J. theor. Biol. (2002) **215,** 491–501
doi:10.1006/jtbi.2001.2516, available online at http://www.idealibrary.com on IDEAL®

Thermodynamics of Systems of Biochemical Reactions

R. A. Alberty

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

(*Received on* 4 *January* 2001, *Accepted in revised form on* 5 *December* 2001)

When a reaction system described in terms of species is in a certain state, the Gibbs energy G provides the means for determining whether each reaction will go to the right or the left, and the equilibrium composition of the whole system can be calculated using G. When the pH is specified, a system of biochemical reactions is described in terms of reactants, like ATP (a sum of species), and the transformed Gibbs energy G' provides the means for determining whether each reaction will go to the right or the left. The equilibrium composition of the whole system can be calculated using G'. Since metabolism is complicated, the thermodynamics of systems of reactions like glycolysis and the citric acid cycle can also be considered at specified concentrations of coenzymes like ATP, ADP, NAD_{ox}, and NAD_{red}. This is of interest because coenzymes tend to be in steady states because they are involved in many reactions. When the concentrations of coenzymes are constant, the further transformed Gibbs energy G'' provides the means for calculating whether each reaction will go to the right or the left, and the equilibrium composition of the whole system can be calculated using G''. Under these conditions, a metabolic reaction system can be reconceptualized in terms of sums of reactants; for example, glycolysis can be represented by $C_6 = 2C_3$, where C_6 is the sum of the reactants with six carbon atoms and C_3 is the sum of the reactants with three carbon atoms. These calculations can also be described by use of semigrand partition functions. Semigrand partition functions have the advantage of containing all the thermodynamic information on a series of reactions at specified pH or at specified pH and specified concentrations of coenzymes.

© 2002 Elsevier Science Ltd. All rights reserved.

1. Introduction

It is a common misconception that the thermodynamics of biochemical reaction systems at constant pH can be analysed with the Gibbs energy G. After all, does not the Gibbs energy provide the criterion for spontaneous change and equilibrium at constant T and P? Yes, but when you hold the pH constant in addition to T and P, the system will in general go to a different equilibrium composition at each pH. When the pH is specified, there are three independent intensive variables T, P, and pH, and a Legendre transform has to be used to define a new thermodynamic potential that is minimized at equilibrium. This is not a new idea because it goes back to what Gibbs did in the 1870s. The internal energy U provides the criterion for spontaneous change at specified entropy and volume (an isolated system), but these are not convenient variables to hold constant in the laboratory, and so Gibbs subtracted two products of conjugate variables from the internal energy to define what we now call the Gibbs energy:

$$G = U + PV - TS. \tag{1}$$

0022-5193/02/$35.00/0

© 2002 Elsevier Science Ltd. All rights reserved.

This is a Legendre transform, and it produces a new thermodynamic potential G with T and P as its natural independent variables.

When the pH is specified, the criterion for spontaneous change and equilibrium is provided by a transformed Gibbs energy G' defined by the following Legendre transform (Alberty, 1992a,b):

$$G' = G - n_c(H)\mu(H^+), \tag{2}$$

where the conjugate variables are $n_c(\mathrm{H})$, the amount of the hydrogen component, and $\mu(\mathrm{H}^+)$, the chemical potential of hydrogen ions determined by the specification of the pH. The concept of a component is very important, and it is discussed in Section 2.4. At a specified pH, the various protonated forms of ATP, that is ATP^{4-}, HATP^{3-}, and $\mathrm{H_2ATP}^{2-}$, have fixed relative concentrations at equilibrium, and so we can calculate a standard transformed Gibbs energy of formation of the entity ATP. This makes it possible to calculate the contribution that ATP makes to any reaction in which it is involved. Alberty & Goldberg (1992) calculated the standard transformed thermodynamic properties of the ATP series at specified pH and pMg.

Recently, Alberty (2001d) has placed on the web mathematical functions in Mathematica® (Wolfram Research, Inc., 100 World Trade Center, Champaign, IL 61820-7237) that can be used to calculate standard transformed Gibbs energies of formation of 116 biochemical reactants (Alberty, 1998a,b) at any pH in the range pH 5–9 and any ionic strength in the range 0–0.35 M at 298.15 K. Thus, the apparent equilibrium constant K' can be calculated for any reaction between these 116 reactants. The size of this table can be increased a great deal using literature values for apparent equilibrium constants. The usefulness of such a table will increase exponentially with the number of reactants in the table. Goldberg & Tewari (1993–1999) have published a series of papers on evaluated literature data on apparent equilibrium constants of enzyme-catalysed reactions.

