The Physical Behaviour of Macro-molecules with BIOLOGICAL FUNCTIONS

S. P. SPRAGG

THE PHYSICAL BEHAVIOUR OF MACROMOLECULES WITH BIOLOGICAL FUNCTIONS

S. P. Spragg
Department of Chemistry,
University of Birmingham

JOHN WILEY & SONS

Chichester · New York · Brisbane · Toronto

British Library Cataloguing in Publication Data:

Spragg, S. P.
The physical behaviour of macromolecules with biological functions. —(Monographs in molecular biophysics and biochemistry).
1. Macromolecules—Measurement
I. Title II. Series
574.1'924 QP801.P64 80-40280

ISBN 0 471 27784 3

Photosetting by Thomson Press (India) Limited, New Delhi
Printed and bound in England by The Pitman Press Limited, Bath, Avon

Series preface

The term 'molecular biophysics' means different things to different people. To some the subject matter is covered by X-ray structure studies of molecules of biological interest, while others believe that the application of one of several different physical techniques to the study of biological systems would come under that title. Personally, I consider Karl Pearson's definition most appropriate: 'Biophysics is the study of the physical laws which govern the behaviour of biological systems'. Obviously, physical techniques are usually required to study physical laws, but the emphasis—if one considers biophysics as part of biological science—must be on the 'behaviour of a biological system'. All too often techniques have ended up as 'methods in search of a problem'. One could remark, somewhat cynically, that while most techniques have a limited life, fundamental biological problems never do get solved. This is probably true because people ask open-ended questions like 'what is the mechanism of enzyme action?' or 'how does muscle contraction work?'

The complexity of biological systems requires specific questions at each level of organization. Asking a good question is surely more than half the battle of doing good research. I am afraid it is difficult to resist using the opportunity of writing an introduction like this, to pontificate on one's own philosophy of research. I believe that the most difficult balance to hold in a scientist's division of labour is between the acquisition of broad vistas and concentration on a specialized and specific question at any one moment in time. No single series of experiments can tell one how rhodopsin transmits information about the absorption of a photon to the brain. However, the specialized experiments which one carries out with rhodopsin are only likely to contribute to our knowledge about visual perception if they are planned in the context of many other aspects of this field of endeavour. The problem with technique oriented biophysics is that, while it often concentrates on a particular question with each system, it does not permit sufficiently wide acquaintance with the behaviour of the system.

The above comments must not be taken to mean that I believe there is no other way. It merely points to the fact that in spite of the exciting results obtained by crystallographers or NMR spectroscopists, for instance, they have not solved the problem of 'how enzymes work'. Ideally one should ask a question and then look for a technique which can be used to answer it, rather than have a technique and look for a question it can answer. It is with this in mind that the publishers have

asked a number of potential authors to write on biophysical questions, highlighting problems and surveying the range of approaches towards obtaining solutions for them. It is difficult to be sure at this stage which ones of those invited will produce a manuscript. The series will produce some volumes which attempt to answer questions about problems which have a wide application to the behaviour of biological systems. While this may appear to contradict my emphasis on concentration on one system at the time, such volumes will underline the importance of using a wide range of approaches to study those properties of molecules and their interactions which give information about the 'physical laws which govern the behaviour of biological systems'. One of the most important of these problems is the way macromolecules, like proteins and nucleic acids, behave in the 'real' concentrated solutions containing many other small ions which constitute their natural environment. The present volume presents a serious effort to treat this aspect of molecular biophysics. Other volumes, which are in an advanced state of preparation, will be concerned with the thermodynamics of ligand binding and other biochemical equilibria (Edsall and Gutfreund) and with the physical and chemical mechanisms of the action of calcium ions in their ubiquitous physiological functions (Campbell).

H. GUTFREUND

A note on symbols

$\mathcal{N}$ Avogadro's number (number of molecules in one mole).

In general, all other symbols are defined as they appear but, in order to maintain usage according to convention for some symbols, confusion is possible between chapters. An attempt has been made to maintain conformity throughout and the exceptions are listed below.

f refers to activity coefficient in Chapter 1 and occasionally in Chapter 2; this symbol is also used to describe the molecular frictional coefficient in Chapters 2 and 3.

M_2 the subscript 2 refers to the property of the macromolecule in a ternary mixture containing (1) solvent, (2) macromolecule and (3) buffer ions; in Chapter 4, it refers to the property of the dimer.

