Complexation to macromolecules with a large number of sites

Josep Lluís Garcés, Francesc Mas, and Jaume Puy

Citation: J. Chem. Phys. **111**, 2818 (1999); doi: 10.1063/1.479559
View online: http://dx.doi.org/10.1063/1.479559
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v111/i6
Published by the American Institute of Physics.

Complexation to macromolecules with a large number of sites

Josep Lluís Garcés and Francesc Mas[a)]
Physical Chemistry Department, Facultat de Química, Barcelona University (UB), C/Martí i Franquès, 1, E-08028 Barcelona, Catalonia, Spain

Jaume Puy
Chemistry Department, Escola Tècnica Superior d'Enginyeria Agrària (ETSEA), Lleida University (UdL), Av. Rovira Roure, 177, E-25198 Lleida, Catalonia, Spain

(Received 3 November 1998; accepted 11 May 1999)

This paper presents an approach based on the saddle-point approximation to study the equilibrium interactions between small molecules and macromolecules with a large number of sites. For this case, the application of the Darwin–Fowler method results in very simple expressions for the stoichiometric equilibrium constants and their corresponding free energies in terms of integrals of the binding curve plus a correction term which depends on the first derivatives of the binding curve in the points corresponding to an integer value of the mean occupation number. These expressions are simplified when the number of sites tends to infinity, providing an interpretation of the binding curve in terms of the stoichiometric stability constants. The formalism presented is applied to some simple complexation models, obtaining good values for the free energies involved. When heterogeneous complexation is assumed, simple expressions are obtained to relate the macroscopic description of the binding, given by the stoichiomeric constants, with the microscopic description in terms of the intrinsic stability constants or the affinity spectrum. © *1999 American Institute of Physics.* [S0021-9606(99)50530-8]

I. INTRODUCTION

The study of the equilibria and kinetics of the interactions between small moleceules and macromolecules provides important information about many biological and environmental processes; typically, the interactions between biological macromolecules (proteins, DNA) and small molecules are crucial for understanding many metabolic routes;[1,2] in the environment, the complexation of heavy metals with fulvic and humic compounds in soils or in natural waters determines to a large extent their bioavailability, toxicity, and mobility.[3,4]

A macromolecule usually contains several complexing sites, to which small molecules can be bound, ranging from two sites (as is the case in many proteins)[5] to a very large number (as in polymeric complexation).[1,6,7] The presence of a large number of complexing sites can enormously complicate the interpretation of the experimental binding data because of the great number of chemical species and physicochemical phenomena involved (heterogeneity, positive and negative cooperativity, polyelectrolytic behavior, steric effects, conformational changes, linkage or competition effects, etc).[3,8]

A general way to describe the complexation characteristics starts defining ideal complexation as the model in which independent and homogeneous sites are assumed,[5,9,10] and considering the deviations from this ideal case, which are clearly recognized in the typical plots.[11–13] Several magnitudes, such as the Hill coefficient,[11] the binding capacity, introduced by di Cera,[2,5] or the activity coefficients of free and bound sites,[9] have been used to quantify such deviations.

The large number of sites present in many macromolecules complicates the fitting of the stoichiometric constants if no hypothesis on the complexation model is imposed. In some cases the value of the stability constants obtained by a nonlinear fitting of the Adair equation can be unstable and depend strongly on the experimental errors.[14] Some general properties of different magnitudes related to coverage data can be extremely useful for improving this fitting, such as the symmetrical properties of the binding curve[2,5] as well as some properties of the activity coefficients of free and bound sites and some characteristics of the average equilibrium function.[9] Even when the fitting of the Adair equation is successful, much microscopic information is lost in the global analysis of complexation. The free energy corresponding to a stoichiometric equilibrium is an average energy of all the microscopic species involved (with a fixed number of bound small molecules). The description of site-specific effects demands resolution of more parameters than those yielded by the global description. Local coverage data are required for a description of binding and linkage effects taking place at individual sites of a multisite macromolecule.[15]

Some approximate procedures to model the macromolecular complexation processes have also been developed. The most common start assuming a model of complexation. In many cases, if positive cooperativity is not detected, a model with heterogeneous and independent sites is assumed, which involves superposition of local Langmuirian isotherms. Some procedures to fit the affinity spectrum with no *a priori* assumption of the discrete or continuous set of affinities involved have been developed.[4,16,17] Klotz demonstrated that all the complexation models can be expressed as a summation of Langmuirian isotherms, if imaginary stabil-

a)Author to whom correspondence should be addressed. Electronic mail: fmas@qf.ub.es; Fax: 34-93-402 12 31.

ity constants are allowed.[18] The demonstration is based on the factorization of the macrocanonical partition function. Nevertheless, as Klotz himself recognizes, imaginary stability constants have no physical meaning.

Another important approximation, valid when the number of sites is large enough, supposes that thermodynamical limit conditions can be applied to a macromolecule.[1,9] Darwin and Fowler demonstrated[19] that as the number of terms of a partition function increases, its value can be approximated by taking the maximum term. The assumption of thermodynamical limit conditions, although restrictive, allows to obtain analytical expressions which can be easily used to fit the binding curve.

This work describes a treatment for systems with high number of sites, obtaining the thermodynamical limit value for the stoichiometric stability constants and a first corrective term suited for cases with a number of sites large enough to apply the hypothesis on which the Darwin–Fowler method (saddlepoint approximation) is based. Section II introduces the nomenclature and some definitions used in the work. Section III describes the application of the Darwin–Fowler method to macromolecular complexation. The discussion of the limiting case of very large number of sites and some consequences of such approximation on the values of the stability constants and on the physical interpretation of the binding curve are included in Sec. IV. The results obtained in Secs. III and IV are applied in Sec. V to simple complexations models; the homogeneous model with interactions between bound sites and the heterogeneous ideal model. The mathematical details are described in the Appendix.

II. THERMODYNAMICAL TREATMENT OF THE COMPLEXATION OF A SMALL MOLECULE TO A MACROMOLECULE

In order to avoid ambiguities, let us assume that:

(i) A complexing site S is a set of coordinating groups of a macromolecule, P, to which a small molecule (M) (known in the biochemical literature as ligand) can bind;
(ii) M_jP labels the chemical species with the same number j of bound ligands, regardless of the specific sites complexed;
(iii) All full covered macromolecules have the same number, s, of bound ligands;
(iv) M, MS, and S label the so called formal species defined by their concentrations c_M (concentration of free ligand), $c_{MS}\equiv\Sigma_{j=0}^{s} j c_{M_jP}$ (concentration of bound ligand) and $c_S\equiv\Sigma_{j=0}^{s}(s-j)c_{M_jP}$ (concentration of free sites of the macromolecule);
(v) Ideal dilute solution behavior for the real species M_jP $(j=0,...,s)$ is assumed.

Let us represent the sequential complexation of the macromolecule as

$$M_{j-1}P + M \overset{K_j}{\Leftrightarrow} M_jP \quad j=1,...,s, \tag{1}$$

where the stability constants for these processes, K_j $(j=1,...,s)$, are known as stoichiometric constants.

