Biophysical Journal Volume 65 July 1993 164–170

The Binding Capacity Is a Probability Density Function

Enrico Di Cera and Zhen-Qing Chen*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110 USA

ABSTRACT The binding capacity of a system, or equivalently, the fluctuations of the number of ligands bound around the average value defined by the binding isotherm, can be regarded as a probability density function for the chemical potential of the ligand. The first moment of this density function is the mean ligand activity as defined by Wyman and gives the average free energy (in *kT* units) of binding per site. The second moment is directly related to the cooperativity of the system. These and higher moments can be obtained from numerical integration of experimental data in a direct way. An analytical expression for the moment generating function shows that the *N* independent coefficients of the partition function of a system containing *N* sites are uniquely defined by the first *N* moments of the binding capacity.

INTRODUCTION

It is well known from basic principles of statistical thermodynamics that the analysis of the fluctuations of extensive quantities of a system at equilibrium provides much information on the ensemble properties of the system (1). Thermodynamic response functions related to these fluctuations can be accessed experimentally in a direct way. The heat capacity of the system provides information on thermal transitions (1), while light scattering experiments reveal the magnitude of the fluctuations of the number of molecules of a given component due to changes in its chemical potential (2). The thermodynamic description of binding and adsorption is also based on the analysis of fluctuations as determined experimentally. A number of experimental approaches can be used to monitor binding isotherms, the fundamental properties of which have been the subject of many outstanding treatments of binding and cooperativity, especially by Hill (3), Schellman (4), and Wyman (5). The slope of the binding isotherm is a measure of the variance of the distribution of ligated intermediates and is directly linked to the fluctuations of ligated sites around their mean value. This quantity defines the binding capacity (6) and can be measured experimentally in a direct way using the thin-layer cell introduced by Gill (7). It is the purpose of this paper to draw attention to some statistical properties of the binding capacity that can be exploited in practical applications. Consideration of these properties makes the binding capacity a key thermodynamic quantity of grand canonical ensembles, equally if not more important than other quantities accessible experimentally, such as the binding isotherm, that have been the focus of previous thermodynamic discussions (3–5). Here we demonstrate that basic aspects of binding and cooperativity become transparent when a rather intriguing property of the binding capacity is taken into account. The generality of our treatment carries over to any other thermodynamic system whose partition function can be cast in terms of a polynomial of finite order, as typically seen for a grand canonical ensemble.

The partition function for a system containing N sites for a ligand X, whose activity is x, is given by the following polynomial expansion.

$$Z = 1 + A_1x + A_2x^2 + \ldots + A_Nx^N \tag{1}$$

Here x is related to the chemical potential of the ligand, μ, through the familiar relation $kT \ln x = \mu$, while Z stands for the German *Zustandsumme*, or sum-over-states, as originally defined by Planck. Specifically, Z is interpreted in the Guggenheim sense as the partition function of a generalized grand canonical ensemble with independent variables μ, P, and T. This ensemble differs from a grand canonical ensemble in the fact that P, rather than V, is held constant through external constraints (2). The coefficient A_j is the overall equilibrium constant for the transition from the unligated form of the system to its jth ligated intermediate. Finally, it is implicitly assumed in Eq. 1 that the system does not change its aggregation state upon ligation. Once the explicit form of Z is available, all quantities of interest can be arrived at by differentiation. The average number of ligated sites is

$$\bar{X} = \mathrm{d} \ln Z/\mathrm{d} \ln x \tag{2}$$

and is bounded from zero to N. The variance of the distribution of ligated sites (3–5), or else the binding capacity of the macromolecule (6), is given by the following.

$$B = \mathrm{d}\bar{X}/\mathrm{d} \ln x = \mathrm{d}^2\ln Z/\mathrm{d} \ln x^2 \tag{3}$$

The properties of B encapsulate basic principles of thermodynamic stability. For example, B is always positive since an increase/decrease of ligand activity can only increase/decrease the amount of ligand bound. Also, for any pair of thermodynamically conjugate variables, α and β, where α is an extensive quantity (say V, S, or the amount $\bar{Y}$ of a second ligand bound to the system) and β is the conjugate intensity

---

Received for publication 2 December 1992 and in final form 24 February 1993.

*Zhen-Qing Chen is presently at the Department of Mathematics, University of California at San Diego, La Jolla, CA, 92093-0112.

This work was supported by National Science Foundation Grant DMB91-04963. Enrico Di Cera is an Established Investigator of the American Heart Association. This paper is dedicated to the memory of Prof. Stanley J. Gill.

($-P$, T, or μ_Y), the inequality $B_\beta > B_\alpha$ holds (6, 8). Other properties of B are related to cooperativity. The Hill coefficient (3, 9)

$$n_H = \frac{\mathrm{d}\ln[\bar{X}/(N-\bar{X})]}{\mathrm{d}\ln x} = \frac{N}{\bar{X}(N-\bar{X})}B = B/B_0 \quad (4)$$

is a measure of cooperativity and is directly related to B. In a reference system where all sites are alike and independent the Hill coefficient is identically equal to one, regardless of the value of x. For such a system $B = \bar{X}(N-\bar{X})/N = B_0$. Hence, the value of n_H for an arbitrary system is the ratio of the binding capacity of the system with respect to that of a reference system containing the same number of sites. In addition to these basic properties, the function B has a number of statistical properties that have not been pointed out before.

