**BPC 01326** 

# Cooperative ligand-lattice binding

## Approximate Gaussian binding distribution

## Johannes Reiter and Irving R. Epstein

Department of Chemistry, Brandeis University, Waltham, MA 02254, U.S.A.

Received 13 June 1988
Revised manuscript received 15 September 1988
Accepted 26 September 1988

Cooperativity; Ligand binding; Infinite lattice approximation

Nearest-neighbor cooperative binding of a ligand covering n sites and binding with equilibrium constant K and cooperativity factor  $\omega$  to a large molecule with m binding sites ( $m \gg n\omega$ ,  $n/\omega$ ) can be approximately described by a Gaussian distribution  $P(q-q_{\max})$ , where q is the number of ligands bound and  $q_{\max}$  the most probable value of q. The variance of the Gaussian is equal to the derivative  $dq_{\max}/d\ln(L)$ , where L is the free ligand concentration. This variance,  $\sigma^2$ , is a complicated function of  $q_{\max}$ . However, in the limits of very large cooperativity,  $\omega \gg 1$ , very large anticooperativity,  $\omega \ll 1$ , or noncooperativity,  $\omega \approx 1$ , simpler expressions for  $\sigma^2$  can be given. For  $q_{\max} = m/(n+1)$ , where the most probable number of bound ligands equals the number of free binding sites,  $\sigma^2$  has a particularly simple form:  $\sigma^2 = 2m\omega^{1/2}/(n+1)^3$ . The Gaussian and the infinite lattice approximations for the average number of ligands bound are good approximations only if  $\sigma$  is much smaller than the number of binding sites. The variance may therefore provide an easy check on the validity of the infinite lattice approximation, which is commonly used to analyze experimental binding data.

#### 1. Introduction

In many biochemical systems the binding of ligands to a large molecule (lattice) with multiple binding sites, such as a nucleic acid strand, is cooperative. Therefore, there is a considerable interest in calculating cooperative binding equilibria [1]. Most often, a one-dimensional nearest-neighbor interaction model is used to describe the binding [2-6]. In such models, the concentration (probability) distribution of lattices with q ligands bound can easily be found [3,6,7]. For large lattices, this unimodal binding distribution is well approximated by its maximum. Solutions for the most probable number of ligands bound have been given, for instance, by Zasedatelev et al. [3]

Correspondence address: I.R. Epstein, Department of Chemistry, Brandeis University, Waltham, MA 02254, U.S.A.

using a combinatorial method and by McGhee and Von Hippel [5] with a method based on conditional probabilities.

Many experiments are performed with relatively short lattices, but investigators often use the equations derived for infinite lattices to analyze their results. Differences between short lattices and the infinite lattice approximation have usually been attributed to end effects, i.e., the left-most and right-most ligands on a short linear lattice have no neighbor or free binding site on one side, in contrast to the situation on an infinite lattice [8]. Also, if the ligand occupies n binding sites, the (n-1) sites at either end of the lattice cannot serve as the site to which one end of the ligand binds. Another, generally neglected, source of error is that the binding distribution of small lattices is broad, making the assumption of a narrow, unimodal distribution a poor approximation. For instance, as first demonstrated for the helix-coil equilibrium in DNA by Applequist and Damle [9], small lattices with strong cooperative binding can even have bimodal distributions, i.e., two maxima [7,10,11].

In this work, we examine under what circumstances the binding distribution is well approximated by a Gaussian. In such cases, the standard deviation of the Gaussian serves as an estimate for the width of the binding distribution.

#### 2. Results

# 2.1. Binding distribution for the nearest-neighbor cooperativity model

The model most often used to describe cooperative binding assumes cooperativity only between nearest neighbors. A ligand which covers n sites binds with equilibrium constant  $K\omega^{nn}$ , where K is the equilibrium constant, ω the cooperativity factor, and nn the number of nearest neighbors of the ligand; for a one-dimensional lattice, nn may be 0, 1 or 2. The equilibrium concentration of the free ligand, L, is taken to be constant for a given distribution. The binding distribution has been given by Zasedatelev et al. [3] for an infinite lattice. Epstein [6] and Reiter and Epstein [7] give expressions for finite linear and circular lattices, respectively. We follow the nomenclature of Reiter and Epstein. In the limit of a large lattice, where q-1 can be replaced by q, and m-nq+1by m - nq, the concentration (probability) distribution is:

