6 Macromolecules

Small molecules contain at least two, but less than ten atoms. Common examples of large molecules are the chlorophylls, whose molar mass is still less than 1000. Macromolecules have molar masses in excess of 10^3 or 10^4; there is no common definition.

Naturally occurring macromolecules are cellulose, proteins and polypeptides (e.g. enzymes), and polynucleotides (e.g. DNA, deoxyribo nucleic acid). Artificial macromolecules include polymers such as Nylon, polyethylene, polystyrene or Teflon, which are made by joining smaller molecular groups, the monomers. Macromolecules can be also formed from smaller molecular groups by intermolecular forces as molecular clusters or inclusion compounds. The special supermolecules in biological systems with a specific function are called molecular functional units.

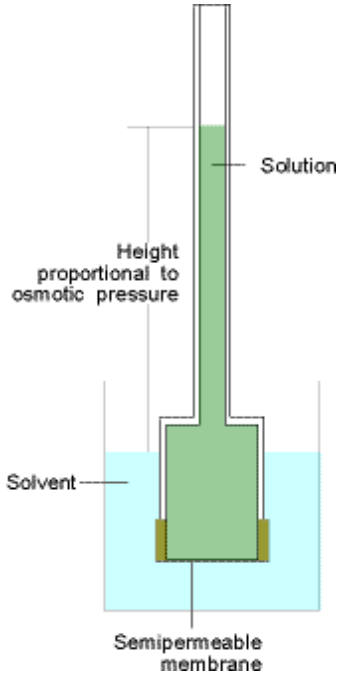
The methods of molecular physics are used on macromolecules just as they are for smaller molecules. A special chapter entitled "Macromolecules" can be justified by

- special methods for the examination of macromolecules and special structural conformations of macromolecules or
- by the numerous current research areas in the world of macromolecules, as a challenge for the molecular physicist.

The methodical approach is described in the textbook "Physical Chemistry, 1999" by P.W. Atkins (referred to as Atkins E6). For example, synthetic polymers are polydisperse, which means that they contain a mixture of macromolecules with different chain lengths and molar masses. We therefore need a definition of an average molar mass, which will be given in the next subchapter 6.1. But for the illustration of the modern world of macromolecules we will finally accept the thematically based description of macromolecules by H. Haken and H.Ch. Wolf in their textbook "Molecular Physics and Elements of Quantum Chemistry, 1995" (referred to as Haken and Wolf) for the subchapters 6.2 - 6.6.

6.1 Average Molar Masses

The *number-average molar mass* can be determined by osmometry. The phenomenon of osmosis (Greek οσμοσ: push) is the tendency of the pure solvent, when separated by a semipermeable membrane (permeable to the solvent but not to the solute) from more concentrated solution, to diffuse through into that solution. The osmotic pressure, Π, is the pressure on the solution which is necessary to stop the influx of the solvent molecules. See at left Fig. 23.1 taken from Atkins.



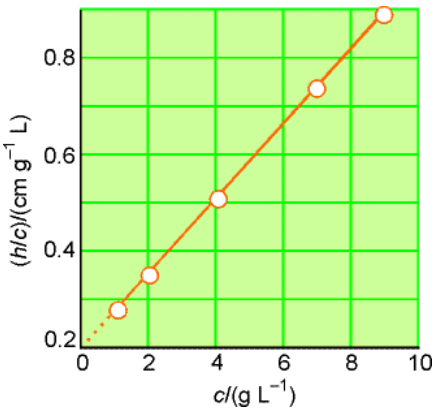
Without proof, we introduce the van't Hoff equation $\Pi = c_i\, RT$, where $c_i = n_i/V$ (mole per volume) denotes the molar concentration of the solute *i*. The van't Hoff equation holds for ideal solutions and can be expanded to real solutions by a virial-like expansion

$$\Pi = c_i\, RT\, [1 + B\, c_i + \ldots\,], \qquad (6.01)$$

where the constant *B* is called the osmotic virial coefficient.