THE BINDING CAPACITY IS A PROBABILITY DENSITY FUNCTION

If measurements of $\bar{X}$ are available, it is always possible to obtain information on the coefficient A_N of the partition function. The Wyman integral equation (Eq. 5)

$$\langle \ln x \rangle = N^{-1}\int_0^N \ln x\,\mathrm{d}\bar{X} = -N^{-1}\ln A_N = \ln x_m \quad (5)$$

can be solved in a rather straightforward way to obtain the mean value of $\ln x$. The quantity $x_m = A_N^{-1/N}$ gives the value of x, where the two limiting configurations with the N sites free or bound have equal probability of occurrence. The logarithm $\ln x_m$ represents the average work (in kT units) done in ligating one site. Solution of the integral Eq. 5 only demands numerical integration of the data. On the other hand, information on the other coefficients of the partition function necessarily requires nonlinear least-squares analysis of the data. There is no other set of relationships such as Eq. 5 which allows, at least in principle, to uniquely determine the coefficients of Z from measurements of $\bar{X}$. Such a set of relationships, if it existed, would provide an important and direct connection between the partition function and experimental measurements.

The importance of the integral Eq. 5 can readily be appreciated by writing down the equivalent form

$$\langle \ln x \rangle = N^{-1}\int_{-\infty}^{\infty} B\ln x\,\mathrm{d}\ln x = -N^{-1}\ln A_N = \ln x_m \quad (6)$$

using the definition of B. The change of variable $y = \ln x$ leads to the following.

$$\langle y \rangle = N^{-1}\int_{-\infty}^{\infty} yB\,\mathrm{d}y \quad (7)$$

This is a rather interesting form, for it suggests that B, divided by N, is the *probability density function* associated with the independent variable y. Calculation of $\ln x_m$ is equivalent to calculation of the first moment of B. All higher moments

$$\langle y^k \rangle = N^{-1}\int_{-\infty}^{\infty} y^k B\,\mathrm{d}y \quad (8)$$

can be arrived at in an analogous way. The existence of a set of relationships such as Eq. 8 draws attention to the fact that measurements of B, or $\bar{X}$, can be used to compute not only $\ln x_m$, but all other moments. Whether such considerable amount of information stored in the function B is of any relevance in a thermodynamic discussion is contingent on the connection between the moments of B and the coefficients of the partition function. In order to demonstrate such a connection, Eq. 8 needs to be solved explicitly, which is a nontrivial problem for $k > 1$. Let us define the moment-generating function

$$G(\omega) = \int_{-\infty}^{\infty} e^{\omega y}B\,\mathrm{d}y \quad (9)$$

Then, by definition

$$\langle y^k \rangle = N^{-1}\left(\frac{\partial^k G(0)}{\partial \omega^k}\right). \quad (10)$$

We will see that the generating function $G(\omega)$ exists for $0 < \omega < 1$. A necessary condition for the existence of $G(\omega)$ is provided by the integral

$$\int_{-\infty}^{\infty} B\,\mathrm{d}y = N \quad (11)$$

which exists and is bounded. No generating function can be associated with $\bar{X}$, since the analogous integral would necessarily diverge. From a mathematical point of view the partition function is a polynomial of degree N with real and positive coefficients. The partition function is also analytical everywhere on the positive real axis, which is the only one of physical significance (1). In view of this fact, it is always possible to cast Z in terms of N factors as follows

$$Z = (1+\alpha_1 x)(1+\alpha_2 x)\ldots(1+\alpha_N x) \quad (12)$$

where α_j is the jth coefficient of the factorization related to the jth root of Z. The coefficients can be real and positive or pairs of complex conjugate. Using this property of Z one can write X and B as a linear combination of N similar terms, i.e.,

$$\bar{X} = \sum_{j=1}^{N} \alpha_j x/(1+\alpha_j x) \quad (13)$$

$$B = \sum_{j=1}^{N} \alpha_j x/(1+\alpha_j x)^2. \quad (14)$$

Hence, the solution of Eq. 9 is itself a linear combination of

N similar terms

$$g(\omega) = \int_{-\infty}^{\infty} e^{\omega y} \frac{\alpha e^y}{(1 + \alpha e^y)^2} \mathrm{d}y \tag{15}$$

and convergence of the integral necessarily demands $0 < \omega < 1$. If α is real, then the transformation $Y = \alpha e^y / (1 + \alpha e^y)$ yields at once

$$\begin{aligned} g(\omega) &= \int_0^1 \alpha^{-\omega} Y^{\omega} (1 - Y)^{-\omega} \mathrm{d}Y \\ &= \alpha^{-\omega} B(1 + \omega, 1 - \omega) = \alpha^{-\omega} \pi\omega \operatorname{cosec} \pi\omega \end{aligned} \tag{16}$$

