19 Materials 1: macromolecules and aggregates

Macromolecules exhibit a range of properties and problems that illustrate a wide variety of physical chemical principles. They need to be characterized in terms of their molar mass, their size, and their shape. However, the molecules are so large and the solutions they form depart so strongly from ideality, that techniques for accommodating these departures from ideality need to be developed. Another major problem concerns the influences that determine the shapes of the molecules. We consider a range of influences in this chapter, beginning with a structureless random coil and ending with the structurally precise forces that operate in polypeptides and nucleic acids. Atoms, small molecules, and macromolecules can form large assemblies that are held together by one or more of the molecular interactions described in Chapter 18. These assemblies, which include colloids and biological membranes, exhibit some of the typical properties of molecules but have their own characteristic features.

There are macromolecules everywhere, inside us and outside us. Some are natural: they include polysaccharides such as cellulose, polypeptides such as protein enzymes, and polynucleotides such as deoxyribonucleic acid (DNA). Others are synthetic: they include **polymers** such as nylon and polystyrene that are manufactured by stringing together and (in some cases) cross-linking smaller units known as **monomers**. Life in all its forms, from its intrinsic nature to its technological interaction with its environment, is the chemistry of macromolecules.

Macromolecules give rise to special problems that include the shapes and the lengths of polymer chains, the determination of their sizes, and the large deviations from ideality of their solutions. Natural macromolecules differ in certain respects from synthetic macromolecules, particularly in their composition and the resulting structure, but the two share a number of common properties. We concentrate on these common properties here. Another level of complexity arises when small molecules aggregate into large particles in a process that is called 'self-assembly' and give rise to aggregates. One example is the assembly of haemoglobin from four myoglobin-like polypeptides. A similar type of aggregation gives rise to a variety of disperse phases, which include colloids. The properties of these disperse phases resemble to a certain extent the properties of solutions of macromolecules, and we describe their common attributes in the final part of this chapter.

Determination of size and shape

X-ray diffraction techniques (Chapter 20) can reveal the position of almost every heavy atom (that is, every atom other than hydrogen) even in very large molecules. However, there are several reasons why other techniques must also be used. In the first

place, the sample might be a mixture of molecules with different chain lengths and extents of cross-linking, in which case sharp X-ray images are not obtained. Even if all the molecules in the sample are identical, it might prove impossible to obtain a single crystal, which is essential for diffraction studies because only then does the electron density (which is responsible for the scattering) have a large-scale periodic variation. Furthermore, although work on proteins and DNA has shown how immensely interesting and motivating the data can be, the information is incomplete. For instance, what can be said about the shape of the molecule in its natural environment, a biological cell? What can be said about the response of its shape to changes in its environment?

19.1 Mean molar masses

A pure protein is **monodisperse**, meaning that it has a single, definite molar mass (although there may be small variations, such as one amino acid replacing another, depending on the source of the sample). A synthetic polymer, however, is **polydisperse**, in the sense that a sample is a mixture of molecules with various chain lengths and molar masses. The various techniques that are used to measure molar masses result in different types of mean values of polydisperse systems.

The mean obtained from the determination of molar mass by osmometry (Section 5.5) is the **number-average molar mass**, $\bar{M}_n$, which is the value obtained by weighting each molar mass by the number of molecules of that mass present in the sample:

$$\bar{M}_n = \frac{1}{N}\sum_i N_i M_i \tag{19.1}$$

where N_i is the number of molecules with molar mass M_i and there are N molecules in all. Viscosity measurements give the **viscosity-average molar mass**, $\bar{M}_v$, light-scattering experiments give the **weight-average molar mass**, $\bar{M}_w$, and sedimentation experiments give the ***Z*-average molar mass**, $\bar{M}_Z$. (The name is derived from the z-coordinate used to depict data in a procedure for determining the average.) Although such averages are often best left as empirical quantities, some may be interpreted in terms of the composition of the sample. Thus, the weight-average molar mass is the average calculated by weighting the molar masses of the molecules by the mass of each one present in the sample:

$$\bar{M}_w = \frac{1}{m}\sum_i m_i M_i \tag{19.2}$$

In this expression, m_i is the total mass of molecules of molar mass M_i and m is the total mass of the sample. Because $m_i = N_i M_i / N_A$, we can also express this average as

$$\bar{M}_w = \frac{\sum_i N_i M_i^2}{\sum_i N_i M_i} \tag{19.3}$$

This expression shows that the weight-average molar mass is proportional to the mean square molar mass. Similarly, the Z-average molar mass can be interpreted in terms of the mean cubic molar mass:

$$\bar{M}_Z = \frac{\sum_i N_i M_i^3}{\sum_i N_i M_i^2} \tag{19.4}$$

Example 19.1 *Calculating number and mass averages*

Determine the number-average and the weight-average molar masses for a sample of poly(vinyl chloride) from the following data:

Molar mass interval/ (kg mol^{-1})	Average molar mass within interval/(kg mol^{-1})	Mass of sample within interval/g
5–10	7.5	9.6
10–15	12.5	8.7
15–20	17.5	8.9
20–25	22.5	5.6
25–30	27.5	3.1
30–35	32.5	1.7

Method The relevant equations are eqns 19.1 and 19.2. Calculate the two averages by weighting the molar mass within each interval by the number and mass, respectively, of the molecule in each interval. Obtain the numbers in each interval by dividing the mass of the sample in each interval by the average molar mass for that interval. Because the number of molecules is proportional to the amount of substance (the number of moles), the number-weighted average can be obtained directly from the amounts in each interval.

Answer The amounts in each interval are as follows:

Interval	5–10	10–15	15–20	20–25	25–30	30–35
Molar mass/(kg mol^{-1})	7.5	12.5	17.5	22.5	27.5	32.5
Amount/mol	1.3	0.70	0.51	0.25	0.11	0.052
					Total:	2.92

The number-average molar mass is therefore

$$\begin{aligned}\bar{M}_n/(\text{kg mol}^{-1}) &= \frac{1}{2.92}(1.3\times 7.5 + 0.70\times 12.5 + 0.51\times 17.5 + 0.25\times 22.5 \\ &\quad + 0.11\times 27.5 + 0.052\times 32.5) \\ &= 13\end{aligned}$$

The weight-average molar mass is calculated directly from the data after noting that the total mass of the sample is 37.6 g:

$$\begin{aligned}\bar{M}_w/(\text{kg mol}^{-1}) &= \frac{1}{37.6}(9.6\times 7.5 + 8.7\times 12.5 + 8.9\times 17.5 + 5.6\times 22.5 \\ &\quad + 3.1\times 27.5 + 1.7\times 32.5) \\ &= 16\end{aligned}$$

Note the significantly different values of the two averages. In this instance, $\bar{M}_w/\bar{M}_n = 1.2$.

Self-test 19.1 Evaluate the *Z*–average molar mass of the sample. [19 kg mol^{-1}]

The ratio $\bar{M}_w/\bar{M}_n$ is called the **heterogeneity index** (or 'polydispersity index'). In the determination of protein molar masses we expect the various averages to be the

same because the sample is monodisperse (unless there has been degradation). A synthetic polymer normally spans a range of molar masses and the different averages yield different values. Typical synthetic materials have $\bar{M}_w/\bar{M}_n \approx 4$. The term 'monodisperse' is conventionally applied to synthetic polymers in which this index is less than 1.1; commercial polyethylene samples might be much more heterogeneous, with a ratio close to 30. One consequence of a narrow molar mass distribution for synthetic polymers is often a higher degree of three-dimensional long-range order in the solid and therefore higher density and melting point. The spread of values is controlled by the choice of catalyst and reaction conditions. In practice, it is found that long-range order is determined more by structural factors (branching, for instance) than by molar mass.

Average molar masses may be determined by osmotic pressure of polymer solutions . The upper limit for the reliability of membrane osmometry is about 1000 kg mol^{-1}. A major problem for macromolecules of relatively low molar mass (less than about 10 kg mol^{-1}) is their ability to percolate through the membrane. One consequence of this partial permeability is that membrane osmometry tends to overestimate the average molar mass of a polydisperse mixture. Several techniques for the determination of molar mass and polydispersity that are not so limited include mass spectrometry, laser light scattering, ultracentrifugation, electrophoresis, and viscosity measurements.

19.2 Mass spectrometry

Mass spectrometry is among the most accurate techniques for the determination of molar masses. The procedure consists of ionizing the sample in the gas phase and then measuring the mass-to-charge number ratio (m/z) of all ions. Macromolecules present a challenge because it is difficult to produce gaseous ions of large species without fragmentation. However, two new techniques have emerged that circumvent this problem: **matrix-assisted laser desorption/ionization** (MALDI) and **electrospray ionization**. We shall discuss **MALDI-TOF mass spectrometry**, so called because the MALDI technique is coupled to a time-of-flight (TOF) ion detector.

Figure 19.1 shows a schematic view of a MALDI-TOF mass spectrometer. The macromolecule is first embedded in a solid matrix that often consists of an organic material such as *trans*-3-indoleacrylic acid and inorganic salts such as sodium chloride or silver trifluoroacetate. This sample is then irradiated with a pulsed laser, such as a nitrogen laser. The laser energy ejects electronically excited matrix ions, cations, and neutral macromolecules, thus creating a dense gas plume above the sample surface. The macromolecule is ionized by collisions and complexation with small cations, such as H^+, Na^+, and Ag^+.

In the TOF spectrometer, the ions are accelerated over a short distance d by an electrical field of strength $\mathcal{E}$ and then travel through a drift region of length l. The time, t, required for an ion of mass m and charge number z to reach the detector at the end of the drift region is (see the *Justification*):

$$t = l\left(\frac{m}{2ze\mathcal{E}d}\right)^{1/2} \tag{19.5}$$

where e is the fundamental charge. Because d, l, and $\mathcal{E}$ are fixed for a given experiment, the time of flight, t, of the ion is a direct measure of its m/z ratio, which is given by:

$$\frac{m}{z} = 2e\mathcal{E}d\left(\frac{t}{l}\right)^2 \tag{19.6}$$

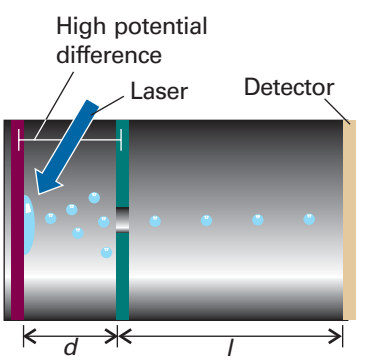


Fig. 19.1 A matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer. A laser beam ejects macromolecules and ions from the solid matrix. The ionized macromolecules are accelerated by an electrical potential difference over a distance d and then travel through a drift region of length l. Ions with the smallest mass to charge ratio (m/z) reach the detector first.

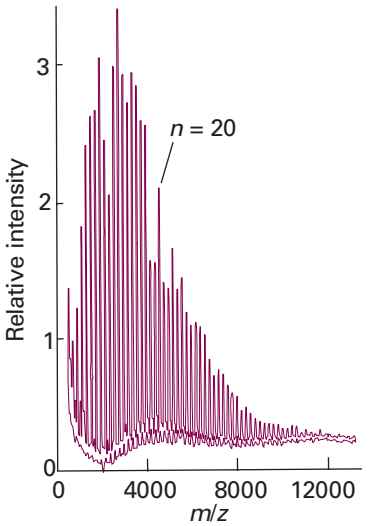


Fig. 19.2 MALDI-TOF spectrum of a sample of poly(butylene adipate) with $\bar{M}_n = 4525$ g mol^{-1} (Adapted from Mudiman *et al.*, *J. Chem. Educ.*, **74**, 1288 (1997).)

Justification 19.1 *The time of flight of an ion in a mass spectrometer*

Consider an ion of charge ze and mass m that is accelerated from rest by an electric field of strength $\mathcal{E}$ applied over a distance d. The kinetic energy, E_K, of the ion is

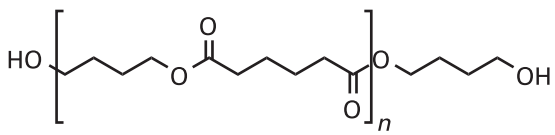
$$E_K = \tfrac{1}{2}mv^2 = ze\mathcal{E}d$$

where v is the speed of the ion. The drift region, l, and the time of flight, t, in the mass spectrometer are both sufficiently short that we can ignore acceleration and write $v = l/t$. Then substitution into this equation gives

$$\tfrac{1}{2}m\left(\frac{l}{t}\right)^2 = ze\mathcal{E}d$$

Rearrangement of this equation gives eqn 19.6.

Figure 19.2 shows the MALDI-TOF mass spectrum of a polydisperse sample of poly(butylene adipate) (PBA, **1**). The MALDI technique produces mostly singly charged molecular ions that are not fragmented. Therefore, the multiple peaks in the spectrum arise from polymers of different lengths, with the intensity of each peak being proportional to the abundance of each polymer in the sample. Values of $\bar{M}_n$, $\bar{M}_w$, and the heterogeneity index can be calculated from the data. It is also possible to use the mass spectrum to verify the structure of a polymer, as shown in the following example.



1

Example 19.2 *Interpreting the mass spectrum of a polymer*

The mass spectrum in Fig. 19.2 consists of peaks spaced by 200 g mol^{-1}. The peak at 4113 g mol^{-1} corresponds to the polymer for which $n = 20$. From these data, verify that the sample consists of polymers with the general structure given by (**1**).

Method Because each peak corresponds to a different value of n, the molar mass difference, ΔM, between peaks corresponds to the molar mass, M, of the repeating unit (the group inside the brackets in **1**). Furthermore, the molar mass of the terminal groups (the groups outside the brackets in **1**) may be obtained from the molar mass of any peak by using

$$M(\text{terminal groups}) = M(\text{polymer with } n \text{ repeating units}) - n\Delta M - M(\text{cation})$$

where the last term corresponds to the molar mass of the cation that attaches to the macromolecule during ionization.

Answer The value of ΔM is consistent with the molar mass of the repeating unit shown in (**1**), which is 200 g mol^{-1}. The molar mass of the terminal group is calculated by recalling that Na^+ is the cation in the matrix:

$$M(\text{terminal group}) = 4113 \text{ g mol}^{-1} - 20(200 \text{ g mol}^{-1}) - 23 \text{ g mol}^{-1} = 90 \text{ g mol}^{-1}$$

The result is consistent with the molar mass of the —$O(CH_2)_4OH$ terminal group (89 g mol^{-1}) plus the molar mass of the —H terminal group (1 g mol^{-1}).

Self-test 19.2 What would be the molar mass of the $n = 20$ polymer if silver trifluoroacetate were used instead of NaCl in the preparation of the matrix?

[4198 g mol^{-1}]

19.3 Laser light scattering

Large particles scatter light very efficiently. A familiar example is the light scattered by specks of dust in a sunbeam. Therefore, light scattering is a convenient method for the characterization of polymers, large aggregates (such as colloids), and biological systems from proteins to viruses. Unlike mass spectrometry, laser light scattering measurements may be performed in nearly intact samples; often the only preparation required is filtration of the sample.

(a) General principles of light scattering

When the oscillating electric field of electromagnetic radiation interacts with the electrons in a particle, an oscillating dipole moment develops with a magnitude proportional to the polarizability of the particle and the strength of the field (Section 18.1). Elastic light scattering is observed as the oscillating dipoles in the particle radiate at the same frequency as the frequency of the exciting electromagnetic radiation. The term *elastic* refers to the fact that the incident and scattered photons have the same frequency and hence the same energy. If the medium is perfectly homogeneous, as in a perfect crystal, the scattered waves interfere destructively in all directions except the direction of propagation of the exciting radiation. If the medium is inhomogeneous, as in an imperfect crystal or a solution of macromolecules, radiation is scattered into other directions as well.

Scattering of light by particles with diameters much smaller than the wavelength of the incident radiation is called **Rayleigh scattering** (Fig. 19.3). This type of scattering has several characteristic features.

1 The intensity of scattered light is proportional to λ^{-4}, so shorter wavelength radiation is scattered more intensely than longer wavelengths.

2 The intensity of scattered light is proportional to the molar mass of the particle.

3 The intensity of scattered light depends on the scattering angle θ (Fig. 19.3). In practice, data are collected at several angles to the incident laser beam (Example 19.3).

4 For very dilute solutions excited by plane-polarized light, the **Rayleigh ratio**, R_θ, a measure of the intensity of scattered light at a given scattering angle θ, is defined as

$$R_\theta = \frac{I}{I_0} \times \frac{r^2}{\sin^2\phi} \tag{19.7}$$

where I is the intensity of scattered light, I_0 is the intensity of incident light, r is the distance between the sample and the detector, ϕ is the angle between the plane of polarization of the incident beam and the plane defined by the incident and scattered beams (see the inset in Fig. 19.3).

For a solution of a polymer of mass concentration c_P, the Rayleigh ratio may be written as

$$R_\theta = K P_\theta c_P \bar{M}_w, \text{ with } \quad K = \frac{4\pi^2 n_{r,0}^2 V (\mathrm{d}n_r/\mathrm{d}c_P)^2}{\lambda^4 N_A} \tag{19.8}$$

Here $n_{r,0}$ is the refractive index of the pure solvent (see *Comment* 18.6 and *Appendix* 3), $(\mathrm{d}n/\mathrm{d}c_P)$ is the change in refractive index of the solution with concentration of polymer, V is the volume of the sample, and N_A is Avogadro's constant. The parameter P_θ is the **structure factor**, which takes into account the fact that scattering may occur from different sites of the same molecule and interference between scattered rays becomes important when the wavelength of the incident radiation is comparable to the size of the scattering particles. When the molecule is much smaller than the wavelength of incident radiation, $P_\theta \approx 1$. However, when the size of the molecule is

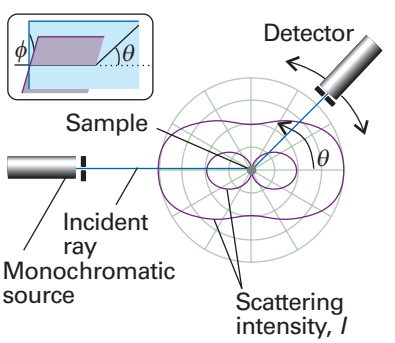


Fig. 19.3 Rayleigh scattering from a sample of point-like particles. The intensity of scattered light depends on the angle θ between the incident and scattered beams. The inset shows the angle ϕ between the plane of polarization of the incident beam and the plane defined by the incident and scattered beams. In a typical experimental arrangement, $\phi = 90°$.

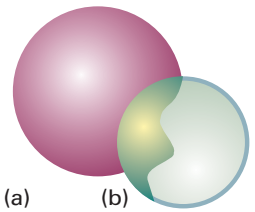


Fig. 19.4 (a) A spherical molecule and (b) the hollow spherical shell that has the same rotational characteristics. The radius of the hollow shell is the radius of gyration of the molecule.

Synoptic table 19.1* Radius of gyration

	M/(kg mol^{-1})	R_g/nm
Serum albumin	66	2.98
Polystyrene	3.2×10^3	50†
DNA	4×10^3	117

* More values are given in the *Data section*.
† In a poor solvent.

about one-tenth the wavelength of the incident radiation, we show in *Further information* 19.1 that

$$P_\theta \approx 1 - \frac{16\pi^2 R_g^2 \sin^2 \frac{1}{2}\theta}{3\lambda^2} \tag{19.9}$$

where R_g is the **radius of gyration** of the macromolecule, the radius of a thin hollow spherical shell of the same mass and moment of inertia as the molecule (Fig. 19.4 and Section 19.8). Table 19.1 lists some experimental values of R_g.

Illustration 19.1 *Why is the sky blue whereas clouds are white?*

We expect from eqn 19.9 that when the particles are very small and $P_\theta \approx 1$, the medium scatters light of shorter wavelengths much more efficiently than light of longer wavelengths. This effect accounts for the colour of a cloudless sky: the N_2 and O_2 molecules in the atmosphere are much smaller than the wavelengths of visible electromagnetic radiation, so blue light is scattered preferentially. We also see clouds because light scatters from them, but they look white, not blue. In clouds, the water molecules group together into droplets of a size comparable to the wavelength of light, and scatter cooperatively. Although blue light scatters more strongly, more molecules can contribute cooperatively when the wavelength is longer (as for red light), so the net result is uniform scattering for all wavelengths: white light scatters as white light. This paper looks white for the same reason. As a result, the scattering intensity is distorted from the form characteristic of small-particle, Rayleigh scattering, and the distortion is taken into account by values of P_θ that differ from 1.

(b) Scattering by non-ideal solutions of polymers

The preceding discussion shows that structural properties, such as size and the weight-average molar mass of a macromolecule, can be obtained from measurements of light scattering by a sample at several angles θ relative to the direction of propagation on an incident laser beam. However, eqn 19.8 applies only to ideal solutions. In practice, even relatively dilute polymer dispersions can deviate considerably from ideality. Being so large, macromolecules displace a large quantity of solvent instead of replacing individual solvent molecules with negligible disturbance. In thermodynamic terms, the displacement and reorganization of solvent molecules implies that the entropy change is especially important when a macromolecule dissolves. Furthermore, its great bulk means that a macromolecule is unable to move freely through the solution because the molecule is excluded from the regions occupied by other solute molecules. There are also significant contributions to the Gibbs energy from the enthalpy of solution, largely because solvent–solvent interactions are more favourable than the macromolecule–solvent interactions that replace them. To take deviations from ideality into account, it is common to rewrite eqn 19.8 as

$$\frac{Kc_P}{R_\theta} = \frac{1}{P_\theta \bar{M}_w} + Bc_P \tag{19.10}$$

where B is an empirical constant analogous to the osmotic virial coefficient and indicative of the effect of excluded volume.

For most solute–solvent systems there is a unique temperature (which is not always experimentally attainable) at which the effects leading to non-ideal behaviour cancel and the solution is virtually ideal. This temperature (the analogue of the Boyle

temperature for real gases) is called the **θ-temperature** (theta temperature). At this temperature, B is zero. As an example, for polystyrene in cyclohexane the θ-temperature is approximately 306 K, the exact value depending on the average molar mass of the polymer. A solution at its θ-temperature is called a **θ-solution**. Because a θ-solution behaves nearly ideally, its thermodynamic and structural properties are easier to describe even though the molar concentration is not low. In molecular terms, in a theta solution the molecules are in an unperturbed condition, whereas in other solutions expansion of the coiled molecule takes place as a result of interactions with the solvent.

Example 19.3 *Determining the size of a polymer by light scattering*

The following data for a sample of polystyrene in butanone were obtained at 20°C with plane-polarized light at $\lambda = 546$ nm.

θ/°	26.0	36.9	66.4	90.0	113.6
R_θ/m^2	19.7	18.8	17.1	16.0	14.4

In separate experiments, it was determined that $K = 6.42 \times 10^{-5}$ mol m^5 kg^{-2}. From this information, calculate R_g and $\bar{M}_w$ for the sample. Assume that B is negligibly small and that the polymer is small enough that eqn 19.9 holds.