This process of making Legendre transforms can be carried a step further to obtain a more global view of systems of biochemical reactions like glycolysis, the pyruvate dehydrogenase reaction, and the citric acid cycle. In making thermodynamic calculations on systems of enzyme-catalysed reactions, it is useful to take advantage of the fact that coenzymes like ATP, ADP, P_i, NAD_{ox}, and NAD_{red} are in steady states. When their concentrations are constant, the following Legendre transform can be used to define a further transformed Gibbs energy G'' by subtracting five products of conjugate variables (Alberty, 1993, 2000):

$$\begin{aligned} G'' = G' &- n'_c(\mathrm{ATP})\mu'(\mathrm{ATP}) - n'_c(\mathrm{ADP})\mu'(\mathrm{ADP}) \\ &- n'_c(\mathrm{P}_i)\mu'(\mathrm{P}_i) - n'_c(\mathrm{NAD}_{ox})\mu'(\mathrm{NAD}_{ox}) \\ &- n'_c(\mathrm{NAD}_{red})\mu'(\mathrm{NAD}_{red}), \end{aligned} \tag{3}$$

where $n'_c(\mathrm{ATP})$ is the amount of the ATP component and $\mu'(\mathrm{ATP})$ is the specified transformed chemical potential of ATP at the specified pH. When the pH is specified, the abbreviation NAD_{ox} is used rather than NAD^+, which tends to suggest that electric charges should be balanced, and NAD_{red} is used rather than NADH, which tends to suggest that hydrogen atoms should be balanced. When the concentrations of coenzymes are specified, the relative concentrations of reactants like glucose, G6P, F6P, and FBP are fixed at equilibrium, and so we can calculate the standard further transformed Gibbs energies of formation of the entity $\mathrm{C_6}$, that is the thermodynamic properties of the equilibrium mixture of the four reactants (glucose, G6P, F6P, and FBP). Similarly, the standard further transformed Gibbs energy of formation of the entity $\mathrm{C_3}$ can be calculated. $\mathrm{C_3}$ represents the pseudoisomer group of reactants in glycolysis with three carbon atoms. Thus, specifying the concentrations of coenzymes has reduced glycolysis to a single reaction (Alberty, 2000):

$$\mathrm{C_6} = 2\mathrm{C_3}. \tag{4}$$

Now, the equilibrium composition in glycolysis can be calculated by simply solving a quadratic equation. Since $[\mathrm{C_6}]_{eq}$ and $[\mathrm{C_3}]_{eq}$ can be calculated, the equilibrium concentrations of all the reactants in glycolysis can be calculated for the specified pH, [ATP], [ADP], $[\mathrm{P}_i]$, $[\mathrm{NAD}_{ox}]$, and

[NAD_{red}]. This is more complete than dealing with net reactions because all the information on the system of reactions is included.

This introduction has shown what it means to treat a system of reactions on three levels. Level 1 is based on the Gibbs energy G and the use of species. At this level the concentration of hydrogen ions is a dependent variable, which has to be calculated from the equilibrium constants and conservation equations. Level 2 is based on the transformed Gibbs energy G' at specified pH (and perhaps free concentrations of metal ions that are bound reversibly) and uses reactants like ATP, which are sums of species. Level 3 is based on the further transformed Gibbs energy G'' at specified concentrations of coenzymes and deals with sums of reactants like C_6 and C_3 in glycolysis. The advantage of calculations at levels 2 and 3 is that there are fewer reactants to deal with, so that a broader overview of the reaction system is obtained. Yet none of the thermodynamic information is lost.

Callen (1985) pointed out that "The choice of variables in terms of which a given problem is formulated, while a seemingly innocuous step, is often the most critical step in the solution." This article is an example of Callen's statement.

Before considering calculations of equilibrium compositions at levels 2 and 3 in more detail, it is important to point out that linear algebra makes it much easier to make the necessary calculations and the use of a mathematical application in a personal computer is really necessary. The calculations described here have been made with Mathematica^{R}, which is very convenient for this purpose.

2. Calculation of Thermodynamic Properties of Species, Reactants, and Reactions

2.1. LEVEL 1

Level 1 is the familiar thermodynamics of species based on the Gibbs energy G. These calculations are discussed in every introductory textbook on physical chemistry, but there is one way to obtain a broader view that is essential for the discussion of levels 2 and 3. When isomers of a species are in equilibrium and are involved in chemical reactions, the standard Gibbs energy of formation of the isomer group $\Delta_f G^0(iso)$ can be calculated using

$$\Delta_f G^0(iso) = -RT \ln \sum_{i=1}^{N_{iso}} \exp(-\Delta_f G_i^0/RT), \qquad (5)$$

where N_{iso} is the number of pseudoisomers in the group. This is a kind of partition function. The equilibrium mole fractions r_i within the pseudoisomer group are given by

$$r_i = \exp((\Delta_f G^0(iso) - \Delta_f G_i^0)/RT). \qquad (6)$$

This is a kind of Boltzmann distribution. These standard Gibbs energies of formation of isomer groups can be used like standard Gibbs energies of species in the following equation to calculate the equilibrium constant K of a chemical reaction involving the isomer group:

$$\Delta_r G^0 = \sum v_i \Delta_f G_i^0 = -RT \ln K, \qquad (7)$$

where the v_i are the stoichiometric numbers (positive for products and negative for reactants). In treating reactions in dilute aqueous solutions, it is convenient to take the thermodynamic properties to be functions of ionic strength and to write equilibrium constant expressions in terms of concentrations. In other words, the extended Debye–Hückel equation is used to make adjustments for the ionic strength.