Contents

Introduction

Experience over years of teaching and research has shown that one major challenge to workers studying the behaviour of macromolecules is to provide and utilize adequate explanations of the concentration dependence of their hydrodynamic and thermodynamic properties. Undergraduates and researchers are always pointing out that most thermodynamic and hydrodynamic concepts are best applied when working at infinite dilution and this is hardly a realistic experimental condition. Deviations between theory and experiment caused by working at finite concentrations are corrected through empirical or, at their best, semi-empirical coefficients. This approach suggests that these deviations are an annoyance which must be tolerated but are not included in theories. Possibly a more realistic approach would be to enquire whether these deviations from ideality, which are small for globular proteins, carry useful information about the state of the molecule in solution and, if so, is this reflected by biological properties of the macromolecule?

The standard way of dealing with these deviations is through virial coefficients, but transforming these into molecular properties is only covered adequately for neutral molecules using the concept of excluded volume. Furthermore, the coefficients are only suitable for discussing thermodynamic non-ideality but not hydrodynamic non-ideality, so they cannot be applied to discussion of concentration dependence of sedimentation coefficients nor to describe the complete deviation of diffusion coefficients. In many respects, the behaviour of macromolecules is very dependent on concentration, and this means that the full understanding of the deviations is necessary for two applications: (*a*) the exact correction of experimental observations on macromolecules when working at normal concentrations can only be made by allowing for the relatively large excluded volume of these molecules compared with, say, sucrose; and (*b*) a full understanding of their physico-chemical behaviour at finite concentrations must include concentration dependence if these data are to be used for descriptions of biological activity. Of these two reasons, the second is possibly the most fundamental, since it seems likely that the forces that cause the concentration dependence will be reflected in the biological activity. Even for this early part of the book, it seems necessary to justify this statement. One of the major problems in biology is to describe detailed processes by which cells or molecules recognize likeness (or foreignness): macroscopic experiences of this activity are many and

range in complexity from rejection of grafts or organs in medicine through specificity of antigen–antibody reactions to purification by crystallization. Qualitative statements about recognition are difficult to extend down to the molecular level. Furthermore, since the expressions of molecular recognition normally occur through non-covalent interactions, it seems reasonable to explore the weaker physical forces between molecules. Thus, in solutions, this means understanding how the detailed structural features of the molecule modify the organization and activity of the solvent, as the frame of reference moves from the solute molecule out into the bulk solvent. Explanations of solute–solvent interactions are generally based on relative energies, and present trends are towards using these energies to support models given at the molecular level. This movement from rigorous thermodynamic interpretation of macroscopic observations to views more concerned with describing molecular motions is a continuing development in biochemistry. Studying diffusional processes at the molecular level is one part of this development. It was with these views in mind that this book was written.

Many authoritative texts are available which present accounts of the basic theories describing the solution properties of macromolecules; thus, detailed theory is often not presented in this text. Of the many books which could form 'necessary reading' are the one by Tanford (1961) and the recent one by Marshall (1978). One disappointing aspect of textbooks for students of all levels is the division that seems to exist between giving detailed descriptions of theories and outlining the flaws in experimental work which lead to frustration when testing or implementing theory. For this reason, I have tried to include examples of situations where inexperience or insufficiencies in the theory have led to false interpretations; it should be added that later work has sorted out these problems, otherwise we would not know that the original ideas carried flaws (normally the experimental data are reliable). These examples are obviously limited to problems that I have met in my own research, and consequently are not a full coverage for all techniques, but I hope the message conveyed is clear—always assume Murphy's law applies to experiments, i.e. that if results can be ambiguous or fogged by uncertainties, then this will occur!

Assuming experiments do give reliable results on the activities of macromolecules in solution, then what use is this to describing biological activities; in other words, what is the overall message of this book? Basically, it is directed to two categories of scientist: to biochemists, the message is not to underestimate the role of long-range molecular forces when describing the role of macromolecules in biology; and to biophysical chemists, it is to continue to provide exact models for molecular interpretations of concentration dependence. It became clear while compiling this book that a frontal attack on these forces and their relationship to concentration dependence was not a useful approach. Their discussion must be tied to the theoretical and experimental descriptions of the actual parameters involved. For this reason, these are discussed at some length. In fact, fewer words are devoted to discussing concentration dependence than to painting the background upon which the effects work. Nevertheless, the evidence shows that

forces between macromolecules range over much larger distances than often thought; this evidence comes from the consideration of the dependences of sedimentation and diffusion coefficients. It is well known that the energies are weak (but may be larger than thermal energy), but it is less well known that the drop in energy with distance between macromolecules is relatively gradual. These may be interesting observations to physical chemists, but it is reasonable for biochemists to question whether this knowledge is important to the biological activity of macromolecules.

