strength (in terms of $\Delta U$) such as tetrahydrothiophene–$I_2$[5] and TMA–$SO_2$ should behave so differently in this respect.

We believe that further progress toward understanding solvent effects on the spectral parameters of CT complexes will depend upon two factors: (1) the acquisition of additional accurate data for both weak and strong complexes in the vapor phase; and (2) development of methods for the study of particularly weak complexes in solution, which may remove possible inadequacies in the interpretation of the solution data.[20] The present study demonstrates the value of a combination of spectral and nonspectral methods to obtain reliable spectral parameters in the vapor phase. For systems in which donor, acceptor and complex are not very volatile, additional techniques must be used. One new method, employing a mixture of polyiodides as a constant iodine activity source,[21] has been successfully used to study the diethyl ether–iodine adduct; the technique is apparently applicable to systems in which either very weak or moderately strong complexes are present.

**Acknowledgment.** This work was supported in part by National Science Foundation Grant No. GP-8029.

(20) P. J. Trotter and M. W. Hanna, *J. Amer. Chem. Soc.*, **88**, 3724 (1966).

(21) J. Childs, J. Grundnes, and S. D. Christian, in preparation.

# Statistical Theory of Cooperative Binding to Proteins. The Hill Equation and the Binding Potential[1]

**Henry d'A. Heck**

*Contribution from the Department of Chemistry, University of California, Berkeley, California 94720.* *Received May 29, 1970*

**Abstract:** The Hill equation and the binding potential are useful methods of describing the cooperative binding of ligands to proteins. Starting from the formal theory of solutions developed by McMillan and Mayer, statistical mechanical versions of both of these classical expressions are here derived. The fractional occupation of protein sites by ligand is expanded in powers of protein concentration, and is used to derive explicit expressions for the effects of intermolecular (protein–protein) interactions on both the Hill equation and the binding potential. The theory leads to a new definition of the apparent free energy of interaction between sites on a single molecule in terms of the free energy of a ligand-transfer reaction between two macromolecules, and provides insight into the significance of the slopes of Hill plots in terms of ligand-transfer processes. The results indicate that, in most cases, the Hill plot parameters may be expected to be influenced to only a minor or negligible extent by intermolecular forces.

The object of this paper is to discuss an empirical equation of A. V. Hill[2] and the binding potential of Wyman[3] from the point of view of McMillan–Mayer[4] solution theory. The Hill equation, which has often been used to describe binding of ligands to proteins,[5] can be written as

$$\bar{Y}_i/(1 - \bar{Y}_i) = Ka_i^n \qquad (1)$$

where $\bar{Y}_i$ is the fractional occupation of $i$ sites by ligand $i$ at ligand activity $a_i$. The exponent $n$ is, in general, a function of $a_i$, but is usually found to be essentially constant over a fairly wide range around the midpoint of the binding curve.[5]

As originally presented by Hill, this equation was a partially successful attempt to describe the cooperative binding of oxygen to hemoglobin. Its modern importance is derived from the demonstration that the Hill equation yields useful thermodynamic information about homotropic[6] reactions, of whatever origin, in *any* system.[5]

If logarithms are taken of both sides of this equation, the resulting expression forms the basis for the well-known Hill plot.[5] Two important quantities can be extracted from such a plot: (1) the minimum value for the decrease in work per site required to saturate the macromolecule with ligand, which results from cooperative interactions among the sites; and (2) the slope of the Hill plot, $n$, at the midpoint of titration, which is a measure of the cooperativity of the binding reaction. The two quantities are closely related, and, in fact, one is a function of the other.[5] The determination of the first of these two quantities from a Hill plot and its connection with the second are not based on the assumption of a particular cooperative model, but arise from general thermodynamic considerations.[5]

In the following sections, the methods of statistical mechanics are used to investigate more closely the significance of the two quantities which are derived from Hill plots. The approach taken is, in essence, an extension of the elegant theory of protein solutions pub-

(1) This work was supported by Grant No. 5-ROI-AM13164-02 from the National Institute of Arthritis and Metabolic Diseases, U. S. Public Health Service.

(2) A. V. Hill, *J. Physiol. (London)*, 40, ivP (1910).

(3) J. Wyman, *J. Mol. Biol.*, 11, 631 (1965).

(4) W. G. McMillan and J. E. Mayer, *J. Chem. Phys.*, 13, 276 (1945).

(5) J. Wyman, *Advan. Protein Chem.*, 19, 223 (1964).

(6) Homotropic is used for interactions between sites which bind the same type of ligand.

lished by T. L. Hill,[7,8] and is thus quite general. The dependence of the Hill plot parameters on intermolecular (protein–protein) interactions is derived. In addition, a statistical mechanical version of the binding potential is presented which includes explicitly the influence of intermolecular interactions on this function. In the limit of infinitely dilute protein solutions, the equation for the microscopic binding potential is formally identical with the corresponding function for the macroscopic thermodynamic case.[3]

## The Hill Equation in the McMillan–Mayer Theory

We begin with some results of the McMillan–Mayer theory for polyatomic multicomponent solutions in order to define briefly the notation to be employed in later sections.

We restrict ourselves to solutions containing one type of protein and one type of solvent ion or molecule whose binding to the protein is considered explicitly. The absolute activity of this single solvent species is denoted by $\lambda = \exp(\mu/kT)$. The activities of the remaining solvent species are signified by $\Lambda$. Let $z_s = \exp(\mu_s/kT)$ be the absolute activity of a protein molecule, with $t$ sites, which contains $s$ bound ligand molecules. We take the point of view that a given protein molecule to which $s$ ligand molecules are bound has accessible to it the full set of energy states associated with all possible distributions of the $s$ ligand molecules on the sites, that is, that the $t!/[s!(t-s)!]$ protein subspecies are in "isomeric" equilibrium.