2.2. LEVEL 2

It can be shown (Alberty, 1992b,1998a) that the standard transformed Gibbs energy of formation $\Delta_f G_i'^0$ of a species at a specified pH can be calculated from its standard Gibbs energy of formation $\Delta_f G_i^0$ by using

$$\Delta_f G_i'^0 = \Delta_f G_i^0 - N_H(i) \times (\Delta_f G^0(H^+) + RT \ln 10^{-pH}), \qquad (8)$$

where $N_H(i)$ is the number of hydrogen atoms in species i and $pH = -\log[H^+]$. As mentioned above, thermodynamic properties like $\Delta_f G_i^0$ and $\Delta_f G_i'^0$ are taken to be functions of the ionic strength.

When a biochemical reactant is made up of several species (for example, ATP^{4-} $HATP^{3-}$, and H_2ATP^{2-}) that are pseudoisomers at a specified pH, the standard transformed Gibbs energy of formation $\Delta_f G'^0(iso)$ of the pseudoisomer group at equilibrium can be calculated from the standard transformed Gibbs energies of formation [see Eq. (8)] of the several species by the use of

$$\Delta_f G'^0(iso) = -RT\ln \sum_{i=1}^{N'_{iso}} \exp(-\Delta_f G'^0_i/RT), \qquad (9)$$

where N'_{iso} is the number of pseudoisomers in the group. This is a kind of partition function. This equation is like eqn (5), but transformed Gibbs energies of formation of species at a specified pH are involved. The equilibrium mole fractions r_i of species within the pseudoisomer group at a specified pH are given by

$$r_i = \exp((\Delta_f G'^0(iso) - \Delta_f G'^0_i)/RT). \qquad (10)$$

This is a kind of Boltzmann distribution.

If the standard transformed Gibbs energies of formation of all the reactants in a biochemical reaction are known, the apparent equilibrium constant K' for the reaction at a specified pH can be calculated using

$$\Delta_r G'^0 = \sum v'_i \Delta_f G'^0_i = -RT\ln K', \qquad (11)$$

where a prime is used on the stoichiometric numbers to distinguish them from the stoichiometric numbers of the underlying reactions in terms of species. K is a function of T, P, and ionic strength, but K' is a function of T, P, pH, and ionic strength.

2.3. LEVEL 3

When concentrations of coenzymes, like ATP and ADP, are held constant, a Legendre transform is used to define a further transformed Gibbs energy G'' that provides the criterion for spontaneous change and equilibrium (Alberty, 1993, 2000). The concentrations of a number of coenzymes can be specified, but these reactants must be components, rather than non-components. The concept of components is discussed in the next section. The Legendre transform that is used to define the further transformed Gibbs energy G'' is

$$G'' = G' - \sum_{i=1}^{C_s} n'_c(i)\mu'(i), \qquad (12)$$

where $n'_c(i)$ is the amount of component i in the system and C_s is the number of components with specified concentrations. This is a generalized version of eqn (3). It can be shown that the standard further transformed Gibbs energy of formation $\Delta_f G''^0_i$ of a reactant at specified pH and specified concentrations of certain coenzymes can be calculated from its standard transformed Gibbs energy of formation $\Delta_f G'^0_i$ at specified pH by using

$$\Delta_f G''^0_i = \Delta_f G'^0_i - \sum_{i=1}^{C_s} N_{cj}(i)(\Delta_f G'^0_i(j) + RT\ln[j]), \qquad (13)$$

where $N_{cf}(i)$ is the number of coenzymes j involved in reactant i. Note that this equation is like eqn (8). When reactants are pseudoisomers, the standard further transformed Gibbs energy of formation $\Delta_f G''^0_i(iso)$ of the pseudoisomer group can be calculated by using

$$\Delta_f G''^0_i(iso) = -RT\ln \sum_{i=1}^{N''_{iso}} \exp(-\Delta_f G''^0_i/RT), \qquad (14)$$

where N''_{iso} is the number of pseudoisomers in the group. This equation is like Eq. (9), but it deals with the further transformed thermodynamic properties at specified concentrations of coenzymes. The equilibrium mole fractions r_i within the pseudoisomer group at specified concentrations of coenzymes can be calculated by using

$$r_i = \exp((\Delta_f G''^0(iso) - \Delta_f G''^0(i))/RT). \qquad (15)$$

If the standard further transformed Gibbs energies of formation of all the pseudoisomer groups in a reaction are known, the apparent equilibrium constant K'' at the specified concentrations of coenzymes like ATP and ADP can be

calculated using

$$\Delta_r G''^0 = \sum v_i'' \Delta_f G_i''^0 = -RT \ln K'', \quad (16)$$

where the double prime is used on the stoichiometric numbers to distinguish them from the stoichiometric numbers of the underlying reactions in terms of reactants. Equilibrium compositions can be calculated at level 3 using apparent equilibrium constants K''.