The most common techniques used in the study of the macromolecular binding are based on measures of c_M. Once this value and that of the total concentrations are known it is possible to calculate the mean occupation number, ν, the coverage, θ, or the average equilibrium function, K_c, at a concentration c_M, as[3,5]

$$\nu(c_M)\equiv\frac{c_{MS}}{c_{T,S}}s, \tag{2}$$

$$\theta(c_M)\equiv\frac{\nu}{s}=\frac{c_{MS}}{c_{T,S}}, \tag{3}$$

$$K_c(c_M)=\frac{c_{MS}}{c_M c_S}=\frac{1}{c_M}\left(\frac{\nu}{s-\nu}\right)=\frac{1}{c_M}\left(\frac{\theta}{1-\theta}\right), \tag{4}$$

where $c_{T,S}$ is the total concentration of macromolecular sites.

The equilibrium relationships corresponding to Eqs. (1) allow us to relate the mean occupation number to a polynomial in terms of c_M, the Adair equation,[2,14] which can be written as

$$\nu(c_M)=\frac{\Sigma_{j=1}^{s} j b_j c_M^j}{\Sigma_{j=0}^{s} b_j c_M^j}=\left(\frac{\partial \ln \Xi(c_M)}{\partial \ln(c_M)}\right), \tag{5}$$

where the coefficients of the polynomial, $b_j\equiv K_1K_2...K_j$ $(j=1,...,s;\ b_0\equiv 1)$, are the Adair coefficients, and $\Xi(c_M)\equiv\Sigma_{j=0}^{s} b_j c_M^j$ is the macrocanonical partition function. The Adair coefficients lead directly to the stoichiometric constants as $K_j=(b_j/b_{j-1})$.

At very low concentrations of the complexing agent, all complexation systems behave like the ideal one.[9,20] The concentration of the formal species do not differ among different complexation models, and the average equilibrium function tends to a constant given by

$$\lim_{c_M\to 0} K_c(c_M)\equiv K. \tag{6}$$

Because of this limiting behavior, a formalism in which K plays the role of a thermodynamical constant for the process $M+S\Leftrightarrow MS$ has been developed, the activity coefficients for formal species relating the concentration of formal species to the concentration that would have been present if the complexation had been ideal. Moreover, at very high concentrations of the complexing agent, the average equilibrium function tends to

$$\lim_{c_M\to\infty} K_c(c_M)\equiv sK_s, \tag{7}$$

which allows us to compute the last stoichiometric stability constant K_s, if s is known.

It has also been reported[9] that the macrocanonical partition function, $\Xi(c_M)$, is in fact a polynomial of the product Kc_M. Therefore, K is a normalization constant for the coefficients of the macrocanonical partition function, and a new set of coefficients $\{a_j\}$ can be defined as

$$b_j\equiv a_jK^j \tag{8}$$

so that $\{a_j\}$ are independent of K. Thus, Eq. (5) can always be rewritten in terms of $\{a_j\}$ if a suitable change of the

concentration scale is chosen. Therefore, systems with different K can be compared and reduced to the case $K=1$ with a suitable election of the concentration units. In the following we assume that $K=1$ without loss of generality.

For ideal complexation, the expressions for K_j and b_j reduce to[5,9]

$$b_j^{\mathrm{id}}=\binom{s}{j}K^j;\quad a_j^{\mathrm{id}}=\binom{s}{j};\quad K_j^{\mathrm{id}}=\frac{s-j+1}{j}K, \tag{9}$$

where the superscript (id) refers to the ideal complexation case. Any complexation model can be tackled by defining an excess equilibrium constant K_j^E as

$$K_j\equiv K\frac{s-j+1}{j}K_j^E \tag{10}$$

and, labeling $\Delta_j G^0$ the standard Gibbs energy for the complexation equilibrium of the j-ligand,

$$\Delta_j G^0=-RT\ln K_j \tag{11}$$

an excess Gibbs energy can defined as

$$\Delta_j G^E\equiv\Delta_j G^0-(\Delta_j G^0)^{\mathrm{id}}=-RT\ln K_j^E \tag{12}$$

indicating the deviation from the ideal complexation case.

The Gibbs energy involved in the binding of j ligands to the naked macromolecule is given by

$$G_j^0\equiv\sum_{i=1}^{j}\Delta_i G^0=-RT\ln b_j. \tag{13}$$

Finally, we will refer to the binding capacity, $B(c_M)$, a function of c_M introduced by di Cera[2,5] and defined as

$$B(c_M)\equiv\frac{d\nu}{d\ln c_M}. \tag{14}$$

Di Cera has proved[5] that the function B is always positive, and an increase or decrease of c_M leads to an increase or decrease of B, respectively. In fact, the binding capacity plays the role of a physical response function such as heat capacity or compressibility, and its positive character is related to the stability principles embodied by the second law of thermodynamics.[21]

III. APPLICATION OF THE SADDLE POINT APPROXIMATION TO THE DETERMINATION OF EQUILIBRIUM CONSTANTS

We are interested in obtaining simple approximate expressions for b_j which can be useful to avoid unstabilities in the fitting of the Adair equation (5).[14] For high s values, a usual way to deal with problems in statistical mechanics is to apply the classical approximation of Darwin and Fowler,[19] which allows a drastical simplification of the expressions involved.

Let us integrate expression (5) for the mean number of occupation. Thus,

$$\Xi(c_M)\equiv\sum_{j=0}^{s}b_j c_M^j=\exp(sF(c_M)), \tag{15}$$

where

$$F(c_M)\equiv\int_0^{c_M}\frac{\theta(c_M)}{c_M}dc_M=\int_{-\infty}^{\ln c}\theta(c_M)d\ln c_M. \tag{16}$$

Equations (15) and (16) relate the macrocanonical partition function Ξ at a concentration c_M to the area covered by the curve $\theta(c_M)/c_M$ vs c_M.

The Wyman integral equation[2,5] can be easily derived from Eqs. (15) and (16). The integration, by parts, of Eq. (16) leads to

$$\ln\left(\frac{\Xi}{c_M^{\nu}}\right)=-\int_0^{\nu}\ln c_M d\nu \tag{17}$$

which in the limit of $c_M\to\infty$ yields b_s,

$$\ln b_s=-\int_0^{s}\ln c_M d\nu. \tag{18}$$

Equation (18), known as the Wyman integral equation indicates that b_s can be obtained from the area of the curve $\ln c_M$ vs ν.

We are now trying to extend the Wyman integral Eq. (18) for b_s to any b_j. Since $\Xi(c_M)$ is a polynomial in c_M, the coefficients $\{b_j\}$ can be expressed using the Laurent development[22] as

$$b_j=\frac{1}{2\pi i}\oint\frac{\exp(sF(z))}{z^{j+1}}dz, \tag{19}$$

where z is a complex variable and the integration takes place over a path containing $z=0$. Let us define $\phi_j(z)$ as

$$\phi_j(z)\equiv\frac{\exp(sF(z))}{z^{j+1}}. \tag{20}$$

The Appendix shows that if z takes values on a complex plane, the function $\phi_j(z)$ has a local minimum in the real direction and a local maximum in the imaginary direction at $z=c_{j+1}$ (where c_{j+1} indicates the value of c_M corresponding to $\nu=j+1$) whenever $j<s-1$; i.e., $z=c_{j+1}$ is a saddle point. This is an interesting mathematical property because the integration (19) can be easily performed on a circle with center $z=0$ and radius $r=c_{j+1}$; since $z=c_{j+1}$ is the highest maximum on such a circle (see Appendix), the major contribution to the value of b_j, is located close to $z=c_{j+1}$, in the imaginary direction.

