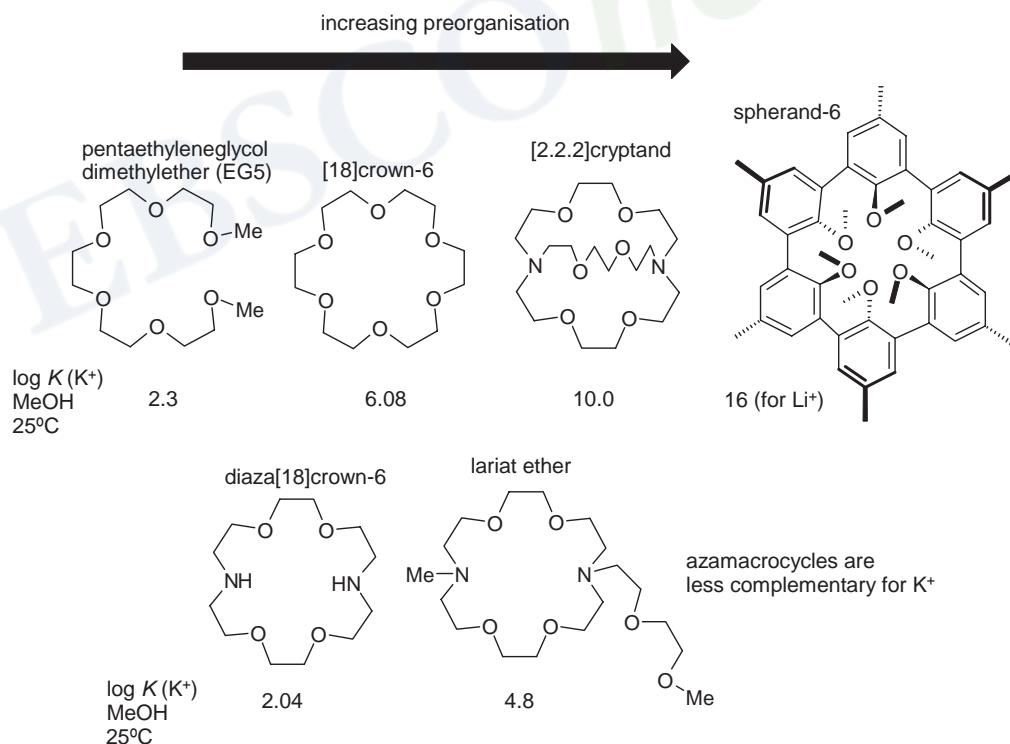


and both complexation and decomplexation are rapid. Solvation enhances the effects of preorganisation since the solvation stabilisation of the unbound host is often greater than the case when it is wrapped around the guest, effectively presenting less surface area to the surrounding medium.

In addition to the degree of host preorganisation, the other principal factor in determining the affinity of a host for a guest is *complementarity*. In order to bind, a host must have binding sites that are of the correct electronic character (polarity, hydrogen bond donor/acceptor ability, hardness or softness *etc.*) to complement those of the guest. Hydrogen bond donors must match acceptors, Lewis acids must match Lewis bases and so on. Furthermore, those binding sites must be spaced out on the host in such a way as to make it possible for them to interact with the guest in the binding conformation of the host molecule. If a host fulfils these criteria, it is said to be *complementary*. The principle of complementarity has been summed up by Donald Cram: 'To complex, hosts must have binding sites which cooperatively contact and attract binding sites of guests without generating strong nonbonded repulsions.'

The combined effects of preorganisation and complementarity are startlingly illustrated by a comparison of the binding constants under standard conditions for the alkali metal complexes shown in Figure 1.12. All of the hosts bind through six ether oxygen atoms. The fairly hard (non-polarisable) oxygen donors are complementary to fairly hard alkali metal cations such as  $K^+$ . However, the stability constants range over nearly 14 orders of magnitude, reflecting the increasing preorganisation of the oxygen atom donor array. The amine nitrogen atoms in some hosts do not significantly enhance the binding because the softer amine is not complementary for alkali metal cations. Thus replacing two



**Figure 1.12** Comparison of the effects of preorganisation and complementarity on the magnitudes of the binding constant of polyether hosts for alkali metal cations. The figure for  $Li^+$  is given for the highly preorganised spherand-6 since it is too small to accommodate  $K^+$ .

oxygen atoms in [18]crown-6 with two secondary amine nitrogen atoms in diaza[18]crown-6 lowers the binding constant to below the value found for the podand EG5.

## 1.7 Thermodynamic and Kinetic Selectivity, and Discrimination

✦ Schneider, H.-J. and Yatsimirsky, A. K., 'Selectivity in supramolecular host-guest complexes', *Chem. Soc. Rev.*, 2008, **37**, 263–277.

The goal of supramolecular host design, both in nature (enzymes, transport proteins *etc.*) and in artificial systems, is the achievement of selectivity; some kind of differentiation of different guests. In the blood, the iron haem transport protein haemoglobin is fine-tuned to selectively take up O<sub>2</sub> in the presence of N<sub>2</sub>, water and CO<sub>2</sub>, and even substances such as CO, which normally bind very strongly to iron. We can readily assess the affinity of a host for a particular receptor by its binding constant (Section 1.4). In thermodynamic terms, selectivity is simply the ratio of the binding constant for one guest over another:

$$\text{Selectivity} = \frac{K_{\text{Guest1}}}{K_{\text{Guest2}}} \quad (1.31)$$

This kind of selectivity tends to be the most easy to achieve because it is highly susceptible to manipulation by intelligent application of concepts such as the lock and key analogy, preorganisation and complementarity, coupled with a detailed knowledge of the host–guest interactions. So, we can say that [18]crown-6, with a binding constant for K<sup>+</sup> of 10<sup>6</sup> M<sup>-1</sup>, is 100-fold selective for K<sup>+</sup> over Na<sup>+</sup>, which it binds with a binding constant of only *ca.* 10<sup>4</sup> M<sup>-1</sup> under the same conditions. There is another kind of selectivity, however, which relates to the rate of transformation of competing substrates along a reaction path. This is *kinetic selectivity* and is the basis for directing the flow of directional processes such as supramolecular (enzymatic) catalysis and guest sensing and signalling. In this sense, it is the guest that is transformed *fastest*, rather than the one that is bound the strongest, that the system is said to be selective for. Indeed, in such time-resolved processes, large binding constants are inhibitory to the system since kinetics are slowed down. Many biochemical enzymes are kinetically selective and examination of their structures reveals that while they are perfectly complementary for the desired (sometimes transitory) state of the guest at any given instant, they are not generally preorganised in a rigid way since this would preclude rapid catalysis. In artificial systems, the engineering of time-resolved selectivity (as in the design of enzyme mimics, Chapter 12) is a much more difficult process since it requires the adaptation of the host to the changing needs of the guest as the system proceeds along its reactive pathway.

We should also distinguish between guest selectivity and inter-guest discrimination. While thermodynamic selectivity relates to the magnitude of observed binding constants, discrimination is applied to the magnitude of other observable results of often highly specific host-guest interactions. Good examples are fluorescent or colorimetric molecular sensing. The guest that is bound most strongly is not necessarily the guest that gives the largest change in colour or in fluorescent emission intensity. This is because the changes in light absorption or emission may result from a particular, guest-specific host-guest interaction, rather than being directly proportional to binding affinity. Thus a host or sensing ensemble may effectively *discriminate* between two potential guests even if their binding constants are similar. The concept of guest discrimination is particularly interesting in the context of binding patterns by arrays of different hosts (for a fuller discussion see Section 11.3.3).<sup>14</sup>

## 1.8 Nature of Supramolecular Interactions

➔ Anslyn, E. V. and Dougherty, D. A., *Modern Physical Organic Chemistry*, University Science Books, Sausalito, CA, USA, 2006, pp. 162–168.

In general, supramolecular chemistry concerns non-covalent bonding interactions. The term ‘non-covalent’ encompasses an enormous range of attractive and repulsive effects. The most important, along with an indication of their approximate energies, are explained below. When considering a supramolecular system it is vital to consider the interplay of all of these interactions and effects relating both to the host and guest as well as their surroundings (*e.g.* solvation, ion pairing, crystal lattice, gas phase *etc.*).

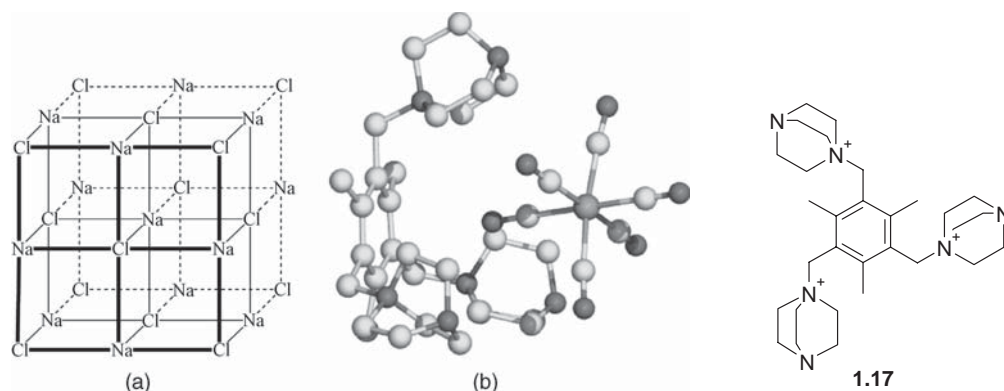
### 1.8.1 Ion–ion Interactions

Ionic bonding is comparable in strength to covalent bonding (bond energy = 100–350 kJ mol<sup>-1</sup>). A typical ionic solid is sodium chloride, which has a cubic lattice in which each Na<sup>+</sup> cation is surrounded by six Cl<sup>-</sup> anions (Figure 1.13a). It would require a large stretch of the imagination to regard NaCl as a supramolecular compound but this simple ionic lattice does illustrate the way in which an Na<sup>+</sup> cation is able to organise six complementary donor atoms about itself in order to maximise non-covalent ion–ion interactions. Note that this kind of lattice structure breaks down in solution because of solvation effects to give species such as the labile, octahedral Na(H<sub>2</sub>O)<sub>6</sub><sup>+</sup>.

A much more supramolecular example of ion–ion interactions is the interaction of the *tris*(diazabicyclooctane) host (**1.17**), which carries a 3+ charge, with anions such as [Fe(CN)<sub>6</sub>]<sup>3-</sup> (Figure 1.13b).<sup>15</sup>

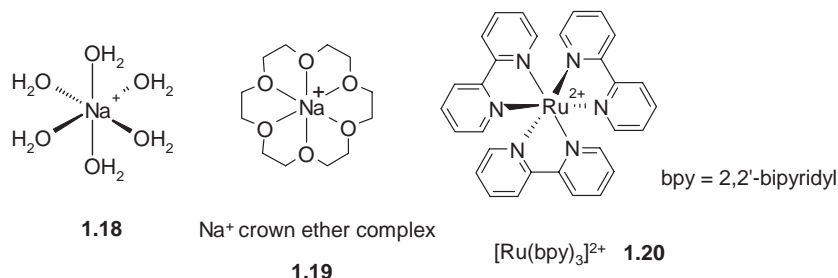
### 1.8.2 Ion–Dipole Interactions

The bonding of an ion, such as Na<sup>+</sup>, with a polar molecule, such as water, is an example of an ion–dipole interaction, which range in strength from *ca.* 50 – 200 kJ mol<sup>-1</sup>. This kind of bonding is seen both in the solid state and in solution. A supramolecular analogue is readily apparent in the structures of the complexes of alkali metal cations with macrocyclic (large ring) ethers termed crown ethers



**Figure 1.13** (a) The NaCl ionic lattice. (b) Supramolecular ion–ion interactions exemplified by the interaction of the organic cation **1.17** with [Fe(CN)<sub>6</sub>]<sup>3-</sup>.

(Chapter 3) in which the ether oxygen atoms play the same role as the polar water molecules, although the complex is stabilised by the chelate effect and the effects of macrocyclic preorganisation. The oxygen lone pairs are attracted to the cation positive charge.



Ion–dipole interactions also include coordinative bonds, which are mostly electrostatic in nature in the case of the interactions of nonpolarisable metal cations and hard bases. Coordinate (dative) bonds with a significant covalent component, as in  $[\text{Ru}(\text{bpy})_3]^{2+}$ , are also often used in supramolecular assembly and, as we will see in Chapters 10 and 11, the distinction between supramolecular and molecular species can become rather blurred.

### 1.8.3 Dipole–Dipole Interactions

Alignment of one dipole with another can result in significant attractive interactions from matching of either a single pair of poles on adjacent molecules (type I) or opposing alignment of one dipole with the other (type II) (Figure 1.14) with energies in the range 5–50  $\text{kJ mol}^{-1}$ . Organic carbonyl compounds show this behaviour well in the solid state and calculations have suggested that type II interactions have an energy of about 20  $\text{kJ mol}^{-1}$ , comparable to a moderately strong hydrogen bond. The boiling point of ketones such as acetone (56 °C), however, demonstrates that dipole–dipole interactions of this type are relatively weak in solution.

### 1.8.4 Hydrogen Bonding

☞ Jeffrey, G. A., *An Introduction to Hydrogen Bonding*, Oxford University Press: Oxford, 1997.

Hydrogen bonding has tremendous effects on molecular properties. It is the strong hydrogen bonding in water that makes its boiling point of 100 °C some 160 °C higher than the heavier  $\text{H}_2\text{S}$ , simply

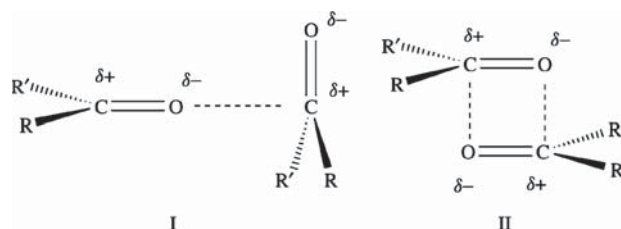
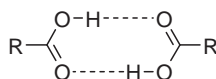


Figure 1.14 Dipole–dipole interactions in carbonyls.

because of the more polar nature of the O–H bonds. Similarly at room temperature (298 K) butanone gas ( $C_4H_8O$ ) which has a hydrogen bonding carbonyl (C=O) functionality is a factor of  $1.7 \times 10^4$  more soluble in water than the non-polar gaseous butane ( $C_4H_{10}$ ). Conversely the inhibition of firefly luciferase activity (the reaction that causes the firefly's glow) by butane is a factor of 74 times more efficient than butanone in water. The increased solvation of the butanone prevents it from blocking the enzyme hydrophobic binding site.

A hydrogen bond may be regarded as a particular kind of dipole–dipole interaction in which a hydrogen atom attached to an electronegative atom (or electron withdrawing group) is attracted to a neighbouring dipole on an adjacent molecule or functional group. Hydrogen bonds are commonly written D–H...A and usually involve a hydrogen atom attached to an electronegative atom such as O or N as the donor (D) and a similarly electronegative atom, often bearing a lone pair, as the acceptor (A). There are also significant hydrogen bonding interactions involving hydrogen atoms attached to carbon, rather than electronegative atoms such as N and O (electronegativities: C: 2.55, H: 2.20, N: 3.04, O: 3.44). Because of its relatively strong and highly directional nature, hydrogen bonding has been described as the ‘masterkey interaction in supramolecular chemistry’.<sup>1</sup> Normal hydrogen bonds typically range in strength from *ca.* 4–60 kJ mol<sup>-1</sup>, although certain highly acidic compounds such as  $HF_2^-$  have hydrogen bond energies up to 120 kJ mol<sup>-1</sup>. An excellent example of hydrogen bonding in supramolecular chemistry is the formation of carboxylic acid dimers (**1.21**), which results in the shift of the  $\nu(OH)$  infrared stretching frequency from about 3400 cm<sup>-1</sup> to about 2500 cm<sup>-1</sup>, accompanied by a significant broadening and intensifying of the absorption. Typically hydrogen bonded O...O distances are 2.50–2.80 Å in length, though interactions in excess of 3.0 Å may also be significant. Hydrogen bonds to larger atoms such as chloride are generally longer, and may be weaker as a consequence of the reduced electronegativity of the larger halide acceptor, although the precise strength of the hydrogen bonds is greatly dependent on its environment. Hydrogen bonds are ubiquitous in supramolecular chemistry. In particular, hydrogen bonds are responsible for the overall shape of many proteins, recognition of substrates by numerous enzymes, and (along with  $\pi$ - $\pi$  stacking interactions) for the double helix structure of DNA (Sections 2.9 and 10.3).



1.21

Hydrogen bonds come in a wide range of lengths, strengths and geometries and can be divided into three broad categories, the properties of which are listed in Table 1.5. A strong interaction is somewhat similar in character to a covalent bond, whereby the hydrogen atom is close to the centre-point of the donor and acceptor atoms. Strong hydrogen bonds are formed between a strong acid and a good hydrogen bond acceptor, for example in the  $H_5O_2^+$  ion or in complexes of ‘proton sponge’ (Section 4.7.3), which are practically linear with the hydrogen atom between the two electronegative atoms. Moderate strength hydrogen bonds are formed between neutral donor and neutral acceptor groups *via* electron lone pairs, for example the self-association of carboxylic acids, or amide interactions in proteins. Moderate hydrogen bond interactions do not have a linear geometry but are slightly bent. Hydrogen bonds commonly deviate from linearity and their angular distribution is influenced by statistical factors. A ‘conical correction’ for statistical effects often appears in the analysis of hydrogen bond angle distributions, particularly from searches of the Cambridge Structural Database (Section 8.4). A linear hydrogen bond requires a fixed position of the hydrogen atom in relation to the acceptor, whereas non-linear hydrogen bonds have many possible positions that form a conical shape around the linear position. Larger bond angles result in a larger cone, and therefore there are more possible positions for the bond

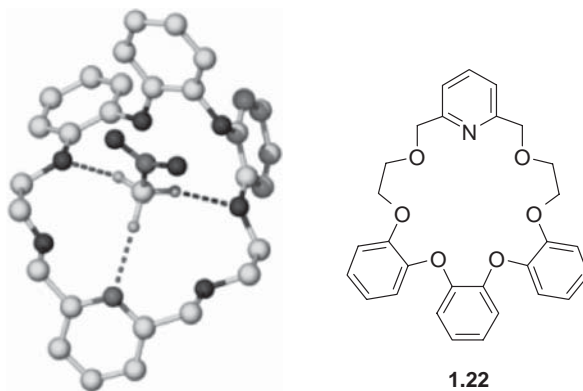
**Table 1.5** Properties of hydrogen bonded interactions (A–H = hydrogen bond acid, B = hydrogen bond base).

	Strong	Moderate	Weak
A–H...B interaction	Mainly covalent	Mainly electrostatic	Electrostatic
Bond energy (kJ mol <sup>-1</sup> )	60–120	16–60	<12
Bond lengths (Å)			
H...B	1.2–1.5	1.5–2.2	2.2–3.2
A...B	2.2–2.5	2.5–3.2	3.2–4.0
Bond angles (°)	175–180	130–180	90–150
Relative IR vibration shift (stretching symmetrical mode, cm <sup>-1</sup> )	25%	10–25%	<10%
<sup>1</sup> H NMR chemical shift downfield (ppm)	14–22	<14	?
Examples	Gas phase dimers with strong acids/bases	Acids	Minor components of bifurcated bonds
	Proton sponge	Alcohols	C–H hydrogen bonds
	HF complexes	Biological molecules	O–H...π hydrogen bonds

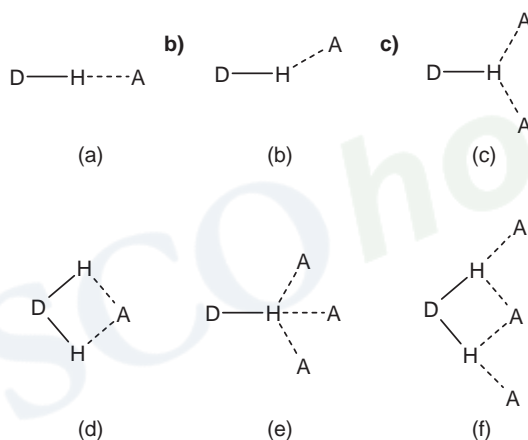
to occur in. In the case of hydrogen bonds between neutral species, it is generally accepted that there is a direct correlation between hydrogen bond strength (in terms of formation energy) and the crystallographically determined distance between hydrogen bond donor and acceptor. Distances tend to be shorter in 'charge assisted' hydrogen bonds involving ions. The strength of similar hydrogen bonds can be very different between various systems and is not necessarily correlated with the Brønsted acidity of the proton (hydrogen) donor. It depends on the type of electronegative atom to which the hydrogen atom is attached and the geometry that the hydrogen bond adopts. Scales of hydrogen bond acidity and basicity exist that partially quantify these effects.<sup>16</sup> A single, strong hydrogen bond per molecule may be sufficient to determine solid-state structure and exert a marked influence on the solution and gas phases, although hydrogen bonding is much more significant in non-polar solvents than in water, where hydrogen bond donors and acceptors are highly solvated and hydrophobic interactions tend to dominate. Weaker hydrogen bonds play a role in structure stabilisation and can be significant when large numbers act in concert. They tend to be highly non-linear and involve unconventional donors and acceptors such as C–H groups, the π-systems of aromatic rings or alkynes or even transition metals and transition metal hydrides (Section 8.9).

While C–H donor hydrogen bonds are at the weaker end of the energy scale of hydrogen bonds, the presence of electronegative atoms near the carbon can enhance significantly the acidity of the C–H proton, resulting in a significant dipole. An elegant example of C–H...N and C–H...O hydrogen bonds is the interaction of the methyl group of nitromethane with the pyridyl crown ether shown in Figure 1.15.<sup>17</sup>

The types of geometries that can be adopted in a hydrogen bonding complex are summarised in Figure 1.16. These geometries are termed *primary hydrogen bond interactions*, this means that there is a direct interaction between the donor group and the acceptor group. There are also *secondary interactions* between neighbouring groups that must be considered. The partial charges on adjacent atoms can either increase the binding strength by virtue of attraction between opposite charges or decrease the affinity due to repulsion between like charges. Figure 1.17 shows two situations in which arrays of hydrogen bond donors and acceptors are in close proximity. An array of three donors (DDD) facing

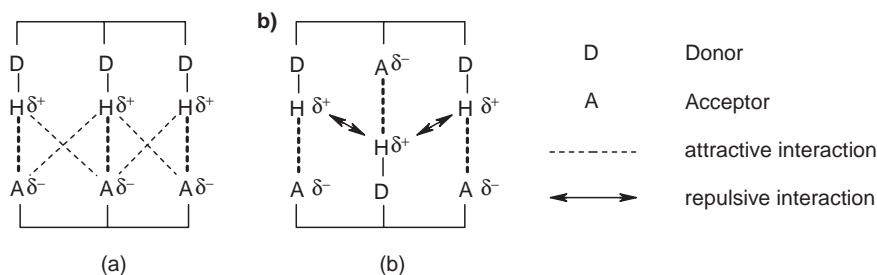


**Figure 1.15** X-ray crystal structure showing C—H $\cdots$ N (2.21Å) and C—H $\cdots$ O (2.41Å, average) hydrogen bonding in a complex of crown ether **1.22** with nitromethane.<sup>17</sup>

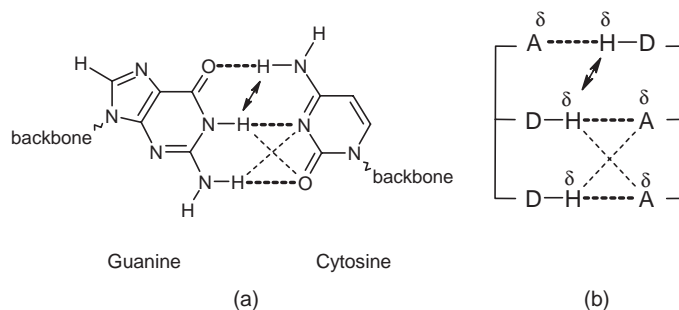


**Figure 1.16** Various types of hydrogen bonding geometries; (a) linear (b) bent (c) donating bifurcated (d) accepting bifurcated (e) trifurcated (f) three centre bifurcated.

three acceptors (AAA) (Figure 1.17a) has only attractive interactions between adjacent groups and therefore the binding is enhanced in such a situation. Mixed donor/acceptor arrays (ADA, DAD) suffer from repulsions by partial charges of the same sign being brought into close proximity by the primary interactions (Figure 1.17b).



**Figure 1.17** (a) Secondary interactions providing attractions between neighbouring groups between DDD and AAA arrays (primary interactions in bold) and (b) repulsions from mixed donor/acceptor arrays (ADA and DAD).



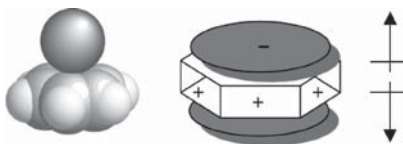
**Figure 1.18** (a) Primary and secondary hydrogen bond interactions between guanine and cytosine base pairs in DNA and (b) a schematic representation.

A real-life example of hydrogen bonding is the double helix of DNA. There are many hydrogen bond donors and acceptors holding base pairs together, as illustrated between the nucleobases cytosine (C) and guanine (G) in Figure 1.18. The CG base pair has three primary interactions (*i.e.* traditional hydrogen bonds) and also has both attractive and repulsive secondary interactions.

### 1.8.5 Cation- $\pi$ Interactions

➔ Ma, J. C., and Dougherty, D., 'The cation- $\pi$  interaction', *Chem. Rev.*, 1997, **97**, 1303–1324.