In answering this question, it is necessary to recall one fact that continually impresses physical scientists, namely the small range of temperatures that a biological system can tolerate. In itself, this observation can suggest either that the system is finely tuned for activity or that weak energies play important roles in regulating the system: both would mean that small changes in energy supplied to the system would produce effects far outside normal physical experiences. Observations show that radical shifts in the chemical status of natural macromolecules are produced by relatively small energy differences. For example, tetrameric haemoglobin readily dissociates into two dimers by dilution, while deoxyhaemoglobin only dissociates marginally even in excessively low dilutions, yet the overall difference in energy between the two contrasting states is only about $20\,kJ\,mol^{-1}$ (or about eight times thermal energy at normal temperatures). Other equally surprising effects of low energy differences are commonplace with biological macromolecules, and Blow (1974) has suggested that energies that are considered reasonable are about $20\,kJ\,mol^{-1}$ whereas those that are strong are about $80\,kJ\,mol^{-1}$ (this reference was to the relatively high energy of interaction between trypsin and antitrypsin agents).

These lines of reasoning can be extended to structural work where large biological changes are often reflected by marginal changes in molecular structure: Perutz (1975) found that the structure of a thermally stable version of haemoglobin had only ten residues different in the β-chain from the common haemoglobin A. This change induced a few extra non-polar interactions which in turn produced no more than $5\,kJ\,mol^{-1}$ extra energy for the thermostable haemoglobin. These findings all indicate that interactions through forces of relatively low energies are biologically important, and the concentration dependence of hydrodynamic parameters may be one way of measuring some of these low interactive energies. In this respect, it is interesting to find that macromolecular structures determined by crystallography are in themselves not leading now to many new concepts of activity of natural macromolecules (summarized by Mathews, 1977). Instead, attempts are being made to relate these structures to proportions of the surface exposed to the solvent (Lee and Richards, 1971; Chotia, 1974; Wodak and Janin, 1978) in attempts to incorporate concepts such as hydrophobicity (Kauzmann, 1959) into the structure. Does this mean that hydrodynamic and crystallographic studies are converging through the acceptance that a major factor in determining biological activity is the long-range weak forces of solute–solvent–solute interactions? Certainly, Anderson (1976a, b) would agree with this view, since he suggested that recognition of biological

foreignness (e.g. an antibody recognizing an antigen molecule in a sea of other proteins) occurs through paratactic (or complementary) surface interactions, so pointing to surface tension as a determining influence (surface tension is a relatively weak force produced by disturbed solvent–solvent interactions at a phase interface).

Attempts to measure the interactions between macromolecules in solution have in the past been made through macroscopic observations. Then, using rigorous thermodynamic theory, attempts have been made to extrapolate down to the molecular level. This path is difficult and the models far from unique; hence, the move to measuring molecular motions through more direct procedures. One of these procedures involves the measurement and molecular interpretation of diffusion using the technique of quasi-elastic light scattering. With this method, it is possible to measure diffusion coefficients relatively accurately and quickly, so making it possible to study routinely the effects that small changes in molecular structure have on the motions of the molecules. Absolute values of the coefficients are possibly less important than the differences produced by varying treatments or changing concentrations. It is this latter phenomenon that provides hitherto untapped source of information for measuring the weak energies of interaction between molecules in solution. The complex theories that relate molecular models to the measurements are at present limited and are unable to cover many of the real situations pertinent to biology; nevertheless, these will improve. In fact, the situation could be likened to X-ray diffraction, which, 20 years ago, was an unknown quantity to most biochemists, but is now accepted without question, even at the undergraduate level. It is for these reasons that the discussions on diffusion (Chapter 3) are longer than would normally be expected and the description of quasi-elastic light scattering is given in some detail.

