CHAPTER 4
BIOLOGICAL MATERIALS

The typical "habitat" of biological systems is made of electrolyte solutions, i.e., water solutions containing ions as solutes. We begin this chapter by characterizing the basic properties of these solutions in an elementary way. The reader will appreciate that, at this elementary level of description, several main points of this chapter are the same as those considered for describing the basic properties of semiconductors, e.g., thermal ionization, electroneutrality, diffusion and drift, and the mass action law. Of course, there are obvious differences, as we move from the realm of solid-state matter to the realm of liquid matter. In liquid matter *physical bonds* play a relevant role and therefore they will be considered again (see Chap. 1) and in more detail. Once the elementary properties of electrolyte solutions are summarized, the main biological actors will be taken into account, i.e., nucleic acids, proteins, and lipids. The cell membrane will be also described and finally a short overview of the eucaryotic animal cell will be given.

In accordance with the purpose of the book, we will not probe deeply in our description of these fundamental macromolecules. Advanced books are quoted in the proper sections of the chapter.

4.1 PHYSICAL BONDS REVISITED

As anticipated in Chap. 1, physical bonds usually lack the strong directionality and specificity of the covalent bonds considered in describing interactions inside a silicon crystal. Physical bonds are perfectly appropriate for holding molecules together in liquids, where the molecules can move about and rotate, while still remaining bonded to each other. In contrast to covalent binding, in physical binding the electrons of the different atoms do not merge, but are only "perturbed" by the bond. Nevertheless, physical binding forces, such as the Coulomb force, can be as strong as the covalent forces.

Physical binding forces display a vast range of strength, from "strong" charge-charge interactions, down to "weak" dipole-induced dipole interactions. A few relevant examples taken from this range will be considered in the following sections.

4.1.1 Ion-Ion Interactions

The interaction energy (or free energy) for the Coulomb interaction between two ions is given by

$$E(R) = \frac{z_1 z_2 q^2}{4\pi\varepsilon_0\varepsilon_r R} \tag{4.1}$$

where ε_0 is the permettivity of vacuum, ε_r is the relative permittivity of the medium, R is the distance between the two charges, and z_1, z_2 are integer positive or negative numbers, named *valence,* indicating how many elementary charges q are carried by the two ions. For example, the valence is $z = +1$ for monovalent *cations* such as K^+, $z = -1$ for monovalent *anions,* such as Cl^-, $z = +2$ for divalent cations such as Mg^{2+}, and so on.

The intensity of the *Coulomb force* is given by

$$F(R) = -\frac{dE(R)}{dR} = \frac{z_1 z_2 q^2}{4\pi\varepsilon_0\varepsilon_r R^2} \tag{4.2}$$

Let us calculate the strength of the interaction between two isolated ions (e.g., Na^+ and Cl^-) in contact in vacuum ($\varepsilon_r = 1$). In this simple case R is the sum of the two ionic radii (about 0.28 nm) and the binding energy is

$$E(R) \simeq -8.4 \times 10^{-19}\, J \simeq -5.2 \text{ eV} \tag{4.3}$$

The negative sign means that the binding is spontaneous, i.e., that positive work $W = -E$ must be exerted to break the bond. The numerical value of E is about 200 times kT at room temperature. This value of the binding energy is in the same range of the covalent ones.

Two isolated ions represent a simplified approximation for estimating the mean energy of an ionic bond in a salt lattice. Nevertheless, this approximation gives the right order of magnitude. A more appropriate calculation, made by considering the 12 next-nearest neighbors in a cubic NaCl lattice, is less than twice the approximate value.

By assuming $\varepsilon_r = 80$ (water relative permittivity) and letting the distance R between charges be equal to the sum of the radii, then Eq. (4.1) can be utilized to give an approximate estimate of the positive free energy ΔE necessary to separate Na^+ and Cl^- in water.

This energy can be used to estimate the solubility (mole fraction) X_s of ions in water forming a saturated solution in equilibrium with the solid can be then approximated by

$$X_s \simeq e^{-\Delta E/kT} \tag{4.4}$$

A derivation of Eq. (4.4) and further considerations on the ionic bond can be found in Ref. 1.

4.1.2 Ion-Dipole Interactions

The electrostatic interaction between an ion ($Q = zq$) and a dipole is depicted schematically in Fig. 4.1. The total interaction energy will be the sum of the Coulomb energy of Q with $-Q_d$ at B and of Q with $+Q_d$ at A:

$$E(R) = -\frac{QQ_d}{4\pi\varepsilon_0\varepsilon_r}\left[\frac{1}{\text{OB}} - \frac{1}{\text{OA}}\right] \tag{4.5}$$

where

$$\text{OA} = [(R + \tfrac{1}{2} l\cos\theta)^2 + (\tfrac{1}{2} l \sin\theta)^2]^{1/2} \tag{4.6}$$

$$\text{OB} = [(R - \tfrac{1}{2} l\cos\theta)^2 + (\tfrac{1}{2} l \sin\theta)^2]^{1/2} \tag{4.7}$$

l and R being the dipole length and the length of the segment OM, respectively.

At separations $R > l$, we can approximate $\text{OA} \simeq R + \tfrac{1}{2} l\cos\theta$ and $\text{OB} \simeq R - \tfrac{1}{2} l\cos\theta$. Moreover, by neglecting $(l^2/4)\cos^2\theta$ in comparison with R^2, we obtain

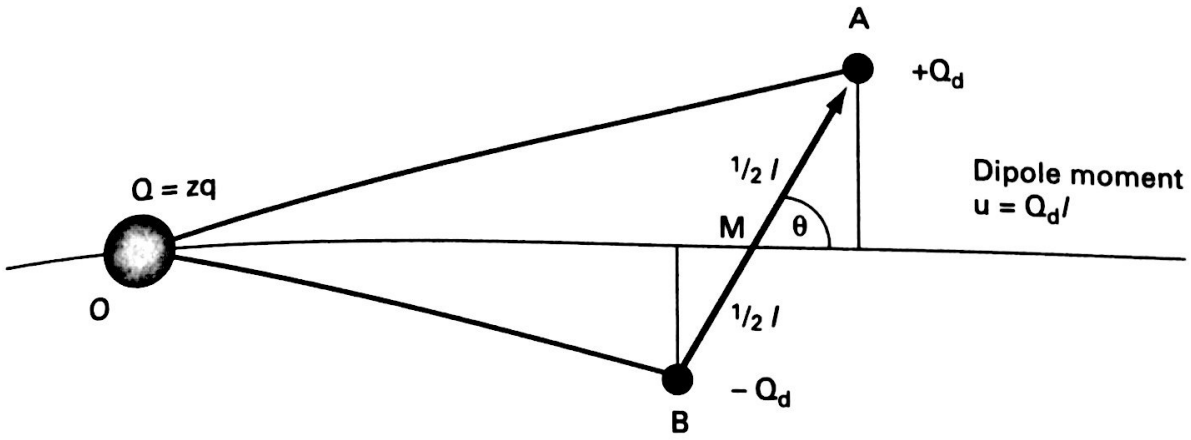

FIGURE 4.1 Ion-dipole configuration resulting in Eq. (4.5).

$$E(R) \simeq -\frac{zqu \cos \theta}{4\pi\varepsilon_0\varepsilon_r R^2} \tag{4.8}$$

We can use Eq. (4.8) to calculate the interaction between an Na^+ ion and a water dipole in *vacuum* (i.e., in a cloud). A drastic simplification for the water dipole is to consider it as a spherical molecule of radius 0.14 nm and with a point dipole of moment 1.85 D (1 D = 1 debye = 3.336×10^{-30} C-m). Thus, for an Na^+ ion (valence $z = 1$, radius = 0.09 nm) near a water molecule, the maximum interaction energy will be given by Eq. (4.8) for $\theta = 0°$.

$$E(R, \theta = 0°) \simeq -1.6 \times 10^{-19} \text{ J} \tag{4.9}$$

The energy value in Eq. (4.9) is about $40kT$ at room temperature. For the small divalent cation Mg^{2+} ($z = 2$, radius = 0.06 nm) the value rises to about $100kT$.

When ion-water interactions take place in *bulk water,* the above energies will be reduced by a factor of about 80 (the bulk water relative permittivity value); even then the strength of interaction will exceed kT at room temperature for small divalent (or multivalent) ions and it is not negligible for small monovalent ions. We can conclude that small or multivalent ions in water will tend to orient the water dipoles around them, with $\theta = 0°$ near cations, and $\theta = 180°$ near anions. As a result cations and anions in solution have a number of water molecules orientationally bound to them. The ions are said to be *hydrated* and the number of bound water molecules (typically 4 to 6) is known as the *hydration number.* The bound dipoles form a *hydration shell.* The reader should realize that there is a continuous exchange of bound dipoles with bulk water. The lifetime of a bound water dipole can range, at room temperature, from about 10^{-9} s for small monovalent cations such as Li^+, to 10^{-8} s for Ca^{2+}, and to 10^{-6} to 10^{-5} s for Mg^{2+}.

Large monovalent anions (such as Cl^-, $z = -1$, radius = 0.18 nm) are generally less hydrated than cations, and the lifetimes are comparable to those of two bulk water molecules (about 10^{-11} s).[1] Hydration modifies the dielectric properties of water and, consequently, the short-range Coulomb interaction among ions.

We will see in subsequent sections that water bonds also *polyelectrolytes,* such as DNA and proteins.

4.1.3 Dipole-Dipole Interaction and Hydrogen Bonding

When two polar molecules are near each other, there is a dipole-dipole interaction. It can be shown[1] that for two point dipoles of moment u_1 and u_2 at a distance R (see Fig. 4.2) the interaction energy is

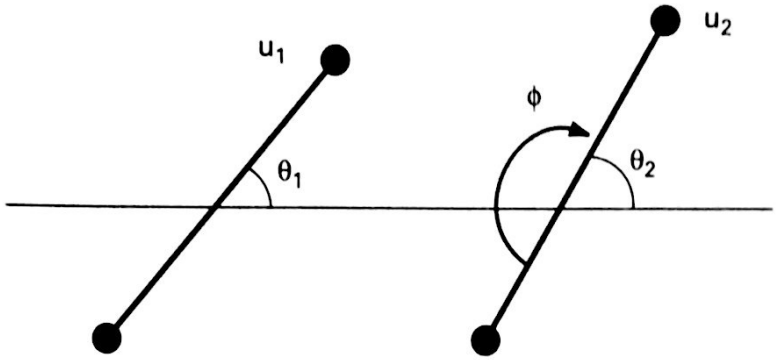

FIGURE 4.2 Dipole-dipole configuration resulting in Eq. (4.10).

$$E(R, \theta_1, \theta_2, \phi) = -\frac{u_1 u_2}{4\pi\varepsilon_0\varepsilon_r R^3}[2\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2\cos\phi] \qquad (4.10)$$

For dipole moments in the order of 1 D, typical energies are weaker than kT at room temperature. An *exception* is given by the bond F^-—H^+, O^-—H^+, N^-—H^+.

F, O, and N are the most *electronegative* elements, that is, they tend to strongly attract the molecular electronic orbitals when making a covalent bond with other elements. This fact results in a very strong dipole when the very small, electron-depleted H atom (almost a bare proton) is bonded.

