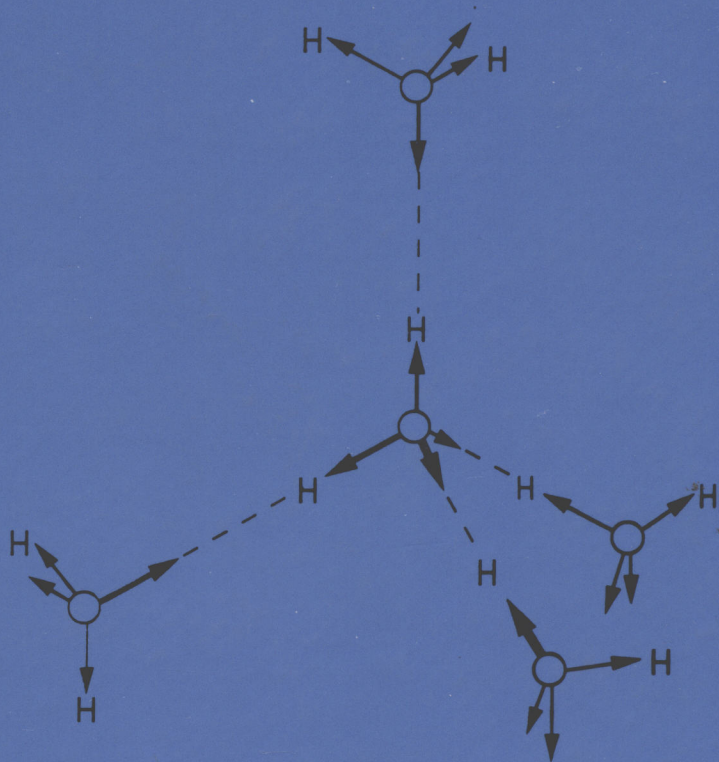# *Statistical Thermodynamics for Chemists and Biochemists*

Arieh Ben-Naim

Statistical Thermodynamics for Chemists and Biochemists

Arieh Ben-Naim

Department of Physical Chemistry
The Hebrew University of Jerusalem
Jerusalem, Israel

Plenum Press • *New York and London*

Printed in the United States of America

To Merlie

Preface

This book was planned and written with one central goal in mind: to demonstrate that statistical thermodynamics can be used successfully by a broad group of scientists, ranging from chemists through biochemists to biologists, who are not and do not intend to become specialists in statistical thermodynamics. The book is addressed mainly to graduate students and research scientists interested in designing experiments the results of which may be interpreted at the molecular level, or in interpreting such experimental results. It is not addressed to those who intend to practice statistical thermodynamics *per se*.

With this goal in mind, I have expended a great deal of effort to make the book clear, readable, and, I hope, enjoyable. This does not necessarily mean that the book as a whole is easy to read. The first four chapters are very detailed. The last four become progressively more difficult to read, for several reasons. First, presuming that the reader has already acquired familiarity with the methods and arguments presented in the first part, I felt that similar arguments could be skipped later on, leaving the details to be filled in by the reader. Second, the systems themselves become progressively more complicated as we proceed toward the last chapter. Thus, mixtures of liquids are more complicated than pure liquids, aqueous solutions are much more complicated than simple mixtures, and aqueous systems containing biomolecules are far more complicated than simple aqueous solutions. As a rule, I tried to develop in detail only those topics that are new, leaving to the reader any details similar to those encountered in previous chapters.

Chapter 1 presents a brief introduction to statistical thermodynamics. Here the basic rules of the game are summarized and some simple results pertaining to ideal gases are presented. The reader is presumed to be familiar with the basic elements of statistical thermodynamics and classical thermodynamics.

Chapter 2 contains several applications of these tools to very simple systems. Except for section 2.10, the material presented here is contained in most standard introductory textbooks in statistical thermodynamics. Section 2.10 is a detailed treatment of a chemical equilibrium affected by the adsorption of a ligand. The results of this section are applied mainly in Chapter 3, but some more general conclusions also appear in later chapters such as 5, 7, and 8.

Chapter 3 deals with slightly more complex systems, e.g., proteins having a few adsorption sites. The proteins are still considered independent, but the ligands on the sites may be dependent or independent. Here the tools of statistical thermodynamics are applied to investigate the behavior of various adsorption models that are presumed to mimic systems of interest in biology.

In these relatively simple solvable models we also introduce several concepts which normally appear in the context of the theory of liquids, such as the analogue of the solvation process, the pair correlation function and potential of average force, triplet

correlations, and the nonadditivity of the triplet potential of average force. All these can be studied exactly within the framework of these models, in terms of the molecular parameters of the models. Particular attention is devoted to elaborating on the various definitions, origins, and manifestations of the cooperativity effects. Gaining familiarity with these concepts within the simplest models should help the reader to understand them when they reappear within the context of the theory of liquids.