$$P(q,c) = \begin{bmatrix} q \\ c \end{bmatrix} \begin{bmatrix} m-nq \\ c \end{bmatrix} (KL\omega)^{q} \omega^{-c}$$
 (1)

where m is the lattice length, i.e., the number of binding sites, q the number of ligands bound to the lattice,  $0 \le q \le m/n$ , and c the number of ligand clusters, which can take values between 0 and the minimum of q and m-nq, the latter being the number of free binding sites. For a large lattice, i.e., one with  $m \gg n\omega$  and  $m \gg n/\omega$ , the distribution possesses a single maximum at  $q_{\text{max}}$  ligands bound. If the distribution has a narrow maximum, the average number of ligands bound

can be set equal to that maximum. For a given  $q_{\text{max}}$ , the distribution in c has only one maximum.  $q_{\text{max}}$  and  $c_{\text{max}}$  can be obtained by approximating the factorials in eq. 1 with the Stirling formula and setting the derivatives with respect to q and c equal to zero. Since the variables q and c are discrete, we calculate the maxima by setting the ratios P(q+0.5,c)/P(q-0.5,c) and P(q,c+0.5)/P(q,c-0.5) equal to unity. Using the approximation

$$(D+\Delta)!/(D-\Delta)! \approx D^{2\Delta}e^{\Delta/D}, \qquad D \gg \Delta \quad (2)$$

we obtain with our first condition:

$$FQ = P(q+0.5,c)/P(q-0.5,c)$$
=  $q(m-nq-c)^{n} KL\omega/[(q-c)(m-nq)^{n}]$ 
= 1 (3)

and with the second condition:

$$FC = P(q,c+0.5)/P(q,c-0.5)$$
  
=  $(q-c)(m-nq-c)/(c^2\omega) = 1$  (4)

In both cases we have neglected the exponential terms in agreement with our assumption of a large lattice. Only one of the two solutions for c in eq. 4 is physically reasonable. On inserting this value of  $c_{\max}$  into eq. 3, we can solve numerically to obtain  $q_{\max}$  [3]. Again, only one solution of  $q_{\max}$  is physically reasonable. McGhee and Von Hippel [5] give a more intuitive derivation of equivalent equations with their method of conditional probabilities.

Eq. 4 gives the maximal value of c for any given q, independent of KL. We reserve the symbol  $c_{\max}$  for the maximal value of c corresponding to  $q_{\max}$ . We also note that while in general  $q_{\max}$  can only be obtained numerically for a given KL and a given  $\omega$ , if  $q_{\max}$  and  $\omega$  are chosen first,  $c_{\max}$  and KL can be calculated analytically.

#### 2.2. Gaussian approximation

Our approach will be to obtain a two-dimensional Gaussian approximation to the exact distribution (eq. 1) and then to sum (actually integrate) that expression over c to generate a Gaussian binding distribution for q. Since we seek the width of the distribution with respect to q, we choose

 $c_{\max}$  to maximize the probability for the given  $q_{\max}$ , and we express the deviation,  $\Delta c$ , from  $c_{\max}$  as a function of  $\Delta q$ . Near  $q_{\max}$  we then have the following distribution:

$$\frac{P(q_{\text{max}} + \Delta q, c_{\text{max}} + \Delta c)}{P(q_{\text{max}}, c_{\text{max}})}$$

$$= \frac{(q_{\text{max}} + \Delta q)!}{(q_{\text{max}})!} \frac{(q_{\text{max}} - c_{\text{max}})!}{(q_{\text{max}} - c_{\text{max}} + \Delta q - \Delta c)!}$$

$$\times \frac{(c_{\text{max}}!)^2}{(c_{\text{max}} + \Delta c)!^2} \frac{(m - nq_{\text{max}} - n\Delta q)!}{(m - nq_{\text{max}})!}$$

$$\times \frac{(m - nq_{\text{max}} - c_{\text{max}})!}{(m - nq_{\text{max}} - n\Delta q - c_{\text{max}} - \Delta c)!}$$

$$\cdot (KL\omega)^{\Delta q} \omega^{-\Delta c} \tag{5}$$

where KL has to be calculated as a function of  $\omega$ ,  $q_{\max}$  and  $c_{\max}$ . A linear approximation to the change in the maximal value of c as q is varied can be easily calculated from FC (eq. 4) using implicit differentiation:

$$dc/dq = [-m + 2nq - (n-1)c]/$$

$$[2c(1-\omega) + (n-1)q - m]$$
 (6)

The linear approximation is always valid if  $\omega \Delta q \ll m$  and  $\Delta q/\omega \ll m$ . This can be seen by expanding  $\Delta c$  as a power series in  $\Delta q$  using eq. 4; for most values of  $q_{\rm max}$  the condition  $\Delta q \ll m$  is sufficient. In the following, we always assume that eq. 6 is evaluated at  $q_{\rm max}$  and  $c_{\rm max}$ .