To fix the ideas, let us consider a collection of water molecules: in each of them the presence of O^-—H^+ bonds generates a very strong dipole and therefore each dipole (i.e., each H_2O molecule) strongly interacts with other dipoles (i.e., other H_2O molecules). Because of this strong dipole-dipole interaction, water is said to be an *associate liquid.* A similarly strong dipolar interaction could be also made by H^+ if other electronegative groups, such as —N or C=O, were present.

In all of these situations, a very strong type of directional dipole-dipole interaction is generated, which can be in the order of 10 to 40 $kJmol^{-1}$. Such a dipole-dipole interaction is named *hydrogen bond.* This is the bond which allows the organization of the secondary structures of proteins and which stabilizes the DNA double helix (see Sec. 4.5).

4.2 WATER AND ELECTROLYTE SOLUTIONS

In a very rudimentary sense, a 150-mM NaCl water solution at 37°C (310 K) is the basic environment of a mammalian cell. Under these conditions water is very poorly ionized (as a silicon crystal is) and NaCl is totally dissociated in hydrated Na^+ cations and hydrated Cl^- anions. Other chemical species can then be added, thus making the situation more complex.

4.2.1 Water as a Collection of Partially Ionized Molecules

Under standard conditions, 1 liter of water is equal to 1 kg of water and it contains (1000/18) = 55.6 moles = 335×10^{23} water molecules. At room temperature (300 K), a very small fraction of these dipolar molecules is thermally ionized, according to the *generation-recombination scheme:*

$$[H_2O] \underset{k_2}{\overset{k_1}{\rightleftarrows}} [H^+] + [OH^-] \qquad (4.11)$$

The square brackets indicate concentrations (*molarity,* i.e., moles/liter, *M*). According to the *mass action law,* Eq. (4.11) corresponds to the differential equations.

$$\frac{d}{dt}[H_2O] = -k_1[H_2O] + k_2[H^+][OH^-] \tag{4.12a}$$

$$\frac{d}{dt}[H^+] = k_1[H_2O] - k_2[H^+][OH^-] \tag{4.12b}$$

$$\frac{d}{dt}[OH^-] = \frac{d}{dt}[H^+] \tag{4.12c}$$

with the constraint that the sum of the number of un-ionized molecules and of the number of H^+—OH^- pairs is, at any time, a fixed constant (see also Sec. 5.4.1).

Under equilibrium conditions ($d/dt = 0$), Eqs. (4.12*a*) to (4.12*c*) reduce to the equality

$$k_1[H_2O] = k_2[H^+][OH^-] \tag{4.13}$$

or

$$\frac{[H^+][OH^-]}{[H_2O]} = \frac{k_1}{k_2} = K_{eq} \tag{4.14}$$

If we assume that the fraction of ionized molecules is very small, we can approximate $[H_2O]$ with the total concentration of water and incorporate it into the equilibrium constant, i.e.,

$$[H^+][OH^-] = K_w \tag{4.15}$$

Under standard conditions

$$K_w = 10^{-14}\ M^2 \tag{4.16}$$

That is,

$$[H^+] = [OH^-] = 10^{-7}\ M \tag{4.17}$$

or

$$\frac{10^{-7}}{10^3} 6.02 \times 10^{23} = 6.02 \times 10^{13}\ \text{ions/cm}^3 \tag{4.18}$$

We wish to remind the reader that, under the same standard conditions (see Chap. 2),

$$n_i = p_i = 1.45 \times 10^{10}\ \text{carriers/cm}^3 \tag{4.19}$$

It is customary to translate Eq. (4.17) into the expression

$$pH = 7 \tag{4.20}$$

where, of course

$$pH = -\log [H^+] \tag{4.21}$$

The notion of pH can be extended to other water solutions where the presence of other chemicals unbalances Eqs. (4.16) and (4.17). Chemicals which are essential to understand the properties of biological molecules are *acids* and *bases.* A great many of the low-molecular-weight metabolites and macromolecular components of living cells are acids and bases, whose main feature is the potential to ionize. According to the Brönsted concept of

acids and bases, an *acid is a substance that donates protons* (hydrogen ions), and a *base is a substance that accepts protons.*[2]

Acids and bases can be *strong* or *weak*. A strong acid is a substance that ionizes almost 100 percent in aqueous solutions. For example HCl in solution is essentially 100 percent ionized to H_3O^+ and Cl^-; i.e.,

$$HCl + H_2O \rightarrow H_3O^+ + Cl^- \tag{4.22}$$

where H_3O^+ (the *hydronium ion,* or conjugate acid of water) is the actual form of the hydrogen ion (i.e., a proton) in solution. The ionization of HCl can be represented as the simple dissociation

$$HCl \rightarrow H^+ + Cl^- \tag{4.23}$$

and, for all practical purposes, H_3O^+ and H^+ can be used interchangeably. A strong base is a substance that ionizes almost 100 percent in aqueous solutions yielding OH^- ions.

Potassium hydroxide is an example of a strong inorganic base that ionizes according to

$$KOH \rightarrow K^+ + OH^- \tag{4.24}$$

Strong acids and strong bases affect the H^+ and OH^- concentration of an aqueous solution in exactly the same way as acceptors and donors unbalance the concentration of electrons and holes in a silicon crystal.

Under the same approximations utilized in Chap. 2, the reader can easily verify that a 1-mM concentration of HCl will shift the H^+ concentration from 10^{-7} M to 10^{-3} M. Assuming that the water H^+ and OH^- generation process is not affected (i.e., $[H] \times [OH^-] = 10^{-14}\ M^2$), the new OH^- concentration is

$$[OH^-] = 10^{-11} M \tag{4.25}$$

The analogy with the doping of a semiconductor is apparent. Of course, if the strong acid/base concentration is equal or even smaller than 10^{-7} M, then the final result is obtained by solving a quadratic equation. The reader should remember that this is also the case for silicon doping comparable to the intrinsic carrier concentration. This analogy is illustrated in Example 4.1.

Example 4.1 Find the pH of a 10^{-8}-M solution of HCl.

Answer By indicating $[H^+] = X$, the second order equation

$$X^2 - 10^{-8} X - 10^{-14} = 0 \tag{E4.1}$$

should be solved, which gives

$$[H^+] = 10.51 \times 10^{-8}\ M \tag{E4.2}$$

or

$$pH = 6.98 \tag{E4.3}$$

In contrast to strong acids and bases, *weak* acids and *weak* bases ionize to a limited extent in aqueous solutions. The effects of their partial ionization on the solution pH is analyzed in the next section.

4.2.2 Partial Ionization and Buffer Solutions

In an aqueous solution a weak acid HA (e.g., acetic acid, CH_3—COOH) ionizes to a limited extent as follows:

$$\underset{[\text{Conjugate acid}]_1}{HA} + \underset{[\text{conjugate base}]_2}{H_2O} \rightleftharpoons \underset{[\text{conjugate base}]_1}{A^-} + \underset{[\text{conjugate acid}]_2}{H_3O^+} \tag{4.26}$$

The proton released from the weak acid HA is accepted by water to form H_3O^+. It should be noted that, according to the Brönsted definition, HA is a conjugated acid and A^- is the corresponding conjugated base. Every time a Brönsted acid loses a proton, then a Brönsted base is produced. The substance that accepts the proton is a different Brönsted base; by accepting the proton, another Brönsted acid is produced. Thus, in every ionization of an acid or base, two conjugate acid–conjugate base pairs are involved [see Eq. (4.26)].

Equation (4.26) is a reversible reaction described by the equilibrium constant

$$K = \frac{[H_3O^+][A^-]}{[HA][H_2O]} \tag{4.27}$$

If the "practically" constant value of $[H_2O]$ is incorporated into K and $[H^+]$ is substituted for $[H_3O^+]$, Eq. (4.27) becomes

$$K_a = \frac{[H^+][A^-]}{[HA]} \tag{4.28}$$

which corresponds to the dissociation reaction:

$$HA \rightleftharpoons H^+ + A^- \tag{4.29}$$

One of the few simple examples of a weak inorganic base is given by ammonia:

$$NH_4OH \rightleftharpoons NH_4^+ + OH^- \tag{4.30}$$

Among organic bases, amines (R—NH_2, where R is a generic radical) play an important role. The reader is referred to Ref. 2 for more on this topic.

Considering again a weak acid, the dissociation of HA yields H^+ and A^- in equal concentrations. Therefore we can write

$$K_a = \frac{[H^+][A^-]}{[HA]} = \frac{[H^+]^2}{[HA]} \tag{4.31}$$

If we assume that the degree of ionization is small so that [HA] is practically equal to the total concentration of the acid $[HA]_T$, and if we further assume that the contribution of water to $[H^+]$ is negligible, then the final $[H^+]_F$ concentration is simply given by

$$[H^+]_F = (K_a[HA]_T)^{1/2} \tag{4.32}$$

or

$$pH = \tfrac{1}{2}(pK_a + p[HA]) \tag{4.33}$$

where $pK_a = -\log K_a$ and $p[HA]_T = -\log [HA]_T$.

Equation (4.31) can be rearranged into

$$\log[H^+]_F = \log K_a + \log \frac{[HA]_T}{[A^-]} \tag{4.34}$$

or

$$pH = pK_a + \log \frac{[A^-]}{[HA]_T} \tag{4.35}$$

Equation (4.35) is a useful starting point for the calculation of the *capacity* of a *buffer*.

Conjugate acids and conjugate bases can form *pH-buffer* solutions, which are of great relevance for any biological system. A pH buffer is a substance (or a mixture of substances) that allows a solution to maintain a near constant pH upon the addition of small amounts of H^+ or OH^- ions to it. Amino acids and proteins, which will be considered in Sec. 4.4.1, are examples of biological buffers. The capacity β of a buffer can be defined as the number of moles per liter of H^+ and OH^- required to cause a given change in pH.

It can be shown[2] that

$$\beta = \frac{2.3[A^-][HA]}{[A^-] + [HA]} \tag{4.36}$$

or

$$\beta = \frac{2.3\, K_a[H^+][C]}{(K_a + [H^+])^2} \tag{4.37}$$

where C is the total concentration of buffer components (i.e., $[C] = [A^-] + [HA]$).

We conclude this section with a definition quite important for subsequent descriptions of amino acids and proteins, i.e., the definition of *polyprotic acids.*

A polyprotic acid ionizes in successive steps, according to the following scheme

$$\begin{array}{ccc} H_2A & \overset{k_{a1}}{\rightleftharpoons} & H^+ + HA^- \\ & & \Big\Updownarrow k_{a2} \\ & & H^+ + A^{2-} \end{array} \tag{4.38}$$

where

$$K_{a1} = \frac{[H^+][HA^-]}{[H_2A]} \tag{4.39}$$

$$K_{a2} = \frac{[H^+][A^{2-}]}{[HA^-]} \tag{4.40}$$

When an intermediate ion of a polyprotic acid is dissolved in water, it undergoes both ionization as an acid and ionization as a base or hydrolysis. There will be a specific pH value at which species with no net charge are present. This is the *isoelectric point* (pI) of that specific polyprotic acid. This definition will be further considered in Sec. 4.4.1 in a discussion of amino acids.

4.2.3 Ionic Strength

The *ionic strength* of an electrolyte solution is defined as

$$I = \tfrac{1}{2}\sum_{i=1}^{n} c_i z_i^2 \tag{4.41}$$

where c_i is the concentration of the ith ion, z_i is its valence, and the summation extends over all the ions in the solution.

It is easy to verify that for a single 1:1 electrolyte (e.g., NaCl) the ionic strength is

equal to its concentration. A 2:2 totally ionized salt (such as Mg_2SO_4) has an ionic strength which is 4 times that of a 1:1 salt with the same concentration.