Chapter 4 is devoted to one-dimensional (1-D) systems. Although none of the models treated here represents a real system, the study of 1-D models is very useful in gaining insight into various phenomena that do occur in reality, such as the helix–coil transition, phase separation, and the temperature of maximum density of liquid water. These are all real phenomena which may be mimicked by extremely simple and artificial models. The main reward of studying these models comes from their solvability. Section 7.1 treats the helix–coil transition theory as a classic example of the application of the 1-D techniques to solve a problem that arises in physical biochemistry. In reality, however, the helix–coil transition occurs in aqueous solution. The solvent might affect the process to the extent that the "vacuum theory" may not be relevant to the actual process taking place in aqueous solution. (This aspect of the problem is deferred to Chapter 8.)

In Chapter 5 we begin with the theory of the liquid state. Except for section 5.11, we do not survey the various theories suggested for the liquid state. Instead, we focus on the fundamental concepts of molecular distribution functions, their properties, and their relation to thermodynamic quantities.

Chapter 6 is the extension of Chapter 5 to include mixtures of two or more liquids. The most important concepts here are ideal behavior and small deviations from it. Most of the treatment is based on the Kirkwood–Buff theory of solutions. The derivation and a sample application of this powerful theory are presented in detail. We also present the elements of the McMillan–Mayer theory, which is more limited in application. Its main result is the expansion of the osmotic pressure in power series in the solute density. The most useful part of this expansion is the first-order deviation from ideal dilute behavior, a result that may also be obtained from the Kirkwood–Buff theory.

Chapters 5 and 6 may be viewed as introductory to Chapters 7 and 8, which deal with the more complex and more important aqueous solutions. Chapter 7 is devoted to pure liquid water and dilute aqueous solutions of simple solutes. There is a vast literature dealing with theoretical and experimental aspects of these systems. Only the minimum requirements for understanding the outstanding properties of this liquid and its solution are presented here. The emphasis is not on surveying the various theoretical approaches, but on fundamental concepts such as solvation, the structure of water, structural changes induced by a solute, hydrophobic and hydrophilic interactions, and the like. All of these concepts are used to treat the more complicated systems in Chapter 8.

Chapter 8 is the heart of the entire book. Because of the extreme complexity of the systems treated here, we cannot expect any exact solution to any problem. However, statistical thermodynamics can contribute much toward interpreting experimental results, suggesting new experiments, and correlating various aspects of the behavior of these systems.

Some of the "vacuum theories" treated earlier in the book are repeated in Chapter 8. Examples are chemical equilibrium, allosteric phenomena, and helix–coil transition. The general procedure to transform a "vacuum theory" into a "solution theory" is developed. Then we emphasize possible large solvent effects that can significantly alter the "vacuum theory," especially when the solvent is water. A detailed account of the thermodynamics of protein folding and protein association is also presented.

The style of the book is mainly didactic, emphasizing the methods rather than elaborating on specific examples. Occasionally, some numerical examples are presented in the form of a table or a figure. These are included as illustrations only, to give an idea of the order of magnitude of certain quantities. The numerical values are cited in the same units as they appeared in the literature. No attempt was made to use a unified set of units throughout the book.

In covering a wide range of topics, it is inevitable to run out of letters to denote different quantities. Thus, G is used for Gibbs energy throughout the book, but it is also used in the Kirkwood–Buff integrals in the form G_{ij}, as well as being a subscript to denote a gaseous ligand. Appropriate warning comments are made whenever confusion on the part of the reader would otherwise be anticipated.

The criterion used to choose the topics covered in this book was their usefulness in application to problems in chemistry and biochemistry. Thus cluster expansion methods for a real gas, although very useful for the development of the theory of real gases *per se*, was judged not useful except for the second virial coefficient. Similarly, the statistical mechanical extensions of the theory of ionic solutions beyond the Debye–Hückel limiting law were judged not useful in actual applications. Some important topics may have been missed either because of my lack of familiarity with them or because I failed to appreciate their potential usefulness. I would be grateful to receive comments or criticism from readers on this matter or on any other aspect of this book.

Thanks are due to Drs. D. Kramer, G. Haran, R. Mazo, and M. Mezei for reading various parts of the manuscript and offering helpful comments and suggestions. I am particularly grateful to Ms. Merlie Figura for typing the entire manuscript; without her gracious help this book could never have been readied for publication.

<div style="text-align: right">Arieh Ben-Naim</div>

Jerusalem, Israel

Contents

The Fundamental Tools

1.1. INTRODUCTION

Statistical thermodynamics (ST) is a mathematical tool that bridges the gap between the properties of individual molecules and the macroscopic thermodynamic properties of bulk matter.

A microscopic description of a system of N spherical particles requires the specification of $3N$ coordinates and $3N$ momenta. If N is of the order of 10^{23}, such a detailed description is clearly impractical.