It follows from the definition of chemical potential that, at equilibrium

$$z_s = \lambda^s z \tag{2}$$

where $z$ is the absolute activity of a protein molecule with $s = 0$.

With these definitions, the equivalent of the grand canonical ensemble partition function for a multicomponent solute in an osmotic system of volume $V$ can be written as [7,9,10]

$$\exp(\pi V/kT) = \sum_{\mathbf{m}\geq 0}\left[\prod_s \frac{(H_s z\lambda^s)^{m_s}}{m_s!}\right] \times \int \exp\left[\frac{-w^{(\mathbf{m})}(\{\mathbf{m}\}, \lambda, \Lambda, z=0)}{kT}\right] \mathrm{d}\{\mathbf{m}\} \tag{3}$$

In this equation, $\pi$ represents the osmotic pressure difference across a membrane separating a solution containing solute species and solvent species on one side from one containing solvent species on the other. Only solvent species are capable of permeating the membrane.

$H_s$ is an effective partition function for a single solute molecule of species $s$ in the solvent and includes interactions among the $s$ ligand molecules bound to the protein. For a single molecule of species $s$, $H_s$ is defined as[10]

$$H_s = q(s)/V\gamma_s^\circ(\lambda, \Lambda) \tag{4}$$

where $\gamma_s^\circ$ is the activity coefficient for this species at infinite dilution in the solvent. The influence of solvent is implicit both in the single-molecule partition function, $H_s$, and in the potential of average force on a fixed set of solute molecules $\mathbf{m}$ immersed in the outside solution which contains solvent species only, $w^{(\mathbf{m})}(\{\mathbf{m}\}, \lambda, \Lambda, z = 0)$. The set of coordinates (translational and rotational) specifying the position, orientation, and conformation of each member of this set of solute molecules is denoted by $\{\mathbf{m}\}$.

We hereafter take $w^{(\mathbf{m})}(0)$ to mean $w^{(\mathbf{m})}(\{\mathbf{m}\}, \lambda, \Lambda, z = 0)$ and employ $w_x^{(\mathbf{m})}(0)$ to represent the spatial potential of average force[9] on molecules of the set $\mathbf{m}$, which includes averaging over all rotational configurations in the set, but which depends, in general, on the particular solute species comprising the set. In a fluid, $w_x^{(\mathbf{m})}(0)$ approaches zero when the molecules of the set are widely separated.[9,11]

The fractional occupation of sites is

$$\bar{Y} = \frac{\lambda[\partial \exp(\pi V/kT)/\partial\lambda]_{\Lambda,z,T,V,\gamma^\circ,w}}{tz[\partial \exp(\pi V/kT)/\partial z]_{\lambda,\Lambda,T,V}} \tag{5}$$

$$= \frac{1}{t}\left[\frac{\sum_{s=0}^{t} sH_s\lambda^s}{\sum_{s=0}^{t} H_s\lambda^s} - \rho\left(\frac{\partial B_2^*}{\partial \ln\lambda}\right)_{T,\gamma^\circ,w} - \frac{\rho^2}{2}\left(\frac{\partial B_3^*}{\partial \ln\lambda}\right)_{T,\gamma^\circ,w} - \ldots\right] \tag{6}$$

where $\rho$ is the total density of solute molecules and $B_2^*$ and $B_3^*$ are, respectively, second and third virial coefficients for the solute.[10]

From eq 6 we obtain directly for the Hill equation

$$\frac{\bar{Y}}{1-\bar{Y}} = \frac{\sum_s sH_s\lambda^s - \sum_s H_s\lambda^s(\rho\partial B_2^*/\partial \ln\lambda + \ldots)}{\sum_s (t-s)H_s\lambda^s + \sum_s H_s\lambda^s(\rho\partial B_2^*/\partial \ln\lambda + \ldots)} \tag{7}$$

In the limit $\rho \to 0$

$$\frac{\bar{Y}}{1-\bar{Y}} = \frac{\sum_s sH_s\lambda^s}{\sum_s (t-s)H_s\lambda^s} \tag{8}$$

Equation 7 gives the first-order corrections to the Hill equation for finite protein concentrations. We shall discuss the importance of these and other corrections in a later section.

(7) T. L. Hill, *J. Chem. Phys.*, **23**, 623 (1955).
(8) T. L. Hill, *ibid.*, **23**, 2270 (1955).
(9) T. L. Hill, "Statistical Mechanics." McGraw-Hill, New York, N. Y., 1956, pp 262–285.
(10) T. L. Hill, "An Introduction to Statistical Thermodynamics," Addison-Wesley, Reading, Mass., 1960, pp 353–362.