It is important to keep in mind that ionic strength is a property of the solution and is not a property of any particular ion in the solution. Only the *net* charge of an ion is used in calculating ionic strength. Therefore, an un-ionized compound (e.g., a weak acid) or a dipolar compound does not contribute toward the ionic strength of a solution. An increase in the ionic strength of a solution implies an increase in the electrostatic screening by ions of the macromolecules present in the solution. Biological complexes held together by electrostatic interactions can be separated into their single macromolecular components simply by increasing the ionic strength of the solution. Moreover, because of this screening effect, proteins and DNA tend to assume a folded conformation when the solution ionic strength is increased, as will be discussed in the following sections.

4.3 OPTICAL PROPERTIES OF MOLECULES IN SOLUTION

The Beer-Lambert law [see Eq. (2.6)] can be usefully utilized for estimating the concentration of molecular species in solution. For the reader's convenience we introduce again here the equation utilized to deduce the intensity I_{tr} of the light of frequency $\nu > \nu_{th}$, transmitted by a slab of semiconductor of thickness l, exposed to light of intensity I_o

$$I_{tr}(l) = I_o e^{-\alpha(\nu)l} \tag{4.42}$$

We can adapt Eq. (4.42) to a solution inside a tube of thickness l, and containing a solute of concentration c, by writing

$$I_{tr}(l) = I_o e^{-\varepsilon(\nu)cl} \tag{4.43}$$

where ε, known as the *molar extinction coefficient* of the solute, is a function of the frequency ν of the incident light. Equation (4.43) deserves two comments:

- Equation (4.43) is correct only if the tube containing the solution and the solvent does not absorb any of the incident light. If this is not the case, then the intensity of light, transmitted by the same tube filled with the solvent only, has to play the role of I_o in Eq. (4.43).
- In the case of a typical light-absorbing solute (i.e., a *dye*), there is not a threshold frequency ν_{th} in the same sense as it was defined for semiconductors. This observation will be clarified later in this section.

As already indicated in Chap. 2, the intensity of the absorbed light is

$$I_{abs}(l) = I_o(1 - e^{-\varepsilon(\nu)lc}) \tag{4.44}$$

The corresponding absorbed energy will be partially dissipated in heat and partially given back radiatively as fluorescence light according to

$$I_{fl}(\nu') = QI_{abs} = QI_o(1 - e^{-\varepsilon(\nu)lc}) \tag{4.45}$$

where Q is the *fluorescence quantum yield* ($0 \leq Q \leq 1$).

As already discussed for silicon (see Sec. 2.1.3), the frequency ν' of the fluorescence emission is smaller than the frequency ν of the exciting light:

$$\nu' < \nu \tag{4.46}$$

Note that the preferential direction of the exciting light is completely lost by the fluorescence light, which is emitted isotropically in all directions. If the product εcl is sufficiently small, then Eq. (4.45) can be approximated by

$$I_{fl}(\nu') \simeq QI_o\varepsilon cl \tag{4.47}$$

which gives a direct linear relationship between solute concentration and fluorescence intensity. A qualitative description of the fluorescence phenomenon at the molecular level is as follows: a photon of frequency ν is absorbed by a molecule M, with the consequent transition of an electron of the molecule from its ground state to an excited one. The excited molecule M* then decays to its ground state either by fluorescence emission or by heat production. The process is described by the kinetic relation

$$\frac{d}{dt}[\mathrm{M}^*] = -(k_F + k_H)[\mathrm{M}^*] \tag{4.48}$$

where k_F and k_H are the rate constants of the radiation (fluorescence) and radiationless heat processes and [M*] represents the concentration of excited molecules.

In other words, a population of excited molecules M* decays to its ground state according to

$$\mathrm{M}^* = \mathrm{M}_0^* e^{-t/\tau} \tag{4.49}$$

where

$$\tau = \frac{1}{k_F + k_H} \tag{4.50}$$

It can be easily shown that

$$Q = \frac{k_F}{k_F + k_H} \tag{4.51}$$

The lifetime τ is typically in the order of 10^{-9} to 10^{-7} s, depending both on the molecular structure of the absorbing species and on the physicochemical properties of the solution.

Having introduced a molecular view of the absorption/emission process, we can now consider in some detail the absorption/emission spectra of molecules in solution.

We saw in Chap. 1 that hydrogen atoms absorb photons of discrete frequencies, and consequently their absorption spectrum is made of lines. Semiconductors (e.g., silicon), because of their energy band structure, absorb photons of all the frequencies (up to a certain value) greater than a threshold ν_{th} and therefore their absorption spectrum is flat (up to a certain value), starting from ν_{th} (see Chap. 2). Molecules in solution have absorption spectra made of one or more bell-shaped curves.[3] Let us consider a simple molecule, i.e., a diatomic one which can be approximated by a harmonic oscillator. In the first place we can say that the energy of the molecule will depend on the *electronic state,* i.e., the set of orbitals that the electrons in the molecule occupy.

For a given state, the energy will depend on the distance between the nuclei of the two atoms. For each electronic state, the molecule will have a set of allowed levels of *vibrational energy,* which could be approximately calculated by considering the quantum equivalent of the classical oscillator model.[4]

The energy separation between vibrational levels is of course much smaller than the overall energy separation between electronic levels. Finally, we should also consider a quantized *rotational energy;* with states which correspond to sets of even more closely spaced lines clustered above each vibrational level. A schematic representation of molecular energy levels is given in Fig. 4.3, where a comparison with the room temperature value of kT is given. By making use of the Boltzmann distribution, readers can easily con-

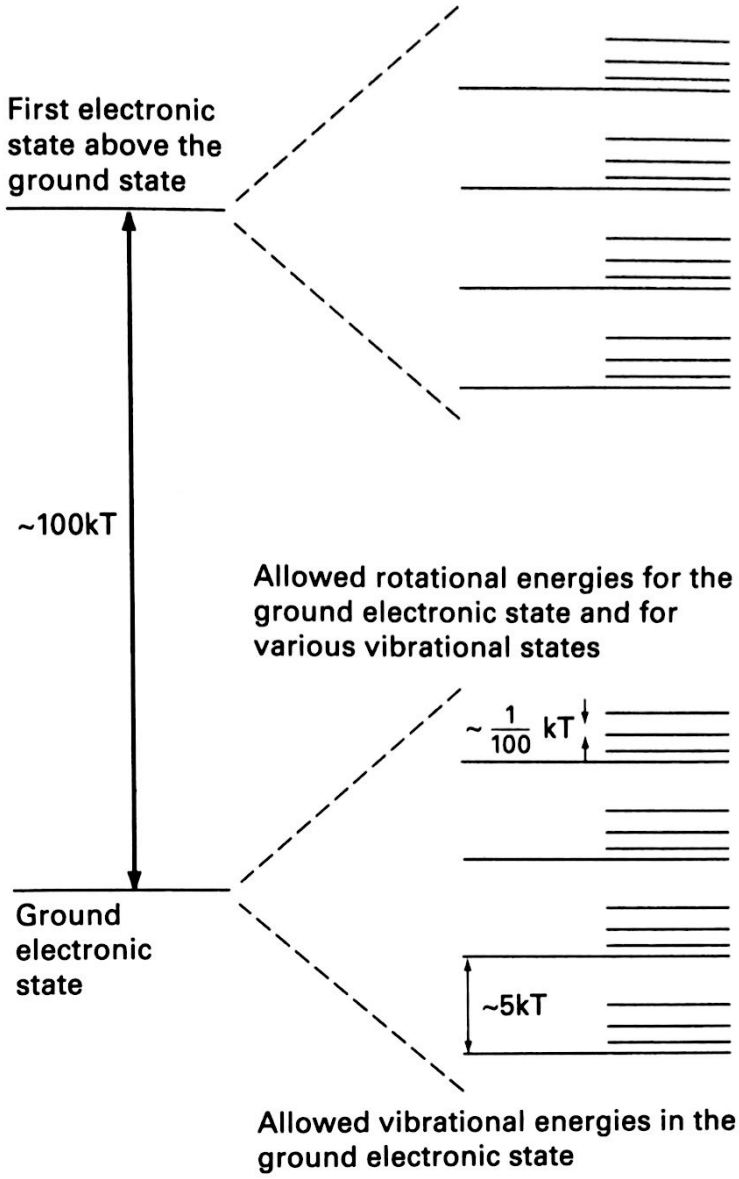

FIGURE 4.3 Schematic representation of typical molecular energy levels. A value of kT corresponding to room temperature is assumed.

vince themselves that at room temperature only the lowest vibrational level of the ground electronic level should be occupied. On the contrary, however, all the rotational levels of this ground level, separate by $^1/_{100}\ kT$ only, are expected to be fully occupied.

By transforming differences in energy levels into photon frequencies, it can be appreciated why infrared light is being used to generate *vibrational spectra* of molecules. Infrared photon absorption at room temperature causes transitions from the lowest vibrational level of the ground electronic state to higher vibrational levels of the same ground state. A nonlinear molecule with n atoms will have $(3n - 6)$ fundamental modes of vibration. Among the biological molecules (to be considered in the next section), even for a simple substance such as an amino acid, the number of modes is large and consequently a quite complex vibrational spectrum is obtained. Transitions in the visible region of the spectrum are relatively low-energy electronic transitions. They are typical of biological macromolecules containing metal ions. Generally speaking, proteins and nucleic acids, which will be considered in detail in the next sections, present bell-shaped absorption

bands in the near ultraviolet (200 to 400 nm). According to the scheme of Fig. 4.3, these bands can be interpreted as being made by the absorption of photons allowing any transition from the ground electronic state to one of the several vibrorotational states of the first excited electronic level, with a probability, and consequently an absorption intensity, dictated by quantum mechanical rules.[3]

Only a few amino acids (e.g., *tryptophan*) present appreciable fluorescence emissions. In most cases, biological macromolecules such as proteins and DNA are made fluorescent by binding appropriate fluorescent dyes to them.

4.4 BIOLOGICAL MOLECULES—PROTEINS

Up to now we have considered water, ions, and small metabolites. In a typical mammalian cell, with a volume of about 4×10^{-9} cm^3, their weight contribution is 70 percent, 1 percent, and 3 percent of the total, respectively. The remaining 26 percent of the weight is roughly shared among proteins (18 percent), nucleic acids (1 percent), lipids (5 percent), and sugars (2 percent). These are the molecules peculiar to living matter. Their organization in a cell is very sophisticated, and it seems to follow a *bottom-up* design, just opposite to the *top-down* process typical of present-day microelectronics.

The amount of research work on these biological molecules, especially on proteins and nucleic acids, is enormous, and a detailed analysis of their properties is out of the scope of this book. Under these premises, a few features of biological molecules, relevant to our *bioelectronic approach,* will be described in the following, beginning with proteins.

4.4.1 Proteins

Proteins are very sophisticated devices that play many fundamental roles in living systems, including catalysis, mechanical resistance, and recognition. The functional properties of proteins depend upon their three-dimensional structures.

The three-dimensional (3-D) structures arise because particular sequences of *amino acids* fold to generate complex architectures in space. The prediction of the 3-D protein structure, starting from the knowledge of its amino acid sequence, is the main issue of *protein engineering.*

As already stated, proteins are made of amino acids. There are 20 amino acids (plus a few rare ones). All of the 20 amino acids have in common a central carbon atom (C_α) to which a hydrogen atom (H), an amino group (NH_2), and a carboxyl group (COOH) are attached. A side chain is linked to the C_α atom through its fourth valency (Fig. 4.4). Any of the 20 amino acids differs from the others because of a different side chain.