Method Substituting the result of eqn 19.9 into eqn 19.8 we obtain, after some rearrangement:

$$\frac{1}{R_\theta} = \frac{1}{Kc_P\bar{M}_w} + \left(\frac{16\pi^2R_g^2}{3\lambda^2}\right)\left(\frac{1}{R_\theta}\sin^2\tfrac{1}{2}\theta\right)$$

Hence, a plot of $1/R_\theta$ against $(1/R_\theta)\sin^2\frac{1}{2}\theta$ should be a straight line with slope $16\pi^2R_g^2/3\lambda^2$ and y-intercept $1/Kc_P\bar{M}_w$.

Answer We construct a table of values of $1/R_\theta$ and $(1/R_\theta)\sin^2\frac{1}{2}\theta$ and plot the data (Fig. 19.5).

$10^2 \times R_\theta^{-1}$/m^{-2}	5.06	5.32	5.83	6.25	6.96
$(10^3/R_\theta)\sin^2(\frac{1}{2}\theta)$/m^{-2}	2.56	5.33	17.5	31.3	48.7

The best straight line through the data has a slope of 0.391 and a y-intercept of 5.06×10^{-2}. From these values and the value of K, we calculate $R_g = 4.71 \times 10^{-8}$ m = 47.1 nm and $\bar{M}_w = 987$ kg mol^{-1}.

A more accurate method for more concentrated samples consists of conducting a series of experiments where R_θ against θ data are obtained for several c_P values. From analysis of the entire data set, the R_g, $\bar{M}_w$, and B values are obtained.

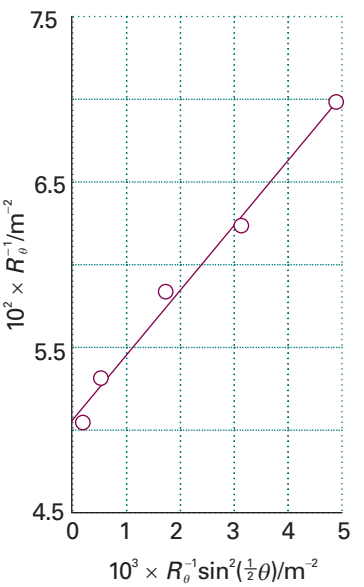


Fig. 19.5 Plot of the data for Example 19.3.

Self-test 19.3 The following data for an aqueous solution of a protein with $c_P = 2.0$ kg m^{-3} were obtained at 20°C with laser light at $\lambda = 532$ nm:

θ/°	15.0	45.0	70.0	85.0	90.0
R_θ/m^2	23.8	22.9	21.6	20.7	20.4

In a separate experiment, it was determined that $K = 2.40 \times 10^{-2}$ mol m^5 kg^{-2}. From this information, calculate the radius of gyration and the molar mass of the protein. Assume that B is negligibly small and that the protein is small enough that eqn 19.9 holds. [$R_g = 39.8$ nm; $M = 498$ kg mol^{-1}]

Synoptic table 19.2* Diffusion coefficients in water at 20°C

	$M/(\text{kg mol}^{-1})$	$D/(\text{m}^2\,\text{s}^{-1})$
Sucrose	0.342	4.59×10^{-10}
Lysozyme	14.1	1.04×10^{-10}
Haemoglobin	68	6.9×10^{-11}
Collagen	345	6.9×10^{-12}

* More values are given in the *Data section*.

Synoptic table 19.3* Frictional coefficients and molecular geometry†

a/b	Prolate	Oblate
2	1.04	1.04
3	1.18	1.17
6	1.31	1.28
8	1.43	1.37
10	1.54	1.46

* More values and analytical expressions are given in the *Data section*.
† Entries are the ratio f/f_0, where $f_0 = 6\pi\eta c$, where $c = (ab^2)^{1/3}$ for prolate ellipsoids and $c = (a^2b)^{1/3}$ for oblate ellipsoids; $2a$ is the major axis and $2b$ is the minor axis.

(c) Dynamic light scattering

A special laser scattering technique, **dynamic light scattering**, can be used to investigate the diffusion of polymers in solution. Consider two polymer molecules being irradiated by a laser beam. Suppose that at a time t the scattered waves from these particles interfere constructively at the detector, leading to a large signal. However, as the molecules move through the solution, the scattered waves may interfere destructively at another time t' and result in no signal. When this behaviour is extended to a very large number of molecules in solution, it results in fluctuations in light intensity that depend on the diffusion coefficient, D, which is a measure of the rate of molecular motion and is given by the **Stokes–Einstein relation** (which is discussed further in Section 21.9e):

$$D = \frac{kT}{f} \tag{19.11}$$

where f is the **frictional coefficient**, a measure of the forces that retard a molecule's motion. Table 19.2 lists some typical values of D. For a spherical particle of radius a in a solvent of viscosity η (see Section 19.6), the frictional coefficient is given by **Stokes's relation**:

$$f = 6\pi a\eta \tag{19.12}$$

If the molecule is not spherical, we use appropriate values of f given in Table 19.3. Hence, dynamic light scattering measurements give the diffusion coefficient and molecular size, in cases where the molecular shape is known. For dilute monodisperse systems of random coils, it has been found empirically that D is related to the molar mass M of the polymer by:

$$D = \beta_D M^{-0.6} \tag{19.13}$$

The coefficient β_D is obtained by determining D at fixed viscosity and temperature for a variety of standard samples with known molar masses. As we should expect, bulky polymers of high molar mass migrate more slowly (have a lower diffusion coefficient) through a solvent than polymers of low molar mass.

19.4 Ultracentrifugation

In a gravitational field, heavy particles settle towards the foot of a column of solution by the process called **sedimentation**. The rate of sedimentation depends on the strength of the field and on the masses and shapes of the particles. Spherical molecules (and compact molecules in general) sediment faster than rod-like and extended molecules. When the sample is at equilibrium, the particles are dispersed over a range of heights in accord with the Boltzmann distribution (because the gravitational field competes with the stirring effect of thermal motion). The spread of heights depends on the masses of the molecules, so the equilibrium distribution is another way to determine molar mass.

Sedimentation is normally very slow, but it can be accelerated by **ultracentrifugation**, a technique that replaces the gravitational field with a centrifugal field. The effect can be achieved in an ultracentrifuge, which is essentially a cylinder that can be rotated at high speed about its axis with a sample in a cell near its periphery (Fig. 19.6). Modern ultracentrifuges can produce accelerations equivalent to about 10^5 that of gravity ('10^5 g'). Initially the sample is uniform, but the 'top' (innermost) boundary of the solute moves outwards as sedimentation proceeds.

(a) The rate of sedimentation

A solute particle of mass m has an effective mass $m_{\text{eff}} = bm$ on account of the buoyancy of the medium, with

$$b = 1 - \rho v_s \tag{19.14}$$

where ρ is the solution density, v_s is the partial specific volume of the solute ($v_s = (\partial V/\partial m_B)_T$, with m_B the total mass of solute), and ρv_s is the mass of solvent displaced per gram of solute. The solute particles at a distance r from the axis of a rotor spinning at an angular velocity ω experience a centrifugal force of magnitude $m_{\text{eff}} r\omega^2$. The acceleration outwards is countered by a frictional force proportional to the speed, s, of the particles through the medium. This force is written fs, where f is the frictional coefficient (Section 19.3). The particles therefore adopt a **drift speed**, a constant speed through the medium, which is found by equating the two forces $m_{\text{eff}} r\omega^2$ and fs. The forces are equal when

$$s = \frac{m_{\text{eff}} r\omega^2}{f} = \frac{bmr\omega^2}{f} \tag{19.15}$$

The drift speed depends on the angular velocity and the radius, and it is convenient to define the **sedimentation constant**, S, as

$$S = \frac{s}{r\omega^2} \tag{19.16}$$

Then, because the average molecular mass is related to the average molar mass $\bar{M}_n$ through $m = \bar{M}_n/N_A$,

$$S = \frac{b\bar{M}_n}{fN_A} \tag{19.17}$$

On substituting the Stokes relation for spherical molecules (eqn 19.12), we obtain

$$S = \frac{b\bar{M}_n}{6\pi a\eta N_A} \tag{19.18}$$

and S may be used to determine either $\bar{M}_n$ or a. Again, if the molecules are not spherical, we use the appropriate value of f given in Table 19.3. As always when dealing with macromolecules, the measurements must be carried out at a series of concentrations and then extrapolated to zero concentration to avoid the complications that arise from the interference between bulky molecules.

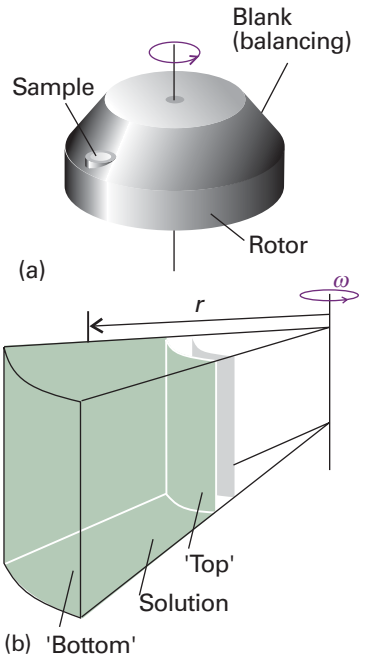


Fig. 19.6 (a) An ultracentrifuge head. The sample on one side is balanced by a blank diametrically opposite. (b) Detail of the sample cavity: the 'top' surface is the inner surface, and the centrifugal force causes sedimentation towards the outer surface; a particle at a radius r experiences a force of magnitude $mr\omega^2$.

Example 19.4 *Determining a sedimentation constant*

The sedimentation of the protein bovine serum albumin (BSA) was monitored at 25°C. The initial location of the solute surface was at 5.50 cm from the axis of rotation, and during centrifugation at 56 850 r.p.m. it receded as follows:

t/s	0	500	1000	2000	3000	4000	5000
r/cm	5.50	5.55	5.60	5.70	5.80	5.91	6.01

Calculate the sedimentation coefficient.

Method Equation 19.16 can be interpreted as a differential equation for $s = \mathrm{d}r/\mathrm{d}t$ in terms of r; so integrate it to obtain a formula for r in terms of t. The integrated expression, an expression for r as a function of t, will suggest how to plot the data and obtain from it the sedimentation constant.

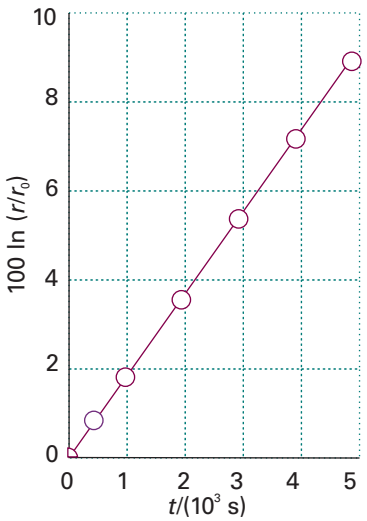


Fig. 19.7 A plot of the data in Example 19.4.

Answer Equation 19.16 may be written

$$\frac{dr}{dt} = r\omega^2 S$$

This equation integrates to

$$\ln\frac{r}{r_0} = \omega^2 St$$

It follows that a plot of $\ln(r/r_0)$ against t should be a straight line of slope $\omega^2 S$. Use $\omega = 2\pi\nu$, where ν is in cycles per second, and draw up the following table:

t/s	0	500	1000	2000	3000	4000	5000
$10^2 \ln(r/r_0)$	0	0.905	1.80	3.57	5.31	7.19	8.87

The straight-line graph (Fig. 19.7) has slope 1.78×10^{-5}; so $\omega^2 S = 1.79 \times 10^{-5}\ s^{-1}$. Because $\omega = 2\pi \times (56\,850/60)\ s^{-1} = 5.95 \times 10^3\ s^{-1}$, it follows that $S = 5.02 \times 10^{-13}$ s. The unit 10^{-13} s is sometimes called a 'svedberg' and denoted Sv; in this case $S = 5.02$ Sv.

Self-test 19.4 Calculate the sedimentation constant given the following data (the other conditions being the same as above):

t/s	0	500	1000	2000	3000	4000	5000
r/cm	5.65	5.68	5.71	5.77	5.84	5.9	5.97

[3.11 Sv]

At this stage, it appears that we need to know the molecular radius a to obtain the molar mass from the value of S. Fortunately, this requirement can be avoided by drawing on the Stokes–Einstein relation (eqn 19.11) between f and the diffusion coefficient, D. The average molar mass is then:

$$\bar{M} = \frac{SRT}{bD} \tag{19.19}$$

where we are not specifying which mean molar mass because the average obtained depends on technical details of the experiment. The result in eqn 19.19 is independent of the shape of the solute molecules. It follows that we can find the molar mass by combining measurements of S and D by ultracentrifugation and dynamic light scattering, respectively.

(b) Sedimentation equilibria

The difficulty with using sedimentation rates to measure molar masses lies in the inaccuracies inherent in the determination of diffusion coefficients of polydisperse systems. This problem can be avoided by allowing the system to reach equilibrium, for the transport property D is then no longer relevant. As we show in the *Justification* below, the weight-average molar mass can be obtained from the ratio of concentrations of the macromolecules at two different radii in a centrifuge operating at angular frequency ω:

$$\bar{M}_w = \frac{2RT}{(r_2^2 - r_1^2)b\omega^2} \ln\frac{c_2}{c_1} \tag{19.20}$$

An alternative treatment of the data leads to the *Z*–average molar mass. The centrifuge is run more slowly in this technique than in the sedimentation rate method to avoid having all the solute pressed in a thin film against the bottom of the cell. At these slower speeds, several days may be needed for equilibrium to be reached.

Justification 19.2 *The weight-average molar mass from sedimentation experiments*

The distribution of particles is the outcome of the balance between the effect of the centrifugal force and the dispersing effect of diffusion down a concentration gradient. The kinetic energy of a particle of effective mass m at a radius r in a rotor spinning at a frequency ω is $\frac{1}{2}m\omega^2r^2$, so the total chemical potential at a radius r is $\bar{\mu}(r) = \mu(r) - \frac{1}{2}m\omega^2r^2$, where $\mu(r)$ is the contribution that depends on the concentration of solute. The condition for equilibrium is that the chemical potential is uniform, so

$$\left(\frac{\partial\bar{\mu}}{\partial r}\right)_T = \left(\frac{\partial\mu}{\partial r}\right)_T - M\omega^2 r = 0$$

To evaluate the partial derivative of μ, we write

$$\left(\frac{\partial\mu}{\partial r}\right)_T = \left(\frac{\partial\mu}{\partial p}\right)_{T,c}\left(\frac{\partial p}{\partial r}\right)_{T,c} + \left(\frac{\partial\mu}{\partial c}\right)_{T,p}\left(\frac{\partial c}{\partial r}\right)_{T,p} = Mv\omega^2 r\rho + RT\left(\frac{\partial \ln c}{\partial r}\right)_{T,p}$$

The first result follows from the fact that $(\partial\mu/\partial p)_T = V_m$, the partial molar volume, and $V_m = Mv$. It also makes use of the fact that the hydrostatic pressure at r is $p(r) = p(r_0) + \frac{1}{2}\rho\omega^2(r^2 - r_0^2)$, where r_0 is the radius of the surface of the liquid in the sample holder (that is, the location of its meniscus), with ρ the mass density of the solution. The concentration term stems from the expression $\mu = \mu^\ominus + RT \ln c$. The condition for equilibrium is therefore

$$M r\omega^2(1 - v\rho) - RT\left(\frac{\partial \ln c}{\partial r}\right)_{T,p} = 0$$

and therefore, at constant temperature,

$$\mathrm{d} \ln c = \frac{M r\omega^2(1 - v\rho)\mathrm{d}r}{RT}$$

This expression integrates to eqn 19.20.

19.5 Electrophoresis

Many macromolecules, such as DNA, are charged and move in response to an electric field. This motion is called **electrophoresis**. Electrophoretic mobility is a result of a constant drift speed, s, reached by an ion when the driving force $ze\mathcal{E}$ (where, as usual, ze is the net charge and $\mathcal{E}$ is the field strength) is matched by the frictional force fs. The drift speed (which is treated in detail in Section 21.7) is then:

$$s = \frac{ze\mathcal{E}}{f} \tag{19.21}$$

Therefore, the mobility of a macromolecule in an electric field depends on its net charge, size (and hence molar mass), and shape. The latter two factors are implied by the dependence of s on f.

The drift speeds attained by polymers in traditional electrophoresis methods are rather low; as a result, several hours are often necessary to effect good separation of

complex mixtures. According to eqn 19.21, one way to increase the drift speed is to increase the electric field strength. However, there are limits to this strategy because very large electric fields can heat the large surfaces of an electrophoresis apparatus unevenly, leading to a non-uniform distribution of electrophoretic mobilities and poor separation.

In **capillary electrophoresis**, the sample is dispersed in a medium (such as methylcellulose) and held in a thin glass or plastic tube with diameters ranging from 20 to 100 μm. The small size of the apparatus makes it easy to dissipate heat when large electric fields are applied. Excellent separations may be effected in minutes rather than hours. Each polymer fraction emerging from the capillary can be characterized further by other techniques, such as MALDI-TOF.

IMPACT ON BIOCHEMISTRY

I19.1 Gel electrophoresis in genomics and proteomics

Advances in biotechnology are linked strongly to the development of physical techniques. The ongoing effort to characterize the entire genetic material, or **genome**, of organisms as simple as bacteria and as complex as *Homo sapiens* will lead to important new insights into the molecular mechanisms of disease, primarily through the discovery of previously unknown proteins encoded by the deoxyribonucleic acid (DNA) in genes. However, decoding genomic DNA will not always lead to accurate predictions of the amino acids present in biologically active proteins. Many proteins undergo chemical modification, such as cleavage into smaller proteins, after being synthesized in the cell. Moreover, it is known that one piece of DNA may encode more than one active protein. It follows that it is also important to describe the **proteome**, the full complement of functional proteins of an organism, by characterizing directly the proteins after they have been synthesized and processed in the cell.

The procedures of **genomics** and **proteomics**, the analysis of the genome and proteome, of complex organisms are time-consuming because of the very large number of molecules that must be characterized. For example, the human genome contains about 30 000 genes and the number of active proteins is likely to be much larger. Success in the characterization of the genome and proteome of any organism will depend on the deployment of very rapid techniques for the determination of the order in which molecular building blocks are linked covalently in DNA and proteins.

An important tool in genomics and proteomics is **gel electrophoresis**, in which biopolymers are separated on a slab of a porous gel, a semirigid dispersion of a solid in a liquid. Because the molecules must pass through the pores in the gel, the larger the macromolecule the less mobile it is in the electric field and, conversely, the smaller the macromolecule the more swiftly it moves through the pores. In this way, gel electrophoresis allows for the separation of components of a mixture according to their molar masses. Two common gel materials for the study of proteins and nucleic acids are agarose and cross-linked polyacrylamide. Agarose has large pores and is better suited for the study of large macromolecules, such as DNA and enzyme complexes. Polyacrylamide gels with varying pore sizes can be made by changing the concentration of acrylamide in the polymerization solution. In general, smaller pores form as the concentration of acrylamide is increased, making possible the separation of relatively small macromolecules by **polyacrylamide gel electrophoresis** (PAGE).

The separation of very large pieces of DNA, such as chromosomes, by conventional gel electrophoresis is not effective, making the analysis of genomic material rather difficult. Double-stranded DNA molecules are thin enough to pass through gel pores, but long and flexible DNA coils can become trapped in the pores and the result is impaired mobility along the direction of the applied electric field. This problem can be avoided with **pulsed-field electrophoresis**, in which a brief burst of the electric

field is applied first along one direction and then along a perpendicular direction. In response to the switching back and forth between field directions, the DNA coils writhe about and eventually pass through the gel pores. In this way, the mobility of the macromolecule can be related to its molar mass.

We have seen that charge also determines the drift speed. For example, proteins of the same size but different net charge travel along the slab at different speeds. One way to avoid this problem and to achieve separation by molar mass is to denature the proteins in a controlled way. Sodium dodecyl sulfate is an anionic detergent that is very useful in this respect: it denatures proteins, whatever their initial shapes, into rods by forming a complex with them. Moreover, most protein molecules bind a constant number of ions, so the net charge per protein is well regulated. Under these conditions, different proteins in a mixture may be separated according to size only. The molar mass of each constituent protein is estimated by comparing its mobility in its rod-like complex form with that of a standard sample of known molar mass. However, molar masses obtained by this method, often referred to as **SDS-PAGE** when polyacrylamide gels are used, are not as accurate as those obtained by MALDI-TOF or ultracentrifugation.

Another technique that deals with the effect of charge on drift speed takes advantage of the fact that the overall charge of proteins and other biopolymers depends on the pH of the medium. For instance, in acidic environments protons attach to basic groups and the net charge is positive; in basic media the net charge is negative as a result of proton loss. At the **isoelectric point**, the pH is such that there is no net charge on the biopolymer. Consequently, the drift speed of a biopolymer depends on the pH of the medium, with $s = 0$ at the isoelectric point (Fig. 19.8). **Isoelectric focusing** is an electrophoresis method that exploits the dependence of drift speed on pH. In this technique, a mixture of proteins is dispersed in a medium with a pH gradient along the direction of an applied electric field. Each protein in the mixture will stop moving at a position in the gradient where the pH is equal to the isoelectric point. In this manner, the protein mixture can be separated into its components.

The separation of complicated mixtures of macromolecules may be difficult by SDS-PAGE or isoelectric focusing alone. However, the two techniques can be combined in **two-dimensional (2D) electrophoresis**. In a typical experiment, a protein mixture is separated first by isoelectric focusing, yielding a pattern of bands in a gel slab such as the one shown in Fig. 19.9a. To improve the separation of closely spaced bands, the first slab is attached to a second slab and SDS-PAGE is performed with the electric field being applied in a direction that is perpendicular to the direction in which isoelectric focusing was performed. The macromolecules separate according to their molar masses along this second dimension of the experiment, and the result is that spots are spread widely over the surface of the slab, leading to enhanced separation of the mixture's components (Fig. 19.9b).

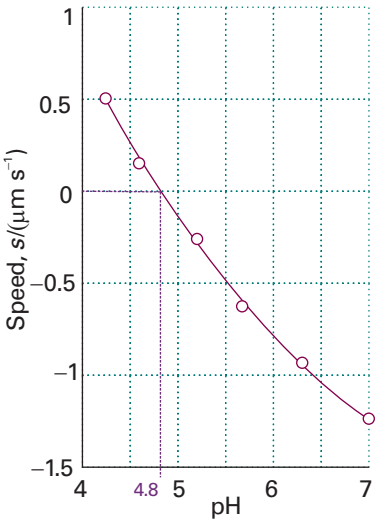


Fig. 19.8 The plot of drift speed of the protein bovine serum albumin in water against pH. The isoelectric point of the macromolecule corresponds to the pH at which the drift speed in the presence of an electric field is zero.