Now the osmotic pressure Π is measured at a series of mass concentrations, γ_i / g L^{-1}, and a plot of Π/γ_i against γ_i is used to determine the molar mass, M_i, of the solute i. With $c_i = \gamma_i / M_i$ and the hydrostatic pressure $\Pi = \rho\, g\, h$, where g ≈ 9,81 m s^{-2} and ρ can be approximated by the density of the solvent, we get from equ. (6.01) for the height h of the solution level above the solvent level

$$\frac{h}{\gamma_i} = \frac{RT}{\rho\, g\, M_i}\left(1 + \frac{B\gamma_i}{M_i} + \ldots\right) = \frac{RT}{\rho\, g\, M_i} + \frac{RT\, B\gamma_i}{\rho\, g\, M_i^2} + \ldots \qquad (6.02)$$

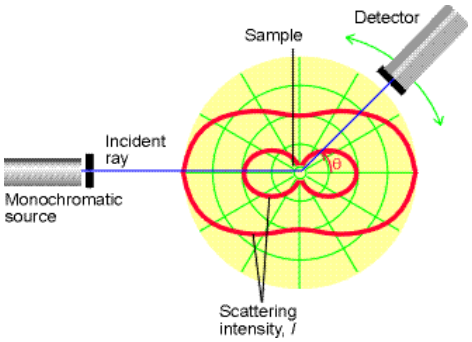


A corresponding plot taken from Atkins Fig. 7.25 describes the solution of PVC in cyclohexanone at 298 K and is given at left. Note that the mass concentration is denoted as c in the figure instead of γ in the text. The latter is in agreement with IUPAC. The figure gives the intercept 0,21 ≈ $h/\gamma_i = RT / \rho\, g\, M_i$ yielding a molar mass of PVC as 1,2 ×10^2 kg mol^{-1}.

PVC is polydisperse. The obtained molar mass is an average value. Equation (6.01) shows that the osmotic pressure is proportional to the molar concentration c_i of the solute i. The molar concentration c_i and the number of particles per volume N_i are connected by the relation $N_i = N_A\, c_i$ with the Avogadro number N_A. Therefore, the molar mass obtained by osmometry is the *number-average molar mass*

$$M_{\text{number}} = \frac{\sum_i N_i M_i}{\sum_i N_i}. \qquad (6.03)$$

Without proof of the used equations we present a simple procedure for the estimation of the volume of a macromolecule:
The osmotic virial coefficient B introduced in equ. (6.01) is related to the excluded volume of the solution (into which the centre of a dilute molecule cannot penetrate): $B = ½\, N_A\, v_{\text{excluded}}$. The excluded volume is eight times the volume of the molecule itself: $v_{\text{excluded}} = 8\, v_{\text{molecule}}$ The value of B can be determined from equ. (6.02) dividing the slope by the intercept; see the figure above.



Light diffraction methods are used to determine the *mass-averaged* molar mass. Rayleigh scattering is caused by particles with diameters much smaller than the wavelength of the incident radiation. The Rayleigh scattering from a point-like particle is presented at right, Fig. 23.8 taken from Atkins. The scattering Intensity I depends on the scattering angle θ, on the incident intensity I_0, on the molar concentration c of the solute, on the mass averaged molar mass M_{mass}, on the forth power of the frequency of the radiation and the angle dependence $g(\theta) = 1 + \cos^2\theta$ for unpolarized light (outer trace in the figure above)and $g(\theta) = \cos^2\theta$ for plane-polarized light (inner trace):

$$I(\theta) \approx I_0\, c\, M_{\text{mass}}\, \nu^4\, g(\theta). \qquad (6.04)$$

Without deriving the Rayleigh scattering law, equ. (6.04), we can make the following hand-waving argument: The electric field of a dipole with the displacement of charges $r = r_0\, e^{i\omega t}$ is proportional to their acceleration $d^2r/dt^2 = -\omega^2\, r_0\, e^{i\omega t}$. Since the radiation power is proportional to the square of the electric field, we get finally the frequency in the forth power, see equ. (4.99) in chapter 4. This frequency dependence causes the blue light of the sky and the red light of the sun at sunrise or sunset.

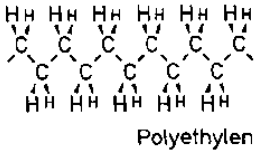
Equation (6.04) contains the average mass of the scattering atom. Therefore, scattering experiments for the determination of a average molar mass of macromolecules yield the mass average or weight-average mass, which can be calculated by weighting the molar masses M_i of the molecules by the mass of each m_i present in the polydisperse material of the total mass m:

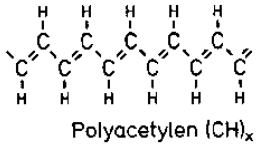
$$M_{\text{mass}} = \frac{\sum_i m_i M_i}{m} = \frac{\sum_i N_i M_i^2}{\sum_i N_i M_i}\,. \tag{6.05}$$