Transition metal cations such as  $\text{Fe}^{2+}$ ,  $\text{Pt}^{2+}$  *etc.* are well known to form complexes with olefinic and aromatic hydrocarbons such as ferrocene [ $\text{Fe}(\text{C}_5\text{H}_5)_2$ ] and Zeise's salt [ $\text{PtCl}_3(\text{C}_2\text{H}_4)^-$ ]. The bonding in such complexes is strong and could by no means be considered non-covalent, since it is intimately linked with the partially occupied *d*-orbitals of the metals. Even species such as  $\text{Ag}^+ \cdots \text{C}_6\text{H}_6$  have a significant covalent component. The interaction of alkaline and alkaline earth metal cations with C=C double bonds is, however, a much more non-covalent 'weak' interaction, and is suggested to play an important role in biological systems. For example, the interaction energy of  $\text{K}^+$  and benzene in the gas phase is about  $80 \text{ kJ mol}^{-1}$  (Figure 1.19). By comparison, the association of  $\text{K}^+$  with a single water molecule is similar at  $75 \text{ kJ mol}^{-1}$ . The reason  $\text{K}^+$  is more soluble in water than in benzene is related to the fact that many water molecules can interact with the potassium ion, whereas only a few bulkier benzene molecules can fit around it. The interaction of nonmetallic cations such as  $\text{RNH}_3^+$  with double bonds may be thought of as a form of  $\text{X-H} \cdots \pi$  hydrogen bond.



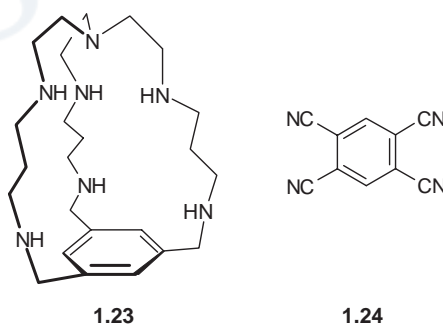
**Figure 1.19** Schematic drawing of the cation- $\pi$  interaction showing the contact between the two. The quadrupole moment of benzene, along with its representation as two opposing dipoles, is also shown.



## 1.8.6 Anion- $\pi$ Interactions

- Berryman, O. B., Bryantsev, V. S., Stay, D. P., Johnson, D. W. and Hay, B. P., 'Structural criteria for the design of anion receptors: The interaction of halides with electron-deficient arenes', *J. Am. Chem. Soc.*, 2007, **129**, 48–58.

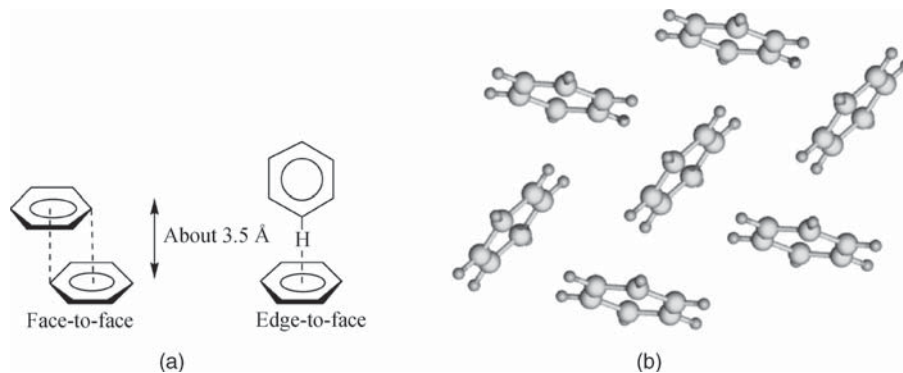
Cation- $\pi$  interactions (see Section 1.8.5) have been known for many years, however it is only relatively recently that there has been interest in anion- $\pi$  interactions. Intuitively, the interaction of an anion with  $\pi$ -electron density seems like it should be repulsive and indeed the affinity of the aromatic ring containing cryptand **1.23** for halides rapidly falls off in the order  $F^- \gg Cl^- > Br^- \sim I^-$  because of anion- $\pi$  repulsions in the case of the larger halides, with all except  $F^-$  showing a constant anion-ring centroid distance of *ca.* 3.7 Å.<sup>18</sup> However, there is a charge difference between an overall neutral aromatic ring and an anion and therefore in principle the possibility exists for an electrostatic attraction. Work by Kochi<sup>19</sup> has shown that anions form stable charge transfer complexes with a variety of electron deficient aromatic compounds such as **1.24**. The formation constants for the anion-aromatic complexes are in the range 1–10 M<sup>-1</sup> and there is a linear correlation between the energy of the charge transfer band in the electronic spectrum and the formal reduction potential of the aromatic compound. This is referred to as a Mülliken correlation and provides strong evidence for the charge transfer nature of the interaction. The charge transfer also results in strong red or yellow colourations for the complexes and a number have been characterised by X-ray crystallography. The crystal structures reveal that the anion sits in a offset fashion at the edge of the aromatic rings rather than above the centroid with anion-carbon distances as short as 2.93 Å for tetrachloro *o*-quinone and Br<sup>-</sup>, shorter than the sum of the a van der Waals radii (3.55 Å). Similar short anion-carbon contacts have been noted for organometallic calixarene derivatives such as **4.34** (see Chapter 4) where the aromatic ring bears a significant positive charge. Anion- $\pi$  interactions have also been implicated as controlling elements in self-assembly reactions of Ag(I) complexes with  $\pi$ -acidic aromatic rings.<sup>20</sup>



## 1.8.7 $\pi$ - $\pi$ Interactions

- Hunter, C. A., Lawson, K. R., Perkins, J. and Urch, C. J., 'Aromatic interactions', *J. Chem. Soc., Perkin Trans. 2* 2001, 651–669.

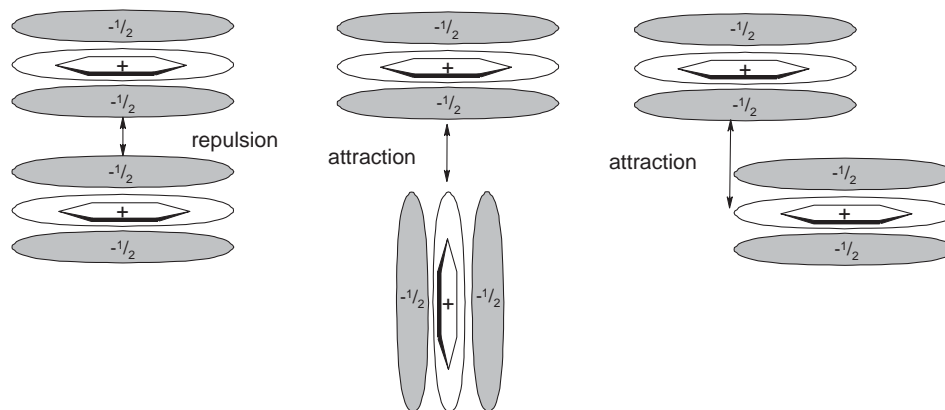
Aromatic  $\pi$ - $\pi$  interactions (sometimes called  $\pi$ - $\pi$  stacking interactions) occur between aromatic rings, often in situations where one is relatively electron rich and one is electron poor. There are two general types of  $\pi$ -interactions: face-to-face and edge-to-face, although a wide variety of intermediate geometries are



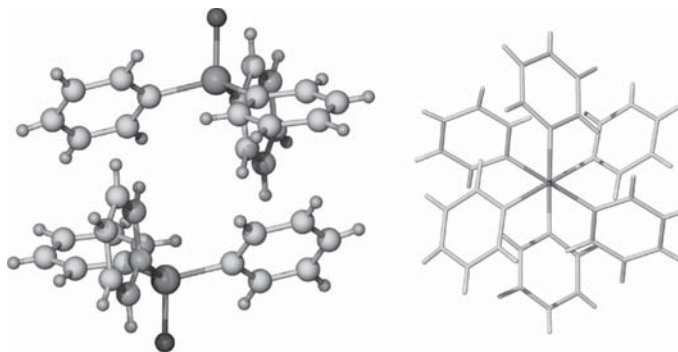
**Figure 1.20** (a) Limiting types of  $\pi$ - $\pi$  interaction. Note the offset to the face-to-face mode (direct overlap is repulsive). (b) X-ray crystal structure of benzene showing herringbone motif arising from edge-to-face interactions.

known (Figure 1.20a). Face-to-face  $\pi$ -stacking interactions are responsible for the slippery feel of graphite and its useful lubricant properties. Similar  $\pi$ -stacking interactions between the aryl rings of nucleobase pairs also help to stabilise the DNA double helix. Edge-to-face interactions may be regarded as weak forms of hydrogen bonds between the slightly electron deficient hydrogen atoms of one aromatic ring and the electron rich  $\pi$ -cloud of another. Strictly they should not be referred to as  $\pi$ -stacking since there is no stacking of the  $\pi$ -electron surfaces. Edge-to-face interactions are responsible for the characteristic herringbone packing in the crystal structures of a range of small aromatic hydrocarbons including benzene (Figure 1.20b).

Sanders and Hunter have proposed a simple model based on competing electrostatic and van der Waals influences, in order to explain the variety of geometries observed for  $\pi$ - $\pi$  stacking interactions and to predict quantitatively the interaction energies. Their model is based on an overall attractive van der Waals interaction (Section 1.8.8), which is proportional to the contact surface area of the two  $\pi$ -systems. This attractive interaction dominates the overall energy of the  $\pi$ - $\pi$  interaction and may be regarded as an attraction between the negatively charged  $\pi$ -electron cloud of one molecule and the positively charged  $\sigma$ -framework of an adjacent molecule. The relative orientation of the two interacting molecules is determined by the electrostatic repulsions between the two negatively charged  $\pi$ -systems (Figure 1.21).



**Figure 1.21** Interacting  $\pi$ -quadrupoles.



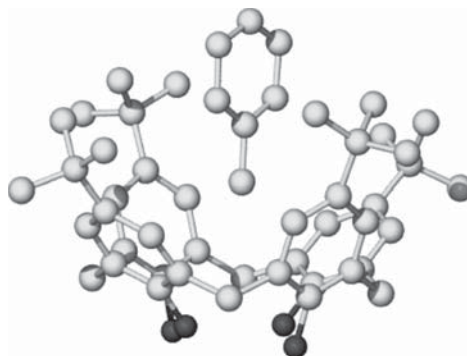
**Figure 1.22** Sixfold edge-to-face phenyl embrace in  $\text{ClGePh}_3$ .<sup>23</sup>

Sanders and Hunter stress the importance of the interactions between individual pairs of atoms rather than molecules as a whole and, while their approach has been relatively successful, there is still a great deal of current debate over the nature of  $\pi$ - $\pi$  stacking interactions. In particular, work involving substituent effects suggests that London dispersion forces might play a more important role than electrostatic interactions.<sup>21</sup>  $\pi$ -stacking interactions are of considerable interest and importance in the crystal structures of both organic and coordination compounds and have a marked influence on solution binding *via* the hydrophobic effect (Section 1.9.1). Edge-to-face  $\pi$ -interactions give rise to common motifs such as the sixfold phenyl embrace often found in compounds containing three or more aromatic rings, such as metal complexes of  $\text{PPh}_3$  (Figure 1.22). A survey of  $\pi$ -interactions in crystalline coordination compounds found that a slipped (parallel displaced) interaction is the most common with the vector between the ring centroids forming an angle of about  $20^\circ$  and aromatic ring centroid-centroid distances of up to 3.8 Å. This parallel-displaced structure is thought to have a contribution from  $\pi$ - $\sigma$  attractive interactions that increases with increasing offset.<sup>22</sup>

## 1.8.8 Van der Waals Forces and Crystal Close Packing

⚡ Kitaigorodskii, A. I., *Organic Chemical Crystallography* (Originally published in Russian by Press of the Academy of Sciences of the USSR, Moscow, 1955). Consultants Bureau: New York, 1961.

Van der Waals interactions arise from the polarisation of an electron cloud by the proximity of an adjacent nucleus, resulting in a weak electrostatic attraction. They are nondirectional and hence possess only limited scope in the design of specific hosts for selective complexation of particular guests. In general, van der Waals interactions provide a general attractive interaction for most 'soft' (polarisable) species with an interaction energy proportional to the surface area of contact. In supramolecular chemistry, they are most important in formation of 'inclusion' compounds in which small, typically organic molecules are loosely incorporated within crystalline lattices or molecular cavities, *e.g.* the inclusion of toluene within the molecular cavity of the *p*-*tert*-butylphenol-based macrocycle, *p*-*tert*-butylcalix[4]arene (Section 6.2.2 and Figure 1.23).<sup>24</sup> Strictly, van der Waals interactions may be divided into dispersion (London) and exchange-repulsion terms. The dispersion interaction is the attractive component that results from the interactions between fluctuating multipoles (quadrupole, octupole *etc.*) in adjacent molecules. The attraction decreases very rapidly with distance ( $r^{-6}$  dependence) and is additive with every bond in the molecule contributing to the overall interaction energy. The exchange-repulsion defines molecular shape and balances dispersion at short range, decreasing with the twelfth power of interatomic separation.



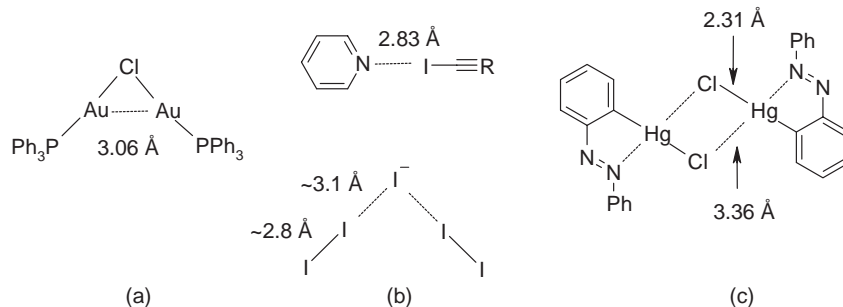
**Figure 1.23** X-ray crystal structure of a typical van der Waals inclusion complex *p*-*tert*-butylcalix[4]arene-toluene.<sup>24</sup>

In examination of solid state (*i.e.* crystal) structures the need to achieve a close packed arrangement is also a significant driving force. This has been summed up in the truism ‘Nature abhors a vacuum’, but, according to the close packing theory of Kitaigorodsky, it is simply a manifestation of the maximisation of favourable isotropic van der Waals interactions. Kitaigorodsky’s theory tells us that molecules undergo a shape simplification as they progress towards dimers, trimers, higher oligomers, and ultimately crystals. This means that one molecule dovetails into the hollows of its neighbours so that a maximum number of intermolecular contacts are achieved, rather like the popular computer game *Tetris*. Very few solid-state structures are known to exhibit significant amounts of ‘empty’ space. Those which do (*e.g.* zeolites, Section 9.2) possess a very rigid framework that is able to resist implosion under what amounts effectively to an enormous pressure differential between atmospheric pressure and the empty crystal pore or channel. Such materials often exhibit very interesting and useful properties in catalysis and separation science.

### 1.8.9 Closed Shell Interactions

☛ Pyykkö, P., ‘Strong closed-shell interactions in inorganic chemistry’, *Chem. Rev.*, 1997, **97**, 597–636.

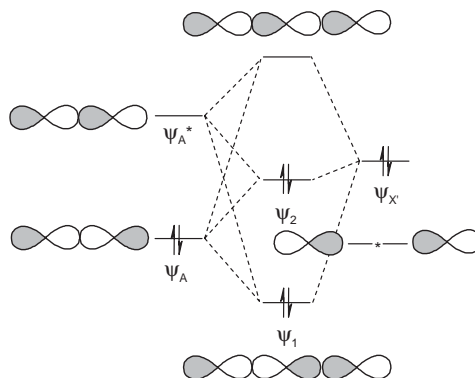
Atoms with unfilled electron shells form strong, covalent bonds. Ions generally have closed valence electron shells but experience strong attractions between oppositely charged pairs. We would not intuitively expect closed shell atoms of neutral or like charges to form significant interactions, however in some cases they do. These interactions are termed closed shell interactions and include secondary bonding interactions,<sup>25</sup> metalphilic interactions<sup>26</sup> and halogen bonding.<sup>27</sup> Closed shell interactions are broadly comparable in strength with moderate strength hydrogen bonds and are thought to arise from electron correlation effects, significantly strengthened by relativistic effects in heavy elements, particularly gold (where they are termed *aurophilic interactions*). Thus closed shell interactions are most pronounced for heavy metals with examples reported for electron configurations from  $d^8$  to  $d^{10}s^2$ , and the heavier halogens with halogen bonding strength decreasing in the order  $I > Br > Cl \gg F$ . Structural studies have shown that halogen bonds of type  $D \cdots X-C$  (where D is an electron-pair donor and X is a halogen electron pair acceptor) have a well-defined, linear geometry ( $160-180^\circ$ ) with  $D \cdots X$  distances considerably less than the sum of the van der Waals radii of D and X. The most obvious example is the  $I^- \cdots I_2$  interaction found in the  $I_3^-$  ion which has an energy of *ca.* 200 kJ mol<sup>-1</sup>, and indeed halogen bonds have been known since the discovery of  $Me_3N \cdots Br_2$  in 1896. The geometries adopted by halogen bonding are influenced by ‘polar flattening’, the anisotropic distribution of electron density about halogen and some other polarisable atoms, however they represent a genuine attractive interaction (see further discussion in Section 8.4)



**Figure 1.24** Examples of closed shell interactions (a) Auophilic interactions in  $[\text{Au}_2(\mu\text{-Cl})(\text{PPh}_3)_2](\text{ClO}_4)$ , (b) halogen bonding in  $\text{pyridine}\cdots\text{I-CCR}$  and  $\text{I}_3^-$  and (c) secondary bonding in  $[\{\text{HgCl}(\text{C}_6\text{H}_4\text{N}_2\text{Ph})\}_2]$ .

Auophilic interactions experience significant relativistic shortening and the  $\text{Au}\cdots\text{Au}$  distances which are in the range 2.8–3.0 Å are typically shorter than silver(I) analogues. The relativistic factor comes from the fact that electrons moving near highly charged nuclei have a velocity close to that of the speed of light and therefore experience a relativistic mass that is larger than their resting mass. The increased mass causes a decrease in the orbital radius and hence a decrease in atomic radius that is particularly pronounced for gold. Auophilic interactions are ubiquitous in linear, 2-coordinate  $\text{Au}(\text{I})$  complexes with many examples having the ‘A-frame’ geometry as in  $[\text{Au}_2(\mu\text{-Cl})(\text{PPh}_3)_2](\text{ClO}_4)$ ,  $\text{Au}\cdots\text{Au} = 3.06$  Å, Figure 1.24.<sup>28</sup> Interestingly in the compound  $[\text{Au}_2(\text{dmpm})_3](\text{ClO}_4)_2$  which exhibits an  $\text{Au}\cdots\text{Au}$  distance of 3.05 Å, it proved possible to measure a Raman vibrational stretching band for the  $\text{Au}\cdots\text{Au}$  bond at  $79\text{ cm}^{-1}$ . The energy of this band *increases* to  $165\text{ cm}^{-1}$  on UV irradiation suggesting that auophilic interaction is strengthened in the excited state.<sup>29</sup>

Secondary bonding (a term coined by Alcock in 1972) is a closed shell interaction of type  $\text{X-A}\cdots\text{X}'$  where X is commonly a heavier halogen or chalcogenide element (Cl, Br, S *etc.*). Secondary bonds closely resemble hydrogen bonds except that A is often a multi-valent heavy atom such as Hg, Tl, Sn, Pb, Sb, Bi, Se or Te instead of hydrogen. The  $\text{X-A}$  bond is a normal covalent bond and while the  $\text{A}\cdots\text{X}'$  is a closed shell interaction involving donation of a non-bonding lone pair on  $\text{X}'$  into an antibonding  $\sigma^*$  orbital of the  $\text{X-A}$  bond, Figure 1.25. Secondary bonding is a very significant interaction in determining the solid state structures of heavy element compounds.



**Figure 1.25** Molecular orbital description of secondary bonding and related interactions showing the  $n \rightarrow \sigma^*$  interaction.

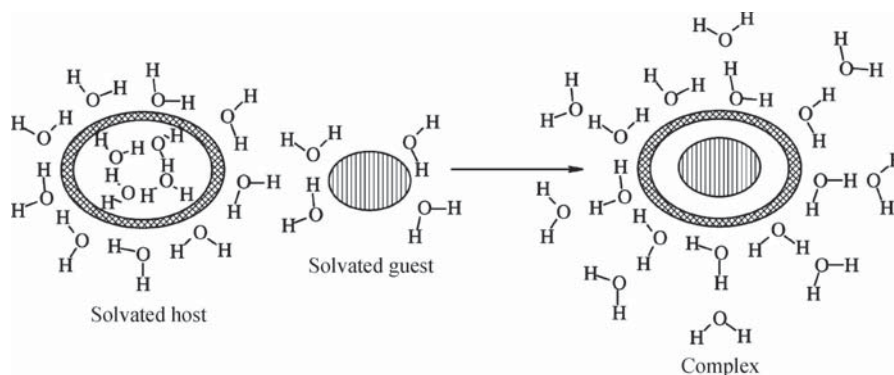
## 1.9 Solvation and Hydrophobic Effects

### 1.9.1 Hydrophobic Effects

☛ Southall, N. T.; Dill, K. A and Haymet, A. D. J., 'A view of the hydrophobic effect', *J. Phys. Chem. B*, 2002, **106**, 521–533.

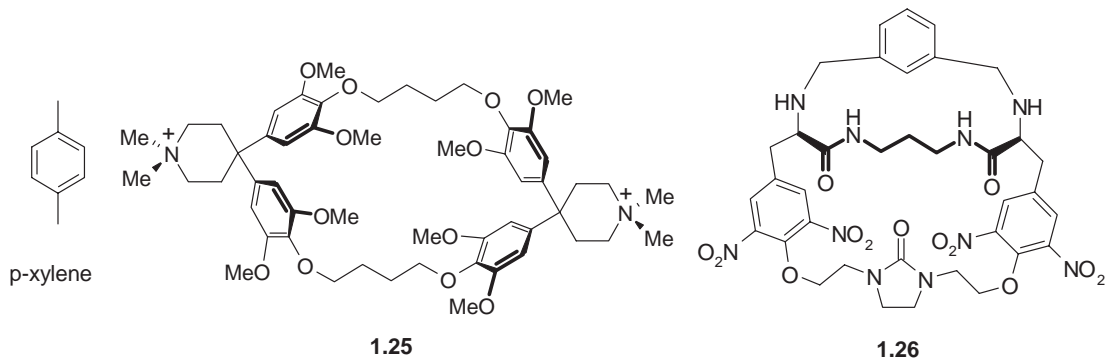
Although occasionally mistaken for a force, hydrophobic effects generally relate to the exclusion from polar solvents, particularly water, of large particles or those that are weakly solvated (*e.g.* via hydrogen bonds or dipolar interactions). The effect is obvious in the immiscibility of mineral oil and water. Essentially, the water molecules are attracted strongly to one another resulting in a natural agglomeration of other species (such as non-polar organic molecules) as they are squeezed out of the way of the strong inter-solvent interactions. This can produce effects resembling attraction between one organic molecule and another, although there are in addition van der Waals and  $\pi$ - $\pi$  stacking attractions between the organic molecules themselves. The hydrophobic effect is very important in biological systems in the creation and maintenance of protein and polynucleotide structure and in the maintenance of phospholipid bilayer cell walls. Hydrophobic effects are of crucial importance in the binding of organic guests by cyclodextrins and cyclophane hosts in water (Chapter 6) and may be divided into two energetic components: enthalpic and entropic. The enthalpic hydrophobic effect involves the stabilisation of water molecules that are driven from a host cavity upon guest binding. Because host cavities are often hydrophobic, intracavity water does not interact strongly with the host walls and is therefore of high energy. Upon release into the bulk solvent, it is stabilised by interactions with other water molecules. The entropic hydrophobic effect arises from the fact that the presence of two (often organic) molecules in solution (host and guest) creates two 'holes' in the structure of bulk water. Combining host and guest to form a complex results in less disruption to the solvent structure and hence an entropic gain (resulting in a lowering of overall free energy). The process is represented schematically in Figure 1.26.

As an example, consider the binding of the guest *p*-xylene by the cyclophane host **1.25** (part of a class described in more detail in Section 6.5). The binding constant in water is  $9.3 \times 10^3 \text{ M}^{-1}$ . At 293K, the complexation free energy,  $\Delta G^\circ$ , is  $-22 \text{ kJ mol}^{-1}$  which divides into a favourable enthalpic stabilisation,  $\Delta H^\circ = -31 \text{ kJ mol}^{-1}$ , and an unfavourable entropic component,  $T\Delta S^\circ = -9 \text{ kJ mol}^{-1}$ . In this case it is the enthalpic contribution to the hydrophobic binding that dominates. The enthalpic contribution is too great to result from attractive forces between host and guest (which experience only weak  $\pi$ -stacking



**Figure 1.26** Hydrophobic binding of organic guests in aqueous solution.

and van der Waals interactions) and thus must arise from specific solvent–solvent forces. In methanol solvent, the enthalpic component is reduced greatly as a result of weaker solvent–solvent interactions.



## 1.9.2 Solvation

➔ Smithrud, D. B., Sanford, E. M., Chao, I., *et al.*, ‘Solvent effects in molecular recognition’, *Pure Appl. Chem.*, 1990, **62**, 2227–2236.