A more global view of glycolysis can be obtained by specifying the concentrations of coenzymes ATP, ADP, P_i, NAD_{ox}, and NAD_{red}. Doing this reduces glycolysis to $C_6 = 2C_3$, where C_6 is the pseudoismer group of reactants with six carbon atoms and C_3 is the pseudoisomer group with three carbon atoms. When the concentrations of coenzymes in the pyruvate dehydrogenase reaction and the citric acid cycle are specified, quantitative equilibrium calculations can be made with

$$C_3 + C_4 = C'_6 + C_1, \quad (17)$$

$$C'_6 = C_5 + C_1, \quad (18)$$

$$C_5 = C_4 + C_1, \quad (19)$$

where C_4, C_5, and C'_6 represent pseudoisomer groups of reactants in the citric acid cycle with four, five, and six carbon atoms, respectively. The prime on C'_6 is used to differentiate this pseudoisomer group from the pseudoisomer group C_6 in glycolysis. C_3 is pyruvate plus other reactants with three carbon atoms, and C_1 represents the sum of $CO_2(aq)$, H_2CO_3, HCO_3^-, and CO_3^{2-}. Using these three reactions, the equilibrium concentrations of all of the reactants can be calculated at specified concentrations of the coenzymes.

Note that the same pattern of equations is used at levels 1, 2, and 3, but the reaction equations are written differently. At level 1, this paper emphasizes the Gibbs energy, but there are corresponding entropies S and enthalpies H. At level 2, there are corresponding transformed entropies S'' and transformed enthalpies H'. At level 3, there are corresponding further transformed entropies S'' and further transformed enthalpies H''. At each level there is a full set of Maxwell equations, Gibbs–Helmholtz equations, and Gibbs–Duhem equations, and so there are lots of quantitative relations between these thermodynamic properties, which are not discussed here.

2.4. CONCEPT OF COMPONENTS

Components are important because they are conserved in a closed reaction system, and conservation equations provide powerful constraints on what can happen in a reaction system. In chemical reactions, atoms of elements and electric charges are conserved, but, for a chemical reaction system the element and charge conservation equations may be redundant. Only an independent set of element and charge conservation equations is used in calculating an equilibrium composition. The number C of independent conservation equations for a chemical reaction system is given by $C = N - R$, where N is the number of species and R is the number of independent reactions.

When the pH is specified for a system of enzyme-catalysed reactions, hydrogen atoms and charges are not conserved in the reaction system, and so you might expect that there would be one less component. But this is usually not the case. In fact, the number of components may be significantly greater. The reason is that enzyme mechanisms may introduce constraints that behave like atom conservations; that is, they lead to linear relations between amounts of reactants. The number C' of independent conservation equations for a system of biochemical reactions at a specified pH is given by $C' = N' - R'$, where N' is the number of reactants (sums of species) and R' is the number of independent biochemical reactions. What is conserved in addition to elements other than hydrogen? The answer is various groups of atoms. It is somewhat arbitrary as to what groups are chosen, but linear algebra provides the means for selecting different sets of conservation equations. The conservation equations can be written in terms of reactants, and the N' reactants can be divided into C' conserved groups of atoms and R' non-conserved groups of atoms because $N' = C' + R'$. When the pH is specified, the thermodynamic property conjugate to $\mu(H^+)$ is the amount $n_c(H)$ of the hydrogen component in the system.

When the concentrations of certain coenzymes are held constant, in addition to the pH, these coenzymes are not conserved in the reaction system, just like hydrogen atoms are not conserved in the reaction system when the pH is specified. In glycolysis when the concentrations of ATP, ADP, P_i, NAD_{ox}, and NAD_{red} are held constant, there is only one component left, namely carbon, and a single reaction, which is represented by $C_6 = 2C_3$. The number C'' of components is given by $C'' = N'' - R''$, where N'' is the number of reactants (like C_6 and C_3) and R'' is the number of independent reactions. In the glycolysis case, $C'' = 2 - 1 = 1$. Thus, there is a limit on the number of concentrations that can be held constant in an equilibrium calculation; there has to be at least one component left unspecified. The complete Legendre transform at specified concentrations of coenzymes yields the Gibbs–Duhem equation for the system, which shows that the intensive variables for a reaction system are not all independent. When the steady-state concentration of a coenzyme is specified, the thermodynamic property conjugate to the transformed chemical potential of a coenzyme is the amount n_c of free coenzyme plus "bound" coenzyme in the system.

The conservation matrix **A** for a system of chemical reactions has a column for each species and a row for each element, but redundant rows are deleted. When a Gaussian reduction is performed on this conservation matrix, the resulting matrix is in canonical form with an identity matrix in the left side. The number of rows is equal to the number of components, but the groups of atoms selected as components will depend on the ordering of the columns and the rows. The stoichiometric number matrix **v** for a system of chemical reactions has a column for each independent reaction and a row for each species. These two types of matrices are related by $\mathbf{Av} = \mathbf{0}$, where **0** is a zero matrix. Thus, **v** is a basis for the null space of **A**. It is necessary to say "a basis for" because **A** and **v** can be written in various ways.

When the pH is specified, the row for hydrogen atoms in the conservation matrix is deleted, and this causes some of the columns to become redundant. In the new conservation matrix $\mathbf{A}'$, the columns apply to reactants (sums of species), rather than species. Thus, a broader overview of the system is obtained. The corresponding stoichiometric number matrix $\mathbf{v}'$ has a column for each independent biochemical reaction and a row for each reactant, rather than each species. These two types of matrices are related by $(\mathbf{v}')^{\mathrm{T}}(\mathbf{A}')^{\mathrm{T}} = \mathbf{0}$, where T indicates the transpose. Thus, a basis for conservation matrix $\mathbf{A}'$ can be calculated from a stoichiometric number matrix $\mathbf{v}'$.

