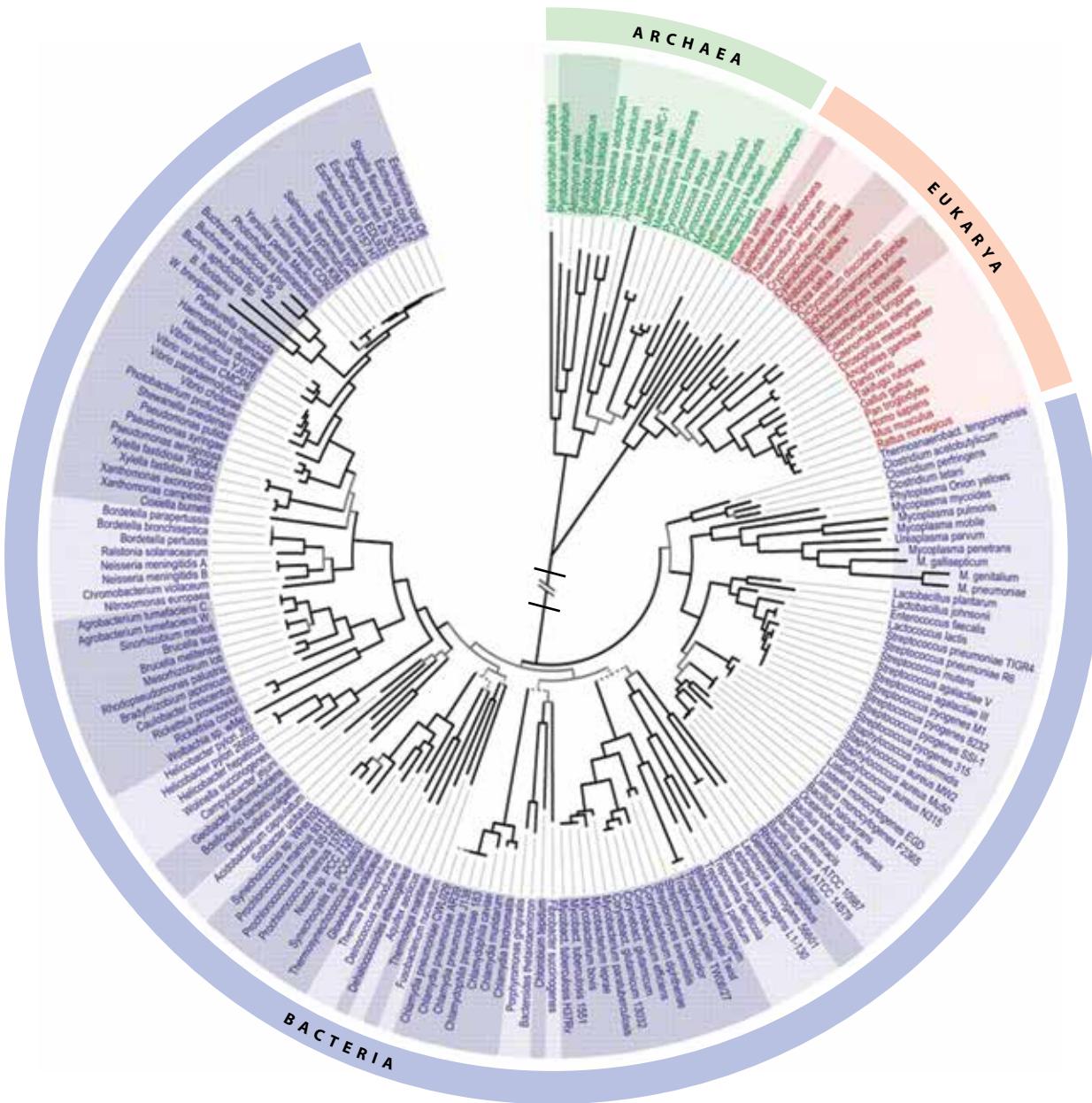
Mitochondria and chloroplasts are of particular relevance to this book through their roles in generating the energy