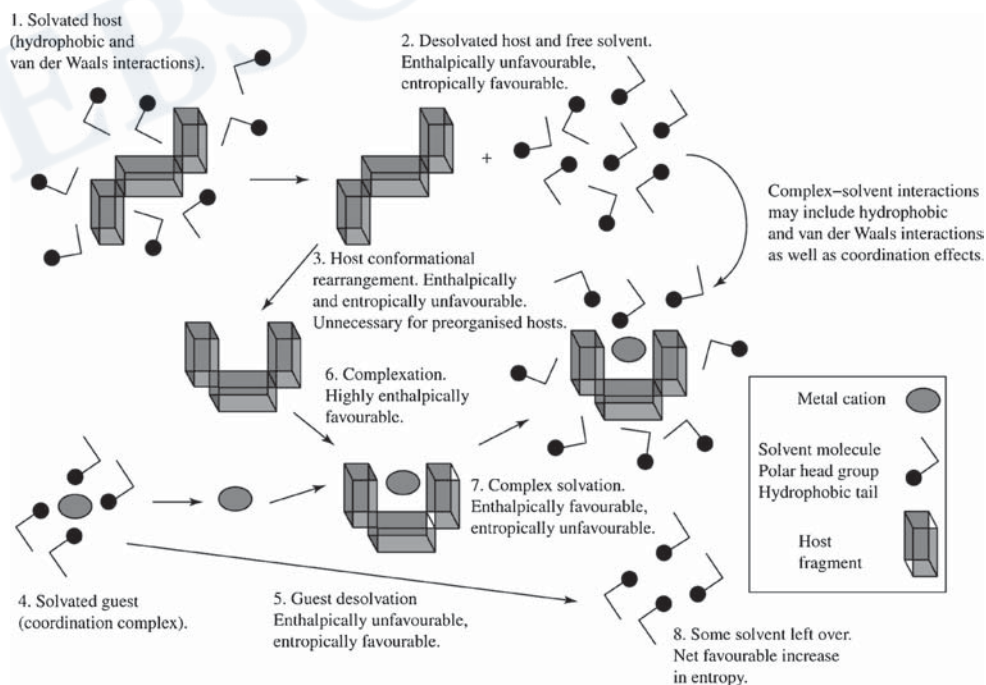
The importance of solvent in supramolecular chemistry can hardly be overstated. In the solid state solvent is often included as a guest in the crystal lattice and usually mediates the nucleation and deposition of a crystalline (or otherwise) compound from solution. In solution all complexation phenomena are in competition with solvation interactions and the solvent is almost invariably in a huge molar excess. Polar solvents, particularly water, compete very effectively for binding sites, particularly hydrogen bonding functionality, making hydrophobic (or solvophobic) effects of paramount importance. In non-polar solvents and in the gas phase specific host–guest dipolar and hydrogen bonding interactions are much more significant. It is thus essentially meaningless to discuss the magnitude of binding constants without mention of solvent and impossible to compare binding constants or even relative affinity across different solvent media. Indeed a common ‘trick’ to differentiate the affinity of a host for various guests is to lower the apparent binding constants by moving to a more competitive (generally more polar) solvent. Thus binding constants that are too high to measure in one solvent become lower and hence experimentally accessible in a more competitive one. An example of the influence of solvent of binding constant is shown in Table 1.6. Binding is clearly enhanced in non-polar solvents with a dramatic maximum value for 1,1,2,2-tetrachloroethane coming from the fact that this solvent is too large to enter the macrobicyclic cavity and hence it does not compete for the guest binding site.

So far we have regarded the host–guest binding process as being the interaction of a more-or-less preorganised host with a naked guest. In reality both host and guest are highly solvated in solution and the solvation stabilisation of the free host may well be significantly different from its interactions to solvent in the complexed state, particularly if there is a significant conformational change (induced fit) on binding. A fuller picture of both the energetics and kinetics of the complexation process must take into account the desolvation of both host and guest upon binding and the resolution of the resulting complex, often with release of free solvent and consequent formation of solvent–solvent interactions, Figure 1.27.

Unfortunately specific solvation effects are very difficult to understand, although molecular mechanics simulations have recently gone some way towards modelling complexation phenomena

**Table 1.6** Influence of solvent on the binding constant of host **1.26** for the neutral organic molecule imidazole (298 K).

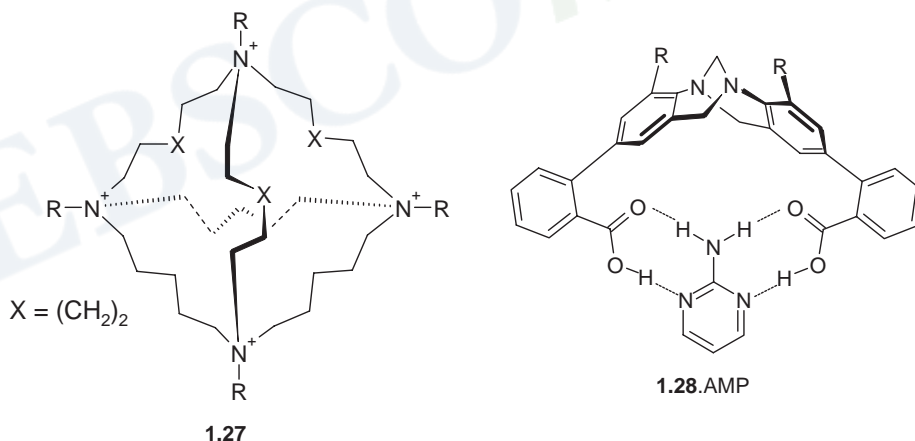
Solvent	Solvent type	$K_{11}$ ( $M^{-1}$ )
$CH_2Cl_2$	Non-polar	240
$CHCl_3$	Non-polar	
	H-bond acidic	490
$CH_3CCl_3$	Non-polar	8161
$CHCl_2CHCl_2$	Non-polar	
	Larger size	128,000
tetrahydrofuran (THF)	Non-polar, coordinating	29.0
2-Me-THF	Non-polar, coordinating	77.0
2,5-Me <sub>2</sub> -THF	Non-polar, coordinating	185
2,2-Me <sub>2</sub> -THF	Non-polar, coordinating	156
2,2,5,5-Me <sub>4</sub> -THF	Non-polar, coordinating	
	Sterically hindered	1067
tetrahydropyran	Non-polar, coordinating	104
1,4-dioxane	Non-polar, coordinating	87
<i>tert</i> -butyl methyl ether	Non-polar, coordinating	566
<i>iso</i> -propanol	Polar, protic	13
<i>tert</i> -butyl alcohol	Polar, protic	66
acetonitrile	Polar, aprotic, coordinating	No association

**Figure 1.27** Solvation considerations during the host-guest complexation of a metal cation.



in solution. For example a molecular dynamics study of halide anion inclusion complexes of a macrotricyclic tetrahedral host **1.27** compared halide binding in 'dry' and 'wet' forms of an ionic liquid solvent. In the 'dry' ionic liquid the uncomplexed halides are surrounded by 4–5 1-butyl-3-methylimidazolium cations whose binding mode changes from hydrogen bonding to facial coordination going from  $F^-$  to  $I^-$ . In the wet solvent the first shell organic ionic liquid cations are all displaced by water molecules, while other halides exhibit a mixture. The solvation of the host and its halide complexes mainly involves  $PF_6^-$  anions in the 'dry' medium, and additional water molecules in the 'wet' ionic liquid. The calculations predict that the anion binding selectivity is different in the two different media. In the 'dry' ionic liquid,  $F^-$  is preferred over the other halides but in aqueous solution **1.27** is selective for  $Cl^-$ . In the 'wet' ionic liquid, there is no  $F^- / Cl^-$  selectivity, highlighting the importance of even small amounts of water on the complexation selectivity.<sup>30</sup>

An experimental comparison of the effect of 'wet' and 'dry' solvent has been carried out for host **1.28** which is capable of binding 2-aminopyrimidine (APy) to give a 1:1 complex **1.28**·APy. The binding constant in dry chloroform is up to  $2.4 \times 10^4 M^{-1}$ , although this decreases significantly with temperature. Saturation of the chloroform with water does not result in a significant diminution of the host-guest affinity, as measured by the overall complexation free energy. This is because the competition for the binding sites by the water (which decreases the enthalpic contribution to the binding) is compensated by a more favourable entropy term associated with the release of water upon organisation of the complex (*cf.* the hydrophobic effect in neat water). A fortuitous case of cancellation of enthalpic and entropic effects.<sup>31</sup>



## 1.10 Supramolecular Concepts and Design

### 1.10.1 Host Design

➔ Lehn, J.-M. 'Perspectives in supramolecular chemistry—from molecular recognition towards self-organisation', *Pure Appl. Chem.*, 1994, **66**, 1961–1966.

In order to design a host that will selectively bind a particular guest, we make use of the chelate and macrocyclic effects as well as the concept of complementarity (matching of host and guest steric and electronic requirements) and, crucially, host preorganisation.

The first step in host design is a clear definition and careful consideration of the target. This leads immediately to conclusions about the properties of the new host system. If a metal cation is to be the guest (Chapter 3), then its size (ionic radius), charge density and ‘hardness’ (Section 3.1.2) are important (e.g. soft donor atoms such as sulfur are suited to the binding of soft guests such as  $\text{Hg}^{2+}$ ,  $\text{Ag}^+$  and  $\text{Pb}^{2+}$ ). For anion complexation (Chapter 4), these factors also affect spherical anions such as chloride, bromide *etc.*, but for nonspherical anionic guests, other factors such as shape, charge and hydrogen bond donor characteristics come into play. Organic cations and anions may require hosts with both hydrophilic and hydrophobic regions, while neutral molecule guests may lack specific ‘handles’ such as polar groups that can strongly interact with the host.

Having defined parameters such as the required host size, charge, character of the donor atoms *etc.*, the intellectual process of ligand tailoring can begin. The key concept in this process is organisation. Host–guest interactions occur through binding sites. The type and number of binding sites must be selected in a fashion that is most complementary to the characteristics of the binding sites of the guest (recall the definition of a guest as the partner with divergent binding sites), and these binding sites must be arranged on a (usually) organic scaffold or framework of suitable size to accommodate the guest. Binding sites should be spaced somewhat apart from one another to minimise repulsions between them, but arranged so that they can all interact simultaneously with the guest. The more favourable interactions there are, the better. The most stable complexes are generally obtained with hosts that are preorganised for guest binding, thus where there is no entropically and enthalpically unfavourable rearrangement on binding that reduces the overall free energy of complexation. Such hosts are ideally ‘sinks’ for their guests in which binding is entirely irreversible. This kind of complexation is ideal, for example, for the removal of toxic metal ions from polluted water. Hosts that bind guests less strongly (*i.e.* there is some equilibrium between bound and unbound species) find applications as sensors and carriers in which event sequences such as ‘bind–detect–release’ or ‘bind–transport–release’ are needed.

The nature of the organic framework of the host itself, whether lipophilic or hydrophilic, plays a fundamental role in host behaviour. This determines the solubility characteristics of the host and its complexes. The thickness of the ligand and the ease of access to the binding pocket, cavity or cleft affect both thermodynamic stability and binding kinetics. Addition of side arms may enhance lipophilicity (e.g. long alkyl chains) or encourage interaction with some external entity such as a polymer support or biomolecule. Such hosts are used to transport radioactive isotopes to targeted regions of the human body for medical radioimaging purposes or to develop artificial ‘enzyme mimics’ (Chapter 12).

Overall, the thorough application of these very broad principles has been generalised into what has been referred to as ‘complete coordination chemistry’, encompassing both supramolecular and classical (Werner) inorganic coordination chemistry.

### 1.10.2 Informed and Emergent Complex Matter

➔ Lehn, J.-M., ‘Toward complex matter: Supramolecular chemistry and self-organization’, *Proc. Nat. Acad. Sci. USA* 2002, **99**, 4763–4768.

We saw in Figure 1.2c that supramolecular chemistry is not just about solid state or solution host–guest chemistry but increasingly emphasises self-assembly and the construction of multi-nanometre scale devices and ultimately materials based on nanometre-scale components (a nanometre is  $10^{-9}$  of a metre). Strict supramolecular self-assembly (Chapter 10) involves the spontaneous formation of a multi-component aggregate under thermodynamically controlled conditions based on information encoded within the individual building blocks (referred to as ‘tectons’) themselves. The aggregate might comprise only one kind of molecule (as in the multiple copies of the same protein that comprise

the coat of many viruses such as the tobacco mosaic virus) or more than one type. In the latter case the different components are usually mutually complementary. Strict self-assembly implicitly carries with it the notion of reversibility of inter-component bond formation in that the final aggregate is the most thermodynamically stable structure under the prevailing conditions. This means that there is an inbuilt error-checking mechanism – malformed aggregates are less stable than the true minimum energy structure and decompose and are reassembled correctly if sufficient time is allowed to pass. Self-assembly processes may be regarded as the result of a series of supramolecular templated steps that build the aggregate. In turn, the interactions and synergy of self-assembled components, leading to the emergence of collective behaviour and properties, may be regarded as *self-organisation*. The hierarchical sequence of *templation* leading to *self-assembly* leading to *self-organisation* represents a supramolecular information concept in which all the information needed to produce a complex, functioning device or material with some measure of sophisticated responsiveness to external stimuli is contained in the molecular components themselves.

When we talk of the *emergence* of a self-organised system we are talking very specifically about complex properties arising over time as a result of non-linear and perhaps unpredictable interactions between the molecular components. *Emergence* is a powerful, if controversial topic<sup>32</sup> that is part of the field of complexity science and cuts across a very broad range of disciplines. Emergence is in some sense the opposite of *convergence*, another possible outcome of templation and self-assembly, in which a single stable structure results from the interactions of the molecular components. A good example of an emergent structure in the everyday world is the complex network of tunnels and vents that regulate the environment within a termite mound (Figure 1.28). The mound is not planned by the termites and does not arise from a predictable template. It emerges from the individual, synergic efforts of the individual termites over time. From a chemical point of view it is possible to exert a high degree of control over the structure and shape and hence information content of small molecules (*i.e.* around 1 nm or less size) using chemical synthetic techniques. Self-organisation then leads to multi-nanometre scale structures and systems that, because of their complexity and large number of components, cannot be made in a linear one-step-at-a-time fashion. Emergent structures and properties arising from the self-assembly and self-organisation of molecules on a multi-nanometre scale are ubiquitous in Biochemistry (*e.g.* self-replication and enzymic catalysis) and are increasingly being studied by chemists in artificial, abiotic (= non biochemical) systems.

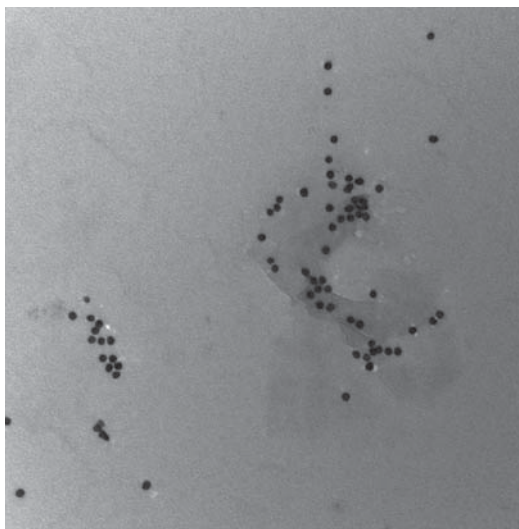


**Figure 1.28** A termite mound in Australia's Northern Territory. Termite mounds are classic examples of Natural emergent structures.

### 1.10.3 Nanochemistry

➔ Ozin, G. and Arsenault, A. *Nanochemistry*, Royal Society of Chemistry: Cambridge, 2005.

We can refer to the synthesis and study of chemical systems with features and functionality on the multi-nanometre length scale as *nanochemistry* and materials with features of size of the order of 1–100 nm as *nanomaterials*. Very broadly there are two approaches to the nanoscale dimension – ‘synthesising-up’ and ‘engineering-down’. The engineering down approach includes the latest in modern techniques for producing electronic components and originates in a bulk sense. Engineering down to the nanoscale (*nanotechnology*) involves doing the same sorts of things that an engineer or artisan does on a macroscopic scale but using specialised techniques in order to miniaturise. In contrast the synthesising-up approach (*nanochemistry*) is modelled on biology, particularly biological self-assembly, and aims to produce nanoscale functional components (perhaps with molecular device or molecular scale computing applications in mind) by chemical synthesis. Indeed the very first reports of functional molecular computing using supramolecular species have already begun to appear.<sup>33</sup> Geff Ozin of the University of Toronto, one of the leading proponents of nanochemistry and the synthesis of nanomaterials has defined nanomaterials as materials whose properties change according to their size, or the size of their components. An excellent example is gold *nanoparticles*. Nanoparticles are tiny fragments of a material such as metallic gold. They are typically more or less spherical in shape (Figure 1.29) and may have a regular, faceted crystalline morphology and structure (in which case they are referred to as *nanocrystals*). Nanoparticles are typically 2–30 nm in radius and very interestingly often exhibit very intense colours. You may already have seen such colours in red or purple suspensions of gold colloids. The colour arises from a visible absorption termed a *surface plasmon resonance* absorption and it arises from the collective motion of free electrons around across the surface of the nanoparticle. Crucially the wavelength of this absorption (and hence the observed colour) depend on the size of the nanoparticle.



**Figure 1.29** TEM micrograph showing gold nanoparticles. Note the regular shape and uniform size distribution. Scale bar = 100 nm.

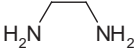
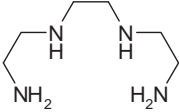
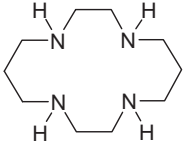
We can also include in the definition of nanochemistry molecular systems exhibiting designed or emergent nanometer-scale features, functionality or properties. For example, interlocked molecules termed *catenanes* can exhibit complex nanoscale molecular motions somewhat analogous to the mechanical interlocking of gears (Section 11.5). As we will see in Chapter 15 there are also hybrid systems coupling molecular hosts to nanomaterials such as nanoparticles. It is also possible to carry out nanochemistry using a variety of scanning probe microscopy based techniques such as Scanning Tunnelling Microscopy (STM) and Atomic Force Microscopy (AFM). Because they operate on the nanometer scale these techniques give unique insight into molecular behaviour. In addition scanning probe microscope tips may be used to carry out molecular manipulations and redox chemistry. Tips may even be modified by attaching a probe molecule and using molecular recognition to examine a surface. Chemical reactions may even be carried out by physically pushing individual molecules into close mutual proximity using an STM tip.<sup>34</sup> Chemical assembly thus represents one facet of the preparation and study of nanoscale materials and there is a continuum of increasing complexity and decreasing predictability and control between molecular and materials properties. Within this continuum we can make a useful distinction between situations where an observed property is an emergent consequence of the interactions between molecules or ions and where it is a distinctive molecular property. In this book we aim to show the full series of steps from the grass roots of simple intermolecular interactions through to emergent nanomaterials and nanochemical systems.

## Summary

- Supramolecular Chemistry involves the chemistry of the non-covalent bond.
- Non-covalent bonds include ionic and dipolar interactions, hydrogen bonds, aromatic interactions ( $\pi-\pi$ , cation- $\pi$  and anion- $\pi$ ), closed shell interactions and van der Waals interactions.
- Supermolecules generally comprise a host component with convergent binding sites and a guest component with divergent binding sites.
- In solid-state host-guest or clathrate compounds the guest is included within a gap in the packing of host molecules.
- Self-assembling systems do not involve hosts and guests but rather self-complementary molecules or complementary partners (tectons).
- The affinity of a host for a guest is measured by the binding constant; selective hosts have a high binding constant for one particular guest.
- Supramolecular chemistry makes use of 'generalised' coordination chemistry and binding cooperativity.
- The traditional picture of 'lock-and-key' binding is generally less appropriate than an induced-fit approach.
- Perhaps the most important concepts in supramolecular host design and preorganisation and complementarity, which encapsulate more traditional concepts such as the macrocyclic effect.
- Solvation and desolvation effects are of tremendous importance in assessing binding equilibria.
- Modern supramolecular chemistry is progressing towards concepts such as molecular information content and intermolecular interaction algorithms, leading to the self-organisation of increasingly complex, informed matter with emergent properties that arise spontaneously in the multi nanometre length scale.

## Study Problems

- 1.1 Thermodynamic parameters for the reaction of  $\text{Cu}^{2+}_{(\text{aq})}$  with various ligands are given below (aqueous solution, 25 °C). Use these data to calculate the binding constants ( $\log K$ ) for the resulting 1:1 metal-to-ligand complexes. Explain the differences in stability observed.

Ligand	$\Delta H^\circ$ (kJ mol <sup>-1</sup> )	$T\Delta S^\circ$ (kJ mol <sup>-1</sup> )
	-105	7.1
	-90.4	24.3
	-76.6	64.0

- 1.2 Give a concise definition of the term 'supramolecular chemistry'. Explain the distinction between molecular and supramolecular interactions. Illustrate your answer with examples of supramolecular interactions and discuss their relative importance.
- 1.3 Draw up a relative scale of the strengths of non-molecular interactions using the information given in Section 1.8. Correlate this with a second scale of the importance of these interactions in supramolecular design of host for metal cations, taking into account factors such as directionality, ease of incorporation into host frameworks, and propensity of binding enhancement *via* multiple binding sites. Would this ranking change if you were to design hosts for anions or neutral organic molecules? What interactions might be important in designing a host for the following species: methane, benzene, methanol, phenol, ammonia, Cl<sup>-</sup>, Na<sup>+</sup> and Ni<sup>2+</sup>?
- 1.4 Using the timeline given in Table 1.1 suggest what may be some of the most important discoveries in supramolecular chemistry. Why do you think it has taken so long for the topic to evolve as a separate discipline?

## Suggested Further Reading

- Atwood, J. L. and Steed, J. W. (eds), *Encyclopedia of Supramolecular Chemistry*, Marcel Dekker, New York, NY, USA, 2004.
- Atwood, J. L. and Steed, J. W. (eds.), *Organic Nanostructures*, Wiley-VCH: Weinheim, 2008.
- Anslyn, E. V. and Dougherty, D. A., *Modern Physical Organic Chemistry*, University Science Books: Sausalito, California, USA, 2006.
- Balzani, V., Credi, A. and Venturi, M. *Molecular Devices and Machines: A Journey into the Nanoworld*, Wiley-VCH: Weinheim, 2003.
- Cragg, P. J., *Practical Supramolecular Chemistry*, John Wiley & Sons, Ltd., Chichester, UK, 2006.
- Desiraju, G. R. (ed.), *The Crystal as a Supramolecular Entity*, John Wiley & Sons, Ltd., Chichester, UK, 1996.
- Haiduc, I. and Edelman, F. T., *Supramolecular Organometallic Chemistry*, Wiley-VCH, Weinheim, Germany, 1999.
- Lehn, J.-M., *Supramolecular Chemistry*, VCH, Weinheim, Germany, 1995.
- Lehn, J.-M., Atwood, J. L., Davies, J. E. D., MacNicol, D. D. and Vögtle, F. (eds), *Comprehensive Supramolecular Chemistry*, Pergamon, Oxford, UK, 1996.
- Ozin, G. A. and Arsenault, A. C., *Nanochemistry*, The Royal Society of Chemistry, Cambridge, UK, 2005.

- Schneider, H.-J. and Yatsimirski, A. K., *Principles and Methods in Supramolecular Chemistry*, Wiley-VCH, Weinheim, Germany, 2000.
- Sessler, J. L., Gale, P. A. and Cho, W.-S., *Anion Receptor Chemistry*, Royal Society of Chemistry: Cambridge, 2006.
- Schalley, C. A. (ed.), *Analytical Methods in Supramolecular Chemistry*, Wiley-VCH: Weinheim, 2007.
- Steed, J. W., Turner, D. R. and Wallace, K. J., *Core Concepts in Supramolecular Chemistry and Nanochemistry*, John Wiley & Sons, Ltd., Chichester, UK, 2007.