On the other hand, the same system of N particles at equilibrium can be described by only a few thermodynamic parameters, such as volume, temperature, and pressure. Such a drastic reduction in the number of parameters is achieved by averaging over all possible locations and momenta of all the particles involved in the system. The rules employed in averaging are contained in the theory of statistical thermodynamics. This remarkable theory provides a set of relationships between thermodynamic quantities on the one hand and molecular quantities on the other. These relationships are presented in this chapter, and will be referred to as our "rules of the game"; the ultimate proof of their validity is provided *a posteriori* by the success of ST in predicting the thermodynamic quantities of a real system from knowledge of the molecular properties of its constituent particles.

Naturally, one may require a proof of the validity of these rules based on some other, more fundamental postulates. Indeed, such proofs are available in standard textbooks on ST. However, since our aim in this book is the applications of ST rather than its development, we shall not elaborate on the derivation of the rules from more fundamental postulates. The reader should be aware of the fact that no matter how deep one seeks to look for the foundations of the theory, at some point one must accept some postulates which can also be viewed as the "rules of the game." Therefore we start this book with what we believe to be the most convenient set of rules. The proof, or rather the confidence in their validity, will be achieved through their application to systems of interest in chemistry and biochemistry.

In this chapter we also present without derivation some elementary results of ST, e.g., the thermodynamics of ideal gases. These are derived in any elementary textbook on ST. They relate the thermodynamics of an ideal-gas system to the molecular properties of single molecules. For any real system one must use as input not only the properties of single molecules, but also some information on the extent of interactions among a group of a few molecules. These interactions are presumed to be known either from theory or from experiments. In practice, however, the interactions among even simple molecules are not known. Therefore it is a common practice to employ "model potential functions" to describe the interaction between two or more molecules and to use these

model functions as input in the theory. In section 1.7 we present a short survey of such model potential functions used in systems of interest in this book.

Both thermodynamics and statistical thermodynamics deal with the same quantities, such as temperature, pressure, energy, and entropy, yet there are some fundamental differences in the way the two theories handle these quantities.

Thermodynamics provides *general* relationships among thermodynamic quantities. For instance, we have the well-known relationships

$$\left(\frac{\partial S}{\partial P}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_P, \qquad \frac{C_V}{T} = \left(\frac{\partial S}{\partial T}\right)_V, \qquad V = \left(\frac{\partial G}{\partial P}\right)_T. \tag{1.1.1}$$

These are universal relationships in the sense that they apply to any system at equilibrium.

On the other hand ST deals with thermodynamic quantities that are pertinent to *specific* systems. The ultimate goal of ST is to calculate the thermodynamic quantities of a specific system in terms of its molecular properties. For example, the entropy of an ideal gas of simple particles can be computed from the well-known expression

$$S = Nk \ln\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \frac{e^{5/2}}{\rho}\right], \tag{1.1.2}$$

where N and T are the number of particles and the temperature, respectively, k is the Boltzmann constant, h is the Planck constant, and ρ is the number density ($\rho = N/V$). In Eq. (1.1.2) the entropy is given explicitly in terms of the thermodynamic variables T, V, N as well as in terms of the molecular parameters, in this case the molecular mass m. Thus we have $S = S(T, V, N; m)$. Clearly, for different gases we have different functions $S(T, V, N)$ depending on the parameter m. In more general cases we might obtain a function of the form $S = S(T, V, N; a_1 \cdots a_n)$ where $a_1 \cdots a_n$ are the molecular parameters characterizing the specific system under consideration.

In thermodynamics we often encounter "constants" that are characteristic of a certain equilibrium condition. For instance, for the chemical reaction

$$A + B \rightleftarrows C + D \tag{1.1.3}$$

at equilibrium in an ideal gas system, we have

$$K(T) = \frac{P_C P_D}{P_A P_D}, \tag{1.1.4}$$

where P_i is the partial pressure of the species i and K is the equilibrium constant. K depends on T but not on the pressure of the system. Thermodynamics does not offer a method of computing the equilibrium constant $K(T)$. Such a computation is possible for a specific system of A, B, C, and D using the methods of ST.

In thermodynamics we usually do not specify the choice of the independent variables used to describe our system. By writing $(\partial S/\partial T)_{V,N}$ it is implicitly assumed that S is viewed as a function of the independent variables T, V, N. Clearly, if the symbol S is used for the *value* of the entropy of a given system, then S is independent of the choice of the independent variables that describe the system. However, if S denotes the *function* then it becomes essential to specify the choice of the independent variables. Thus, $S(T, V, N)$, $S(T, P, N)$, and $S(T, V, \mu)$ are all different functions of the independent variables. These different functions are not equivalent. For example, having the functions $S(E, V, N)$ or $A(T, V, N)$ or $G(T, P, N)$, one can derive all other thermodynamic quantities of the system using standard thermodynamic relationships. This is not the case if we