needed to support life. The main properties of mitochondria are summarized here; chloroplasts are covered in Chapter 15. Eukaryotic cells typically contain 1000–2000 mitochondria, which are delimited by an outer and an inner membrane. The inner membrane encloses a compartment called the **mitochondrial matrix** (Figure 1.5.1). Its surface area is enlarged by the cristae, whose number correlates with the respiratory intensity of the cell or tissue; for example, there are many more in heart mitochondria than in those of skeletal muscle. The sides of the inner membrane facing the matrix and inter-membrane space are termed the N and P sides respectively, because the **membrane potential** is negative on the inside (matrix) and positive on the outside (inter-membrane space). The inner membrane is essentially impermeable to ions and polar molecules of all kinds but allows the passage of O_2 and CO_2 . In contrast, the outer membrane is largely permeable to ions and small molecules (<5–10 kDa) because it contains protein pores, *mitochondrial porins*, among them the voltage-dependent anion channel (VDAC).

A mitochondrion contains from one to many copies of a circular DNA, which varies widely in size between species and encodes a variety of RNAs (tRNAs, rRNAs, etc.) and proteins. The existence of circular mitochondrial DNA and the presence of bacterial-type ribosomes, together with many other similarities to bacterial cells, have led to the *endosymbiont theory* (see the main text) that mitochondria and chloroplasts arose from the engulfment of bacteria by primordial eukaryotic cells.

Figure 1.5.1 The structure of a mitochondrion. (A) Electron micrograph of a thin section of a mitochondrion. (B) Schematic diagram of a mitochondrion, showing the different compartments contained within the outer membrane. The respiratory chain complexes and ATP synthases (Chapter 15) are arranged vectorially in the inner membrane. The ATP synthases span the inner membrane and cluster as dimers (depicted as yellow spheres) along the sharp folds in the cristae. The matrix contains many enzymes of oxidative metabolism, including those involved in the citric acid cycle and fatty acid degradation (Chapter 9), plus ribosomes and molecules of mitochondrial DNA. (A, from B. Alberts et al., Molecular Biology of the Cell, 6th ed. New York: Garland Science, 2015.)





Archaea more closely resemble eukaryotes than bacteria in some key features

The first archaea to be described happened to inhabit extreme environments such as hot springs or salt lakes, and so they were perceived as **extremophiles**. However, it is now known that the great majority of archaeal species inhabit moderate environments such as the oceans or the soil. It has been estimated that archaea account for ~40% of the Earth's biomass and are major players in its carbon and nitrogen cycles. Our understanding of their physiology and their contribution to the planet's global ecology is still rudimentary: only a tiny fraction of known archaeal species have been isolated and cultivated. Some archaea, in particular methanogens, inhabit the guts of humans and ruminants, but there are no examples yet of archaeal pathogens or parasites.

In archaea, and to a lesser extent in bacteria, a variety of metabolic pathways enables them to exploit a wide range of energy sources. In phylogenetic trees constructed by comparing metabolic enzymes, the division between archaea and bacteria is not well defined; this appears to be a consequence of numerous gene transfer events between these two kingdoms. However, the machinery of gene expression (DNA replication, transcription and

Figure 1.40 A phylogenetic tree derived from the genomes of fully sequenced organisms. The representatives of Archaea are shown in green, Eukarya in red, and Bacteria in blue. (Adapted from F.D. Cicerelli et al., *Science* 311:1283–1297, 2006. With permission from American Association for the Advancement of Science.)