(11) It should, perhaps, be stressed in order to avoid confusion that the present approach differs from that taken by T. L. Hill.[7] The potential of average force employed in the present paper is related to potentials of average force in the Hill theory by

$$w^{(\mathbf{m})}(0) = \mathcal{W}_{\mathbf{m}}(0) + \sum_s [W^{(1)}((1_s),0) + \ldots + W^{(1)}((m_s),0)] \quad (\alpha)$$

The $\mathcal{W}_{\mathbf{m}}(0)$ are potentials of average force which vanish without integration over rotational coordinates when the solute molecules are widely separated in a fluid. The configuration integrals in eq 43 are related to those of the Hill theory according to

$$\int_V \exp[-w_x^{(\mathbf{m})}(0)/kT]\mathrm{d}\{\mathbf{m}\}_x = \int \mathrm{d}\{\mathbf{m}\}_\theta \prod_{s,i_s} \exp[-W^{(1)}((i_s),0)/kT] \int \exp[-\mathcal{W}_{\mathbf{m}}(0)/kT]\mathrm{d}\{\mathbf{m}\}_x \quad (\beta)$$

**Apparent Interaction Energy**

It may be observed from eq 8 that, as $\lambda \rightarrow 0$

$$\ln [\bar{Y}/(1 - \bar{Y})] = \ln (H_1/tH_0) + \ln \lambda \qquad (9)$$

On the other hand, as $\lambda \rightarrow \infty$

$$\ln [\bar{Y}/(1 - \bar{Y})] = \ln (tH_t/H_{t-1}) + \ln \lambda \qquad (10)$$

The slope of the Hill plot at the two extremities is unity.

These two limiting equations define the asymptotes of the Hill plot. The vertical distance between the asymptotes at any $\lambda$ is proportional to the minimum value for the decrease in work per site required to saturate the macromolecule with ligand,[5] and is termed the apparent interaction energy. When the final asymptote lies above the initial one, this energy is conventionally taken to be positive.[5] From eq 9 and 10, this quantity is

$$\Delta F_I^\circ = kT \ln t^2 \frac{H_0 H_t}{H_1 H_{t-1}} \qquad (11)$$

$$= kT \ln t^2 \frac{q_0 q_t}{q_1 q_{t-1}} \frac{\gamma_1^\circ \gamma_{t-1}^\circ}{\gamma_0^\circ \gamma_t^\circ} \qquad (12)$$

It is evident that the ratio of partition functions in eq 11 corresponds to the equilibrium constant in solution for the transfer of a ligand molecule between two protein molecules according to

$$P_1 + P_{t-1} \rightleftharpoons P_0 + P_t \qquad (13)$$

where $P_s$ refers to a protein species with $s$ bound ligand molecules.

The entropy change for reaction 13 will involve a configurational contribution associated with the possible ways in which the given numbers of ligand molecules on each protein molecule can be distributed among $t$ sites. For the reaction as written, the configurational entropy change is

$$\Delta S^\circ_{\text{config}} = k \ln 1/t^2 \qquad (14)$$

Thus, if the net free energy change for this transfer reaction arising from enthalpic and further entropic effects is zero, then it is clear from eq 12 and 14 that $\Delta F_I^\circ$ will vanish, except for the contribution of activity coefficients at infinite dilution in the solvent. We may deduce, therefore, if the ratio of the latter is essentially equal to unity, that the apparent interaction energy represents just the *negative* of the net free energy change, apart from configurational entropy factors, for the reaction depicted in (13).

The interpretation of eq 11 in terms of reaction 13 thus leads to a new definition of the apparent interaction energy. This quantity can be directly related to the free energy of a specific ligand transfer reaction. It should be noted that an elementary rearrangement of eq 11, on which the present interpretation is based, leads to an equation which is formally identical with one derived by Wyman[12] from classical thermodynamic considerations. This shows that the present interpretation of the apparent interaction energy is based as well in classical as in statistical thermodynamics.

It is, incidentally, easy to prove that if the sites are independent but unequal, and each is capable of binding one ligand molecule, then the apparent interaction energy will be negative, and the final asymptote of the Hill plot will lie below the initial one. To prove this, let each protein molecule have $\tau_1$ sites of type 1, $\tau_2$ of type 2, ..., $\tau_m$ of type $m$. Define effective partition functions for a *site* of type $j$ to which zero or one ligand molecule is bound as $h_j(0)$ or $h_j(1)$. Then, if there is a total of $s$ ligand molecules bound to a protein molecule, of which $s_1$ are bound to sites of type 1, $s_2$ are bound to sites of type 2, etc., we have that

$$H_s\lambda^s = \sum_s \left[ \prod_{i=1}^{m} \frac{h_i(0)^{\tau_i - s_i} [h_i(1)\lambda]^{s_i}}{(\tau_i - s_i)! s_i!} \right] \qquad (15)$$

where the zero of energy for intramolecular interactions is arbitrary. The sum is over all sets $s = s_1, s_2, \ldots, s_m$ satisfying the restrictions

$$\sum_{i=1}^{m} s_i = s \quad \sum_{i=1}^{m} \tau_i = t \qquad (16)$$

If $s$ is allowed to assume all possible values from 0 to $t$, we obtain from the binomial theorem that

$$\sum_{s=0}^{t} H_s\lambda^s = \prod_{i=1}^{m} \frac{[h_i(0) + h_i(1)\lambda]^{\tau_i}}{\tau_i!} \qquad (17)$$

In the limit $\rho \rightarrow 0$, the corresponding equation to (6) is

$$\begin{aligned} \bar{Y} &= \sum_s s H_s \lambda^s / t \sum_s H_s \lambda^s \\ &= (\partial \ln \textstyle\sum_s H_s \lambda^s / \partial \ln \lambda)_{z,T,V,\gamma^\circ} / t \\ &= \sum_i \tau_i h_i(1)\lambda / [h_i(0) + h_i(1)\lambda] t \end{aligned} \qquad (18)$$

whence

$$\frac{\bar{Y}}{1 - \bar{Y}} = \frac{\sum_i \tau_i h_i(1)\lambda/[h_i(0) + h_i(1)\lambda]}{\sum_i \tau_i h_i(0)/[h_i(0) + h_i(1)\lambda]} \qquad (19)$$

If all the sites are identical as well as independent, this equation reduces to

$$\frac{\bar{Y}}{1 - \bar{Y}} = \frac{h(1)\lambda}{h(0)} = k\lambda \qquad (20)$$

where $k$ is the equilibrium constant for binding to a single site.