To calculate the approximate Gaussian distribution corresponding to eq. 1 we take  $q = q_{\text{max}} + \Delta q$  and  $c = c_{\text{max}} + \Delta q \cdot \text{d}c/\text{d}q + \Delta c$ , where eq. 6 is used for the derivative that gives the linear change in  $c_{\text{max}}$  and  $\Delta q$  and  $\Delta c$  express the deviations from the maxima of the distribution. We evaluate eq. 5 with the approximation

$$(D + \Delta)!/D! \approx D^{\Delta}e^{\Delta(\Delta+1)/(2D)}, \qquad D \gg \Delta$$
 (7)

Since we calculate small deviations from  $q_{\rm max}$ , we cannot neglect the exponential terms from the approximation (eq. 2) in FC and FQ (eqs. 3 and 4), but have to incorporate them into KL by setting

$$P(q_{\text{max}} + 0.5, c_{\text{max}} + 0.5 \cdot dq/dc) /$$

$$P(q_{\text{max}} - 0.5, c_{\text{max}} - 0.5 \cdot dq/dc) = 1$$
(8)

using the approximation, eq. 2, and keeping the exponential terms. Calculating KL from eq. 8 ensures that  $q_{\text{max}}$  and  $c_{\text{max}}$  are indeed the maximal values within the given approximation. Using this value of KL in eq. 5 together with eqs. 6 and 7 cancels all factors except the exponential terms in  $\Delta q^2$ ,  $\Delta c^2$  and  $\Delta q \Delta c$ :

$$P(q_{\text{max}} + \Delta q, c_{\text{max}} + \Delta q \cdot dc/dq + \Delta c)/$$

$$P(q_{\text{max}}, c_{\text{max}})$$

$$= \exp(-\Delta q^2/2\sigma_q^2 - \Delta c^2/2\sigma_c^2 - \Delta q\Delta c/\sigma_{qc}^2) \quad (9)$$

with

$$1/\sigma_q^2 = -1/q + 2(dc/dq)^2/c$$

$$+ (1 - dc/dq)^2/(q - c) - n^2/(m - nq)$$

$$+ (n + dc/dq)^2/(m - nq - c) \qquad (10a)$$

$$1/\sigma_c^2 = 1/(q - c) + 1/(m - nq - c) + 2/c \qquad (10b)$$

$$1/\sigma_{qc}^2 = -(1 - dc/dq)/(q - c)$$

$$+ (n + dc/dq)/(m - nq - c)$$

$$+ 2(dc/dq)/c \qquad (10c)$$

All quantities in eqs. 10 are to be evaluated at  $q_{\rm max}$  and  $c_{\rm max}$ .  $\Delta q$  can take on positive and negative values and is physically restricted to the range  $0 \le q_{\rm max} + \Delta q \le m/n$ . Also,  $0 < q_{\rm max} < m/n$ , because no Gaussian exists if  $q_{\rm max}$  equals zero or m/n.

The quantity we seek is the variance in the binding distribution, i.e., in the distribution P(q) that remains after integrating eq. 9 over all possible values of  $\Delta c$ . This integration is greatly facilitated by the observation that, thanks to eqs. 4 and 6, the coupling term (eq. 10c) vanishes. Since  $1/\sigma_c^2$  (eq. 10b) is independent of both  $\Delta q$  and  $\Delta c$ , the integration over  $\Delta c$  changes neither the form nor the variance of the Gaussian distribution for  $q=q_{\rm max}+\Delta q$ . Using the fact that the right-hand side of eq. 10c is zero, we can simplify this variance (eq. 10a) to give

$$1/\sigma_q^2 = (1 - dc/dq)/(q - c) + n(n + dc/dq)/(m - nq - c) - 1/q - n^2/(m - nq)$$
(11)