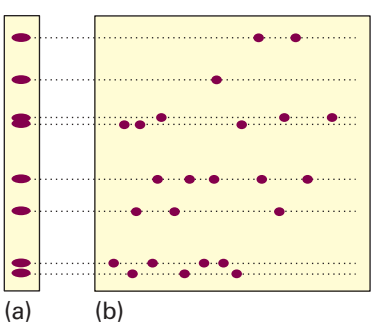


Fig. 19.9 The experimental steps taken during separation of a mixture of biopolymers by two-dimensional electrophoresis. (a) Isoelectric focusing is performed on a thin gel slab, resulting in separation along the vertical direction of the illustration. (b) The first slab is attached to a second, larger slab and SDS-PAGE is performed with the electric field oriented in the horizontal direction of the illustration, resulting in further separation by molar mass. The dashed horizontal lines show how the bands in the two-dimensional gel correspond to the bands in the gel on which isoelectric focusing was performed.

19.6 Viscosity

The formal definition of viscosity is given in Section 21.4; for now, we need to know that highly viscous liquids flow slowly and retard the motion of objects through them. The presence of a macromolecular solute increases the viscosity of a solution. The effect is large even at low concentration, because big molecules affect the fluid flow over an extensive region surrounding them. At low concentrations the viscosity, η, of the solution is related to the viscosity of the pure solvent, η_0, by

$$\eta = \eta_0(1 + [\eta]c + \cdots) \tag{19.22}$$

The **intrinsic viscosity**, $[\eta]$, is the analogue of a virial coefficient (and has dimensions of 1/concentration). It follows from eqn 19.22 that

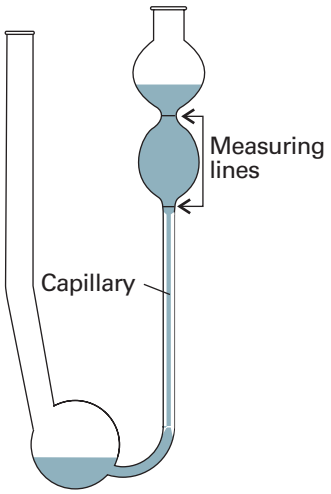


Fig. 19.10 An Ostwald viscometer. The viscosity is measured by noting the time required for the liquid to drain between the two marks.

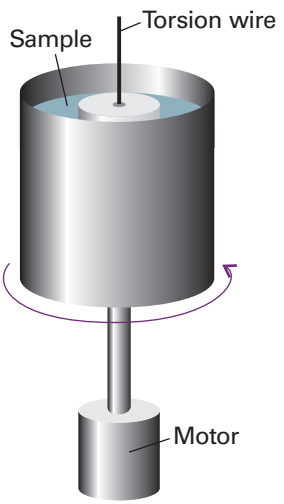


Fig. 19.11 A rotating rheometer. The torque on the inner drum is observed when the outer container is rotated.

Synoptic table 19.4* Intrinsic viscosity

	Solvent	θ/°C	K/(cm^3 g^{-1})	a
Polystyrene	Benzene	25	9.5×10^{-3}	0.74
Polyisobutylene	Benzene	23	8.3×10^{-2}	0.50
Various proteins	Guanidine hydrochloride + $HSCH_2CH_2OH$		7.2×10^{-3}	0.66

* More values are given in the *Data section*.

$$[\eta] = \lim_{c\to 0}\left(\frac{\eta - \eta_0}{c\eta_0}\right) = \lim_{c\to 0}\left(\frac{\eta/\eta_0 - 1}{c}\right) \tag{19.23}$$

Viscosities are measured in several ways. In the **Ostwald viscometer** shown in Fig. 19.10, the time taken for a solution to flow through the capillary is noted, and compared with a standard sample. The method is well suited to the determination of $[\eta]$ because the ratio of the viscosities of the solution and the pure solvent is proportional to the drainage time t and t_0 after correcting for different densities ρ and ρ_0:

$$\frac{\eta}{\eta_0} = \frac{t}{t_0} \times \frac{\rho}{\rho_0} \tag{19.24}$$

(In practice, the two densities are only rarely significantly different.) This ratio can be used directly in eqn 19.23. Viscometers in the form of rotating concentric cylinders are also used (Fig. 19.11), and the torque on the inner cylinder is monitored while the outer one is rotated. Such **rotating rheometers** (some instruments for the measurement of viscosity are also called rheometers) have the advantage over the Ostwald viscometer that the shear gradient between the cylinders is simpler than in the capillary and effects of the kind discussed shortly can be studied more easily.

There are many complications in the interpretation of viscosity measurements. Much of the work is based on empirical observations, and the determination of molar mass is usually based on comparisons with standard, nearly monodisperse sample. Some regularities are observed that help in the determination. For example, it is found that θ solutions of macromolecules often fit the **Mark–Kuhn–Houwink–Sakurada equation**:

$$[\eta] = K\bar{M}_v^a \tag{19.25}$$

where K and a are constants that depend on the solvent and type of macromolecule (Table 19.4); the viscosity-average molar mass, $\bar{M}_v$, appears in this expression.

Example 19.5 *Using intrinsic viscosity to measure molar mass*

The viscosities of a series of solutions of polystyrene in toluene were measured at 25°C with the following results:

c/(g dm^{-3})	0	2	4	6	8	10
η/(10^{-4} kg m^{-1} s^{-1})	5.58	6.15	6.74	7.35	7.98	8.64

Calculate the intrinsic viscosity and estimate the molar mass of the polymer by using eqn 19.25 with $K = 3.80 \times 10^{-5}$ dm^3 g^{-1} and $a = 0.63$.

Method The intrinsic viscosity is defined in eqn 19.23; therefore, form this ratio at the series of data points and extrapolate to $c = 0$. Interpret $\bar{M}_v$ as $\bar{M}_v/(\text{g mol}^{-1})$ in eqn 19.25.

Answer We draw up the following table:

$c/(\text{g dm}^{-3})$	0	2	4	6	8	10
η/η_0	1	1.102	1.208	1.317	1.43	1.549
$100[(\eta/\eta_0) - 1]/(c/\text{g dm}^{-3})$		5.11	5.2	5.28	5.38	5.49

The points are plotted in Fig. 19.12. The extrapolated intercept at $c = 0$ is 0.0504, so $[\eta] = 0.0504\ \text{dm}^3\ \text{g}^{-1}$. Therefore,

$$\bar{M}_v = \left(\frac{[\eta]}{K}\right)^{1/a} = 9.0 \times 10^4\ \text{g mol}^{-1}$$

Self-test 19.5 Evaluate the viscosity-average molar mass by using the second plotting technique. [90 kg mol^{-1}]

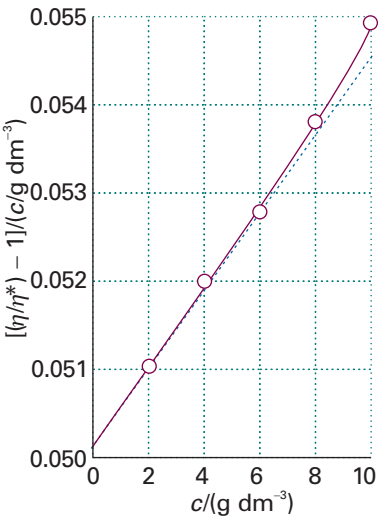


Fig. 19.12 The plot used for the determination of intrinsic viscosity, which is taken from the intercept at $c = 0$; see Example 19.5.

In some cases, the flow is non-Newtonian in the sense that the viscosity of the solution changes as the rate of flow increases. A decrease in viscosity with increasing rate of flow indicates the presence of long rod-like molecules that are orientated by the flow and hence slide past each other more freely. In some somewhat rare cases the stresses set up by the flow are so great that long molecules are broken up, with further consequences for the viscosity.

Structure and dynamics

The concept of the 'structure' of a macromolecule takes on different meanings at the different levels at which we think about the arrangement of the chain or network of monomers. The term **configuration** refers to the structural features that can be changed only by breaking chemical bonds and forming new ones. Thus, the chains —A—B—C— and —A—C—B— have different configurations. The term **conformation** refers to the spatial arrangement of the different parts of a chain, and one conformation can be changed into another by rotating one part of a chain around a bond.

19.7 The different levels of structure

The **primary structure** of a macromolecule is the sequence of small molecular residues making up the polymer. The residues may form either a chain, as in polyethylene, or a more complex network in which cross-links connect different chains, as in cross-linked polyacrylamide. In a synthetic polymer, virtually all the residues are identical and it is sufficient to name the monomer used in the synthesis. Thus, the repeating unit of polyethylene is —CH_2CH_2—, and the primary structure of the chain is specified by denoting it as —$(CH_2CH_2)_n$—.

The concept of primary structure ceases to be trivial in the case of synthetic copolymers and biological macromolecules, for in general these substances are chains formed from different molecules. For example, proteins are **polypeptides** formed from different amino acids (about twenty occur naturally) strung together by the **peptide link**, —CONH—. The determination of the primary structure is then a highly

Comment 19.1
More rigorously, the repeating unit in polyethylene is —CH_2— and the substance is polymethylene. However, the advantage of regarding the repeating unit as —CH_2CH_2— and naming it after its monomer is that derivatives, —$CHXCH_2$—, are seen to belong to the same family.

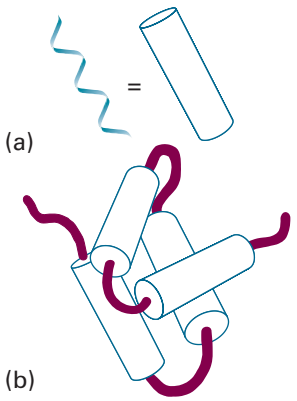


Fig. 19.13 (a) A polymer adopts a highly organized helical conformation, an example of a secondary structure. The helix is represented as a cylinder. (b) Several helical segments connected by short random coils pack together, providing an example of tertiary structure.

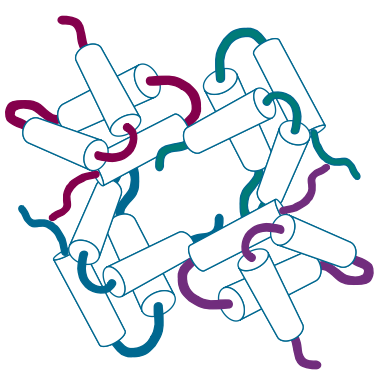


Fig. 19.14 Several subunits with specific tertiary structures pack together, providing an example of quaternary structure.

complex problem of chemical analysis called **sequencing**. The **degradation** of a polymer is a disruption of its primary structure, when the chain breaks into shorter components.

The **secondary structure** of a macromolecule is the (often local) spatial arrangement of a chain. The secondary structure of an isolated molecule of polyethylene is a random coil, whereas that of a protein is a highly organized arrangement determined largely by hydrogen bonds, and taking the form of random coils, helices (Fig. 19.13a), or sheets in various segments of the molecule. The loss of secondary structure is called **denaturation**. When the hydrogen bonds in a protein are destroyed (for instance, by heating, as when cooking an egg) the structure denatures into a random coil.

The **tertiary structure** is the overall three-dimensional structure of a macromolecule. For instance, the hypothetical protein shown in Fig. 19.13b has helical regions connected by short random-coil sections. The helices interact to form a compact tertiary structure.

The **quaternary structure** of a macromolecule is the manner in which large molecules are formed by the aggregation of others. Figure 19.14 shows how four molecular subunits, each with a specific tertiary structure, aggregate together. Quaternary structure can be very important in biology. For example, the oxygen-transport protein haemoglobin consists of four subunits that work together to take up and release O_2.

19.8 Random coils

The most likely conformation of a chain of identical units not capable of forming hydrogen bonds or any other type of specific bond is a **random coil**. Polyethylene is a simple example. The random coil model is a helpful starting point for estimating the orders of magnitude of the hydrodynamic properties of polymers and denatured proteins in solution.

The simplest model of a random coil is a **freely jointed chain**, in which any bond is free to make any angle with respect to the preceding one (Fig. 19.15). We assume that the residues occupy zero volume, so different parts of the chain can occupy the same region of space. The model is obviously an oversimplification because a bond is actually constrained to a cone of angles around a direction defined by its neighbour (Fig. 19.16). In a hypothetical one-dimensional freely jointed chain all the residues lie in a straight line, and the angle between neighbours is either 0° or 180°. The residues in a three-dimensional freely jointed chain are not restricted to lie in a line or a plane.

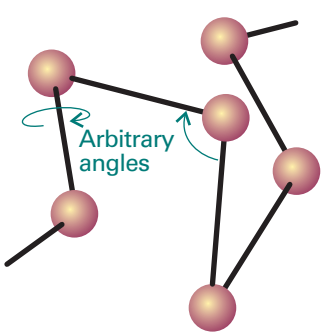


Fig. 19.15 A freely jointed chain is like a three-dimensional random walk, each step being in an arbitrary direction but of the same length.

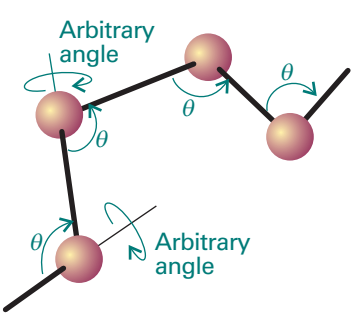


Fig. 19.16 A better description is obtained by fixing the bond angle (for example, at the tetrahedral angle) and allowing free rotation about a bond direction.

(a) Measures of size

As shown in the following *Justification*, we can deduce the probability, P, that the ends of a one-dimensional freely jointed chain composed of N units of length l are a distance nl apart:

$$P = \left(\frac{2}{\pi N}\right)^{1/2} \mathrm{e}^{-n^2/2N} \tag{19.26}$$

This function is plotted in Fig. 19.17 and can be used to calculate the probability that the ends of a three-dimensional freely jointed chain lie in the range r to $r + \mathrm{d}r$. We write this probability as $f(r)\mathrm{d}r$, where

$$f(r) = 4\pi\left(\frac{a}{\pi^{1/2}}\right)^3 r^2 \mathrm{e}^{-a^2r^2} \qquad a = \left(\frac{3}{2Nl^2}\right)^{1/2} \tag{19.27}$$

In some coils, the ends may be far apart whereas in others their separation is small. Here and elsewhere we are ignoring the fact that the chain cannot be longer than Nl. Although eqn 19.29 gives a nonzero probability for $r > Nl$, the values are so small that the errors in pretending that r can range up to infinity are negligible.

An alternative interpretation of eqn 19.27 is to regard each coil in a sample as ceaselessly writhing from one conformation to another; then $f(r)\mathrm{d}r$ is the probability that at any instant the chain will be found with the separation of its ends between r and $r + \mathrm{d}r$.

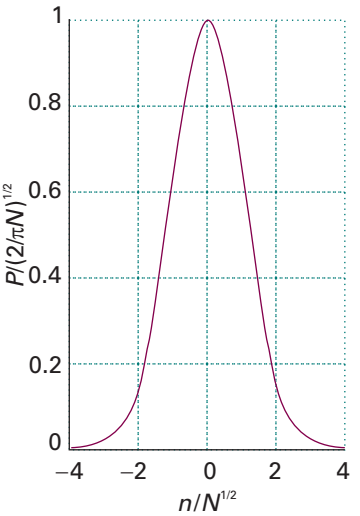


Fig. 19.17 The probability distribution for the separation of the ends of a one-dimensional random coil. The separation of the ends is nl, where l is the bond length.

Justification 19.3 *The one-dimensional freely jointed chain*

Consider a one-dimensional freely jointed polymer. We can specify the conformation of a molecule by stating the number of bonds pointing to the right (N_R) and the number pointing to the left (N_L). The distance between the ends of the chain is $(N_R - N_L)l$, where l is the length of an individual bond. We write $n = N_R - N_L$ and the total number of bonds as $N = N_R + N_L$.

The number of ways W of forming a chain with a given end-to-end distance nl is the number of ways of having N_R right-pointing and N_L left-pointing bonds. There are $N(N-1)(N-2)\ldots 1 = N!$ ways of selecting whether a step should be to the right or the left. If N_L steps are to the left, $N_R = N - N_L$ will be to the right. However, we end up at the same point for all $N_L!$ and $N_R!$ choices of which step is to the left and which to the right. Therefore

$$W = \frac{N!}{N_L!N_R!} = \frac{N!}{\{\frac{1}{2}(N+n)\}!\{\frac{1}{2}(N-n)\}!} \tag{19.28}$$

The probability that the separation is nl is

$$P = \frac{\text{number of polymers with } N_R \text{ bonds to the right}}{\text{total number of arrangements of bonds}}$$

$$= \frac{N!/N_R!(N-N_R)!}{2^N} = \frac{N!}{\{\frac{1}{2}(N+n)\}!\{\frac{1}{2}(N-n)\}!2^N}$$

When the chain is compact in the sense that $n \ll N$, it is more convenient to evaluate $\ln P$: the factorials are then large and we can use Stirling's approximation (Section 16.1a) in the form

$$\ln x! \approx \ln(2\pi)^{1/2} + (x + \tfrac{1}{2})\ln x - x$$

The result, after quite a lot of algebra, is

$$\ln P = \ln\left(\frac{2}{\pi N}\right)^{1/2} - \tfrac{1}{2}(N+n+1)\ln(1+\nu) - \tfrac{1}{2}(N-n+1)\ln(1-\nu) \tag{19.29}$$

where $\nu = n/N$. For a compact coil ($\nu \ll 1$) we use the approximation $\ln(1 \pm \nu) \approx \pm\nu - \frac{1}{2}\nu^2$ and so obtain

$$\ln P \approx \ln\left(\frac{2}{\pi N}\right)^{1/2} - \tfrac{1}{2}N\nu^2$$

which rearranges into eqn 19.26.

Self-test 19.6 Provide the algebraic steps that lead from eqn 19.28 to eqn 19.29.

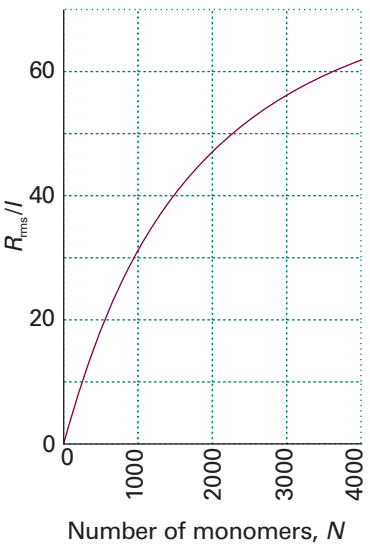


Fig. 19.18 The variation of the root mean square separation of the ends of a three-dimensional random coil, R_{rms}, with the number of monomers.

There are several measures of the geometrical size of a random coil. The **contour length**, R_c, is the length of the macromolecule measured along its backbone from atom to atom. For a polymer of N monomer units each of length l, the contour length is

$$R_c = Nl \tag{19.30}$$

The **root mean square separation**, R_{rms}, is a measure of the average separation of the ends of a random coil: it is the square root of the mean value of R^2. We show in the following *Justification* that:

$$R_{rms} = N^{1/2}l \tag{19.31}$$

We see that, as the number of monomer units increases, the root mean square separation of its end increases as $N^{1/2}$ (Fig. 19.18), and consequently its volume increases as $N^{3/2}$. The result must be multiplied by a factor when the chain is not freely jointed (see below).

Justification 19.4 *The root mean square separation of the ends of a freely jointed chain*

In *Appendix* 2 we see that the mean value $\langle X \rangle$ of a variable X with possible values x is given by

$$\langle X \rangle = \int_{-\infty}^{+\infty} x f(x)\mathrm{d}x$$

where the function $f(x)$ is the *probability density*, a measure of the distribution of the probability values over x, and dx is an infinitesimally small interval of x values. The mean value of a function $g(X)$ can be calculated with a similar formula:

$$\langle g(X) \rangle = \int_{-\infty}^{+\infty} g(x) f(x)\mathrm{d}x$$

To apply these concepts to the calculation of the root mean square separation of the ends of a random coil, we identify $f(r)\mathrm{d}r$ as the probability that the ends of the chain lie in the range $R = r$ to $R = r + \mathrm{d}r$. It follows that the general expression for the mean nth power of the end-to-end separation (a positive quantity that can vary from 0 to $+\infty$) is

$$\langle R^n \rangle = \int_0^{\infty} r^n f(r)\mathrm{d}r$$

To calculate R_{rms}, we first determine $\langle R^2 \rangle$ by using $n = 2$ and $f(r)$ from eqn 19.27:

$$\langle R^2 \rangle = 4\pi\left(\frac{a}{\pi^{1/2}}\right)^3 \int_0^{\infty} r^4 \mathrm{e}^{-a^2r^2}\mathrm{d}r = 4\pi\left(\frac{a}{\pi^{1/2}}\right)^3 \times \frac{3\pi^{1/2}}{8a^5} = \frac{3}{2a^2}$$

where we have used the standard integral

$$\int_0^\infty x^4 e^{-a^2x^2} dx = \frac{3}{2a^2}$$

When we use the expression for a in eqn 19.27 we obtain:

$$\langle R^2 \rangle = \frac{3}{2} \times \left(\frac{2Nl^2}{3} \right) = Nl^2$$

The root mean square separation follows from

$$R_{rms} = \langle R^2 \rangle^{1/2} = N^{1/2} l$$

Self-test 19.7 Calculate the mean separation of the ends of a freely jointed chain of N bonds of length l. *Hint*. You will need the standard integral $\int_0^\infty x^3 e^{-a^2x^2} dx = \frac{1}{2}a^4$.

$$[\langle R \rangle = \left(\frac{8N}{3\pi} \right)^{1/2} l]$$

Another convenient measure of size is the **radius of gyration**, R_g, which we encountered in Section 19.3a. It is calculated formally from the expression:

$$R_g = \frac{1}{N} \left(\tfrac{1}{2} \sum_{ij} R_{ij}^2 \right)^{1/2} \tag{19.32}$$

where R_{ij} is the separation of atoms i and j. The radius of gyration of the coil also increases as $N^{1/2}$:

$$R_g = \left(\frac{N}{6} \right)^{1/2} l \tag{19.33}$$

The radius of gyration may also be calculated for other geometries. For example, a solid uniform sphere of radius R has $R_g = (\frac{3}{5})^{1/2} R$, and a long thin uniform rod of length l has $R_g = l/(12)^{1/2}$ for rotation about an axis perpendicular to the long axis.

The random coil model ignores the role of the solvent: a poor solvent will tend to cause the coil to tighten so that solute–solvent contacts are minimized; a good solvent does the opposite. Therefore, calculations based on this model are better regarded as lower bounds to the dimensions for a polymer in a good solvent and as an upper bound for a polymer in a poor solvent. The model is most reliable for a polymer in a bulk solid sample, where the coil is likely to have its natural dimensions.