From eq 19, when $\lambda \rightarrow 0$

$$\ln \frac{\bar{Y}}{1 - \bar{Y}} = \ln \frac{\sum_i \tau_i h_i(1)/h_i(0)}{t} + \ln \lambda \qquad (21)$$

and, when $\lambda \rightarrow \infty$

$$\ln \frac{\bar{Y}}{1 - \bar{Y}} = \ln \frac{t}{\sum_i \tau_i h_i(0)/h_i(1)} + \ln \lambda \qquad (22)$$

As before, the slope of the Hill plot approaches unity at the two extreme values of $\lambda$.

The difference between the two limiting asymptotes, multiplied by $kT$, gives the apparent interaction energy. This difference is, from eq 21 and 22

$$\ln \frac{t^2}{\left(\sum_i \tau_i \frac{h_i(0)}{h_i(1)}\right)\left(\sum_i \tau_i \frac{h_i(1)}{h_i(0)}\right)} \qquad (23)$$

If the denominator in (23) is greater than the numerator, then the apparent interaction energy is negative. Let $h_j(1)/h_j(0) = k_j$; then, from the definition of $t$, the denominator in (23) is as shown in eq 24 and 25.

If the sites are identical as well as independent, then $k_i = k_j$, and the apparent interaction energy from (23) is zero. Otherwise, the last member in eq 25 is positive definite, which proves the assertion.

(12) J. Wyman, *J. Amer. Chem. Soc.*, **89**, 2202 (1967).

$$(\sum_i \tau_i/k_i)(\sum_i \tau_i k_i) = t^2 + \sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \tau_i\tau_j\left(\frac{k_i}{k_j} + \frac{k_j}{k_i} - 2\right) \quad (24)$$

$$= t^2 + \sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \frac{\tau_i\tau_j}{k_i k_j} \times (k_i - k_j)^2 \quad (25)$$

We shall consider the correction terms in eq 30, which apply at finite protein concentrations, in the following section.

It may be observed from (31) that the equation for the slope of the Hill plot is a ratio of two polynomials in $\lambda$, each of the $(2t - 1)$th degree. There are, in general, two ways in which this ratio can be equal to unity: (1) term-by-term cancellation, *i.e.*, for every value of $s$, the corresponding coefficients of $\lambda^s$ in the numerator and denominator of eq 31 are equal; (2) coincidental cancellation, *i.e.*, for two or more values of $s$, the corresponding coefficients of $\lambda^s$ in the numerator and denominator are unequal, but the total summations for all $s$ are equal. Of these two possibilities, only the first can give rise to a slope which is unity at all values of $\lambda$. We shall concern ourselves with this possibility in the following discussions.

$$n = \frac{t\left[\sum_{s,k}^{(e,o)} (s - 2k)^2 H_k H_{s-k}\lambda^s - \left(\sum_s H_s\lambda^s\right)\left(\rho\frac{\partial^2 B_2^*}{\partial(\ln\lambda)^2} + \frac{\rho^2}{2}\frac{\partial^2 B_3^*}{\partial(\ln\lambda)^2} + \cdots\right)\right]}{\left[\sum_s sH_s\lambda^s - \left(\sum_s H_s\lambda^s\right)\left(\rho\frac{\partial B_2^*}{\partial \ln\lambda} + \cdots\right)\right]\left[\sum_s (t - s)H_s\lambda^s + \left(\sum_s H_s\lambda^s\right)\left(\rho\frac{\partial B_2^*}{\partial \ln\lambda} + \cdots\right)\right]} \quad (30)$$

$$n = \frac{t\sum_{s,k}^{(e,o)} (s - 2k)^2 H_k H_{s-k}\lambda^s}{\sum_{s,k}^{(e,o)} [ts - 2k(s - k)]H_k H_{s-k}\lambda^s + \sum_{s_{even}=2}^{2t-2} (s/4)(2t - s)H_{s/2}^2\lambda^s} \quad (31)$$

The above proof can be easily extended to the case of a heterogeneous population of macromolecules, or of a macromolecule existing in more than one conformational form with differing affinities for ligand, but with no equilibrium between the forms. If, however, there is equilibrium between the forms, then the apparent interaction energy may take on positive values. Indeed, as shown by Wyman,[12] in this instance the final asymptote of the Hill plot must lie above the initial one for the special case where the sites are all identical and independent. The proof presented by Wyman applies in the present microscopic theory; we need only convert the ratio of partition functions in eq 11 to the equivalent form

$$\frac{H_0 H_t}{H_1 H_{t-1}} = \frac{K_t}{K_1 K_{t-1}} \quad (26)$$

in which $K_j$ is the apparent equilibrium constant for binding $j$ molecules of ligand to the macromolecule. Then, for the special case just cited, it has been shown[12] that $(K_1 K_{t-1}/K_t) < t^2$, which means that the apparent interaction energy is positive.