Amino acids are joined end to end during protein synthesis by the formation of covalent bonds named *peptide bonds.* The bond formation is enzymatically catalyzed. The carboxyl group of one amino acid is *condensed* with the amino group of the next by elimination of a molecule of water, yielding a peptide bond NH—CH—CO. This process is repeated as the chain elongates. The formation of a succession of peptide bonds (from the first-unbonded-amino terminus to the last-unbonded-carboxyl terminus) generates the backbone of a protein, from which the various side chains project (see Fig. 4.4). This sequence of amino acids is the protein *primary structure.* The number of amino acids in a protein can range from a few tens to thousands.

The 20 amino acids are polyprotic acids and they can be divided into three classes, according to the properties of the side chain: class I, with an apolar (i.e., hydrophobic) side chain; class II with a polar side chain; and class III with an acid or basic side chain. The

amino acid *glycine* has only a hydrogen atom as a side chain and can be considered to constitute a fourth class. With the exception of glycine, the four groups attached to the C_α atom are chemically different. Therefore all amino acids (except glycine) are chiral forms, which can exist in left (L-) and right (D-) forms. Proteins of eucaryotic cells (i.e., cells provided with a nucleus) are all made of L-amino acids.

The structures of an apolar amino acid (*alanine*) and of an acidic amino acid (*aspartic acid*) are compared in Fig. 4.5. Note that the amino and carboxyl groups of the main chain are shown in their ionized form, i.e., NH_3^+ and COO^-, respectively.

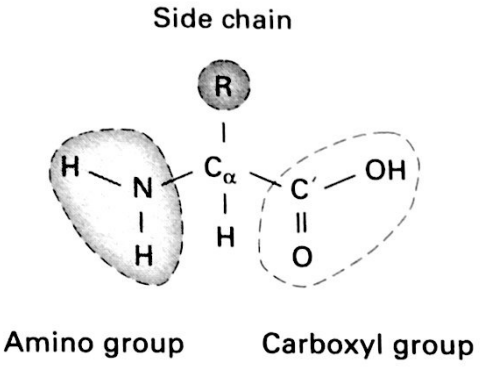

FIGURE 4.4 Schematic representation of an amino acid.

As Fig. 4.5 shows, the side chain of the aspartic acid has an extra carboxyl group. As a consequence, the aspartic acid has an isoelectric point pI smaller than 7 and, at physiological pH (around pH = 7.4), it is negatively charged. Figure 4.6 summarizes the pH-dependent amino acid charge. Other amino acids present an extra amino group (for example, *lysine*), have an isoelectric point greater than 7, and bear a positive charge at physiological pH.

As the reaction schemes of Fig. 4.6 suggest, at a given pH the primary structure of a protein is a complex sequence of negative charges, positive charges, and hydrophobic and polar regions, according to the side chains of the amino acid sequence. Moreover, it should be noted that H and the highly electronegative elements O and N are periodically repeated along the backbone and therefore hydrogen bonding (see Sec. 4.1.3) can be expected. A specific sequence of amino acids in a protein can wind on itself via H bonding, forming an helix (called *α helix* by L. Pauling, who first described it in 1951).

The length of α helices varies considerably in proteins, ranging from 4 or 5 amino acids to over 40. Helix formation/disruption is a phenomenon driven by several physicochemical parameters, such as pH, temperature, ionic strength and medium polarity.

The formation of an α helix is a well-studied phenomenon, and the classical mathematical tools of statistical mechanics are appropriate for describing it. As indicated in Fig. 4.7, a helix turn is formed via hydrogen bonding between O of amino acid i and H of amino acid $i + 4$. This helix turn forces six conformation angles along the chain to assume fixed values. The propagation of the helix implies fixing two angles per turn. This is a clear example of a *cooperative phenomenon,* which is characterized by the fact that the *starting of the process is more energy-demanding than the advancing of it.* In other words, the coil-to-helix transition is a nonlinear phenomenon.

If we plot the fraction of helix turns θ versus any physicochemical parameter s (e.g., pH), by varying where a transition is induced, then we get a sigmoidal curve (Fig. 4.8)

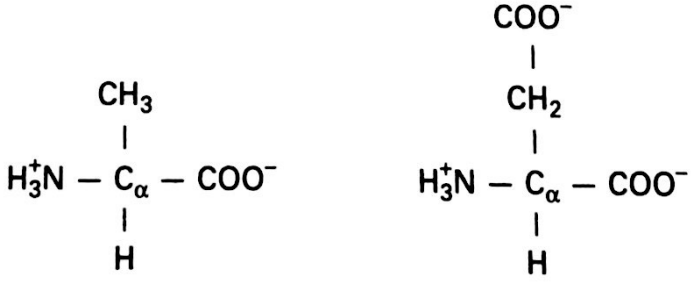

FIGURE 4.5 Structures of *alanine* (left) and *aspartic* acid (right).

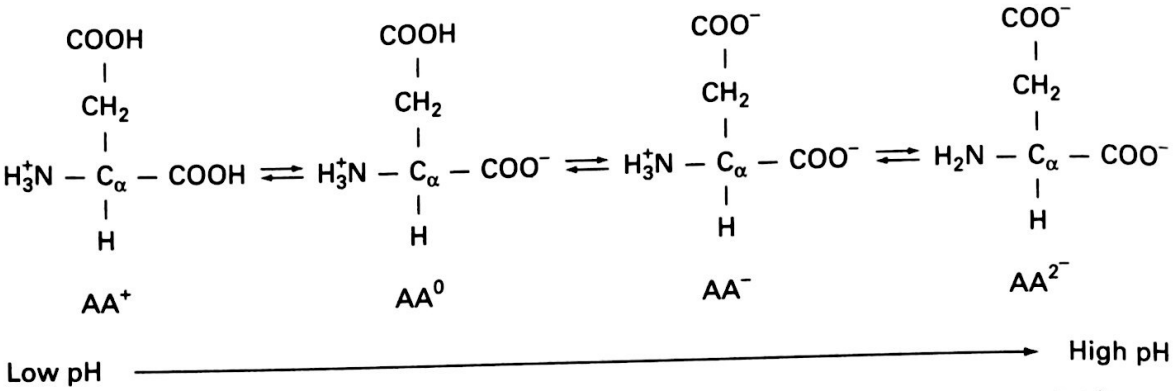

FIGURE 4.6 At pH = pI = 2.98 the net charge of the amino acid (AA) *aspartic acid* is zero (AA^0). The amino acid is positively charged (AA^+) at low pH and negatively charged (AA^-, AA^{2-}) at high pH. This is a general behavior, which can be applied both to amino acids and proteins.

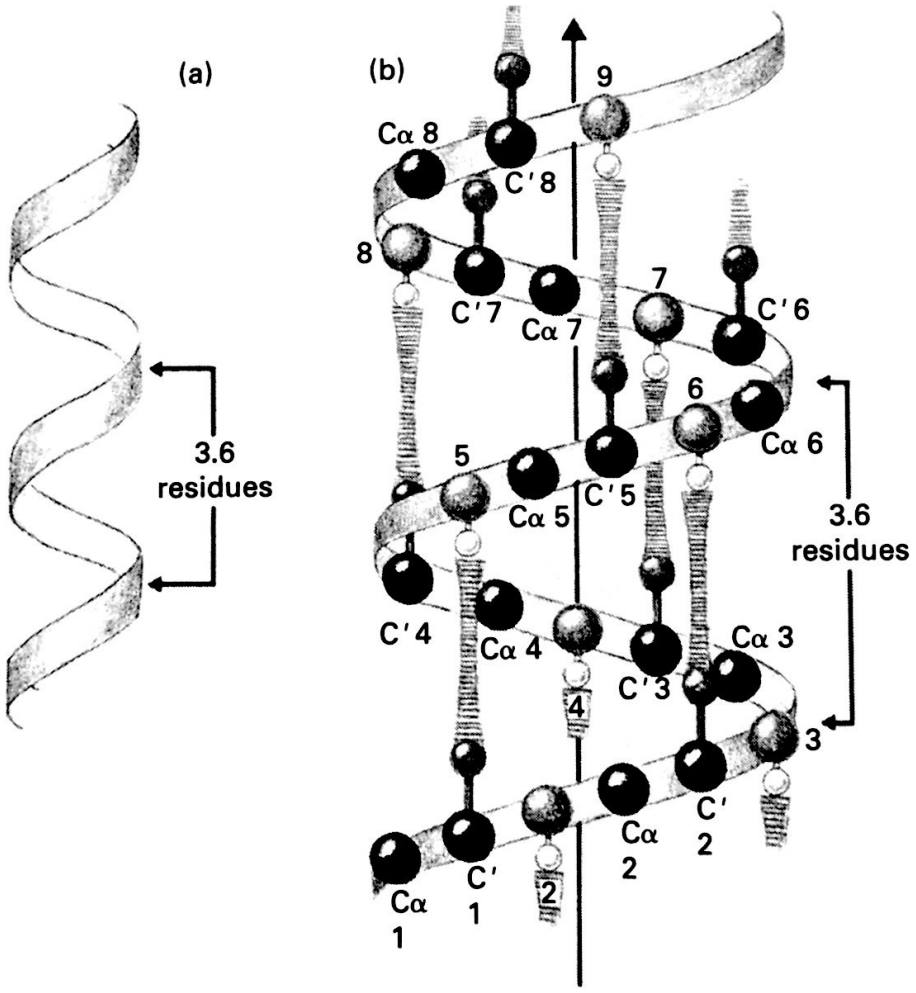

FIGURE 4.7 (*a*) Idealized diagram of the path of the main chain in a protein α helix. There are 3.6 residues per turn in an α helix, which corresponds to 0.54 nm (0.15 nm per residue); (*b*) the path with approximate positions for main chain atoms and hydrogen bonds included. (*Adapted from C. Branden and J. Tooze.*[6] *Used by permission.*)

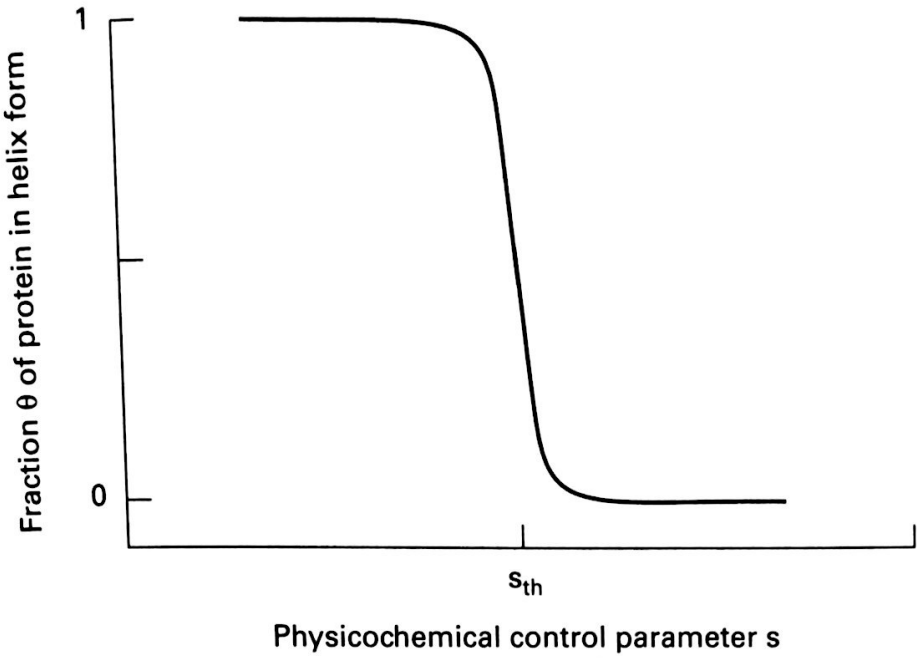

FIGURE 4.8 The coil-to-helix transition is a cooperative phenomenon. The fraction of helicity is "almost" one for s smaller than the threshold value s_{th} and it "jumps" to zero for $s > s_{th}$.

with a threshold value s_{th}.[5] This curve is an approximation of a two-state function, which is the fingerprint of any (electronic or biological) device appropriate for information processing.