Performing the integral (19) as described in the Appendix Eq. (A14) leads to the following value for b_j:

$$\ln b_j=-\int_0^{j+1}\ln c_M d\nu-\frac{1}{2}\ln\left(2\pi\frac{1}{c_{j+1}^2}\left[\frac{d\nu}{d\ln c_M}\right]_{c_M=c_{j+1}}\right), \tag{21}$$

which, using the definition of b_j in terms of K_j, becomes

$$\ln K_j = -\int_j^{j+1} \ln c_M d\nu - \frac{1}{2}\ln\left\{ \left(\frac{c_j}{c_{j+1}}\right)^2 \frac{\left(\frac{d\nu}{d\ln c_M}\right)_{c_M=c_{j+1}}}{\left(\frac{d\nu}{d\ln c_M}\right)_{c_M=c_j}} \right\}. \quad (22)$$

Expression (22) provides a method to obtain an estimation of the first $s-2$ stoichiometric stability constants from experimental data of $\ln c_M$ vs ν. The first term in the rhs of Eq. (22) corresponds to the area under the curve $\ln c_M$ vs the mean occupation number, ν, between $\nu=j$ and $\nu=j+1$. Therefore, for a large number of binding sites, the calculation of the stoichiometric constants only needs a small interval of points around $\nu=j$. The constants K_{s-1} and K_s cannot be calculated by this method because the function $\phi_j(z)$ does not have a saddle point for $j=s-1$ and $j=s$ as described in the Appendix. Obviously, both K_s and b_s can be calculated as indicated in Eqs. (7) and (18), respectively. When K_s and b_s are known, it is possible to evaluate the values of b_{s-1}, and also K_{s-1}, the remaining constant.

IV. THERMODYNAMICAL LIMITING CONDITIONS ($s\to\infty$)

For many systems, the number of sites per macromolecule is high enough to obtain a good description taking the limit $s\to\infty$, an assumption which is known as the thermodynamical limit. We will particularize the results (21) and (22) to these conditions in order to deduce simple expressions for the set of stoichiometric constants of complexation and to provide a physical interpretation of the binding curve.

A. An expression for the set of stability constants $\{K_j\}$ in the limit $s\to\infty$

In the limit $s\to\infty$, the second term of the rhs of Eq. (22) disappears. This is easily seen if Eq. (22) is written in terms of θ and it is divided by s. For large values of s, it becomes

$$\lim_{s\to\infty}\frac{1}{s}\ln K_j = -\lim_{s\to\infty}\frac{1}{s}\frac{\Delta_j G^0}{RT} \simeq -\int_{\theta_j}^{\theta_{j+1}} \ln c_M d\theta, \quad (23)$$

and the standard Gibbs energy, $\Delta_j G^0$, can be calculated by computing the area under the curve $\ln c_M$ vs ν or θ between the points ($\ln c_j$, $\nu=j$ or θ_j) and ($\ln c_{j+1}$, $\nu=j+1$ or θ_{j+1}).

Furthermore, for very large values of s, expression (23) can be simplified applying the mean value theorem to the integral

$$\int_{\theta_j}^{\theta_{j+1}} \ln c_M d\theta = \frac{1}{s}\int_j^{j+1} \ln c_M d\nu = \frac{1}{s}\{(1-\varphi)\ln c_j + \varphi \ln c_{j+1}\}; \quad 0\leqslant\varphi\leqslant 1. \quad (24)$$

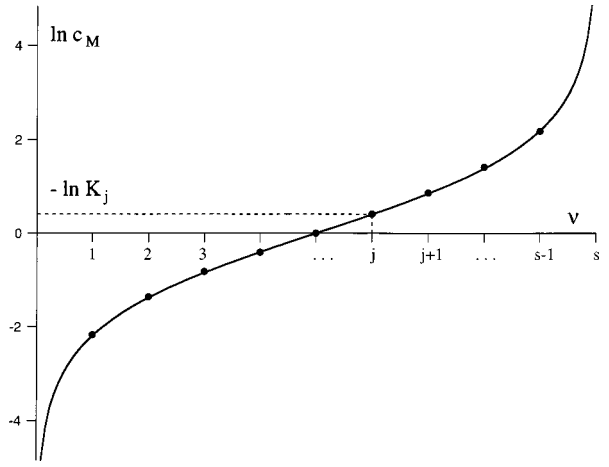

FIG. 1. Schematic plot of the inverse of the binding curve, $\ln c_M$ vs ν. The value of $\ln c_M$ at the integer values of $\nu=j$ yields directly, in the limit $s\to\infty$, $-\ln K_j$.

Since the interval of integration in Eq. (24) decreases as s increases, we can assume $\varphi\approx 0$ for large values of s; therefore, Eq. (23) becomes

$$\lim_{s\to\infty}\ln K_j = -\frac{\Delta_j G^0}{RT} \simeq -\ln[c_M]_{\nu=j} = -\ln c_j. \quad (25)$$

Hence, under thermodynamical limiting conditions, the value of the standard Gibbs energy associated with K_j is the opposite of the logarithm of c_j. Expression (25) can also be rewritten as

$$\lim_{s\to\infty} K_j = \left[\frac{1}{c_M}\right]_{\nu=j} = \frac{1}{c_j} \quad (26)$$

providing a very simple method for obtaining the set of stoichiometric constants and their corresponding Gibbs energies in a system with a large number of sites.

Some remarks of the result (26) should be emphasized:

(i) This result provides a simple physical interpretation to the inverse of the binding curve ($\ln c_M$ vs ν); the value of $\ln c_M$ at $\nu=j$ gives directly the standard Gibbs energy involved in the complexation of the jth ligand, $\Delta_j G^0$ (see Fig. 1). Likewise, a plot of $1/c_M$ vs ν leads directly, for integer values of ν, to the stoichiometric stability constants.

(ii) If thermodynamical limiting conditions are fulfilled, K_j must decrease with j for any complexation model, reaching the values provided by Eq. (26). Now, we remark the generality of this result. Actually, if K_j increases with j for $s\to\infty$, $(K_{j+1}/K_j)\geqslant 1$, Eq. (26) yields $c_{j+1}\leqslant c_j$. This means that the free ligand concentration corresponding to $\nu=j+1$ is lower than the value of c_M corresponding to $\nu=j$ and leads immediately to a negative value of the slope ν vs $\ln c_M$ ($B<0$). Therefore, Eq. (26) is directly related to the thermodynamical stability condition ($B>0$) for $s\to\infty$.

(iii) According to Eq. (25), we can write

$$\ln K_{j+1}-\ln K_j=\ln c_j-\ln c_{j+1}\approx-\left[\frac{d\ln c_M}{d\nu}\right]_{\nu=j}=-\left[\frac{1}{B}\right]_{\nu=j} \tag{27}$$

and

$$\frac{\Delta_{j+1}G^0-\Delta_jG^0}{RT}=-\left[\frac{1}{B}\right]_{\nu=j}. \tag{28}$$

Thus, for $s\to\infty$, the inverse of the B value can be interpreted as a measure of the difference between the affinity of the ligand $j+1$ and the ligand j, i.e., as the standard Gibbs energy needed to complex a new ligand, relative to the standard Gibbs energy needed by the last complexed one.