When the concentrations of coenzymes are specified in addition to the pH, the rows in the conservation matrix for coenzymes are deleted, and this causes some of the columns to become redundant (Alberty, 2000). In the new conservation matrix $\mathbf{A}''$, the columns apply to reactants like C_3 that are sums of reactants, rather than reactants. Thus, a still broader overview of the system is obtained. There is a corresponding stoichiometric matrix, as indicated by $(\mathbf{v}'')^{\mathrm{T}}(\mathbf{A}'')^{\mathrm{T}} = 0$. Thus, a basis for conservation matrix $\mathbf{A}''$ can be calculated from a stoichiometric number matrix $\mathbf{v}''$.

2.5. CALCULATIONS OF EQUILIBRIUM COMPOSITIONS OF SYSTEMS

When there are two or more reactions, the calculation of the equilibrium composition requires an iteration (Newton–Raphson) to obtain the equilibrium composition. Krambeck (1991) wrote computer programs for doing this in APL and more recently has translated them into Mathematica$^{\mathrm{R}}$. The program equcalcc requires the conservation matrix for the system and the Gibbs energies or transformed Gibbs energies for the reactants. Since the input is in matrix form, calculations can be made on system of any size. The program equcalcrx requires the stoichiometric number matrix for the system and equilibrium constants or apparent equilibrium constants for a set of independent reactions. This latter program should be used for biochemical-reaction systems involving H_2O as a reactant because of the special convention that the activity of H_2O is taken as unity in dilute aqueous solutions (Alberty, 2001a). These programs are available in digital form on the web (Alberty, 2001d).

3. Applications to Metabolism

A more global view of the thermodynamics of a system of biochemical equations can be obtained by specifying the steady-state concentrations of coenzymes like ATP and ADP. This also makes it possible to study the effects of these steady-state concentrations on equilibrium compositions. When concentrations of coenzymes are specified, the reaction system can be considered as being made up of pseudoisomer groups, each made up of several reactants. The values of apparent equilibrium constants $\mathbf{K}''$ for reactions between pseudoisomer groups can be calculated if the standard transformed Gibbs energies of formation of reactants are known at the desired pH or the apparent equilibrium constants are all known.

Glycolysis involves ten biochemical reactions and 16 reactants. Water is not counted as a reactant in writing the stoichiometric number matrix or the conservation matrix for reasons mentioned above (Alberty, 2001a). Thus, at specified pH there are six components because $C' = N' - R' = 16 - 10 = 6$. From a chemical viewpoint this is a surprise because the reactants involve only C, H, O, N, and P. Since H and O are not conserved at specified pH in dilute aqueous solution, there are only three conservation equations based on elements. Thus, three additional conservation relations arise from the mechanisms of the enzyme-catalysed reactions in glycolysis. When it is assumed that ATP, ADP, P_i, NAD_{ox}, and NAD_{red} are in steady states, the ten reactions of glycolysis are reduced to one ($C_6 = 2C_3$) and the 16 reactants are reduced to 2. The pseudoisomer group C_6 is made up of glucose, G6P, F6P, and FBP. The pseudoisomer group C_3 is made up of the reactants containing three carbon atoms. The calculations of $\Delta_f G''^0(C_6)$ and $\Delta_f G''^0(C_3)$ are accomplished using eqns (13) and (14). The number $N_{cj}(i)$ of coenzymes j involved in reactant i is most easily calculated by using a Gaussian reduction of the conservation matrix for glycolysis. To do this, glucose, ATP, ADP, P_i, NAD_{ox}, and NAD_{red} are put first in the columns, followed by the rest of the reactants, ending with pyruvate. The calculation of the equilibrium composition in glycolysis only involves solving a quadratic equation. The equilibrium concentrations of all the $16-5=11$ reactants can then be calculated. The equilibrium composition has been calculated in this way for the first five reactions of glycolysis at 298.15 K, pH 7, 0.25 M ionic strength, [ATP] = 0.0001 M, and [ADP] = 0.01 M. This only used the quadratic formula. Then, the calculation was repeated for [ATP] = 0.01 M and [ADP] = 0.01 M. The conclusion (Alberty, 2001e) was "It is perhaps surprising that raising the concentration of ATP by a factor of 100 makes so little difference."

As shown in eqns (17–19), the pyruvate dehydrogenase reaction plus the citric acid cycle can be reduced to three reactions and five reactants. More complicated metabolic reaction systems can be studied in the same way.

4. Use of Semigrand Partition Functions

The grand canonical partition function of Gibbs applies to a one-component system at specified T and V that is in contact with a large reservoir of the one species through a membrane. Thus, the chemical potential of this species in the system has the same value as in the large reservoir. After Gibbs, statistical mechanics also used semigrand partition functions for systems containing two or more species. In a system represented by a semigrand partition function, one or more species are available through a semipermeable membrane while the other species are confined to the system. In contrast with the grand canonical partition function, the pressure on the system can be held constant rather than the volume. This kind of semigrand partition function corresponds with a system of enzyme-catalysed reactions in contact with a reservoir of hydrogen ions at a specified pH.