Hydrogen bonds can also be formed among several regions of a polypeptide chain. This results in a structure known as *β sheet.* Different regions of a sequence of amino acids can be organized in either of these local regular structures, which represent the secondary structure of a protein. The packing of these structures into one or several compact globular units, called *domains*, determines the tertiary structure of a protein. The sequence of primary, secondary, and tertiary structures is sketched in Fig. 4.9.

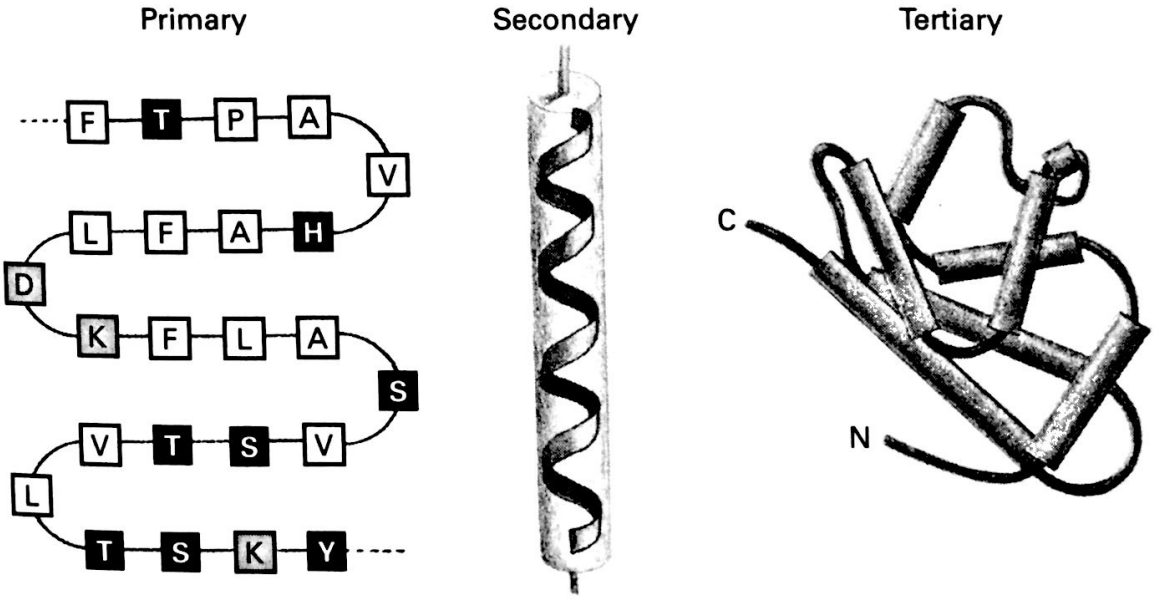

FIGURE 4.9 Structures of a protein. (*Adapted from C. Branden and J. Tooze.*[6] *Used by permission.*)

Protein structures are built up by combinations of secondary structural elements, α helices and β strands, according to a bottom-up design. α helices and β strands typically form the core regions of globular proteins and are connected to the protein surface by loops.[5] A protein may also contain several polypeptide chains arranged in a *quaternary structure.* By formation of tertiary and quaternary structures, amino acids far apart in the primary sequence are brought close together in three dimensions to form a *functional* region (an active site). In conclusion, in a very complex and only partially understood process, the bottom-up procedure organizes structures in such a way as to obtain highly sophisticated functions. A few examples will be described in the discussion of the plasma membrane in Sec. 4.7. Enzyme catalysis will be considered in Chap. 5.

4.5 NUCLEIC ACIDS

There are two kinds of nucleic acids in a living cell: *deoxyribonucleic acid* (DNA) and *ribonucleic acid* (RNA). Both of them bear negative charges at physiological pH. The term *nucleic* suggests their presence in the nucleus of a cell, but they can be found also outside the nucleus of an eucaryotic cell: DNA is present in mitochondria and RNA is present in various forms in the cytoplasm (see Sec. 4.8). A few key features of these biological macromolecules will be summarized in the following.

4.5.1 DNA and RNA Structures

A DNA chain is a long, unbranched polymer composed of four types of subunits, called *deoxyribonucleotides.* Each nucleotide is formed by a *base,* a *five-carbon sugar,* and one *phosphate group.* There are four kinds of bases. They are all nitrogen-containing ring compounds, either purines [the bases *adenine* (A) and *guanine* (G)] or pyrimidines [the bases *cytosine* (C) and *thymine* (T)]. The name *base* arises from the fact that these molecules can combine with H^+ in acidic solutions. Nucleotides are joined together by a covalent bond called *phosphodiester linkage.* The result is a linear chain, which can be identified simply by indicating the bases symbols as in Fig. 4.10.

Specific hydrogen bonding between G (large purine base) and C (smaller pyrimidine base) and between A (large purine base) and T (smaller pyrimidine base) generates base pairing between complementary linear chains. The result is the well-known DNA *double helix,* described in 1953 by Watson and Crick on the basis of x-ray diffraction data obtained by Wilkinson and Franklin. In a DNA molecule two antiparallel strands that are complementary in their nucleotide sequence are paired in a right-handed double helix with about 10 nucleotide pairs per helix turn. This is the so-called *B form* of DNA (see Fig. 4.11)

Two other forms have been described, namely A and Z forms. The B form is considered to be the typical form assumed by DNA inside the nucleus of a living cell.

A typical mammalian cell contains about one meter of DNA (3×10^9 nucleotide pairs). In other words, inside a cell nucleus (diameter in the order of 5 to 10 μm) there is a collection of 6×10^9 symbols, belonging to four categories only (A, C, G, T). Specific subsets of this collection represent the *code* utilized to synthesize different proteins. The process of synthesis involves RNA.

Inside a cell there are three major species of RNA present in several copies, namely: *messenger* RNA (mRNA); *transfer* RNA (tRNA); *ribosomal* RNA (rRNA). Any RNA

··· — A — C — T — A — G — C — ···

FIGURE 4.10 Symbolic representation of a DNA strand.

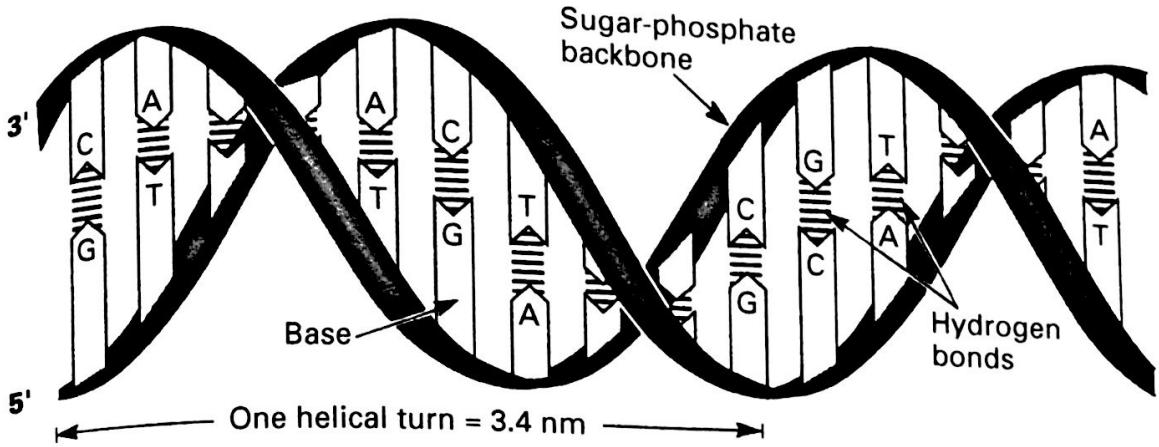

FIGURE 4.11 In a B-DNA molecule two antiparallel strands that are complementary in their base sequence are paired in a right-handed double helix held together by hydrogen bonds.

molecule is formed again by nucleotides, where the sugar is now *ribose* (and not deoxyribose, as in DNA) and one of the pyrimidine bases is *uracil* instead of thymine.

DNA is a single huge macromolecule lying inside the core of the cell as protected as the source code of a very important software program. In contrast, there are several small molecules of RNA (a transfer RNA is in the order of 80 nucleotides long) which behave like executable programs. The three different types of RNA have a different role in the process of translating subsets of DNA (*genes*) into proteins. The understanding of this translation procedure is one of the major scientific achievements of this century. The detailed description of this procedure is obviously out of the scope of this book. A short summary will be given in the next section, for the sake of completeness. The reader interested in a deeper description of protein synthesis is referred to Ref. 7.

4.5.2 Outline of Protein Synthesis

A simplified collection of key points concerning the process follows.

1. The synthesis of proteins involves copying specific regions of DNA (the genes) into polynucleotides of RNA. Molecules of RNA are synthesized by a process known as DNA *transcription*, by which one of the two strands of DNA acts as a template for generating a sequence of RNA, named *messenger* RNA, as symbolically shown in Fig. 4.12.

One can imagine RNA nucleotides floating around the appropriate DNA sequence, where the base-pairing abilities (hydrogen bonding) of incoming nucleotides are tested. Appropriate RNA nucleotides stay near the template long enough to be covalently linked together through the catalytic action of an enzyme known as RNA *polymerase*. When the sequence to be translated is terminated, the mRNA molecule breaks off from the DNA template.

2. Then (in eucaryotes) RNA molecules are spliced to remove *intron* (i.e., noncoding) sequences and only coding sequences (*exons*) are left (see Fig. 4.13).

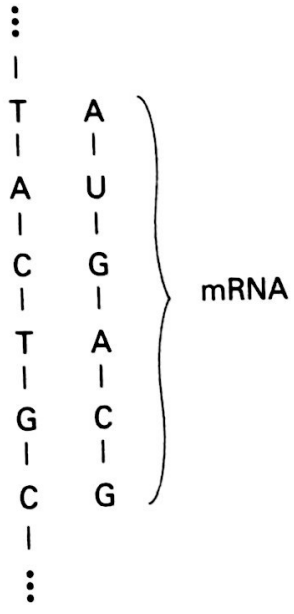

FIGURE 4.12 A symbolic sketch of mRNA synthesis.

3. mRNA moves to the cytoplasm, to interact with molecular devices known as *ribosomes* and to allow protein synthesis. Sequences of mRNA nucleotides are "analyzed" in sets of three and translated into amino acids.

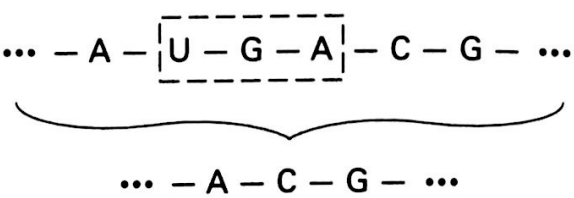

FIGURE 4.13 Symbolic representation of intron removal.