## References

1. Lehn, J.-M., *Supramolecular Chemistry*. 1 ed.; VCH: Weinheim, 1995;
2. Cram, D. J., Preorganisation – from solvents to spherands. *Angew. Chem., Int. Ed. Engl.* 1986, **25**, 1039–1134.
3. Lloyd, N. C., Morgan, H. W., Nicholson, B. K., Ronimus, R. S., The Composition of Ehrlich's Salvarsan: Resolution of a century-old debate. *Angew. Chem. Int. Ed.* 2005, **44**, 941–944.
4. Koshland, D. E., The Key-Lock Theory and the Induced Fit Theory. *Angew. Chem., Int. Ed.* 1995, **33**, 2375–2378.
5. Bencini, A., Bianchi, A., Garcia-España, E., Micheloni, M., Ramirez, J. A., Proton coordination by polyamine compounds in aqueous solution. *Coord. Chem. Rev.* 1999, **188**, 97–156.
6. Hynes, M. J., EQNMR: A computer program for the calculation of stability constants from nuclear magnetic resonance chemical shift data. *J. Chem. Soc., Dalton Trans.* 1993, 311–312.
7. Szalay, L., Farkas, V., Vass, E., Hollósi, M., Móczár, I., Pintér, Á., Huszthy, P., Synthesis and selective lead(II) binding of achiral and enantiomerically pure chiral acridono-18-crown-6 ether type ligands. *Tetrahedron Asymmetry* 2004, **15**, 1487–1493.
8. Sessler, J. L., Gross, D. E., Cho, W.-S., *et al.*, Calix[4]pyrrole as a chloride anion receptor: Solvent and counteraction effects. *J. Am. Chem. Soc.* 2006, **128**, 12281–12288.
9. Williams, D. H., Westwell, M. S., Aspects of weak interactions. *Chem. Soc. Rev.* 1998, **27**, 57–63.
10. Metzger, A., Lynch, V. M., Anslyn, E. V., A synthetic receptor selective for citrate. *Angew. Chem., Int. Ed. Engl.* 1997, **36**, 862–865.
11. Rebek Jr., J., Binding forces, equilibria, and rates—new models for enzymic catalysis. *Acc. Chem. Res.* 1984, **17**, 258–264.
12. Hamacek, J., Borkovec, M., Piguet, C., Simple thermodynamics for unravelling sophisticated self-assembly processes. *Dalton Trans.* 2006, 1473–1490.
13. Cabbiness, D. K., Margerum, D. W., Macrocyclic effect on the stability of copper(II) tetramine complexes. *J. Am. Chem. Soc.* 1969, **91**, 6540–6541.
14. Folmer-Andersen, J. F., Kitamura, M., Anslyn, E. V., Pattern-based discrimination of enantiomeric and structurally similar amino acids: An optical mimic of the mammalian taste response. *J. Am. Chem. Soc.* 2006, **128**, 5652–5653.
15. Garratt, P. J., Ibbett, A. J., Ledbury, J. E., O'Brien, R., Hursthouse, M. B., Malik, K. M. A., Molecular design using electrostatic interactions. 1. Synthesis and properties of flexible tripodand tri- and hexa-cations with restricted conformations. Molecular selection of ferricyanide from ferrocyanide. *Tetrahedron* 1998, **54**, 949–968.
16. Laurence, C.; Berthelot, M., Observations on the strength of hydrogen bonding. *Persp. Drug Disc. Des.* 2000, **18**, 39–60.
17. Weber, E., Franken, S., Puff, H., Ahrendt, J., Enclave inclusion of nitromethane by a new crown host—X-ray crystal-structure of the inclusion complex and host selectivity properties. *J. Chem. Soc., Chem. Commun.* 1986, 467–469.
18. Ilioudis, C. A., Tocher, D. A., Steed, J. W., A highly efficient, preorganized macrobicyclic receptor for halides based on CH $\cdots$  and NH $\cdots$ anion interactions. *J. Am. Chem. Soc.* 2004, **126**, 12395–12402.
19. Rosokha, Y. S., Lindeman, S. V., Rosokha, S. V., Kochi, J. K., Halide recognition through diagnostic 'anion- $\pi$ ' interactions: Molecular complexes of Cl $^-$ , Br $^-$ , and I $^-$  with olefinic and aromatic  $\pi$  receptors. *Angew. Chem., Int. Ed.* 2004, **43**, 4650–4652.
20. Schottel, B. L., Chifotides, H. T., Shatrak, M., *et al.*, Anion- $\pi$  interactions as controlling elements in self-assembly reactions of Ag(I) complexes with  $\pi$ -acidic aromatic rings. *J. Am. Chem. Soc.* 2006, **128**, 5895–5912.
21. Kim, E.-I., Paliwal, S., Wilcox, C. S., Measurements of molecular electrostatic field effects in edge- to-face aromatic interactions and CH- $\pi$  interactions with implications for protein folding and molecular recognition. *J. Am. Chem. Soc.* 1998, **120**, 11192.
22. Janiak, C., A critical account on  $\pi$ - $\pi$  stacking in metal complexes with aromatic nitrogen-containing ligands. *J. Chem. Soc., Dalton Trans.* 2000, 3885–3896.

23. Prince, P. D., McGrady, G. S., Steed, J. W., Weak interactions induce asymmetry in the crystal structures of triaryl derivatives of group 14 elements. *New J. Chem.* 2002, **26**, 457–461.
24. Andreetti, G. D., Ungaro, R., Pochini, A., Crystal and molecular structure of cyclo{quarter[(5-t-butyl-2-hydroxy-1,3-phenylene)methylene]} toluene (1:1) clathrate. *J. Chem. Soc., Chem. Commun.* 1979, 1005–1007.
25. Alcock, N. W., Countryman, R. M., Secondary bonding .1. Crystal and molecular-structures of  $(C_6H_5)_2 I_X$  (X=Cl, Br, or I). *J. Chem. Soc., Dalton Trans.* 1977, 217–219.
26. Laguna, A., Fernández, E. J., López-de-Luzuriaga, J. M., 'Auophilic interactions', in *Encyclopedia of Supramolecular Chemistry*, Atwood, J. L., Steed, J. W., eds. Marcel Dekker: New York, 2004; Vol. 1, pp. 82–87.
27. Metrangolo, P., Resnati, G., 'Halogen bonding', in *Encyclopedia of Supramolecular Chemistry*, Atwood, J. L., Steed, J. W., eds. Marcel Dekker: New York, 2004; Vol. 1, pp. 628–635.
28. Jones, P. G., Sheldrick, G. M., Uson, R., Laguna, A.,  $\mu$ -Chloro-bis(triphenylphosphine)digold(I) perchlorate dichloromethane solvate. *Acta Crystallogr., Sect. B* 1980, **36**, 1486–1488.
29. Leung, K. H.; Phillips, D. L.; Mao, Z.; Che, C. M.; Miskowski, V. M.; Chan, C. K., Electronic excited states of  $[Au_2(dmpm)_3](ClO_4)_2$  (dmpm= bis(dimethylphosphine)methane) *Inorg. Chem.* 2002, **41**, 2054–2059.
30. Chaumont, A., Wipff, G., Halide anion solvation and recognition by a macrotricyclic tetraammonium host in an ionic liquid: a molecular dynamics study. *New J. Chem.* 2006, **30**, 537–545.
31. Adrian, J. C., Wilcox, C. S., Chemistry of synthetic receptors and functional-group arrays .15. Effects of added water on thermodynamic aspects of hydrogen-bond-based molecular recognition in chloroform. *J. Am. Chem. Soc.* 1991, **113**, 678–680.
32. Goldstein, J., Emergence as a construct: history and issues. *Emergence: Complexity and Organization* 1999, **1**, 49–72.
33. Collier, C. P., Wong, E. W., Belohradsky, M., *et al.*, Electronically configurable molecular-based logic gates. *Science* 1999, **285**, 391–394.
34. Hla, S. W., Bartels, L., Meyer, G., Rieder, K. H., Inducing all steps of a chemical reaction with the scanning tunneling microscope tip: Towards single molecule engineering. *Phys. Rev. Lett.* 2000, **85**, 2777–2780.



# 2

## The Supramolecular Chemistry of Life

*'Nature that fram'd us of four elements,  
Warring within our breasts for regiment,  
Doth teach us all to have aspiring minds:  
Our souls, whose faculties can comprehend  
The wondrous Architecture of the world.....'*

Christopher Marlowe (1564–1593), *Conquests of Tamburlaine*

EBSCOhost®

## 2.1 Biological Inspiration for Supramolecular Chemistry

8 Nelson, D. L. and Cox, M. M., *Lehninger Principles of Biochemistry*. 4th ed.; W. H. Freeman: New York, 2004.

Much of the inspiration and origins of supramolecular chemistry comes from the chemistry found in living biological systems. Sometimes incredibly complex, sometimes elegantly simple, Nature has evolved an enormous amount of highly specific, hierarchical, selective and cooperative chemistry that enables living systems to maintain themselves in a dynamic equilibrium with their environment and to feed, respire, reproduce and respond to external stimuli. In biological chemistry, the supramolecular hosts are the receptor sites of enzymes, genes, antibodies of the immune system, and ionophores. The guests are substrates, inhibitors, co-factors, drugs or antigens. These components variously exhibit supramolecular properties such as molecular recognition, self-assembly, self-organisation, self-replication and kinetic and thermodynamic complementarity. The vast majority of these properties rely upon supramolecular interactions such as coordination (ion–dipole) bonds, hydrogen bonds and  $\pi$ – $\pi$  stacking discussed in Section 1.8. Biological systems are, therefore, supramolecular systems *par excellence*. A great deal of effort in supramolecular chemistry has been expended in attempts to model or mimic biological processes such as the catalysis of organic chemical reactions by enzymes, or the selective transport of metal cations or molecular substrates such as O<sub>2</sub>. As part of this process, our understanding of biological systems has grown enormously, but it is also fair to say that the molecular and supramolecular chemistry of human endeavour, as they stand now, are still a long way away in scale, scope and functionality from their biochemical analogues. The fact that Nature exhibits such a rich and efficient natural supramolecular chemistry is, however, an enormous encouragement and motivation to continue to seek ever more sophisticated abiotic (nonbiological) analogues and indeed to attempt to develop synthetic systems capable of carrying out transformations or possessing properties not found naturally. In this chapter, we give a brief and highly selective overview of some of the more important biological chemistry of relevance to supramolecular chemists by way of a very brief introduction to the extensive synthetic systems that we will be looking at in the rest of this book. Subsequent chapters will deal with the ways in which synthetic and model systems mimic these biological processes and how insight has been gained into biochemistry by the study of supramolecular compounds, as well as the enormous diversity of entirely non-biological supramolecular chemistry.

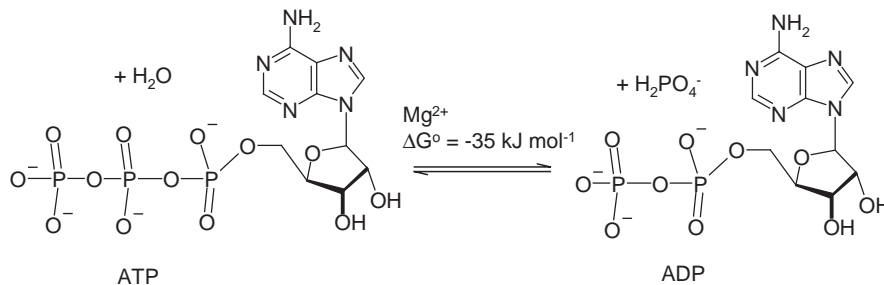
## 2.2 Alkali Metal Cations in Biochemistry

### 2.2.1 Membrane Potentials

Energy is vital to life. Plants get energy from sunlight (photosynthesis), humans get energy from food, which we oxidise to CO<sub>2</sub> and water. Energy is used in respiration, a process by which energy from food is transformed and stored as the chemical bond energy of ATP (adenosine triphosphate). Strictly, ATP has a 4– ionic charge, balanced by alkaline and alkaline earth metal cations. In biological notation this is often omitted. ATP is capable of long-term energy storage and is transported to any areas where energy is needed to drive endergonic (energy-consuming) reactions such as muscle contraction. The energy is released from ATP by a class of enzymes called ATPases, of which Na<sup>+</sup>/K<sup>+</sup>-ATPase is perhaps the most important example. One mole of ATP releases 35 kJ of energy, according to the

---

*Supramolecular Chemistry, 2nd edition* J. W. Steed and J. L. Atwood  
© 2009 John Wiley & Sons, Ltd



**Scheme 2.1** Energy-releasing dephosphorylation of ATP to ADP and dihydrogen phosphate.  $\text{Mg}^{2+}$  acts as a catalyst for the reaction.

reaction shown in Scheme 2.1. Note that while the ATP molecule is relatively complex, it is only the triphosphate tail that is changed during the course of the reaction. Breaking of the terminal phosphate ester P—O bond gives the dihydrogen phosphate anion ( $\text{H}_2\text{PO}_4^-$ , often referred to as ‘inorganic phosphate’,  $\text{P}_i$ ) and adenosine diphosphate, ADP.

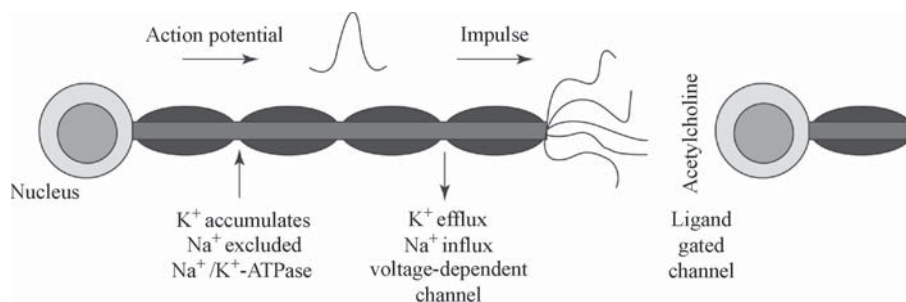
The  $\text{Na}^+/\text{K}^+$ -ATPase enzyme is an example of a *transmembrane* enzyme, *i.e.* an enzyme that exists in the phospholipid membrane (wall) of a biological cell. As part of the process of consuming ATP, the  $\text{Na}^+/\text{K}^+$ -ATPase enzyme transports the alkali metal cations  $\text{Na}^+$  and  $\text{K}^+$  from one side of the cell membrane to the other. Effectively, the enzyme scavenges  $\text{Na}^+$  from the inside of the cell and transports it to the outside, against the prevailing concentration gradient. Simultaneously,  $\text{K}^+$  is transported into the cell. Thus, in the intracellular fluid there is a high concentration of  $\text{K}^+$ ; outside there is a high concentration of  $\text{Na}^+$  (Table 2.1). This uneven distribution of alkali metal cations across the cell membrane is a highly important and necessary feature and results in a transmembrane electrical potential, rather like a battery. This potential difference across the cell is used, amongst other things, in information transfer in nerve cells (Figure 2.1).

The actual amount of charge separation across a cell membrane is very small (the number of  $\text{M}^+$  ions is equal on either side of the membrane). Such a potential difference could, in principle, have been set up by separating  $\text{Na}^+$  and  $\text{Cl}^-$  across a membrane. However, such an actual separation of oppositely charged ions would require much more energy because of the large electrostatic forces between such ions. In fact, the resulting chemical potential arising from the different identities of the alkali metal ions ( $\text{Na}^+$  and  $\text{K}^+$ ) is sufficient to generate the required signal.

The most important requirement for utilisation of this kind of ionic diffusion as a means to information transfer is the maintenance of the non-equilibrium ionic concentration gradient. This is a relatively unstable state – it requires energy to counteract the natural entropy-increasing flow back to equilibrium. This is best illustrated by the *pump storage* model. Ions are actively ‘pumped’ through

**Table 2.1** Some examples of biochemical  $\text{Na}^+$  and  $\text{K}^+$  distributions.

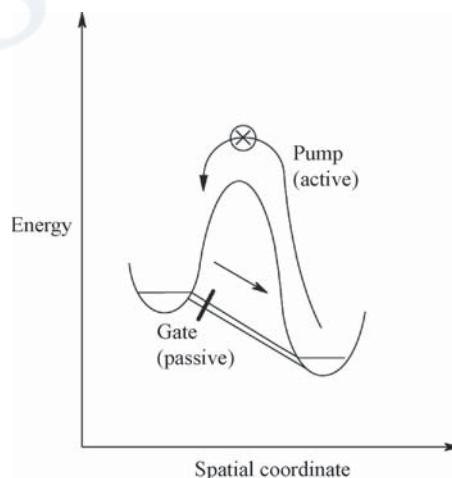
Location	Concentration/mmol $\text{kg}^{-1}$	
	$\text{K}^+$	$\text{Na}^+$
Human intracellular fluid ( <i>e.g.</i> erythrocytes)	92	11
Human extracellular fluid ( <i>e.g.</i> blood plasma)	5	152
Squid nerve (inside)	300	10
Squid nerve (outside)	22	440



**Figure 2.1** Mode of signal transduction of nerve system. Under concentration gradients, generated by  $\text{Na}^+/\text{K}^+$ -ATPase, opening of an ion channel causes the passive efflux of  $\text{K}^+$  and the influx of  $\text{Na}^+$  resulting in a small burst of electrical current (nerve impulse) and a change in the membrane potential. At the end of the nerve cell (axon) the electrical signal is transformed into a chemical signal by triggering the ejection of a hormone such as acetylcholine (Section 2.7). The hormone, in turn, triggers the opening of a ligand-gated ion channel in the next nerve axon and restarts the nerve impulse as an electrical current by allowing passive flow of  $\text{K}^+$  and  $\text{Na}^+$  across the next membrane.

the biological membrane against the concentration gradient until a stationary, non-equilibrium state is reached. Stimulation results in the rapid, passive flow of ions back to equilibrium *via* the operation of *gate functions* (Figure 2.2).

The importance of maintaining precise concentration gradients is highlighted by the severe effects of metabolic disorders involving alkali metal cations. For example, high sodium intake is linked intimately with the development of high blood pressure; on the other hand, aged



**Figure 2.2** The pump storage model. Metal ions are actively pumped by  $\text{Na}^+/\text{K}^+$ -ATPase (an energy-consuming process) from regions of low concentration to regions of high concentration against the concentration gradient until a dynamic non-equilibrium state is reached in which active pumping is balanced by accidental diffusion. Upon activation by the appropriate hormone, selectively gated ion channels open allowing the passive (and therefore rapid) flow of ions back towards equilibrium, resulting in current flow.

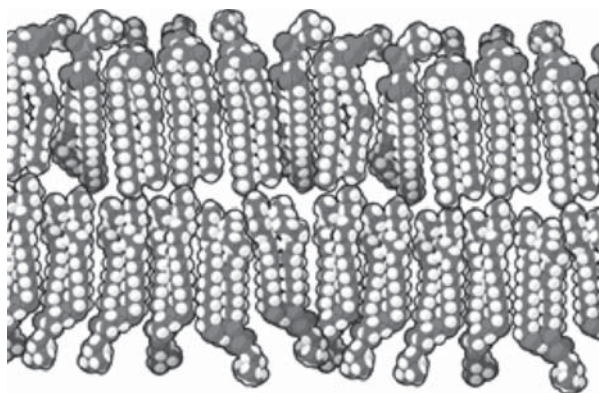
organisms have some difficulty in preventing the excretion of the very labile  $K^+$  because of disturbed membrane permeabilities.

## 2.2.2 Membrane Transport

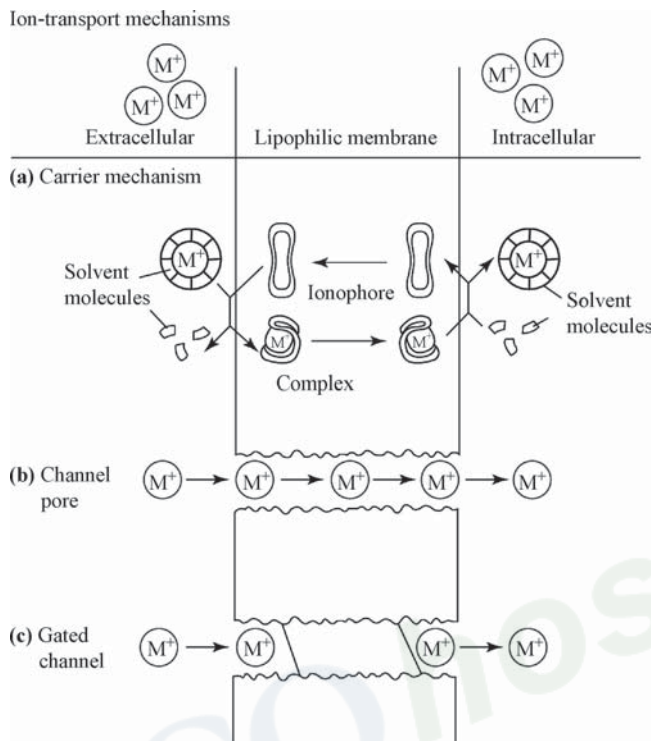
So how does an alkali metal get from the inside of the cell to the outside? A cell membrane consists of hydrophilic (water soluble) phosphate head groups attached to a long lipid (fatty) tail and is thus an example of an amphiphile (Section 13.2.1). In the body's aqueous environment, the hydrophilic head groups are attracted to the surrounding medium (*e.g. via* hydrogen bonding and dipolar interactions) while the organic tail is repelled. This results in a bilayer arrangement in which the organic components are all hidden away from the solvent while the hydrophilic portions face out. Anything that is to pass through the cell wall must therefore be able to pass this lipophilic (fat-soluble) region (Figure 2.3).

Sodium and potassium cations are not at all lipophilic. They cannot effectively diffuse through the cell wall unless something makes them lipophilic or a nonlipophilic pathway is created for them. There are two main possible methods of such passive cation transport along a concentration gradient: transport by some kind of lipophilic carrier, or controlled passage through a hydrophilic channel in the membrane (Figure 2.4).

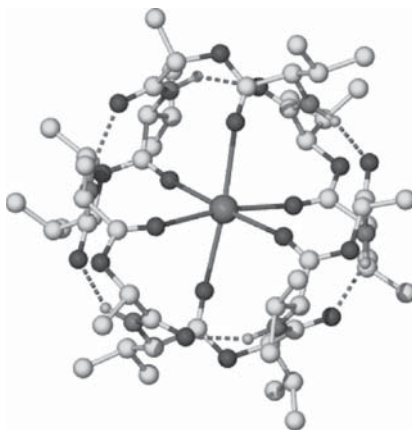
Transport of metal ions *via* the carrier mechanism involves a carrier ligand that is able both to bind selectively to the metal cation and to shield it from the lipophilic region of the membrane. Such ion carriers are termed *ionophores*. The natural products valinomycin (**2.1**) and nonactin (**2.2**) are among the best known. Valinomycin was first isolated from the bacterium *Streptomyces fulvissimus* in 1955, and was established in 1967 to catalyse the exchange of  $K^+$  and  $H^+$  across the membrane of mitochondria within cells *via* a carrier mechanism without affecting  $Na^+$  concentration. Chemically, valinomycin is a cyclic depsipeptide made up of a threefold repetition of four amino acid residues: L-valine (Val), D-hydroxyisovaleric acid (Hyi), D-valine and L-lactic acid (Lac) (**2.3**). Hydrogen bonding of type  $N-H \cdots O=C$  to both ester and amide carbonyl groups plays an important role in the conformation of valinomycin, where it helps the peptide chain wrap around the metal cation, contributing to its degree of preorganisation and stabilising the bound conformation. Valinomycin and nonactin are both selective for  $K^+$  because they are able to fold in on themselves in order to produce an approximately octahedral array of hard (*i.e.* non-polarisable, according to the hard and soft acids and bases (HSAB) theory; see Section 3.1.2) carbonyl oxygen atom donors exactly suited to fit an ion of the size of  $K^+$ . Rubidium and caesium are too large, whereas the ionophore cannot contract tightly enough to bind to  $Na^+$ . The X-ray crystal structure of the  $K^+$  complex of valinomycin is shown in Figure 2.5. It can



**Figure 2.3** Schematic diagram of a phospholipid biological membrane (5–6 nm in width).



**Figure 2.4** (a) Carrier, (b) channel and (c) gated channel mechanisms of ion transport across a biological membrane. The carrier encapsulates the alkali metal cation, stripping away most or all of the water molecules bound to it. The carrier then presents a lipophilic surface, moving across the membrane and releasing the ion back into aqueous solution at the other end. The channel is, primarily, an aqueous hole running through the membrane. Ions can traverse swiftly through the channel without losing their solvent sphere throughout most of their journey, as long as the gate (activated by potential changes or hormones) is open and the ion can pass through the selectivity filter (which discriminates between  $Na^+$  and  $K^+$ ). (Reproduced by permission of John Wiley & Sons, Ltd).



**Figure 2.5** X-ray crystal structure of the  $K^+$  complex of valinomycin. For a summary of the technique of X-ray crystallography, see Box 2.1.

**Box 2.1** X-ray Crystallography of Supramolecular Compounds

Massa, W. and Gould, R. O., *Crystal Structure Determination*. 2nd ed.; Springer: Heidelberg, 2004.

The primary objective of an X-ray diffraction experiment is to obtain a detailed picture of the contents of the crystal at the atomic level, as if you were viewing it through an extremely powerful microscope. Experimentally this process consists of a series of intensity measurements of X-ray beams diffracted from a small, single crystal sample. In the best case the experiment results in a detailed knowledge of the positions of all the atoms in the molecule and hence detailed knowledge of the molecular structure as a whole, including bond lengths, angles and intermolecular contacts. The experiment also gives insight into thermal motion in the solid state. These results are of particular importance to supramolecular chemists because they give direct information on intermolecular interactions such as hydrogen bonding and ion–dipole interactions, as well as the steric fit of a receptor and a substrate (host and guest). Because X-ray crystallography relies upon diffraction by the electron density in the crystal, however, it is not very good at locating hydrogen atoms accurately (after all, they have only one electron each). These are important in hydrogen bonding studies and, if extremely precise information is required, they are usually located using single crystal neutron diffraction (see Box 8.1).

So, why don't we just use a very powerful microscope and view atoms directly? The answer is that atoms are simply too small. The wavelength of visible light falls between 400 and 700 nm. In comparison, a typical bond distance between two atoms (*e.g.* carbon) is about 0.15 nm, so we need radiation of a much smaller wavelength, comparable to that of the interatomic separations: X-ray radiation. The problem with radiation of very short wavelength is that it is impossible to manufacture a lens powerful enough to refocus it. This is referred to as the phase problem. Another way of describing the phase problem is that crystal structure determination calculations require information about the amplitude of diffracted X-rays (the so-called structure factor amplitude,  $F$ ) which is proportional to the square root of the observed X-ray intensity. Unfortunately  $F$  can be both positive and negative and there is no way of directly determining the sign of  $F$  from a measurement of intensity, *i.e.*  $F^2$  because  $|-F|^2$  is the same as  $|+F|^2$ . This loss of  $+/-$  phase information is the phase problem. The phase problem is solved by developing a mathematical model of the structure based on clues from the experimental data and chemical intuition. This is used to calculate the molecules' X-ray diffraction patterns. The calculated prediction is compared with the observed experimental data and improved by least-squares refinement. The experiment is complete when the agreement between calculated and observed data is as good as possible. This agreement is measured by the  $R$  (residual) factor and the standard uncertainties (errors, often referred to as estimated standard deviations) on the individual bond lengths and angles. For a good structure determination, the  $R$  factor should be around 5% or less.