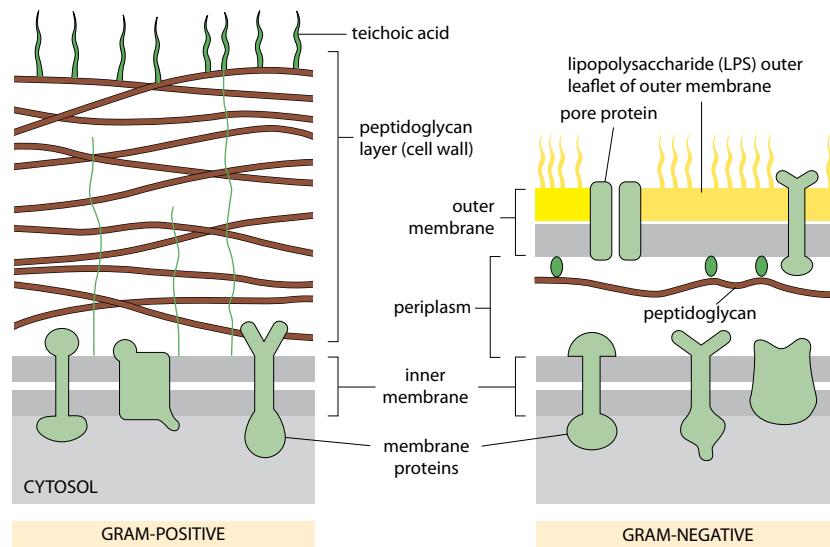


Figure 1.41 The cell envelope architectures of Gram-positive and Gram-negative bacteria. The dye introduced by H.C. Gram in 1864 binds to the thick peptidoglycan layers of Gram-positive bacteria, producing a purple color. Gram-negative bacteria, with less peptidoglycan, bind less stain and are colored pink. The green tendrils in the Gram-positive cell wall are the polyglyceridephosphate chains of lipoteichoic acid molecules. (Adapted from B. Alberts et al., Molecular Biology of the Cell, 6th ed. New York: Garland Science, 2015.)

translation), although fundamentally similar, exhibits distinct differences. In particular, the replication machinery of archaea resembles that of eukaryotes, whereas archaeal ribosomes superficially resemble bacterial ribosomes but at the level of ribosomal RNA and protein structure are more closely related to eukaryotic ribosomes.

Major differences exist between archaeal and bacterial cell envelopes

Almost all bacteria possess a peptidoglycan layer as part of their cell envelope. In Gram-positive bacteria this layer is thick (20–100 nm) and covers the plasma membrane; in Gram-negative bacteria it is thin (<10 nm) and located in the periplasm (Figure 1.41). **Peptidoglycan**, also called **murein**, is a sheetlike polymer formed by linear chains of two amino sugars, *N*-acetyl glucosamine and *N*-acetyl muramic acid, linked by a β -(1,4) glycosidic bond and cross-linked by short (4–5 aa) peptides. Archaeal cells lack peptidoglycan layers, but a few species have an analog called pseudopeptidoglycan or pseudomurein (Figure 1.42). It is different in chemical composition (*N*-acetylglucosamine and *N*-acetyltaulosaminuronic acid linked by a β -(1,3) glycosidic bond), but otherwise resembles bacterial peptidoglycans. In many archaea, the plasma membrane is reinforced and protected by regularly arrayed (glyco)proteins forming surface layers (S-layers) (Figure 1.43 and Figure 1.44). Some S-layers have been implicated in determining cell shape. S-layers unrelated in primary structure have also been found in some bacteria, most of them Gram-positive.

Whereas the membranes of bacteria and eukarya are composed mainly of glycerol-ester lipids, archaeal membranes are made of glycerol-ether lipids. Typically, lipids are amphipathic with a hydrophilic ‘head group’ and hydrophobic ‘tails.’ In eukaryotic and bacterial phospholipids, a tail is made of long fatty acid chains, but archaeal hydrophobic tails