Table 1

Particular solutions evaluated at  $q = q_{max}$ 

	Timex		
	Спях	dc/dq	$\sigma_q^2$
$q_{\max} = \frac{m}{(n+1)}$	$\frac{m}{(n+1)(1+\sqrt{\omega})}$	$\frac{-(n-1)}{2(1+\sqrt{\omega})}$	$\frac{2m\sqrt{\omega}}{(n+1)^3}$
s = 1	$\frac{(m-nq)q}{m-(n-1)q}$	$\frac{(m-2nq)(m-(n-1)q)+(n-1)(m-nq)q}{(m-(n-1)q)^2}$	$\frac{q(m-nq)(m-(n-1)q)}{m^2}$
£ \$	$\frac{b(bu-w)}{b}$	$\frac{m-2nq}{2\sqrt{(m-nq)q\omega}}$	$\frac{2\sqrt{\omega}\left(q\left(m-nq\right)\right)^{3/2}}{m^2}$
$\omega \ll 1$ $0 < q < m/(n+1)$ $ q - m/(n+1)  \gg 1$	$q - \frac{q^2 \omega}{(m - (n+1)q)}$	$1 - \frac{2q(m - (n+1)q) + (n+1)q^2}{(m - (n+1)q)^2} \omega$	$\frac{q(m-nq)(m-(n+1)q)}{m^2}$
$\omega \ll 1$ $m > q > m/(n+1)$ $ q-m/(n+1)  \gg 1$	$m - nq + \frac{\left(m - nq\right)^2 \omega}{\left(m - \left(n + 1\right)q\right)}$	$-n + \frac{-2n(m-nq)(m-(n+1)q) + (n+1)(m-nq)^{2}}{(m-(n+1)q)^{2}}$	$-\frac{q(m-nq)(m-(n+1)q)}{m^2}$

We observe that the variance just calculated is identical to the total derivative  $dq_{max}/d \ln(L)$ , i.e., the tangent to the titration curve, obtained by implicit differentiation of eq. 3, i.e.,

$$dq_{\text{max}}/d\ln(L) = \sigma_a^2 \tag{12}$$

An alternative route to eq. 12 utilizes the relations

$$\partial \{ P(q + \Delta q) / P(q) \} / \partial \ln(L)$$

$$= \Delta q \cdot \{ P(q + \Delta q) / P(q) \}$$
(13)

where eq. 1 for P(q,c) is summed over c and evaluated at  $q = q_{\text{max}}$  to give P(q), and

$$\frac{\partial \{P(q + \Delta q)/P(q)\}}{\partial q} = -\Delta q \{P(q + \Delta q)/P(q)\}/\sigma_q^2 \tag{14}$$

where the approximate Gaussian distribution summed over c and evaluated at  $q = q_{\text{max}}$  is used. We then equate the two distributions and assume the applicability of the implicit function theorem.

If the number of lattices is large, the distribution (eq. 1) gives the concentration of lattices with

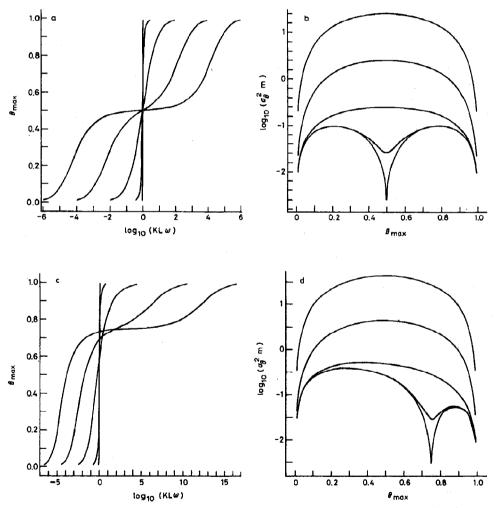


Fig. 1. Titration curves calculated with the infinite lattice approximation for cooperativity factors  $\omega$  of  $10^4$ ,  $10^2$ , 1,  $10^{-2}$  and  $10^{-4}$  are shown in panels a and c. The steeper curves correspond to higher cooperativity factors. In panels b and d, the variance of the approximate Gaussian binding distribution is shown, multiplied by the lattice length m. The variance is related to the derivatives of the curves in panels a and c, i.e.,  $(m/n)\sigma_\theta^2 = d\theta_{\max}/d \ln(L)$  (see text). The curves calculated for higher cooperativities have larger variances. (a,b) n=1; (c,d) n=3.

q ligands bound (and c clusters). Alternatively, it may be viewed as the probability for an individual lattice out of the ensemble of lattices to have q ligands bound. In the first case, the normalized P(q,c) sum to the total lattice concentration, in the second to unity. If both the ligand and the lattice concentrations are very small, a stochastic description must be used.

