ELSEVIER

Biophysical Chemistry 64 (1997) 7–13

Biophysical Chemistry

Thermodynamics, molecules and the Gibbs conference

John A. Schellman *

Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA

Received 28 May 1996; revised 15 July 1996; accepted 15 July 1996

Abstract

Molecular concepts have been gradually incorporated into chemical thermodynamics. This is done by applying standard thermodynamics manipulations to systems which are restricted by the assumption of a specific molecular model or mechanism. Here we trace the development of this procedure and its applications from Avogadro's hypothesis to modern work on site-directed mutagenesis.

Keywords: Avogadro's hypothesis; Gibbs conferences; Molecular mechanisms; Site-directed mutagenesis; Thermodynamics

1. Introduction

Modern thermodynamics does not appear to fit its classic definition. Webster [1] defines thermodynamics as "The science that deals with the relationship of heat and mechanical energy and the conversion of one into the other". This is clearly a pre-Gibbsian point of view. Kirkwood and Oppenheim [2], known for the precision and rigor of their methods, broaden this definition to "Thermodynamics is a macroscopic phenomenological discipline concerned with a description of the gross properties of systems of interest". Two features are discernible from these definitions. Thermodynamics deals with the relations of macroscopic quantities and is a science which has no explicit dependence on the atomic–molecular nature of matter.

How are these descriptions to be reconciled with current studies in physical biochemistry and biophysics? The laboratory next to my office is carrying on studies on the thermodynamics of cavities in proteins. One sees in the literature such titles as "The thermodynamics of phosphate transfer", "The thermodynamics of the hydrogen bond and its stabilizing influence on protein structure", "The thermodynamics of the hydrophobic interaction and its influence on protein–protein interactions", "The thermodynamics of the enzyme–substrate complex in the serine proteases", etc. Far from being non-molecular, it would be a fair statement that practically all papers with thermodynamic content, which are presented at the Gibbs conferences, have as their aim the search for molecular information.

2. Thermodynamics and molecular concepts

The encroachment of molecular concepts into thermodynamics began at about the same time that the first two laws of thermodynamics were definitively enunciated. Prior to 1860, when Cannizzarro

* Phone: (541) 346 6093; Fax: (541) 346 5891; E-mail JOHN@JAS1.UOREGON.EDU

PII S0301-4622(96)02206-5

drew attention to the molecular concept of Avogadro, the general gas law could only be stated as a proportionality, $PV \propto T$, since the proportionality constant varied from substance to substance for a given mass. By adding the assumption that PV is also proportional to the number of molecules, the proportionality could be converted to the general gas law, $PV = nRT$. R is a universal constant, and n is proportional to the number of molecules in the gas. When thermodynamics is applied to this equation of state, we have a system of thermodynamic relations in which molecularity is deeply embedded. The work of Dulong and Petit on the heat capacity of solids eventually introduced atomic considerations into solid state thermodynamics, and bewilderingly (at the time) established the "gas constant" as a key quantity in their results.

Twenty-five years later van't Hoff pursued the analogy of dilute solutions to gases at low pressure [1]. Starting out with the transformed gas law, $\pi = RTC$, where π is the osmotic pressure, he was able to derive correct limiting formulas for the freezing points, boiling points and vapor pressure of solutions and to provide thermodynamic derivations for the law of mass action and the laws of Henry and Raoult.

Van't Hoff's results were all approximate. He was unaware that Gibbs had already laid down the foundations for a rigorous discussion of solutions [4] in his first statement of the combined first and second laws for reversible processes:

$$dE = TdS - PdV + \mu_1 dm_1 + \mu_2 dm_2 + \cdots + \mu_n dm_n$$

The quantities of components, m_j, are not given a fixed definition, but Gibbs opens the door to the possibility of a molar, i.e. molecular, definition. "The units by which we measure the substance of which we regard the given mass as composed may each be chosen independently. To fix our ideas for the purpose of a general discussion we may suppose all substances measured by weight or mass. Yet, in special cases, it may be more convenient to adopt chemical equivalents as the units of component substances." Gibbs was even able to show that the chemical potential of a component has a logarithmic dependence on its concentration in the limit of high dilution [5]. From this starting point all of van't Hoff's results are easily obtained.