## Slope of the Hill Plot

The slope of the Hill plot is, by definition

$$n = \frac{\mathrm{d}\ln \bar{Y}/(1 - \bar{Y})}{\mathrm{d}\ln\lambda} = \frac{\lambda}{\bar{Y}(1 - \bar{Y})}\frac{\mathrm{d}\bar{Y}}{\mathrm{d}\lambda} \quad (27)$$

It is convenient, in writing a general equation for the slope of the Hill plot, to introduce the following operators. Let

$$\sum_{s,k}^{(t_{even})} = \sum_{s_{odd}=1}^{t-1}\sum_{k=0}^{(s-1)/2} + \sum_{s_{odd}=t+1}^{2t-1}\sum_{k=s-t}^{(s-1)/2} + \sum_{s_{even}=2}^{t}\sum_{k=0}^{(s-2)/2} + \sum_{s_{even}=t+2}^{2t-2}\sum_{k=s-t}^{(s-2)/2} \quad (28)$$

$$\sum_{s,k}^{(t_{odd})} = \sum_{s_{odd}=1}^{t}\sum_{k=0}^{(s-1)/2} + \sum_{s_{odd}=t+2}^{2t-1}\sum_{k=s-t}^{(s-1)/2} + \sum_{s_{even}=2}^{t-1}\sum_{k=0}^{(s-2)/2} + \sum_{s_{even}=t+1}^{2t-2}\sum_{k=s-t}^{(s-2)/2} \quad (29)$$

and denote by $\Sigma_{s,k}^{(e,o)}$ either eq 28 or eq 29 corresponding, respectively, to whether $t$ is even or odd. Then, using eq 7 and 27, $n$ is found by eq 30. In the limit $\rho \to 0$, eq 31 is obtained.

Insight into the physical significance of $n$ is obtained if, for any $s$, we investigate the ratio of coefficients of $\lambda^s$ in the numerator and denominator of eq 31. In both numerator and denominator, the coefficients of $\lambda^s$ are seen to involve sums of products of partition functions, $H_k H_{s-k}$, for single protein molecules to which are bound $k$ and $(s - k)$ molecules of ligand, respectively. Such products are, of course, overall partition functions for two distinguishable protein molecules associated, respectively, with the given numbers of ligand molecules. To illustrate, suppose that $t = 4$ and $s = 3$. Then the ratio of coefficients of $\lambda^3$ in the numerator and denominator of eq 31 is

$$\frac{36H_0H_3 + 4H_1H_2}{12H_0H_3 + 8H_1H_2} = \frac{9\left(\frac{H_0H_3}{H_1H_2}\right) + 1}{3\left(\frac{H_0H_3}{H_1H_2}\right) + 2} = \frac{9 + \left(\frac{H_1H_2}{H_0H_3}\right)}{3 + 2\left(\frac{H_1H_2}{H_0H_3}\right)} \quad (32)$$

In this example, and in general, division of numerator and denominator by any one of the $H_k H_{s-k}$'s leads to an equivalent ratio expressed in terms of equilibrium constants for transfer of one or more ligand molecules from one of the pair of protein molecules to the other member of the pair.

Now, just as in the discussion in the preceding section, the free energies of these transfer reactions will involve contributions arising from differences in configurational entropy between reactants and products. If the net effect of enthalpic and any further entropic contributions to the free energies of the transfer reactions is zero, the magnitude of the equilibrium constants will be just that expected from statistical considerations, except for the effects of activity coefficients. Should the overall effects of the latter on the equilibrium

constants be negligible, or cancel, it will be found for every $s$ that the ratio of coefficients of $\lambda^s$ in the numerator and denominator of eq 31 reduces exactly to unity. Consequently, the Hill plot will have unit slope for all values of $\lambda$.

We shall use the above example as an illustration. The statistically expected value for the equilibrium constant $H_0H_3/H_1H_2$ is

$$\frac{H_0H_3}{H_1H_2} = \frac{1}{6}\frac{\gamma_1^\circ\gamma_2^\circ}{\gamma_0^\circ\gamma_3^\circ} \tag{33}$$

Hence, if the ratio of activity coefficients at infinite dilution in the solvent is essentially equal to unity, eq 32 becomes

$$\frac{9(1/6) + 1}{3(1/6) + 2} = \frac{9 + 6}{3 + 12} = 1 \tag{34}$$

The exact cancellation of numerator and denominator for every value of $s$ corresponds most naturally to the situation of identical and independent sites, although, of course, this is not the only possible interpretation: a combination of positive interactions and unlike sites could produce the same result. On the other hand, if the slope of the Hill plot is not equal to unity at any point, it means that the equilibrium constants for ligand transfer (*e.g.*, eq 33) do not have their statistical values, and that interactions or inequalities among the sites, or both, are manifest.

The maximum value that the slope of the Hill plot may attain corresponds to the hypothetical case of "infinite cooperativity." In this case, there are only two types of solute molecules present in solution: those with no ligand molecules bound, and those for which every site is occupied with ligand. Then it can be easily seen that eq 31 reduces to

$$n = \frac{t^3H_0H_t\lambda^t}{t^2H_0H_t\lambda^t} = t \tag{35}$$

This same result was obtained by Wyman[5] using a classical thermodynamic approach.