B. Relationship between the average equilibrium function, K_c, and the stoichiometric constants K_j in the limit $s\to\infty$

According to Eq. (10),

$$K_j^E=\frac{j}{s-j+1}\frac{K_j}{K}=\frac{j/s}{1-j/s+1/s}\frac{K_j}{K}. \tag{29}$$

For s large enough the term $1/s$ can be neglected

$$\lim_{s\to\infty}K_j^E\simeq\frac{j}{s-j}\frac{K_j}{K} \tag{30}$$

and using Eq. (26) for the value of K_j and Eq. (4) for the average equilibrium function,

$$\lim_{s\to\infty}K_j^E\simeq\frac{1}{K}\left[\frac{1}{c_M}\frac{\nu}{s-\nu}\right]_{c_M=c_j}=\left[\frac{K_c}{K}\right]_{c_M=c_j}. \tag{31}$$

In terms of Δ_jG^E, Eq. (31) is rewritten as

$$\lim_{s\to\infty}\Delta_jG^E=-RT\left[\ln\left(\frac{K_c}{K}\right)\right]_{c_M=c_j}. \tag{32}$$

Thus, for a high enough s value, that of K_c/K when the mean number of occupation is an integer is merely the value of the excess stability constant K_j^E, and its logarithm is proportional to the excess Gibbs energy.

V. APPLICATION TO SOME PARTICULAR COMPLEXATION MODELS

In this section, we apply the results above described to some simple models of complexation; an homogeneous complexation with interaction between sites, and an heterogeneous case of complexation without interaction between sites. Both cases are widely used in the literature.

A. Complexation with interaction between neighboring sites

A simple model to describe systems with interaction between bound species considers a fixed interaction energy δ between the nearest neighboring occupied sites (1D Ising model). The excess stability constant (12), within the 1D-mean-field approximation, is given by[9,23–25]

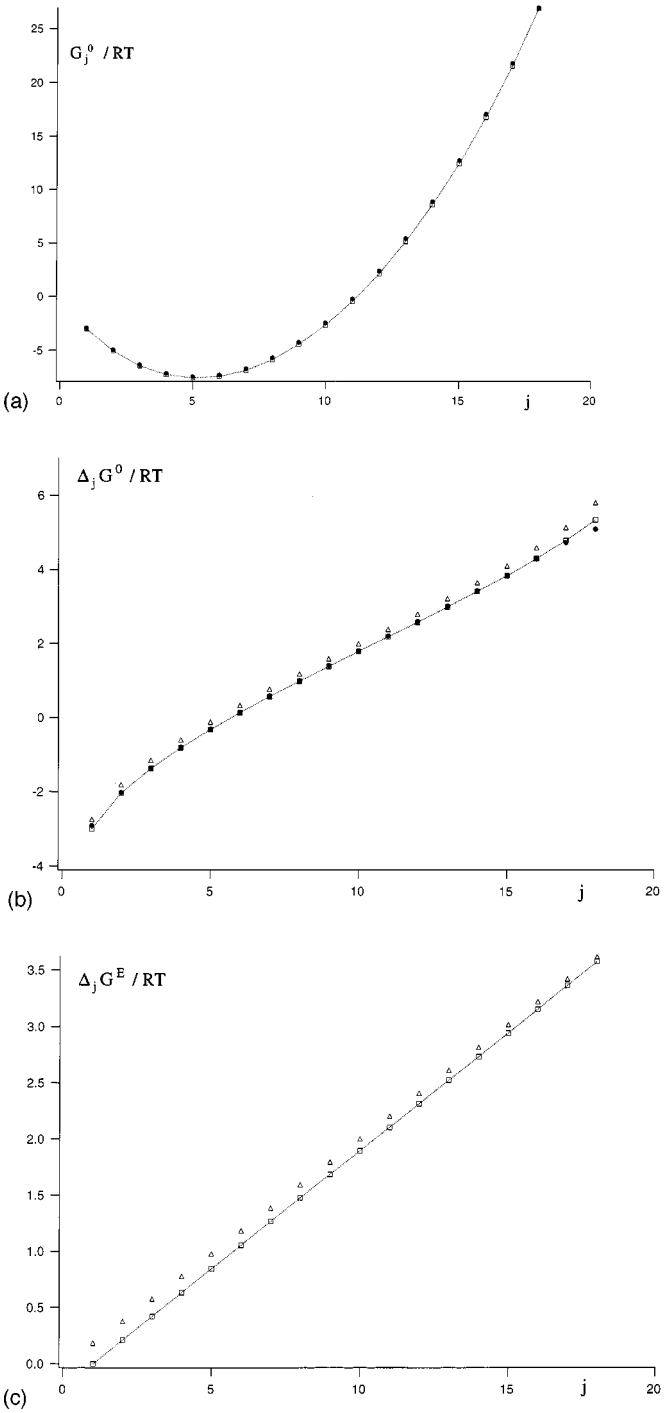

FIG. 2. Exact values [in dotted lines with marker □ (empty square)] of the Gibbs energies G_j^0 (a), increments of Gibbs energies (Δ_jG^0) (b), and excess Gibbs energies (Δ_jG^E) (c), for the interaction model of complexation with interaction parameter $\beta\delta=2$ (negative cooperativity), number of sites $s=20$ and $K=1\ \mathrm{M}^{-1}$, compared with those obtained by using the different approximations proposed in this work; marker ● (filled circle) for approximation given by Eqs. (21) (a) and (22) (b); marker △ (empty triangle) for approximations given by Eqs. (25) (b) and (32) (c).

$$K_j^E=\exp\left(-\beta\delta\frac{2(j-1)}{s-1}\right), \tag{33}$$

where $\beta=(1/k_BT)$, thus showing that K_j depends exponentially on j.

As described in Sec. II, results obtained for systems with any value of K can be compared with those present here if a suitable change of the concentration scale is performed. As K is merely a concentration scale factor we have used $K=1$ in the present calculations.

As usual, two contributions (an ''enthalpic'' or ''energetic'' effect and an ''entropic'' effect) can be considered in $G_j^0\,(=H_j^0-TS_j^0)$ and $\Delta_j G^0\,(=\Delta_j H^0-T\Delta_j S^0)$. Taking $K=1$, the enthalpic contribution only includes the effect of the lateral interactions (positive for repulsive interactions and negative for attractive ones). Concerning the entropic contribution, while $j<s/2$, S_j^0 increases as j increases and then S_j^0 decreases for $j>s/2$ because the number of accessible microstates for the binding of j ligands to the naked macromolecule decreases. Accordingly, $\Delta_j S^0$ is a monotonically decreasing function taking positive values for $j<s/2$ and negative ones for $j>s/2$ regardless the sign of δ.

Figures 2(a), 2(b), and 2(c) show the exact G_j^0, $\Delta_j G^0$, and $\Delta_j G^E$ $(j=1,...,s-2)$ values obtained with Eq. (33) for a case with repulsive interactions $(\beta\delta=2)$ and $s=20$ (as an estimation of the thermodynamic limit conditions) together with the results obtained from the approximate expressions (21), (22) proposed in this work and their corresponding limits when $s\to\infty$, Eqs. (25) and (32). The exact G_j^0 values are in good agreement to those obtained with the approximate expression (21), as Fig. 2(a) shows. A good estimation is also obtained for $\Delta_j G^0$ (related to the stoichiometric equilibrium constant, K_j) using the approximate expressions (22) and (25) in Fig. 2(b) and a comparison between the exact $\Delta_j G^E$ values and those obtained from the average equilibrium function, K_c, by using the approximate expression (32), is shown in Fig. 2(c). Decreasing s (see Fig. 3, where the same calculations as Fig. 2 have been performed for a smaller number of sites, $s=10$), the thermodynamical limit approximate expressions begin to deviate, especially Eq. (25) for $\Delta_j G^0$ [Fig. 3(b)], but the corrective term given by Eq. (21) for G_j^0 [Fig. 3(a)] and by Eq. (22) for $\Delta_j G^0$ [Fig. 3(b)], yield still accurate results.