where $B(1+\omega, 1-\omega)$ is Euler's Beta function (10). If α is complex, then $g(\omega)$ is obtained by properly choosing an integration path in the complex plane. Letting $z = \alpha e^y$ in Eq. 15 one has

$$g(\omega) = \int_{\gamma} \alpha^{-\omega} \frac{z^{\omega}}{(1 + z)^2} \mathrm{d}z \tag{17}$$

where γ denotes integration along the semi-line in the direction of α in the complex plane. Since the function $z^{\omega}/(1 + z)^2$ is holomorphic in the region I_r^R shown in Fig. 1, the path integral along a closed contour C must vanish. Hence,

$$\int_C \alpha^{-\omega} \frac{z^{\omega}}{(1 + z)^2} \mathrm{d}z = \int_{I_1} + \int_{I_2} + \int_{I_3} + \int_{I_4} = 0. \tag{18}$$

The integral I_1 computed in the complex plane along the semi-line in the direction of α for $R \rightarrow \infty$ and $r \rightarrow 0$ is the required solution for $g(\omega)$. The integral I_2 is such that

$$\begin{aligned} I_2 &= \left| \int_{I_2} \alpha^{-\omega} \frac{z^{\omega}}{(1 + z)^2} \mathrm{d}z \right| \leq \int_{I_2} |\alpha^{-\omega}| \frac{|z^{\omega}|}{|1 + z^2|} |\mathrm{d}z| \\ &\leq \int_0^{\pi} |\alpha^{-\omega}| \frac{R^{1+\omega}}{(R - 1)^2} \mathrm{d}\theta = \frac{\pi |\alpha^{-\omega}| R^{1+\omega}}{(R - 1)^2} \end{aligned} \tag{19}$$

and goes to zero for $R \rightarrow \infty$, since $\omega < 1$. The integral I_3 involves the positive real axis only. The integral I_4 is such that

$$\begin{aligned} I_4 &= \left| \int_{I_4} \alpha^{-\omega} \frac{z^{\omega}}{(1 + z)^2} \mathrm{d}z \right| \leq \int_{I_4} |\alpha^{-\omega}| \frac{|z^{\omega}|}{|1 + z|^2} |\mathrm{d}z| \\ &\leq \int_0^{\pi} |\alpha^{-\omega}| \frac{r^{1+\omega}}{(1 - r)^2} \mathrm{d}\theta = \pi |\alpha^{-\omega}| \frac{r^{1+\omega}}{(1 - r)^2} \end{aligned} \tag{20}$$

and goes to zero for $r \rightarrow 0$. The solution for $g(\omega)$ is set by the value of $-I_3$, i.e.,

$$g(\omega) = \int_0^{\infty} \alpha^{-\omega} \frac{x^{\omega}}{(1 + x)^2} \mathrm{d}x = \alpha^{-\omega} \pi\omega \operatorname{cosec} \pi\omega \tag{21}$$

and is identical with that found in the case of α real and positive. The solution for $G(\omega)$ in the general case of N sites is therefore

$$G(\omega) = \pi\omega \operatorname{cosec} \pi\omega \sum_{j=1}^{N} \alpha_j^{-\omega} \tag{22}$$

which provides the analytical form to be used in practical applications.

Once the generating function $G(\omega)$ is known, then all moments $\langle y^k \rangle$ can be obtained by differentiation using Eq. 10. Consider the Taylor expansion of $G(\omega)$ as follows.

$$G(\omega) = f(\omega) h(\omega) \tag{23}$$

$$f(\omega) = \pi\omega \operatorname{cosec} \pi\omega \tag{24}$$

$$h(\omega) = \alpha_1^{-\omega} + \alpha_2^{-\omega} + \ldots + \alpha_N^{-\omega} \tag{25}$$

The Taylor expansion of $f(\omega)$ is as follows.

$$f(\omega) = 1 + \sum_{k=1}^{\infty} \frac{2(2^{2k-1} - 1) |B_{2k}| (\pi\omega)^{2k}}{(2k)!} \tag{26}$$

where B_{2k} is the $2k$th Bernoulli number (10). The Taylor expansion of $h(\omega)$ is

$$\begin{aligned} h(\omega) &= N - \sum_{j=1}^{N} \omega \ln \alpha_j + \sum_{j=1}^{N} \frac{\omega^2 (\ln \alpha_j)^2}{2!} - \sum_{j=1}^{N} \frac{\omega^3 (\ln \alpha_j)^3}{3!} \\ &\quad + \ldots = N - C_1 \omega + C_2 \omega^2/2! - C_3 \omega^3/3! + \ldots \end{aligned} \tag{27}$$