*A Typical Experiment*

1. Prepare single crystal sample (slow evaporation, recrystallisation *etc.*):
  - Homogeneous single crystal without defects or cracks.
  - Size 0.1–0.8 mm edge length, ideally spherical or at least equidimensional.
2. Examine under optical microscope to check for imperfections.
3. Preliminary X-ray photographs (polaroid film or electronic area detector).
  - Check crystal quality.
  - Determine unit cell dimensions and symmetry information (the unit cell is the basic building block of the crystal).
4. Collect intensity data (1000–100 000 data points depending on molecular size).
5. Develop an approximate model of the structure (referred to as *solving* the structure).
  - Application of Patterson or Direct methods to 'guess' approximate phases.
  - Stereochemistry of molecule in outline.
  - Bond lengths to  $\pm 0.1 \text{ \AA}$  ( $1 \text{ \AA}$  (Ångstrom) =  $10^{-10} \text{ m}$  = 0.1 nm).

*(continued)*

**Box 2.1** (Continued)

6. Refine structure by least squares method.
  - Optimise model (set of atom coordinates and thermal parameters) to get best fit.
  - Bond lengths to  $\pm 0.005\text{\AA}$ .
  - Precise stereochemistry.
7. Convert coordinates to tables of bond lengths and angles, and draw picture.
8. Deposit the results in CIF format in the Cambridge Structural Database (CSD, Section 8.4)

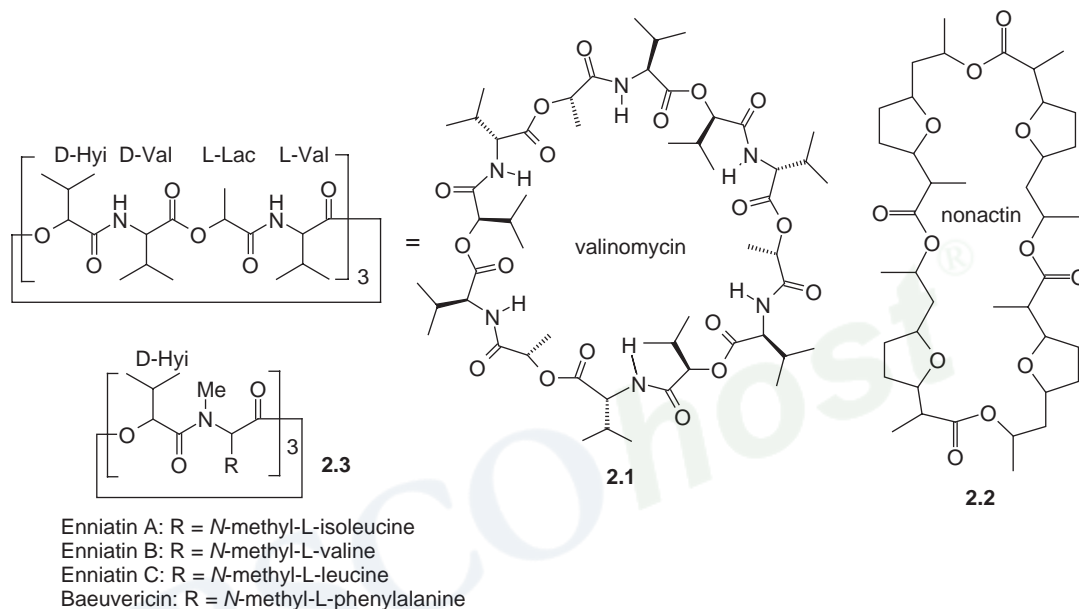
This process becomes more complicated as the size of the molecule or array to be studied increases. Most supramolecular compounds are of a size intermediate between traditional 'small molecules' and large biomolecules such as proteins and it is usually possible to obtain precise information on them fairly readily. As the size of the molecule of molecular assembly increases, however, the number of data measured and parameters that must be fitted rises dramatically. This complexity is often accompanied by increasing problems in obtaining good-quality crystalline samples. Large molecules, particularly those encountered in supramolecular chemistry that contain cavities or are of awkward shapes and fit together poorly, often contain occluded solvent molecules. These can diffuse out of the crystal lattice during the X-ray experiment causing loss of crystallinity and hence loss of diffraction intensity. They can also move around resulting in a 'smeared' (disordered) averaged electron density, which is difficult to model. Even in cases where solvent is not present, poorly packed crystals result in weak diffraction. The X-ray crystallography of even higher molecular weight samples, such as proteins, is much more complicated, requiring the collection of a number of closely related sets of data, often from many crystals. In protein work, problems are encountered with X-ray damage to the sample, weak diffraction, and difficulty in identifying the various molecular fragments. Both supramolecular and protein crystallography have been revolutionised by the advent of area detectors such as charge-coupled device (CCD) instruments (Figure 2.6) (X-ray detectors that are capable of measuring many data points simultaneously), which enhance vastly both the speed and sensitivity of the experiment. Moreover synchrotron sources such as Diamond in the UK are becoming relatively accessible and are able to obtain single crystal X-ray data on extremely small samples often only a few tens of microns in size. In addition modern X-ray diffraction work, especially that involving hydrogen bonded or solvated supramolecular species, is generally carried out at very low temperatures (100–150 K) to reduce atomic motion and prevent diffusional solvent loss.



**Figure 2.6** A modern CCD diffractometer. Note the circular area detector on the left, which acts as a very sensitive electronic, reusable equivalent of photographic film, allowing many data points to be collected simultaneously. (Photograph courtesy of Nonius.).



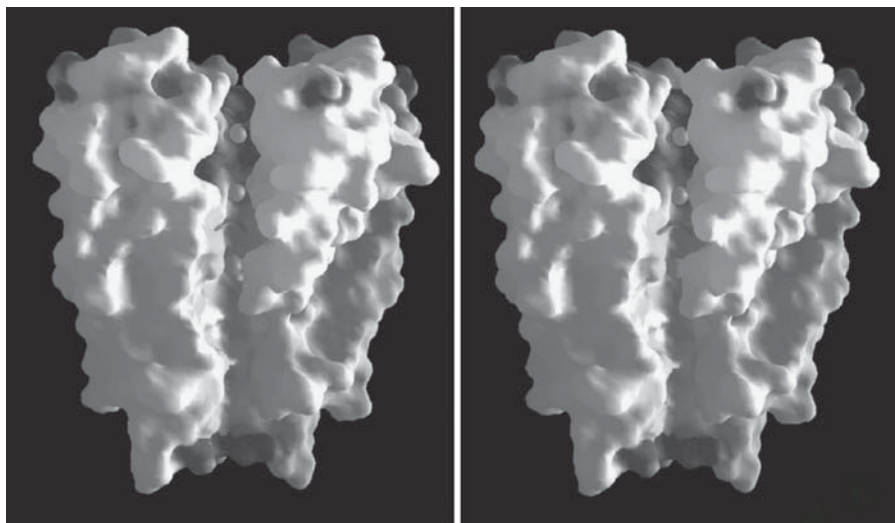
be seen clearly that the interaction of the hydrophilic carbonyl oxygen atoms with the central  $K^+$  ion causes the lipophilic *iso*-propyl groups to point outwards thus exposing a primarily hydrocarbon coated outer sheath to the surrounding medium. The remaining amide functionalities act to 'zip up' the molecule *via* intramolecular hydrogen bonds, ensuring the  $K^+$  ion is encased entirely in a lipophilic exterior as it crosses the membrane. Both valinomycin and nonactin are potent antibiotics because of their ability to perturb transmembrane ionic balance in bacteria.



Related closely to valinomycin are the enniatins (**2.3**), which are made up of only half the number of amino acid binding units (see Box 2.2 for amino acid structure). The enniatins transport alkali metal cations and alkaline earth metal ions, although they are much less selective than valinomycin. Note that the amide nitrogen atoms are methylated, precluding the possibility of hydrogen bonding. Binding constants for a range of naturally occurring ionophores are given in Table 2.2. Note that many of the ionophores bind strongly to  $K^+$  while only monesin actually binds  $Na^+$  with any selectivity. We will return to the origins of the high affinity of these ligands for  $K^+$  and the factors affecting selectivity in general in the next chapter.

**Table 2.2** Log  $K_{11}$  values for alkali metal ion binding by naturally occurring ionophores in methanol solvent at 25 °C.

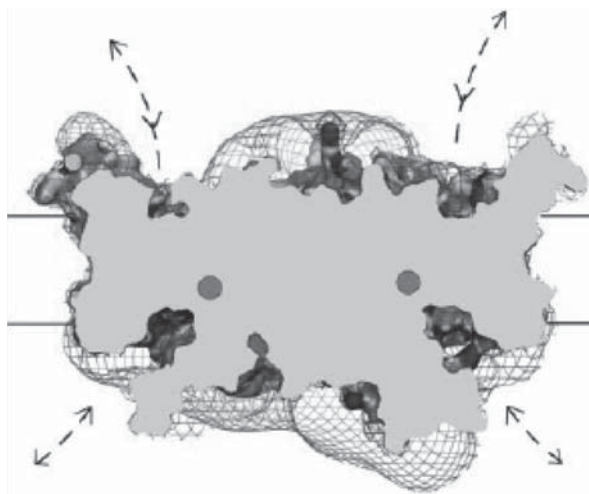
Ligand	$Li^+$	$Na^+$	$K^+$	$Rb^+$	$Cs^+$
Valinomycin	<0.7	0.67	4.90	5.26	4.42
Monactin	<0.3	2.60	4.38	4.38	3.30
Enniatin B	1.28	2.42	2.92	2.74	2.34
Nigericin	–	4.7	5.6	5.0	–
Monesin	3.6	6.5	5.0	4.3	3.6



**Figure 2.7** Cutaway stereoview of the X-ray crystal structure of the  $K^+$  channel of *Streptomyces lividans*. The upper and lower ends of the channel are regions of high negative charge density while the central portion comprises hydrophobic amino acid side chains. Positively charged regions are on the outer surface, while the spheres represent  $K^+$  ion positions. (Reproduced with permission from [2] MacKinnon). See plate section for colour version of this image.

While important, the sequential desolvation–complexation–transport–decomplexation mechanism of ionophore-mediated cation transport is far too slow for effective use in the generation of nerve impulses. In contrast, passage of ions through ion channels results in transport close to the diffusion limits (about  $10^8$  ions per channel per second). The X-ray crystal structure determination on the  $K^+$  channel from the microorganism *Streptomyces lividans*<sup>2</sup> reveals the basis for both the remarkable ion throughput rate and the vital  $10^5$ -fold selectivity of the channel for  $K^+$  over  $Na^+$  (Figure 2.7). Both ends of the ion channel are surrounded by negatively charged amino acid residues, which have the effect of raising the local concentration of cations. The majority of the pore consists of a wide tunnel (18 Å in length), opening out into a central cavity of about 10 Å in diameter at the centre of the pore. Potassium cations can traverse both tunnel and cavity without shedding their hydration sphere. It has been suggested that the large central cavity overcomes the electrostatic destabilisation resulting from the low dielectric constant of the bilayer simply by surrounding the ion with polarisable water. The pore then narrows into a selectivity filter lined with polar carbonyl oxygen atoms, prearranged for binding a metal ion with the radius of  $K^+$ . Rubidium, which is only slightly bigger (see Box 3.1), also passes readily through the channel. Most importantly, the ring of metal cation binding sites is held open in a ‘spring-loaded’ fashion by the aryl moieties of tryptophanyl and tyrosyl residues positioned like a cuff of the selectivity filter (see Box 2.2 for amino acid structures). The binding site is thus kept open so that it cannot contract sufficiently to bind to  $Na^+$ , hence the origin of its selectivity.

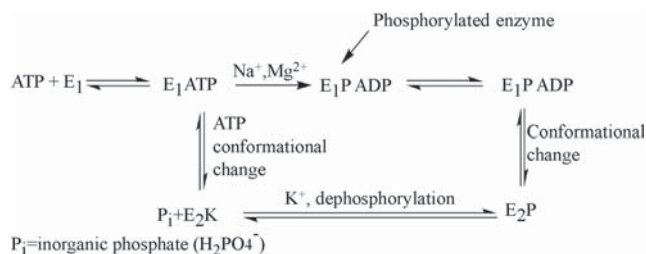
Just as there are cation channels, there are also trans-membrane channels involved in the transport of biologically important anions such as  $Cl^-$ . The crystal structure of the CIC chloride channel from *Salmonella typhimurium* was reported in 2002.<sup>3</sup> Along with the determination of the *Streptomyces lividans* potassium channel structure, this work won a share of the 2003 Nobel prize in chemistry for Roderick MacKinnon (Howard Hughes Medical Institute, New York, USA). Chloride channels catalyse the flow of chloride across cell membranes and play a significant role in functions such as



**Figure 2.8** Surface electrostatic potential map of the CIC chloride channel dimer. The channel is sliced in half to show the pore entryways (but not the full extent of their depth) on the extracellular (above) and intracellular (below) sides of the membrane (represented by horizontal lines). The  $\text{Cl}^-$  ions are shown as spheres and dashed lines highlight the pore entryways. (reproduced with permission from Reference 3).

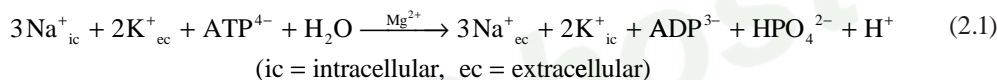
the regulation of pH, volume homeostasis, organic solute transport, cell migration, cell proliferation and differentiation. In skeletal muscle, CIC chloride channels stabilise the resting membrane potential while in the kidney they act in conjunction with a  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Cl}^-$  co-transporter and  $\text{K}^+$  channels to produce transepithelial fluid and electrolyte transport. The structure of the CIC chloride channel is based on two identical subunits and the channel functions as a dimer. In contrast to the  $\text{K}^+$  channel structure, each subunit has its own individual aqueous pore and chloride selectivity filter (Figure 2.8). In the  $\text{K}^+$  channel the pore occurs at the interface between four protein subunits. The anion binding region is surrounded by positively charged N-terminus regions of the protein, complementary to  $\text{Cl}^-$ . The channel also conducts  $\text{Br}^-$  well, however, like  $\text{Rb}^+$  and  $\text{Cs}^+$ , bromide is not a significant bulk biological anion. Another chloride channel of interest to supramolecular chemists is the *cystic fibrosis transmembrane conductance regulator* (CFTR). The CFTR is found in the epithelial (*i.e.* surface) cells of a range of major organs, particularly the lungs and skin. Like the CIC channel its role is chloride ion transport, in this case from out of the epithelial cell to the surrounding mucus. This process causes an electrical gradient and the movement of  $\text{Na}^+$  ions in the same direction to balance the charge ( $\text{Na}^+ / \text{Cl}^-$  symport). This results in the absorption of water by the mucus maintaining the correct level of fluidity. In the genetic disease cystic fibrosis (CF) mutations in the CFTR reduces the ability of the affected cells, predominantly epithelial cells lining the lung and other mucosal surfaces, to transport chloride effectively with consequences for fluid transport. Defective fluid transport leads to an underhydration of mucus secretions which then obstructs organ passages and causes the widespread and serious organ malfunction associated with the disease. The development of artificial chloride channels to treat the disease is of considerable interest in supramolecular chemistry (Section 12.8).<sup>4</sup>

The crystal structure of  $\text{Na}^+/\text{K}^+$ -ATPase, the enzyme responsible for maintaining  $\text{Na}^+/\text{K}^+$  non-equilibrium in most eucaryotic organisms was determined 2007.<sup>5</sup> The enzyme is difficult to purify since it is membrane-bound and has a high molecular mass (294 kDa). It is a heterodimer made up of



**Scheme 2.2** Mechanism of the enzymatic action of ATPase.

two pairs of a large ( $\alpha$ ) and small ( $\beta$ ) peptide, *i.e.* ( $\alpha\beta$ )<sub>2</sub>. Each  $\alpha\beta$  pair is active. The  $\beta$ -peptide is a glycopeptide with a molecular weight of about 50 000 Da which is a special feature of K<sup>+</sup>-counter-transporting ATPases. Its job is the routing of the  $\alpha$ -subunit to the plasma membrane and binding of the K<sup>+</sup> ions. The  $\alpha$ -peptide is a membrane-spanning protein with a molecular weight of about 100 000 Da, and exists in different conformations in the presence of Na<sup>+</sup> (conformation E<sub>1</sub>) or K<sup>+</sup> (E<sub>2</sub>). It is involved in Na<sup>+</sup>-dependent phosphorylation of the  $\beta$ -carboxyl group of an aspartic acid residue, a process that also requires Mg<sup>2+</sup>. Conversely, loss of the phosphate requires K<sup>+</sup>. The overall process is:



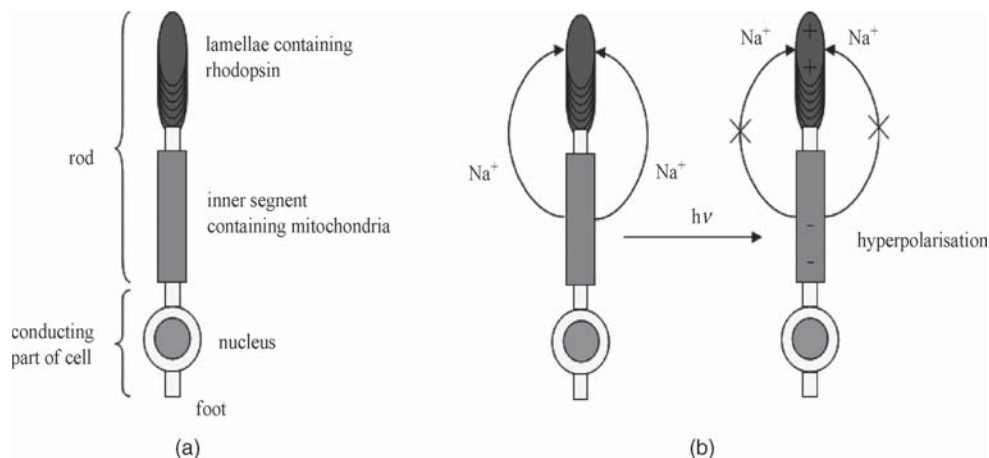
The mechanism of the enzymatic action of the ATPase, in isolation from its transport function, may be considered as shown in Scheme 2.2.

The system may be regarded as involving a Na<sup>+</sup>/Mg<sup>2+</sup> co-catalysed phosphorylation step and a K<sup>+</sup> catalysed dephosphorylation. Each phosphorylation/dephosphorylation step involves a ‘pseudorotation’ of an Mg<sup>2+</sup>-stabilised 5-coordinate intermediate, resulting in transport of the alkali metal cations. The cation transport ability of the enzyme is a direct result of the enzymatic reactivity of the protein. There are three binding sites with high Na<sup>+</sup> affinity and two with K<sup>+</sup> affinity (occupied by Rb<sup>+</sup> in the crystal structure determination). The structure (which is of the E<sub>2</sub>K state of the system) reveals that carboxy end of the  $\alpha$ -subunit is held in a pocket in between transmembrane helices and acts as an unusual regulating element that controls sodium affinity and may be influenced by the membrane potential.

### 2.2.3 Rhodopsin: A Supramolecular Photonic Device

☛ Pepe, I. M., ‘Rhodopsin and phototransduction’, *J. Photochem. Photobiol. B-Biology* 1999, **48**, 1–10.

Supramolecular alkali metal cation transport is also of importance in other signal generating processes, such as the stimulation by visible light of rod and cone cells in the retina of the eye. The rod cells responsible for black and white vision are better understood and bear all of the characteristics of a functional supramolecular device. The structure of a typical rod is shown in Figure 2.9a. The light-harvesting portion of a rod cell contains a photosensitive reddish pigment called rhodopsin (commonly known as *visual purple*). This comprises a complex protein (opsin) conjugated with a simple aldehyde of vitamin A (carotene) called retinine. The absorption spectrum of rhodopsin is sensitive to the kinds of wavelengths that occur at low lighting levels, and on photoexcitation the connection between the opsin and retinine components (an azomethine function, C=NH<sup>+</sup>) undergoes a simple *cis* to *trans* isomerism forming bathorhodopsin. This results in the degradation of cyclic guanosine monophosphate



**Figure 2.9** (a) Light-responsive rod cell from the human retina. The photosensitive pigment rhodopsin is located in the outer lamellae. The foot makes contact with the optic nerve. (b) Operation of the rod cell.

(cGMP), which is necessary to the maintenance of the ever-present energy-consuming flow of  $\text{Na}^+$  ions that comprise the cell's 'dark current', and hence  $\text{Na}^+$  transmembrane channels close. This, in turn, causes a marked hyperpolarisation within the cell and leads to amplification of the original light signal and the generation of a nerve impulse (Figure 2.9b).

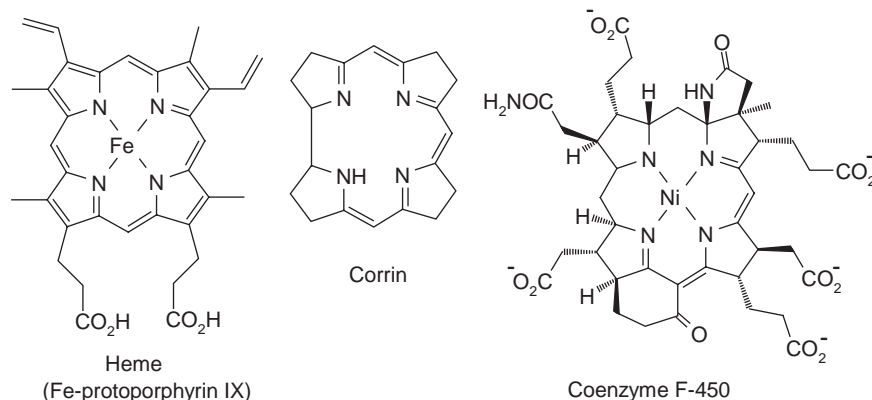
Rhodopsin absorbs light at 498 nm while the three pigments responsible for trichromatic colour vision in the cone cells absorb at 425, 530, and 560 nm. Interestingly the chromophore in these receptors is the same compound – a protonated Schiff base of 11-cis-retinal. Logically, therefore, the absorption wavelength of these pigments must be related to their supramolecular environment within their protein home. Interestingly recent calculations show that the main role of the protein binding pocket is to stabilise the pigments as they undergo the isomerisation reaction to bathorhodopsin and it is the counter anion that has by far the most significant effect on the pigment absorption maximum.<sup>6</sup>

Examples such as this show the importance of alkali metal cation binding and transport in biochemistry, and a great deal of effort has been expended in supramolecular chemistry in attempts to understand natural cation binding and transport of the ionophore and channel type and to develop artificial systems capable of similar selectivities and reactivities. We will take a close look at many of these compounds in Chapter 3.

## 2.3 Porphyrins and Tetrapyrrole Macrocycles

Macrocyclic (large-ring) ligands are common in biochemistry, and are able to bind even substitutionally labile metal ions by virtue of the chelate and macrocyclic effects. This stabilisation is especially important for alkali metal cation binding, in a similar way to ionophores such as valinomycin and non-actin. Tetrapyrrole macrocycles are also important in binding to both transition and main group metal ions in an enormous variety of situations in which strong, size-selective complexation is required. Tetrapyrroles possess extensive redox chemistry because the conjugated ring framework itself may be readily reduced. Common examples are:

- chlorophylls, which contain otherwise labile  $\text{Mg}^{2+}$ , and are important as energy harvesters in photosynthesis;
- cobalamins, the active form of vitamin  $\text{B}_{12}$ , which involves a partially conjugated 'corrin' ring system (Figure 2.10);



**Figure 2.10** Biological tetrapyrrole macrocyclic compounds.

- haem complexes, which consist of an iron centre and a substituted porphyrin ligand, as in iron protoporphyrin IX (Figure 2.10), the active  $O_2$  binding site found in haemoglobin; haems are also found in a variety of enzymes such as cytochromes, which rely upon  $O_2$  as a substrate;
- a porphyrinoid nickel(II) complex, coenzyme F450 (Figure 2.10c), found in methane-producing microorganisms; the independence of the function of this species from proteins makes it a good candidate for ‘first-hour’ catalysis, *i.e.* the origins of life.

Tetrapyrrole macrocycles such as haems and corrin exhibit a number of special features that make them extremely important:

- The underlying planar (or near-planar) ring system is very stable.
- As tetradentate chelate ligands that, after deprotonation, carry a single (corrin) or double negative charge, the tetrapyrrole can bind even highly labile metal ions. The complex can dissociate only if all metal ligand bonds are broken at the same time (*cf.* chelate and macrocyclic effects).
- Macrocyclic ligands are usually quite selective with regard to the ionic radius of the metal ion they will bind (those that are bound preferentially are those that fit the cavity). Tetrapyrroles are especially selective in this regard because of their rigidity, which arises from the conjugated network of double bonds.
- Most tetrapyrroles contain a conjugated  $\pi$ -system. The Hückel rule for aromatic systems is obeyed for porphyrins which feature 18  $\pi$ -electrons in the inner 16-membered ring (aromatic systems contain  $4n + 2$   $\pi$ -electrons; for fully conjugated tetrapyrroles  $n = 4$ ). This goes some way to account for the thermal stability of the ring system. It also often results in intense colours, hence the designation ‘pigments’ for many tetrapyrrole-containing biological species. Furthermore, it means that the oxidation and reduction products that form part of biological processes can be quite stable because of delocalisation of the resulting charges.
- The macrocycle contains four coordinating atoms in a planar arrangement, leaving two available sites on an octahedral metal centre available to bind the substrate and a regulating ligand.
- Tetrapyrroles are able to distort to give a ‘domed’ (rather than planar) coordination environment for large metal ions such as high-spin Fe(II). Reversible stabilisation of this high-spin state is vital in the action of haemoglobin. In contrast, a planar tetrapyrrole system gives a high ligand field splitting and hence low-spin complexes.