Since  $c_{\rm max}$  is the solution of a quadratic, and only one solution makes physical sense, we can always calculate the variance of the Gaussian for a given  $q_{\rm max}$  with the equations above. In the special case

$$q_{\max} = m - nq_{\max} = m/(n+1) \tag{15}$$

where the number of bound ligands equals the number of free binding sites, the distribution in c is the square of a binomial distribution, and a simple solution can be obtained for the variance (table 1). Simple solutions are also available for  $\omega = 1$ ,  $\omega \gg 1$  and  $\omega \ll 1$  (see table 1).

To compare different sized lattices and ligands, it is useful to calculate the variance of the fraction of occupied binding sites,  $\theta = nq/m$  ( $0 \le \theta \le 1$ ). We then have  $\sigma_{\theta}^{\ 2} = (n\sigma_{q}/m)^{2}$  and  $\mathrm{d}\theta_{\mathrm{max}}/\mathrm{d}\ln(L) = (n/m) \,\mathrm{d}q_{\mathrm{max}}/\mathrm{d}\ln(L)$ .  $\sigma_{\theta}^{\ 2}$  is inversely proportional to the lattice length, while  $\mathrm{d}\theta_{\mathrm{max}}/\mathrm{d}\ln(L)$  is independent of the lattice length in the limit of a large lattice.

Simple solutions for KL can be given for  $\theta_{\text{max}} = 0.5$  and for  $\theta_{\text{max}} = n/(n+1)$  ( $q_{\text{max}} = m/(n+1)$ ) in certain cases by substituting  $c_{\text{max}}$  from table 1 into eq. 3. We add these solutions for completeness when discussing the titration curves in fig. 1. Schneider et al. [12] have investigated the case n = 1; for n > 1, various results are given in refs. 2-6.

Fig. 1 shows the variance of the distribution in  $\theta$  (multiplied by m) or of the derivative of  $\theta_{\text{max}}$  with respect to  $\ln(L)$  (multiplied by n). For large cooperativities the distribution is broad, and the variance and the steepness of the transition are proportional to the square root of the cooperativity factor. The latter is well known as all-or-none behavior at large cooperativities. The maximum variance occurs at

$$\theta_{\text{max}} = 0.5, \qquad \sigma_{\theta}^2 = (n\omega)^{1/2}/(4m).$$

It is proportional by a factor depending only on n to the value at

$$\theta_{\text{max}} = n/(n+1), \qquad \sigma_{\theta}^2 = 2n^2\omega^{1/2}/(m(n+1)^3).$$

Using  $c_{\text{max}}$  from table 1, we find  $KL\omega = 1$  for  $\theta_{\text{max}} = 0.5$  at large cooperativities and  $KL\omega = (1 + \omega^{-1/2})^{n-1}$  for  $\theta_{\text{max}} = n/(n+1)$  at all cooperativities. The latter value of  $KL\omega$  also approaches unity for large cooperativities, as the titration curve becomes nearly vertical when  $KL \approx 1/\omega$  for  $\omega$  large.

In the noncooperative case, the variance takes on its maximum at  $\theta_{\text{max}} \le 0.5$ , and may be found from the equation for the variance given in table 1. At  $\theta_{\text{max}} = 0.5$ ,  $KL = (1 + 1/n)^{n-1}/n$  and  $\sigma_{\theta}^2 = (n+1)/8m$ ; at  $\theta_{\text{max}} = n/(n+1)$ ,  $KL = 2^{n-1}$  and  $\sigma_{\theta}^2 = 2n^2/(m(n+1)^3)$ .