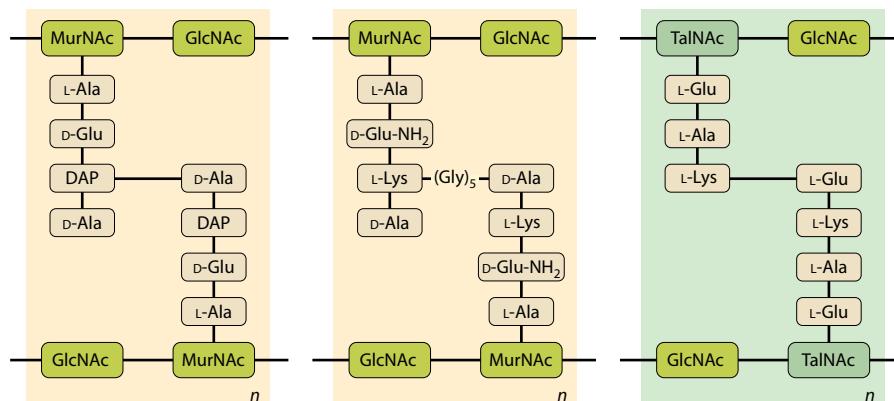


Figure 1.42 Cross-linking patterns of bacterial murein and archaeal pseudomurein. In peptidoglycan (murein), the sugar derivatives *N*-acetylmuramic acid (MurNAc) and *N*-acetylglucosamine (GlcNAc) form a linear backbone cross-linked by short peptides. *n* denotes the number of units in the backbone (many). Some archaea have an evolutionarily unrelated cell envelope component referred to as pseudopeptidoglycan or pseudomurein (right). Here GlcNAc and acetyltaulosaminuronic acid (TalNAc) alternate; unlike murein, pseudomurein lacks D-amino acids.

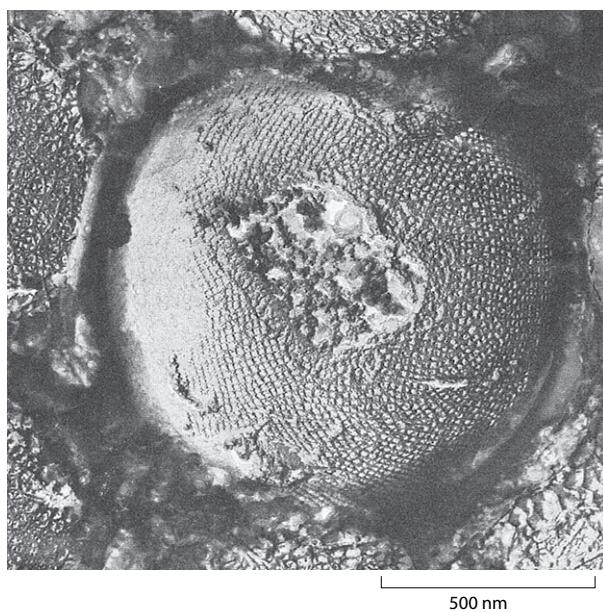


Figure 1.43 Tetragonal S-layer of the archaeon *Desulfurococcus mobilis*.

Electron micrograph of the outer surface of a cell exposed by freeze-etching and contrasted by applying a thin layer of evaporated metal. The cell is covered with a loosely ordered tetragonal lattice with a periodicity of 18 nm. (From I. Wildhaber, U. Santarius and W. Baumeister, *J. Bacteriol.* 169:5563–5568, 1987. With permission from American Society for Microbiology.)

are isoprenoid in nature, often containing methyl branches or cyclopentane rings. Some archaea, in particular extreme thermophiles, contain tetraether lipids that have two polar headgroups with their 2,3-*sn*-glycerol moieties covalently linked by the C₄₀ isoprenoid chains. These tetraether lipids can form membranous monolayers. The presence of tetraether lipids and of isoprenoid tails, often with cyclopropane or cyclohexane rings and ether bonds replacing ester bonds, are all thought to increase membrane stability, enabling the organisms to tolerate extremes of temperature, salt concentration, or pH (Figure 1.45).