Statistical mechanics is often thought of as a way to predict the thermodynamic properties of molecules from their microscopic properties such as masses of atoms, bond distances, vibration frequencies, etc. Of course, for aqueous solutions of large molecules this is not currently possible. But statistical mechanics provides more than this because it provides a complementary way of looking at thermodynamics. The statistical mechanical partition function for a system contains all the thermodynamic information on

a system. The transformed Gibbs energy G' for a biochemical reaction system at specified pH is given by

$$G' = -kT \ln \Gamma', \tag{20}$$

where Γ' is the semigrand partition function and k is the Boltzmann constant (R/N_A, where N_A is the Avogadro constant). The further transformed Gibbs energy G'' for a biochemical system at specified pH and specified concentrations of certain coenzymes is given by

$$G'' = -kT \ln \Gamma'', \tag{21}$$

where Γ'' is the corresponding semigrand partition function. A semigrand partition function can be written for a biochemical reaction system, and so G'' can be calculated as indicated. All the remaining thermodynamic properties of the system can be calculated by taking partial derivatives of G'' or Γ''.

The forms of semigrand partition functions for biochemical reactions systems can be illustrated, starting with an aqueous solution of a weak acid and its basic form at a specified pH (Alberty, 2001b, c). The semigrand partition function Γ' for this system is given by

$$\begin{aligned}\Gamma'(T, P, \mathrm{pH}, N'_{iso}) &= \{\exp(-\beta\mu_1)\exp(-N_{\mathrm{H}1}\ln(10)\mathrm{pH}) \\ &\quad + \exp(-\beta\mu_2)\exp(-N_{\mathrm{H}2}\ln(10)\mathrm{pH})^{N'_{iso}} \\ &= \{\exp(-\beta\mu'_1) + \exp(-\beta\mu'_2)^{N'_{iso}},\end{aligned} \tag{22}$$

where μ_1 and μ_2 are the chemical potentials of species 1 and 2 and $\mu_i' = \mu_i - N_{\mathrm{H}i}\mu(\mathrm{H}^+)$. $N_{\mathrm{H}1}$ and $N_{\mathrm{H}2}$ are the numbers of hydrogen atoms in these two species. N_{iso}' is the number of molecules in the pseudoisomer group. In statistical mechanics, it is customary to use numbers of molecules N_i rather than amounts $n_i = N_i/N_A$, and to use $\beta = 1/kT$, where k is the Boltzmann constant. Equation (22) is the form of the partition function that applies at zero ionic strength. This sum of exponential terms for a pseudoisomer group may contain many terms and is really a Laplace transform. Statistical mechanics shows the advantages in putting as many intensive properties as possible in the exponential terms.

Substituting eqn (22) into eqn (20) yields

$$\begin{aligned}G' &= -N'_{iso}kT \ln\{\exp(-\beta\mu_1)\exp(-N_{\mathrm{H}1}\ln(10)\mathrm{pH}) \\ &\quad + \exp(-\beta\mu_2)\exp(-N_{\mathrm{H}2}\ln(10)\mathrm{pH})\} \\ &= -N'_{iso}kT \ln\{\exp[-\beta\mu'_1] \\ &\quad + \exp[-\beta\mu'_2]\} = N'_{iso}\mu'_{iso}.\end{aligned} \tag{23}$$

The transformed chemical potential for the pseudoisomer group (that is HA plus A^-) is given by

$$\mu'_{iso} = -kT \ln\{\exp[-\beta\mu'_1] + \exp[-\beta\mu'_2]\}, \tag{24}$$

which can also be written as

$$\mu'_{iso} = \mu'^0_{iso} = -kT \ln[A], \tag{25}$$

where $[\mathrm{A}] = [\mathrm{HA}] + [\mathrm{A}^-]$. Equation (24) leads to

$$\mu'^0_{iso} = -kT \ln\{\exp[-\beta\mu'^0_1] + \exp[-\beta\mu'^0_2]\}. \tag{26}$$

This can be demonstrated by substituting $\mu'_1 = \mu'^0_1 + kT\ln[\mathrm{HA}]$, and $\mu'_2 = \mu'^0_2 + kT\ln[\mathrm{HA}^-]$ into eqn (22) and by using eqn (10).

For a system containing two pseudoisomer groups A and B, the semigrand partition function is given by

$$\begin{aligned}\Gamma'(T,\ P,\ \mathrm{pH}, N'_{isoA},\ N'_{isoB}) \\ = \{\exp(-\beta\mu'_{isoA})\}^{N'_{isoA}} \times \{\exp(-\beta\mu'_{isoB})\}^{N'_{isoB}}.\end{aligned} \tag{27}$$

Substituting this into $G' = -kT\ln\Gamma'$ yields

$$\begin{aligned}G' &= -kT \ln[\{\exp(-\beta\mu'_{isoA})\}^{N'_{isoA}} \\ &\quad \times \{\exp(-\beta\mu'_{isoB})\}^{N'_{isoB}}]. \\ &= N'_{isoA}\mu'_{isoA} + N'_{isoB}\mu'_{isoB}.\end{aligned} \tag{28}$$

If pseudoisomer groups A and B are involved in a reaction

$$A = 2B \tag{29}$$

and there are initially $(N_{\mathrm{iso}A})_0$ molecules of A, the transformed Gibbs energy of the system at

any time in the reaction is given by

$$G' = (N_{isoA})_0 \mu'_{isoA} + (2\mu'_{isoB} - \mu'_{isoA})\xi', \qquad (30)$$

where ξ' is the extent of reaction. The transformed Gibbs energy is minimized at equilibrium, and so

$$\Delta G'/\mathrm{d}\xi' = 2\mu'_{isoB} - \mu'_{isoA} = 0, \qquad (31)$$

which is the equilibrium condition.