Only the future will decide whether these conjectures are borne out by experimental evidence, but it is a fact that most biologically active proteins are formed from oligomeric complexes containing several individual chains in the complex. The forces holding these chains together vary in strength from strong (but less strong than covalent forces) to relatively weak, and the view that even the weaker forces that occur between the oligomeric macromolecules moving free in solution is really an extension of these forces.

Summary

An attempt has been made to present descriptions of physical ideas with the minimum of mathematical derivaiton. With the exceptions of Sections 1.4.3, 3.3.2 and 3.4.2, only a moderate knowledge of mathematics and physics is required to appreciate the concepts. The three sections mentioned could be omitted from the first reading, but in any case a full understanding of the material in these sections will need supplementing through the reviews and books quoted in the text. The text contains sufficiently detailed references for workers to get the details if they so wish. This approach was chosen in order to present ideas to biochemists, who are

familiar with handling proteins and nucleic acids and who can pose detailed biological questions, but who do not wish to survey the literature, which covers physico-chemical ideas, in order to examine these ideas. Thus, Chapter 1 provides a summary of many of the physico-chemical concepts used later in the book. Concepts concerning solute–H_2O interactions, virial coefficients, equilibrium constants, and origins of molecular forces are discussed at a variety of levels.

In Chapter 2 are given details of methods used to weigh macromolecules. The discussion covers experimental interpretations and precautions, concentration dependence of sedimentation coefficients, and relative mass. The problems raised by heterogeneity and solvation are discussed through results which were originally misinterpreted because of insufficient information.

In Chapter 3 is given a detailed discussion of macromolecular behaviour at the molecular level. This is followed by describing ideas that lead to discussion on solvated shapes and flexing of macromolecules. The main part of this chapter is concerned with translational and rotational movements of mass as measured by diffusion. It is through interpretations of these movements in a solvent that one can begin to recognize the importance of solute–solvent interaction on the behaviour of macromolecules. The experimental realization of these studies is through diffusion, and for this reason methods of measurement are discussed, paying particular attention to the role played by quasi-elastic light scattering in this work. Other methods such as dielectric relaxation are briefly discussed in order to maintain continuity with other texts.

Finally, Chapter 4 gives a survey of the specialized form of concentration dependence which can be described best through direct macromolecular association–dissociation reactions. The sources of errors and misinterpretations of earlier work are given for a few cases, together with a description of the influence that the strengths of interaction have on the measurements for proteins: these are particularly 'sticky' molecules.

I should finish this summary by stating that I have tried to avoid the propagation of misinterpretations, but I plead human frailty if I have committed the same sin of logic that befell a well known cartoon character. In the first picture of the series is shown a fluttering object in the distance and the character saying, 'Look at that butterfly, it has flown all the way from Brazil'. The next picture showed the companion examining the object and saying, 'This is no butterfly, it is a leaf'. The final picture showed the character saying, 'How did a leaf fly all the way from Brazil?' I feel sure most readers will agree that it is very easy to be funnelled into this type of argument when details obscure the true facts.

Acknowledgements

Extensive talks and discussions with many friends have been incorporated in the text and it would be difficult to list all these people. All I can do is to thank them and apologize to them if they feel I have incorporated their ideas in this text without due recognition: the subconscious is a powerful but forgetful force. There are, however, five people whom I want to thank specially: Professor G. A. Gilbert

and Mrs L. Gilbert, who have suffered my inadequacies stoically over many years as well as acting as very intelligent sounding boards; Dr N. G. Anderson, for providing persuasive logic to clear many involved biological concepts; Dr J. M. Creeth, for carefully reading and criticizing this work — he offered many valuable suggestions and guided my reasoning in many places; finally, Professor H. Gutfreund, who has criticized, cajoled, and encouraged me to put some ideas down on paper. I have to thank NATO for the generous financial assistance which allowed me to work with Dr N. G. Anderson while my thoughts on solute interactions were related to biological problems. I also wish to thank the editors and publishers of the following journals for letting me incorporate various figures in the text: *Analytical Biochemistry*, *Biochemistry*, *Biophysical Journal*, *Biophysical Chemistry*, *Biopolymers*, *Chemical Society*, *Journal of Biological Chemistry*, *Journal of Molecular Biology*, *Journal of Physical Chemistry*, *Methods in Enzymology*, and *Nature*, *New Biology*.