Eukaryotic cell organelles probably arose by engulfing bacteria

Most scenarios for the origin of eukaryotic cells accept the **endosymbiont theory** advocated by Lynn Margulis. This proposes that their organelles arose from the internalization of one organism by another, followed by their symbiotic association (see Figure 1.38). The supporting arguments are strong for mitochondria and chloroplasts: both replicate by a process similar to the binary fission of bacteria; both are surrounded by two membranes; in both cases their ribosomes more closely resemble those of bacteria than those of eukaryotes, particularly for chloroplasts; and both have circular genomes distinct from the linear DNA of nuclear chromosomes. However, their genomes are much smaller than those of bacteria: human mitochondrial DNA (16,569 bp) has only 37 genes (encoding two ribosomal RNAs, 22 tRNAs, and 13 proteins). Chloroplast DNA is generally larger; for example, that of *Arabidopsis* (154,478 bp) has about 120 genes (encoding 4 ribosomal RNAs, 37 tRNAs, and potentially 87 proteins). Most of the present-day mitochondrial and chloroplast proteins are encoded in nuclear DNA, and it appears that, during evolution, many genes were

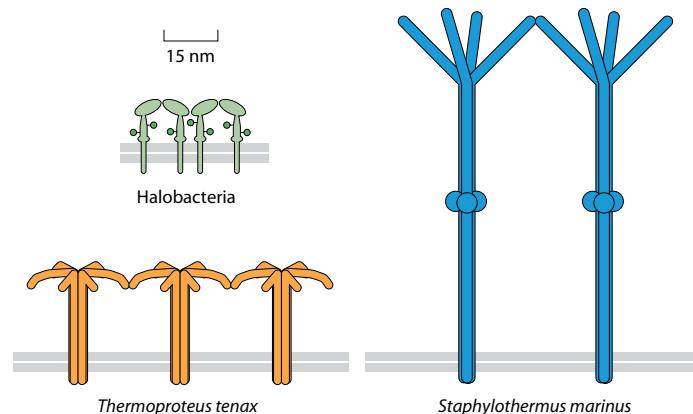


Figure 1.44 Schematic representation of various archaeal S-layers.

The respective repeating elements are drawn to scale. The S-layer proteins or glycoproteins are membrane-anchored and form rooflike structures delimiting a quasi-periplasmic space that may harbor extracellular proteins such as a protease bound to the long stalk of the *Staphylothermus marinus* S-layer protein named tetrabrachion.

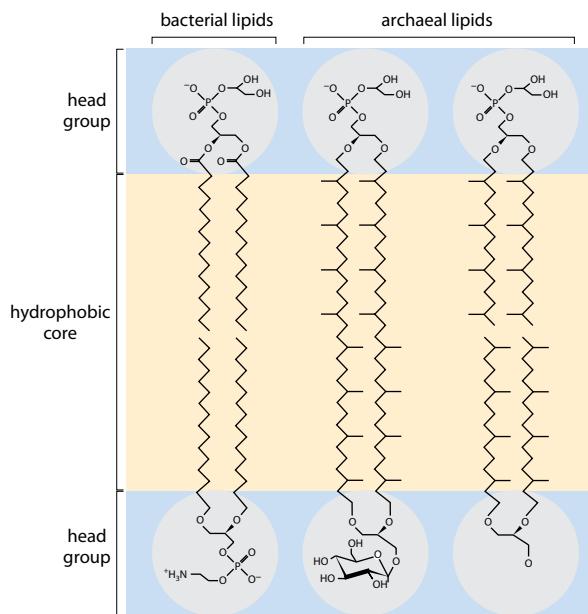


Figure 1.45 Comparison of bacterial and archaeal lipids. In bacteria and eukarya, the glycerol moiety is ester-linked to an *sn*-glycerol 3-phosphate backbone, whereas in archaea the isoprenoid side chains are ether-linked to an *sn*-glycerol 1-phosphate moiety. The bilayer-forming lipids in bacteria are phosphatidylglycerol (upper) and phosphatidylethanolamine (lower). The archaeal lipid can either be a C₄₀ tetraether forming a continuous monolayer (left) or C₂₀ diether lipids forming a bilayer (right). In bacterial lipids, the hydrophobic core is formed by the fatty acid hydrocarbon chains, whereas in archaea it consists of isoprenoids, often with methyl branches. (Adapted from S.-V. Albers and B.H. Meyer, *Nat. Rev. Microbiol.* 9:414–426, 2011. With permission from Macmillan Publishers Ltd.)

transferred from organelle to nucleus and the organellar DNAs shrank in size. These events are thought to have happened more than 1.5 billion years ago, when substantial amounts of oxygen entered the atmosphere from the activity of bacterial cells engaged in photosynthesis (Section 15.3).