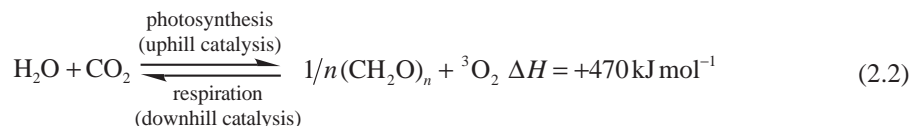
In the ensuing sections we will examine some porphyrin and tetrapyrrole complexes in greater detail.

## 2.4 Supramolecular Features of Plant Photosynthesis

Allen, J. P. and Williams, J. C., 'Photosynthetic reaction centers', *FEBS Lett.* 1998, **438**, 5–9.

### 2.4.1 The Role of Magnesium Tetrapyrrole Complexes

Photosynthesis is the process by which organic matter (reduced carbon) is produced from CO<sub>2</sub> and water by a process of uphill (energy-consuming) catalysis using sunlight — a readily available, clean, though rather 'diluted', form of energy.

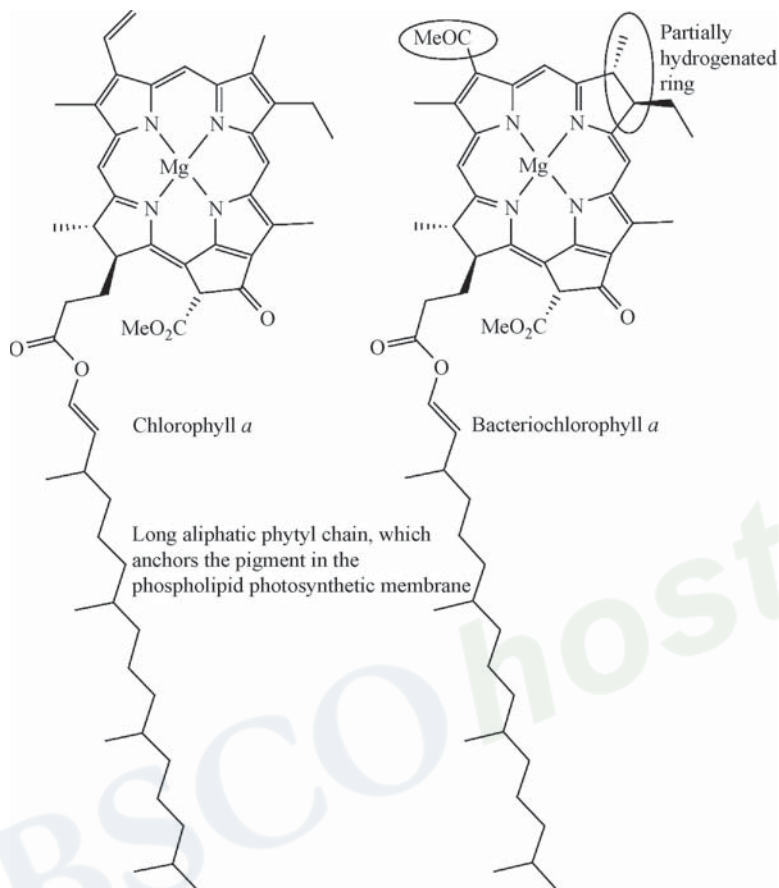


Photosynthesis is carried out by green plants and certain kinds of bacteria and algae, and may involve the oxidation of other substrates such as H<sub>2</sub>S or H<sub>2</sub> instead of water, depending on the conditions under which the organism evolved. In green plants, the photosynthetic output is about 1 g of glucose per hour per 1 m<sup>2</sup> of leaf surface area. Although this represents a relatively inefficient utilisation of the available radiation energy (<1 per cent) the global turnover is tremendous: about 200 billion tons of carbohydrate equivalents (CH<sub>2</sub>O)<sub>n</sub> are produced from CO<sub>2</sub> each year (Equation 2.3).

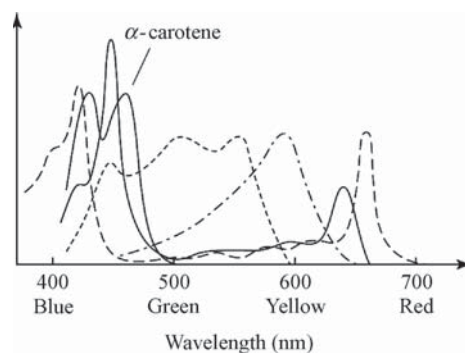
Sunlight available at the Earth's surface includes the wavelength range visible to the human eye (380–750 nm), but there is also a significant contribution from light of longer wavelengths (and therefore lower energy) up to about 1000 nm. Efficient photosynthetic transformation of light of this range of energies requires a number of different pigments (*i.e.* light receptors or chromophores), each sensitive to a particular part of the spectrum. These pigments include chlorophyll *a* and bacteriochlorophyll *a* (Figure 2.11) as well as a number of related receptors, all based on tetrapyrrole macrocycles both with and without a metal cation guest. The pigments are positioned in a highly folded photosynthetic membrane with a large surface area and therefore a high cross-section for photon capture.

Chlorophylls contain a fully conjugated tetrapyrrole  $\pi$ -system (18  $\pi$  electrons) with a low-energy  $\pi$ - $\pi^*$  transition. The extinction coefficient is high (about 10<sup>5</sup> M<sup>-1</sup> cm<sup>-1</sup>) at both long- and short-wavelength ends of the spectrum. The complementary colours blue (after short-wavelength absorption) and yellow (after long-wavelength absorption) combine to give the characteristic green colour of fresh leaves ( $\lambda_{\text{max}}$  455 and 630 nm). Bacteriochlorophylls have two partially hydrogenated pyrrole rings and their absorption is consequently shifted to longer wavelengths, reaching the near infrared (IR) region. Carotenoids and open chain tetrapyrrole molecules, such as phycobilins, complement the chlorophyll pigments so that a broad range of absorption is covered (Figure 2.12). At the end of each growth period non-green carotenoid pigments become visible (resulting in autumn colours) as the relatively unstable chlorophylls decompose (*e.g.* phycoerythrin,  $\lambda_{\text{max}}$  455, 510 and 555 nm; phycocyanin, 595 nm).

Because of the relatively slow rate of light absorption in diffuse sunlight, the majority (>98 per cent) of pigments are used as light-harvesting or antenna devices, absorbing and transferring energy to the actual reaction centres. This means that there has to be efficient, spatially-orientated transfer of the absorbed energy. This energy transfer is made possible by a well-defined antenna network of pigments that are able to funnel the light energy in an energy transfer 'cascade' (Figure 2.13a). Essentially, the energy transfer proceeds *via* overlap of the emission bands of the source with the absorption bands of the acceptor (Figure 2.13b).

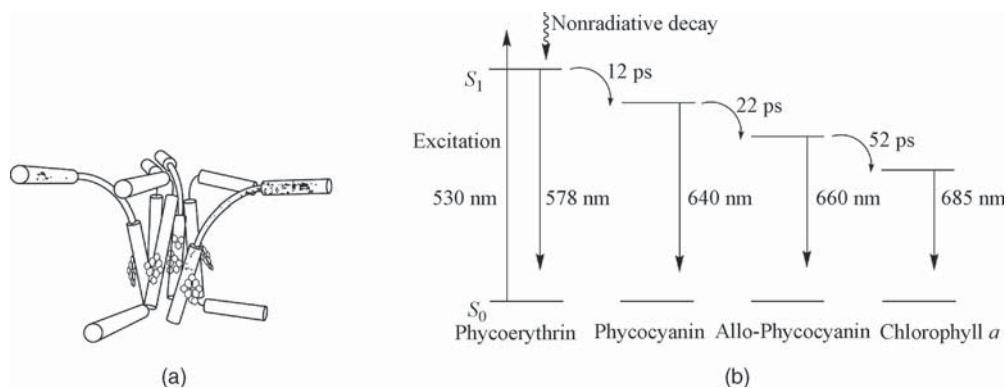


**Figure 2.11** Structures of chlorophyll *a* and bacteriochlorophyll *a*.



**Figure 2.12** Absorption spectra of various pigments from algae and plants: chlorophyll *a* (---), chlorophyll *b* (—)  $\alpha$ -carotene (— · —), phycocyanin (· · · · ·), phycoerythrin (- -). (Reproduced by permission of John Wiley & Sons, Ltd).





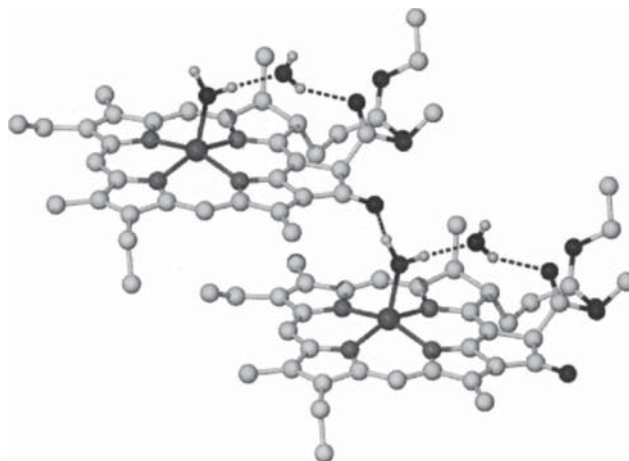
**Figure 2.13** (a) Spatially-orientated antenna network of light-harvesting pigments. (Reproduced by permission of John Wiley & Sons, Ltd). (b) Energy-transfer cascade for antenna pigments in light-harvesting complexes of the algae *Porphyridium cruetum*.

The role of the coordinated  $Mg^{2+}$  ion in chlorophylls is to contribute to the arrangement of the pigments. Pigments are anchored spatially in their correct place, to some extent, by the long phytol side chain, buried deep in the photosynthetic membrane. However, in order to fix the pigment firmly, the free axial coordination site of the normally square pyramidal  $Mg^{2+}$  is bound to polypeptide side chain ligands, giving two points of anchorage of the pigment and hence a well-defined spatial orientation. In the case of the photosynthetic reaction centres of purple bacteria this is a histidine nitrogen atom.  $Mg^{2+}$  is particularly suited to this role for the following reasons:

- high natural abundance (consistent with this noncatalytic 'bulk' function);
- lack of redox activity (redox activity at the metal is incompatible with interpigment electron transfer);
- strong tendency for penta- or hexacoordination;
- suitable ionic radius;
- small spin-orbit coupling constant (spin-orbit coupling results in the formation of long-lived triplet excited states by intersystem crossing, resulting in light- or heat-producing processes competing with chemical reactions).

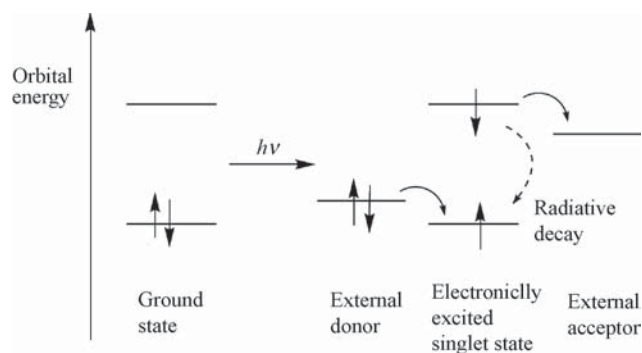
The X-ray crystal structure of ethyl chlorophyllide (in which the phytol group is replaced with a simple ethyl substituent) (Figure 2.14) shows the tendency of the  $Mg^{2+}$  towards a 5- or 6-coordinate environment. The axial site of the polypyrrole-bound  $Mg^{2+}$  ion is bound to a water molecule, which, in turn, hydrogen bonds to the carbonyl group on adjacent molecules, forming a hydrogen-bonded solid-state polymer. This is a good example of solid-state supramolecular organisation, but this kind of direct linkage probably does *not* occur in the real system.

The energy collected by light-harvesting antennae systems and funnelled to the photosynthetic reaction centre is used to produce a spatial charge separation. In other words, an electron is promoted to an excited state and used to effect a chemical reduction before it has the chance to undergo a biochemically useless radiative decay back to its ground state. In simple reaction centres, such as that of the purple bacteria *Rhodospseudomonas viridis*, the excited electron is transferred to an external acceptor. In higher evolved organisms, the 'hole' left behind is also utilised to oxidise a substrate by the involvement of an external donor. Ultimately this results in the production of  $O_2$  from water (Figure 2.15).

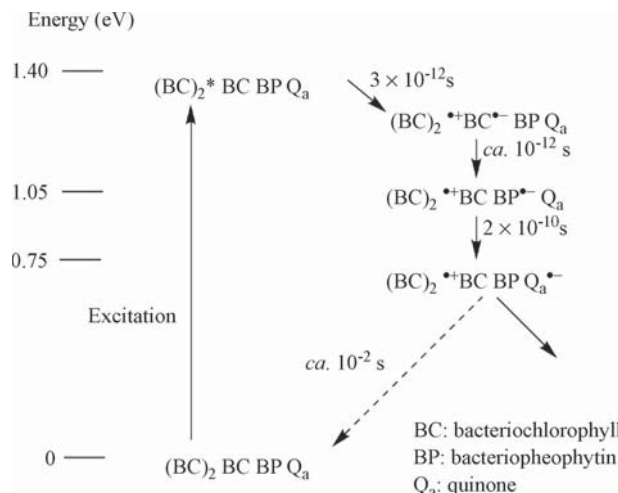


**Figure 2.14** X-ray crystal structure of ethyl chlorophyllide show in the 5-coordinate  $Mg^{2+}$  ion.<sup>7</sup>

In *Rps. viridis*, the reaction centre is situated in a polyprotein complex that spans the photosynthetic membrane. It consists of a nearly  $C_2$  symmetric arrangement in which a bacteriochlorophyll dimer,  $(BC)_2$ , sits on the symmetry axis. Electronic excitation of this dimer (called  $P_{960}$ —a pigment with a long-wavelength absorption maximum at 960 nm) results in an electronically excited state. This is termed *primary charge separation*. Following this excitation, one energetically elevated electron is transferred to the primary acceptor, a monomeric bacteriochlorophyll, before radiative decay can occur. The reduced monomeric BC molecule transfers an electron in turn to the secondary acceptor, a bacteriopheophytin (BP; a bacteriochlorophyll molecule without a coordinated metal centre). This is known as *secondary charge separation* and results in the spatial removal of the electron from its source (Scheme 2.3). The absence of the metal centre makes the molecule easier to reduce since the ionic bond to the  $Mg^{2+}$  ion from the dianionic ligand leaves a significant amount of electron density on the ligand. In effect, replacing the  $Mg^{2+}$  ion with protons results in a more covalent situation. The electron is then transferred from the BP radical anion to a *para*-quinone,  $Q_a$ , such as menaquinone (2.4), reducing it to a semiquinone radical anion. This in turn reduces a more labile quinone,  $Q_b$ , such as ubiquinone (2.5), *via* a high-spin, six-coordinate iron(II) centre.  $Q_b$  is not bound tightly to the

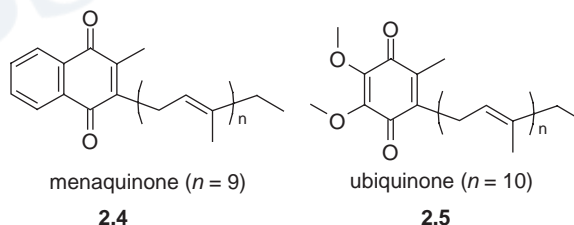
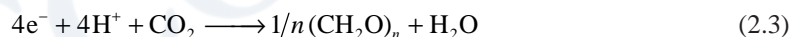


**Figure 2.15** Fate of the photoexcited electron and the resulting positive hole.



**Scheme 2.3** Charge separation at the photosynthetic reaction centre of *Rps. viridis*.

protein, but exchanges with quinones in the ‘quinone pool’ of the membrane. The resulting electron gradient finally gives rise to a coupled  $H^+$  gradient and photosynthetic phosphorylation (ATP synthesis) takes place. In higher organisms, further ‘dark’ reactions take place resulting in the reduction of  $CO_2$  (this is known as the Calvin cycle):

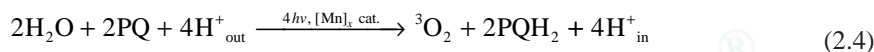


In *Rps. viridis*, the radical cation,  $(BC)_2^{*+}$ , which remained after the initial charge separation is reduced after a relatively long time through regulated electron flow *via* one or more haem centres of cytochrome proteins. In higher organisms, this hole is the beginning of substrate oxidation, which requires an additional photosynthetic system including a manganese containing oxidation catalyst (Section 2.4.2).

The key feature of photosynthesis is the ability to carry charge spatially away from an excited state reaction centre before the usually highly efficient and biochemically useless recombination can take place. In photosynthetic systems, charge separation occurs about  $10^8$  times faster than recombination, a ratio that is impossible to reach in ‘normal’ chemical reactions. This phenomenon is achieved by the spatial anchoring of the components at particular orientations to one another within a non-polar region of the membrane anchored protein, thus preventing free diffusion and allowing a vectorial uphill chemical reaction.

## 2.4.2 Manganese-Catalysed Oxidation of Water to Oxygen

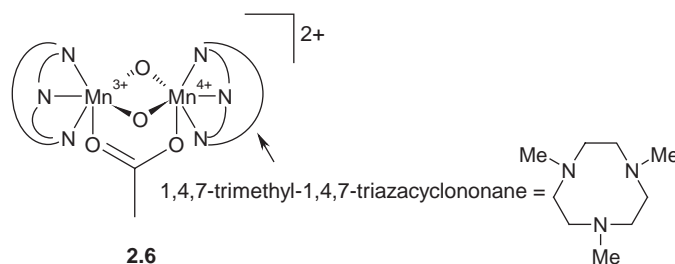
In higher plants, there are two coupled, separately excitable photosystems (PS), labelled PS I and PS II. PS I is based around a pigment ( $P_{700}$ ) with an absorption at 700 nm; which receives an excitation from PS II and, ultimately, goes on to reduce  $\text{CO}_2$ . PS II resembles the simpler bacterial photosynthetic reaction centre (Scheme 2.3), but is based around a higher-energy absorbing pigment,  $P_{680}$ . Promotion of an electron in PS II, and its use in ATP synthesis, results in a strongly oxidising hole being left behind. The high energies involved mean that this hole may be used to oxidise tyrosine to a tyrosine radical cation (+0.95 V), which in turn oxidises a hydroquinone (plastoquinone,  $\text{PQH}_2$ ) to the quinone form (PQ). The reactive PQ oxidises water in a four-electron redox step, catalysed by the so-called OEC (oxygen-evolving complex), which is an enzyme containing four Mn centres. In the process, four protons are transferred from the outside of the membrane to the inside, resulting in a proton gradient. The overall reaction is:

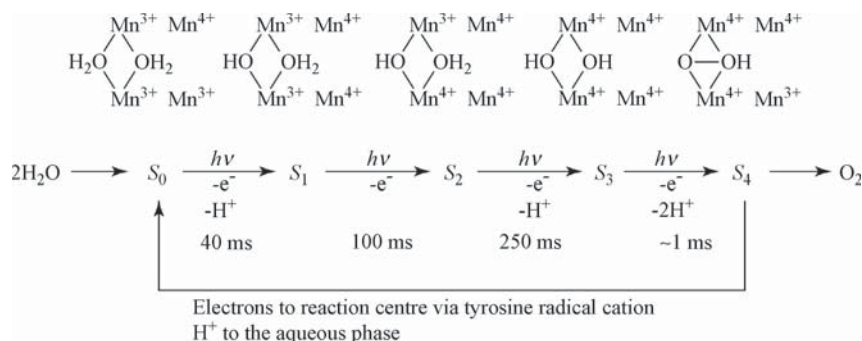


Overall, PS II consists of the following necessary elements (those involved directly with water oxidation are shown in **bold**):

- ~200 Antenna chlorophylls
- ~50 Carotenoids
- 1 Reaction centre,  $P_{680}$
- 2 Chlorophylls
- 2 Pheophytins (primary acceptor)
- 2 **Plastoquinones**
- 2 **Tyrosine residues** (primary donor)
- 4 **Mn centres**
- 1  $\text{Ca}^{2+}$  (role still relatively obscure, but may be related to structure stabilisation or regulation)
- several  $\text{Cl}^-$
- 1 Cytochrome  $b_{559}$

It is well established, however, that the OEC contains four Mn centres in one 33 kDa subunit. Extended X-ray absorption fine structure (EXAFS) measurements indicate Mn—Mn distances of 2.70 and 3.30 Å, suggesting, respectively, oxide and carboxylate bridges linking pairs of metal centres. Supporting evidence comes from an extensive number of small structural model compounds such as **2.6** that mimic these properties successfully. During the process of oxidation of two water molecules to  $\text{O}_2$ , the  $\text{Mn}_4$  cluster passes through a total of five different states, referred to as  $S_0$  to  $S_4$ , during the course of a few hundred milliseconds. Structural changes are small between each of the five states, consistent with a low activation energy, but there are significant changes in oxidation state.

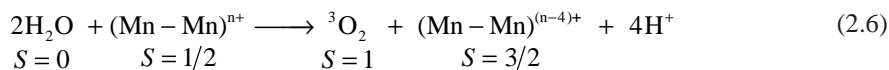
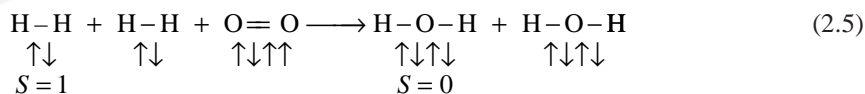




**Figure 2.16** Changes in manganese oxidation state in the conversion of water to  $O_2$ .

Figure 2.16 represents one possibility for the changes occurring in the five  $S$  states. Evidence from EXAFS spectroscopy confirms a change in oxidation state going from  $S_0$  to  $S_1$  and to  $S_2$ , but there are conflicting reports concerning the change to  $S_3$ . It may involve oxidation of a histidine ligand instead of a Mn centre. There is good evidence from EPR spectroscopy that the  $S_2$  state involves a single unpaired electron, consistent with an odd total number of electrons and antiferromagnetic coupling.

The very short-lived  $S_4$  state contains a peroxide  $O_2^{2-}$  ligand and one of the Mn(IV) centres may have been reduced back to Mn(III) to give back the odd number of electrons observed for  $S_2$  ( $S_4$  has two electrons fewer than  $S_2$ , and so if  $S_2$  is odd,  $S_4$  must also be odd). The fact that the system involves an odd number of unpaired electrons is important since oxygen must be released in its triplet  $^3O_2$  state, *i.e.* with an even number of unpaired electrons. A match between the catalyst and the oxygen might result in irreversible binding, as is the case with the reactions of most transition metals with  $O_2$ . Luckily, the spin flipping required to make the two systems compatible is a formally forbidden, and hence slow, process. This ‘spin inhibition’ is similar to that for the reaction of  $H_2$  and  $O_2$ , which does not proceed without a bond breaking catalyst or radical initiator.



So, what makes manganese so suitable as an oxidation catalyst? This question may be summarised by the following points:

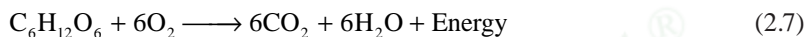
- It is known that fresh manganese dioxide,  $\text{MnO}_{2-x}\cdot n\text{H}_2\text{O}$ , a nonstoichiometric mixed-valent (+IV, +III) system, acts as a good heterogeneous catalyst for the decomposition of  $\text{H}_2\text{O}_2$  to  $\text{O}_2$  and water.
- Manganese was readily bioavailable in seawater at the time of the evolution of photosynthesis (about  $3 \times 10^9$  years ago) and now abounds on the sea bottom in the form of oxidic manganese nodules with about 20 per cent Mn content.
- Manganese has a large variety of stable or metastable oxidation states (+II, +III, +IV, +VI, +VII).
- Manganese exhibits labile binding of ligands.
- Manganese has a strong preference for high-spin states because of small  $d$  orbital splitting.

In the four-electron oxidation of water to O<sub>2</sub>, the polymanganese system acts as (i) an electron reservoir, accumulating charge in an exactly controlled fashion at physiologically high redox potential and (ii) as a non-<sup>3</sup>O<sub>2</sub>-retaining catalyst.

The role of the essential Cl<sup>-</sup> ions may be that of temporary ligands for the Mn centres in between O<sub>2</sub> release and water binding. Usually this role is fulfilled by the solvent, but in this case the solvent (water) is also the substrate. Cl<sup>-</sup> is unlikely to be oxidised itself ( $E^\circ +1.36$  V for oxidation to Cl<sub>2</sub>) but more sensitive peptide residues may be damaged if the Cl<sup>-</sup> were not present to prevent their binding.

## 2.5 Uptake and Transport of Oxygen by Haemoglobin

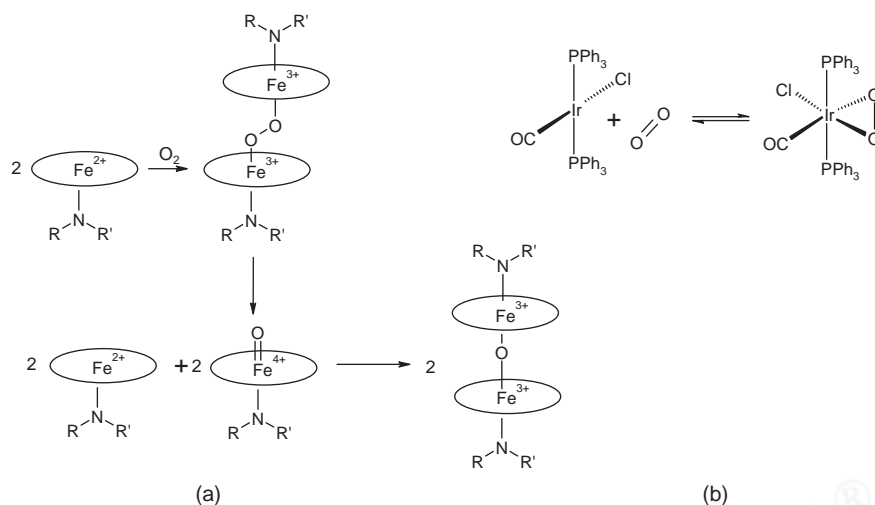
Dioxygen, O<sub>2</sub>, is a vital part of the metabolism of higher organisms that respire aerobically. It is used to metabolically oxidise sugars such as glucose and sucrose (Equation 2.7) with an associated release of energy. The energy from this controlled ‘cold combustion’ is used in ATP synthesis.