This three-by-three grouping is the basis of the famous *genetic code* which gives the rule for translating a sequence of three bases into a specific amino acid, e.g., {A—C—G} → threonine. Such sequences are named *codons*. There are $4^3 = 64$ possible codons corresponding to about 20 amino acids. This fact implies that the genetic code is *degenerate* (i.e., the code is redundant). The process of "analysis" is physically performed in connection with the ribosomes and it involves various tRNA molecules (each molecule bearing a different amino acid).

The reader is invited to compare this sophisticated chain of events with the most demanding sequence of processes utilized in the field of silicon technology (i.e., doping, oxidation, patterning, etc.). To stimulate further considerations (and readings[7]) we point out the impression that everything seems to be *loosely coupled* in this biological scheme while, on the contrary, everything is *tightly coupled* in any technological scheme.

4.5.3 A Short Note on Computation with DNA

Hopefully, in the previous sections we have been able to transmit to the reader the impression that cell processes, such as protein synthesis, are indeed very sophisticated ones. The physical procedure of DNA synthesis and the generation and recognition of complementary strands suggest by themselves remarkable combinatory properties: a specific sequence of 10 nucleotides will be complementary to exactly 1 over 4^{10} possible sequences and, in a way, it will "sort it out" by double stranding. Interestingly enough, the possibility of "real" computation with DNA has been seriously considered in the 1990s. In accordance with the bioelectronic approach of this book we very briefly expose the reader to this idea, even if, on the basis of the results available up to 1997, its value (i.e., a breakthrough or simply a clever scientific exercise) is not clear yet.

The idea, suggested by L. Adelman[8] in 1994, is to use the enormous parallelism of solution-phase chemistry to solve a computational problem, which is known as the Hamilton path problem. A path through a graph is said to be Hamiltonian if it visits each vertex exactly once. In a way, this is the starting point for solving more complex problems, such as the well-known traveling salesman problem, which looks for the minimum path. Adelman has approached the Hamiltonian path problem in a biological context where each vertex and edge of the graph can be represented by a short synthetic sequence of nucleotides. The binding together of chosen oligonucleotides representing vertices results in DNA molecules that encode the solution to a Hamiltonian in path problem.

In the experiments performed by Adelman, various copies of different oligonucleotides were generated by taking advantage of the DNA *ligation reaction,* catalyzed by the the enzyme *ligase.* Amplification was obtained by using the polymerase chain reaction. The "computation" required 7 days of laboratory work. Further details can be found in Ref. 8.

4.6 PHOSPHOLIPIDS ORGANIZATION

Several fatty acids are present in a living organism. They include: *tryglicerides, steroids* (e.g., cholesterol), and *phospholipids*. Phospholipids are self-assembling molecules which constitute the major component of cell membranes.

Phospholipid structure is characterized by a polar head (containing a phosphoric acid) and two hydrophobic fatty acid tails. This asymmetrical (hydrophilic-hydrophobic) design makes these molecules perfectly suited to self-assemble themselves in solution into organized structures. This is a clear example of the bottom-up scheme pursued by biological organisms to function.

The polar part of a phospholipid easily makes physical bonds with water dipoles and/or ions present in solution. On the contrary, the hydrophobic tails can make only weak hydrophobic bondings with other fatty acid chains.

On the basis of these simple considerations, three major space organizations of phospholipid assemblies can be qualitatively predicted:

1. An assembly of phospholipids gently dissolved on the surface of a water solution will tend to self organize into a *monolayer* with the heads interacting with the solution and all the tails parallel to each other out in the gas phase on the top of the solution. This can be considered as an example of a molecular insulator, a few nanometers thick (typically 2 to 3 nm), which can be deposited on the surface of a solid material appropriately dipped into the solution.

2. An assembly of phospholipids forced inside a water solution will tend to self-organize into small (4 to 6 nm) drops, with the external surface formed by the phosphorous heads and the core by packed, water-excluding, hydrophobic tails. This self-organized structure is named *micelle.*

3. A self-sealing spheroidal bilayer represents a third way of satisfying the hydrophilic/hydrophobic rule. This is a more complex system which achieves the fundamental result of separating an outer water solution from an inner water solution, with two polar (charged) surfaces facing the two solutions. The result is a system (we could call it a "device") known as *liposome.* Liposomes are of relevance for at least two reasons: First, biocompatible liposomes can be loaded with drugs and then injected into the blood. In this way the drug can be released inside the body through the slowly leaking bilayer of the liposome with a time scale that can be, to some extent, programmed by the experimenter. Second, and more important for the purposes of this book, liposomes are the structural basis for all cellular membranes (see Sec. 4.7). The three described self-organized structures are sketched in Fig. 4.14.

Finally, before moving to the membrane structures, let us introduce *glycolipids.* Like phospholipids, these molecules are composed of a hydrophobic region, containing two long hydrocarbon tails, and a polar region which contains one or more sugar residues. Glycolipids belong to the family of *glycoconiugates* which include also proteins covalently linked to sugars. Glycoconiugates play a fundamental role as *receptors* (i.e., *molecular sensors*) on the outer cell membrane.

4.7 CELL MEMBRANE

Every cell is separated from the outside world by a membrane. A cell membrane represents a further step in the bottom-up design, resulting from the assemblage of lipids, pro-

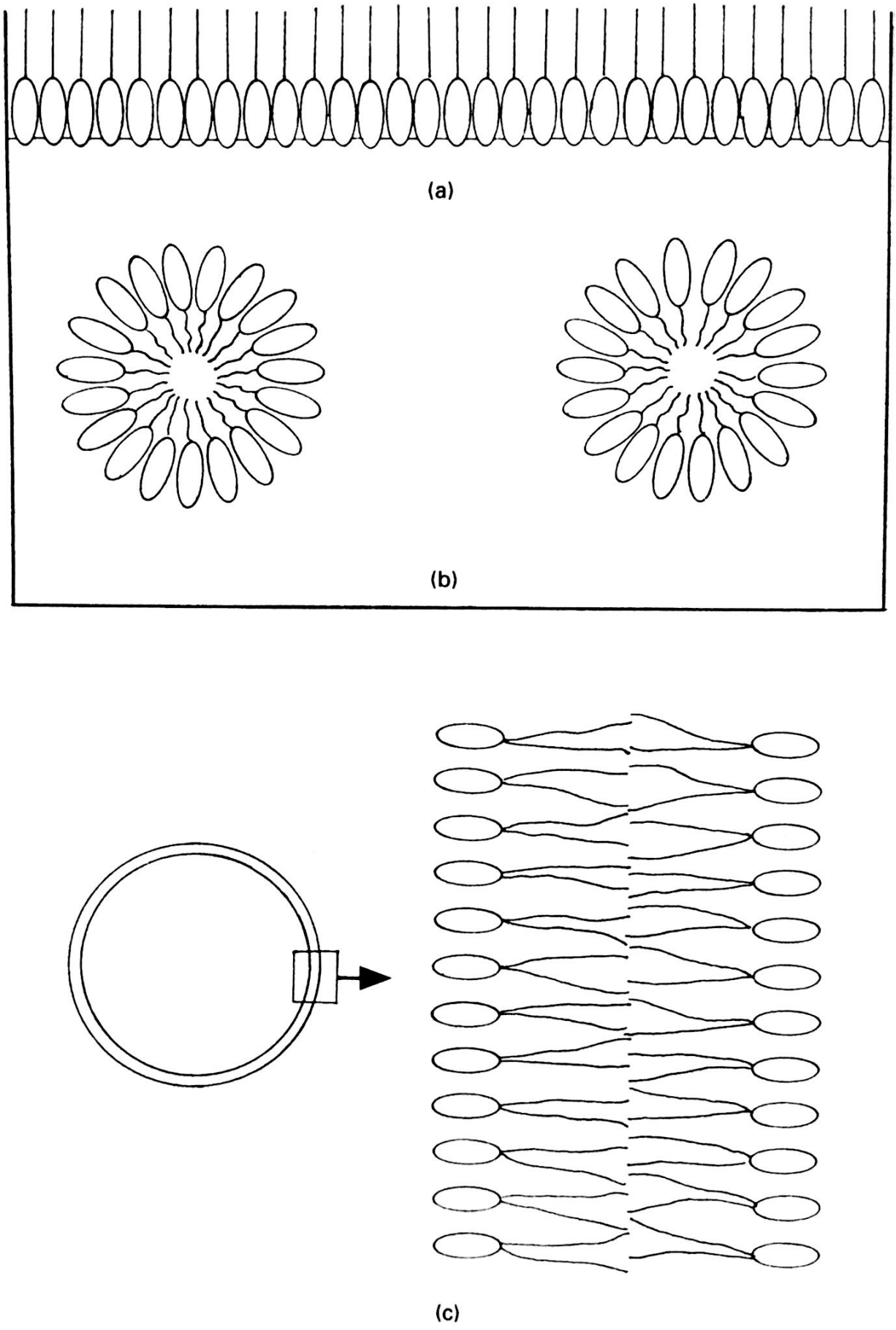

FIGURE 4.14 (*a*) Lipid monolayer, (*b*) micelle, and (*c*) liposome. (*Kindly provided by Maddalena Grattarola.*)

teins, and glycoconiugates. The cell membrane is much more than a simple physical interface. In the language of this book it should be considered a device (better, an array of devices) rather than a kind of material. It is shortly introduced and it will be considered again when dealing with membrane transport (Chap. 7), and with neurons (Chap. 11).

4.7.1 The Cell Membrane Structure

The membrane of a typical cell of a multicellular organism is known as *plasma membrane.* It encloses the cell and maintains the essential differences between the *cytosol* (i.e., the inside of the cell) and the extracellular environment. All the plasma membranes have a common general structure: each is a very thin film of lipid and protein molecules, held together mainly by noncovalent interactions. Cell membranes are dynamic, fluid structures and most of their molecules are able to move about in the plane of the membrane. In some cases this movement can be described as a random one, known as *Brownian motion* (see Chap. 5). The lipid molecules (mostly phospholipids) are arranged as a continuous double layer about 5 nm thick. This is the same structure already described for liposomes. The lipid bilayer furnishes the basic structure of the membrane and plays the role of a relatively impermeable barrier to the passage of most water-soluble molecules. From this point of view, the lipid bilayer can be considered as an *insulator.* Moreover, in consideration of the fact that some of the lipid heads bear a charge and that the membrane is surrounded by two conducting electrolyte solutions, the lipid layer, plus the conducting inner and outer media, can be schematized as a *capacitor.*

Protein molecules "dissolved" in this two-dimensional solvent mediate most of the functions of the membrane, for example, transporting ions across it (*ion channels*), catalyzing membrane-associate reactions (*enzymes*), or acting as sensors of external signals, allowing the cell to change its behavior in response to environmental cues (*receptors*).

Ion channels, enzymes, and receptors act as a multifunctional array of "fluid devices" which continuously reorganize themselves as a function of external physicochemical signals and internal ones, the latter coming from the genes.

Several kinds of phospholipids form, together with cholesterol, the lipid bilayer of any plasma membrane. As a rule, negatively charged lipids are located in the inner monolayer, and therefore there is a significant difference in charge between the two halves of the bilayer.