CHAPTER 1

Setting the scene

Biological activity of macromolecules is normally associated with a solvated molecule either in suspension or in true solution. Thus, if one is to describe biological activity in physico-chemical terms, the interaction of biological macromolecules with their aqueous environment becomes a major determinant. The steps between carrying out an experiment and realizing from the results useful information on biological activity are often long and tortuous, involving the interpretation of complex interactions between all the molecules; sometimes it appears that the details cloud the overall picture. If one is to maintain clarity, then a good starting point would be from a discussion of the interactions of polyatomic molecules in solution. A full and detailed description of these interactions cannot be made for most molecules because this would involve complete formulation of wave equations using the rules of quantum mechanics and statistical thermodynamics. At their best, our physico-chemical descriptions tend to stop some distance from this and theoretical discussions falter after interpreting the macroscopic thermodynamic function. The strength (and weakness) of thermodynamics lies in its independence of molecular models, and its rigorous application for interpreting macroscopic results is inarguable. While accepting this limitation, from thermodynamics have flowed many useful models which provide interpretations of the dynamic activities of molecules; these have been constructed from solid models using coordinates determined by X-ray crystallography. These models rarely allow extrapolation to the quantized energy levels measured by spectroscopy. This chapter cannot do justice to the wide range of detailed theory which links spectroscopy with thermodynamics or X-ray crystallography with biological activity, but it attempts to summarize some of the components which are relevant to understanding the interactions between a macromolecular solute with the excess solvent molecules.

1.1 Aqueous solutions

First, one must consider the physical characteristics of water. Natural macromolecules are dissolved in water in concentrations which, in the biological cell, span several orders of magnitude. Many enzymes are present in relatively low average concentrations and these can be treated as very dilute ideal solutions, but others are present in concentrations which are far removed from ideal (for example,

haemoglobin at 30–40% w/v, and muscle filaments and DNA in the nucleus where it is doubtful whether the term 'solution' can be applied). Throughout this range of concentrations, the introduction of the macromolecules into water produces marked changes in its macroscopic properties. Thermodynamic arguments show that these changes are reflected in the activities of the solutes. Normally, the liquid phase of water is considered to consist of monomers and polymers interacting through the well known hydrogen bond. Hydrogen bonding raises the boiling and melting points of water when compared with similar compounds formed from elements in the same group of the periodic table (e.g. H_2S). Furthermore, the fact that the size of the water polymers is not constant, but is a constantly changing dynamic equilibrium, means that any description of water must be through statistical averages.

The iceberg concept and hydrophobicity

Our basic ideas on the properties of aqueous solutions come from studying the effect that small solutes have on the properties of water. For example, introducing into water a solute which is not capable of forming hydrogen bonds (e.g. CH_4) excludes water from the volume occupied by the solute. This induces shells of organized water molecules arranged concentrically around the solute: the number of shells is ill defined, but rehydration studies suggest the existence of primary, secondary, tertiary, and more shells around all natural macromolecules dissolved in water (Texter, 1978). In colloid chemistry, these well defined lower layers are known as the Stern layer and the more diffuse layers the Gouy layers. Thermodynamically, this would be considered an entropy-based contribution to the energy, because the number of degrees of freedom of water in the region of methane molecules has been reduced. The effect is to produce a region where spatial reorganization of the hydrogen bonds of water is forbidden because of the presence of the methane. Instead, bonding can only occur around the surface of the methane to produce a relatively stable shell of water arranged in a configuration similar to that found in some forms of ice. Restricting the movement of water in this region profoundly changes its properties, producing an increased dielectric constant than in the bulk phase. This simple picture of the influence that a neutral molecule has on the structure of water was first proposed by Frank and Evans (1945) and, despite its age, the simplicity of the model has attracted many theorists who have used it to describe the activity of aqueous solutions. Basically, it provides an explanation of the negative enthalpy and entropy that occurs when a paraffin is mixed with water. Further, it fits the observation that dissolving paraffins in water is accompanied by a contraction of volume, which Nemethy and Scheraga (1962) pointed out was consistent with packing water in the large vacancies formed in ice under low external pressures (several well defined ice structures can be formed by varying the external pressures during crystallization). Simple extensions of this model to include solutions of biological macromolecules cannot be explored in detail here, but it is possible to discuss the factors and conclusions which follow from the model.