In the case of anticooperativity, the variance at  $\theta_{\text{max}} = n/(n+1)$  is still proportional to the square root of the cooperativity factor, but attains a minimum at that value. This minimum results from the special stability of the configuration in which vacant sites alternate with bound ligands. In fact, the binding of an n-site anticooperative ligand is roughly equivalent to the noncooperative binding of a ligand which covers n+1 binding sites [5] with the same equilibrium constant K. The (n+1)-st site allows for the effects of the anticooperativity. For example, in table 1, we see that the variance calculated for large anticooperativities can be obtained from the expression for the noncooperative case simply by replacing n by n+1. This expression is not valid near  $\theta_{max} =$ n/(n+1), where the variance has to be calculated from eq. 11.

#### 2.3. Short lattices

We focus on short circular lattices, because for circular lattices the influence of the width of the binding distribution can be observed independent of end effects. Short linear lattices will show both the effects noted here and the additional end effects discussed earlier. Since the Gaussian is a continuous distribution, it should be discretized by integrating from q-1/2 to q+1/2 to calculate the probability of having q ligands bound.

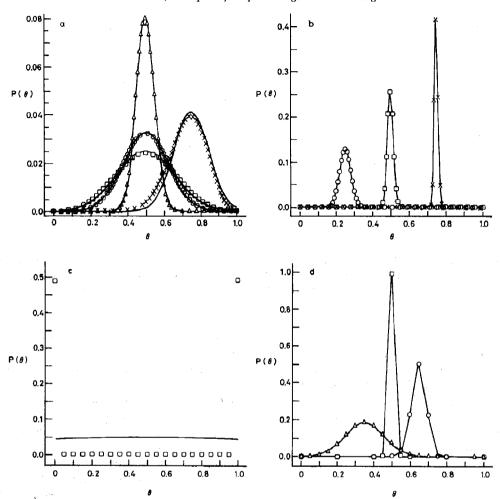


Fig. 2. Exact binding distributions for finite (small) circular lattices are compared with Gaussian approximations (solid lines). KL has been chosen such that  $\theta_{\text{max}}$  of the Gaussian equals the average number of bound ligands calculated with the exact binding distribution. Panel a: ( $\square$ ) m=100, n=1,  $\omega=100$ ,  $\theta_{\text{max}}=0.5$ ; ( $\bigcirc$ ) m=300, n=3,  $\omega=100$ ,  $\theta_{\text{max}}=0.5$ ; ( $\bigcirc$ ) m=100, n=1,  $\omega=1$ ,  $\theta_{\text{max}}=0.5$ ; panel b: ( $\square$ ) m=100, n=1,  $\omega=0.01$ ,  $\theta_{\text{max}}=0.5$ ; ( $\bigcirc$ ) m=100, n=1,  $\omega=0.01$ ,  $\theta_{\text{max}}=0.25$ ; ( $\bigcirc$ ) m=300, n=3,  $\omega=0.01$ ,  $\theta_{\text{max}}=0.75$ ; panel c: ( $\square$ ) m=20, n=1,  $\omega=10^4$ ,  $\theta_{\text{max}}=0.5$ ; panel d: ( $\square$ ) m=20, n=1,  $\omega=10^{-4}$ ,  $\theta_{\text{max}}=0.5$ ; ( $\square$ ) m=20, n=1,  $\omega=10^{-4}$ ,  $\theta_{\text{max}}=0.5$ ; ( $\square$ ) m=20, n=1,  $\omega=10^{-4}$ ,  $\omega=10^{-4}$ 

We have simply used the probability as calculated from eq. 9. If the width of the distribution is very small, the maximum may be larger than unity (e.g., in the case of strong anticooperativity). In such cases, we scale eq. 9 to make the discrete probabilities for all permissible numbers of bound ligands sum to unity.

The Gaussian is a useful approximation to the binding distribution for short circular lattices if the width of the distribution is much smaller than the lattice length (fig. 2). In noncooperative and anticooperative binding, even very short lattices are well approximated (fig. 2d). Very short lattices with strong cooperative binding cannot be approximated by a Gaussian. Fig. 2c shows a case with a bimodal binding distribution. The standard deviation of the approximate Gaussian is much larger than the lattice length.