(b) Conformational entropy

The random coil is the least structured conformation of a polymer chain and corresponds to the state of greatest entropy. Any stretching of the coil introduces order and reduces the entropy. Conversely, the formation of a random coil from a more extended form is a spontaneous process (provided enthalpy contributions do not interfere). As shown in the *Justification* below, we can use the same model to deduce that the change in **conformational entropy**, the statistical entropy arising from the arrangement of bonds, when a coil containing N bonds of length l is stretched or compressed by nl is

$$\Delta S = -\tfrac{1}{2} kN \ln\{(1+\nu)^{1+\nu}(1-\nu)^{1-\nu}\} \qquad \nu = n/N \tag{19.34}$$

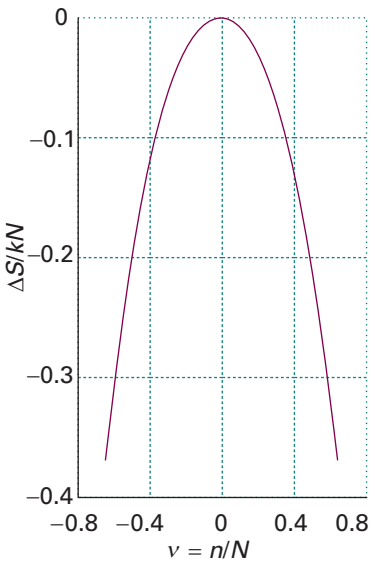


Fig. 19.19 The change in molar entropy of a perfect elastomer as its extension changes; $\nu = 1$ corresponds to complete extension; $\nu = 0$, the conformation of highest entropy, corresponds to the random coil.

This function is plotted in Fig. 19.19, and we see that minimum extension corresponds to maximum entropy.

Justification 19.5 *The conformational entropy of a freely jointed chain*

The conformational entropy of the chain is $S = k \ln W$, where W is given by eqn 19.28. Therefore,

$$S/k = \ln N! - \ln\{\tfrac{1}{2}(N+n)\}! - \ln\{\tfrac{1}{2}(N-n)\}!$$

Because the factorials are large (except for large extensions), we can use Stirling's approximation to obtain

$$S/k = -\ln(2\pi)^{1/2} + (N+1)\ln 2 + (N+\tfrac{1}{2})\ln N - \tfrac{1}{2}\ln\{(N+n)^{N+n+1}(N-n)^{N-n+1}\}$$

The most probable conformation of the chain is the one with the ends close together ($n = 0$), as may be confirmed by differentiation. Therefore, the maximum entropy is

$$S/k = -\ln(2\pi)^{1/2} + (N+1)\ln 2 + \tfrac{1}{2}\ln N$$

The change in entropy when the chain is stretched or compressed by nl is therefore the difference of these two quantities, and the resulting expression is eqn 19.34.

(c) Constrained chains

The freely jointed chain model is improved by removing the freedom of bond angles to take any value. For long chains, we can simply take groups of neighbouring bonds and consider the direction of their resultant. Although each successive individual bond is constrained to a single cone of angle θ relative to its neighbour, the resultant of several bonds lies in a random direction. By concentrating on such groups rather than individuals, it turns out that for long chains the expressions for the root mean square separation and the radius of gyration given above should be multiplied by

$$F = \left(\frac{1-\cos\theta}{1+\cos\theta}\right)^{1/2} \tag{19.35}$$

For tetrahedral bonds, for which $\cos\theta = -\frac{1}{3}$ (that is, $\theta = 109.5°$), $F = 2^{1/2}$. Therefore:

$$R_{rms} = (2N)^{1/2} l \qquad R_g = \left(\frac{N}{3}\right)^{1/2} l \tag{19.36}$$

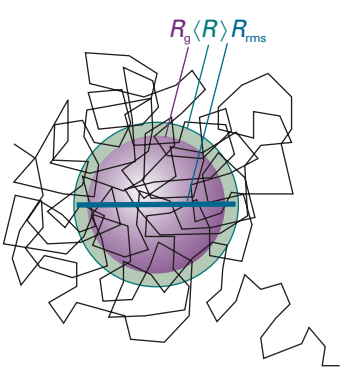


Fig. 19.20 A random coil in three dimensions. This one contains about 200 units. The root mean square distance between the ends (R_{rms}), the mean radius, and the radius of gyration (R_g) are indicated.

Illustration 19.2 *The dimensions of a polymer chain*

Consider a polyethylene chain with $M = 56$ kg mol^{-1}, corresponding to $N = 4000$. Because $l = 154$ pm for a C—C bond, we find $R_{rms} = 14$ nm and $R_g = 5.6$ nm (Fig. 19.20). This value of R_g means that, on average, the coils rotate like hollow spheres of radius 5.6 nm and mass equal to the molecular mass.

The model of a randomly coiled molecule is still an approximation, even after the bond angles have been restricted, because it does not take into account the impossibility of two or more atoms occupying the same place. Such self-avoidance tends to swell the coil, so (in the absence of solvent effects) it is better to regard R_{rms} and R_g as lower bounds to the actual values.

19.9 The structure and stability of synthetic polymers

Synthetic polymers are classified broadly as *elastomers*, *fibres*, and *plastics*, depending on their **crystallinity**, the degree of three-dimensional long-range order attained in the solid state. An **elastomer** is a flexible polymer that can expand or contract easily upon application of an external force. Elastomers are polymers with numerous cross-links that pull them back into their original shape when a stress is removed. A **perfect elastomer**, a polymer in which the internal energy is independent of the extension of the random coil, can be modelled as a freely jointed chain. We saw in Section 19.8b that the contraction of an extended chain to a random coil is spontaneous in the sense that it corresponds to an increase in entropy; the entropy change of the surroundings is zero because no energy is released when the coil forms. In the following *Justification* we also see that the restoring force, F, of a one-dimensional perfect elastomer is

$$F = \frac{kT}{2l}\ln\left(\frac{1+\nu}{1-\nu}\right) \qquad \nu = n/N \tag{19.37a}$$

where N is the total number of bonds of length l and the polymer is stretched or compressed by nl. This function is plotted in Fig. 19.21. At low extensions, when $\nu \ll 1$:

$$F \approx \frac{\nu kT}{l} = \frac{nkT}{Nl} \tag{19.37b}$$

and the sample obeys Hooke's law: the restoring force is proportional to the displacement (which is proportional to n). For small displacements, therefore, the whole coil shakes with simple harmonic motion.

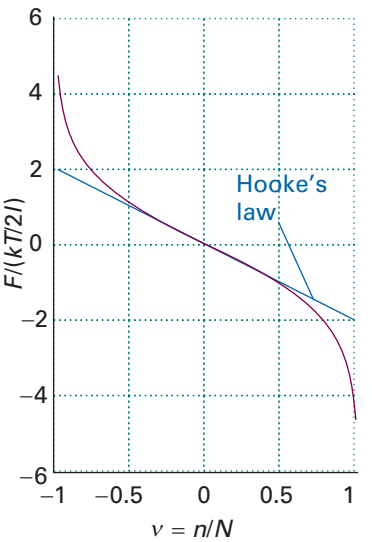


Fig. 19.21 The restoring force, F, of a one-dimensional perfect elastomer. For small strains, F is linearly proportional to the extension, corresponding to Hooke's law.

Justification 19.6 *Hooke's law*

The work done on an elastomer when it is extended through a distance dx is Fdx, where F is the restoring force. The change in internal energy is therefore

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + F\mathrm{d}x$$

It follows that

$$\left(\frac{\partial U}{\partial x}\right)_{T,V} = T\left(\frac{\partial S}{\partial x}\right)_{T,V} + F$$

In a perfect elastomer, as in a perfect gas, the internal energy is independent of the dimensions (at constant temperature), so $(\partial U/\partial x)_{T,V} = 0$. The restoring force is therefore

$$F = -T\left(\frac{\partial S}{\partial x}\right)_{T,V}$$

If now we substitute eqn 19.34 into this expression (we evade problems arising from the constraint of constant volume by supposing that the sample contracts laterally as it is stretched), we obtain

$$F = -\frac{T}{l}\left(\frac{\partial S}{\partial n}\right)_{T,V} = \frac{T}{Nl}\left(\frac{\partial S}{\partial \nu}\right)_{T,V} = \frac{kT}{2l}\ln\left(\frac{1+\nu}{1-\nu}\right)$$

as in eqn 19.37a.

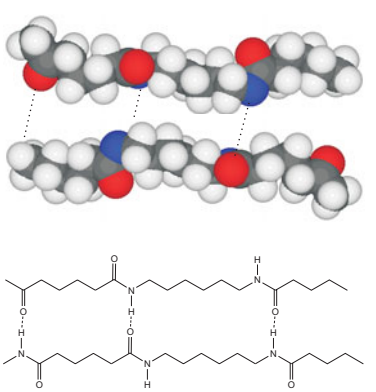


Fig. 19.22 A fragment of two nylon-66 polymer chains showing the pattern of hydrogen bonds that are responsible for the cohesion between the chains.

A **fibre** is a polymeric material that owes its strength to interactions between chains. One example is nylon-66 (Fig. 19.22). Under certain conditions, nylon-66 can be

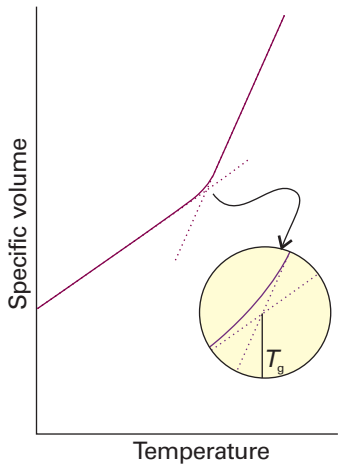


Fig. 19.23 The variation of specific volume with temperature of a synthetic polymer. The glass transition temperature, T_g, is at the point of intersection of extrapolations of the two linear parts of the curve.

prepared in a state of high crystallinity, in which hydrogen bonding between the amide links of neighbouring chains results in an ordered array.

A **plastic** is a polymer that can attain only a limited degree of crystallinity and as a result is neither as strong as a fibre nor as flexible as an elastomer. Certain materials, such as nylon-66, can be prepared either as a fibre or as a plastic. A sample of plastic nylon-66 may be visualized as consisting of crystalline hydrogen-bonded regions of varying size interspersed amongst amorphous, random coil regions. A single type of polymer may exhibit more than one characteristic, for to display fibrous character, the polymers need to be aligned; if the chains are not aligned, then the substance may be plastic. That is the case with nylon, poly(vinyl chloride), and the siloxanes.

The crystallinity of synthetic polymers can be destroyed by thermal motion at sufficiently high temperatures. This change in crystallinity may be thought of as a kind of intramolecular melting from a crystalline solid to a more fluid random coil. Polymer melting also occurs at a specific **melting temperature**, T_m, which increases with the strength and number of intermolecular interactions in the material. Thus, polyethylene, which has chains that interact only weakly in the solid, has $T_m = 414$ K and nylon-66 fibres, in which there are strong hydrogen bonds between chains, has $T_m = 530$ K. High melting temperatures are desirable in most practical applications involving fibres and plastics.

All synthetic polymers undergo a transition from a state of high to low chain mobility at the **glass transition temperature**, T_g. To visualize the glass transition, we consider what happens to an elastomer as we lower its temperature. There is sufficient energy available at normal temperatures for limited bond rotation to occur and the flexible chains writhe. At lower temperatures, the amplitudes of the writhing motion decrease until a specific temperature, T_g, is reached at which motion is frozen completely and the sample forms a glass. Glass transition temperatures well below 300 K are desirable in elastomers that are to be used at normal temperatures. Both the glass transition temperature and the melting temperature of a polymer may be measured by differential scanning calorimetry (*Impact* I2.1). Because the motion of the segments of a polymer chain increase at the glass transition temperature, T_g may also be determined from a plot of the specific volume of a polymer (the reciprocal of its mass density) against temperature (Fig. 19.23).

Comment 19.2

As we shall discuss further in Chapter 20, a metallic conductor is a substance with an electrical conductivity that decreases as the temperature is raised. A semiconductor is a substance with an electrical conductivity that increases as the temperature is raised.

IMPACT ON TECHNOLOGY

I19.2 Conducting polymers

We have just seen how the structure of a polymer chain affects its mechanical and thermal properties. Now we consider the electrical properties of synthetic polymers.

Most of the macromolecules and self-assembled structures considered in this chapter are insulators, or very poor electrical conductors. However, a variety of newly developed macromolecular materials have electrical conductivities that rival those of silicon-based semiconductors and even metallic conductors. We examine one example in detail: **conducting polymers**, in which extensively conjugated double bonds facilitate electron conduction along the polymer chain. The Nobel Prize in chemistry was awarded in 2000 to A.J. Heeger, A.J. McDiarmid, and H. Shirakawa for their pioneering work in the synthesis and characterization of conducting polymers.

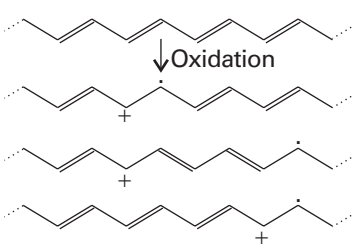


Fig. 19.24 The mechanism of migration of a partially localized cation radical, or polaron, in polyacetylene.

One example of a conducting polymer is polyacetylene (Fig. 19.24). Whereas the delocalized π bonds do suggest that electrons can move up and down the chain, the electrical conductivity of polyacetylene increases significantly when it is partially oxidized by I_2 and other strong oxidants. The product is a **polaron**, a partially localized cation radical that does not delocalize but rather travels through the chain, as shown in Fig. 19.24. Oxidation of the polymer by one more equivalent forms either **bipolarons**, a di-cation that moves as a unit through the chain, or **solitons**, two separate cation

radicals that move independently. Polarons and solitons contribute to the mechanism of charge conduction in polyacetylene.

Conducting polymers are slightly better electrical conductors than silicon semiconductors but are far worse than metallic conductors. They are currently used in a number of devices, such as electrodes in batteries, electrolytic capacitors, and sensors. Recent studies of photon emission by conducting polymers may lead to new technologies for light-emitting diodes and flat-panel displays. Conducting polymers also show promise as molecular wires that can be incorporated into nanometre-sized electronic devices.

19.10 The structure of proteins

A protein is a polypeptide composed of linked α-amino acids, $NH_2CHRCOOH$, where R is one of about 20 groups. For a protein to function correctly, it needs to have a well defined conformation. For example, an enzyme has its greatest catalytic efficiency only when it is in a specific conformation. The amino acid sequence of a protein contains the necessary information to create the active conformation of the protein from a newly synthesized random coil. However, the prediction of the conformation from the primary structure, the so-called *protein folding problem*, is extraordinarily difficult and is still the focus of much research.

(a) The Corey–Pauling rules

The origin of the secondary structures of proteins is found in the rules formulated by Linus Pauling and Robert Corey in 1951. The essential feature is the stabilization of structures by hydrogen bonds involving the peptide link. The latter can act both as a donor of the H atom (the NH part of the link) and as an acceptor (the CO part). The **Corey–Pauling rules** are as follows (Fig. 19.25):

1 The four atoms of the peptide link lie in a relatively rigid plane.

The planarity of the link is due to delocalization of π electrons over the O, C, and N atoms and the maintenance of maximum overlap of their p orbitals.

2 The N, H, and O atoms of a hydrogen bond lie in a straight line (with displacements of H tolerated up to not more than 30° from the N—O vector).

3 All NH and CO groups are engaged in hydrogen bonding.

The rules are satisfied by two structures. One, in which hydrogen bonding between peptide links leads to a helical structure, is a *helix*, which can be arranged as either a right- or a left-handed screw. The other, in which hydrogen bonding between peptide links leads to a planar structure, is a *sheet*; this form is the secondary structure of the protein fibroin, the constituent of silk.

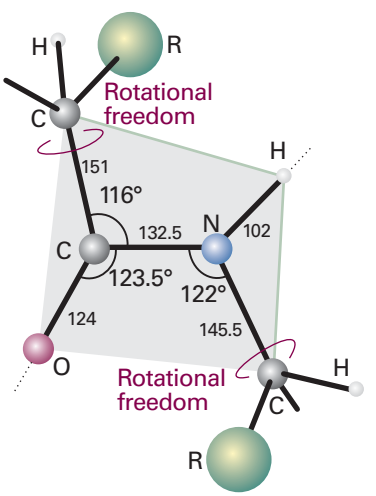


Fig. 19.25 The dimensions that characterize the peptide link (bonds in picometres). The C—NH—CO—C atoms define a plane (the C—N bond has partial double-bond character), but there is rotational freedom around the C—CO and N–C bonds.

(b) Conformational energy

A polypeptide chain adopts a conformation corresponding to a minimum Gibbs energy, which depends on the **conformational energy**, the energy of interaction between different parts of the chain, and the energy of interaction between the chain and surrounding solvent molecules. In the aqueous environment of biological cells, the outer surface of a protein molecule is covered by a mobile sheath of water molecules, and its interior contains pockets of water molecules. These water molecules play an important role in determining the conformation that the chain adopts through hydrophobic interactions and hydrogen bonding to amino acids in the chain.

The simplest calculations of the conformational energy of a polypeptide chain ignore entropy and solvent effects and concentrate on the total potential energy of all the

interactions between nonbonded atoms. For example, these calculations predict that a right-handed α helix of L-amino acids is marginally more stable than a left-handed helix of the same amino acids.

To calculate the energy of a conformation, we need to make use of many of the molecular interactions described in Chapter 18, and also of some additional interactions:

1 *Bond stretching*. Bonds are not rigid, and it may be advantageous for some bonds to stretch and others to be compressed slightly as parts of the chain press against one another. If we liken the bond to a spring, then the potential energy takes the form of Hooke's law

$$V_{\text{stretch}} = \tfrac{1}{2}k_{\text{stretch}}(R - R_e)^2 \tag{19.38}$$

where R_e is the equilibrium bond length and k_{stretch} is the force constant, a measure of the stiffness of the bond in question.

2 *Bond bending*. An O—C—H bond angle (or some other angle) may open out or close in slightly to enable the molecule as a whole to fit together better. If the equilibrium bond angle is θ_e, we write

$$V_{\text{bend}} = \tfrac{1}{2}k_{\text{bend}}(\theta - \theta_e)^2 \tag{19.39}$$

where k_{bend} is the force constant, a measure of how difficult it is to change the bond angle.

Self-test 19.8 Theoretical studies have estimated that the lumiflavin isoalloazine ring system has an energy minimum at the bending angle of 15°, but that it requires only 8.5 kJ mol^{-1} to increase the angle to 30°. If there are no other compensating interactions, what is the force constant for lumiflavin bending?

[6.27×10^{-23} J deg^{-2}, equivalent to 37.7 J mol^{-1} deg^{-2}]

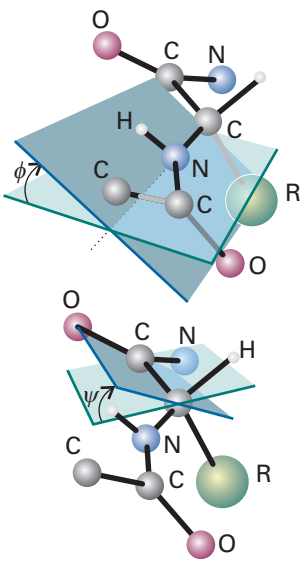


Fig. 19.26 The definition of the torsional angles ψ and ϕ between two peptide units. In this case (an α-L-polypeptide) the chain has been drawn in its all-*trans* form, with $\psi = \phi = 180°$.

3 *Bond torsion*. There is a barrier to internal rotation of one bond relative to another (just like the barrier to internal rotation in ethane). Because the planar peptide link is relatively rigid, the geometry of a polypeptide chain can be specified by the two angles that two neighbouring planar peptide links make to each other. Figure 19.26 shows the two angles ϕ and ψ commonly used to specify this relative orientation. The sign convention is that a positive angle means that the front atom must be rotated clockwise to bring it into an eclipsed position relative to the rear atom. For an all-*trans* form of the chain, all ϕ and ψ are 180°. A helix is obtained when all the ϕ are equal and when all the ψ are equal. For a right-handed helix, all $\phi = -57°$ and all $\psi = -47°$. For a left-handed helix, both angles are positive. The torsional contribution to the total potential energy is

$$V_{\text{torsion}} = A(1 + \cos 3\phi) + B(1 + \cos 3\psi) \tag{19.40}$$

in which A and B are constants of the order of 1 kJ mol^{-1}. Because only two angles are needed to specify the conformation of a helix, and they range from −180° to +180°, the torsional potential energy of the entire molecule can be represented on a **Ramachandran plot**, a contour diagram in which one axis represents ϕ and the other represents ψ.

4 *Interaction between partial charges*. If the partial charges q_i and q_j on the atoms i and j are known, a Coulombic contribution of the form $1/r$ can be included (Section 18.3):

$$V_{\text{Coulomb}} = \frac{q_i q_j}{4\pi\varepsilon r} \tag{19.41}$$

where ε is the permittivity of the medium in which the charges are embedded. Charges of $-0.28e$ and $+0.28e$ are assigned to N and H, respectively, and $-0.39e$ and $+0.39e$ to O and C, respectively. The interaction between partial charges does away with the need to take dipole–dipole interactions into account, for they are taken care of by dealing with each partial charge explicitly.

5 *Dispersive and repulsive interactions*. The interaction energy of two atoms separated by a distance r (which we know once ϕ and ψ are specified) can be given by the Lennard-Jones (12,6) form (Section 18.5):

$$V_{\mathrm{LJ}} = \frac{C}{r^{12}} - \frac{D}{r^{6}} \tag{19.42}$$

6 *Hydrogen bonding*. In some models of structure, the interaction between partial charges is judged to take into account the effect of hydrogen bonding. In other models, hydrogen bonding is added as another interaction of the form

$$V_{\mathrm{H\ bonding}} = \frac{E}{r^{12}} - \frac{F}{r^{10}} \tag{19.43}$$

The total potential energy of a given conformation (ϕ,ψ) can be calculated by summing the contributions given by eqns 19.38–19.43 for all bond angles (including torsional angles) and pairs of atoms in the molecule. The procedure is known as a **molecular mechanics** simulation and is automated in commercially available molecular modelling software. For large molecules, plots of potential energy against bond distance or bond angle often show several local minima and a global minimum (Fig. 19.27). The software packages include schemes for modifying the locations of the atoms and searching for these minima systematically.

The structure corresponding to the global minimum of a molecular mechanics simulation is a snapshot of the molecule at $T = 0$ because only the potential energy is included in the calculation; contributions to the total energy from kinetic energy are excluded. In a **molecular dynamics** simulation, the molecule is set in motion by heating it to a specified temperature, as described in Section 17.6b. The possible trajectories of all atoms under the influence of the intermolecular potentials correspond to the conformations that the molecule can sample at the temperature of the simulation. At very low temperatures, the molecule cannot overcome some of the potential energy barriers given by eqns 19.38–19.43, atomic motion is restricted, and only a few conformations are possible. At high temperatures, more potential energy barriers can be overcome and more conformations are possible. Therefore, molecular dynamics calculations are useful tools for the visualization of the flexibility of polymers.

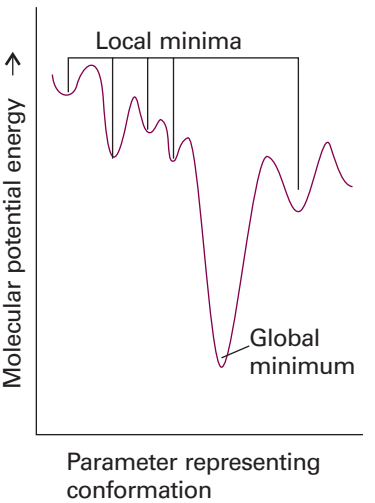


Fig. 19.27 For large molecules, a plot of potential energy against the molecular geometry often shows several local minima and a global minimum.