The link between thermodynamic behavior and molecular concepts became even stronger as a result of Arrhenius' theory of ionization. "Molecules" of NaCl and KOH completely dissociate in solution; those of H_2CO_3 and H_2SO_4 partially dissociate. To lead to fruitful relations the chemical potential of an ionizing substance must be written as the sum of the chemical potentials of its anions and cations.

As a result of these advances, expressions involving the gas constant, the number of moles, or partial molar properties are liberally distributed over most pages of textbooks of chemical thermodynamics. There is no real inconsistency with the definitions given above. To be useful thermodynamics must be applied to real systems which have their own laws of behavior. The mole concept is fundamental to the laws of gases, the colligative properties of solutions and to chemical reactions and equilibria. It is the fusion of these disciplines which leads to what we call chemical thermodynamics.

Complete confirmation of the appropriateness of a union of thermodynamics with the molecular description of matter has been supplied by statistical mechanics, where it is shown that combining molecular mechanics with probability postulates leads to valid thermodynamic conclusions. So accurate is the statistical mechanical method for certain problems, that theoretical calculations are considered to be more accurate than thermodynamic measurements. An important result is the modern tendency to think of all thermochemical quantities in terms of molecular mechanisms.

Important adjuncts to the molecular aspects of thermodynamics were the additivity laws discovered (or at least emphasized) during the 1930s and exploited up to the present time. It was found that the heat capacities, entropies, enthalpies, etc. of chain molecules were often linear functions of the number of units in the chain [6]. Examples were high polymers, aliphatic hydrocarbons, aliphatic alcohols, aliphatic carboxylic acids, and many others. In this way it was possible to ascribe an entropy, enthalpy,

[1] For a modern discussion, with references, of the contributions of van't Hoff and Arrhenius, see Chapter 1 of Servos [3].

heat capacity or free energy to a unit in the chains such as a $-CH_2-$ group or an isoprene unit. Some of the applications had a theoretical basis such as the work of Pitzer and Gwinn [7], but most were purely empirical. Additivity rules have been extended to branched as well as linear chains and to a variety of chemical groups. In addition, the concept of additivity of bond energies has led to a simple method of approximating the energies of chemical reactions. During WWII, while the physics community was busy at Los Alamos, Oakridge and Hanford, chemical teams were calculating the enthalpy of reaction of substances and mixtures which were potential explosives and propellants.

Up until about the time of that war, thermodynamics was mainly in the hands of professionals: investigators whose primary research programs were devoted to thermodynamics itself. Though there still remains an extensive group of practitioners, there is more and more emphasis on the use of thermodynamics as a research tool to solve mechanistic problems. This has been particularly true in the biophysical sciences, where thermodynamics is often regarded as a technique to help solve problems, not as the primary aim. As we have seen, thermodynamic applications to molecular problems require external models of some kind: equations of state, observational laws, or pattern discovery. The current procedure is to invent models to match a particular problem which is under investigation.

We can illustrate with a very old example: the titration of a molecule possessing more than one acid or base group. Experiment shows that the titration is not accurately governed by the intrinsic binding constants for protons. The problem was studied by Bjerrum [8] who proposed that in addition to the standard free energy of ligation, one must include a term for the interaction of the charges on the species. Instead of the binding polynomial for independent sites