We have already shown that if the sites are independent, unequal, and bind a single ligand molecule, then the apparent interaction energy must be negative. It does not necessarily follow from the latter that the slope of the Hill plot must be everywhere less than unity. We can, however, prove that, given the former restrictions, this is indeed so for all finite values of $\lambda$. As in the preceding section, we employ partition functions for sites of type $j$ and denote them by $h_j(0)$ and $h_j(1)$, depending on whether the sites are empty or occupied by ligand, respectively. Then, from eq 18 and 27

$$n = \frac{t\sum_i \tau_i \dfrac{h_i(0)h_i(1)\lambda}{[h_i(0) + h_i(1)\lambda]^2}}{\left(\sum_i \tau_i \dfrac{h_i(0)}{[h_i(0) + h_i(1)\lambda]}\right)\left(\sum_i \tau_i \dfrac{h_i(1)\lambda}{[h_i(0) + h_i(1)\lambda]}\right)} \tag{36}$$

It is readily shown, by application of l'Hospital's rule, that eq 36 approaches unity in the limits $\lambda \to 0$ and $\lambda \to \infty$. To show that $n$ is less than unity at all finite values of $\lambda$, we need only prove that

$$\left(\sum_i \tau_i \frac{h_i(0)}{[h_i(0) + h_i(1)\lambda]}\right)\left(\sum_i \tau_i \frac{h_i(1)\lambda}{[h_i(0) + h_i(1)\lambda]}\right) - t\sum_i \tau_i \frac{h_i(0)h_i(1)\lambda}{[h_i(0) + h_i(1)\lambda]^2} > 0 \tag{37}$$

The fractions of free and occupied sites of type $j$ are, respectively

$$f_j(0) = h_j(0)/[h_j(0) + h_j(1)\lambda] \tag{38}$$

$$f_j(1) = h_j(1)\lambda/[h_j(0) + h_j(1)\lambda] \tag{39}$$

Then, from the definition of $t$, we may write the left-hand side of eq 37 as

$$\left[\sum_i \tau_i f_i(0)\right]\left[\sum_i \tau_i f_i(1)\right] - \left(\sum_i \tau_i\right)\left[\sum_i \tau_i f_i(0) f_i(1)\right] = \sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \tau_i\tau_j[f_i(0) - f_j(0)][f_j(1) - f_i(1)] = \sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \tau_i\tau_j[f_i(0) - f_j(0)]^2 \tag{40}$$

where the last term follows from the fact that $f_k(0) + f_k(1) = 1$.

If all of the sites are identical as well as independent, then $f_i(0) = f_j(0)$ for all $i$ and $j$, and eq 40 is equal to zero, which means that the slope is equal to unity. Otherwise, eq 40 is necessarily greater than zero, which constitutes the proof.

In the event that the protein molecules can assume more than one conformational form, and equilibrium exists between the forms, the slope of the Hill plot may become greater than unity. In fact, if the sites are identical and independent, the slope of the Hill plot must be greater than unity, as shown by Watts-Tobin.[13] The proof of Watts-Tobin can be applied in the present instance. We may define a macroscopic binding polynomial[3,12] for the special case just cited as

$$N_\lambda = \sum_{i=1}^{r} \nu_i(1 + k_i\lambda)^t \tag{41}$$

in which $\nu_i$ is the fraction of macromolecules in the $i$th conformation. The relationship to the present theory is through $k_i = H_1^{(i)}/tH_0^{(i)}$, the intrinsic affinity constant for binding to a single site of that conformation. It follows by the method of Watts-Tobin that the slope of the Hill plot must be everywhere greater than unity.

## Effects of Protein–Protein Interactions

The influence of protein–protein interactions on the apparent interaction energy and slope of the Hill plot can be given explicitly in terms of configuration integrals.

Define

$$c = \sum_s H_s\lambda^s$$

$$e = \sum_s sH_s\lambda^s \tag{42}$$

$$f = \sum_s s^2H_s\lambda^s$$

$$\int_{ss'} = \int_V \exp[-w_x^{(s,s')}(0)/kT]\,d(1s)_x\,d(1s')_x$$

$$\int_{ss's''} = \int_V \exp[-w_x^{(s,s',s'')}(0)/kT] \times d(1s)_x\,d(1s')_x\,d(1s'')_x \tag{43}$$

(13) R. J. Watts-Tobin, *J. Mol. Biol.*, **23**, 305 (1967).

etc.

$$C_1 = cV$$
$$C_2 = \sum_{s,s'} H_s H_{s'} \lambda^{s+s'} \int_{ss'} \quad (44)$$
$$C_3 = \sum_{s,s',s''} H_s H_{s'} H_{s''} \lambda^{s+s'+s''} \int_{ss's''}$$

etc.

$$E_1 = eV$$
$$E_2 = \sum_{s,s'} (s + s') H_s H_{s'} \lambda^{s+s'} \int_{ss'} \quad (45)$$
$$E_3 = \sum_{s,s',s''} (s + s' + s'') H_s H_{s'} H_{s''} \lambda^{s+s'+s''} \int_{ss's''}$$

etc.

$$F_1 = fV$$
$$F_2 = \sum_{s,s'} (s + s')^2 H_s H_{s'} \lambda^{s+s'} \int_{ss'} \quad (46)$$
$$F_3 = \sum_{s,s',s''} (s + s' + s'')^2 H_s H_{s'} H_{s''} \lambda^{s+s'+s''} \int_{ss's''}$$

etc.