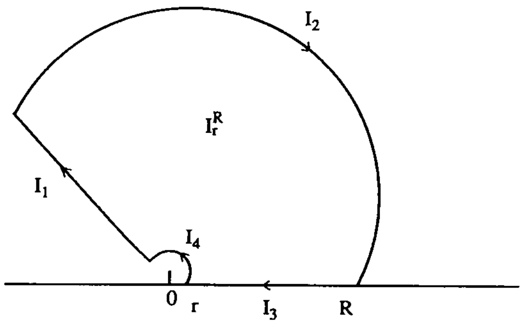

FIGURE 1 Integration path in the complex plane for the integral in Eq. 17. The argument of the integral is a holomorphic function in I_r^R, and therefore the path integral vanishes. The integration path I_1 develops along the semi-line in the direction of α in the complex plane and provides the required solution in the limit $R \rightarrow \infty$ and $r \rightarrow 0$. The integral I_2 vanishes for $R \rightarrow \infty$, and likewise the integral I_4 vanishes for $r \rightarrow 0$. The value of I_1 is therefore equal to the integral I_3 from zero to ∞ along the real axis only.

Hence,

$$\langle y \rangle = -N^{-1}C_1 \tag{28}$$

$$\langle y^2 \rangle = \pi^2/3 + N^{-1}C_2 \tag{29}$$

$$\langle y^3 \rangle = -\pi^2 C_1 - N^{-1}C_3 \tag{30}$$

$$\langle y^4 \rangle = 7\pi^4/15 + 2\pi^2 C_2 + N^{-1}C_4 \tag{31}$$

and so on. The first moment equals ln x_m. The second moment is proportional to the sum of squares of the logarithm of each coefficient α, and in general the kth moment is expressed as the sum of all even/odd powers of each ln α term up to order k, for k even/odd. It is straightforward to see from Eqs. 28–31 that the C values are uniquely defined in terms of the moments and themselves uniquely define the sum of the various powers of ln α_j values. Hence,

THEOREM: *The first N moments of the binding capacity uniquely define the N independent coefficients of the partition function.*

The connection between the statistical properties of the binding capacity and the partition function of the system is now evident.

As an example consider the case of $N = 2$. The partition function and the first two moments of the binding capacity are

$$Z = 1 + A_1 x + A_2 x^2 = (1 + \alpha_1 x)(1 + \alpha_2 x) \tag{32}$$

$$m_1 = \langle \ln x \rangle = -(\ln \alpha_1 + \ln \alpha_2)/2 \tag{33}$$

$$m_2 = \langle \ln^2 x \rangle = \pi^2/3 + (\ln^2 \alpha_1 + \ln^2 \alpha_2)/2 \tag{34}$$

Hence, if m_1 and m_2 are known it follows that

$$\alpha_1 = \exp(-m_1 + \Delta^{1/2}) \tag{35}$$

$$\alpha_2 = \exp(-m_1 - \Delta^{1/2}) \tag{36}$$

$$\Delta = m_2 - m_1^2 - \pi^2/3 = \sigma^2 - \pi^2/3. \tag{37}$$

Note that the value $\sigma^2 = \pi^2/3$ provides the separation for the α values to be real or complex conjugate. We shall see in the next section that this is the value of the variance of the B distribution for a noncooperative system. The desired values of A_1 and A_2 are as follows.

$$A_1 = \alpha_1 + \alpha_2 = 2\exp(-m_1)\cosh(\Delta^{1/2}) \tag{38}$$

$$A_2 = \alpha_1 \alpha_2 = \exp(-2m_1) \tag{39}$$

The values of the first two moments uniquely define the two independent coefficients of the partition function.

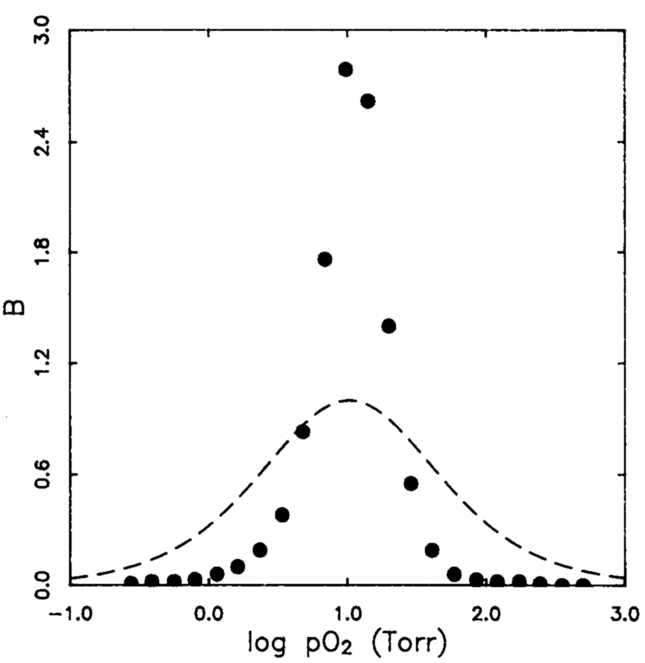

FIGURE 2 Oxygen binding capacity of human hemoglobin ($N = 4$) at pH 6.95 plotted versus the logarithm of the oxygen partial pressure (11). The B distribution of a reference system with $N = 4$ and the same mean ligand activity is given by comparison as a discontinuous line.