$$1 + (K_1 + K_2)[H^+] + K_1K_2[H^+]^2$$

he proposed the relation

$$1 + (K_1 + K_2)[H^+] + K_1K_2\, e^{-w/RT}\,[H^+]^2$$

w was calculated by applying Coulomb's law to the bound protons. (The Debye–Hückel theory which permitted the calculation of screened potentials had not yet been published.) Since $K_1K_2\, e^{-w/RT} = \exp[-(\Delta G_1^\circ + \Delta G_2^\circ + w)/RT]$, this was tantamount to adding an interaction free energy to the sum of the standard free energies of the two ligands. Bjerrum's idea is correct but the results are not very accurate because of molecular flexibility, neglect of screening, and the inaccuracies of the macroscopic Coulomb's law at distances of a few Ångströms. A more modern approach would be to perform the titrations, measure w as an experimental parameter (preferably as a function of temperature) and then seek molecular models which will account for the observations. An even more modern approach (akin to site-directed mutagenesis) would be to add or delete titratable groups and measure interactions by observing differences in titration curves. This method is in fact the oldest and was accomplished by Wegscheider [9] in 1895! See [10]. The main point is that one is now measuring the thermodynamics of an interaction, not a system or a mole of substance or a group. Interactions are also observable from the deviations of dilute gases and dilute solutions from ideality.

Hemoglobin presents the problem in a much more subtle form. Back in the 1920s it was shown by Adair [11] that the four phenomenological binding constants are not derivable from four identical intrinsic constants. Current research combines structural studies, mutants generated by molecular biological techniques, and thermodynamic studies to evaluate the interaction parameters for the different pairs of binding sites. The problem is made more interesting (and complicated) by the allosteric effects of diphosphoglycerate, CO_2 and pH. Many structural, biochemical, genetic, and physical techniques are being used for this problem, but the thermodynamic results are the real measure of the effectiveness of the mechanism of oxygen retention and release by hemoglobin [12].

Another very widespread application is to the transitions of biopolymers. For proteins, a common approach to the unfolding transition has been the two-state model. The condition for the applicability of this model is that in the transition region practically all molecules are in either the unfolded or the folded state. The concentrations of intermediates are undetectably small. If one has two identifiable states

one can define changes in the thermodynamic functions between them: $\Delta\overline{G}$, $\Delta\overline{H}$, $\Delta\overline{S}$, $\Delta\overline{C}p$. (Note that these are partial molar quantities, the only way one can assign an extensive property to a component of a solution!) A common way to model the thermal transitions of proteins is to assume a form for the change in heat capacity. This choice is influenced by the large temperature coefficient of the heat capacity for hydrophobic interactions. Two temperature integrations lead to a long-established formula for $\Delta\overline{G}$ which can be used in the van't Hoff analysis of the transition to derive values for the parameters (see, for example, [13]).

In the above we are evidently not dealing with the thermodynamic definition of a "state" which relies on a listing of its intensive variables: pressure, temperature, density, composition, etc. The transformations take place at a molecular level. In the absence of precipitation they are invisible to the human senses and require a physical technique that has been shown to be sensitive to the transformation (calorimetry, CD, UV, or IR bands, absorbance, fluorescence, NMR). What we mean by a state of a protein molecule is most easily defined in statistical mechanical terms. We consider a single protein molecule surrounded by solvent components and write the partition function as a sum of two parts

$$Q = Q_u + Q_f \qquad (1)$$

where Q_f is the sum over all states of the system (not just the protein molecule) in which the protein molecule has its native, folded structure. Solvent conformation is of major importance. Q_f includes the region surrounding the lowest free-energy conformation because of thermal fluctuations of the native state. The native state is defined as the distribution over the region of phase space of Q_f. Q_u spans the unfolded state and is the sum over all states in which the chain behaves like a random polymer. This includes the transient formation of helices, folded segments, etc. By the definition of a two-state transition, these two regions do not intersect. Then the probability of the folded form is given by Q_f/Q and the equilibrium constant for unfolding by $K = Q_u/Q_f$.

This theoretical definition is a great aid to thought but is little help in the laboratory. In practice the native and unfolded states of proteins are defined by the experimental technique that is used. For example, studies establish that the native protein has a characteristic negative circular dichroism at 222 nm. On unfolding this is changed to another value at 222 nm. When studying a transition, we assume that there is a linear relation between the observed CD and the fraction unfolded. In this case the states are defined by their CD spectra. One could use calorimetry, NMR, fluorescence, IR, side-chain CD, etc. Each of these techniques has a body of theory behind it to aid in the interpretation. In principle, all these methods could lead to different results. Protein chemists have in general been aware of this problem and it is common practice to observe the same transition with a variety of methods. For the cleanest cases all these criteria are in agreement. A strict criterion for a two-state transition is that within the transition zone, all properties which are expected to be additive (absorption spectra, circular dichroism, extensive thermodynamic properties such as H, S, V) should be linear combinations of the properties of the two states.

