Introduction

Molecular biophysics at the beginning of the twenty-first century: from ensemble measurements to single-molecule detection

The ideal biophysical method would be capable of measuring atomic positions in molecules in vivo. It would also permit visualisation of the structures that form throughout the course of conformational changes or chemical reactions, regardless of the time scale involved. At present there is no single experimental technique that can yield this information.

A brief history and perspectives

Molecular biology was born with the double-helix model for DNA, which provided a superbly elegant explanation for the storage and transmission mechanisms of genetic information (Fig. 1). The model by J. D. Watson and F. H. C. Crick and supporting fibre diffraction studies by M. H. F Wilkins, A. R. Stokes, and H. R Wilson, and R. Franklin and R. G. Gosling, published in a series of papers in the 25 April, 1953 issue of Nature, marked a major triumph of the physical approach to biology.

The Watson and Crick model was based only in part on data from X-ray fibre diffraction diagrams. The patterns, which demonstrated the presence of a helical structure of constant pitch and diameter, could not provide unequivocal proof for a more precise structural model. One of the ‘genius’ aspects of the discovery was the realisation that A–T and G–C base pairs have identical dimensions; as the rungs of the double-helix ladder, they give rise to a constant diameter and pitch. From a purely ‘diffraction physics’ point of view, a variety of helical models was compatible with the fibre diffraction diagram, and other authors proposed an alternative model for DNA, the so-called ‘side-by-side model’, coupling two single DNA helices. This shows that if molecular biology were to be established, it was important to obtain the structure of biological molecules in more detail than was possible from fibre diffraction. Considering the dimensions involved, about 1 Å (0.1 nm) for the distance between atoms, X-ray crystallography appeared to be the only suitable method. Major obstacles remained to be overcome such as obtaining suitable crystals, coping with the large quantity of data required to describe the positions of all the atoms in a macromolecule, and solving the phase problem.
Introduction

(a) DNA

(b) (c) (d) X-ray fibre diffraction of the B-form of DNA. The figures are facsimiles from the original papers of Watson and Crick (1953) and Franklin and Gosling (1953).

Protein crystals had already been obtained in the 1930s. It was not until 1957, however, that Max Perutz and John Kendrew found a way to solve the crystallographic phase problem by isomorphic substitution using heavy-atom derivatives. This permitted the structure of myoglobin to be solved in sufficient detail to describe how the molecule was folded. The difficulties encountered with protein crystallisation, and the labour intensive nature of the crystallographic study itself (this was before powerful computers and long calculations were essentially performed by 'post-doctoral hands') appeared to doom protein crystallography to providing rare, unique information on the three-dimensional structure of a very few biological macromolecules. Structural molecular biologists, therefore, continued the development and improvement of methods that do not provide atomic resolution but have complementary advantages for the study of macromolecular structures. These methods, at the boundary between thermodynamics and structure, had already played crucial roles in the century before the discovery of the double helix. The discovery of biological macromolecules is itself tightly interwoven with the application of physical concepts and methods to biology (biophysics).

The application of physics to tackle problems in biology is certainly older than its definition as biophysics. The Encyclopædia Britannica suggests that the study of bioluminescence by Athanasius Kircher in the seventeenth century might be considered as one of the first biophysical investigations. Kircher showed that an extract made from fireflies could not be used to light houses. The relation between biology and what would become known as electricity has preoccupied physicists for centuries. Isaac Newton, in the concluding paragraph of his Principia (1687), reflected that ‘... all sensation is excited, and the members of animal bodies move at the command of the will, namely, by the vibrations of this Spirit, mutually propagated along the solid filaments of the nerves, from the outer organs of sense to the brain, and from the brain into the muscles. But these are things that cannot be explained in few words, nor are we furnished with that sufficiency of experiments which is required to an accurate determination and demonstration of the laws by which this electric and elastic Spirit operates.’ One hundred years later, Luigi Galvani and Alessandro Volta performed the experiments on frogs’
legs that would lead to the invention of the electric battery. They also laid the foundations of the science of electrophysiology, even though, because of the excitement caused by the electric battery it was well into the nineteenth century before the study of animal electricity was developed further, notably by Emil Heinrich Du Bois-Reymond. Another nineteenth century branch of biophysics, however, that dealing with diffusion and osmotic pressure in solutions, would later overlap with physical chemistry, and is more directly relevant to the discovery and study of biological macromolecules. The first papers published in Zeitschrift für Physikalische Chemie (1887) were concerned with reactions in solution, because biological processes essentially take place in the aqueous environment inside living cells.