(c) Helices and sheets

A right-handed **α-helix** is illustrated in Fig. 19.28. Each turn of the helix contains 3.6 amino acid residues, so the period of the helix corresponds to 5 turns (18 residues). The pitch of a single turn (the distance between points separated by 360°) is 544 pm. The N—H···O bonds lie parallel to the axis and link every fourth group (so residue i is linked to residues $i - 4$ and $i + 4$). All the R groups point away from the major axis of the helix.

Figure 19.29 shows the Ramachandran plots for the helical form of polypeptide chains formed from the nonchiral amino acid glycine (R = H) and the chiral amino acid L-alanine ($R = CH_3$). The glycine map is symmetrical, with minima of equal depth at $\phi = -80°$, $\psi = +90°$ and at $\phi = +80°$, $\psi = -90°$. In contrast, the map for L-alanine is unsymmetrical, and there are three distinct low-energy conformations (marked I, II, III). The minima of regions I and II lie close to the angles typical of right- and

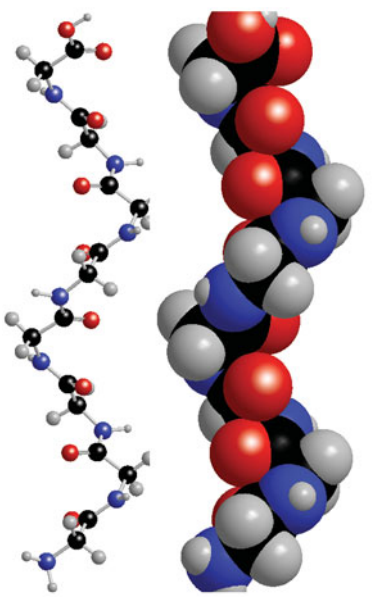


Fig. 19.28 The polypeptide α helix, with poly-L-glycine as an example. Carbon atoms are shown in green, with nitrogen in blue, oxygen in red, and hydrogen atoms in grey. There are 3.6 residues per turn, and a translation along the helix of 150 pm per residue, giving a pitch of 540 pm. The diameter (ignoring side chains) is about 600 pm.

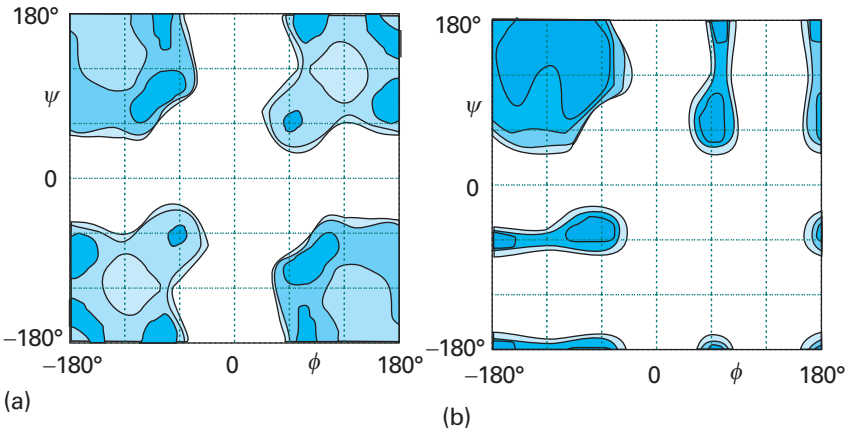


Fig. 19.29 Contour plots of potential energy against the torsional angles ψ and ϕ, also known as Ramachandran plots, for (a) a glycyl residue of a polypeptide chain and (b) an alanyl residue. The darker the shading is, the lower the potential energy. The glycyl diagram is symmetrical, but regions I and II in the correspond to right- and left-handed helices, are unsymmetrical, and the minimum in region I lies lower than that in region II. (After D.A. Brant and P.J. Flory, *J. Mol. Biol.* **23**, 47 (1967).)

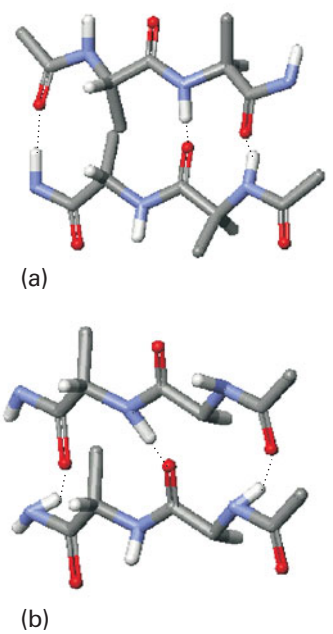


Fig. 19.30 The two types of β-sheets: (a) antiparallel ($\phi = -139°$, $\psi = 113°$), in which the N—H—O atoms of the hydrogen bonds form a straight-line; (b) parallel ($\phi = -119°$, $\psi = 113°$ in which the N—H···· O atoms of the hydrogen bonds are not perfectly aligned.

Comment 19.3

The web site contains links to sites where you may predict the secondary structure of a polypeptide by molecular mechanics simulations. There are also links to sites where you may visualize the structures of proteins and nucleic acids that have been obtained by experimental and theoretical methods.

left-handed helices, but the former has a lower minimum. This result is consistent with the observation that polypeptides of the naturally occurring L-amino acids tend to form right-handed helices.

A **β-sheet** (also called the **β-pleated sheet**) is formed by hydrogen bonding between two extended polypeptide chains (large absolute values of the torsion angles ϕ and ψ). Some of the R groups point above and some point below the sheet. Two types of structures can be distinguished from the pattern of hydrogen bonding between the constituent chains.

In an **anti-parallel β-sheet** (Fig. 19.30a), $\phi = -139°$, $\psi = 113°$, and the N—H····O atoms of the hydrogen bonds form a straight line. This arrangement is a consequence of the antiparallel arrangement of the chains: every N—H bond on one chain is aligned with a C—O bond from another chain. Antiparallel β-sheets are very common in proteins. In a **parallel β-sheet** (Fig. 19.30b), $\phi = -119°$, $\psi = 113°$, and the N—H····O atoms of the hydrogen bonds are not perfectly aligned. This arrangement is a result of the parallel arrangement of the chains: each N—H bond on one chain is aligned with a N—H bond of another chain and, as a result, each C—O bond of one chain is aligned with a C—O bond of another chain. These structures are not common in proteins.

Circular dichroism (CD) spectroscopy (Section 14.2) provides a great deal of information about the secondary structure of polypeptides. Consider a helical polypeptide. Not only are the individual monomer units chiral, but so is the helix. Therefore, we expect the α-helix to have a unique CD spectrum. Because β-sheets and random coils also have distinguishable spectral features (Fig. 19.31), circular dichroism is a very important technique for the study of protein conformation.

(d) Higher-order structures

Covalent and non-covalent interactions may cause polypeptide chains with well defined secondary structures to fold into tertiary structures. Subunits with well defined tertiary structures may interact further to form quaternary structures.

Although we do not know all the rules that govern protein folding, a few general conclusions may be drawn from X-ray diffraction studies of water-soluble natural proteins and synthetic polypeptides. In an aqueous environment, the chains fold in such a way as to place nonpolar R groups in the interior (which is often not very accessible to solvent) and charged R groups on the surface (in direct contact with the polar solvent). A wide variety of structures can result from these broad rules. Among them, a **four-helix bundle** (Fig. 19.32), which is found in proteins such as cytochrome b_{562} (an electron transport protein), forms when each helix has a nonpolar region along its length. The four nonpolar regions pack together to form a nonpolar interior.

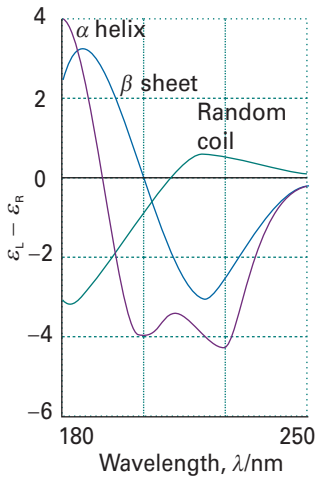


Fig. 19.31 Representative CD spectra of polypeptides. Random coils, α-helices, and β-sheets have different CD features in the spectral region where the peptide link absorbs.

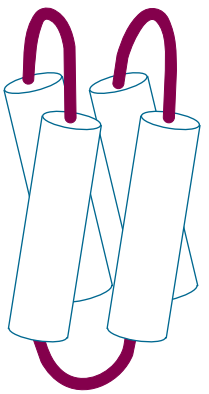
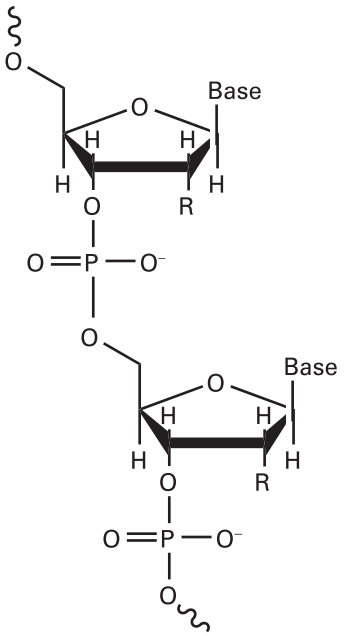


Fig. 19.32 A four-helix bundle forms from the interactions between nonpolar aminoacids on the surfaces of each helix, with the polar aminoacids exposed to the aqueous environment of the solvent.



2 D-ribose (R = OH) and 2'-deoxy-D-Ribose (R = H)

Similarly, interconnected β-sheets may interact to form a ***β*-barrel** (Fig. 19.33), the interior of which is populated by nonpolar R groups and which has an exterior rich in charged residues. The retinol-binding protein of blood plasma, which is responsible for transporting vitamin A, is an example of a β-barrel structure.

Factors that promote the folding of proteins include covalent disulfide (—S—S—) links, Coulombic interactions between ions (which depend on the degree of protonation of groups and therefore on the pH), hydrogen bonding, van der Waals interactions, and hydrophobic interactions (Section 18.4g). The clustering of nonpolar, hydrophobic, amino acids into the interior of a protein is driven primarily by hydrophobic interactions.

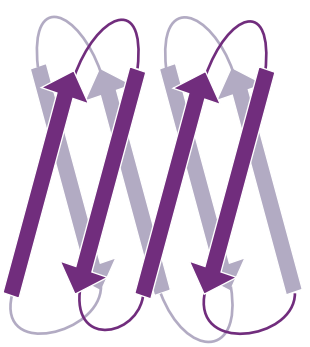
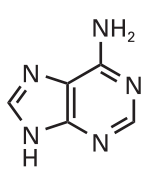


Fig. 19.33 Eight anti-parallel β-sheets, each represented by a purple arrow and linked by short random coils fold together as a β-barrel. Nonpolar aminoacids are in the interior of the barrel.

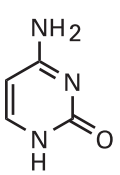
19.11 The structure of nucleic acids

Nucleic acids are key components of the mechanism of storage and transfer of genetic information in biological cells. Deoxyribonucleic acid (DNA) contains the instructions for protein synthesis, which is carried out by different forms of ribonucleic acid (RNA). In this section, we discuss the main structural features of DNA and RNA.

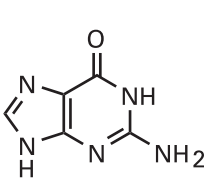
Both DNA and RNA are *polynucleotides* (**2**), in which base–sugar–phosphate units are linked by phosphodiester bonds. In RNA the sugar is β-D-ribose and in DNA it is β-D-2-deoxyribose (as shown in **2**). The most common bases are adenine (A, **3**), cytosine (C, **4**), guanine (G, **5**), thymine (T, found in DNA only, **6**), and uracil (U, found in RNA only, **7**). At physiological pH, each phosphate group of the chain carries a



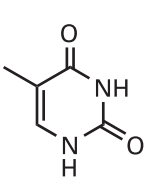
3 Adenine, A



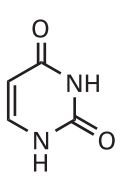
4 Cytosine, C



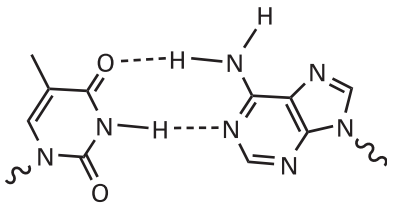
5 Guanine, G



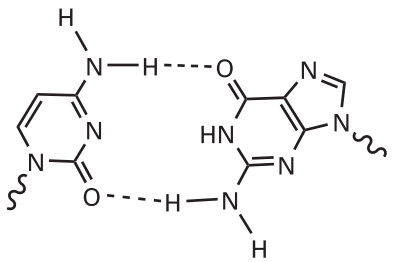
6 Thymine, T



7 Uracil, U



8 A–T base pair



9 C–G base pair

negative charge and the bases are deprotonated and neutral. This charge distribution leads to two important properties. One is that the polynucleotide chain is a **polyelectrolyte**, a macromolecule with many different charged sites, with a large and negative overall surface charge. The second is that the bases can interact by hydrogen bonding, as shown for A—T (**8**) and C—G base pairs (**9**). The secondary and tertiary structures of DNA and RNA arise primarily from the pattern of this hydrogen bonding between bases of one or more chains.

In DNA, two polynucleotide chains wind around each other to form a double helix (Fig. 19.34). The chains are held together by links involving A—T and C—G base pairs that lie parallel to each other and perpendicular to the major axis of the helix. The structure is stabilized further by interactions between the planar π systems of the bases. In B-DNA, the most common form of DNA found in biological cells, the helix is right-handed with a diameter of 2.0 nm and a pitch of 3.4 nm. Long stretches of DNA can fold further into a variety of tertiary structures. Two examples are shown in Fig. 19.35. Supercoiled DNA is found in the chromosome and can be visualized as the twisting of closed circular DNA (ccDNA), much like the twisting of a rubber band.

The extra —OH group in β-D-ribose imparts enough steric strain to a polynucleotide chain so that stable double helices cannot form in RNA. Therefore, RNA exists primarily as single chains that can fold into complex structures by formation of A—U and G—C base pairs. One example of this structural complexity is the structure of transfer RNA (tRNA), shown schematically in Fig. 19.36 in which base-paired regions are connected by loops and coils. Transfer RNAs help assemble polypeptide chains during protein synthesis in the cell.

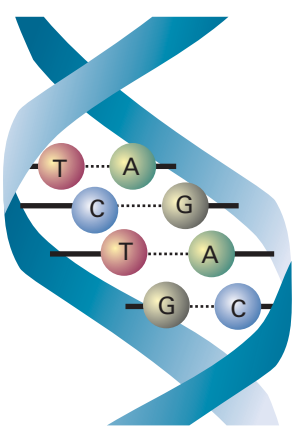


Fig. 19.34 DNA double helix, in which two polynucleotide chains are linked together by hydrogen bonds between adenine (A) and thymine (T) and between cytosine (C) and guanine (G).

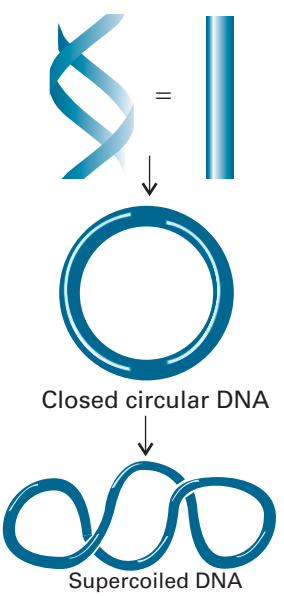


Fig. 19.35 A long section of DNA may form closed circular DNA (ccDNA) by covalent linkage of the two ends of the chain. Twisting of ccDNA leads to the formation of supercoiled DNA.

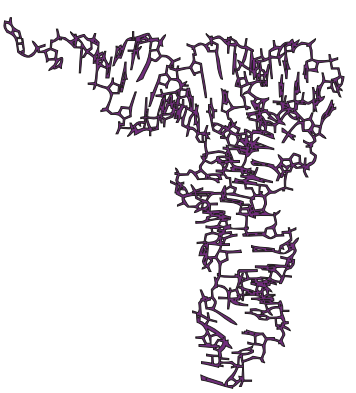


Fig. 19.36 Structure of a transfer RNA (tRNA).

19.12 The stability of proteins and nucleic acids

The loss of their natural conformation by proteins and nucleic acids is called **denaturation**. It can be achieved by changing the temperature or by adding chemical agents. Cooking is an example of thermal denaturation. When eggs are cooked, the protein albumin is denatured irreversibly, collapsing into a structure that resembles a random coil. One example of chemical denaturation is the 'permanent waving' of hair, which is a result of the reorganization of the protein keratin in hair. Disulfide cross-links between the chains of keratin render the protein and, hence hair fibres, inflexible. Chemical reduction of the disulfide bonds unravels keratin, and allows hair to be shaped. Oxidation re-forms the disulfide bonds and sets the new shape. The 'permanence' is only temporary, however, because the structure of newly formed hair is genetically controlled. Other means of chemically denaturing a protein include the addition of compounds that form stronger hydrogen bonds than those within a helix or sheet. One example is urea, which competes for the NH and CO groups of a polypeptide. The action of acids or bases, which can protonate or deprotonate groups involved in hydrogen bonding or change the Coulombic interactions that determine the conformation of a protein, can also result in denaturation.

Closer examination of thermal denaturation reveals some of the chemical factors that determine protein and nucleic acid stability. Thermal denaturation is similar to the melting of synthetic polymers (Section 19.9). Denaturation is a **cooperative process** in the sense that the biopolymer becomes increasingly more susceptible to denaturation once the process begins. This cooperativity is observed as a sharp step in a plot of fraction of unfolded polymer versus temperature (*Impact* I16.1). The melting temperature, T_m, is the temperature at which the fraction of unfolded polymer is 0.5 (Fig. 19.37).

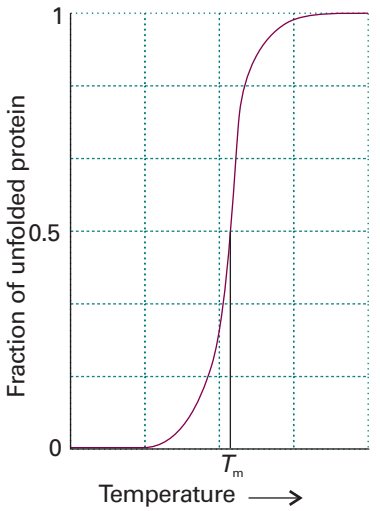


Fig. 19.37 A protein unfolds as the temperature of the sample increases. The sharp step in the plot of fraction of unfolded protein against temperature indicated that the transition is cooperative. The melting temperature, T_m, is the temperature at which the fraction of unfolded polymer is 0.5.

A DNA molecule is held together by hydrogen bonding interactions between bases of different chains and by **base-stacking**, in which dispersion interactions bring together the planar π systems of bases. Each G—C base pair has three hydrogen bonds whereas each A—T base pair has only two. Furthermore, experiments show that stacking interactions are stronger between C—G base pairs than between A—T base pairs. It follows that two factors render DNA sequences rich in C—G base pairs more stable than sequences rich in A—T base pairs: more hydrogen bonds between the bases and stronger stacking interactions between base pairs.

Proteins are relatively unstable towards chemical and thermal denaturation. For example, T_m = 320 K for ribonuclease T_1 (an enzyme that cleaves RNA in the cell), which is not far above the temperature at which the enzyme must operate (close to body temperature, 310 K). More surprisingly, the Gibbs energy for the unfolding of ribonuclease T_1 at pH 7.0 and 298 K is only 19.5 kJ mol^{-1}, which is comparable to the energy required to break a single hydrogen bond (about 20 kJ mol^{-1}). Therefore, unlike DNA, the stability of a protein does not increase in a simple way with the number of hydrogen bonding interactions. While the reasons for the low stability of proteins are not known, the answer probably lies in a delicate balance of all intra- and intermolecular interactions that allow a protein to fold into its active conformation, as discussed in Section 19.10.

Self-assembly

Much of the material discussed in this chapter also applies to aggregates of particles that form by **self-assembly**, the spontaneous formation of complex structures of molecules or macromolecules held together by molecular interactions, such as

Coulombic, dispersion, hydrogen bonding, or hydrophobic interactions. We have already encountered a few examples of self-assembly, such as the formation of liquid crystals (*Impact I6.1*), of protein quaternary structures from two or more polypeptide chains, and of a DNA double helix from two polynucleotide chains. Now we concentrate on the specific properties of additional self-assembled systems, including small aggregates that are at the heart of detergent action and extended sheets like those forming biological cell membranes.

19.13 Colloids

A **colloid**, or **disperse phase**, is a dispersion of small particles of one material in another. In this context, 'small' means something less than about 500 nm in diameter (about the wavelength of visible light). In general, colloidal particles are aggregates of numerous atoms or molecules, but are too small to be seen with an ordinary optical microscope. They pass through most filter papers, but can be detected by light-scattering and sedimentation.

(a) Classification and preparation

The name given to the colloid depends on the two phases involved. A **sol** is a dispersion of a solid in a liquid (such as clusters of gold atoms in water) or of a solid in a solid (such as ruby glass, which is a gold-in-glass sol, and achieves its colour by light scattering). An **aerosol** is a dispersion of a liquid in a gas (like fog and many sprays) or a solid in a gas (such as smoke): the particles are often large enough to be seen with a microscope. An **emulsion** is a dispersion of a liquid in a liquid (such as milk).

A further classification of colloids is as **lyophilic**, or solvent attracting, and **lyophobic**, solvent repelling. If the solvent is water, the terms **hydrophilic** and **hydrophobic**, respectively, are used instead. Lyophobic colloids include the metal sols. Lyophilic colloids generally have some chemical similarity to the solvent, such as —OH groups able to form hydrogen bonds. A **gel** is a semirigid mass of a lyophilic sol in which all the dispersion medium has penetrated into the sol particles.

The preparation of aerosols can be as simple as sneezing (which produces an imperfect aerosol). Laboratory and commercial methods make use of several techniques. Material (for example, quartz) may be ground in the presence of the dispersion medium. Passing a heavy electric current through a cell may lead to the sputtering (crumbling) of an electrode into colloidal particles. Arcing between electrodes immersed in the support medium also produces a colloid. Chemical precipitation sometimes results in a colloid. A precipitate (for example, silver iodide) already formed may be dispersed by the addition of a peptizing agent (for example, potassium iodide). Clays may be peptized by alkalis, the OH^- ion being the active agent.

Emulsions are normally prepared by shaking the two components together vigorously, although some kind of emulsifying agent usually has to be added to stabilize the product. This emulsifying agent may be a soap (the salt of a long-chain carboxylic acid) or other **surfactant** (surface active) species, or a lyophilic sol that forms a protective film around the dispersed phase. In milk, which is an emulsion of fats in water, the emulsifying agent is casein, a protein containing phosphate groups. It is clear from the formation of cream on the surface of milk that casein is not completely successful in stabilizing milk: the dispersed fats coalesce into oily droplets which float to the surface. This coagulation may be prevented by ensuring that the emulsion is dispersed very finely initially: intense agitation with ultrasonics brings this dispersion about, the product being 'homogenized' milk.