THE VARIANCE OF *B* AS A MEASURE OF COOPERATIVITY

The first moment of B contains information on the average free energy of binding (in kT units) per site. This quantity per se provides no information on cooperativity. Cooperativity is reflected by the steepness of the binding curve and hence by the binding capacity. This situation is illustrated in Fig. 2 for the case of oxygen binding to human hemoglobin (11). A reference (noncooperative) system containing the same number of sites is depicted by a discontinuous line. This system has been assigned the same value of mean ligand activity, and hence m_1, as the hemoglobin. The striking difference between the two systems suggests that the variance of the B distribution may be a good measure of cooperativity. The variance of B is by definition

$$\sigma^2 = m_2 - m_1^2 = \langle \ln^2 x \rangle - \langle \ln x \rangle^2 \tag{40}$$

and contains the first two moments. Since m_1 contains no information on cooperativity, the value of σ^2 is set by the value of m_2. In principle, one could shift the B distributions in Fig. 2 along the ln x axis, without affecting their shape, solely to make $m_1 = 0$. The shift would not affect the value of σ^2, and therefore Eq. 40 can be interpreted as the variance of a "normalized" B distribution shifted along the ln x axis to make its mean equal to zero. Hence, the variance of B corresponds to the second moment of its normalized form.

In order to understand the connection between cooperativity and m_2 it is useful to explore first the properties of a noncooperative reference system where all sites are alike and independent. A system containing N such sites has a partition function

$$Z = (1 + \alpha x)^N \tag{41}$$

and a generating function

$$G(\omega) = N\alpha^{-\omega}\pi\omega \operatorname{cosec} \pi\omega. \tag{42}$$

The first two moments and the variance are

$$m_1 = \langle y \rangle = -\ln \alpha \tag{43}$$

$$m_2 = \langle y^2 \rangle = \pi^2/3 + \ln^2\alpha \tag{44}$$

$$\sigma_0^2 = \pi^2/3 \tag{45}$$

where the subscript indicates the value of σ^2 for the reference, noncooperative system. Not surprisingly, σ^2 is independent of N and α. In the general case where the macromolecule has a partition function such as Eq. 1 one has

$$m_1 = \langle y \rangle = -N^{-1}\ln(\alpha_1\alpha_2 \ldots \alpha_N) \tag{46}$$

$$\begin{aligned} m_2 &= \langle y^2 \rangle \\ &= \pi^2/3 + N^{-1}(\ln^2\alpha_1 + \ln^2\alpha_2 + \ldots + \ln^2\alpha_N) \end{aligned} \tag{47}$$

$$\begin{aligned} \sigma^2 = \frac{\pi^2}{3} &+ (2N^2)^{-1} \sum_{i=1}^{N}\sum_{j=1}^{N} (\ln \mathrm{r}_i - \ln \mathrm{r}_j)^2 \\ &- N^{-1}\sum_{j=1}^{N} \theta_j^2. \end{aligned} \tag{48}$$

Here each coefficient α_j has been expressed uniquely in polar coordinates in the complex plane with a radius r_j and an angle θ_j. The expression above reveals the mathematical "driving forces" for cooperativity. The first term on the right-hand side is the variance of a reference system, σ_0^2, for which all r values are identical and the angles are zero. The variance of the B distribution in the general case is controlled by two opposing forces. One is the heterogeneity of the norm of the various α coefficients which tends to increase the value of σ^2 with respect to σ_0^2. The other is the sum of the squares of the angles in complex plane formed by the α values which tends to decrease the value of σ^2 with respect to σ_0^2. Therefore a necessary, although not sufficient, condition for cooperativity is that at least two coefficients α in the partition function are complex conjugate. The necessary and sufficient condition for cooperativity is that the force due to the angles exceeds the one reflecting heterogeneity of the norm of the coefficients α. The variance is minimized when the heterogeneity of the norms of the α values tends to zero with finite values of the angles. On the other hand, the variance is maximized when the heterogeneity of the norms of the α values is finite and the angles vanish.

Consider the case of a highly cooperative system where interactions are such that only the two limiting configurations of the system exist, i.e., the ones where all sites are either free or bound. Although somewhat unrealistic, this case is rather instructive for our discussion. The partition function for this system can be written as follows.

$$Z = 1 + A_N x^N = (1 + \alpha_1 x)(1 + \alpha_2 x) \ldots (1 + \alpha_N x) \tag{49}$$

with

$$\alpha_{k+1} = A_N^{1/N} e^{i\pi(N-1-2k)/N} \qquad (k = 0, 1 \ldots N-1) \tag{50}$$

The coefficients α are distributed around a circle in the complex plane with radius $A_N^{1/N}$. The norm r_j of α_j is equal to $A_N^{1/N}$ and is the same for all α values. Hence, the second term of the right-hand side of Eq. 48 vanishes. The angle associated with α_{k+1} is

$$\theta_{k+1} = \pi(N - 1 - 2k)/N \qquad (k = 0, 1 \ldots N-1) \tag{51}$$

and it is straightforward to prove that the first two moments and the variance of the B distribution are as follows.