As Figs. 2 and 3 correspond to a repulsive case, for low values of j, G_j^0 decreases with increasing j because the entropic contribution predominates over the enthalpic one, whereas for intermediate and large values of j, the repulsive interaction becomes more important and G_j^0 increases even to positive values. Accordingly, $\Delta_j G^0$ and $\Delta_j G^E$ always increase with j and, therefore, K_j and K_c always decrease with j, as expected.

When positive cooperativity is considered (see Fig. 4, where $\beta\delta=-1$) only the entropic contribution for $j>s/2$ contributes to increase G_j^0. So, G_j^0 decreases with increasing j until a j value larger than that found in the repulsive case. For this attractive case, the agreement between the real and approximate values of G_j^0 is not so good [Fig. 4(a)] as it is in the repulsive case (Fig. 2), especially for the lowest and the highest values of j. This can be explained from the relationship between the coverage θ and the ligand concentration c_M, which under thermodynamical limiting conditions is

$$\frac{\theta}{Kc_M}=(1-\theta)\exp(-2\beta\delta\theta). \tag{34}$$

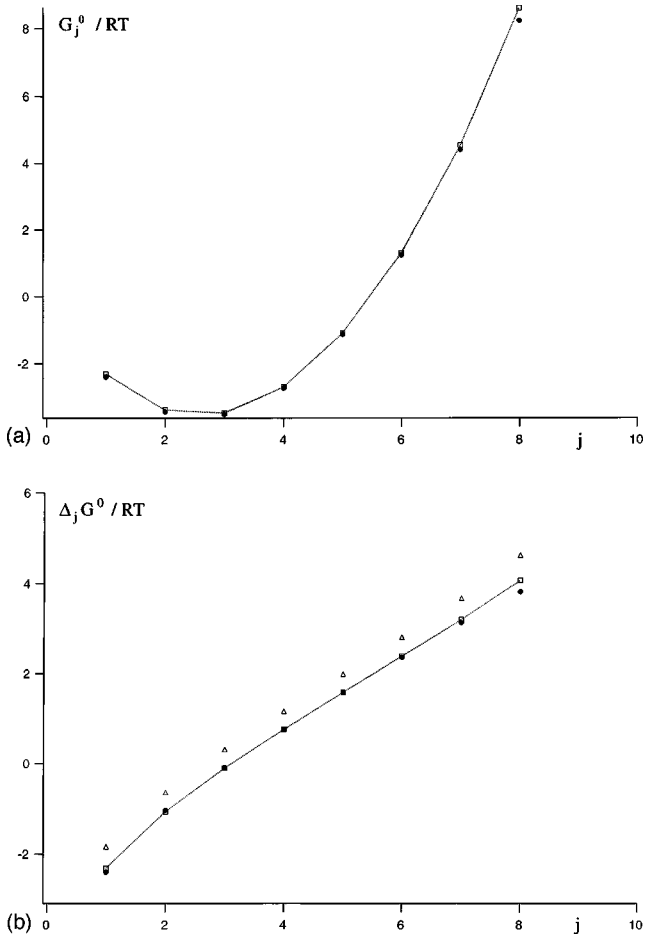

FIG. 3. As in Fig. 2(a) and Fig. 2(b) but with the number of sites $s=10$.

The condition $B>0$ [the binding capacity defined in Eq. (14)], applied to Eq. (34), leads to

$$\beta\delta>-2 \tag{35}$$

which indicates that the system is thermodynamically stable for $s\to\infty$ only when $\beta\delta>-2$. This is a consequence of the use of the mean-field approximation (which implies a phase transition at $\beta\delta=-2$ for the 1D-Ising model).[19,21] Since approximations (21)–(22), (25), and (32) have been deduced using the thermodynamic limit as the starting point, their values will diverge from the real ones as we approach $\beta\delta=-2$, since the model used leads to a non realistic description of the system if $\beta\delta<-2$ as s increases. This difficulty can be avoided taking into account the correlations between the interactions of different ligands in the macromolecule, which we neglected with the mean-field approximation.

B. Heterogeneous complexation: Relationship between the stoichiometric and the intrinsic stability constants

Due to its direct microscopic interpretation, the heterogeneous model is widely used especially when there is some previous physical information suggesting the consideration of sites with different affinities.

Let us consider a system with m kinds of sites, and let s_i be the number of sites of type i with an intrinsic stability

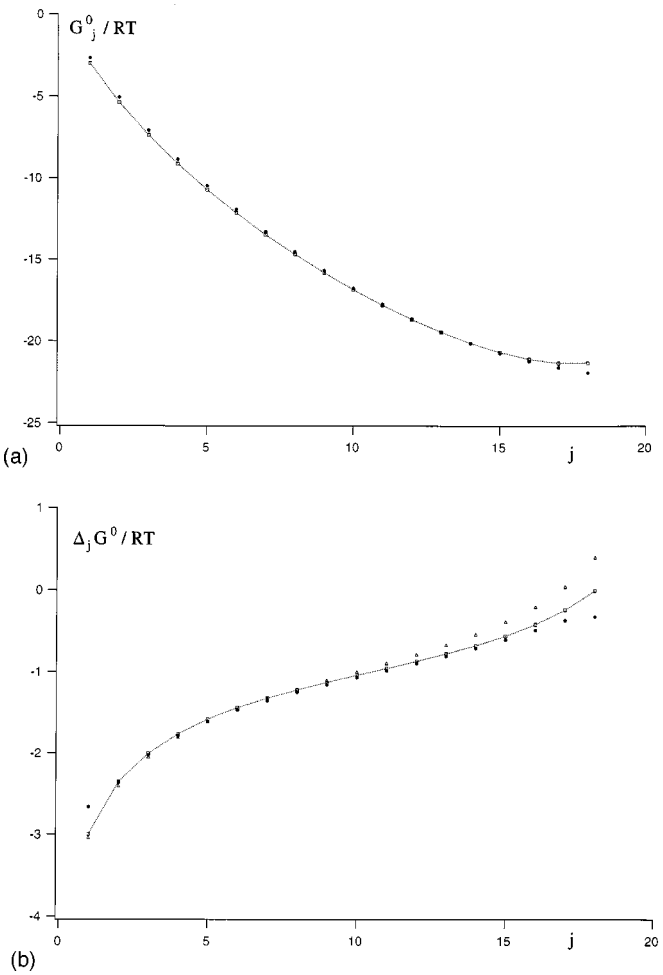

FIG. 4. As in Fig. 3 but with $\beta\delta=-1$ (positive cooperativity) and $s=20$.

constant k_i $(i=1,...,m)$. Assuming a Langmuirian local isotherm, the mean occupation number can be related to the ligand concentration, without assuming thermodynamical limiting conditions, by

$$\nu(c_M)=\sum_{i=1}^{m} s_i \frac{k_i c_M}{1+k_i c_M};\ \sum_{i=1}^{m} s_i = s. \tag{36}$$

The relationship between the stoichiometric constants and the intrinsic stability constants described by Klotz[10] is cumbersome for large s values. Nevertheless, expression (25), valid for the stoichiometric stability constants in the limit $s\to\infty$, together with Eq. (36), provide a very simple expression for such relationship, valid for large s values,

$$\lim_{s\to\infty} K_j=\frac{1}{c_j}=\sum_{i=1}^{m}\frac{s_i}{j}\frac{k_i}{1+k_i c_j}. \tag{37}$$

Equation (37) indicates that the stoichiometric stability constants K_j can be considered as an average of the intrinsic stability constants,

$$\lim_{s\to\infty} K_j=\frac{1}{j}\sum_{i=1}^{m} w_{i,j}k_i;$$
$$w_{i,j}\equiv\frac{s_i}{1+k_i c_j}=s_i(1-\theta_{i,j})=s_i-\nu_{i,j}, \tag{38}$$

where $w_{i,j}$, the weighting factors, corresponds to the mean number of free sites of the ith type at $\nu=j$ and $\theta_{i,j}\equiv(k_i c_M)/(1+k_i c_M)_{c_M=c_i}$ and $\nu_{i,j}\equiv s_i\theta_{i,j}$ are the coverage and the mean occupation number of the sites of type i at the free ligand concentration c_j.