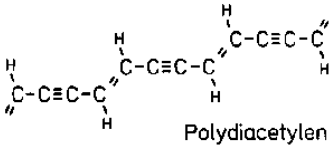
However, the size of macromolecules can be in magnitude of the wavelength of the light. When the wavelength of the incident radiation is comparable to the size of the scattering particles, the so-called Mie-scattering occurs from different sites of the same molecule, and the interference between different rays becomes important. The net amplitude of the coherent scattered radiation is now the sum of the individual amplitudes, and the radiated power is proportional to N^2, if there are N coherent oscillators. In addition, the number of coherent oscillators is proportional to the surface of the sphere of a wavelength what means proportional to λ^2. Therefore, the radiated power is proportional to λ^4, canceling the ν^4 or $1/\lambda^4$ dependence of the isolated Rayleigh oscillator. In conclusion, the scattered radiation of particles with larger size is essentially frequency independent. Fog, clouds but also paper appear white by scattered light.

The difference between the observed scattering intensity for larger molecules and the intensity of a pure Rayleigh scattering (point-like particles) can be described by $I_{\text{observed}} = P\, I_{\text{Rayleigh}}$. For particles which are smaller than the wavelength of the incident light can be shown that $P - 1 \propto R_g^2$, where R_g denotes the radius of gyration. It is the radius of a thin hollow spherical shell of the same mass and moment of inertia as the molecule. This gives another approach for the determination of the size of a macromolecule.

6.2 Polymers


Polyethylen


Polyacetylen $(CH)_x$


Polydiacetylen

The figure at right, Fig. 20.1 taken von Haken and Wolf, shows the molecular structure of three important polymers. Note that ethylene is ethene ($H_2C{=}CH_2$), acetylen is ethin ($HC{\equiv}CH$) and diacetylen is butadiin ($HC{\equiv}C{-}C{\equiv}CH$). Polyethylene (PE), polyvinylchlorid (PVC) and polystyren (PS) are the top polymers with respect to the produced mass.

polyethylene: $(-CH_2-)_n$ with $1\,000 < n < 100\,000$

polyacetylen: $(=CH-)_n$

polydiacetylen: $(\equiv C-HC=)_n$

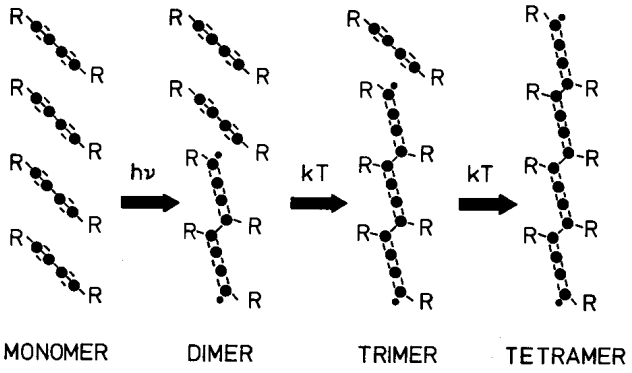


Fig. 20.3. The mechanism of the solid-state polymerisation of diacetylene. The formation of dimers is initiated by irradiation with light; the further progress of the polymerisation is thermally activated as an addition reaction. At the ends of the oligomer chains, the reactive radical electrons of the diradicals which are formed in the course of the polymerisation reaction are indicated as dots. After Sixl

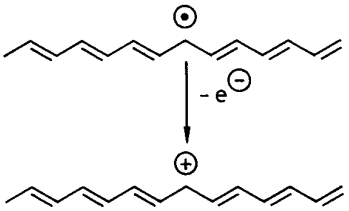


Fig. 20.6. A neutral and a charged "soliton" defect in the conjugated π-electron chain of polyacetylene. The upper part of the picture shows a neutral soliton, consisting of an alternating-bond defect and an unpaired electron (i.e. a free radical), which can move along the conjugated chain. In the lower part of the picture, an electron has been removed from the neutral soliton by oxidation. The remaining positively-charged soliton is likewise delocalized along the chain and is mobile