$$m_1 = -N^{-1}\ln A_N \tag{52}$$

$$m_2 = N^{-2}\pi^2/3 + N^{-2}\ln^2 A_N \tag{53}$$

$$\sigma^2 = N^{-2}\pi^2/3 = N^{-2}\sigma_0^2 \tag{54}$$

Note that the variance is identical to that of a reference system divided by N^2. This is the minimum possible value of σ^2 for a system containing N sites. The maximum value of σ^2 for the same system is obtained when all sites are different and independent of each other. The partition function is then

$$Z = (1 + \alpha_1 x)(1 + \alpha_2 x) \ldots (1 + \alpha_N x) \tag{55}$$

where all coefficients are now real and positive and reflect the equilibrium constants for binding to each individual site. The first two moments and the variance of the B distribution are in this case

$$\langle y \rangle = -N^{-1}\ln(\alpha_1\alpha_2 \ldots \alpha_N) \tag{56}$$

$$\langle y^2 \rangle = \pi^2/3 + N^{-1}(\ln^2\alpha_1 + \ln^2\alpha_2 + \ldots + \ln^2\alpha_N) \tag{57}$$

$$\sigma^2 = \pi^2/3 + (2N^2)^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N} (\ln \alpha_i - \ln \alpha_j)^2. \tag{58}$$

The variance grows without bounds with the heterogeneity of the binding affinity of the individual sites. In general, the ratio $c = \sigma_0/\sigma$ is bounded from 0 to N, this property being remarkably similar to that of the Hill coefficient (3, 9). Hence, the standard deviation of B compared to that of a reference system can be used as a measure of cooperativity.

APPLICATIONS

The operational significance of the B distribution stems from the possibility of deriving important quantities for binding and cooperativity directly from experimental data, without necessarily knowing the detailed form of the partition function. Binding capacity measurements for the biologically relevant case of oxygen binding to human hemoglobin are given in Fig. 2. Calculation of the first two moments directly from the data by numerical integration yields the results depicted

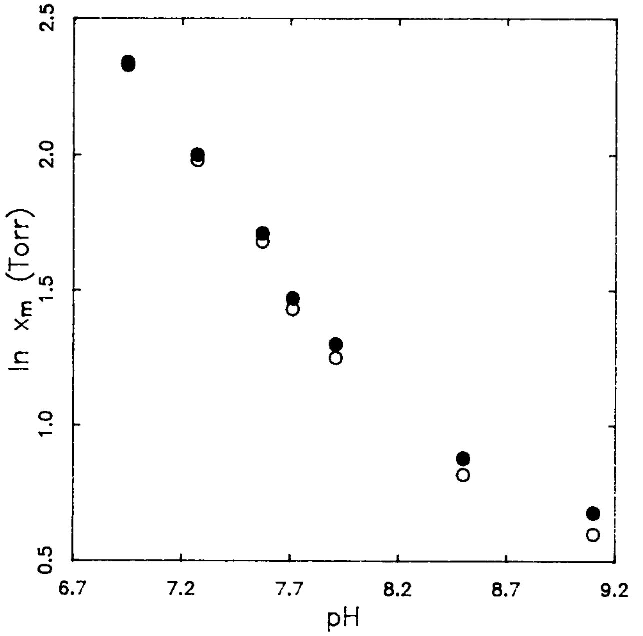

FIGURE 3. Values of ln x_m (○) for oxygen binding to human hemoglobin in the pH range 6.95–9.10 (11). The results of the analysis based on application of Eq. 6 (●) are in very good agreement with those obtained by nonlinear least-squares and show how the mean free energy of oxygenation/heme site changes with pH.

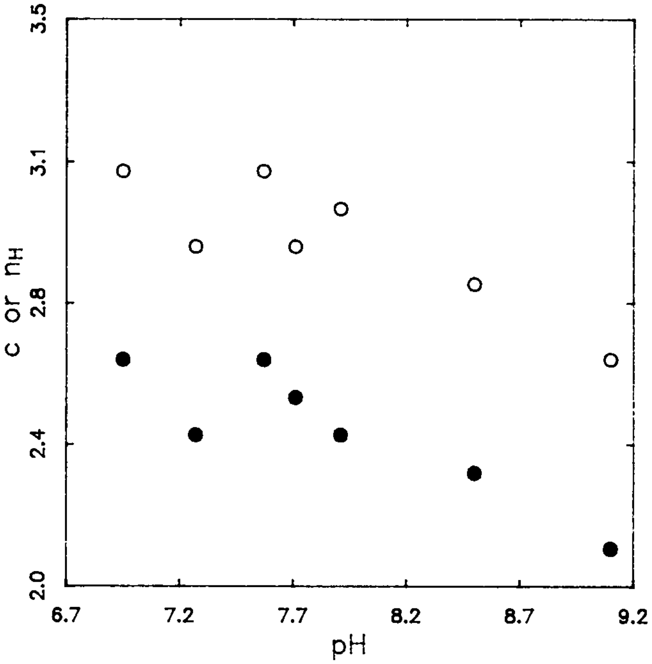

FIGURE 4. Comparison between the values of $c = \sigma_0/\sigma$ (●) computed from the data analyzed in Fig. 3 according to Eq. 8 using the moments of the *B* distribution and the maximum values of the Hill coefficient n_H (○). Although the two measures of cooperativity differ by about 0.4 units, the dependence of c and n_H on pH is remarkably similar.