Thus, K_j are calculated as an average of the intrinsic stability constants k_i with a weighting factor given by the mean number of free sites of type i at the concentrations c_j, respectively. This indicates that the stoichiometric stability constant of the binding of an additional ligand to a macromolecule P is an average of the intrinsic stability constants not yet occupied.

The relationship between K_c and the intrinsic stability constants is easily obtained. From Eqs. (4) and (36) it is obtained[26]

$$K_c(c_M)=\frac{1}{c_M}\left(\frac{\nu}{s-\nu}\right)=\frac{1}{s-\nu}\sum_{i=1}^{m} s_i\frac{k_i}{1+k_i c_M}=\sum_{i=1}^{m} x_i(c_M)k_i, \tag{39}$$

where $x_i\equiv(s_i-\nu_i)/(s-\nu)$ provided that $\Sigma_{i=1}^{m}x_i=1$, and ν_i is the mean occupation number of the sites of ith type at the concentration c_M.

This expression holds independently of the s-value and indicates that the average equilibrium constant is an average of the intrinsic stability constants with a weighting factor given by the fraction of mean number of free sites of type i with respect to the total of free sites at the concentration c_M.

Figures 5(a) and 5(b) consider a heterogeneous system with two types of sites with the same probability $(s_1=s_2)$, but with a different total number of sites s, 16 and 30, respectively. The approximate expressions (22) and (25) are used to reproduce the value of the free energies involved. In general, a good accordance between the exact results and those given by (22) and (25) is obtained, especially for the highest number of sites $(s=30)$. As expressions (22) and (25) work out the results for the stoichiometric constants from the plot $1/c_M$ vs ν [Fig. 1(b)], care must be taken to minimize the experimental error in such plots, because they have direct influence on the stoichiometric constants.

Regarding the results described above:

(i) Equation (37) relates K_j to the intrinsic stability constants k_i. Replacing c_j in Eq. (37) in terms of K_j, we have

$$1=\sum_{i=1}^{m}\frac{s_i}{j}\frac{k_i}{k_i+K_j}. \tag{40}$$

Expression (40) relates the Adair approach, a macroscopic description of the binding, to a microscopic or local formalism widely used in complexation problems; the affinity spectrum method [usually, in this last formalism, the summation involved in Eq. (40) is replaced by an integral when a continuous distribution of intrinsic stability constants is assumed].

(ii) In some cases, it is reasonable to replace the local Langmuirian isotherm in Eq. (36) by other local isotherms (for instance, the Fowler–Guggenheim or Frunkim isotherm) of the kind $f(k_i,c_M)$. Then Eq. (36) could be generalized as

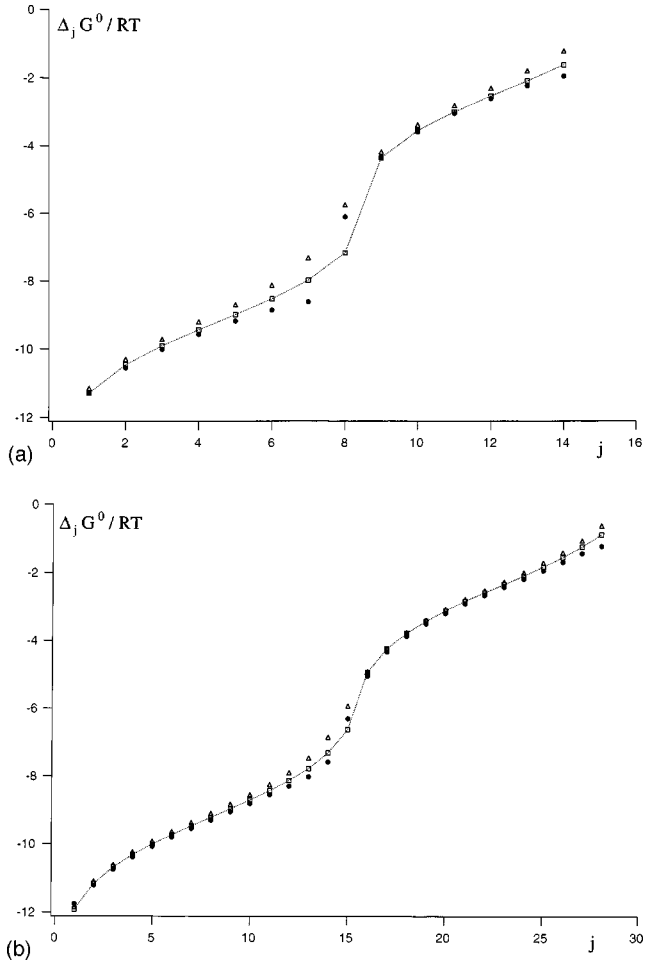

FIG. 5. Exact increments of Gibbs energies ($\Delta_j G^0$) for a heterogeneous system with two kinds of independent sites ($k_1=10\ \mathrm{M}^{-1}$, $k_2=10^4\ \mathrm{M}^{-1}$), compared with those obtained by using the different approximations proposed in this work with markers as in Fig. 2(b) (a) $s_1=s_2=8$; (b) $s_1=s_2=15$.

$$\nu=\sum_{i=1}^{m} s_i f(k_i,c_M). \tag{41}$$

In such a case, a relationship between microscopic and macroscopic parameters like Eq. (40) can also be written as

$$1=\sum_{i=1}^{m} \frac{s_i}{j} f\left(k_i,\frac{1}{K_j}\right) \tag{42}$$

which generalizes the relationship between the Adair approach with the affinity spectrum method given by Eq. (40) for a wide variety of local isotherms used.

VI. CONCLUDING REMARKS

The Darwin–Fowler method, a classical result of statistical thermodynamics, has been reviewed in order to obtain approximate expressions for the stoichiometric constants of the complexation of small ligands to macromolecules with a number of sites so high that thermodynamical limiting conditions are used. The Darwin–Fowler method is based on the so-called saddle point approximation. The goodness of such approximation to analise complexation problems is studied in this work.

Very simple expressions for the value of the first $s-2$ Gibbs energies of the sequential complexation processes have been deduced. These Gibbs energies are calculated directly from the binding curve without making any hypothesis on the model of complexation involved, and contain two contributions; the first one is an integral term corresponding to the area under the curve $\ln c_M$ vs the mean occupation number between j and $j+1$. The second contribution, which can be considered as a corrective term, depends on the binding curve derivatives on the points with an integer value for the mean occupation number. This last term can be neglected for high s values.