With definitions similar to these, the partial derivatives appearing in eq 6 have already been evaluated.[7] Those appearing in eq 30 are easily obtained by an identical procedure. The results are

$$\frac{\partial B_2^*}{\partial \ln \lambda} = \frac{e}{c^2}\left(\frac{C_2}{C_1} - \frac{E_2}{2E_1}\right) \quad (47)$$

$$\frac{\partial B_3^*}{\partial \ln \lambda} = \frac{e}{c^3}\left(\frac{E_3}{3E_1} - \frac{2E_2C_2}{E_1C_1} - \frac{C_3}{C_1} + \frac{4C_2^2}{C_1^2} + \frac{E_2C_1}{E_1} - 2C_2\right) \quad (48)$$

$$\frac{\partial^2 B_2^*}{\partial(\ln \lambda)^2} = \frac{e}{c^2}\left(\frac{2E_2}{C_1} - \frac{3E_1C_2}{C_1^2} + \frac{F_1C_2}{E_1C_1} - \frac{F_2}{2E_1}\right) \quad (49)$$

$$\frac{\partial^2 B_3^*}{\partial(\ln \lambda)^2} = \frac{e}{c^3}\left(\frac{2E_3}{C_1} - \frac{F_3}{3E_1} + \frac{F_1C_3}{E_1C_1} - \frac{4E_1C_3}{C_1^2} + \frac{2F_2C_2}{E_1C_1} + \frac{2E_2^2}{E_1C_1} - \frac{16E_2C_2}{C_1^2} - \frac{4F_1C_2^2}{E_1C_1^2} + \frac{20E_1C_2^2}{C_1^3}\right) \quad (50)$$

If the terms on the right-hand sides of eq 47–50 are assembled over common denominators, it will be seen that the numerators are polynomials in $\lambda$ which include, as coefficients in every term, differences between configuration integrals for clusters which are composed of the same numbers of protein molecules, but which contain different numbers of bound ligand molecules. It follows, therefore, that if the configuration integrals, defined in eq 43, are the same for all values of $s$, $s'$, etc., that is, are independent of the numbers of ligand molecules bound to the protein molecules which comprise the clusters, then the partial derivatives in eq 47–50 must vanish for all values of $s$ and $\lambda$. This result is, of course, to be expected in view of the close connection between virial coefficients and configuration integrals.[9] For this case, therefore, protein–protein interactions can have no effect on the extent of ligand binding, and, therefore, cannot influence either the apparent interaction energy or the slope of the Hill plot.

On the other hand, if the differences between the configuration integrals in eq 47–50 are not zero for any values of $s$, then the partial derivatives will exist at at least some values of $\lambda$. Since, however, the partial derivatives involve differences between configuration integrals, it seems reasonable to expect that the correction terms in eq 7 and 30 will, in most cases, be small. Indeed, T. L. Hill[7] has gathered evidence which indicates that, for the binding of singly charged ions to two different proteins, the correction embodied in eq 47 is of minor or negligible importance. This should be true in general whenever the protein concentration is itself low. An exception could occur if the binding of ligand were accompanied by a large change in the state of ionization of the protein, or if the ligand itself were highly charged, although it should generally be possible to choose conditions such that even these influences are negligible.

**Relation to the Binding Potential**

Wyman[3] has pointed out that the binding potential bears a certain resemblance to the grand canonical partition function of statistical mechanics. The exact connection between the two can be derived readily if we define the binding potential as that function whose partial derivative with respect to the chemical potential of a component yields the amount of that component in chemical combination with the macromolecule.[14] Then, by rearrangement and integration of eq 6, the binding potential is easily shown to be

$$\Pi = kT\left[\ln\left(1 + \frac{H_1}{H_0}\lambda + \frac{H_2}{H_0}\lambda^2 + \ldots + \frac{H_t}{H_0}\lambda^t\right) - \rho(B_2^*{}_{(\lambda)} - B_2^*{}_{(\lambda=0)}) - (\rho^2/2)(B_3^*{}_{(\lambda)} - B_3^*{}_{(\lambda=0)}) - \ldots\right] \quad (51)$$

where $H_j/H_0$ is the equilibrium constant in solution for the reaction

$$P_0 + jX \rightleftharpoons PX_j \quad (52)$$

In the limit $\rho \to 0$

$$\Pi = kT \ln\left(\sum_{s=0}^{t} H_s\lambda^s/H_0\right) \quad (53)$$

Equation 53 is formally identical with Wyman's equation for the binding potential.[15] The correction terms in eq 51 show the effects of finite protein concentrations on this function; the terms involve differences between virial coefficients for protein molecules at ligand activities $\lambda$ and 0. Explicit expressions for the virial coefficients, which give their dependence on $\lambda$, are[7]

$$B_2^* = -\frac{1}{2c}\left(\frac{C_2}{C_1} - C_1\right) \quad (54)$$

$$B_3^* = -\frac{1}{3c^2}\left(\frac{C_3}{C_1} - \frac{3C_2^2}{C_1^2} + 3C_2 - C_1^2\right) \quad (55)$$

etc. The arguments presented by Wyman[3] regarding the factorability of the binding polynomial, which is the macroscopic analog of the polynomial in eq 53, and the

(14) H. d'A. Heck, *ibid.*, 50, 703 (1970).
(15) See eq 19 in ref 3.

relation of factorability to statistical binding, apply here as well as in the macroscopic thermodynamic case.