Studies during the past ten years have revealed the molten globule or A-state of proteins [14]. These appear to be compact molecules having some of the secondary structural characteristics of the native state but an ill-defined tertiary structure. These "states" have less consistent properties than native or unfolded states, but presumably occupy regions of phase space in the intermediate zone between the native and unfolded states.

One of the most direct ways in which molecular structure can be expressed by thermodynamic experiments is the comparison of directed mutations, either natural or synthetic. Pauling et al. [15] were evidently the first to see that the change in a single amino acid in a large protein could significantly change its physical properties and its biological effectiveness. The point of view developed in this essay was initiated about twenty years ago [16] by the following conundrum. Proteins unfold cooperatively. This is interesting, appealing and useful. It permits us to observe beautifully sharp transitions and to use two-state thermodynamics. On the other hand, since everything happens at once, it effectively obscures all information on the interactions which stabilize the protein. However, if one can compare

the transitions of a wild-type protein and a protein with a single mutation, the differences in transition thermodynamics can unequivocally be attributed to the interchange of a single amino acid (see [17] for a recent review). At the present time site-directed mutagenesis is a major industry in almost every branch of modern biochemistry and molecular biology. We discuss it here because it has made it possible to apply thermodynamics to the study of the mechanism of protein stabilization as well as many other problems.

The helix–coil transition is another interesting case. Originally, initiation and propagation parameters were introduced via a linear free-energy relationship [18]. Zimm and Bragg [19] approached the problem in a more general and less specific way by introducing "statistical weights". These are quantities which are intermediate between thermodynamics and statistical mechanics. They are related to the grouping of states in the partition function as in Eqs. (1) and (2) (below). One can identify s with free energy via the formula $s = \exp[-\Delta g_a/RT]$, where Δg_a is the free energy change when one unit of a polypeptide chain is added to an existing helix. One can also view it in terms of partition functions, $s = q_h/q_c$, where q_h is the sum over all states of the system when the peptide unit is in a helix, and q_c is the sum over all states when it is not. Lifson and Oppenheim [20] have provided a detailed analysis of this viewpoint, which was later used in the Lifson–Roig theory [21]. In recent work further parameters (statistical weights) have been added to the theory to take into account capping residues and side-chain interactions [22].

The concept of statistical weights is now frequently used to set up thermodynamic models for systems which are beyond our capacity to solve using statistical mechanics. There is no way in which a genuine partition function (sum over molecular states) can be set up for a complex biopolymer interacting with solvent. Statistical weights divide the molecular phase space into regions which are experimentally identifiable. The total partition function is then the sum over these regions. The important aspect is that these parameters can be evaluated by experiment.

This point of view is being adopted in the field of ligand interactions with biopolymers. The binding polynomial, which plays a central role in Wyman's linkage thermodynamics [23], is frequently, and correctly, referred to as the "partition function". The appropriate partition function for a protein interacting with a number of identical ligands is the semigrand partition function for a highly dilute solution of the biopolymer:

$$\Psi_p = Q_0 + Q_1 e^{\mu/RT} + Q_2 e^{2\mu/RT} + \cdots + Q_n e^{n\mu/RT} \quad (2)$$

where the subscript p indicates that the number of protein molecules is constant, μ is the chemical potential of the ligand, and Q_n is the partition function for a protein to which n ligands are bound. After dividing by Q_0, it is easy to show (by relating both the partition functions and the equilibrium constants to free energies) that this is identical to the binding polynomial $1 + K_1[L] + K_2[L]^2 + \cdots + K_n[L]^n$. [2] [L] is the ligand activity. This result can be applied to more complex cases such as multiple ligands.