The value of the stoichiometric stability constants K_j tends to the inverse of the free ligand concentration at the point of mean occupation number equal to j as the number of sites s increases. This provides a simple interpretation of the average equilibrium function, K_c, which takes the value of the excess stability constant, K_j^E, at integer values of the mean occupation number ($\nu=j$).

The expressions derived are applied to a simple heterogeneous complexation case and to an homogeneous complexation with interaction between bound sites. The results obtained reproduce quite well the real values of the stability constants, and the accuracy increases as s increases.

For the heterogeneous case, in the limit $s\to\infty$, we have derived very simple expressions to relate the stoichiometric stability constants (a macroscopic description of the complexation) with the intrinsic or microscopic stability constants represented in the affinity spectrum.

APPENDIX: APPLICATION OF THE SADDLE POINT APPROXIMATION TO THE MACROMOLECULAR BINDING

Coefficients b_j of the macrocanonical partition function are obtained as the result of the integral (19)

$$b_j=\frac{1}{2\pi i}\oint_C \frac{\exp(sF(z))}{z^{j+1}}dz, \tag{A1}$$

where z takes values in the complex plane and C is a closed path on this plane around $z=0$.

We have defined an auxiliar function $\phi_j(z)$ in Eq. (20) as

$$\phi_j'(z)\equiv\frac{\exp(sF(z))}{z^{j+1}}=\frac{\Xi(z)}{z^{j+1}}=\frac{\Sigma_{i=0}^{s} b_i z^i}{z^{j+1}}$$

$$=\sum_{i=0}^{s} b_i z^{i-j-1}. \tag{A2}$$

Equation (A2) indicates that the function $\phi_j(z)$, the integrand of Eq. (A1), is an analytical function except for $z=0$. Therefore, $\phi_j(z)$, is an harmonic function[22]

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\phi_j(z)=0, \tag{A3}$$

where x and y indicate, respectively, the real and imaginary components of the complex variable z.

Restricting z to real values, whereas $j<s-1$, $\phi_j(z)$ has a minimum along this x-axis at some point $z=z_j$ because $\exp(sF(z))$ is a monotonically increasing function (a polynomial of degree s with all the coefficients b_i being positive) and $1/z^{j+1}$ is a monotonically decreasing function along the real axis. Hence Eq. (A3) for $z=z_j$ leads to

$$\phi_j'(z_j)=\left(\frac{d\phi_j}{dz}\right)_{z=z_j}=0; \quad \left(\frac{\partial^2\phi_j}{\partial x^2}\right)_{z=z_j}>0; \quad \left(\frac{\partial^2\phi_j}{\partial y^2}\right)_{z=z_j}<0; \quad (j=0,...,s-1), \tag{A4}$$

where the first equality is derived from the condition of extreme at $z=z_j$, the second indicates the presence of a minimum in the real direction and the last condition comes from Eq. (A3).

$\phi_j(z)$ is an analytical function and satisfies the Cauchy–Riemann equations,[22] which ensures that the condition of extreme for $\phi_j(z_j)$ yields to $(\partial\phi_j/\partial x)_{z=z_j}=-i(\partial\phi_j/\partial y)_{z=z_j}=0$, which together with the last inequality of Eq. (A4) indicates the simultaneous presence of a maximum for $\phi_j(z)$ along the imaginary direction at $z=z_j$. Therefore, z_j is a saddle point on the complex plane. For $j=s-1$ or $j=s$ there is no minimum for $\phi_j(z)$ along the real axis, since function $\phi_j(z)$ decreases monotonically.

It is easy to compute the real value z_j corresponding to the minimum of $\phi_j(z)$. The condition of extreme, $\phi_j'(z=z_j)=0$, using Eq. (A2) becomes

$$\left(\frac{\partial \ln \Xi(z)}{\partial \ln z}\right)_{z=z_j}=j+1, \tag{A5}$$

which, taking into account that the lhs (≡left hand side) of Eq. (A5) corresponds to the mean occupation number, $\nu(5)$, indicates that at $z=z_j$, $\nu=j+1$. Thus, z_j corresponds to the concentration c_M for which $\nu=j+1$, labeled as c_{j+1},

$$z_j=c_{j+1}; \quad \nu(c_{j+1})=s\,\theta(c_{j+1})=j+1. \tag{A6}$$

To calculate the second derivative of $\phi_j(z)$ at $z=z_j$ we use the auxiliar function

$$g_j(z)\equiv\frac{1}{s}\ln\phi_j(z), \tag{A7}$$

leading to

$$\phi_j''(z=c_{j+1})=s g_j''(c_{j+1})\phi_j(c_{j+1}) = s g_j''(c_{j+1})\exp\{s g_j(c_{j+1})\}, \tag{A8}$$

where it has been used the condition of extreme $\phi_j'(c_{j+1})=g_j'(c_{j+1})=0$.

Equation (A8) gives the concavity of $\phi_j(z)$ at $z=z_j$. It indicates that the concavity increases with s leading to a sharper maximum on the imaginary direction and a deeper minimum on the real direction appear. If we choose the contour of integration of Eq. (A1) to be a circle about $z=0$ with radius c_{j+1}, the contour will pass through z_j in the imaginary direction, in which point and direction the integrand has an extremely sharp maximum.

This is the highest maximum on the circle. The absolute value of $\phi_j(z)$ over the circle with $|z|=c_{j+1}$ is

$$|\phi_j(z)|=\frac{1}{|z|^{j+1}}\left|\sum_{i=0}^{s} b_i z^i\right|, \tag{A9}$$

which reaches its maximum value when all the terms are real. As all the b_i coefficients are real and positive, all the terms are also real and positive if and only if $z=z_j$. Hence, $\phi_j(z)$ reaches the highest maximum along the circle $|z|=c_{j+1}$ over the real axis.

As there is no maximum comparable in height along the contour, the main contribution to the integral (A1) comes from the neighborhood of c_{j+1}. Expanding $g_j(z)$ around $z=c_{j+1}$,

$$g(z)=g(c_{j+1})+\tfrac{1}{2}g''(c_{j+1})(z-c_{j+1})^2+\cdots \tag{A10}$$

replacing Eq. (A10) in Eq. (A1), and using the polar form of z and dz, providing that the circle of integration has the radius $|z|=c_{j+1}$ ($z=c_{j+1}e^{i\varphi}$, $dz=ic_{j+1}e^{i\varphi}d\varphi$), the integral (A1) becomes

$$b_j\cong\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Re}\{\exp(s\{g_j(c_{j+1}) + \tfrac{1}{2}g_j''(c_{j+1})(c_{j+1}e^{i\varphi}-c_{j+1})^2\})c_{j+1}e^{i\varphi}\}d\varphi, \tag{A11}$$

where it has been taken as the real part since b_j is real.

From the definition of $g(z)$ and Eq. (A2), $g''(c_{j+1})$ can be evaluated as

$$g''(c_{j+1})=\frac{1}{s}\frac{1}{c_{j+1}^2}\left(\frac{d\nu}{d\ln c_M}\right)_{c_M=c_{j+1}}. \tag{A12}$$

If the integral (A11) is performed only in a small interval around c_{j+1}, $e^{i\varphi}$ can be replaced by $1+i\varphi$, and taking into account that the only relevant contribution to the integral is located in this small interval, the limits of integration can be extended to $(-\infty,\infty)$.