In the event that the binding of more than one type of ligand molecule to the protein is considered explicitly, the binding potential assumes the form

$$\Pi = kT \ln \left( \sum_{\mathbf{s}} \lambda_x^{s_x} \lambda_y^{s_y} \ldots H_{\mathbf{s}}/H_0 \right) \qquad (56)$$

where $\mathbf{s}$ refers to $s_x$, $s_y$, ... molecules of the different ligands, $H_{\mathbf{s}}$ is the effective partition function for a single molecule of protein to which $\mathbf{s}$ molecules of ligand are bound, and $\lambda_i$ is the absolute activity of ligand species $i$. If $H_{\mathbf{s}}/H_0$ is interpreted as an equilibrium constant for the reaction

$$P_0 + s_x X + s_y Y + \ldots \qquad PX_{s_x}Y_{s_y}\ldots \qquad (57)$$

the polynomial in 56 becomes formally identical with the macroscopic binding polynomial for the case where the number of ligands is greater than one.[16]

(16) See expression 20 of ref 3.

# A Theoretical Study of the Optical Rotatory Properties of Poly-L-tyrosine[1]

**Alice K. Chen and Robert W. Woody***

*Contribution from the W. A. Noyes Laboratory of Chemistry, University of Illinois, Urbana, Illinois 61801.* *Received July 2, 1970*

**Abstract**: The optical rotatory properties of poly-L-tyrosine (PLT) are very different from those of simple polypeptides, and the helix sense of PLT remains uncertain. Calculations of rotational strengths for peptide and side-chain transitions have been performed for four conformations of low potential energy. Two of these are right handed and two are left handed, the conformations of a given screw sense differing in side-chain conformation. The results of these calculations were shown to be qualitatively independent of chain length and of small variations in side-chain conformation. Comparison of calculated and experimental circular dichroism curves indicates that only one of the four conformations is consistent with experiment, and that is the conformation we have denoted as RA. We conclude that poly-L-tyrosine forms a *right-handed* helix. Calculations of Ooi, *et al.*, yield a minimum potential energy for a right-handed helix, but predict a different side-chain conformation (R1). We propose that the RA conformation is lower than the R1 in free energy because of greater side-chain flexibility and hence more positive entropy.

The optical rotatory properties of simple helical polypeptides are now well understood, and unambiguous assignments of helix sense can be made on the basis of optical rotatory dispersion[2] (ORD) or circular dichroism[3] (CD). By simple polypeptides, we mean those which have alkyl side chains, side chains with weak chromophores (carboxyl, ester, or amino groups), or aromatic groups beyond the $\gamma$ carbon (*e.g.*, poly-$\gamma$-benzyl-L-glutamate). Polypeptides such as poly-L-tyrosine (PLT) and poly-L-phenylalanine, which have aromatic rings attached to the $\beta$ carbon, have drastically altered optical rotatory properties, and the helix sense remains uncertain.[4]

The ORD behavior of PLT in the visible region was observed to be anomalous.[5,6] Fasman[7] showed that copolymers of L-tyrosine and L-glutamic acid of varying composition showed a smooth and nearly linear change in the Moffitt-Yang[8] $b_0$ parameter from *ca.* +500, characteristic of PLT, to *ca.* −600, characteristic of poly-L-glutamic acid (PGA). This indicated that L-tyrosine residues fit into the right-handed PGA helix and implied but did not conclusively prove that PLT forms a right-handed $\alpha$ helix.

Subsequently, Fasman *et al.*,[9] extended ORD measurements on PLT down to 227 nm and Beychok and Fasman[10] reported the CD spectrum in the 214–300-nm region. The negative Cotton affect at 224 nm was assigned to the peptide n–$\pi^*$ transition, and because its sign coincided with that of the n–$\pi^*$ Cotton effect in simple polypeptides in the right-handed $\alpha$-helix conformation, a right-handed screw sense of the PLT helix was taken to be confirmed. However, this interpretation can be questioned. The amplitude of the 224-nm Cotton effect in PLT is only about one-third that of the 222-nm band in simple $\alpha$-helical polypeptides.[3] This reduction in magnitude was attributed[10] to band overlap and/or coupling with transitions characteristic of the phenolic side chain. It is possible that coupling with

* Author to whom reprint requests should be addressed at the Department of Chemistry, Arizona State University, Tempe, Ariz. 85281.

(1) (a) Based in part on the Ph.D. Thesis of A. K. Chen, University of Illinois, Urbana, Ill., 1969; presented in part at the 158th National Meeting of the American Chemical Society, New York, N. Y., Sept 1969, Abstract BIOL 275.

(2) (a) P. Urnes and P. Doty, *Advan. Protein Chem.*, **16**, 401 (1961); (b) J. A. Schellman and C. Schellman in "The Proteins," Vol. 2, 2nd ed, H. Neurath, Ed., Academic Press, New York, N. Y., 1964, p 1.

(3) S. Beychok in "Poly-$\alpha$-Amino Acids: Protein Models for Conformational Studies," G. D. Fasman, Ed., Marcel Dekker, New York, N. Y., 1967, p 293.

(4) M. Goodman, G. W. Davis, and E. Benedetti, *Accounts Chem. Res.*, 1, 275 (1968).

(5) A. R. Downie, A. Elliott, and W. E. Hanby, *Nature (London)*, **183**, 110 (1959).

(6) J. D. Coombes, E. Katchalski, and P. Doty, *ibid.*, **185**, 534 (1960).

(7) G. D. Fasman, *ibid.*, **193**, 681 (1962).

(8) W. Moffitt and J. T. Yang, *Proc. Nat. Acad. Sci. U. S.*, **42**, 596 (1956).

(9) G. D. Fasman, E. Bodenheimer, and C. Lindblow, *Biochemistry*, **3**, 1665 (1964).

(10) S. Beychok and G. D. Fasman, *ibid.*, **3**, 1675 (1964).