Phylogenetic analysis based on genome sequences supports the endosymbiont theory and suggests that the bacterial progenitor of mitochondria was an ancestor of *Rickettsia prowazekii*, which has $>10^6$ bp of DNA and is a cause of lice-borne typhus, and that chloroplasts arose from the engulfment of an O₂-producing photosynthetic bacterium resembling a modern cyanobacterium.

The origin of the first cell with a nucleus remains enigmatic. Possibly, it reflects another fusion event, involving bacterial hosts related to the *Planctomycetes* (see above) and archaea as symbionts. The common ancestor of plants and animals was probably a eukaryotic cell that had mitochondria but not chloroplasts, and it is likely that plants and animals evolved multicellularity independently.

Higher eukaryotes have similar numbers of genes as lower eukaryotes but many more regulatory elements

The genomes of archaea and bacteria are similar in size. The smallest (for example, that of *Mycoplasma genitalium*) encode about 500 proteins, the largest (for example, that of *Bacillus cereus*) about 5000, in the same range as the genomes of some single-celled eukaryotes such as yeast (*S. cerevisiae*). Multicellular higher eukaryotes have between 13,600 genes (the fly *Drosophila melanogaster*) and 28,000 genes (the plant *Arabidopsis thaliana*). The human genome contains about 20,000 protein-coding genes—about the same in number (and with largely similar functions) as those of the nematode worm *C. elegans*. However, this worm has only 10³ cells and $\sim 10^8$ bp in its chromosomal DNA, compared with $\sim 10^{14}$ cells and $\sim 3.2 \times 10^9$ bp for an adult human. In general, the ratio of non-protein-coding (formerly referred to, misleadingly, as ‘junk’ DNA) to protein-coding DNA increases with developmental complexity; it is approximately 3:1 in *C. elegans* and 50:1 in humans. Almost all of the genomes of humans and worms are now known to be transcribed, generating vast numbers of non-protein-coding RNAs, and there is a growing awareness that these form the basis of complex regulatory systems that underlie cellular differentiation and development. Thus the role of RNA is by no means confined to that of a messenger for protein synthesis and serving as components (structural and catalytic) of the ribosome (Section 6.2).

REFERENCES

1.1 Expression of the genetic blueprint

- Ahlquist P (2002) RNA-dependent RNA polymerases, viruses, and RNA silencing. *Science* 296:1270–1273.
- Crick F (1970) Central dogma of molecular biology. *Nature* 227:561–563.
- Kempner ES & Miller JH (2003) The molecular biology of *Euglena gracilis*. XV. Recovery from centrifugation-induced stratification. *Cell Motil Cytoskel* 56:219–224.
- Mattick JS (2011) The central role of RNA in human development and cognition. *FEBS Lett* 585:1600–1616.
- Nandkumar J & Cech T (2013) Finding the end: recruitment of telomerase to telomeres. *Nat Rev Mol Cell Biol* 14:69–82.

1.2 Weak forces and molecular interactions

- Frank RAW, Pratap JV, Pei XY et al (2005) The molecular origins of specificity in the assembly of a multienzyme complex. *Structure* 13:1119–1130.
- Janin J, Bahadur RP & Chakrabarti P (2008) Protein–protein interaction and quaternary structure. *Q Rev Biophys* 41:133–180.
- Jung H-I, Bowden SJ, Cooper A et al (2002) Thermodynamic analysis of the binding of component enzymes in the assembly of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *Protein Sci* 11:1091–1100.
- Kuriyan J, Konforti B & Wemmer D (2012) The Molecules of Life: Physical and Chemical Principles. Garland Science.
- Lee B & Richards FM (1971) The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* 55:379–400.
- McDonald IK & Thornton JM (1994) Satisfying hydrogen bonding potential in proteins. *J Mol Biol* 238:777–793.
- PISA may be accessed at www.ebi.ac.uk/pdbe/pisa/
- Zhang Q, Sanner M & Olson AJ (2008) Shape complementarity of protein–protein complexes at multiple resolutions. *Proteins* 75:453–467.