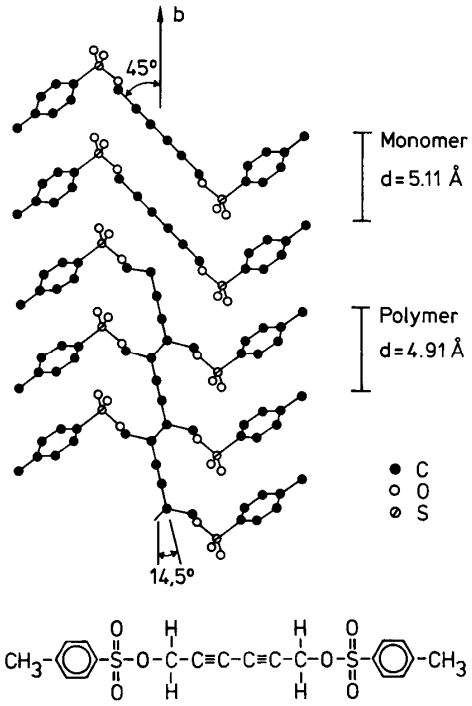


Fig. 20.2. The polymerisation of polydiacetylene. In the upper part of the figure, the X-ray structure of the monomeric diacetylene TS6 is shown (the notation refers to the particular side group, in general denoted by *R*); in the lower part of the figure, the structure of the polymerized crystal is indicated. The protons in the molecule are not shown. The complete formula of the TS6 monomer is given at the bottom of the figure

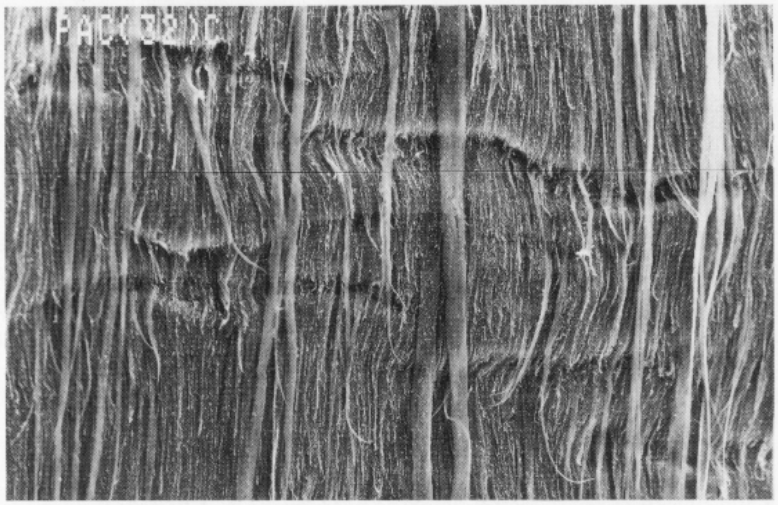


Fig. 20.5. A scanning electron microscope image of a polyacetylene foil. The foil contains more or less disordered strands of polymer molecules. A certain degree of preferred orientation was obtained by stretching the foil. The length of the picture corresponds to 125 μm. (Kindly placed at our disposal by M. Schwoerer)

Figures above and figures and table on the next page were taken from Haken and Wolf.

6.3 Molecular Recognition, Molecules in Confined Geometry

The biologically important mechanism of molecular recognition can be demonstrated using crown ethers. The industrial important heterogeneous catalysis in zeolites is an example how restrictive geometry causes form selective conversion of molecules.

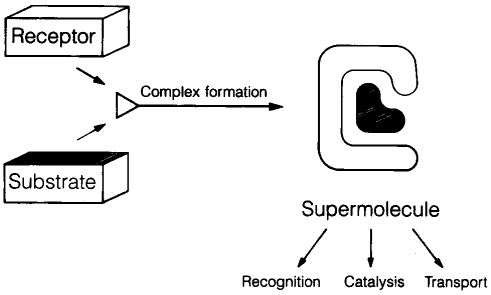
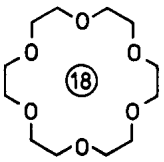
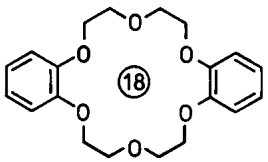


Fig. 20.7. The "receptor" and the "substrate" bind together to form a supermolecule or supermolecular complex. This forms the basis for the concept of of "molecular recognition", important in the catalysis and transport processes of molecular units. After J.M. Lehn

Fig. 20.8. Two crown ether molecules; the C atoms are not drawn in. *Notation*: the ring size, that is the number of bonds in the ring, is first given in square brackets, followed by "crown" or "C" for the class of molecules, and finally the number of donor positions is noted. In the ethers shown, these are the 6 oxygen atoms