At level 3, the semigrand partition function Γ'' for a sum of reactants that are pseudoisomers at specified concentrations of coenzymes is given by a sum of exponential terms raised to a power equal to the number of molecules in the pseudoisomer group.

$$\Gamma'' = \{\exp(-\beta\mu''_1) + \exp(-\beta\mu''_2) + \cdots\}^{N''_{iso}}, \qquad (32)$$

$$G'' = -N''_{iso} kT \ln\{\exp(-\beta\mu''_1) + \exp(-\beta\mu''_2) + \cdots\}. \qquad (33)$$

For glycolysis at specified T, P, pH, [ATP], [ADP], [P_i], [NAD_{ox}], and [NAD_{red}], the semigrand partition function is given by

$$\Gamma'' = \{\exp(-\beta\mu''_6)\}^{N''_6} + \{\exp(-\beta\mu''_3)\}^{N''_3}, \qquad (34)$$

$$G'' = -N''_6\mu''_6 - N''_3\mu''_3. \qquad (35)$$

Note that these two exponential terms are each made up of summations of exponential terms over the C_6 and C_3 reactants, respectively. The standard further transformed Gibbs energies of formation of C_6 and C_3 can be calculated at the desired pH and ionic strength, and the apparent equilibrium constant K'' for $C_6 = 2C_3$ can be calculated. Solving a quadratic equation yields [C_6] and [C_3]. The distribution of reactants within these pseudoisomer groups can then be calculated. The concentrations of species within the reactants can also be calculated. Thus, no thermodynamic information is lost in going to the level 3 calculation. This method can be applied to larger systems by specifying the concentrations of more coenzymes.

5. Discussion

Thus, the use of transformed Gibbs energies leads to the reconceptualization of a thermodynamic system into a system with fewer reactions and fewer reactants. When information about standard Gibbs energies of formation of species is available, thermodynamic calculations at level 1 can show whether reactions written in terms of species can go spontaneously to the right or left for a specified state of the system. It also provides the means to calculate equilibrium compositions in terms of species. When the pH is specified, calculations at level 2 can show whether biochemical reactions can go spontaneously to the right or left for a specified state of the system. It also provides the means to calculate equilibrium compositions in terms of reactants (sums of species). When, in addition, steady-state concentrations of coenzymes are specified, calculations at level 3 can show whether reactions between groups of reactants (like C_6 and C_3) can go spontaneously to the right or left for a specified state of the system, and can yield the equilibrium compositions in terms of pseudoisomer groups like C_6 and C_3. There are no limitations on the size of the systems that can be considered, except that standard transformed thermodynamic properties or apparent equilibrium constants are needed.

When these concepts and equations are applied to glycolysis plus the pyruvate dehydrogenase reaction plus the citric acid cycle, quantitative thermodynamic calculations can be made by working with six reactants, rather than 32, and four reactions, rather than 20. This is what is referred to in the Introduction as a more global view of the thermodynamics of a metabolic system.

This emphasizes the importance of experimental measurements of apparent equilibrium constants of biochemical reactions under clearly specified conditions and the use of these data to make thermodynamic tables or lists of functions that can be used to calculate standard thermodynamic properties of species at specified conditions. Goldberg *et al.* (1993–1999) have critically evaluated the literature data on apparent equilibrium constants and have rated the quality of these measurements. Their compilations show that experimental data are available on about 500 enzyme-catalysed reactions involving about 1000 reactants. As mentioned earlier, a list of

functions has been published (Alberty, 2001d) for 116 reactants so that $\Delta_f G'^0$ at 298.15 K can be calculated at any desired pH in the range 5–9 and any desired ionic strength in the range 0–0.35 M. Many more standard thermodynamic properties of reactants can be calculated from measured apparent equilibrium constants, but rather complicated calculations are required because measurements have been made at various pH values, ionic strengths, and temperatures. It is important to realize that only apparent equilibrium constants K' in the range between about 10^4 and 10^{-4} can be determined experimentally. Many biochemical reactions have apparent equilibrium constants outside this range, and so their apparent equilibrium constants can only be calculated from those of two or more other reactions with K' values in the experimental range. Many more K' values can be calculated from a table of $\Delta_f G'^0$ values than that are required to make the table. Since some biochemical reactions are similar in terms of structure, their apparent equilibrium constants can be inferred from known values of apparent equilibrium constants of other reactions.