in Figs. 3 and 4. The values of ln x_m shown in Fig. 3 as a function of pH (*filled circles*) are in excellent agreement with those derived by nonlinear least-squares analysis (*open circles*). Likewise, the values of $c = \sigma_0/\sigma$ shown in Fig. 4 (*filled circles*) parallel the values of the Hill coefficient (*open circles*) which gives the classical measure of cooperativity.

The applicability of the properties of the *B* distribution extends to any grand canonical ensemble showing saturation effects. The Fermi-Dirac statistics provide an example which can be treated in terms of the ideas developed here. Each quantum level which can exist either free or occupied by a fermion is equivalent to a binding site which can exist either free or bound to the ligand. At constant temperature, the ensemble of quantum levels with energy ϵ_j behaves just like an ensemble of independent binding sites with binding affinity α_j. The Fermi potential μ induces fluctuations of the number of occupied levels, $\partial N/\partial\mu$, around the mean *N*. These fluctuations are the "binding capacity" *B* of the ensemble of quantum states. The first moment of *B* gives the mean Fermi potential (in *kT* units), or else the average work done in occupying a quantum state. The standard deviation of *B* is always greater than σ_0 since the different quantum levels are independent and "bind" a fermion with different affinity.

Another relevant example which demonstrates the importance of the *B* distribution is given by the mean-field treatment of binding processes (1, 12). It is well known that the binding isotherm is in this case a transcendental function of the probability of binding per site $\nu = \overline{X}/N$, i.e. (12),

$$\nu = \alpha e^{\mathcal{L}\nu}x/(1 + \alpha e^{\mathcal{L}\nu}x) \qquad (59)$$

where $\mathcal{L}$ is a coupling parameter reflecting the effect of the Weiss molecular field, while α is the binding affinity per site. It is straightforward to show that for $\mathcal{L} = 4$ the binding capacity $B = d\nu/d \ln x$ diverges at $\nu = ½$ and so does the Hill coefficient (1, 12, 13). The *B* distribution, however, shows some intriguing properties. The generating function for the mean-field approximation based on Eq. 59 is found by application of Eq. 9. Elementary transformations lead to

$$G(\omega) = \int_0^1 \alpha^{-\omega}\nu^{\omega}(1-\nu)^{-\omega}e^{-\mathcal{L}\omega\nu}\,d\nu$$

$$= \alpha^{-\omega}\pi\omega \operatorname{cosec} \pi\omega\Phi(1+\omega; 2; -\mathcal{L}\omega) \qquad (60)$$

where Φ is the degenerate hypergeometric function (10). The generating function of a reference system is obtained for $\mathcal{L} = 0$, as expected. The Taylor expansion of $G(\omega)$ up to second order is

$$G(\omega) = 1 - (\ln\alpha + \mathcal{L}/2)\omega + \{(\ln\alpha + \mathcal{L})\ln\alpha/2$$
$$+ [\pi^2 + \mathcal{L}(\mathcal{L} - 3)]/6\}\omega^2 + \dots \qquad (61)$$

Hence,

$$m_1 = -\ln\alpha - \mathcal{L}/2 \qquad (62)$$

$$m_2 = \pi^2/3 + (\ln\alpha + \mathcal{L})\ln\alpha + \mathcal{L}(\mathcal{L} - 3)/3 \qquad (63)$$