1.3 Protein folding and stability

- Brockwell DJ & Radford SE (2007) Intermediates: ubiquitous species on folding energy landscapes? *Curr Opin Struct Biol* 17:30–37.
- Dill KA, Ozkan SB, Shell MS et al (2008) The protein folding problem. *Annu Rev Biophys* 37:289–316.
- Dobson CM (2003) Protein folding and misfolding. *Nature* 426:884–890.
- Hartl U & Hayer-Hartl M (2009) Converging concepts of protein folding *in vitro* and *in vivo*. *Nat Struct Mol Biol* 16:564–558.
- Kalia Y, Brocklehurst SM, Hipps DS et al (1993) The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *J Mol Biol* 230:323–341.
- Karshikoff A & Ladenstein R (1998) Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. *Prot Eng* 11:867–872.
- Neuweiler H, Sharpe TD, Rutherford TJ et al (2009) The folding mechanism of BBL: plasticity of transition-state structure observed within a ultrafast folding protein family. *J Mol Biol* 390:1060–1073.
- Pauling L (1960) The Nature of the Chemical Bond. Cornell University Press.
- Privalov PL & Gill SJ (1988) Stability of protein structure and hydrophobic interaction. *Adv Prot Chem* 39:191–234.

1.4 Self-assembly and symmetry

- Cansu S & Doruker P (2008) Dimerization affects collective dynamics of triosephosphate isomerase. *Biochemistry* 47:1358–1368.
- Chan WW-C, Kaiser C, Salvo JM et al (1974) Formation of dissociated enzyme subunits by chemical treatment during renaturation. *J Mol Biol* 87:847–852.

- Engström A & Strandberg B (eds) (1969) Symmetry and Function of Biological Systems at the Macromolecular Level: Nobel Symposium 11. Wiley Interscience.
- Goodsell DS & Olson AJ (2000) Structural symmetry and protein function. *Annu Rev Biophys Biomol Struct* 29:105–153.
- Perham RN (1975) Self-assembly of biological macromolecules. *Phil Trans R Soc Lond B* 272:123–136.
- Perica T, Marsh JA, Sousa FL et al (2012) The emergence of protein complexes: quaternary structure, dynamics and allostery. *Biochem Soc Trans* 40:475–491.

1.5 Macromolecular dynamics

- Baxter RH, Ponomarenko N, Srajer V et al (2004) Time-resolved crystallographic studies of light-induced structural changes in the photosynthetic reaction center. *Proc Natl Acad Sci USA* 101:5982–5987.
- Kay LE (2011) Solution NMR spectroscopy of supra-molecular systems, why bother? A methyl-TROSY view. *J Magn Reson* 210:159–170.
- Khalili-Araghi F, Gumbart J, Wen PC et al (2009) Molecular dynamics simulations of membrane channels and transporters. *Curr Opin Struct Biol* 19:128–137.
- Ma BC, Tsai J, Haliloglu T & Nussinov R (2011) Dynamic allostery: linkers are not merely flexible. *Structure* 19:907–917.
- Perrais D & Merrifield CJ (2005) Dynamics of endocytic vesicle creation. *Dev Cell* 9:581–592.
- Torchia DA (2011) Dynamics of biomolecules from picoseconds to seconds at atomic resolution. *J Magn Reson* 212:1–10.
- Tsuruta H & Irving TC (2008) Experimental approaches for solution X-ray scattering and fiber diffraction. *Curr Opin Struct Biol* 18:601–608.
- Zhou J, Ha KS, La Porta A et al (2011) Applied force provides insight into transcriptional pausing and its modulation by transcription factor NusA. *Mol Cell* 44:635–646.