Then, substituting Eq. (A12) into the integral (A11), the evaluation of the resulting quadratic integral yields

$$\ln b_j=-\int_{-\infty}^{\ln c_{j+1}}\nu(c_M)d\ln c_M-(j+1)\ln c_{j+1} - \frac{1}{2}\ln\left\{2\pi\frac{1}{c_{j+1}^2}\left(\frac{d\nu}{d\ln c_M}\right)_{c_M=c_{j+1}}\right\} \tag{A13}$$

and integrating the first term of the rhs of Eq. (A13) by parts, this equation can be written as

$$\ln b_j=-\int_0^{j+1}\ln c_M\,d\nu - \frac{1}{2}\ln\left\{2\pi\frac{1}{c_{j+1}^2}\left(\frac{d\nu}{d\ln c_M}\right)_{c_M=c_{j+1}}\right\}. \tag{A14}$$

GLOSSARY OF SYMBOLS

Latin letters

a_j	dimensionless Adair coefficient
b_j	Adair coefficient
B	binding capacity
c_{M_jP}	molar concentration of species M_jP $(j=1,...,s)$
c_j	molar concentration of free ligands when $\nu=j$
c_s	molar concentration of free sites
c_M	molar concentration of free ligands
c_{MS}	molar concentration of bound sites
$c_{T,S}$	total molar concentration of sites (free and bound)
$f(k,c)$	generalized local isotherm for complexation
$F(c)$	auxiliar function related to the logarithm of the macrocanonical partition function
$g_j(z)$	auxiliar complex function related to $\phi_j(z)$
G_j^0	standard Gibbs energy necessary to complex j ligands
$\Delta_j G^0$	standard Gibbs energy for the process of complexation of the jth ligand
$\Delta_j G^E$	excess Gibbs energy for the process of complexation of the jth ligand
i	specific kind of site in heterogeneous complexation; $i=1,...,m$
j	number of bound ligands
k_B	Boltzmann constant
k_i	intrinsic stability constant of a complexing site of the ith type
K	thermodynamical stability constant for formal species in molar concentration scale (in M^{-1} units), which coincides with the limit of K_c when $c_M \to 0$
K_c	average equilibrium constant
K_j	stoichiometric stability constant in an ideal dilute solution (from $M_{j-1}P$ to M_jP), (in M^{-1} units)
K_j^E	excess equilibrium constant
m	number of different kind of sites in heterogeneous macromolecule
M	free ligand
M_jP	macromolecule species with j bound ligands
P	naked macromolecule
R	gas constant
s	total number of sites in one macromolecule
s_i	number of sites of ith type in the heterogeneous macromolecule
T	absolute temperature
$w_{i,j}$	mean number of the free sites of the ith type
x	real part of the complex variable z
x_i	fraction of free sites of the ith type with respect to the total number of free sites
y	imaginary part of the complex variable z
z	complex variable $z=x+yi$

Greek letters

$\beta=1/k_BT$	inverse of the thermic energy
δ	interaction energy between adjacent bound ligands
$\phi_j(z)$	auxiliar complex function related to the complex macrocanonical partition function
φ	phase of a complex variable (dimensionless real variable)
ν	mean occupation number
ν_i	mean occupation number of sites of the ith type
$\nu_{i,j}$	mean occupation number of sites of the ith type when $\nu=j$
θ	coverage or fraction of occupied sites in a macromolecule
$\theta_{i,j}$	coverage of sites of the ith type when $\nu=j$
Ξ	macrocanonical partition function of the system $\{M_jP; j=0,...,s\}$

Superindices

id	ideal
0	standard state

ACKNOWLEDGMENTS

One of us (J.L.G.) benefitted from a grant from the Generalitat de Catalunya (Catalonia, Spain). The authors acknowledge the support to this research by the Spanish Ministry of Education and Science (DGICYT, Project No. PB 96-0379), by the ''Comissionat d' Universitats i Recerca'' of the Generalitat de Catalunya and by the ''Paeria de Lleida'' (Catalonia, Spain).

[1] C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules* (Freeman, New York, 1980), Chap. 15.
[2] J. Wyman and S. J. Gill, *Binding and Linkage: Functional Chemistry of Biological Macromoleciles* (University Science, Mill Valley, 1990), Chap. 2.
[3] J. Buffle, *Complexation Reactions in Aquatic Systems*, Ellis Horword Series in Analytical Science (Ellis Horwood Limited, Chichester, 1988), Chap. 5.
[4] W. H. van Riemsdijk and L. K. Koopal, in *Environmental Particles*, edited by J. Buffle and H. P. van Leeuwen (Lewis, Michigan, 1992), Vol. 1, Chap. 12.
[5] E. di Cera, *Thermodynamic Theory of Site-Specific Binding Processes in Biological Macromolecules* (Cambridge University Press, New York, 1995), Chap. 2.
[6] C. Tanford, *Physical Chemistry of Macromolecules* (Wiley, New York, 1961).
[7] J. Galceran, J. Salvador, J. Puy, F. Mas, and H. P. van Leeuwen, J. Electroanal. Chem. **391**, 29 (1995).
[8] T. L. Hill, *Cooperativity Theory in Biochemistry: Steady-State and Equilibrium Systems* (Springer, New York, 1984).
[9] J. L. Garcés, F. Mas, J. Puy, J. Galceran, and J. Salvador, J. Chem. Soc., Faraday Trans. **94**, 2783 (1998).
[10] I. M. Klotz, *Ligand–Receptor Energetics: A Guide for the Perplexed* (Wiley, New York, 1997).
[11] A. V. Hill, J. Physiol. (London) **40**, iv (1910).
[12] G. Scatchard, Ann. (N.Y.) Acad. Sci. **51**, 660 (1949).
[13] I. Ruzic, Anal. Chim. Acta **140**, 99 (1982).
[14] R. Brodersen, F. Nielsen, J. C. Christiansen, and K. Andersen, J. Biochem. **169**, 487 (1987).
[15] E. di Cera, Biophys. Chem. **37**, 147 (1990); J. Chem. Phys. **96**, 6515 (1992).
[16] M. Cernik, M. Borkovec, and J. C. Westall, Environ. Sci. Technol. **29**, 413 (1995).
[17] J. L. Garcés, F. Mas, and J. Puy, Environ. Sci. Technol. **32**, 539 (1998).
[18] I. M. Klotz, Proc. Natl. Acad. Sci. USA **90**, 7191 (1993).
[19] K. Huang, *Statistical Mechanics*, 2nd ed. (Wiley, New York, 1987), Chap. 9.
[20] J. L. Garcés, F. Mas, J. Cecília, J. Galceran, J. Salvador, and J. Puy,

Analyst (Cambridge, U.K.) **21**, 1855 (1996).
[21] L. E. Reichl, *A Modern Course in Statistical Physics*, 2nd ed. (Wiley, New York, 1998), Chap. 2.
[22] R. M. Spiegel, *Complex Variables*, 1st ed. (McGraw–Hill, New York, 1967) Spanish translation by McGraw–Hill de México, 1991.
[23] M. V. José and C. Larralde, Math. Biosci. **58**, 159 (1982).
[24] M. V. José, Anal. Biochem. **144**, 494 (1985).
[25] A. R. Wolfe and T. Meehan, J. Mol. Biol. **223**, 1063 (1992).
[26] H. P. van Leuween and J. Buffle, J. Electroanal. Chem. Interfacial Electrochem. **296**, 359 (1990).