The thermal motion of particles in solution (‘Brownian’ motion) was discovered by Robert Brown (1827). The Abbé Nollet, a professor of experimental physics, first described osmotic pressure in the early nineteenth century from experiments using animal bladder membranes to separate alcohol and water. The further study and naming of the phenomenon is credited to the medical doctor and physiologist René J. H. Dutrochet (1828), who recognised the important implication of osmotic phenomena in living systems and firmly believed that basic biological processes could be explained in terms of physics and chemistry. The theory of osmotic pressure was developed by J. Van’t Hoff (1880). George Gabriel Stokes (middle of the nineteenth century) is best known for his fundamental contributions to the understanding of the laws governing particle motion in a viscous medium, but he also named and worked on the phenomenon of fluorescence. The laws of diffusion under concentration gradients were written down by Adolf Fick (1856), by analogy with the laws governing heat flow. The second half of the nineteenth century also saw the discoveries of flow birefringence by James Clerk Maxwell and of electric birefringence in solutions by John Kerr. Both phenomena depend on the existence of large asymmetric solute particles.

Macromolecules, although large as molecules, are still much smaller than the wavelength of light. They could not be seen through direct observation by using microscopes, which had already shown the existence of cells in biological tissue and of structures within the cells such as the chromosomes (from the Greek, meaning ‘coloured bodies’). From the knowledge gained from experiments on solutions it gradually became apparent that the biochemical activity of proteins, studied by Emil Fischer (1882), is due to discrete macromolecules. In 1908, Jean Perrin applied a theory proposed by Albert Einstein (1905) to determine Avogadro’s number from Brownian motion. The theory of macromolecules is due to Werner Kuhn (1930) after Hermann Staudinger (1920) proposed the concept of macromolecules as discrete entities, rather than colloidal structures made up of smaller molecules. The discovery of X-rays by Wilhelm Conrad Röntgen (1895), and its application to atomic crystallography in the 1910s through the work of Peter Ewald, Max von Laue and William H. and W. Laurence Bragg...
laid the ground work for the observation of atomic structural organisation within macromolecules almost half a century later. Theodor Svedberg (1925) made the first direct ‘observation’ of a protein as a macromolecule of well-defined molar mass by using the analytical centrifuge he had invented. In parallel, the atomic theory of matter became accepted as fact. There was rapid progress in X-ray diffraction and crystallography, electron microscopy and atomic spectroscopy. The novel experimental tools, provided by the new understanding of the interactions between radiation and matter, were carefully honed to meet the challenge of biological structure at the molecular and atomic levels. Physicists, encouraged by the example of Max Delbrück, who chose to study the genetics of a bacteriophage (a bacterial virus) as a tractable model in the 1940s, and Erwin Schrödinger’s influential book What is Life? (1944), which discussed whether or not biological processes could be accounted for by the known laws of physics, turned to biological problems in a strongly active way.