[18] crown-6
18C6



Dibenzyl [18] crown-6
DB18C6

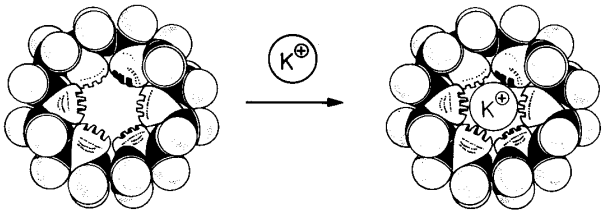
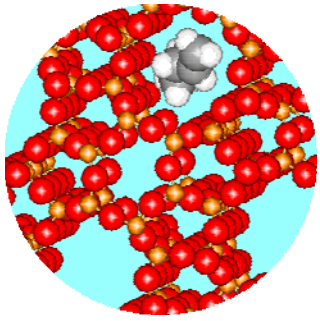


Fig. 20.9. Using a hard-sphere model, it can be shown that a K^+ ion just fits into the cavity of the crown ether [18]crown 6 and forms a complex there. After Vögtle

Table 20.1. A comparison of the diameters of various alkali-metal cations and crown compounds*

Cation	Diameter [pm]	Crown	Diameter [pm]
Li^+	136	[12] C4 (9)	120–150
Na^+	190	[15] C5 (10)	170–220
K^+	266	[18] C6 (7)	260–320
Cs^+	338	[21] C7 (11)	340–430

* After Vögtle



Zeolites are microcrystalline highly porous alumosilicates with an internal surface up to 1000 m^2 per cm^3 of the material. They are molecular sieves, since only molecules smaller as the windows of the cages can be absorbed. The figure at left shows the zeolite ferrierite with a molecule of butene adsorbed in it.

The figures below were taken from Weitkamp and Puppe: Catalysis and Zeolites, 1999.

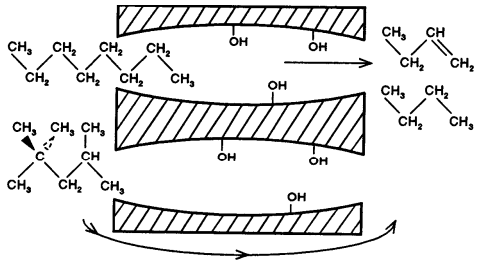


Fig. 5.7. Selective cracking of n-octane in the presence of 2,2,4-trimethylpentane as an example for reactant shape selectivity (adapted from [70])

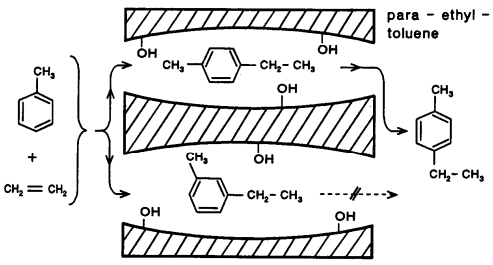


Fig. 5.8. Selective formation of *para*-ethyltoluene in the alkylation of toluene with ethylene as an example for product shape selectivity (adapted from [70])

6.4 Energy Transfer

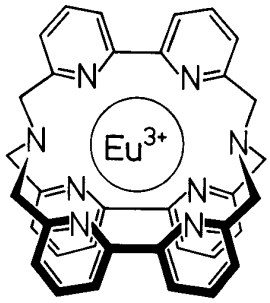


Fig. 20.12. The molecular structure of a europium (III) cryptate with a macrobicyclic polypyridine ligand, enclosing the Eu^{3+} ion. [After J.M. Lehn, Angew. Chem. **100**, 92 (1988)]

A simple example of the absorption of light by one subsystem, transfer of the absorbed energy to another subsystem, and emission of the light at a different wavelength by the second subsystem is found in Europium(III)-cryptate of the macrobiocyclic ligand polypiridine (figures taken from Haken and Wolf).

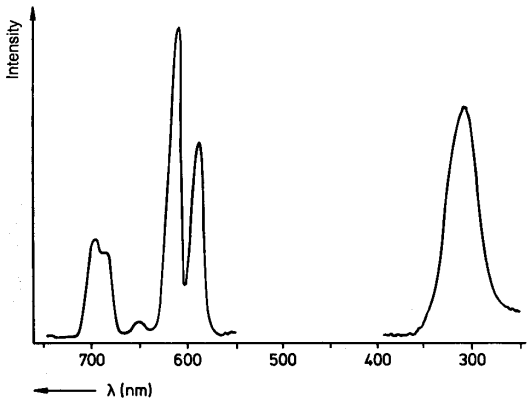
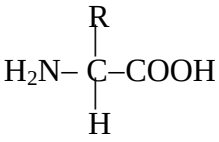