One way to form an aerosol is to tear apart a spray of liquid with a jet of gas. The dispersal is aided if a charge is applied to the liquid, for then electrostatic repulsions

help to blast it apart into droplets. This procedure may also be used to produce emulsions, for the charged liquid phase may be directed into another liquid.

Colloids are often purified by dialysis (*Impact I5.2*). The aim is to remove much (but not all, for reasons explained later) of the ionic material that may have accompanied their formation. A membrane (for example, cellulose) is selected that is permeable to solvent and ions, but not to the colloid particles. Dialysis is very slow, and is normally accelerated by applying an electric field and making use of the charges carried by many colloidal particles; the technique is then called **electrodialysis**.

(b) Structure and stability

Colloids are thermodynamically unstable with respect to the bulk. This instability can be expressed thermodynamically by noting that because the change in Gibbs energy, dG, when the surface area of the sample changes by $d\sigma$ at constant temperature and pressure is $dG = \gamma d\sigma$, where γ is the interfacial surface tension (Section 18.7a), it follows that $dG < 0$ if $d\sigma < 0$. The survival of colloids must therefore be a consequence of the kinetics of collapse: colloids are thermodynamically unstable but kinetically nonlabile.

At first sight, even the kinetic argument seems to fail: colloidal particles attract each other over large distances, so there is a long-range force that tends to condense them into a single blob. The reasoning behind this remark is as follows. The energy of attraction between two individual atoms i and j separated by a distance R_{ij}, one in each colloidal particle, varies with their separation as $1/R_{ij}^6$ (Section 18.4). The sum of all these pairwise interactions, however, decreases only as approximately $1/R^2$ (the precise variation depending on the shape of the particles and their closeness), where R is the separation of the centres of the particles. The sum has a much longer range than the $1/R^6$ dependence characteristic of individual particles and small molecules.

Several factors oppose the long-range dispersion attraction. For example, there may be a protective film at the surface of the colloid particles that stabilizes the interface and cannot be penetrated when two particles touch. Thus the surface atoms of a platinum sol in water react chemically and are turned into —$Pt(OH)_3H_3$, and this layer encases the particle like a shell. A fat can be emulsified by a soap because the long hydrocarbon tails penetrate the oil droplet but the carboxylate head groups (or other hydrophilic groups in synthetic detergents) surround the surface, form hydrogen bonds with water, and give rise to a shell of negative charge that repels a possible approach from another similarly charged particle.

(c) The electrical double layer

A major source of kinetic nonlability of colloids is the existence of an electric charge on the surfaces of the particles. On account of this charge, ions of opposite charge tend to cluster nearby, and an ionic atmosphere is formed, just as for ions (Section 5.9).

We need to distinguish two regions of charge. First, there is a fairly immobile layer of ions that adhere tightly to the surface of the colloidal particle, and which may include water molecules (if that is the support medium). The radius of the sphere that captures this rigid layer is called the **radius of shear** and is the major factor determining the mobility of the particles. The electric potential at the radius of shear relative to its value in the distant, bulk medium is called the **zeta potential**, ζ, or the **electrokinetic potential**. Second, the charged unit attracts an oppositely charged atmosphere of mobile ions. The inner shell of charge and the outer ionic atmosphere is called the **electrical double layer**.

The theory of the stability of lyophobic dispersions was developed by B. Derjaguin and L. Landau and independently by E. Verwey and J.T.G. Overbeek, and is known as

the **DLVO theory**. It assumes that there is a balance between the repulsive interaction between the charges of the electric double layers on neighbouring particles and the attractive interactions arising from van der Waals interactions between the molecules in the particles. The potential energy arising from the repulsion of double layers on particles of radius a has the form

$$V_{\text{repulsion}} = +\frac{Aa^2\zeta^2}{R}e^{-s/r_D} \tag{19.44}$$

where A is a constant, ζ is the **zeta potential**,[1] R is the separation of centres, s is the separation of the surfaces of the two particles ($s = R - 2a$ for spherical particles of radius a), and r_D is the thickness of the double layer. This expression is valid for small particles with a thick double layer ($a \ll r_D$). When the double layer is thin ($r_D \ll a$), the expression is replaced by

$$V_{\text{repulsion}} = \tfrac{1}{2}Aa\zeta^2 \ln(1 + e^{-s/r_D}) \tag{19.45}$$

In each case, the thickness of the double layer can be estimated from an expression like that derived for the thickness of the ionic atmosphere in the Debye–Hückel theory (eqn 5.80):

$$r_D = \left(\frac{\varepsilon RT}{2\rho F^2 I b^\ominus}\right)^{1/2} \tag{19.46}$$

where I is the ionic strength of the solution, ρ its mass density, and $b^\ominus = 1\ \text{mol kg}^{-1}$. The potential energy arising from the attractive interaction has the form

$$V_{\text{attraction}} = -\frac{B}{s} \tag{19.47}$$

where B is another constant. The variation of the total potential energy with separation is shown in Fig. 19.38.

At high ionic strengths, the ionic atmosphere is dense and the potential shows a secondary minimum at large separations. Aggregation of the particles arising from the stabilizing effect of this secondary minimum is called **flocculation**. The flocculated material can often be redispersed by agitation because the well is so shallow. **Coagulation**, the irreversible aggregation of distinct particles into large particles, occurs when the separation of the particles is so small that they enter the primary minimum of the potential energy curve and van der Waals forces are dominant.

The ionic strength is increased by the addition of ions, particularly those of high charge type, so such ions act as flocculating agents. This increase is the basis of the empirical **Schulze–Hardy rule**, that hydrophobic colloids are flocculated most efficiently by ions of opposite charge type and high charge number. The Al^{3+} ions in alum are very effective, and are used to induce the congealing of blood. When river water containing colloidal clay flows into the sea, the salt water induces flocculation and coagulation, and is a major cause of silting in estuaries. Metal oxide sols tend to be positively charged whereas sulfur and the noble metals tend to be negatively charged.

The primary role of the electric double layer is to confer kinetic non-lability. Colliding colloidal particles break through the double layer and coalesce only if the collision is sufficiently energetic to disrupt the layers of ions and solvating molecules, or if thermal motion has stirred away the surface accumulation of charge. This disruption may occur at high temperatures, which is one reason why sols precipitate

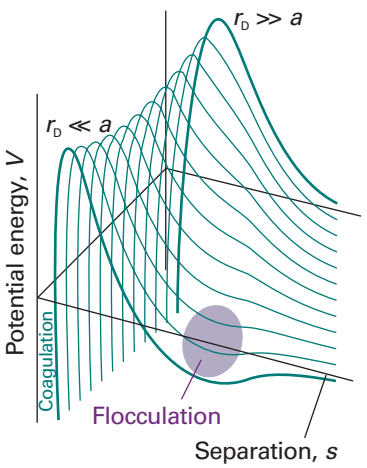


Fig. 19.38 The potential energy of interaction as a function of the separation of the centres of the two particles and its variation with the ratio of the particle size to the thickness a of the electric double layer r_D. The regions labelled coagulation and flocculation show the dips in the potential energy curves where these processes occur.

[1] The actual potential is that of the surface of the particles; there is some danger in identifying it with the zeta potential. See the references in *Further reading*.

when they are heated. The protective role of the double layer is the reason why it is important not to remove all the ions when a colloid is being purified by dialysis, and why proteins coagulate most readily at their isoelectric point.

19.14 Micelles and biological membranes

Surfactant molecules or ions can cluster together as **micelles**, which are colloid-sized clusters of molecules, for their hydrophobic tails tend to congregate, and their hydrophilic heads provide protection (Fig. 19.39).

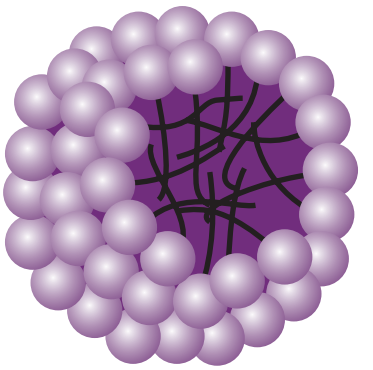


Fig. 19.39 A schematic version of a spherical micelle. The hydrophilic groups are represented by spheres and the hydrophobic hydrocarbon chains are represented by the stalks; these stalks are mobile.

(a) Micelle formation

Micelles form only above the **critical micelle concentration** (CMC) and above the **Krafft temperature**. The CMC is detected by noting a pronounced change in physical properties of the solution, particularly the molar conductivity (Fig. 19.40). There is no abrupt change in properties at the CMC; rather, there is a transition region corresponding to a range of concentrations around the CMC where physical properties vary smoothly but nonlinearly with the concentration. The hydrocarbon interior of a micelle is like a droplet of oil. Nuclear magnetic resonance shows that the hydrocarbon tails are mobile, but slightly more restricted than in the bulk. Micelles are important in industry and biology on account of their solubilizing function: matter can be transported by water after it has been dissolved in their hydrocarbon interiors. For this reason, micellar systems are used as detergents, for organic synthesis, froth flotation, and petroleum recovery.

Non-ionic surfactant molecules may cluster together in clumps of 1000 or more, but ionic species tend to be disrupted by the electrostatic repulsions between head groups and are normally limited to groups of less than about 100. The micelle population is often polydisperse, and the shapes of the individual micelles vary with concentration. Spherical micelles do occur, but micelles are more commonly flattened spheres close to the CMC.

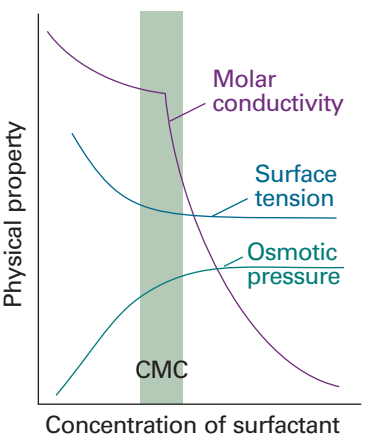


Fig. 19.40 The typical variation of some physical properties of an aqueous solution of sodium dodecylsulfate close to the critical micelle concentration (CMC).

Under certain experimental conditions, a **liposome** may form, with an inward pointing inner surface of molecules surrounded by an outward pointing outer layer (Fig. 19.41). Liposomes may be used to carry nonpolar drug molecules in blood. In concentrated solutions micelles formed from surfactant molecules may take the form of long cylinders and stack together in reasonably close-packed (hexagonal) arrays. These orderly arrangements of micelles are called **lyotropic mesomorphs** and, more colloquially, 'liquid crystalline phases'.

The enthalpy of micelle formation reflects the contributions of interactions between micelle chains within the micelles and between the polar head groups and the surrounding medium. Consequently, enthalpies of micelle formation display no readily discernible pattern and may be positive (endothermic) or negative (exothermic). Many non-ionic micelles form endothermically, with ΔH of the order of 10 kJ per mole of surfactant. That such micelles do form above the CMC indicates that the entropy change accompanying their formation must then be positive, and measurements suggest a value of about +140 J K^{-1} mol^{-1} at room temperature. The fact that the entropy change is positive even though the molecules are clustering together shows that hydrophobic interactions (Section 18.*) are important in the formation of micelles.

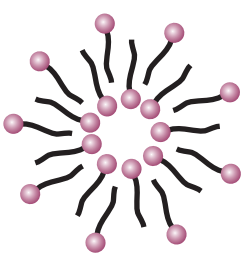
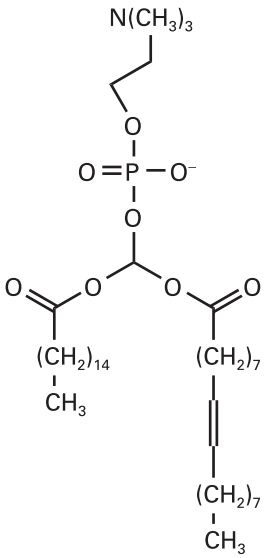


Fig. 19.41 The cross-sectional structure of a spherical liposome.

(b) Membrane formation

Some micelles at concentrations well above the CMC form extended parallel sheets, called **lamellar micelles**, two molecules thick. The individual molecules lie perpendicular to the sheets, with hydrophilic groups on the outside in aqueous solution and



10 Phosphatidyl choline

Comment 19.4

The web site contains links to databases of thermodynamic properties of lipids.

on the inside in nonpolar media. Such lamellar micelles show a close resemblance to biological membranes, and are often a useful model on which to base investigations of biological structures.

Although lamellar micelles are convenient models of cell membranes, actual membranes are highly sophisticated structures. The basic structural element of a membrane is a phospholipid, such as phosphatidyl choline (**10**), which contains long hydrocarbon chains (typically in the range C_{14}–C_{24}) and a variety of polar groups, such as $-CH_2CH_2N(CH_3)_3^+$ in (**10**). The hydrophobic chains stack together to form an extensive bilayer about 5 nm across. The lipid molecules form layers instead of micelles because the hydrocarbon chains are too bulky to allow packing into nearly spherical clusters.

The bilayer is a highly mobile structure, as shown by EPR studies with spin-labelled phospholipids (*Impact* I15.2). Not only are the hydrocarbon chains ceaselessly twisting and turning in the region between the polar groups, but the phospholipid and cholesterol molecules migrate over the surface. It is better to think of the membrane as a viscous fluid rather than a permanent structure, with a viscosity about 100 times that of water. In common with diffusional behaviour in general (see Section 21.*), the average distance a phospholipid molecule diffuses is proportional to the square-root of the time; more precisely, for a molecule confined to a two-dimensional plane, the average distance travelled in a time t is equal to $(4Dt)^{1/2}$. Typically, a phospholipid molecule migrates through about 1 μm (the diameter of a cell) in about 1 min.

All lipid bilayers undergo a transition from a state of high to low chain mobility at a temperature that depends on the structure of the lipid. To visualize the transition, we consider what happens to a membrane as we lower its temperature (Fig. 19.42). There is sufficient energy available at normal temperatures for limited bond rotation to occur and the flexible chains writhe. However, the membrane is still highly organized in the sense that the bilayer structure does not come apart and the system is best described as a liquid crystal (Fig. 19.42a). At lower temperatures, the amplitudes of the writhing motion decrease until a specific temperature is reached at which motion is largely frozen. The membrane is said to exist as a gel (Fig. 19.42b). Biological membranes exist as liquid crystals at physiological temperatures.

Phase transitions in membranes are often observed as 'melting' from gel to liquid crystal by differential scanning calorimetry (*Impact* I2.1). The data show relations between the structure of the lipid and the melting temperature. For example, the melting temperature increases with the length of the hydrophobic chain of the lipid. This correlation is reasonable, as we expect longer chains to be held together more strongly by hydrophobic interactions than shorter chains. It follows that stabilization of the gel phase in membranes of lipids with long chains results in relatively high melting temperatures. On the other hand, any structural elements that prevent alignment of the hydrophobic chains in the gel phase lead to low melting temperatures. Indeed, lipids containing unsaturated chains, those containing some C=C bonds, form membranes

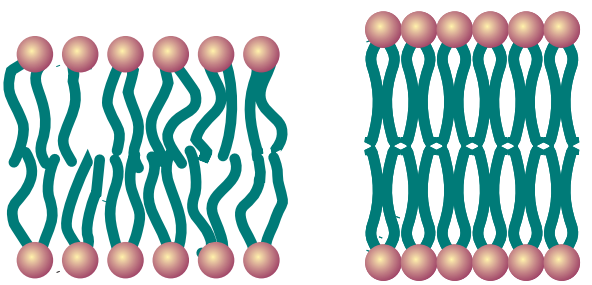
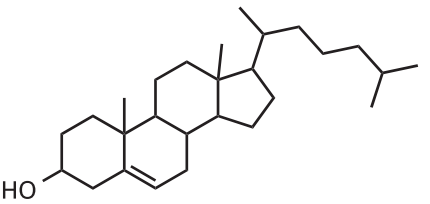


Fig. 19.42 A depiction of the variation with temperature of the flexibility of hydrocarbon chains in a lipid bilayer. (a) At physiological temperature, the bilayer exists as a liquid crystal, in which some order exists but the chains writhe. (b) At a specific temperature, the chains are largely frozen and the bilayer is said to exist as a gel.

with lower melting temperatures than those formed from lipids with fully saturated chains, those consisting of C—C bonds only.

Interspersed among the phospholipids of biological membranes are sterols, such as cholesterol (**11**), which is largely hydrophobic but does contain a hydrophilic —OH group. Sterols, which are present in different proportions in different types of cells, prevent the hydrophobic chains of lipids from 'freezing' into a gel and, by disrupting the packing of the chains, spread the melting point of the membrane over a range of temperatures.



11 Cholesterol

Self-test 19.9 Organisms are capable of biosynthesizing lipids of different composition so that cell membranes have melting temperatures close to the ambient temperature. Why do bacterial and plant cells grown at low temperatures synthesize more phospholipids with unsaturated chains than do cells grown at higher temperatures?

[Insertion of lipids with unsaturated chains lowers the plasma membrane's melting temperature to a value that is close to the lower ambient temperature.]

Peripheral proteins are proteins attached to the bilayer. **Integral proteins** are proteins immersed in the mobile but viscous bilayer. These proteins may span the depth of the bilayer and consist of tightly packed α helices or, in some cases, β sheets containing hydrophobic residues that sit comfortably within the hydrocarbon region of the bilayer. There are two views of the motion of integral proteins in the bilayer. In the **fluid mosaic model** shown in Fig. 19.43 the proteins are mobile, but their diffusion coefficients are much smaller than those of the lipids. In the **lipid raft model**, a number of lipid and cholesterol molecules form ordered structures, or 'rafts', that envelop proteins and help carry them to specific parts of the cell.

The mobility of the bilayer enables it to flow round a molecule close to the outer surface, to engulf it, and incorporate it into the cell by the process of *endocytosis*. Alternatively, material from the cell interior wrapped in cell membrane may coalesce with the cell membrane itself, which then withdraws and ejects the material in the process of *exocytosis*. The function of the proteins embedded in the bilayer, though, is to act as devices for transporting matter into and out of the cell in a more subtle manner. By providing hydrophilic channels through an otherwise alien hydrophobic environment, some proteins act as **ion channels** and **ion pumps** (*Impact* I21.2).

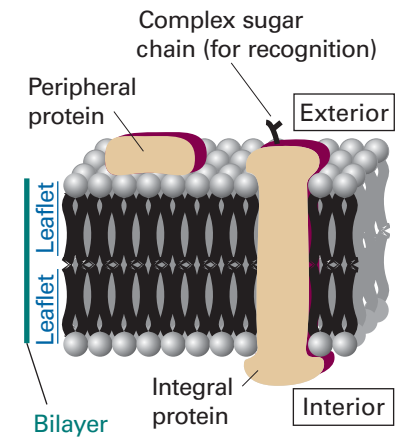


Fig. 19.43 In the fluid mosaic model of a biological cell membrane, integral proteins diffuse through the lipid bilayer. In the alternative lipid raft model, a number of lipid and cholesterol molecules envelop and transport the protein around the membrane.

19.15 Surface films

The compositions of surface layers have been investigated by the simple but technically elegant procedure of slicing thin layers off the surfaces of solutions and analysing their compositions. The physical properties of surface films have also been investigated. Surface films one molecule thick, such as that formed by a surfactant, are called **monolayers**. When a monolayer has been transferred to a solid support, it is called a **Langmuir–Blodgett film**, after Irving Langmuir and Katherine Blodgett, who developed experimental techniques for studying them.

(a) Surface pressure

The principal apparatus used for the study of surface monolayers is a **surface film balance** (Fig. 19.44). This device consists of a shallow trough and a barrier that can be moved along the surface of the liquid in the trough, and hence compress any monolayer on the surface. The **surface pressure**, π, the difference between the surface

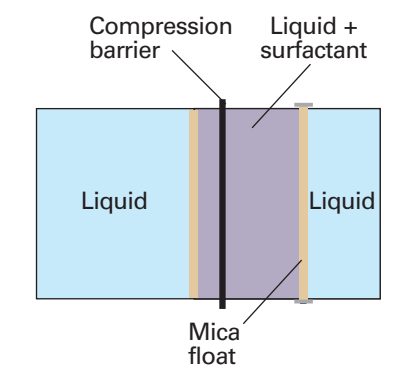


Fig. 19.44 A schematic diagram of the apparatus used to measure the surface pressure and other characteristics of a surface film. The surfactant is spread on the surface of the liquid in the trough, and then compressed horizontally by moving the compression barrier towards the mica float. The latter is connected to a torsion wire, so the difference in force on either side of the float can be monitored.

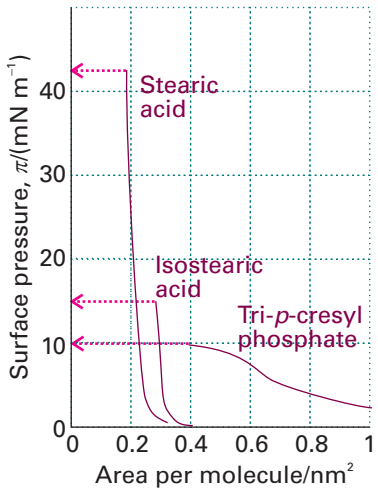
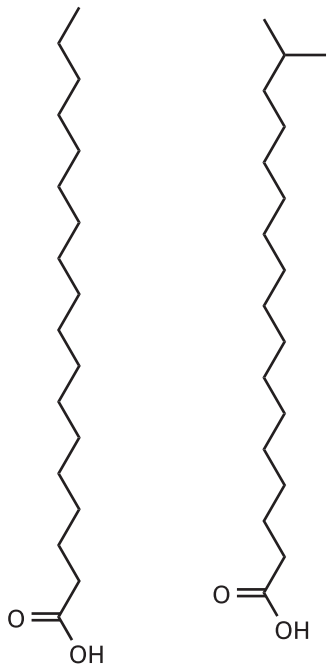
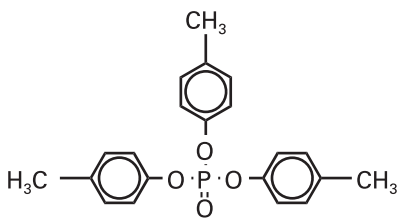


Fig. 19.45 The variation of surface pressure with the area occupied by each surfactant molecule. The collapse pressures are indicated by the horizontal dotted lines.



12 Stearic acid, $C_{17}H_{35}COOH$

13 Isostearic acid, $C_{17}H_{35}COOH$



14 Tri-*p*-cresylphosphate

tension of the pure solvent and the solution ($\pi = \gamma^* - \gamma$) is measured by using a torsion wire attached to a strip of mica that rests on the surface and pressing against one edge of the monolayer. The parts of the apparatus that are in touch with liquids are coated in polytetrafluoroethylene to eliminate effects arising from the liquid–solid interface. In an actual experiment, a small amount (about 0.01 mg) of the surfactant under investigation is dissolved in a volatile solvent and then poured on to the surface of the water; the compression barrier is then moved across the surface and the surface pressure exerted on the mica bar is monitored.