Another asymmetry in lipid distribution among the two halves is related to glycolipids (see Sec. 4.6). These sugar-containing lipid molecules are found exclusively in the extracellular half of the lipid bilayer, where they are thought to self-associate into microaggregates by forming hydrogen bonds with one another. Microaggregates of glycolipids could have some role in interactions of the cell with its surroundings. The most complex of the glycolipids, the *gangliosides* contain oligosaccharides with one or more sialic acid residues, which give gangliosides a net negative charge.[7] This could imply *electrical effects*: their presence will alter the electrical field across the membrane and the concentrations of ions such as Ca^{2+} at the cell surface. Glycolipids can act in general as membrane receptors.

Figure 4.15 schematically shows the asymmetrical distribution of phospolipids and glycolipids.

4.7.2 Membrane Proteins

Roughly speaking, about 50 percent of the mass of a "typical" plasma membrane is made of proteins. This corresponds to 1 protein molecule for every 50 lipid molecules, which

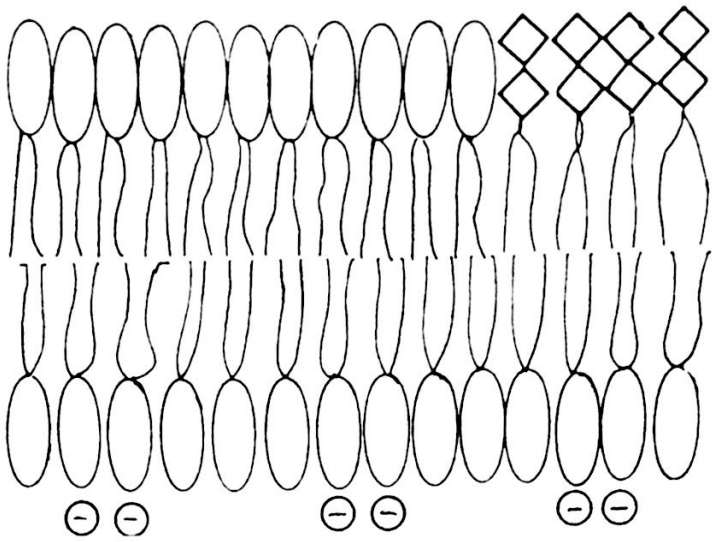

FIGURE 4.15 The asymmetrical distribution of phospholipids and glycolipids in the lipid bilayer. Glycolipids are drawn with double-square polar head groups. Cholesterol (not shown) is thought to be distributed about equally in both monolayers.

are much smaller than the proteins. Like membrane lipids, membrane proteins often have sugar chains attached to them in the extracellular half of the membrane. As a consequence, the surface that the cell presents to the exterior consists largely of carbohydrates, forming a fluid coat known as *glycocalyx.*

Many membrane proteins extend through the lipid bilayer. These proteins present hydrophobic regions (typically organized in α helices) inside the lipid bilayer, interacting with the hydrophobic tails of the lipid molecules, and hydrophilic regions exposed to the electrolyte solution on one or the other side of the membrane. Other membrane proteins are located entirely in the *cytoplasm* (see Sec. 4.8), and are associated with the bilayer by means of covalenty attached fatty acid chains. Similarly, other proteins are entirely exposed to the extracellular side, being attached to the bilayer only by a covalent linkage via sugar groups. All of these proteins, tightly bound to the membrane, are called *integral proteins.*

Some proteins, called *peripheral membrane proteins,* are simply bound to one or the other face of the membrane by noncovalent interactions with other membrane proteins. Most of these proteins can be detached from the membrane by simple exposure to solutions of very high ionic strength and extreme pH, which break protein-protein noncovalent binding but leave the lipid bilayer intact. Figure 4.16 summarizes all the considered cases.

Transmembrane proteins can transport molecules across the membrane, or function on both sides of it. Some cell-surface receptors, for example, are transmembrane proteins that bind signaling molecules in the extracellular space and generate different intracellular signals on the opposite side of the plasma membrane. Several input signals, such as *light* and *neurotransmitters* are recognized and amplified at the plasma membrane level with a common scheme involving similar receptors. These *receptors* are transmembrane proteins with seven helices spanning the lipid bilayer. Light-responding *rhodopsin,* and *acetylcholine*-responding *muscarinic receptors* are two examples of this transmembrane protein family. The signals received by these receptors are *amplified* by a common family of membrane proteins known as *G proteins.* G proteins activate an enzyme at the membrane level, and the final result is the production of a second messenger inside the cell. The described scheme is depicted in Fig. 4.17. Note that receptor, G protein, and enzyme can move independently inside the lipid bilayer.

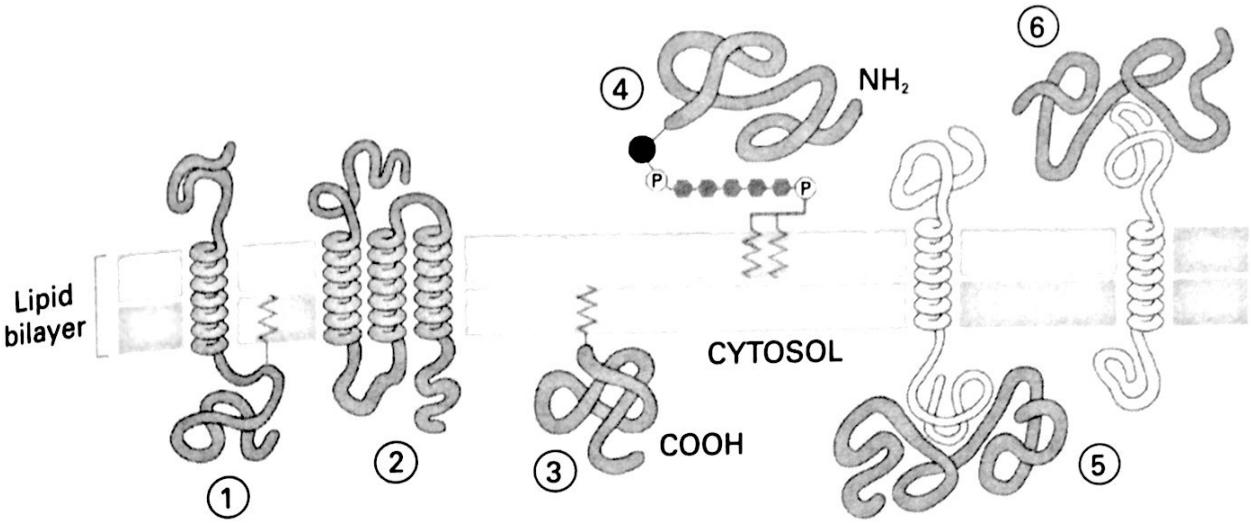

FIGURE 4.16 Six ways in which membrane proteins associate with the lipid bilayer. Most transmembrane proteins are thought to extend across the bilayer as a single α helix (1) or as multiple α helices (2). Some of these "single-pass" and "multipass" proteins have a covalently attached fatty acid chain inserted in the cytoplasmic monolayer (1). Other membrane proteins are attached to the bilayer solely by a covalently attached lipid in the cytoplasmic monolayer (3) or, less often, via an oligosaccharide, to a phospholipid, in the extracellular monolayer (4). Finally, many proteins are attached to the membrane only by noncovalent interactions with other membrane proteins (5 and 6). (*From B. Alberts et al.*[7] *Used by permission.*)

Among the family of transmembrane proteins that transport molecules across the membrane, *ion channels* are protein-based *membrane sensors* which can be activated by chemical or physical inputs. Voltage-gated channels open in response to a change in the intensity of the electric field present inside the plasma membrane. They represent one of the key devices in the operation of neurons, and, in the framework of a "totally mechanistic" viewpoint, they can be considered (together with *synapses*) the molecular basis of thought. A hypothetical view of a voltage-gated channel, taken from the excellent book by Hille,[9] is shown in Fig. 4.18. The cartoon is a kind of mesoscopic view of the channel, where the equivalent volumes of the protein channel are given rather than the detailed sec-

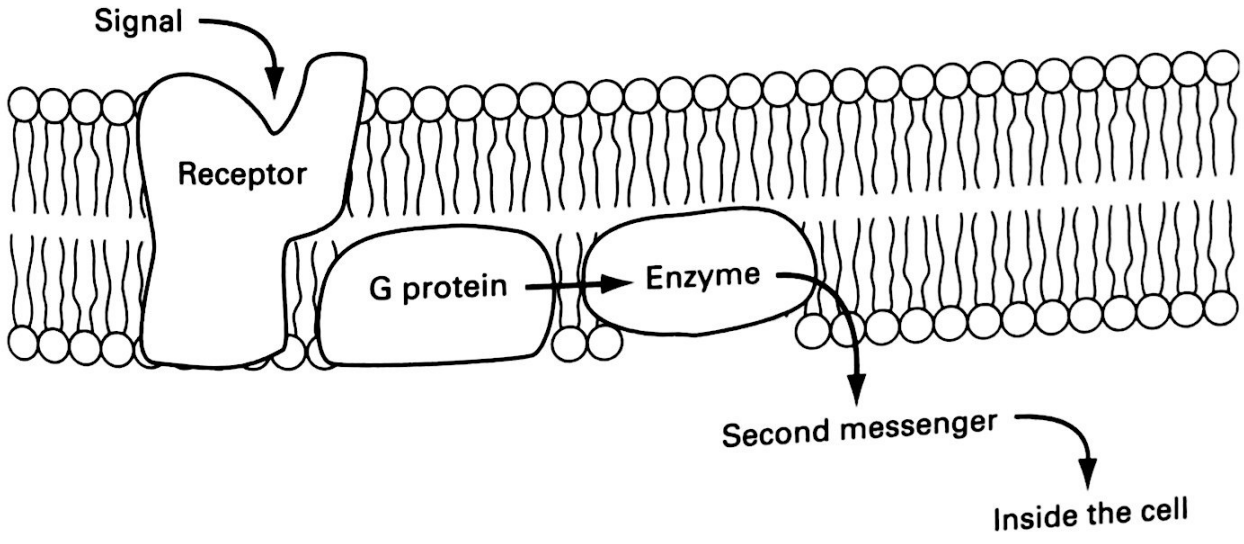

FIGURE 4.17 Sketch of a signaling pathway using G proteins. (*Adapted from B. Hille.*[9] *Used by permission.*)