Scanning and titration calorimetry are purely thermodynamic methods, but because third law entropies are not available for complex protein solutions, a full interpretation must make use of the two-state, or multistate, model (see, for example, [24]). The data which are obtained are heat capacities and enthalpies. For a two-state transition the enthalpy itself may be used to measure the extent of unfolding. Calorimetry has the advantage that it can test the two-state hypothesis by comparing van't Hoff and calorimetric enthalpies. Models are further being used to distinguish hydrophobic and polar interactions in aqueous solution.

One of the better known applications is the use of the additivity concept to estimate the stability of proteins [25]. This has its origins in Kauzmann's formulation of hydrophobicity based on the special thermodynamic behavior of water in the vicinity of

[2] The formula as given is an approximation. The semigrand partition function has V as an independent variable, binding polynomials are usually measured at constant pressure. The formula is accurate for most applications to solutions where changes in free energy are much larger than changes in PV except at very high pressures. See [10] for a full discussion without approximation.

non-polar groups [26]. The free energies of transfer of all the amino acid side-chains (and the peptide group itself) from a non-aqueous to an aqueous environment have been measured or estimated. The cohesive free energy of a compact protein is then estimated by summing these free energies in accordance with the protein structure and amino acid composition. The results have provided a very satisfactory qualitative representation of the stabilizing forces in protein molecules, especially for the hydrophobic interactions. Agreement is not quantitative. Three possible reasons have been considered: the difficulty in modeling the interior of a protein by a uniform phase, assumptions which must be made about residues at the surface, and the incompleteness of the database of transfer data. There is insufficient work on series of side-chains and variation of temperature.

Another approach to transfer and stability makes use of the microscopic surface area of groups exposed to water and other solvents [27]. Known structures or detailed models are necessary to evaluate the surface areas. We are here applying a macroscopic, thermodynamic concept to specific molecular structures.

3. Conclusion

The examples discussed above are derived from the author's experiences and interests. The range of applications could be increased tenfold. They do exemplify the basic principle which is that thermodynamics applied within the framework of a detailed molecular model permits the investigator to take "the broad highway of thermodynamics" [28] to explore the mechanisms of complex biological molecules and assemblies. Others would emphasize the areas of DNA replication and translation, enzyme action, immunology, allosteric interactions of proteins, tRNA and mRNA structure, stabilized intermediates on kinetic pathways, membrane transduction, receptor interactions, gene-expression, and biological assembly to name a few. We chemists and biophysicists found it amusing that Gunther Stent discussed the decline and fall of molecular biology and wrote articles with titles such as "That was the molecular biology that was" in the mid-1960s [29], the idea being that the heroic age of molecular biology was over and significant discoveries lay mainly in the past. For us the game was just about to begin. The early genetic and biochemical studies established *what* was happening, but the *how* and *why*, i.e. the detailed molecular mechanisms, were totally absent. The progression from Avogadro's hypothesis to site-directed mutagenesis has established a framework for thermodynamic studies to play an important part in establishing molecular mechanisms.

There is no inconsistency in the statements made about thermodynamics in the first and second paragraphs at the beginning of this essay. The resolution arises from the combination of classical thermodynamics with non-thermodynamic restrictions, based on models of molecular interactions and mechanisms. The main difference between current practice and that of earlier times is that the field is motivated by the needs of individual research problems, rather than arising from generalizations of a body of collected thermodynamic data.

In recent years the Gibbs conferences have played a very central role in the application and propagation of this type of research. Indeed, the field was well on its way before the first Gibbs meeting in 1987, but there is no other meeting at which the applications of thermodynamics to biological systems are so uniformly and unanimously embraced. There is also the emphasis on young investigators for both the presentation of results and as attendees. At the present time the biosciences are going through a period of very rapid qualitative discovery and expansion. The attractions of becoming a part of the exciting search for novelty and discovery are strong, and we would not be real scientists if we ignored such possibilities when we encounter them. On the other hand, the history of science demonstrates that the quantification of a field is a necessary step in its ultimate formulation. At the Gibbs conferences we are, and we generate, a group of scientists who are making sure that this aspect of our work will not fall into neglect.