Fig. 20.13. *Right*: an excitation spectrum, that is the spectrum of absorbed light which gives rise to an emission at 700 nm, corresponding to the fluorescence of the molecule shown in Fig. 20.12. *Left*: the emission spectrum; it belongs to Eu^{3+}, while the absorption is accomplished by the organic ligand. This follows from the agreement of the excitation spectrum shown with an absorption band of the organic macrocycle. This is an example of intramolecular energy transfer from excited states of the ligand to excited states of the Eu^{3+} ion. The latter emits with a high yield, owing to its screening from the solvent by the organic partner. The measurements were carried out on a 10^{-6} molar solution at room temperature

6.5 Molecules as the Basic Units of Life



The most important molecular components of life are proteins and nucleic acids. Proteins are polymer chains of amino acids, which have an amino group, $–NH_2$, on one end and a carboxyl group, –COOH, at the other. If R is an H atom, the amino acid is called glycine. Amino acids polymerize to polypeptides at their reactive ends by the separation of H_2O.

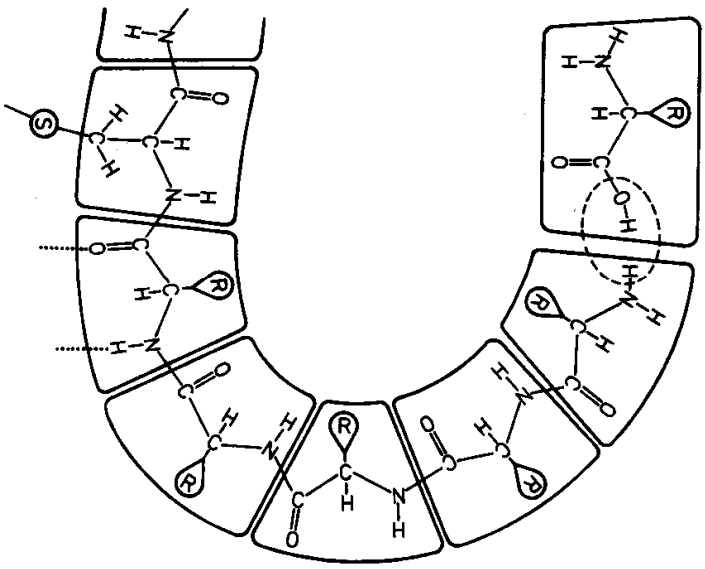


Fig. 20.20. A schematic representation of a peptide chain. Through elimination of a water molecule, amino acids can bind together, forming so-called peptide bonds. The backbone of the chain consists of many identical units. The variable side chains *R* determine the characteristic properties of each protein chain. The symbol *S* indicates here a crosslink to another part of the chain in the form of a disulphide bond. A different type of crosslinking is formed by hydrogen bonds (*dashed lines*). After Dickerson and Geis

60 to 600 of the 20 different amino acids come together by to form a protein by spiralling and folding, see figure at right taken from Haken and Wolf.