At the beginning of the twenty-first century, biophysics is dominated by two methods, X-ray crystallography and NMR, which play the key role in determining three-dimensional structures of biological macromolecules to high resolution. But even if all the protein structures in different genomes were solved, crucial questions would still remain. What is the structure and dynamics of each macromolecule in the crowded environment of a living cell? How does macromolecular structure change during biological activity? How do macromolecules interact with each other in space and time? These questions can be addressed only by the combined and complementary use of practically all biophysical methods. Mass spectrometry can determine macromolecular masses with astonishing accuracy. Highly sensitive scanning and titration microcalorimetry are applied to determine the thermodynamics of macromolecular folding and stability, and are joined by biosensor techniques in the study of binding interactions. There has been a rebirth of analytical ultracentrifugation, with the advent of new, highly precise and automated instrumentation, and it has joined small-angle X-ray and neutron scattering in the study of macromolecular structure and interactions in solution and the role of hydration. A femtosecond time resolution has been achieved for the probing of fast kinetics by optical spectroscopy. Light microscopy combined with fluorescence probes can locate single molecules inside cells. Scanning force microscopy is determining the profile of macromolecular surfaces and their time-resolved changes. Electron microscopy is approaching close to atomic resolution and is most likely to bridge the gap between single-macromolecule and cellular studies. Neutron spectroscopy is providing information on functional dynamics of proteins within living cells. Synchrotron radiation circular dichroism can access a wider wavelength range vacuum ultraviolet for the study of electronic transitions in the polypeptide backbone.

Up to the late 1970s, biophysics and biochemistry had only dealt with large molecular ensembles for which the laws of thermodynamics are readily applicable. One hundred microlitres of a 1 mg/ml solution of haemoglobin, for example,
contains $10^{18}$ protein molecules; a typical protein crystal contains of the order of $10^{15}$ macromolecules. In their natural environment, however, far fewer molecules are involved in any interaction and exciting new methods have been developed that allow the study of single molecules. Single molecules can now be detected and manipulated with hypersensitive spectroscopic and even mechanical probes such as atomic force microscopy, with which a single macromolecule can even be stretched into novel conformations. Conformational dynamics can be measured by single-molecule fluorescence spectroscopy. Fluorescence resonance energy transfer can measure distances between donor and acceptor pairs in single molecules, in vitro or in living cells. Near-field scanning optical microscopy can identify and provide dynamics information on single molecules in the condensed phase.

The historical development of each of the biophysical methods outlined above is discussed in more detail in the corresponding section of the book.

**Languages and tools**

_Physike_ in Greek is the feminine of _physikos_ meaning _natural_. Physics is the science of observing and describing Nature. When one of the authors (J. Z.) was a student at Edinburgh University, physics was taught in the department of Natural Philosophy. The word _philosophy_, love of wisdom, conveyed quite accurately how the wisdom of the observer is brought to bear in science. The observer plays his role through the tools he uses in his experiments and the language he uses to describe his results. Modern science covers so many diverse areas that it is impossible to master an understanding of all the tools and languages involved. Biophysics students are familiar with the language difficulties of trying to communicate with ‘pure’ physicists, on the one hand, and ‘pure’ biologists, on the other, despite decades of interdisciplinary teaching and research in universities. Rather than bemoaning this fact, we should recognise that it reflects the richness and depth of each discipline, expressed in its own sophisticated language, and developed in its own set of observational tools. Clearly, physics and biology have different languages, but it is important to appreciate that within each discipline also there are different languages. Language influences tool development, which in turn contributes to refine the concepts described by language. Biophysicists have to be fluent in the various languages of physics and biology and be able to translate between them accurately. This is a very difficult and sometimes impossible task, as any good language interpreter can testify, each language having its own specificity and view point.

Biophysics deals, to a large extent, with the structure, dynamics and interactions of biological macromolecules. What are biological macromolecules? Their biological activity is described in the language of biochemistry and molecular spectroscopy; they were discovered through their hydrodynamic and thermodynamic properties; they are visualised by their radiation scattering properties,
Introduction

and their pictures are drawn in beautiful colour as physical particles. To each
language there corresponds a set of tools, the instruments and methods of experi-
mental observation. Progress in probing and understanding biological macro-
molecules has undoubtedly been based on advances in the methods used. Phys-
ical tools capable of ever increasing accuracy and precision require a parallel
development in biochemical tools (often themselves of physical basis, like elec-
trophoresis or chromatography, for example) to provide meaningful samples for
study. The word meaningful is a key word in the previous sentence. It refers to the
relevance of the study with respect to biology (from the Greek bios, life, and logos,
word or reason), i.e. biophysics has the goal of increasing our understanding of
life processes. It should be distinguished from biological physics, which deals
with the properties of biological matter, for example to design nanomachines
based on DNA.