Oxygen is produced biogenically by higher plants (Section 2.4.2) and must initially have been a waste product arising from the need to replace the photoexcited electron in the photosynthetic reaction centre. In the primeval oceans, water would have been by far the most abundant electron source. Unfortunately, O<sub>2</sub> is a highly reactive and, to primitive organisms, highly toxic gas. Initially, oxygen was scavenged from the atmosphere by reduced metal ions such as Fe(II) and Mn(II), but about two billion years ago, following the deposition of large amounts of metal oxide sediments, the O<sub>2</sub> content began to rise in the primitive atmosphere from levels of less than one per cent (as it is on lifeless planets and moons today) to 21 per cent of the atmospheric volume. As a result, most primitive organisms would have perished by radical degradation mechanisms and oxidation of metalloenzymes. Only the newly evolved aerobic organisms that had begun to appear as a result of this surplus of high-energy compound would have survived. Today, only a few anaerobic organisms survive: those that live in ecological niches such as the deep ocean, where atmospheric O<sub>2</sub> cannot penetrate, and those that were able to develop defence mechanisms against O<sub>2</sub> and its partially reduced radical products.

In order for aerobic organisms to utilise the reactive O<sub>2</sub> in what is effectively the reverse of photosynthesis, it is necessary for the O<sub>2</sub> to be taken up and transported to the cell mitochondria where respiration with ‘food’ (*i.e.* sugars) occurs without reacting irreversibly or causing radical or oxidative damage along the way. A very sophisticated oxygen uptake and transport protein, haemoglobin, has evolved to carry out this task. Haemoglobin is a tetrameric protein (RMM = 64.5 kDa) containing four myoglobin units (RMM = 17.8 kDa). Each myoglobin unit contains an iron–porphyrin coordination complex called haem or Fe-protoporphyrin IX (Figure 2.10), which is bound to the protein by a coordination interaction between an axial site of the octahedral Fe(II) centre and a nitrogen atom from the proximal protein histidine residue. It is the ability of the iron centre to bind O<sub>2</sub> *reversibly* that is the key to this vital biological system. The remaining axial site on the iron atom is available for oxygen binding, although it may be occupied by a loosely bound water molecule in its resting state.

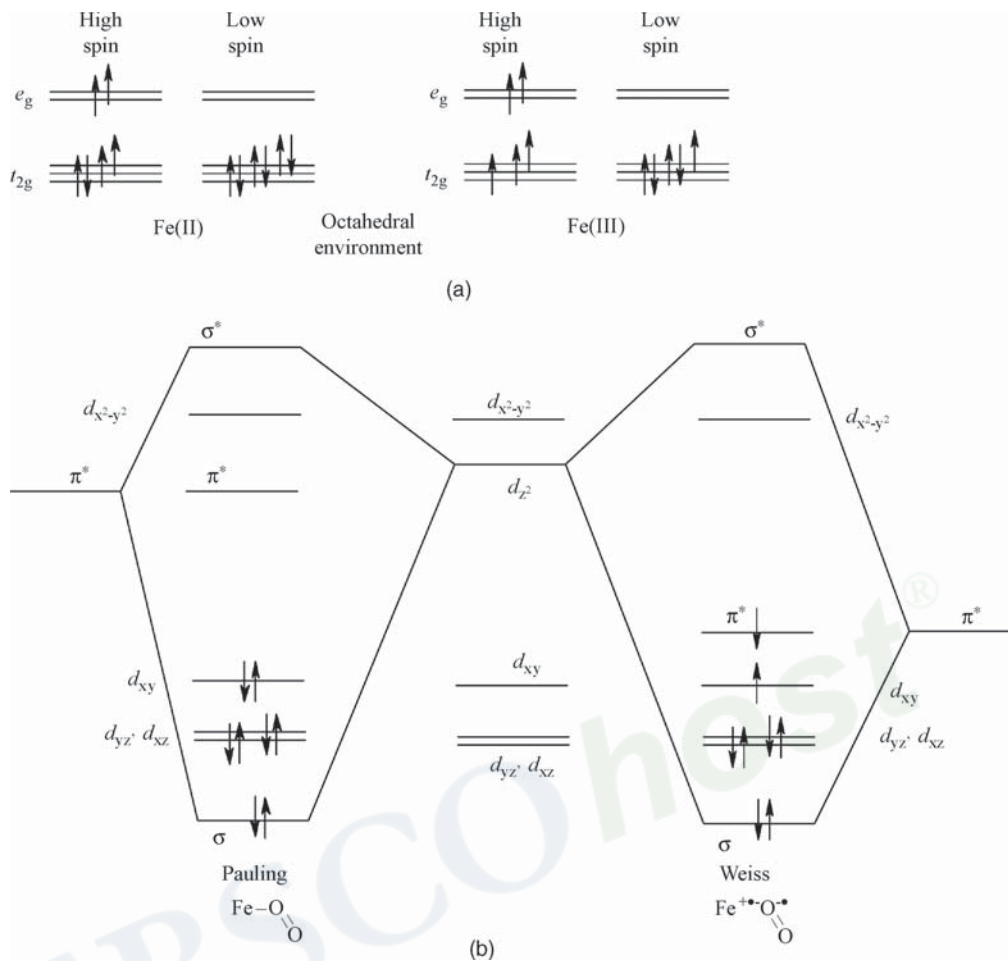
The main problem with the reaction of metal centres with O<sub>2</sub> is its tendency to react irreversibly with oxidation of the metal centre, to form either *cis* dioxo species or  $\mu$ -oxo bridged binuclear complexes (Scheme 2.4a). However, in cases where there is a strong driving force towards lower oxidation states, the reaction is somewhat reversible, *e.g.* reaction with Vaska’s compound (Scheme 2.4b). The role of the haem centre in haemoglobin is to ensure not only that its O<sub>2</sub> binding is reversible, but that both its complexation and release occur rapidly and at the correct concentrations. These concentrations, or partial pressures, must correspond to those found in the lung and intracellular medium, respectively.



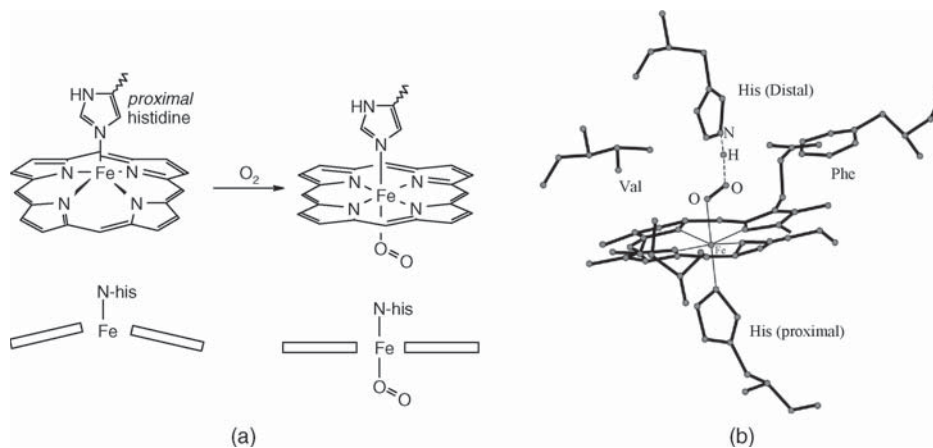
**Scheme 2.4** (a) Irreversible  $\mu$ -oxo dimer formation in nonprotein iron porphyrin complexes. (b) A rare example of reversible  $\text{O}_2$  binding in a simple inorganic complex. The formally Ir(III)/ $\text{O}_2^{2-}$  complex is destabilised by the  $\pi$ -acceptor properties of the phosphine and CO ligands.

Furthermore,  $\text{O}_2$  binding must occur selectively amongst other atmospheric components such as water,  $\text{N}_2$ ,  $\text{CO}_2$  and even excellent ligands for Fe(II) such as CO. Haemoglobin is thus an excellent example of a functional and selective supramolecular receptor.

Before  $\text{O}_2$  binding, the haem units contain high-spin Fe(II) centres of valence electron configuration  $(t_{2g})^4(e_g)^2$ , resulting in a paramagnetic species with four unpaired electrons (Figure 2.17a). The Fe(II) centre is thought to be located significantly out of the plane of the porphyrin unit, a conformational characteristic known as ‘doming’, and is orientated towards the proximal (closest) histidine residue of the surrounding protein. This is a consequence of the large size of high-spin Fe(II) compared with the dimensions of the porphyrin ring (ionic radius =  $0.78\text{\AA}$ ). The binding of  $\text{O}_2$  (in its triplet ground state) has been the subject of much controversy centred around two opposing viewpoints: the Pauling and Weiss models. According to the Weiss model (which is now generally accepted),  $\text{O}_2$  binding is accompanied by a single electron transfer to give a doublet superoxide ligand ( $^2\text{O}_2^{\bullet-}$ ) and a low-spin Fe(III) centre. The smaller ionic radius for Fe(III) (high spin,  $0.65\text{\AA}$ ; low spin,  $0.55\text{\AA}$ ) results in a much better fit for the porphyrin, and no porphyrin ‘doming’ (Scheme 2.5). The Weiss model maintains that the Fe(III) centre is now low-spin ( $t_{2g})^5$ . This should result in paramagnetism both from the one remaining unpaired electron on the Fe(III) centre and the unpaired electron of the superoxide ligand. In fact the oxygenated haem system is found experimentally to be diamagnetic (this clearly rules out a high-spin Fe(III) centre, which would have five unpaired electrons). Weiss suggested a strong degree of spin correlation between the low-spin Fe(III) and  $^2\text{O}_2^{\bullet-}$  unpaired electrons, resulting in virtual diamagnetism for the low-spin case. The transition from high to low spin is explained by the increase in crystal field splitting energy as the Fe centre moves into the plane of the haem unit on  $\text{O}_2$  binding. This is termed the *trigger* mechanism. The alternative explanation, put forward by Pauling, suggests that a spin cross-over occurs such that  $\text{O}_2$  is bound by a low-spin Fe(II) centre in the singlet form  $^1\text{O}_2$  in the more conventional synergic fashion. Both components are diamagnetic and hence the complex as a whole is diamagnetic (Figure 2.17b). Again, the transition from high to low spin is explained by the trigger mechanism. The Weiss model is favoured by the observation of the infrared  $\nu(\text{O}—\text{O})$  stretch at  $1100\text{ cm}^{-1}$ , consistent with  $\text{O}_2^{\bullet-}$  rather than the higher wavenumber expected for the double bonded  $^1\text{O}_2$ . Furthermore, replacement of the iron centre in the haem with cobalt to give coboglobin (which



**Figure 2.17** (a) Crystal field orbital descriptions for high- and low-spin iron(II) and iron(III). (b) Ligand field description of the Pauling and Weiss models for oxygen binding by haemoglobin. The Weiss model involving superoxide binding to an iron(III) centre was eventually proven to be correct.



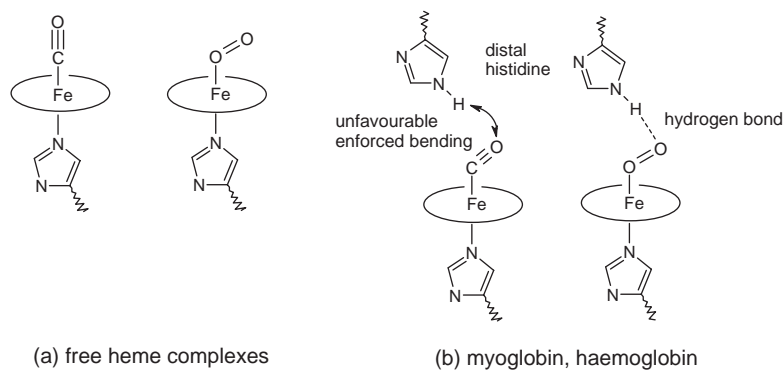
**Scheme 2.5** (a) Porphyrin doming relaxes upon  $\text{O}_2$  binding. (b) Oxygenated haem showing immediate protein environment including stabilising hydrogen bonds (reproduced with permission from Reference 1).



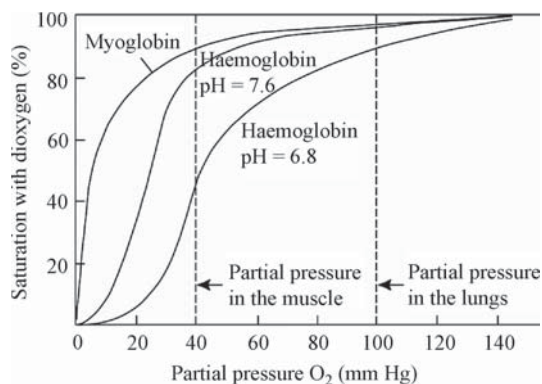
gives a system with one extra electron) gives EPR spectroscopic data that suggests that this extra electron resides on the oxygen. This results in the formulation  $\text{Co(III)—O}_2^{\bullet-}$ , which is in agreement with the Weiss model. X-ray photoelectron spectroscopy also suggests iron has an oxidation state of approximately 3.2 and hence the correct oxidation state of iron is thus the +3 state.

Evidence for the trigger mechanism comes from the X-ray crystal structure of a ‘picket fence porphyrin’ model compound. In this species (containing a 2-methyl imidazole ligand in one axial site as a model for the proximal histidine residue) the deoxy form is found to have a doming corresponding to the Fe ion being  $0.399\text{\AA}$  out of the plane formed by the four porphyrin N atoms.  $\text{O}_2$  binding causes the Fe to move to a position only  $0.086\text{\AA}$  out of the macrocycle (N) plane. An extensive range of porphyrin-based model compounds have been prepared to help understand the mechanism of operation of haemoglobin. These compounds are discussed in Section 12.6.1. It is surprising, however, that the binding of a weak ligand such as  $\text{O}_2$  should induce a spin cross-over from high to low spin at the metal centre. Indeed, porphyrin ligands usually induce a large ligand field, resulting in complexes always being of low spin. Thus the high-spin situation (which is vital for  $^3\text{O}_2$  binding if forbidden spin cross-overs are to be avoided) is something of an anomaly that can be explained only by the porphyrin doming and consequent weakening of the ligand field. The relatively light pull from the incoming  $\text{O}_2$  thus makes all the difference in energy terms between the strained, high-spin deoxygenated state and the low-spin oxygenated state. The porphyrin doming thus represents an example of enzymatic *entatic state*, in which the role of the bulk protein environment is to maintain the active site in a state of high energy, part way along the desired reaction coordinate. Metalloproteins in general are found very rarely with the metal ion in a relaxed, regular coordination environment, but are often balanced finely between two high-energy states requiring only small external influences to effect the cross-over from one to another.

It is well known that gases such as CO or readily adsorbed salts such as  $\text{CN}^-$  are extremely toxic. This is because they bind irreversibly to the Fe in haemoglobin preventing oxygen transport and resulting in rapid suffocation. CO in particular is a much better  $\pi$ -acceptor than  $\text{O}_2$  and hence is bound very much more strongly. The affinity of protein-free haem model systems for CO is much greater than for  $\text{O}_2$ , as expected:  $K_{\text{CO}}/K_{\text{O}_2} = 25\,000$ . In haemoglobin, however, this ratio is reduced to a more favourable 200 enabling the human body to at least tolerate small quantities of CO. This phenomenon is explained by the protein conformation restricting access to the iron binding site such that the geometry of the binding pocket is more suited to the bent  $\text{O}_2$  mode of binding than the linear CO form (Figure 2.18). In the oxygenated haem, the  $\text{O}_2$  molecule is bound through only one O atom with an  $\text{Fe—O—O}$  angle of about  $120^\circ$  (O—O distance  $1.89\text{\AA}$ ) as a consequence of the remaining lone pair on the binding O atom. The haem- $\text{O}_2$  interaction is stabilised by hydrogen bonding of the unbound O atom to a distal histidine NH functionality.



**Figure 2.18** Improved selectivity for  $\text{O}_2$  over CO is enforced by the protein environment.



**Figure 2.19** Oxygen saturation curves for myoglobin and haemoglobin and different pH values. (Reproduced by permission of John Wiley & Sons, Ltd).

One final point concerns the incorporation of four myoglobin units into the overall haemoglobin protein. Clearly myoglobin alone is quite capable of  $O_2$  uptake and transport, so why is it necessary to incorporate it into a tetrameric protein? The answer lies in the mutual interactions of the four myoglobin units in haemoglobin. This has the effect of synergically modifying the  $O_2$  uptake and decomplexation characteristics to make them more suitable for the  $O_2$  partial pressures found in the lungs and muscle. The behaviour of myoglobin and haemoglobin (Figure 2.19) shows that, without these cooperative characteristics, myoglobin would not be able to release  $O_2$  where it is needed in oxygen-deficient tissue such as muscle.

This cooperative behaviour relates to the enhancement of binding at one haem site if another is oxygenated or, if one haem centre loses its  $O_2$ , lowered oxygen affinity at the other iron centres. Long-range influences on binding at one site by binding at another of this type are termed *allosteric* effects and the sigmoidal binding profile is characteristic of their occurrence (see Section 1.5). The result is an all-or-nothing type of behaviour in that as soon as  $O_2$  levels are high enough to induce binding at one site (*i.e.* in the lung) all sites are quickly saturated. Upon  $O_2$  release in the muscle, decomplexation of all four sites follows rapidly upon release of  $O_2$  from any one of them. The mode of operation of this allosteric effect has been described by a 'spring loaded' model in which the haemoglobin as a whole alternates between two states termed *tense* (T) and *relaxed* (R). Hence each haem may be either oxy or deoxy and T or R. Binding of  $O_2$  to the domed deoxy T form causes a reduction in doming to give a strained oxy T form. Strain is released by movements of the proximal histidine residue to give a straighter Fe—N bond. This results in a movement of the protein backbone, which is transmitted to the remaining myoglobin units, encouraging  $O_2$  binding.

## 2.6 Enzymes and Coenzymes

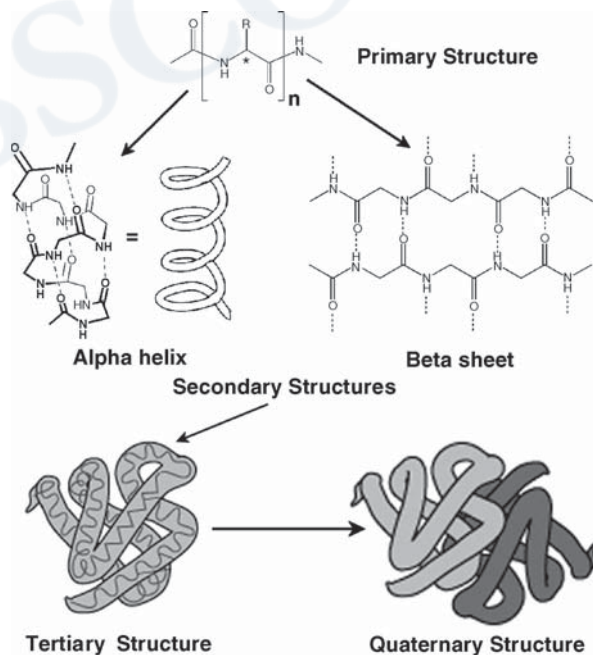
### 2.6.1 Characteristics of Enzymes

- ✦ Kim, D. H., 'Supramolecular aspects of enzymes', in *Comprehensive Supramolecular Chemistry*, J.L. Atwood, J.E.D. Davies, D.D. MacNicol and F. Vögtle (eds), Pergamon: Oxford, 1996, vol. 4, 503–526.

Enzymes are of vital importance and interest in biochemistry. It is enzymes that catalyse essentially all biological processes and, as we will see in Chapter 12, they have stimulated a wealth of interest from supramolecular chemists. Enzymes are dynamic macromolecules with a molecular weight

generally in excess of 10 000 Da. They are made up of polypeptide chains, which in turn are polymers of the protein amino acids (Box 2.2). The polypeptide chains are folded into a unique conformation giving a globular structure incorporating surface clefts and crevices. Substrate binding in enzymatic catalysis occurs within one of these clefts, termed the *active site* which often contains a metal ion. The initial binding is a thermodynamically controlled equilibrium process and is highly selective, usually proceeding on an induced fit basis. Binding occurs by three-dimensional contact between enzyme and substrate, and involves hydrophobic effects, hydrogen bonding, salt bridges (ion–ion) and other forms of intermolecular interaction. Bound substrates experience chemical transformation, sometimes at rates approaching the diffusion limit (the reaction occurs as fast as the substrate reaches the enzyme in accordance with the local concentration gradient and thermal fluctuation; about  $10^8 \text{ s}^{-1} \text{ M}^{-1}$ ). A key concept behind this rate enhancement in enzymatic catalysis is that of the *entatic state*. Bound substrates are held in a way that distorts them making them more like the transition state for the reaction the enzyme catalyses thus lowering the activation barrier for the reaction.

Enzyme structure may be divided into primary, secondary, tertiary and quaternary features, Figure 2.20. The primary structure is the sequence of amino acid residues on the polypeptide chain and is determined by the way in which the enzyme is synthesised. The secondary structure relates to ordering of the chains into discrete units or segments such as  $\alpha$ -helices and  $\beta$ -sheets, while tertiary structure is the way in which the secondary structural features arrange themselves to produce the full globular protein. This occurs *via* hydrogen bonding, stacking interactions and hydrophobic forces, and often involves the participation of water molecules buried deep within the enzyme, where they fill small cavities and act as a ‘glue’ holding secondary features together. In many cases more than one protein strand is involved in a fully functioning enzyme system and the supramolecular association



**Figure 2.20** Enzyme structure can be divided into primary (amino acid sequence), secondary (chain folding), tertiary (full 3D conformation) and quaternary (association of more than one protein chain) features.



**Figure 2.21** Eduard Buchner – father of enzyme chemistry.

of more than one protein molecule is termed quaternary structure (*e.g.* haemoglobin, a tetramer of four myoglobin units). It is the enzyme tertiary and quaternary structure that is responsible for the organisation of the binding site(s); it is also the most flexible component of the enzyme structure, allowing the binding site to deform in response to guest binding. Enzyme tertiary and quaternary structure is well-defined, however, and enzymes are ordered so highly that they can be crystallised. The way in which tertiary and quaternary structure is encoded within the primary amino acid sequence is a complicated and, as yet, unsolved mystery—one that challenges supramolecular chemists to address. Hydrophobic effects, resulting in an overall globular shape, are important, as are steric effects, which may be responsible for most of the organisation. Specific, directional hydrogen bonds then determine which of a small number of globular conformations is the one favoured.

Some of the earliest work on enzymes was by German chemist Eduard Buchner (Figure 2.21) working at Berlin University who in 1897 discovered that yeast extracts that without any living yeast cells were still able to ferment sugar. Buchner named the enzyme that brought about sucrose fermentation *zymase*. His work won him the 1907 Nobel Prize in Chemistry. In line with Buchner's ideas enzymes are generally named according to the reaction they carry out with the suffix *-ase* being added to the name of the substrate. Hence  $\text{Na}^+/\text{K}^+$ -ATPase is an enzyme that uses ATP and lactase is an enzyme that cleaves lactose. Enzymes can also be named after the type of reaction they catalyse, for example DNA polymerase forms polymeric DNA.

There exists a bewildering variety of protein-based enzymes ranging in size from a mere 62 amino acid residues for monomer of 4-oxalocrotonate tautomerase<sup>8</sup> to a whopping 2500 residues for fatty acid synthase.<sup>9</sup> According to the International Union of Biochemistry and Molecular Biology nomenclature for enzymes each enzyme is described by a sequence of four numbers preceded by 'EC'. The first number broadly classifies the enzyme based on its mechanism. The first-level general categories are listed below (see <http://www.chem.qmul.ac.uk/iubmb/enzyme/> for the complete nomenclature).

- EC 1 *Oxidoreductases*: catalyze oxidation/reduction reactions
- EC 2 *Transferases*: transfer a functional group (*e.g.* a methyl or phosphate group)
- EC 3 *Hydrolases*: catalyze the hydrolysis of various bonds
- EC 4 *Lyases*: cleave various bonds by means other than hydrolysis and oxidation
- EC 5 *Isomerases*: catalyze isomerisation changes within a single molecule
- EC 6 *Ligases*: join two molecules with covalent bonds

## 2.6.2 Mechanism of Enzymatic Catalysis

The overall process of enzyme catalysis may be represented by Equation 2.8, where E, S and P represent enzyme, substrate and product, respectively. Note that the enzyme is regenerated after the reaction, as required in a catalytic process.