Finally, a caveat for the overly enthusiastic. The thermodynamic functions of a substance are usually slowly varying functions of the external parameters (T, P, concentrations) which can usually be represented by linear or quadratic relations. As such they are not well suited for making subtle distinctions

between models. The fitting of a model to a set of thermodynamic data does not prove that the model is correct, even in cases where the accuracy is high. Other models might do just as well. In general, a considerable amount of support must be obtained from other experimental approaches and from theory where applicable. In the absence of such support, the only conclusion that can be drawn is that the model is not inconsistent with the data. Any parameters obtained by the analysis using such a model should be clearly reported as provisional.

Acknowledgements

The author wishes to thank Enrico Di Cera and Robert Baldwin for helpful suggestions and advice in preparing this manuscript and Charlotte Schellman whose early interest in species variations of egg-white lysozymes was important in developing our ideas on the synthesis of thermodynamics and molecular concepts. The work was supported by NIH grant GM 20195.

References

[1] Webster's New Universal Dictionary, Simon and Schuster, NY, 1979.
[2] J.G. Kirkwood and I. Oppenheim, Chemical Thermodynamics, McGraw-Hill, New York, 1961.
[3] J.W. Servos, Physical Chemistry from Ostwald to Pauling, Princeton University Press, Princeton, 1990.
[4] J.W. Gibbs, in W.R. Longley and R.G. Van Name, The Collected Works of J. Willard Gibbs, Yale University Press, New Haven, 1948.
[5] J.W. Gibbs, in W.R. Longley and R.G. Van Name, The Collected Works of J. Willard Gibbs, Yale University Press, New Haven, 1948, p. 135.
[6] G.J. Janz, Estimation of Thermodynamic Properties, Academic Press, New York, 1958.
[7] K.S. Pitzer and W.D. Gwinn, J. Chem. Phys., 10 (1942) 428.
[8] N. Bjerrum, Z. Phys. Chem., 106 (1923) 219.
[9] R. Wegscheider, Monatsch. Chem., 16 (1895) 153.
[10] E. Di Cera, Thermodynamic Theory of Site-Specific Binding Processes in Biological Macromolecules, Cambridge University Press, New York, 1995.
[11] G.S. Adair, J. Biol. Chem., 63 (1925) 529.
[12] G.K. Ackers, M.L. Doyle, D. Myers and M.A. Daugherty, Science, 255 (1992) 54.
[13] W.J. Becktel and J.A. Schellman, Biopolymers, 26 (1987) 1859.
[14] O. Ptitsyn, in T.E. Creighton (Ed.), Protein Folding, Freeman, New York, 1992, pp. 243–300.
[15] L. Pauling, H.A. Itano, S.J. Singer and I.C. Wells, Science, 110 (1949) 543.
[16] M.L. Elwell and J.A. Schellman, Biochim. Biophys. Acta, 494 (1977) 367.
[17] B.W. Matthews, Adv. Protein Chem., 46 (1995) 249.
[18] J.A. Schellman, C. R. Trav. Lab. Carlsberg, Ser. Chim., 29 (1955) 230.
[19] B.H. Zimm and J.K. Bragg, J. Chem. Phys., 31 (1959) 526.
[20] S. Lifson and I. Oppenheim, J. Chem. Phys., 33 (1960) 109.
[21] S. Lifson and A. Roig, J. Chem. Phys., 34 (1961) 1963.
[22] A.J. Doig, A. Chakrabartty, T.M. Klingler and R.L. Baldwin, Biochemistry, 33 (1993) 3396.
[23] J. Wyman and S.J. Gill, Binding and Linkage: Functional Chemistry of Biological Macromolecules, University Science Books, Mill Valley, CA, 1990.
[24] W. Pfeil and P.L. Privalov, Biophys. Chem., 4 (1976) 41.
[25] Y. Nozaki and C. Tanford, J. Biol. Chem., 245 (1970) 1646.
[26] W. Kauzmann, Adv. Protein Chem., 14 (1959) 1.
[27] B. Lee and F.M. Richards, J. Mol. Biol., 55 (1971) 379.
[28] G.N. Lewis and M. Randall, Thermodynamics and the Free Energy of Chemical Substances, McGraw-Hill, New York, 1923.
[29] G. Stent, Science, 160 (1968) 390.