$$\sigma^2 = \pi^2/3 + \mathcal{L}^2/12 - \mathcal{L}. \qquad (64)$$

Notice that the value of m_1 coincides with the logarithm of

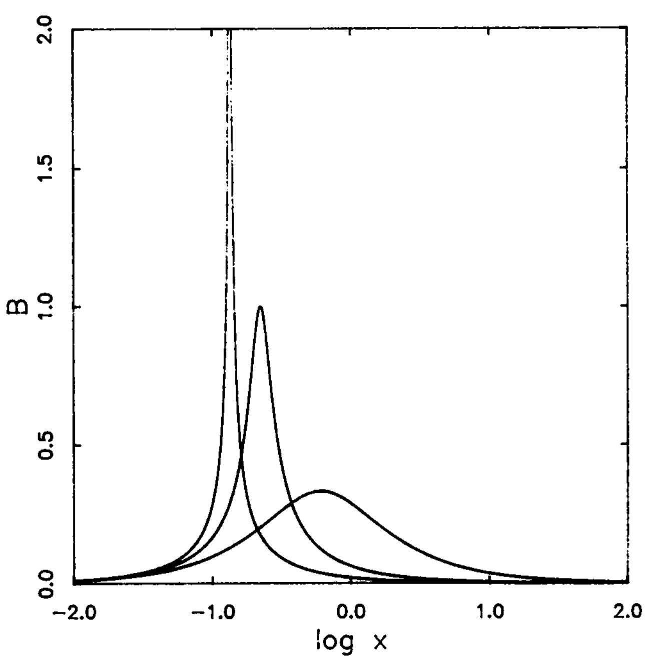

FIGURE 5 B distributions for the mean field approximation. Curves were calculated from the basic relationship: $B = \nu(1 - \nu)/[1 - \mathscr{L}\nu(1 - \nu)]$, with ν expressed as in Eq. 59. The values of $\mathscr{L}$ are (from right to left) 1, 3, and 4, while $\alpha = 1$. The curves progressively shift to the left when changing $\mathscr{L}$, as implied by Eq. 62. Also, the dispersion of the B distribution remains finite even when $\mathscr{L}$ assumes its critical value.

the value of x at half saturation, $x_{1/2}$. This is a consequence of the symmetry of ν (1) and the fact that for a symmetric binding curve $x_{\mathrm{m}} = x_{1/2}$ (14). The interesting result embodied by Eq. 64 is that the variance of the B distribution remains finite even for the critical value of $\mathscr{L} = 4$. The critical value of $c = \sigma_0/\sigma$ is about 2.3, thereby indicating that the critical B distribution has a dispersion of only 2.3 times smaller than that of a reference system. This feature is illustrated in Fig. 5. Another interesting result of our analysis is that the N coefficients (with $N \rightarrow \infty$) of the partition function for the mean-field approximation must be such that

$$\sum_{i=1}^{N} \sum_{j=1}^{N} (\ln \alpha_i - \ln \alpha_j)^2 = 2N^2\mathscr{L}(\mathscr{L}/12 - 1). \quad (65)$$

This result, along with the availability of an analytical expression for the moment generating function Eq. 60, makes it possible to tackle the rather nontrivial problem of finding the exact form of the α values and hence the solution for ν in closed form.

CONCLUSIONS

The definition of x_{m} and the statistical properties of the binding capacity pointed out in the foregoing analysis allow for a new definition of B. Classically, B represents a measure of the fluctuations of the number of ligated sites around the mean value. In this study we have demonstrated that B, divided by N, is a probability density function for $\ln x$. Specifically, B represents the distribution of free energies of binding (in kT units)/site. In fact, the mean value of this distribution is $\ln x_{\mathrm{m}}$, i.e., the average work (in kT units) done in ligating one site of the system. The standard deviation of B gives a measure of cooperativity as the dispersion of these free energy values around $\ln x_{\mathrm{m}}$. Consequently, the ratio B/N for a given value of $x = \xi$ gives exactly the probability that the free energy of binding per site is $kT \ln \xi$.

REFERENCES

1. Hill, T. L. 1960. Introduction to Statistical Thermodynamics. Dover, New York.
2. Stockmayer, W. H. 1950. Light-scattering in multi-component systems. *J. Chem. Phys.* 18:58–61.
3. Hill, T. L. 1985. Cooperativity Theory in Biochemistry. Springer, New York.
4. Schellman, J. A. 1975. Macromolecular Binding. *Biopolymers* 14:999–1013.
5. Wyman, J. 1964. Linked functions and reciprocal effects in hemoglobin: a second look. *Adv. Protein Chem.* 19:223–286.
6. Di Cera, E., S. J. Gill, and J. Wyman. 1988. Binding capacity: cooperativity and buffering in biopolymers. *Proc. Natl. Acad. Sci. USA.* 85:449–452.
7. Di Cera, E., and S. J. Gill. 1988. A new method for the determination of equilibrium constants through binding capacity measurements. *Biophys. Chem.* 32:149–152.
8. Di Cera, E. 1990. Thermodynamics of local linkage effects: contracted partition functions and the analysis of site-specific energetics. *Biophys. Chem.* 37:147–164.
9. Wyman, J., and S. J. Gill. 1990. Binding and Linkage. University Science Books, Mill Valley, CA.
10. Gradshtein, S., and I. M. Ryzhik. 1980. Tables of Integrals, Series and Products. Academic, New York.
11. Di Cera, E., M. L. Doyle, and S. J. Gill. 1988. Alkaline Bohr effect of human hemoglobin A_0. *J. Mol. Biol.* 200:593–599.
12. Di Cera, E. 1992. Mean-field treatment of local binding processes. *J. Chem. Phys.* 96:6515–6522.
13. Stanley, H. E. 1976. Introduction to Phase Transitions and Critical Phenomena. Oxford University Press, New York.
14. Di Cera, E., K.-P. Hopfner, and J. Wyman. 1992. Symmetry conditions for binding processes. *Proc. Natl. Acad. Sci. USA.* 89:2727–2731.