Some typical results are shown in Fig. 19.45. One parameter obtained from the isotherms is the area occupied by the molecules when the monolayer is closely packed. This quantity is obtained from the extrapolation of the steepest part of the isotherm to the horizontal axis. As can be seen from the illustration, even though stearic acid (**12**) and isostearic acid (**13**) are chemically very similar (they differ only in the location of a methyl group at the end of a long hydrocarbon chain), they occupy significantly different areas in the monolayer. Neither, though, occupies as much area as the tri-*p*-cresyl phosphate molecule (**14**), which is like a wide bush rather than a lanky tree.

The second feature to note from Fig. 19.45 is that the tri-*p*-cresyl phosphate isotherm is much less steep than the stearic acid isotherms. This difference indicates that the tri-*p*-cresyl phosphate film is more compressible than the stearic acid films, which is consistent with their different molecular structures.

A third feature of the isotherms is the **collapse pressure**, the highest surface pressure. When the monolayer is compressed beyond the point represented by the collapse pressure, the monolayer buckles and collapses into a film several molecules thick. As can be seen from the isotherms in Fig. 19.45, stearic acid has a high collapse pressure, but that of tri-*p*-cresyl phosphate is significantly smaller, indicating a much weaker film.

(b) The thermodynamics of surface layers

A surfactant is active at the interface between two phases, such as at the interface between hydrophilic and hydrophobic phases. A surfactant accumulates at the interface, and modifies its surface tension and hence the surface pressure. To establish the relation between the concentration of surfactant at a surface and the change in surface tension it brings about, we consider two phases α and β in contact and suppose that the system consists of several components J, each one present in an overall amount n_J. If the components were distributed uniformly through the two phases right up to the interface, which is taken to be a plane of surface area σ, the total Gibbs energy, G, would be the sum of the Gibbs energies of both phases, $G = G(\alpha) + G(\beta)$. However, the components are not uniformly distributed because one may accumulate at the interface. As a result, the sum of the two Gibbs energies differs from G by an amount called the **surface Gibbs energy**, $G(\sigma)$:

$$G(\sigma) = G - \{G(\alpha) + G(\beta)\} \tag{19.48}$$

Similarly, if it is supposed that the concentration of a species J is uniform right up to the interface, then from its volume we would conclude that it contains an amount $n_J(\alpha)$ of J in phase α and an amount $n_J(\beta)$ in phase β. However, because a species may accumulate at the interface, the total amount of J differs from the sum of these two amounts by $n_J(\sigma) = n_J - \{n_J(\alpha) + n_J(\beta)\}$. This difference is expressed in terms of the **surface excess**, Γ_J:

$$\Gamma_J = \frac{n_J(\sigma)}{\sigma} \tag{19.49}$$

The surface excess may be either positive (an accumulation of J at the interface) or negative (a deficiency there).

The relation between the change in surface tension and the composition of a surface (as expressed by the surface excess) was derived by Gibbs. In the following *Justification* we derive the **Gibbs isotherm**, between the changes in the chemical potentials of the substances present in the interface and the change in surface tension:

$$\mathrm{d}\gamma = -\sum_{\mathrm{J}} \Gamma_{\mathrm{J}} \mathrm{d}\mu_{\mathrm{J}} \tag{19.50}$$

Justification 19.7 *The Gibbs isotherm*

A general change in G is brought about by changes in T, p, and the n_{J}:

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p + \gamma\mathrm{d}\sigma + \sum_{\mathrm{J}} \mu_{\mathrm{J}} \mathrm{d}n_{\mathrm{J}}$$

When this relation is applied to G, $G(\alpha)$, and $G(\beta)$ we find

$$\mathrm{d}G(\sigma) = -S(\sigma)\mathrm{d}T + \gamma\mathrm{d}\sigma + \sum_{\mathrm{J}} \mu_{\mathrm{J}} \mathrm{d}n_{\mathrm{J}}(\sigma)$$

because at equilibrium the chemical potential of each component is the same in every phase, $\mu_{\mathrm{J}}(\alpha) = \mu_{\mathrm{J}}(\beta) = \mu_{\mathrm{J}}(\sigma)$. Just as in the discussion of partial molar quantities (Section 5.1), the last equation integrates at constant temperature to

$$G(\sigma) = \gamma\sigma + \sum_{\mathrm{J}} \mu_{\mathrm{J}} n_{\mathrm{J}}(\sigma)$$

We are seeking a connection between the change of surface tension $\mathrm{d}\gamma$ and the change of composition at the interface. Therefore, we use the argument that in Section 5.1 led to the Gibbs–Duhem equation (eqn 5.12b), but this time we compare the expression

$$\mathrm{d}G(\sigma) = \gamma\mathrm{d}\sigma + \sum_{\mathrm{J}} \mu_{\mathrm{J}} \mathrm{d}n_{\mathrm{J}}(\sigma)$$

(which is valid at constant temperature) with the expression for the same quantity but derived from the preceding equation:

$$\mathrm{d}G(\sigma) = \gamma\mathrm{d}\sigma + \sigma\mathrm{d}\gamma + \sum_{\mathrm{J}} \mu_{\mathrm{J}} \mathrm{d}n_{\mathrm{J}}(\sigma) + \sum_{\mathrm{J}} n_{\mathrm{J}}(\sigma)\mathrm{d}\mu_{\mathrm{J}}$$

The comparison implies that, at constant temperature,

$$\sigma\mathrm{d}\gamma + \sum_{\mathrm{J}} n_{\mathrm{J}}(\sigma)\mathrm{d}\mu_{\mathrm{J}} = 0$$

Division by σ then gives eqn 19.50.

Now consider a simplified model of the interface in which the 'oil' and 'water' phases are separated by a geometrically flat surface. This approximation implies that only the surfactant, S, accumulates at the surface, and hence that Γ_{oil} and Γ_{water} are both zero. Then the Gibbs isotherm equation becomes

$$\mathrm{d}\gamma = -\Gamma_{\mathrm{S}} \mathrm{d}\mu_{\mathrm{S}} \tag{19.51}$$

For dilute solutions,

$$\mathrm{d}\mu_{\mathrm{S}} = RT \ln c \tag{19.52}$$

where c is the molar concentration of the surfactant. It follows that

$$\mathrm{d}\gamma = RT\Gamma_{\mathrm{S}} \frac{\mathrm{d}c}{c}$$

at constant temperature, or

$$\left(\frac{\partial\gamma}{\partial c}\right)_T = -\frac{RT\Gamma_{\mathrm{S}}}{c} \tag{19.53}$$

If the surfactant accumulates at the interface, its surface excess is positive and eqn 19.53 implies that $(\partial\gamma/\partial c)_T < 0$. That is, the surface tension decreases when a solute accumulates at a surface. Conversely, if the concentration dependence of γ is known, then the surface excess may be predicted and used to infer the area occupied by each surfactant molecule on the surface.

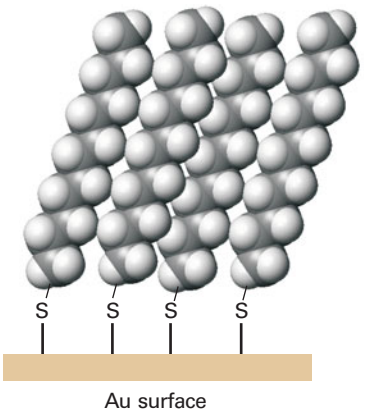


Fig. 19.46 Self-assembled monolayers of alkylthiols formed on to a gold surface by reaction of the thiol groups with the surface and aggregation of the alkyl chains.

IMPACT ON NANOSCIENCE

I19.3 Nanofabrication with self-assembled monolayers

Nanofabrication is the synthesis of *nanodevices*, nanometre-sized assemblies of atoms and molecules that can be used in nanotechnological applications, such as those discussed in *Impact I9.2*. Here we see how molecular self-assembly can be used as the basis for nanofabrication on surfaces . Of current interest are **self-assembled monolayers** (SAMs), ordered molecular aggregates that form a monolayer of material on a surface. To understand the formation of SAMs, consider exposing molecules such as alkyl thiols RSH, where R represents an alkyl chain, to an Au(0) surface. The thiols react with the surface, forming $RS^-Au(I)$ adducts:

$$RSH + Au(0)_n \rightarrow RS^-Au(I) \cdot Au(0)_{n-1} + \tfrac{1}{2}H_2$$

If R is a sufficiently long chain, van der Waals interactions between the adsorbed RS units lead to the formation of a highly ordered monolayer on the surface, as shown in Fig. 19.46. It is observed that the Gibbs energy of formation of SAMs increases with the length of the alkyl chain, with each methylene group contributing 400–4000 J mol^{-1} to the overall Gibbs energy of formation.

The atomic force microscope (*Impact* I9.1) can be used to manipulate SAMs into specific shapes on a surface by digging the microscope tip through the alkyl chains, bringing it into contact with the surface and then moving SAMs around the surface. In one application of the technique, enzymes were bound to patterned SAMs. The experiment shows that it is possible to form nanometre-sized reactors that take advantage of the catalytic properties of enzymes (which we explore in Chapter 23).

Checklist of key ideas

- ☐ 1. A polymer is a compound formed by linking together small molecules. Many proteins (specifically protein enzymes) are monodisperse; a synthetic polymer is polydisperse.
- ☐ 2. The number-average molar mass, $\bar{M}_n$, is the value obtained by weighting each molar mass by the number of molecules of that mass present in the sample; the weight-average molar mass, $\bar{M}_w$, is the average calculated by weighting the molar masses of the molecules by the mass of each one present in the sample; the Z-average molar mass, $\bar{M}_Z$, is the average molar mass obtained from sedimentation measurements.
- ☐ 3. The heterogeneity index of a polymer sample is $\bar{M}_w/\bar{M}_n$.
- ☐ 4. Techniques for the determination of the mean molar masses of macromolecules include mass spectrometry (as MALDI), ultracentrifugation, laser light scattering, and viscometry.
- ☐ 5. The least structured model of a macromolecule is as a random coil; for a freely jointed random coil of contour length Nl, the root mean square separation is $N^{1/2}l$ and the radius of gyration is $R_g = (N/6)^{1/2}l$.
- ☐ 6. The conformational entropy is the statistical entropy arising from the arrangement of bonds in a random coil.
- ☐ 7. The primary structure of a biopolymer is the sequence of its monomer units.
- ☐ 8. The secondary structure of a protein is the spatial arrangement of the polypeptide chain and includes the α-helix and β-sheet.
- ☐ 9. Helical and sheet-like polypeptide chains are folded into a tertiary structure by bonding influences between the residues of the chain.
- ☐ 10. Some macromolecules have a quaternary structure as aggregates of two or more polypeptide chains.
- ☐ 11. Synthetic polymers are classified as elastomers, fibres, and plastics.
- ☐ 12. A perfect elastomer is a polymer in which the internal energy is independent of the extension of the random coil; for small extensions a random coil model obeys a Hooke's law restoring force.
- ☐ 13. Synthetic polymers undergo a transition from a state of high to low chain mobility at the glass transition temperature, T_g.
- ☐ 14. The melting temperature of a polymer is the temperature at which three-dimensional order is lost.
- ☐ 15. A mesophase is a bulk phase that is intermediate in character between a solid and a liquid.
- ☐ 16. A disperse system is a dispersion of small particles of one material in another.
- ☐ 17. Colloids are classified as lyophilic (solvent attracting, specifically hydrophilic for water) and lyophobic (solvent repelling, specifically hydrophobic).
- ☐ 18. A surfactant is a species that accumulates at the interface of two phases or substances and modifies the properties of the surface.

- ☐ 19. The radius of shear is the radius of the sphere that captures the rigid layer of charge attached to a colloid particle.
- ☐ 20. The zeta potential is the electric potential at the radius of shear relative to its value in the distant, bulk medium.
- ☐ 21. The inner shell of charge and the outer atmosphere jointly constitute the electric double layer.
- ☐ 22. Many colloid particles are thermodynamically unstable but kinetically non-labile.
- ☐ 23. Flocculation is the reversible aggregation of colloidal particles; coagulation is the irreversible aggregation of colloidal particles.
- ☐ 24. The Schultze–Hardy rule states that hydrophobic colloids are flocculated most efficiently by ions of opposite charge type and high charge number.
- ☐ 25. A micelle is a colloid-sized cluster of molecules that forms at the critical micelle concentration (CMC) and at the Krafft temperature.
- ☐ 26. A liposome is a vesicle with an inward pointing inner surface of molecules surrounded by an outward pointing outer layer.
- ☐ 27. A lamellar micelle is an extended layer of molecules two molecules thick.
- ☐ 28. A monolayer, a single layer of molecules on a surface. A Langmuir–Blodgett film is a monolayer that has been transferred to a solid support.
- ☐ 29. Surface pressure is the difference between the surface tension of the pure solvent and the solution: $\pi = \gamma^* - \gamma$.
- ☐ 30. The collapse pressure is the highest surface pressure sustained by a surface film.

Further reading

Articles and texts

A.W. Adamson and A.P. Gast, *Physical chemistry of surfaces*. Wiley, New York (1997).

J. Barker, *Mass spectrometry*. Wiley, New York (1999).

C.E. Carraher, Jr., *Seymour/Carraher's polymer chemistry*, Marcel Dekker (2000).

D.F. Evans and H. Wennerström, *The colloidal domain : where physics, chemistry, biology, and technology meet*. Wiley-VCH, New York (1999).

P. Flory, *Principles of polymer chemistry*. Cornell University Press, Ithaca (1953).

C.S. Johnson and D.A. Gabriel, *Laser light scattering*. Dover, New York (1995).

A.R. Leach, *Molecular modelling: principles and applications*. Longman, Harlow (1997).

K.E. van Holde, W.C. Johnson, and P.S. Ho, *Principles of physical biochemistry*. Prentice Hall, Upper Saddle River (1998).

P.C. Heimenz and R. Rajagopalan, *Principles of colloid and surface chemistry*. Marcel Dekker, New York (1997).

Sources of data and information

D.R. Lide (ed.), *CRC Handbook of Chemistry and Physics*, Sections 7 and 13. CRC Press, Boca Raton (2000).

Further information

Further information 19.1 *The Rayleigh ratio*

Here we outline the key steps in the derivation of eqn 19.8 for the Rayleigh ratio. You are encouraged to consult sources in the *Further reading* section for additional details.

The ratio of the intensity of the light scattered by a sample to the intensity of the incident light is (see *Further reading*):

$$\frac{I}{I_0} = \frac{\pi^2\alpha^2}{\varepsilon_r^2\lambda^4}\frac{\sin^2\phi}{r^2} \tag{19.54}$$

where r is the distance between the sample and the detector and ϕ is the angle of observation relative to the z-axis ($\phi = 90°$ in Fig. 19.3). It follows from eqn 19.7 that

$$R_\theta = \frac{\pi^2\alpha^2}{\varepsilon_r^2\lambda^4}$$

The relation between the polarizability and the refractive index, n_r, of a solution is (see *Appendix* 3 for a qualitative explanation and *Further reading* for quantitative details)

$$n_r^2 - n_{r,0}^2 = \frac{\mathcal{N}\alpha}{\varepsilon_r} \tag{19.55}$$

where $n_{r,0}$ is the refractive index of the solvent and $\mathcal{N}$ is the number density of polymer molecules. Because $\mathcal{N} = c_P N_A/M$ (where c_P is the mass concentration of the polymer and M is its molar mass), we have:

$$\alpha = \frac{\varepsilon_r M}{c_P N_A}(n_r^2 - n_{r,0}^2)$$

For a dilute solution, n_r differs little from $n_{r,0}$ and we can write:

$$n_r = n_{r,0} + \left(\frac{dn_r}{dc_P}\right)c_P + \cdots$$

It follows that

$$n_r^2 \approx n_{r,0}^2 + 2n_{r,0}\left(\frac{dn_r}{dc_P}\right)c_P$$

and therefore that

$$\alpha = \frac{2\varepsilon_r n_{r,0} M}{N_A}\left(\frac{dn_r}{dc_P}\right) \qquad (19.56)$$

The Rayleigh ratio for scattering by a single molecule now becomes

$$R_\theta = \frac{4\pi^2 n_{r,0}^2 M^2}{\lambda^4 N_A^2}\left(\frac{dn_r}{dc_P}\right)^2 \qquad (19.57)$$

For a sample of N molecules, we multiply the expression above by $N = c_P N_A V/M$ and, after substituting $\bar{M}_w$ for M, obtain

$$R_\theta = Kc_P\bar{M}_w \qquad K = \frac{4\pi^2 n_{r,0}^2 V(dn_r/dc_P)^2}{\lambda^4 N_A} \qquad (19.58)$$

Equation 19.8 follows after we multiply the right-hand side of this equation by the structure factor P_θ.

Now we derive expressions for the structure factor. If the molecule is regarded as composed of a number of atoms i at distances R_i from a convenient point, interference occurs between the radiation scattered by each pair. The scattering from all the particles is then calculated by allowing for contributions from all possible orientations of each pair of atoms in each molecule. If there are N atoms in the macromolecule, and if all are assumed to have the same scattering power, then it is possible to show that (see *Further reading*)

$$P_\theta = \frac{1}{N^2}\sum_{i,j}\frac{\sin sR_{ij}}{sR_{ij}}, \qquad s = \frac{4\pi}{\lambda}\sin\tfrac{1}{2}\theta \qquad (19.59)$$

where R_{ij} is the separation of atoms i and j, and λ is the wavelength of the incident radiation.

When the molecule is much smaller than the wavelength of the incident radiation in the sense that $sR_{ij} \ll 1$ (for example, if $R = 5$ nm, and $\lambda = 500$ nm, all the sR_{ij} are about 0.1), we can use the expansion $\sin x = x - \frac{1}{6}x^3 + \cdots$ to write

$$\sin sR_{ij} = sR_{ij} - \tfrac{1}{6}(sR_{ij})^3 + \cdots$$

and then

$$P_\theta = \frac{1}{N^2}\sum_{i,j}\left\{1 - \tfrac{1}{6}(sR_{ij})^2 + \cdots\right\} = 1 - \frac{s^2}{6N^2}\sum_{i,j}R_{ij}^2 + \cdots$$

The sum over the squares of the separations gives the radius of gyration of the molecule (through eqn 19.32). Therefore

$$P_\theta \approx 1 - \tfrac{1}{3}s^2R_g^2 = 1 - \frac{16\pi^2 R_g^2 \sin^2\frac{1}{2}\theta}{3\lambda^2}$$

which is eqn 19.9.

Discussion questions

19.1 Distinguish between number-average, weight-average, and Z-average molar masses. Discuss experimental techniques that can measure each of these properties.

19.2 Suggest reasons why the techniques described in the preceding question produce different mass averages.

19.3 Distinguish between contour length, root mean square separation, and radius of gyration of a random coil.

19.4 Identify the terms in and limit the generality of the following expressions: (a) $\Delta S = -\frac{1}{2}kN\ln\{(1+\nu)^{1+\nu}(1-\nu)^{1-\nu}\}$, (b) $R_{rms} = (2N)^{1/2}l$, and (c) $R_g = (N/6)^{1/2}l$.

19.5 Distinguish between molecular mechanics and molecular dynamics calculations. Why are these methods generally more popular in the field of polymer chemistry than the quantum mechanical procedures discussed in Chapter 11?

19.6 It is observed that the critical micelle concentration of sodium dodecyl sulfate in aqueous solution decreases as the concentration of added sodium chloride increases. Explain this effect.

19.7 Explain the physical origins of surface activity by surfactant molecules.

19.8 Discuss the physical origins of the surface Gibbs energy.

19.9 Self-assembled monolayers (SAMs) are receiving more attention than Langmuir–Blodgett (LB) films as starting points for nanofabrication. How do SAMs differ from LB films and why are SAMs more useful than LB films in nanofabrication work?

Exercises

19.1a Calculate the number-average molar mass and the mass-average molar mass of a mixture of equal amounts of two polymers, one having $M = 62$ kg mol^{-1} and the other $M = 78$ kg mol^{-1}.

19.1b Calculate the number-average molar mass and the mass-average molar mass of a mixture of two polymers, one having $M = 62$ kg mol^{-1} and the other $M = 78$ kg mol^{-1}, with their amounts (numbers of moles) in the ratio 3:2.

19.2a The radius of gyration of a long chain molecule is found to be 7.3 nm. The chain consists of C—C links. Assume the chain is randomly coiled and estimate the number of links in the chain.

19.2b The radius of gyration of a long chain molecule is found to be 18.9 nm. The chain consists of links of length 450 pm. Assume the chain is randomly coiled and estimate the number of links in the chain.

19.3a A solution consists of solvent, 30 per cent by mass, of a dimer with $M = 30$ kg mol^{-1} and its monomer. What average molar mass would be obtained from measurement of (a) osmotic pressure, (b) light scattering?

19.3b A solution consists of 25 per cent by mass of a trimer with $M = 22$ kg mol^{-1} and its monomer. What average molar mass would be obtained from measurement of: (a) osmotic pressure, (b) light scattering?

19.4a Evaluate the rotational correlation time, $\tau_R = 4\pi a^3\eta/3kT$, for serum albumin in water at 25°C on the basis that it is a sphere of radius 3.0 nm. What is the value for a CCl_4 molecule in carbon tetrachloride at 25°C? (Viscosity data in Table 21.4 in the *Data section* at the end of this volume; take $a(CCl_4) = 250$ pm.)

19.4b Evaluate the rotational correlation time, $\tau_R = 4\pi a^3\eta/3kT$, for a synthetic polymer in water at 20°C on the basis that it is a sphere of radius 4.5 nm.

19.5a What is the relative rate of sedimentation for two spherical particles of the same density, but which differ in radius by a factor of 10?

19.5b What is the relative rate of sedimentation for two spherical particles with densities 1.10 g cm^{-3} and 1.18 g cm^{-3} and which differ in radius by a factor of 8.4, the former being the larger? Use $\rho = 0.794$ g cm^{-3} for the density of the solution.

19.6a Human haemoglobin has a specific volume of 0.749×10^{-3} m^3 kg^{-1}, a sedimentation constant of 4.48 Sv, and a diffusion coefficient of 6.9×10^{-11} m^2 s^{-1}. Determine its molar mass from this information.

19.6b A synthetic polymer has a specific volume of 8.01×10^{-4} m^3 kg^{-1}, a sedimentation constant of 7.46 Sv, and a diffusion coefficient of 7.72×10^{-11} m^2 s^{-1}. Determine its molar mass from this information.

19.7a Find the drift speed of a particle of radius 20 μm and density 1750 kg m^{-3} which is settling from suspension in water (density 1000 kg m^{-3}) under the influence of gravity alone. The viscosity of water is 8.9×10^{-4} kg m^{-1} s^{-1}.

19.7b Find the drift speed of a particle of radius 15.5 μm and density 1250 kg m^{-3} which is settling from suspension in water (density 1000 kg m^{-3}) under the influence of gravity alone. The viscosity of water is 8.9×10^{-4} kg m^{-1} s^{-1}.

19.8a At 20°C the diffusion coefficient of a macromolecule is found to be 8.3×10^{-11} m^2 s^{-1}. Its sedimentation constant is 3.2 Sv in a solution of density 1.06 g cm^{-3}. The specific volume of the macromolecule is 0.656 cm^3 g^{-1}. Determine the molar mass of the macromolecule.