In fig. 3, tritration curves of small circular lattices calculated with the combinatorial binding

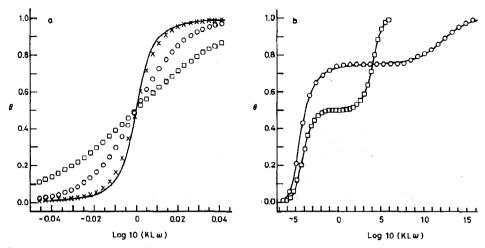


Fig. 3. Titration curves (average value of  $\theta$ ) for finite (small) circular lattices calculated with the exact combinatorial binding distribution are compared with titration curves obtained with the infinite lattice approximation ( $\theta_{\text{max}}$ ; solid lines). Panel a:  $\omega = 10^4$ , n = 1; ( $\square$ ) m = 20; ( $\bigcirc$ ) m = 40; ( $\times$ ) m = 100. Panel b:  $\omega = 10^{-4}$ ; ( $\square$ ) m = 20, n = 1; ( $\bigcirc$ ) m = 60, n = 3.

distribution [7] are compared with those obtained with the infinite lattice approximation. Small lattices with strong anticooperative binding, having a narrow binding distribution, are well approximated by the infinite lattice solution, but small lattices with strong cooperative binding, having broad and even bimodal distributions, show large deviations.

#### 3. Conclusions

The combinatorial distribution describing nearest-neighbor cooperative binding of a ligand to a molecule with many binding sites can be approximated by a Gaussian distribution for large lattices. The variance of the Gaussian is equal to the derivative  $dq_{max}/d \ln(L)$  of the titration curve. Therefore, the Gaussian approximation is most prone to failure where the amount of ligand bound is most sensitive to the free ligand concentration, i.e., where the titration curve is steepest. The approximation is satisfactory if its standard deviation is much smaller than the lattice length. In general, the width of the distribution divided by the lattice length increases with increasing cooperativity and decreases with increasing lattice length. In the case of large cooperativities and short lattices, the exact binding distribution is quite broad and may even be bimodal, so that both the Gaussian and the infinite lattice solution for the most probable number of bound ligands are poor approximations. For short lattices, the Gaussian and the infinite lattice approximations are still satisfactory at very small cooperativities or large anticooperativities. Differences between the finite lattice and the infinite lattice approximation due to end effects and possible corrections have been discussed elsewhere [8]. However, for strong cooperativity and short lattices, the infinite lattice approximation will not give correct results.

The results developed here have a number of potential practical applications. Knowledge of the binding distribution is equivalent to knowledge of the partition function. Hence, when the Gaussian approximation is applicable, thermodynamic quantities may easily be calculated analytically, since algebraic manipulations are straightforward. Secondly, if parameters are obtained from experimental data using a method (e.g., see ref. 3 or 5) based on the infinite lattice approximation, eq. 11 may be used to estimate the variance, which may then be compared with the lattice length. If  $\sigma$  is not significantly less than m, the infinite lattice approximation is not appropriate, and a more exact treatment (e.g., see refs. 6 and 7) must be

employed. Finally, eq. 12 makes possible a direct estimate of the width of the distribution from the titration curves. This information can be used to assess the validity of approximate data treatments and/or to check estimated parameters by comparing the variance obtained from eq. 12 with that given by eq. 11.

### Acknowledgements

This work was supported by the National Science Foundation under grant DMB-8604794. We thank Mark Schumaker for helpful comments on an earlier version of the manuscript.

#### References

- 1 T.L. Hill, Cooperativity theory in biochemistry: steady state and equilibrium systems (Springer, New York, 1985).
- 2 D.M. Crothers, Biopolymers 6 (1968) 575.
- 3 A.S. Zasedatelev, G.V. Gurskii and M.V. Volkenstein, Mol. Biol. 5 (1971) 245.
- 4 G. Schwarz, Eur. J. Biochem. 12 (1970) 442.
- 5 J.D. McGhee and P.H. von Hippel, J. Mol. Biol. 86 (1974) 469.
- 6 I.R. Epstein, Biophys. Chem. 8 (1978) 327.
- 7 J. Reiter and I.R. Epstein, J. Phys. Chem. 91 (1987) 4813.
- 8 C.E. Dateo and I.R. Epstein, Biopolymers 20 (1981) 1651.
- J. Applequist and V.J. Damle, J. Am. Chem. Soc. 87 (1965) 1450.
- 10 D. Pörschke and M. Eigen, J. Mol. Biol. 62 (1971) 361.
- 11 D. Pörschke, in: Chemical relaxation in molecular biology, eds. I. Pecht and R. Rigler (Springer, Berlin, 1977) p. 191.
- 12 F.W. Schneider, C.L. Cronan and S.K. Podder, J. Phys. Chem. 72 (1968) 4563.