1.6 Catalysis

- Cornish-Bowden A (2012) Fundamentals of Enzyme Kinetics, 4th ed. Wiley-VCH.
- Glowacki DR, Harvey JN & Mulholland AJ (2012) Taking Ockham's razor to enzyme dynamics and catalysis. *Nat Chem* 4:169–176.
- Hammes GG, Benkovic SJ & Hammes-Schiffer S (2011) Flexibility, diversity, and cooperativity: pillars of enzyme catalysis. *Biochemistry* 50:10422–10430.
- Hay S & Scrutton NS (2012) Good vibrations in enzyme-catalysed reactions. *Nat Chem* 4:161–168.
- Maps of metabolic pathways may be accessed at www.iubmb-nicholson.org/chart.html
- Moser CC, Chobot SE, Page CC et al (2008) Distance metrics for heme protein electron tunneling. *Biochim Biophys Acta* 1777:1032–1037.
- Klinman JP (2014) The power of integrating kinetic isotope effects into the formalism of the Michaelis–Menten equation. *FEBS J* 281:489–497.

1.7 Signaling and regulatory mechanisms

- Alonso A, Sasin J, Bottini N et al (2004) Protein tyrosine phosphatases in the human genome. *Cell* 117:699–711.
- Fell DA (1992) Metabolic control analysis: a survey of its theoretical and experimental development. *Biochem J* 286:313–330.
- Hofmeyr J-H & Cornish-Bowden A (2000) Regulating the cellular economy of supply and demand. *FEBS Lett* 476:47–51.
- Johnson SA & Hunter T (2005) Kinomics: methods for deciphering the kinome. *Nat Methods* 2:17–25.
- Koshland DE Jr, Nemethy G & Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* 5:365–385.

- Monod J, Wyman J & Changeux JP (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* 12:88–118.
- Swain JF & Giersch LM (2006) The changing landscape of protein allostery. *Curr Opin Struct Biol* 16:102–108.

1.8 Macromolecular crowding

- Bloomfield V, Dalton WO & Van Holde KE (1967) Frictional coefficients of multisubunit structures. I. Theory. *Biopolymers* 5:135–148.
- Dix JA & Verkman AS (2008) Crowding effects on diffusion in solutions and cells. *Annu Rev Biophys* 37:247–263.
- Doster W & Longeville S (2007) Microscopic diffusion and hydrodynamic interactions of hemoglobin in red blood cells. *Biophys J* 93:1360–1368.
- Malmström J, Beck M, Schmidt A et al (2009) Proteome-wide cellular protein concentration of the human pathogen *Leptospira interrogans*. *Nature* 460:762–766.
- McGuffee SR & Elcock AH (2010) Cytoplasm full energy model. doi:10.1371/journal.pcbi.1000694.s014.
- Minton AP (2006) How can biochemical reactions within cells differ from those in test tubes? *J Cell Sci* 119:2863–2869.
- Schachman HK (1959) Ultracentrifugation in Biochemistry. Academic Press.
- Zhou H-X, Rivas G & Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* 37:375–397.

1.9 Cellular compartmentation and evolution

- Archibald JM (2008) The eocyte hypothesis and the origin of eukaryotic cells. *Proc Natl Acad Sci USA* 105:20049–20050.
- Ciccarelli FD, Doerks T, von Mering C et al (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* 311:1283–1287.
- Cox CJ, Foster PG, Hirt RP, Harris SR & Embley TM (2008) The archaeabacterial origin of eukaryotes. *Proc Natl Acad Sci USA* 105:20356–20361.
- McInerney JO, Martin WF, Koonin EV et al (2011) Planctomycetes and eukaryotes: a case of analogy not homology. *BioEssays* 33:810–817.
- Woese CR, Kandler O & Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* 87:4576–4579.