The thermodynamics of biochemical reactions at specified pH and specified concentrations of coenzymes can be represented by semigrand partition functions. Partition functions for reactants at specified pH are sums of exponential terms. For a single reactant at level 2, Γ' has a term for each species that is weighted by a factor that is exponential in $N_H(i)$pH. For a mixture of reactants, the partition function Γ' is a product of partition functions each raised to the power of the number of molecules of the reactant. Thus, the partition function for a mixture of reactants is also a summation of exponential terms. For a sum of reactants at level 3, for example C_6, the partition function Γ'' is a sum of exponential terms, with a term for each reactant (for example, glucose, G6P, F6P, and FBP) in the pseudoisomer group. Each term is weighted by a factor that gives the dependence on [ATP] and [ADP] for the pseudoisomer group being discussed. For a mixture of reactants, like C_6 and C_3, $\Gamma''$is a product of two sums of exponentials, each raised to the power of the number of molecules in each pseudoisomer group.

It is well known that thermodynamic properties obey all the rules of calculus, but this paper shows that the thermodynamics of a system of biochemical reactions is even more mathematical in the sense that it involves further Legendre transforms, conservation matrices, stoichiometric number matrices, and semigrand partition functions. Thermodynamic calculations on systems of biochemical reactions become very complicated, but fortunately personal computers can be used to carry out the calculations. The research described here has been greatly benefitted by the use of Mathematica^R.

I am indebted to Dr Robert N. Goldberg and Prof. Irwin Oppenheim for many helpful discussions. I am indebted to NIH for support of this research by award number 2-RO1-GM48358-05A2.

REFERENCES

ALBERTY, R. A. (1992a). Equilibrium calculations on systems of biochemical reactions. *Biophys. Chem.* **42,** 117–131.

ALBERTY, R. A. (1992b). Calculation of the transformed thermodynamic properties of biochemical reactants at specified pH and pMg. *Biophys. Chem.* **43,** 239–254.

ALBERTY, R. A. (1993). Levels of thermodynamic treatment of biochemical reaction systems. *Biophys. J.* **65,** 1234.

ALBERTY, R. A. (1998a). Calculation of standard transformed Gibbs energies and standard transformed entropies of biochemical reactants. *Arch. Biochem. Biophys.* **353,** 116–130.

ALBERTY, R. A. (1998b). Calculation of standard transformed properties of biophysical reactants and standard apparent reduction potentials of half reactions. *Arch. Biochem. Biophys.* **358,** 25–39.

ALBERTY, R. A. (2000). Calculation of equilibrium compositions of large systems of biochemical reactions. *J. Phys. Chem.* **104B,** 4807–4814.

ALBERTY, R. A. (2001a). Calculation of equilibrium compositions of biochemical reaction systems involving water as a reactant. *J. Phys. Chem.* **105B,** 1109–1114.

ALBERTY, R. A. (2001b). Biochemical reaction equilibria from the point of view of a semigrand partition function. *J. Chem. Phys.* **114,** 8270–8274.

ALBERTY, R. A. (2001c). Systems of biochemical reactions from the point of view of a semigrand partition function. *Biophys. Chem.* **93,** 1–10.

ALBERTY, R. A. (2001d). BasicBiochemData Package in MathSource Electronic Library: http://www.mathsource.com/cgi-bin/msitem? 0211–622.

ALBERTY, R. A. & GOLDBERG, R. N. (1992). Calculation of thermodynamic formation properties for the ATP series at a specified pH and pMg. *Biochemistry* **31,** 10610–10615.

CALLEN, H. B. (1985). *Thermodynamics and an Introduction to Thermostatics*. New York: Wiley.

GOLDBERG, R. N. (1999). Thermodynamics of enzyme-catalyzed reactions: Part 6. 1999 update. *J. Phys. Chem. Ref. Data* **28,** 931–965.

GOLDBERG, R. N., TEWARI, Y. B., BELL, D., FAZIO, D. K., & ANDERSON, E. (1993). Thermodynamics of enzyme catalyzed reactions: Part 1. Oxidoreductases. *J. Phys. Chem. Ref. Data* **22,** 515–582.

GOLDBERG, R. N. & TEWARI, Y. B. (1994). Thermodynamics of enzyme catalyzed reactions: Part 2. Transferases. *J. Phys. Chem. Ref. Data* **23,** 547–617.

GOLDBERG, R. N. & TEWARI, Y. B. (1994). Thermodynamics of enzyme catalyzed reactions: Part 3. Hydrolases. *J. Phys. Chem. Ref. Data* **23,** 1035–1103.

GOLDBERG, R. N. & TEWARI, Y. B. (1995). Thermodynamics of enzyme catalyzed reactions: Part 4. Lyases. *J. Phys. Chem. Ref. Data* **24,** 1669–1698.

GOLDBERG, R. N. & TEWARI, Y. B. (1995). Thermodynamics of enzyme catalyzed reactions: Part 5. Isomerases and ligases. *J. Phys. Chem. Ref. Data* **24,** 1765–1801.

KRAMBECK, F. J. (1991). In: *Chemical Reactions in Complex Systems* (Krambeck, F. J. & Sapre, A. M., eds). NewYork: Van Nostrand Reinhold.