19.8b At 20°C the diffusion coefficient of a macromolecule is found to be 7.9×10^{-11} m^2 s^{-1}. Its sedimentation constant is 5.1 Sv in a solution of density 997 kg m^{-3}. The specific volume of the macromolecule is 0.721 cm^3 g^{-1}. Determine the molar mass of the macromolecule.

19.9a The data from a sedimentation equilibrium experiment performed at 300 K on a macromolecular solute in aqueous solution show that a graph of ln c against r^2 is a straight line with a slope of 729 cm^{-2}. The rotational rate of the centrifuge was 50 000 r.p.m. The specific volume of the solute is 0.61 cm^3 g^{-1}. Calculate the molar mass of the solute.

19.9b The data from a sedimentation equilibrium experiment performed at 293 K on a macromolecular solute in aqueous solution show that a graph of ln c against $(r/\text{cm})^2$ is a straight line with a slope of 821. The rotation rate of the centrifuge was 1080 Hz. The specific volume of the solute is 7.2×10^{-4} m^3 kg^{-1}. Calculate the molar mass of the solute.

19.10a Calculate the radial acceleration (as so many g) in a cell placed at 6.0 cm from the centre of rotation in an ultracentrifuge operating at 80 000 r.p.m.

19.10b Calculate the radial acceleration (as so many g) in a cell placed at 5.50 cm from the centre of rotation in an ultracentrifuge operating at 1.32 kHz.

19.11a A polymer chain consists of 700 segments, each 0.90 nm long. If the chain were ideally flexible, what would be the r.m.s. separation of the ends of the chain?

19.11b A polymer chain consists of 1200 segments, each 1.125 nm long. If the chain were ideally flexible, what would be the r.m.s. separation of the ends of the chain?

19.12a Calculate the contour length (the length of the extended chain) and the root mean square separation (the end-to-end distance) for polyethylene with a molar mass of 280 kg mol^{-1}.

19.12b Calculate the contour length (the length of the extended chain) and the root mean square separation (the end-to-end distance) for polypropylene of molar mass 174 kg mol^{-1}.

Problems*

Numerical problems

19.1 In a sedimentation experiment the position of the boundary as a function of time was found to be as follows:

t/min	15.5	29.1	36.4	58.2
r/cm	5.05	5.09	5.12	5.19

The rotation rate of the centrifuge was 45 000 r.p.m. Calculate the sedimentation constant of the solute.

19.2 Calculate the speed of operation (in r.p.m.) of an ultracentrifuge needed to obtain a readily measurable concentration gradient in a sedimentation equilibrium experiment. Take that gradient to be a concentration at the bottom of the cell about five times greater than at the top. Use $r_{top} = 5.0$ cm, $r_{bott} = 7.0$ cm, $M \approx 10^5$ g mol^{-1}, $\rho v_s \approx 0.75$, $T = 298$ K.

19.3 The concentration dependence of the viscosity of a polymer solution is found to be as follows:

c/(g dm^{-3})	1.32	2.89	5.73	9.17
η/(g m^{-1} s^{-1})	1.08	1.20	1.42	1.73

The viscosity of the solvent is 0.985 g m^{-1} s^{-1}. What is the intrinsic viscosity of the polymer?

19.4 The times of flow of dilute solutions of polystyrene in benzene through a viscometer at 25°C are given in the table below. From these data, calculate the molar mass of the polystyrene samples. Since the solutions are dilute, assume that the densities of the solutions are the same as those of pure benzene. η(benzene) $= 0.601 \times 10^{-3}$ kg m^{-1} s^{-1} (0.601 cP) at 25°C.

c/(g dm^{-3})	0	2.22	5.00	8.00	10.00
t/s	208.2	248.1	303.4	371.8	421.3

19.5 The viscosities of solutions of polyisobutylene in benzene were measured at 24°C (the θ temperature for the system) with the following results:

c/(g/10^2 cm^3)	0	0.2	0.4	0.6	0.8	1.0
η/(10^{-3} kg m^{-1} s^{-1})	0.647	0.690	0.733	0.777	0.821	0.865

Use the information in Table 19.4 to deduce the molar mass of the polymer.

19.6‡ Polystyrene in cyclohexane at 34.5°C forms a θ solution, with an intrinsic viscosity related to the molar mass by $[\eta] = KM^a$. The following data

* Problems denoted with the symbol ‡ were supplied by Charles Trapp and Carmen Giunta.

on polystyrene in cyclohexane are taken from L.J. Fetters, N. Hadjichristidis, J.S. Lindner, and J.W. Mays (*J. Phys. Chem. Ref. Data* **23**, 619 (1994)):

M/(kg mol^{-1})	10.0	19.8	106	249	359	860	1800	5470	9720	56 800
$[\eta]$/(cm^3 g^{-1})	8.90	11.9	28.1	44.0	51.2	77.6	113.9	195	275	667

Determine the parameters K and a. What is the molar mass of a polystyrene that forms a θ solution in cyclohexane with $[\eta] = 100$ cm^3 g^{-1}?

19.7‡ Standard polystyrene solutions of known average molar masses continue to be used as for the calibration of many methods of characterizing polymer solutions. M. Kolinsky and J. Janca (*J. Polym. Sci., Polym. Chem.* **12**, 1181 (1974)) studied polystyrene in tetrahydrofuran (THF) for use in calibrating a gel permeation chromatograph. Their results for the intrinsic viscosity, $[\eta]$, as a function of average molar mass at 25°C are given in the table below. (a) Obtain the Mark–Houwink constants that fit these data. (b) Compare your values to those in Table 19.4 and Example 19.5. How might you explain the differences?

$\bar{M}_v$/(kg mol^{-1})	5.0	10.3	19.85	51	98.2	173	411	867
$[\eta]$/(cm^3 g^{-1})	5.2	8.8	14.0	27.6	43.6	67.0	125.0	206.7

19.8 The concentration dependence of the osmotic pressure of solutions of a macromolecule at 20°C was found to be as follows:

c/(g dm^{-3})	1.21	2.72	5.08	6.60
Π/Pa	134	321	655	898

Determine the molar mass of the macromolecule and the osmotic virial coefficient.

19.9 The osmotic pressure of a fraction of poly(vinyl chloride) in a ketone solvent was measured at 25°C. The density of the solvent (which is virtually equal to the density of the solution) was 0.798 g cm^{-3}. Calculate the molar mass and the osmotic virial coefficient, B, of the fraction from the following data:

c/(g /10^2 cm^3)	0.200	0.400	0.600	0.088	1.000
h/cm	0.48	1.2	1.86	2.76	3.88

19.10 The following table lists the glass transition temperatures, T_g, of several polymers. Discuss the reasons why the structure of the monomer unit has an effect on the value of T_g.

Polymer	Poly(oxymethylene)	Polyethylene	Poly(vinyl chloride)	Polystyrene
Structure	$-(OCH_2)_n-$	$-(CH_2CH_2)_n-$	$-(CH_2{-}CHCl)_n-$	$-(CH_2{-}CH(C_6H_5))_n-$
T_g/K	198	253	354	381

Theoretical problems

19.11 In formamide as solvent, poly(γ-benzyl-L-glutamate) is found by light scattering experiments to have a radius of gyration proportional to M; in contrast, polystyrene in butanone has R_g proportional to $M^{1/2}$. Present arguments to show that the first polymer is a rigid rod whereas the second is a random coil.

19.12 The *kinematic viscosity*, ν, of a fluid is defined as η/ρ, where ρ is the mass density. What are the SI units of kinematic viscosity? Confirm that the drainage times through a narrow tube are governed by the kinematic viscosity by referring to the Poiseuille equation for fluid flow (eqn 21.25) and hence confirm eqn 19.24.

19.13 A polymerization process produced a Gaussian distribution of polymers in the sense that the proportion of molecules having a molar mass in the range M to $M + \mathrm{d}M$ was proportional to $\mathrm{e}^{-(M-\bar{M})^2/2\gamma}\mathrm{d}M$. What is the number average molar mass when the distribution is narrow?

19.14 Show how eqn 19.26 for a one-dimensional freely jointed chain can be used to derive eqn 19.27 for a three-dimensional freely-jointed chain. *Hint.* Write the probability that the ends lie in the range n_x to $n_x + \mathrm{d}n_x$ as $\mathrm{d}P_x = P\mathrm{d}n_x$, with P given in eqn 19.26, and similarly for the other two dimensions. Multiply these probabilities together, and integrate $\mathrm{d}n_x\mathrm{d}n_y\mathrm{d}n_z$ over a shell of thickness $\mathrm{d}n$. Don't count negative integers (that is, divide the volume of the shell by 8, corresponding to the all-positive octant of values).

19.15 Use eqn 19.27 to deduce expressions for (a) the root mean square separation of the ends of the chain, (b) the mean separation of the ends, and (c) their most probable separation. Evaluate these three quantities for a fully flexible chain with $N = 4000$ and $l = 154$ pm.

19.16 Construct a two-dimensional random walk by using a random number generating routine with mathematical software or electronic spreadsheet. Construct a walk of 50 and 100 steps. If there are many people working on the problem, investigate the mean and most probable separations in the plots by direct measurement. Do they vary as $N^{1/2}$?

19.17 Evaluate the radius of gyration, R_g, of (a) a solid sphere of radius a, (b) a long straight rod of radius a and length l. Show that in the case of a solid sphere of specific volume v_s, $R_g/\mathrm{nm} \approx 0.056902 \times \{(v_s/\mathrm{cm^3\,g^{-1}})(M/\mathrm{g\,mol^{-1}})\}^{1/3}$. Evaluate R_g for a species with $M = 100$ kg mol^{-1}, $v_s = 0.750$ cm^3 g^{-1}, and, in the case of the rod, of radius 0.50 nm.

19.18 The effective radius, a, of a random coil is related to its radius of gyration, R_g, by $a = \gamma R_g$, with $\gamma = 0.85$. Deduce an expression for the osmotic virial coefficient, B, in terms of the number of chain units for (a) a freely jointed chain, (b) a chain with tetrahedral bond angles. Evaluate B for $l = 154$ pm and $N = 4000$. Estimate B for a randomly coiled polyethylene chain of arbitrary molar mass, M, and evaluate it for $M = 56$ kg mol^{-1}. Use $B = \frac{1}{2}N_A v_P$, where v_P is the excluded volume due to a single molecule.

19.19 Radius of gyration is defined in eqn 19.32. Show that an equivalent definition is that R_g is the average root mean square distance of the atoms or groups (all assumed to be of the same mass), that is, that $R_g^2 = (1/N)\sum_j R_j^2$, where R_j is the distance of atom j from the centre of mass.

19.20 Consider the thermodynamic description of stretching rubber. The observables are the tension, t, and length, l (the analogues of p and V for gases). Because $\mathrm{d}w = t\mathrm{d}l$, the basic equation is $\mathrm{d}U = T\mathrm{d}S + t\mathrm{d}l$. (The term $p\mathrm{d}V$ is supposed negligible throughout.) If $G = U - TS - tl$, find expressions for $\mathrm{d}G$ and $\mathrm{d}A$, and deduce the Maxwell relations

$$\left(\frac{\partial S}{\partial l}\right)_T = -\left(\frac{\partial t}{\partial T}\right)_l \qquad \left(\frac{\partial S}{\partial t}\right)_T = -\left(\frac{\partial l}{\partial T}\right)_t$$

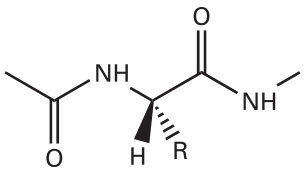
Go on to deduce the equation of state for rubber,

$$\left(\frac{\partial U}{\partial l}\right)_T = t - \left(\frac{\partial t}{\partial T}\right)_l$$

19.21 On the assumption that the tension required to keep a sample at a constant length is proportional to the temperature ($t = aT$, the analogue of $p \propto T$), show that the tension can be ascribed to the dependence of the entropy on the length of the sample. Account for this result in terms of the molecular nature of the sample.

Applications: to biochemistry and technology

19.22 In this problem you will use molecular mechanics software of your instructor's choice to gain some appreciation for the complexity of the calculations that lead to plots such as those in Fig. 19.29. Our model for the protein is the dipeptide (**15**) in which the terminal methyl groups replace the rest of the polypeptide chain. (a) Draw three initial conformers of the dipeptide with R = H: one with $\phi = +75°$, $\psi = -65°$, a second with $\phi = \psi = +180°$, and a third with $\phi = +65°$, $\psi = +35°$. Use a molecular mechanics routine to optimize the geometry of each conformer and measure the total potential energy and the final ϕ and ψ angles in each case. (Although any force field will work satisfactorily, the AMBER force field is strongly recommended, as it is optimized for calculations on biopolymers.) Did all of the initial conformers converge to the same final conformation? If not, what do these



15

final conformers represent? Rationalize any observed differences in total potential energy of the final conformers. (b) Use the approach in part (a) to investigate the case R = CH_3, with the same three initial conformers as starting points for the calculations. Rationalize any similarities and differences between the final conformers of the dipeptides with R = H and R = CH_3.

19.23 Calculate the excluded volume in terms of the molecular volume on the basis that the molecules are spheres of radius a. Evaluate the osmotic virial coefficient in the case of bushy stunt virus, $a = 14.0$ nm, and haemoglobin, $a = 3.2$ nm (see Problem 19.18). Evaluate the percentage deviation of the Rayleigh ratios of 1.00 g/(100 cm^3) solutions of bushy stunt virus ($M = 1.07 \times 10^4$ kg mol^{-1}) and haemoglobin ($M = 66.5$ kg mol^{-1}) from the ideal solution values. In eqn 19.8, let $P_\theta = 1$ and assume that both solutions have the same K value.

19.24 Use the information below and the expression for R_g of a solid sphere quoted in the Problem 19.17, to classify the species below as globular or rod-like.

	M/(g mol^{-1})	v_s/(cm^3 g^{-1})	R_g/nm
Serum albumin	66×10^3	0.752	2.98
Bushy stunt virus	10.6×10^6	0.741	12.0
DNA	4×10^6	0.556	117.0

19.25 Suppose that a rod-like DNA molecule of length 250 nm undergoes a conformational change to a closed-circular (cc) form. (a) Use the information in Problem 19.24 and an incident wavelength $\lambda = 488$ nm to calculate the ratio of scattering intensities by each of these conformations, I_{rod}/I_{cc}, when $\theta = 20°$, 45°, and 90°. (b) Suppose that you wish to use light scattering as a technique for the study of conformational changes in DNA molecules. Based on your answer to part (a), at which angle would you conduct the experiments? Justify your choice.

19.26 In an ultracentrifugation experiment at 20°C on bovine serum albumin the following data were obtained: $\rho = 1.001$ g cm^{-3}, $v_s = 1.112$ cm^3 g^{-1}, $\omega/2\pi = 322$ Hz,

r/cm	5.0	5.1	5.2	5.3	5.4
c/(mg cm^{-3})	0.536	0.284	0.148	0.077	0.039

Evaluate the molar mass of the sample.

19.27 Sedimentation studies on haemoglobin in water gave a sedimentation constant $S = 4.5$ Sv at 20°C. The diffusion coefficient is 6.3×10^{-11} m^2 s^{-1} at the same temperature. Calculate the molar mass of haemoglobin using $v_s = 0.75$ cm^3 g^{-1} for its partial specific volume and $\rho = 0.998$ g cm^{-3} for the density of the solution. Estimate the effective radius of the haemoglobin molecule given that the viscosity of the solution is 1.00×10^{-3} kg m^{-1} s^{-1}.

19.28 The rate of sedimentation of a recently isolated protein was monitored at 20°C and with a rotor speed of 50 000 r.p.m. The boundary receded as follows:

t/s	0	300	600	900	1200	1500	1800
r/cm	6.127	6.153	6.179	6.206	6.232	6.258	6.284

Calculate the sedimentation constant and the molar mass of the protein on the basis that its partial specific volume is 0.728 cm^3 g^{-1} and its diffusion coefficient is 7.62×10^{-11} m^2 s^{-1} at 20°C, the density of the solution then being 0.9981 g cm^{-3}. Suggest a shape for the protein given that the viscosity of the solution is 1.00×10^{-3} kg m^{-1} s^{-1} at 20°C.

19.29 For some proteins, the isoelectric point must be obtained by extrapolation because the macromolecule might not be stable over a very wide pH range. Estimate the pH of the isoelectric point from the following data for a protein:

pH	4.5	5.0	5.5	6.0
Drift speed/(μm s^{-1})	−0.10	−0.20	−0.30	−0.35

19.30 Here we use concepts developed in Chapter 16 and this chapter to enhance our understanding of closed-circular and supercoiled DNA. (a) The average end-to-end distance of a flexible polymer (such as a fully denatured polypeptide or a strand of DNA) is $N^{1/2}l$, where N is the number of groups (residues or bases) and l is the length of each group. Initially, therefore, one end of the polymer can be found anywhere within a sphere of radius $N^{1/2}l$ centred on the other end. When the ends join to form a circle, they are confined to a volume or radius l. What is the change in molar entropy? Plot the function you derive as a function of N. (b) The energy necessary to twist ccDNA by i turns is $\varepsilon_i = ki^2$, with k an empirical constant and i being negative or positive depending on the sense of the twist. For example, one twist ($i = \pm 1$) makes ccDNA resemble the number 8. (i) Show that the distribution of the populations $p_i = n_i/N$ of ccDNA molecules with i turns at a specified temperature has the form of a Gaussian function. (ii) Plot the expression you derived in part (a) for several values of the temperature. Does the curve has a maximum? If so, at what value of i? Comment on variations of the shape of the curve with temperature. (iii) Calculate p_0, p_1, p_5, and p_{10} at 298 K.

19.31 The melting temperature of a DNA molecule can be determined by differential scanning calorimetry (*Impact* I2.1). The following data were obtained in aqueous solutions containing the specified concentration c_{salt} of an soluble ionic solid for a series of DNA molecules with varying base pair composition, with f the fraction of G—C base pairs:

$c_{salt} = 1.0 \times 10^{-2}$ mol dm^{-3}

f	0.375	0.509	0.589	0.688	0.750
T_m/K	339	344	348	351	354

$c_{salt} = 0.15$ mol dm^{-3}

f	0.375	0.509	0.589	0.688	0.750
T_m/K	359	364	368	371	374

(a) Estimate the melting temperature of a DNA molecule containing 40.0 per cent G—C base pairs in both samples. *Hint.* Begin by plotting T_m against fraction of G—C base pairs and examining the shape of the curve. (b) Do the data show an effect of concentration of ions in solution on the melting temperature of DNA? If so, provide a molecular interpretation for the effect you observe.

19.32 The fluidity of a lipid bilayer dispersed in aqueous solution depends on temperature and there are two important melting transitions. One transition is from a 'solid crystalline' state in which the hydrophobic chains are packed together tightly (hence move very little) to a 'liquid crystalline state', in which there is increased but still limited movement of the of the chains. The second transition, which occurs at a higher temperature than the first, is from the liquid crystalline state to a liquid state, in which the hydrophobic interactions holding the aggregate together are largely disrupted. (a) It is observed that the transition temperatures increase with the hydrophobic chain length and decrease with the number of C=C bonds in the chain. Explain these observations. (b) What effect is the inclusion of cholesterol likely to have on the transition temperatures of a lipid bilayer? Justify your answer.

19.33 Polystyrene is a synthetic polymer with the structure —(CH_2—$CH(C_6H_5)$)$_n$—. A batch of polydisperse polystyrene was prepared by initiating the polymerization with t-butyl radicals. As a result, the t-butyl group is

expected to be covalently attached to the end of the final products. A sample from this batch was embedded in an organic matrix containing silver trifluoroacetate and the resulting MALDI-TOF spectrum consisted of a large number of peaks separated by 104 g mol^{-1}, with the most intense peak at 25 578 g mol^{-1}. Comment on the purity of this sample and determine the number of (CH_2—$CH(C_6H_5)$) units in the species that gives rise to the most intense peak in the spectrum .

19.34 A manufacturer of polystyrene beads claims that they have an average molar mass of 250 kg mol^{-1}. Solutions of these beads are studied by a physical chemistry student by dilute solution viscometry with an Ostwald viscometer in both the 'good' solvent toluene and the theta solvent cyclohexane. The drainage times, t_D, as a function of concentration for the two solvents are given in the table below. (a) Fit the data to the virial equation for viscosity,

$$\eta = \eta^*(1 + [\eta]c + k'[\eta]^2c^2 + \cdots)$$

where k' is called the *Huggins constant* and is typically in the range 0.35–0.40. From the fit, determine the intrinsic viscosity and the Huggins constant. (b) Use the empirical Mark–Kuhn–Houwink–Sakurada equation (eqn 19.25) to determine the molar mass of polystyrene in the two solvents. For theta solvents, $a = 0.5$ and $K = 8.2 \times 10^{-5}$ dm^3 g^{-1} for cyclohexane; for the good solvent toluene $a = 0.72$ and $K = 1.15 \times 10^{-5}$ dm^3 g^{-1}. (c) According to a general theory proposed by Kirkwood and Riseman, the root mean square end-to-end distance of a polymer chain in solution is related to $[\eta]$ by $[\eta] = \Phi\langle r^2\rangle^{3/2}/M$, where Φ is a universal constant with the value 2.84×10^{26} when $[\eta]$ is expressed in cubic decimetres per gram and the distance is in metres. Calculate this quantity for each solvent. (d) From the molar masses calculate the average number of styrene ($C_6H_5CH{=}CH_2$) monomer units, $\langle n\rangle$, (e) Calculate the length of a fully stretched, planar zigzag configuration, taking the C—C distance as 154 pm and the CCC bond angle to be 109°. (f) Use eqn 19.33 to calculate the radius of gyration, R_g. Also calculate $\langle r^2\rangle^{1/2} = n^{1/2}$. Compare this result with that predicted by the Kirkwood–Riseman theory: which gives the better fit? (g) Compare your values for M to the results of Problem 19.33. Is there any reason why they should or should not agree? Is the manufacturer's claim valid?

c/(g dm^{-3} toluene)	0	1.0	3.0	5.0
t_D/s	8.37	9.11	10.72	12.52
c/(g dm^{-3} cyclohexane)	0	1.0	1.5	2.0
t_D/s	8.32	8.67	8.85	9.03

19.35‡ The determination of the average molar masses of conducting polymers is an important part of their characterization. S. Holdcroft (*J. Polym. Sci., Polym. Phys.* **29**, 1585 (1991)) has determined the molar mases and Mark–Houwink constants for the electronically conducting polymer, poly(3-hexylthiophene) (P3HT) in tetrahydrofuran (THF) at 25°C by methods similar to those used for nonconducting polymers. The values for molar mass and intrinsic viscosity in the table below are adapted from their data. Determine the constants in the Mark–Kuhn–Houwink–Sakurada equation from these results and compare to the values obtained in your solution to Problem 19.7 .

$\bar{M}_v$/(kg mol^{-1})	3.8	11.1	15.3	58.8
$[\eta]$/(cm^3 g^{-1})	6.23	17.44	23.73	85.28