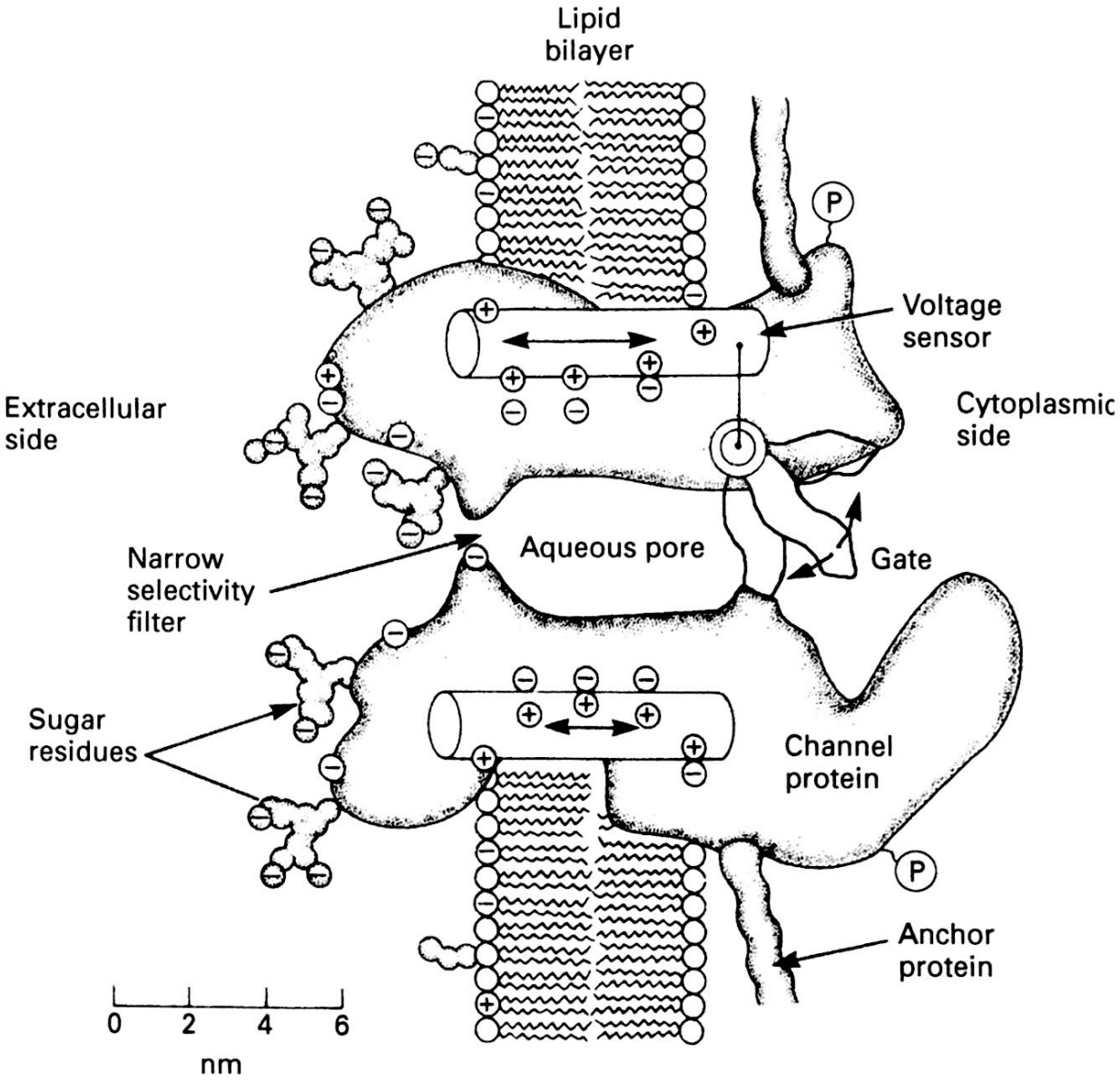

FIGURE 4.18 Working hypothesis for a channel. The channel is drawn as a transmembrane macromolecule with a hole through the center. The external surface of the molecule is glycosylated. (*Adapted from B. Hille.*[9] *Used by permission.*)

ondary structure. The reader should imagine α helix chains repetitively crossing the plasma membrane and connected by long hydrophilic loops projecting into the extracellular and intracellular media.

The protein forming the channel is known to be large, consisting of 1800 to 4000 amino acids, with some hundreds of sugar residues covalently linked to amino acids on the outer surface. The open channel forms a water-filled pore fully extending across the membrane. The pore may narrow in a specific region, where ionic selectivity is established.

Gating requires a conformational change of the pore that moves a "gate" into and out of an occluding position, and it is controlled by a "sensor" (see Fig. 4.18). In the case of a voltage-gated channel, the sensor is a voltage-sensitive one.

We close this section by inviting the engineering-oriented reader to meditate on this system that includes in a volume of a few thousand cubic nanometers a molecular filter (ion selectivity), a molecular sensor, and a molecular actuator (gate).

4.8 AN OVERVIEW OF THE EUCARYOTIC CELL

The *cell* is the functional unit of any living organism. With the exception of bacteria, the cells constituting any organism are called *eucaryotic cells.* They typically comprise an inner region, called *nucleus,* and an exterior region, surrounding the nucleus, called *cytoplasm.* Bacteria also are made of cells (typically a bacterium is just one cell), but these cells are devoid of nucleus. They are named *procaryotic cells.* In principle, any living cell (both eucaryotic and procaryotic) is able to self-replicate. Self-replication may be considered as partial in structures simpler than procaryotic cells, the *viruses,* which are particles consisting of nucleic acids (RNA or DNA) enclosed in a protein coat and able to replicate within a host cell.

Both higher animals and vegetal organisms are made of eucaryotic cells. A "typical" eucaryotic animal cell is structured as follows (see Fig. 4.19):

The *plasma membrane* (see Sec. 4.7) separates the inside of the cell from the outside. The fundamental component of the cytoplasm (the *cytosol*) is a nonideal aqueous solution containing ions (mostly K^+, Na^+, and Cl^-) and organic molecules. The cytoplasm of eucaryotic cells is crossed by a dynamic system of protein filaments known as the *cytoskeleton.* They include *actin filaments* (8 nm diameter), *intermediate filaments* (10 nm diameter), and *microtubules* (25 nm diameter). The cytoskeleton gives the cell shape and the capacity for movement. Flattened membranous sacs and tubes and small membrane-bound vesicles extend throughout the cytoplasm. As a whole, this complex includes the *Golgi apparatus* and the *endoplasmic reticulum.* The Golgi apparatus is involved in the *secretion* of macromolecules. The endoplasmic reticulum (ER) is in physical continuity with the nuclear envelope and it is further divided into smooth and rough ER. Smooth ER function is related to lipid synthesis. The "roughness" of the rough ER is caused by the presence of

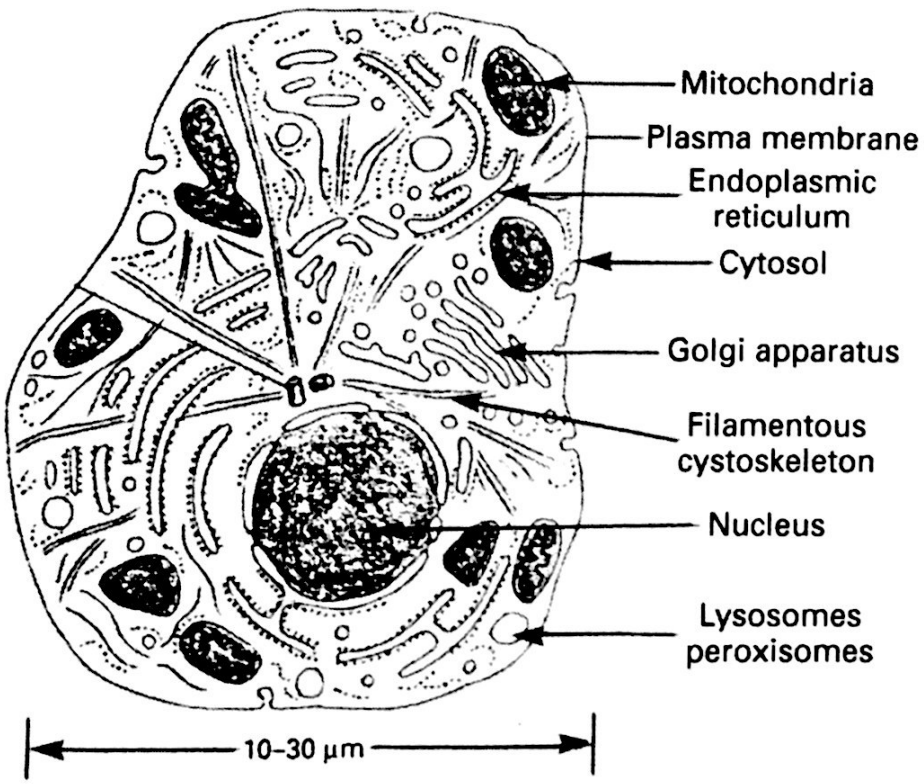

FIGURE 4.19 Section of a generic eucaryotic animal cell.

small (about 50 nm diameter) particles, known as *ribosomes*, which are involved in protein synthesis (see Sec. 4.5.2). The cytoplasm of a metabolically active eucaryotic cell is full of membrane-bound *organelles*, with dimensions in the micrometer range. The most common are *mitochondria*, *lysosomes*, and *peroxisomes*. Mitochondria have intriguing similarities with the procaryotic cell, including the presence of a small amount of "naked" DNA.[7] Mitochondria can be considered the power plants of all eucaryotic cells and they make energy available by combining oxygen with organic material to produce adenosintriphosphate (ATP) molecules.[7] *Lysosomes* are membrane-bound vesicles containing hydrolytic enzymes dealing with intracellular digestions. *Peroxisomes* are membrane-bound vesicles containing oxidative enzymes. Higher plant cells contain also *chloroplasts*, specialized structures dealing with the highly sophisticated cycle of chlorophyll photosynthesis.

The nucleus can be considered the largest organelle in the eucaryotic cell. It is separated from the cytoplasm by a double membrane containing *pores*. Inside the nucleus, DNA is connected in a complex way with basic proteins (*histones*) to form a complex 3-D structure, known as *chromatin*. A small subregion of the nucleus, known as the *nucleolus*, is RNA-rich and involved in the synthesis of ribosomes.

Cells reproduce by duplicating their content and then dividing in two. This process, of course, happens during the development of multicellular eucaryotic organisms (*embryogenesis*) and it also takes place in the adult individual, to replace cells that are lost for several reasons. The capability of self-reproduction is not equally distributed in the cells forming the various *tissues* of an adult individual. The way a cell duplicates its content and then divides is known as the *cell cycle*. The cell cycle has a number of features common to all kind of reproducing (i.e., *cycling*) cells: it is made of *four successive phases*, indicated as G_1, S, G_2, and M phases. The cell cycle begins with a "preparatory period," the G_1 phase, after which the cell enters the S phase. The symbol S stands for DNA synthesis. This is a process of nuclear DNA replication, which is a necessary prerequisite in order to finally produce two daughter cells with the same genetic content of the mother cell. S phase is followed by a gap, the G_2 phase, which somehow allows the cell to prepare to divide. The division process occupies the subsequent *mitotic* (M) phase, during which nuclear chromatin condenses into visible *chromosomes* and the cell eventually physically divides into two new cells (*cytokinesis*). The subsequent cycle starts again with a presynthesis gap, G_1 phase, which allows the cell to monitor its environment and its own size, up to a decisive step that commits it to DNA replication and completion of a new division cycle. It should be noted that at any instant, most of the billions (10^{13} in the human body) cells in an adult complex organism, such as a mammal, are not proliferating, but in a resting state, performing their specialized function while retired from the cell cycle.

A great deal of information concerning the cell cycle of mammalian cells has been obtained by studying cells isolated from the intact animal and growing *in culture*. Cultured cells offer the researcher the opportunity to study living biological systems under controlled environmental conditions produced in the laboratory. Further analysis of this topic is beyond the purpose of the book. The interested reader can find in Ref. 7 a detailed discussion of this topic and of the related topics of *cell lines*, *cell transformation*, and *cell senescence*.

PROBLEMS

4.1 In vacuum a dipole with $l = 0.2$ nm and charge $= 0.5q$ is at a distance $R = 1$ nm from a monovalent cation. Both particles are fixed in space.

a. Calculate the interaction energy for $\theta = 0°$ and $\theta = 90°$.

b. Compare this energy with kT for $T = 300$ K and $T = 1000$ K.