Length and time scales in biology

Biological events occur on a wide range of length and time scales – from the dis-
tance between atoms on the ångström scale to the size of the earth as an ecosys-
tem, from the femtosecond of electronic rearrangements when retinal absorbs
a photon in the first step of vision to the $10^9$ years of evolution. Observation
tools have been developed that are adapted to the different parts of the length
and time scales. The cell represents a central threshold for biological studies
(Fig. 2). With a usual size of the order of 1–10 $\mu$m, cells can be seen in the light

![Fig. 2](http://www.scripps.edu/mb/goodsell/)
microscope. Also, the durations of cellular processes, which are of the order of seconds to minutes can be observed and measured with relative ease. If we imagined diving into a eukaryotic cell through its plasma membrane, we would see other membrane structures that separate distinct compartments like the nucleus or mitochondria, large macromolecular assemblies such as chromatin, ribosomes, chaperone molecular machines or multienzyme complexes. Looking for progressively smaller structures we would find RNA and protein molecules, then peptides and other small molecules, water molecules and ions, and finally the atoms that make them up (Fig. 3).

The smaller the length, the shorter the time, the heavier is the implication of sophisticated physical instrumentation and methods for their experimental observation.

The femtosecond (10\textsuperscript{-15} s) is the shortest time of interest in molecular biology; it corresponds to the time taken by electronic reorganization in the light sensitive molecule, retinal, upon absorption of a photon, in the first step in vision. Time intervals of this order can be measured by laser spectroscopy (the distance covered by light in 1 fs is 3 \times 10\textsuperscript{-7} m, or 300 nm, about one half the wavelength of visible light). Thermal fluctuations are in the picosecond (10\textsuperscript{-12} s) range; DNA unfolds in microseconds; enzyme catalysis rates are of the order of 1000 reactions per second; protein synthesis takes place in seconds etc. The longest time of interest in molecular biology is, in fact, geological time, corresponding to the more than one thousand million years of molecular evolution (Fig. 4).

The structure–function hypothesis

This book describes the application of classical and advanced physical methods to observe biological structure, dynamics and interactions at the molecular level. Intensive research since the 1950s has emphasized the fundamental importance of biological activity at this level. The structure–function hypothesis is the foundation of molecular biology. One of its implications is that if a protein exists today in an organism it is because it fulfils a certain biological function and its ‘structure’ has been selected by evolution. The discovery and study of nucleic acids and proteins as macromolecules with well-defined structures has allowed an unprecedented understanding of processes such as the storage and transmission of genetic information, the regulation of gene expression, enzyme catalysis, immune response or signal transduction. In parallel, it became apparent that we could act on biological processes by acting on macromolecular structures and powerful tools were developed not only to further fundamental scientific understanding but also to apply this knowledge in biotechnology or in drug design pharmacology.

The concept of ‘structure’ should be understood in the broadest sense. The three-dimensional organisation of a protein is not rigid but can adapt to its ligands according to the hypothesis of ‘configurational adaptivity’ or ‘induced fit’. Also,
many proteins have been found that display a highly flexible random-coil conformation under physiological conditions. An intrinsically disordered protein could adopt a permanent structure through binding, but there are cases of proteins with intrinsic disorder that are biologically active while remaining disordered. A large proportion of gene sequences appear to code for long amino acid stretches that are likely either to be unfolded in solution or to adopt non-globular structures of unknown conformation.

Events taking place on the ångström and picosecond scales have profound consequences for life processes over the entire range of length and time scales – from the length and time associated with a cell, via those associated with an organism to those associated with the relation between an organism and its environment. The development of high-throughput techniques for whole genome sequencing, for the analysis of genomic information (bioinformatics), for the identification of all the proteins present in a cell (functional proteomics), for determining how this population responds to external conditions (dynamic proteomics) and for protein structure determination (structural genomics) has opened up a new era in molecular biology whose revolutionary impact still remains to be assessed.