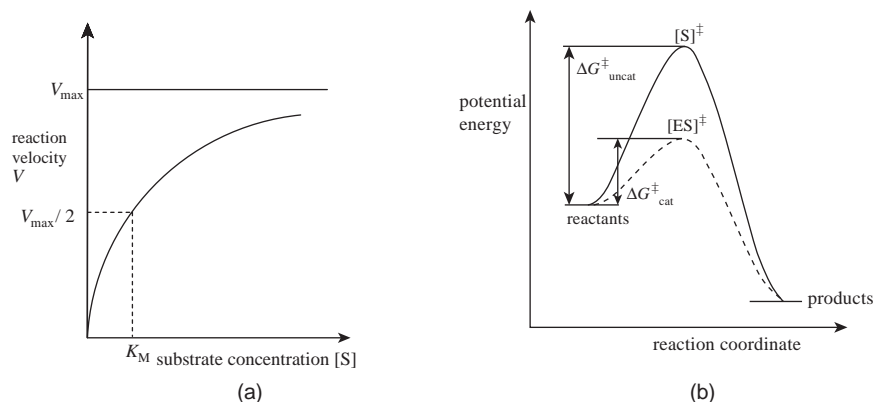


$$k_1/k_{-1} = K_{11} \quad (2.9)$$

The specificity and selectivity of a particular enzyme for competing substrates depends on both rate constants  $k_1$  and  $k_{-1}$  for each substrate (and hence on the equilibrium constant), and  $k_{\text{cat}}$ . This is usually expressed as an overall specificity constant,  $k_{\text{cat}}/K_{11}$ . Thus, the most specific enzymes carry out their catalysis rapidly, without the need for particularly strong binding. Many enzymes adhere to Michaelis-Menten kinetic parameters (named after Leonor Michaelis and Maud Menten who first described a model of enzyme reactivity crucially dependent on the existence of the ES complex shown in Equation 2.8). In the Michaelis-Menten model, the rate of catalysis ( $V$  – defined as the number of moles of product formed per second) increases with increasing concentration of substrate, S. For a constant enzyme concentration,  $V$  is linearly proportional to the concentration of S (denoted  $[S]$ ) when  $[S]$  is small. At high  $[S]$  (i.e. when S is in vast excess compared to the enzyme concentration),  $V$  is nearly independent of  $[S]$ . This gives a plot of  $V$  against  $[S]$  of the type shown in Figure 2.22a. Combining experimental observation with mathematical principles gives the famous Michaelis-Menten equation, Equation 2.10.

$$V = V_{\text{max}} \frac{[S]}{[S] + K_{\text{M}}} \quad (2.10)$$

The quantities  $K_{\text{M}}$  and  $V_{\text{max}}$  are important parameters that can be used to characterise and understand enzymatic reactivity. The  $K_{\text{M}}$  value is the *Michaelis constant* which is defined as the substrate concentration at which the reaction rate is half of its maximal value. Hence  $K_{\text{M}}$  (which varies considerably) is



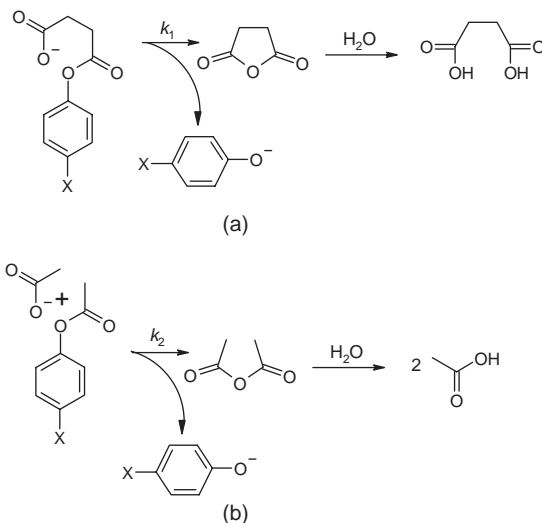
**Figure 2.22** (a) a plot of reaction velocity  $V$  against substrate concentration  $[S]$  for an enzyme that obeys Michaelis-Menten kinetics.  $V_{\text{max}}$  is the maximum reaction velocity and  $K_{\text{M}}$  is the Michaelis constant (b) Chemical reactions profiles for catalysed (···) and uncatalysed reactions (—).

a measure of the relative affinity of an enzyme for its substrate –the *higher* the  $K_M$  value the *lower* the affinity for the substrate. The  $V_{\max}$  parameter describes the maximum rate of product formation under conditions of high substrate concentration (*i.e.*, a large excess over the concentration of enzyme) where all of the active sites of an enzyme are occupied. The  $V_{\max}$  and  $K_M$  values can be determined by varying the concentration of S and looking at the amount of product formed. This information can be expressed as a Lineweaver-Burk plot, namely  $1/V$  vs.  $1/[S]$ . This graph gives a straight line with a y-axis intercept of  $1/V_{\max}$  and a slope of  $K_M/V_{\max}$ . The x-axis intercept gives  $1/K_M$ . This kind of plot can give a significant amount of information about the mechanism of enzyme inhibition by specific compounds.

There are exceptions to Michaelis-Menten behaviour. For example allosteric enzymes which instead of a hyperbolic curve in a  $V$  versus  $[S]$  graph yield a sigmoidal plot (the behaviour is rather like non-catalytic allosteric proteins, such as haemoglobin, Section 2.5. This type of curve can indicate *co-operative binding* of the substrate to the enzyme. We have discussed cooperativity in Section 1.5 (see also Section 10.4.3). In addition, regulatory molecules can further alter the activity of allosteric enzymes.

The key to an understanding of the way in which enzymes are able to achieve such large rate accelerations in the reactions they catalyse may be found in the insight of Linus Pauling, who stated in 1948 that ‘enzymes are molecules that are complementary in structure to the transition states of the reactions they catalyze’. This simple, but profound, assertion relates to the overall transition state theory in chemical reactions, as exemplified by the type of reaction coordinate profile shown in Figure 2.22b. The job of the enzyme catalyst is to lower the energy of the activated enzyme substrate complex  $[ES]^\ddagger$  and hence the free energy of activation,  $\Delta G_{cat}^\ddagger$  relative to the energy of the activated substrate,  $[S]^\ddagger$ , in the uncatalysed reaction. The enzyme must, therefore, be more complementary to the activated substrate than it is to the ground-state species. Indeed, the non-covalent forces involved in substrate binding should be sufficient to distort the substrate such that it proceeds some distance from left to right along the reaction coordinate. In other words, it becomes more like the transition state, hence lowering the activation energy required to form  $[ES]^\ddagger$ . This is a further manifestation of the induced fit model (Section 1.3). Not only does the enzyme undergo a conformational change on substrate binding, but so does the substrate (Figure 1.3b). This conformational change should result in strain at the scissile bond (the one that is to be broken in the catalysed reaction). As an example, the transition state in the deamination reaction of adenosine triphosphate is stabilised by around  $73.2 \text{ kJ mol}^{-1}$  at 310K. Despite the power of this approach, transition state stabilisation is not the only possibility when it comes to explanation of enzymatic rate enhancements. One other topical factor is *dynamical effects*. Rate enhancement from dynamical effects means that the enzyme has evolved to optimise a particular vibrational mode that moves the substrate to the transition state or moves the transition state towards the product. Dynamical effects are controversial and computational analysis does not support their having a significant role, however the idea is an interesting alternative.<sup>10</sup>

Another factor, in addition to transition-state stabilisation, that contributes to enzymatic acceleration is effective concentration. In general, in nonenzymatic chemistry, intramolecular reactions (such as cyclisation) are much faster than their bimolecular analogues because of statistical effects; for example, the intramolecular reaction of phenyl succinate is about  $10^5$  times faster than the bimolecular reaction of acetate with phenyl acetate at 1 M concentration of all species (Scheme 2.6). The ratio  $k_1/k_2$  is termed the effective molarity and its value of  $10^5 \text{ M}$  is clearly much larger than a real concentration (neat acetic acid is about 17 M), even though it notionally represents the acetic acid concentration required to achieve the same rate as the intramolecular equivalent. Compared to bimolecular reactions, unimolecular processes result only in a loss of internal rotational entropy and so may be favoured by as much as  $54\text{--}59 \text{ kJ mol}^{-1}$  on entropic grounds. In enzymes, if the binding of the substrate with the enzyme is strong enough, then the ES complex acts as if it were one molecule, thus intramolecular-type rate accelerations of the order of  $10^9$  are observed. In effect, the entropic ‘price’ required for bringing together two reactants is ‘paid for’ by the substrate binding, not during the reaction itself.



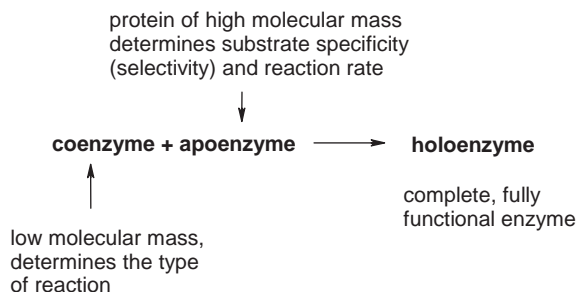
**Scheme 2.6** Intramolecular reactions (a) are much faster than bimolecular equivalents (b).

A final key factor in enzymatic rate acceleration is also the desolvation of the bound substrate, again 'paid for' in energy terms at the complexation stage. The enzyme functional groups and surface interactions effectively replace the water molecules formerly surrounding the substrate. It is often the case that the reactive group attached to the enzyme is one of these solvating moieties, and hence the reactants are brought within one another's first solvation sphere to form a contact pair with high affinity and are held in that way for long enough for the reaction to occur. This process is referred to as *spatiotemporal theory*.

Clearly, enzymes have evolved to carry out these special tasks in a highly selective fashion, and each enzyme family is very different from the others, despite their common amino acid building blocks, according to their particular task. One feature that enzymes share is that they are all very large and complicated molecules. This factor alone almost prohibits any opportunity for a full understanding of enzyme function in model systems since small-molecule models simply cannot begin to approach the complexity of the real species. So, is this large size really necessary, or is it an evolutionary accident? It is possible that the size of enzymes is simply the minimum needed to achieve the required binding site cleft geometry and orientation through the use of folded peptide interactions. Furthermore, the binding site must possess some controlled, well-defined flexibility to accommodate the movements of the substrate as it undergoes chemical transformations. This is one of the reasons why it is constructed using malleable protein tertiary structure. Also perhaps of importance is the large number of functional groups on the enzyme surface remote from the binding site. These may serve to trap potential substrates and enhance their rate of diffusion towards the active site. Often, the very largest enzymes are the most specific, having other polypeptides as their substrates. It has been estimated that an interacting surface area of at least 10 nm<sup>2</sup> is required for such interactions to be of significance. It is likely that small proteins with molecular weights of less than 10 000 Da and total surface areas of under 15 nm<sup>2</sup> simply do not have the available contact area to be effective.

### 2.6.3 Coenzymes

While many enzymes can function by themselves, others work in a modular way with a coenzyme. A coenzyme is a non-enzyme 'helper molecule' that forms one constituent of a biological catalysis system. In order to function, the full system requires the *coenzyme*, an *apoenzyme* and a *substrate*. A chemical

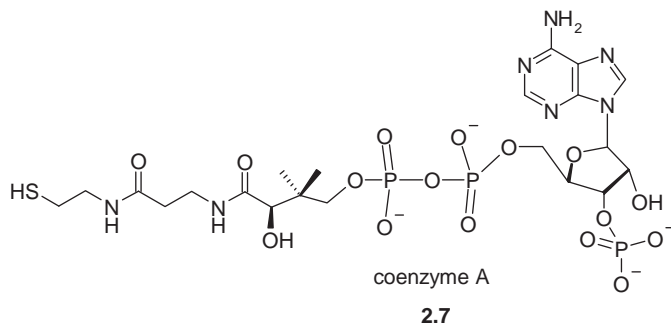


**Scheme 2.7** Modular construction of holoenzymes.

reaction occurs between the coenzyme and the substrate while they are both bound by the apoenzyme. A combined apoenzyme and coenzyme is termed a *holoenzyme* and this is the active species (Scheme 2.7).

Common examples of coenzymes include ATP (the biological energy storage molecule) and NADH (the reduced form of the biological electron carrier) which are a ubiquitous feature of all living systems suggesting that these molecules emerged very early in evolutionary terms. Interestingly the nucleotide adenosine is found in coenzymes that are involved in the catalysis of a broad range of fundamental metabolic reactions such as the transfer of methyl, acyl, and phosphoryl groups, and redox reactions. It has been suggested that this prevalence makes adenosine a chemical fossil remnant of the RNA world in which catalysis was carried out by ribonucleic acid enzymes (ribozymes) rather than proteins, which evolved to bind a restricted set of nucleotides. Adenosine-based coenzymes might have acted as interchangeable adaptors that allowed ribozymes to link to different coenzymes without major modification to existing adenosine-binding regions. We will return to the RNA world theory in Section 12.9.4.

Coenzymes can be either inorganic species such as coenzyme F<sub>450</sub> and or purely organic as in coenzyme A (2.7), the coenzyme that carries acyl groups in the synthesis and oxidation of fatty acids, and is responsible for oxidation of pyruvate in the citric acid cycle. Vitamins are commonly coenzymes or their precursors. Coenzyme A is vitamin B<sub>5</sub> and we will look in detail at cobalamin, vitamin B<sub>12</sub>, by way of an example in the next section. Some examples of vitamin and non-vitamin coenzymes are shown in Table 2.3.



## 2.6.4 The Example of Coenzyme B<sub>12</sub>

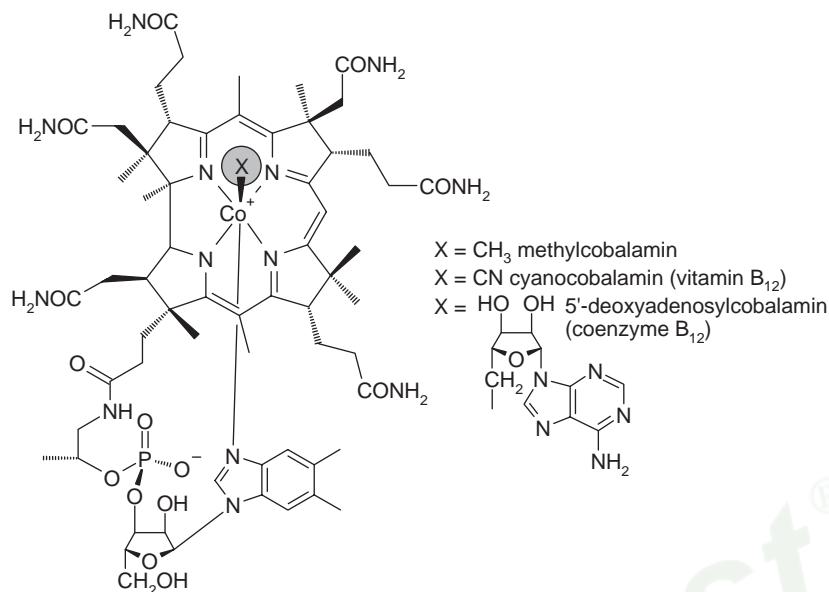
To give an example of how a coenzyme functions we will look at the example of coenzyme B<sub>12</sub> which has proved a popular target for model studies over many years. Coenzyme B<sub>12</sub> and its derivatives such as vitamin B<sub>12</sub> are also based on tetrapyrrole macrocycles, Figure 2.10. The structures of these



**Table 2.3** Role and distribution for some vitamin and non-vitamin coenzymes (adapted from <http://en.wikipedia.org/wiki/Coenzyme>).

Vitamins and derivatives				
Coenzyme	Vitamin	Additional component	Chemical group(s) transferred	Distribution
NAD <sup>+</sup> and NADP <sup>+</sup>	Niacin (B <sub>3</sub> )	ADP	Electrons	Bacteria, archaea and eukaryotes
Coenzyme A	Pantothenic acid (B <sub>5</sub> )	ADP	Acetyl group and other acyl groups	Bacteria, archaea and eukaryotes
Tetrahydrofolic acid	Folic acid (B <sub>9</sub> )	Glutamate residues	Methyl, formyl, methylene and formimino groups	Bacteria, archaea and eukaryotes
Menaquinone	Vitamin K	None	Carbonyl group and electrons	Bacteria, archaea and eukaryotes
Ascorbic acid	Vitamin C	None	Electrons	Bacteria, archaea and eukaryotes
Coenzyme F420	Riboflavin (B <sub>2</sub> )	Amino acids	Electrons	Methanogens and some bacteria
Non-vitamins				
Coenzyme	Chemical group(s) transferred		Distribution	
Adenosine triphosphate	Phosphate group		Bacteria, archaea and eukaryotes	
S-Adenosyl methionine	Methyl group		Bacteria, archaea and eukaryotes	
3'-Phosphoadenosine-5'-phosphosulfate	Sulfate group		Bacteria, archaea and eukaryotes	
Coenzyme Q	Electrons		Bacteria, archaea and eukaryotes	
Cytidine triphosphate	Diacylglycerols and lipid head groups		Bacteria, archaea and eukaryotes	
Nucleotide sugars	Monosaccharides		Bacteria, archaea and eukaryotes	
Glutathione	Electrons		Some bacteria and most eukaryotes	
Coenzyme M	Methyl group		Methanogens	
Coenzyme B	Electrons		Methanogens	
Methanofuran	Formyl group		Methanogens	

compounds are shown in Figure 2.23 and are all based on cobalt(III) complexes of the corrin ring, which, when acting as a ligand, possesses a single negative charge. The remaining charge on the cobalt is balanced by an integral phosphate residue and a substituent, X, which is the enzymatic 'business end' of the molecule. In the active form of the coenzyme, this substituent is a carbanion. Coenzyme B<sub>12</sub> and its derivatives are, therefore, the only naturally occurring true organometallic compounds.

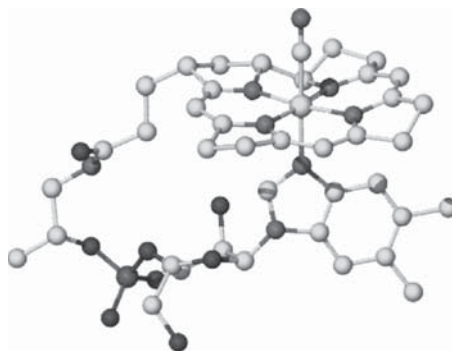


**Figure 2.23** Coenzyme B<sub>12</sub> and its derivatives.

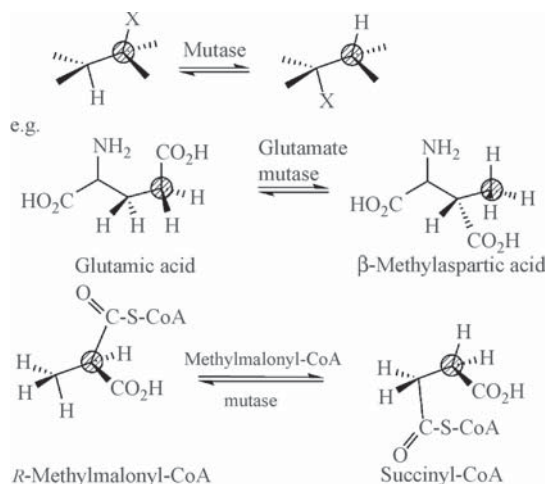
Remarkably the Co—C bond is very stable under physiological conditions (pH 7, oxygenated aqueous solution).

The modular nature of holoenzyme catalysis allows coenzyme B<sub>12</sub> to be used in a variety of biological reactions with different apoenzymes. These reactions can involve homolysis of the Co—C bond to give an alkyl radical resulting in radical induced rearrangements, redox reactivity *via* reduction to Co(II) and Co(I), and alkylations. The nature of the coenzyme determines the type of the reaction, while the nature of the apoenzyme determines the selectivity of the reaction in terms of the substrate and the regioselectivity.

The B<sub>12</sub> class compounds are required in only very small amounts, consistent with the poor bioavailability of cobalt, and are used mainly by microorganisms, which can also synthesise them. In mammals, the B<sub>12</sub>-dependent methylmalonyl-CoA mutase system is of particular importance, however. It is required for amino acid metabolism in the liver and its absence, as a result of genetic defects, is lethal. Indeed, as early as the 1920s it was observed that extracts from animal livers could cure otherwise lethal pernicious anaemia. The presence of cobalt was detected in these liver extracts and this led to



**Figure 2.24** X-ray crystal structure of vitamin B<sub>12</sub> (substituents omitted for clarity).



**Scheme 2.8** 1,2-Shifts by coenzyme B<sub>12</sub> mutases.

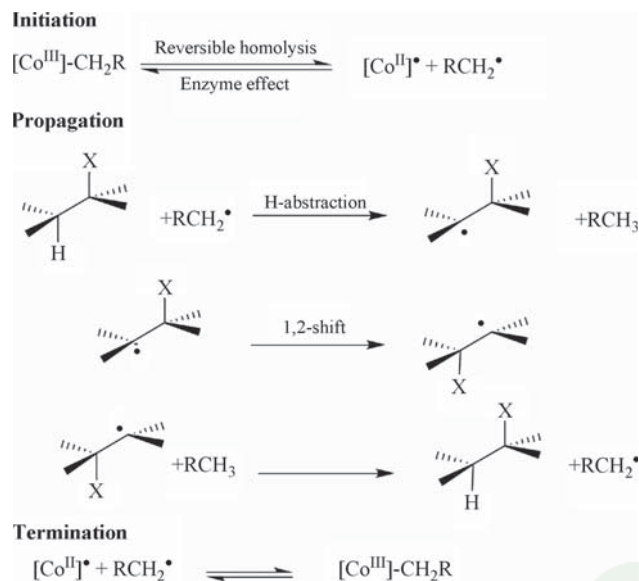
a great deal of work to isolate this unusual cobalt-containing compound, which is present in blood in concentrations of about 0.01 mg L<sup>-1</sup>. Using chromatographic methods and a great deal of perseverance, cyanocobalamin was finally isolated in pure form in 1948. The CN<sup>-</sup> group is not part of the active form of the complex, but is an artefact of the isolation procedure. However, the cyano derivative is still therapeutically useful and the new isolated material was named vitamin B<sub>12</sub>. The X-ray crystal structure of this rather large ‘small molecule’ (RMM 1350 Da) (Figure 2.24) earned the 1964 Nobel Prize in Chemistry for Dorothy Crowfoot-Hodgkin. The structure demonstrated a bent, ‘butterfly’ shape of the corrin ring, which is a consequence of its small size (15-atom ring instead of the usual 16 in porphyrins). This distortion is of relevance to the reactivity of the coenzyme and represents an entatic state structure.

The biochemistry of coenzyme B<sub>12</sub> generally revolves around either mutase enzyme activity, involving functional group migration, notably by stereospecific 1,2-shifts (Scheme 2.8), or methylation by methionine synthetase. The general mechanism for the mutase activity is a radical-based one and has been established by EPR spectroscopy to be of the general form shown in Scheme 2.9.

## 2.7 Neurotransmitters and Hormones

- ➔ D.A. Dougherty, ‘Is the brain ready for physical organic chemistry?’, *J. Phys. Org. Chem.*, 1998, **11**, 334–340.

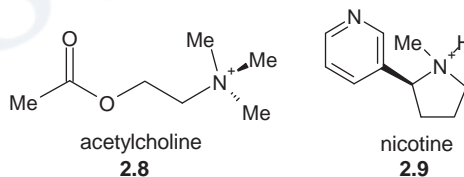
In addition to the bioinorganic chemistry involving metal cations in various forms, there exists a vast amount of biological chemistry that, in a supramolecular sense, involves molecular species and anions as the guests. While the detail of this chemistry is beyond the scope of this book, we must mention in passing the role of two particular classes of biological guest, namely hormones and neurotransmitters, which, broadly speaking, act as messengers and activating agents. Hormones such as the human sex hormones, oestrogen and testosterone, and the pregnancy hormone, progesterone, are household names. A range of testosterone derivatives, the anabolic–androgenic steroids, are capable of altering human physical performance significantly and as a result are highly topical within the context of their illegal use in competitive sports. Similarly, levels of neurotransmitters such as dopamine and acetylcholine may have dramatic effects on brain chemistry and hence behaviour, and



**Scheme 2.9** Radical mechanism for mutase activity.

imbalances are implicated in a number of mental disorders such as schizophrenia and Parkinson's disease, highlighting the chemical nature of even our very thought processes.

To illustrate the supramolecular biological role of such molecular species we will look at the mode of action of acetylcholine (ACh, **2.8**) in particular since this example highlights the profound effect that the non-covalent binding of a small molecular guest may have on a biological system.



Acetylcholine is a key neurotransmitter molecule. Nerve pulses are passed on from nerve fibre to nerve fibre across the synapses (the gaps between nerve cells, about 30–40 nm thick) by transfer of acetylcholine from one nerve cell to the next. The role of acetylcholine is to bind to, and hence open, a gated sodium ion channel (*cf.* Section 2.2), which forms part of the nicotinic acetylcholine receptor protein (nAChR). This large transmembrane protein is composed of five subunits (two labelled  $\alpha$ , and the others labelled  $\beta$ ,  $\delta$ ,  $\gamma$ ; Figure 2.25) with similar amino acid sequences, which encircle a central channel. Acetylcholine binds to the two  $\alpha$ -subunits, causing a conformational change in the channel external region, which allows access to  $\text{Na}^+$  ions located in the synapse. These then flow rapidly into the cell in accordance with the prevailing concentration gradient, resulting in a current flow. The binding of the agonist acetylcholine to its receptor is highly selective and involves cation– $\pi$  interactions (Section 1.8.5) between the quaternary ammonium moiety and a particular tryptophan amino acid residue. Interestingly, acetylcholine activity is affected severely by nicotine (**2.9**), which also possesses a quaternary ammonium residue, and it is likely that this is the basis for nicotine addiction in smokers. Once the membrane has depolarised and the neural current has been generated, acetylcholine is removed by another acetylcholine-binding protein, acetylcholine esterase, which hydrolyses the ester functionality preventing the molecule from binding to nAChR.