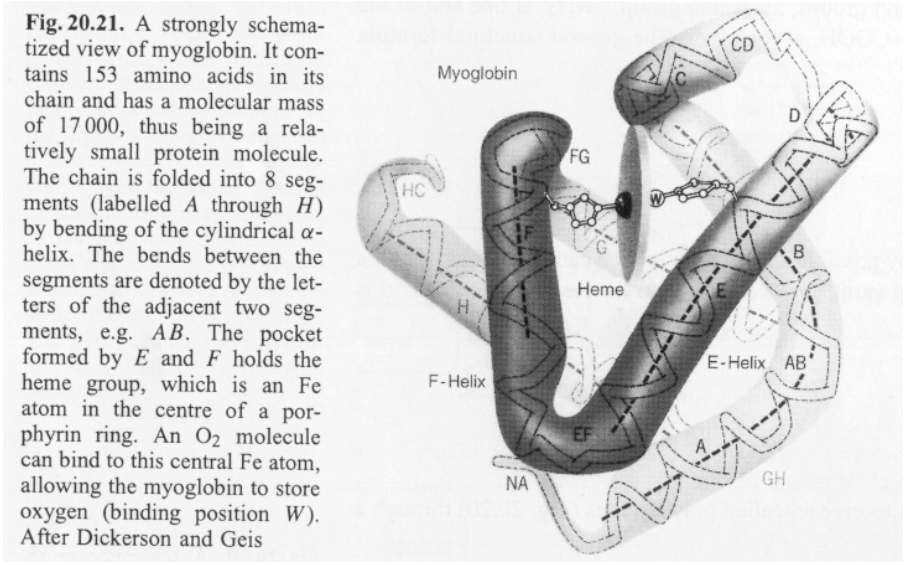


Fig. 20.21. A strongly schematized view of myoglobin. It contains 153 amino acids in its chain and has a molecular mass of 17 000, thus being a relatively small protein molecule. The chain is folded into 8 segments (labelled *A* through *H*) by bending of the cylindrical α-helix. The bends between the segments are denoted by the letters of the adjacent two segments, e.g. *AB*. The pocket formed by *E* and *F* holds the heme group, which is an Fe atom in the centre of a porphyrin ring. An O_2 molecule can bind to this central Fe atom, allowing the myoglobin to store oxygen (binding position *W*). After Dickerson and Geis

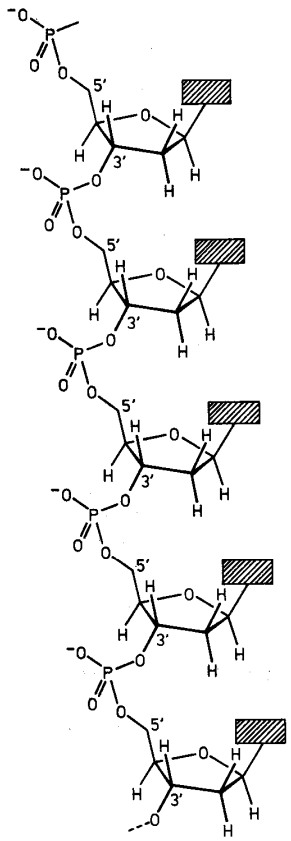


Fig. 20.22. The backbone of deoxyribonucleic acid (DNA) is a long polymer which is made up of alternating phosphate and deoxyribose molecules. They are bound together by an ester bond at the 3′ and 5′ positions of the sugar molecules. The shaded boxes indicate the positions of one of the four possible bases. After Dickerson and Geis

The most important carriers of information are the nucleic acids. DNA (desoxyribo nucleic acids) symbolize the central information storage facility. From there the stored information is passed on with the help of RNA (*ribo nucleic acid*), and is used in the biosynthesis of proteins, see figures taken from Haken and Wolf at left and below.

Fig. 20.24. A schematic representation of DNA. *Left*: a structural model of the deoxyribonucleic acid molecule (DNA). *Right*: some of the structural units of DNA, shown spread out in a two-dimensional manner. The different shapes in the centre indicate how only certain nucleotides can form base pairs by hydrogen bonding

Phosphate			Phosphate
Sugar	Guanine	Cytosine	Sugar
Phosphate			Phosphate
Sugar	Thymine	Adenine	Sugar
Phosphate			Phosphate
Sugar	Adenine	Thymine	Sugar
Phosphate			Phosphate
Sugar	Cytosine	Guanine	Sugar

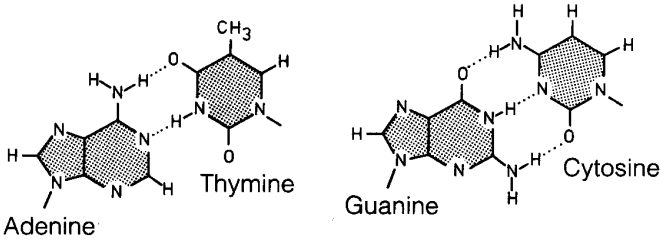


Fig. 20.23. The four bases of DNA, shown joined into pairs by hydrogen bonds

The following text is taken from Haken and Wolf: Molecular Physics and Quantum Chemistry, pages 364-365:

The most important carriers of information for organic life are the nucleic acids, in particular the DNAs, deoxyribonucleic acids. They make up the central information storage units in the nuclei of cells, from which with the help of ribonucleic acids (RNAs) the stored information can be passed on and made use of in the biosynthesis of proteins. The primary contribution which can be made by molecular physicists in the investigation of these fundamental building blocks of life are in the areas of structure determinations and the elucidation of intermolecular binding mechanisms. DNA, deoxyriboneuleic acid, is shown in Figs. 20.22 and 20.23; it is the most important nucleic acid, because it as a rule is used to store and recall genetic information. It is a long-chain polymer made up of molecules of the sugar deoxyribose (the general composition formula of the sugars is $C_x(H_2O)_y$, with x and y between 5 or 6, respectively) together with phosphate groups which alternate with the sugar molecules to form a chain; cf. Fig. 20.22. Each of the individual units is covalently bound to one of the 4 purine or pyrimidine bases, which are symbolized by A (adenine), C (cytosine), G (guanine), and T (thymine); they are all shown in Fig. 20.23. These side groups are thus arranged at regular intervals along the polymer chain. The genetic information is coded in terms of the sequence of the bases along the DNA strand. A group of three bases - a codon - characterizes a particular amino acid, i.e. it gives the prescription to insert that amino acid into a protein chain being synthesized. For coding the 20 different amino acids from which natural proteins are made, there are thus $4^3 = 64$ distinct codons, since 4 different bases can occur in a given position. A second strand, bound to the first strand to form a double helix (by hydrogen bonding between the bases), contains the same information; cf. Fig. 20.24. This information can be called up by messenger RNA, which has a structure similar to that of DNA, and passed on as a recipe for protein biosynthesis; as mentioned, each triplet of bases, i.e. each codon, specifies a particular amino acid that is to be inserted into the protein strand.
The information stored in the codons is transferred to the messenger RNA and brought to the site where protein synthesis is taking place. Figure 20.24 shows again the structure of DNA in a very schematic way.
The contributions of physicists to the elucidation of these mechanisms has been considerable. The most important of them was the determination of the double helix structure of DNA using X-ray diffraction by Crick and Watson in the year 1960. Further significant contributions have been measurements and calculations of the internal motions of biomolecules, of their folding and oscillations, and of their dynamics on being inserted and reformed in living organisms. In this area, there have also been important contributions by means of infra-red and NMR spectroscopies.

6.6 Molecular Functional Units

Many biological processes are based on molecular membranes. These are composed of long chained fatty acid derivatives (lipids) with hydrophilic or hydrophobic end groups (e.g. –OH or $-CH_3$).

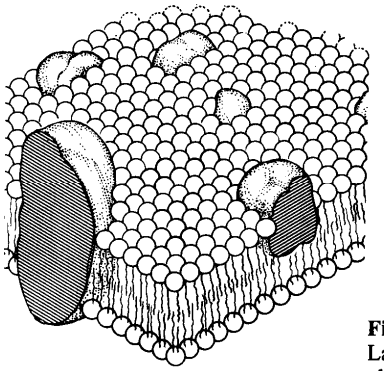


Fig. 20.25. A model of a cell membrane. It consists of a lipid bilayer with embedded protein complexes. The latter can to some extent diffuse between different regions of the membrane and can extend through the membrane on both sides. [After S.J. Singer and G.I. Nicholson, Science **175**, 723 (1972)]

Langmuir-Blodgett layers are models of Molecular membranes. They open many preparative possibilities, and can be examined using many physical methods. Figures taken from Haken and Wolf.

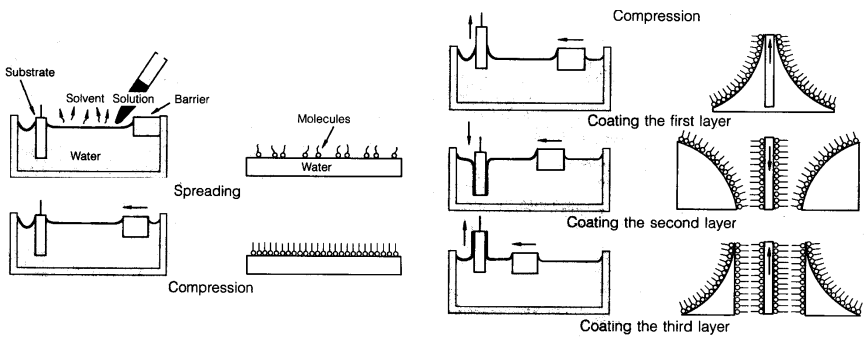


Fig. 20.26. The preparation of Langmuir-Blodgett films. Amphiphilic molecules, that is molecules with a water-soluble and an insoluble end, dissolve two-dimensionally on a water surface, spreading to form a monomolecular layer. They can then be compressed to give a continuous two-dimensional film and can be attached to a substrate by dipping the latter into the water and pulling it out again. This process can be repeated many times. One can thus prepare ordered thin films of a chosen thickness. [see also Nachr. Chem. Techn. **36**, Nr. 10 (1988)]