Biological macromolecule structures usually appear in pictures as static structures. A more precise definition would be ‘ensemble and time-averaged’ structures. The atoms in a macromolecular structure are maintained at their average positions by a balance of forces. Under the influence of thermal energy, the atoms move about these positions. Dynamics, from the Greek *dynamis* meaning strength, pertains to forces. Structure and motions result from forces. It is common usage in biophysics, however, to separate structure from dynamics, considering the first as referring to the length scale (i.e. to the time-averaged configuration) and the second as referring to the time scale (i.e. to energy and fluctuations). The separation into two separate concepts is validated by the fact that the methods used to study structure and dynamics are usually quite separate and specialised. Modern experiments, however, often address both an average structure and how it changes with time.

**Complementarity of physical methods**

We know of the existence of macromolecules only through the methods with which they are observed. No single method, however, provides all the information required on a macromolecule and its interactions. Each method gives a different view of the system in space and time: the methods are complementary.

Biological macromolecules take up their active structures only in a suitable solvent environment. The forces that stabilise them are weak forces (of the order of $kT$, where $k$ is Boltzmann’s constant and $T$ is absolute temperature), which arise in part from interactions with the solvent. The study of biological macromolecules, therefore, cannot be separated from the study of their aqueous solutions. The
Thermodynamics

It is a result of classical thermodynamics that many properties of solutions, such as an increase in boiling point, freezing point depression, and osmotic pressure, depend on the number concentration of the solute. At constant mass concentration, therefore, these thermodynamics parameters vary sensitively with the molecular mass of the solute. Thus, for example, macromolecular masses and interactions have been determined from osmotic pressure measurements.

Macromolecular folding itself and the stabilisation of active biological structures follow strict thermodynamics rules in which interactions with solvent play a determinant role. Sensitive calorimetric measurements of heat capacity as a
Introduction

function of temperature showed very clearly that stabilisation free energy presents a maximum at a temperature close to the physiological temperature, the stability of the folded particle decreasing at lower as well as higher temperatures. The interpretation is the following. The behaviour of the chain surrounded by solvent is much more complex than if it were in a vacuum. Enthalpy may rise, decrease or even not change upon folding, because bonds can be made equally well within the macromolecule and between the chain and solvent components. Similarly for entropy, the loss of chain configuration freedom upon folding may be more than compensated for by a loss of degrees of freedom for the solvent molecules around the unfolded chain, for example through the exposure of apolar groups to water molecules. A water molecule in bulk has the freedom to form hydrogen bonds with partners in all directions. Apolar groups cannot form hydrogen bonds, so that water molecules in their vicinity lose some of their bonding possibilities; their entropy is decreased.

In a protein solution, the heat capacity is strongly dominated by the water, and that of the macromolecules represents a very small part of the measured total. High-precision microcalorimeters were built to allow experiments on protein solutions to be performed. Nevertheless, early calorimetric studies on biological macromolecules concentrated on relatively large effects such as sharp transitions as a function of temperature. They led to a fundamental understanding of the energetics of protein folding. There are now important modern developments in the field. Very sensitive nanocalorimeters have been developed as well as analysis programs to treat the thermodynamics information and relate it to structural data. The energetics of intramolecular conformational changes, of complex formation and of interactions between partner molecules can now be explored in detail for proteins and nucleic acids. We should recall, however, that calorimetry (like all thermodynamics-based methods) provides measurements of an ensemble average over a very large number of particles (typically of the order of 10^{15}), even if results are usually illustrated in a simple way in terms of changes occurring in one particle.

Hydrodynamics

The first hints of the existence of biological macromolecules as discrete particles came from observations of their hydrodynamic behaviour. The language of macromolecular hydrodynamics is the language of fluid dynamics in the special regime of low Reynolds’ numbers. The Reynolds’ number in hydrodynamics is a dimensionless parameter that expresses the relative magnitudes of inertial and viscous forces on a body moving through a fluid. Bodies with the same Reynolds, number display the same hydrodynamics behaviour. Because of this, it is possible, for example, to determine the behaviour of an airplane wing from wind-tunnel studies on a small-scale model. The Reynolds’ numbers of a small fish and a whale are 10^5 and 10^9, respectively.