[CONTRIBUTION FROM THE NAVAL MEDICAL RESEARCH INSTITUTE]

Titration Curves and Ion Binding on Proteins, Nucleic Acids, and Other Macromolecules with a Random Distribution of Binding Sites of Several Types

BY TERRELL L. HILL

RECEIVED JUNE 20, 1956

If a macromolecule has several types of sites for the binding of the same ion (*e.g.*, H^+), if the sites are known or assumed to be arranged in some regular (one, two or three dimensional) array but with sites of different types distributed at random in the array, and if electrostatic (or other) effects are taken care of by nearest neighbor interactions, then the theory of the titration or ion binding curve is available from a slight extension of earlier work on the localized unimolecular adsorption of a gas on a heterogeneous surface. Equations are presented which should be useful in the analysis of experimental results when the above model is appropriate. In addition, a table of effective dielectric constants is included which allows estimation of the magnitude of the electrostatic nearest neighbor interaction for charges on the surface of a sphere (in the presence of electrolyte).

I. Introduction

There are two general approaches possible in taking into account the effect of electrostatic interactions on titration or ion binding curves of macromolecules (*e.g.*, proteins, nucleic acids, linear polyelectrolytes, etc.): one can use a "smeared" net charge; or discrete interactions between nearest neighbor (and possibly higher neighbor) sites can be employed. The first point of view has been most common in the past but a rather widespread increase in interest in the second point of view has become apparent recently.[1]

The treatment of electrostatic effects *via* nearest neighbor interactions is more likely to be appropriate in the presence of electrolyte, because of the consequent reduction in the range of coulombic forces. On the other hand, a disadvantage in the case of almost all proteins, though not in general for polyelectrolytes or nucleic acids, is the lack of knowledge concerning the location of the binding sites relative to one another.

In this paper we consider nearest neighbor interactions in cases in which several types of binding sites are present. Specifically, we assume that the sites are arranged in a regular lattice (*e.g.*, in a linear array for DNA or in a two-dimensional array on the surface of a protein) but that there is a completely random distribution of types of sites among the lattice positions. This model should be useful for synthetic copolyelectrolytes (if the polymerization is believed to occur in a random manner), nucleic acids (if the sequence of bases is assumed to be random) and proteins (in the absence of detailed information about site locations). The statistical problem (*e.g.*, in nucleic acids) when the site distribution is assumed non-random (in some definable manner) is much more difficult.

Fortunately, the closely related problem of localized unimolecular adsorption of a gas on a heterogeneous surface has already been discussed.[2] An easy extension of this earlier work will provide the necessary results for our purposes.

We shall introduce below, for concreteness, two approximations[2,3]: the quasi-chemical (or Bethe) approximation and the Bragg–Williams approximation. Use of the quasi-chemical approximation with the model described above will give us equations which are, in fact, exact for (infinite) linear systems, since (a) the effects of nearest neighbor interactions with a *random* distribution of several types of sites can be shown[2] to be the same as for a system with only one type of site and (b) the quasi-chemical "approximation" is exact for a linear system with one type of site.[3] No exact theory, for one type of site, is available in two (except in the special case that half the sites are filled[3]) or three dimensions.

In computing average binding per site, we shall make the usual approximation[4] of employing equations strictly valid for an infinite number of sites per macromolecule.

II. General Equations

Consider a macromolecule with n types of sites, B_i sites of the ith type, giving a total of B sites, $B = \Sigma_i B_i$. The sites of different types are distributed at random over the B lattice positions. Each site has z nearest neighbor sites.

We take into account interactions (see Fig. 1) between: (a) the ion being bound and the site on which it is bound; (b) ions bound at nearest neighbor sites; (c) an *ion* bound at one site and the *sites* nearest neighbor to it; and (d) nearest neighbor sites. Interaction (a) is included in the intrinsic dissociation constant introduced below. Let the interaction energy (b) (strictly speaking, free energy) for a pair of ions be W. Let the interaction energy (c) between a bound ion and a nearest neighbor site of type i be W_i. The interactions of type (d) make an additive contribution to the free energy of the macromolecule, independent of the extent of binding. We can therefore ignore these

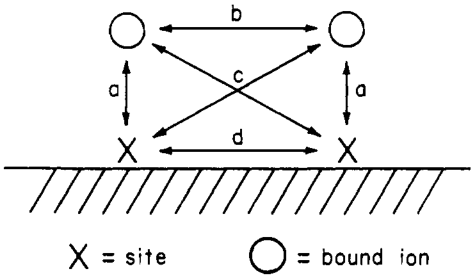

Fig. 1.—Different types of interactions.

---

[1] See, for example, the special issue of *J. Polymer Sci.*, not yet published. (International Symposium on Macromolecular Chemistry, Israel, April 1956).

[2] T. L. Hill, *J. Chem. Phys.*, **17**, 762 (1949).

[3] T. L. Hill, "Statistical Mechanics," McGraw–Hill Book Co., New York, N. Y., 1956.

[4] G. Scatchard, *Ann. New York Acad. Sci.*, **51**, 660 (1949).

interactions in the equations below having to do with binding. In the gas adsorption problem[2] only interactions (a) and (b) were involved. The interactions (c) arise here because both ions and sites may be charged. It is because of interactions (c) that we have to extend the earlier treatment[2] slightly.

The probability that a given site of, say, type i will have for nearest neighbor sites a_1 of type 1, a_2 of type 2, etc., is

$$\frac{z!}{\prod_{s} a_s!} \prod_{s} f_s^{a_s}$$

where

$$f_s = B_s/B = \text{fraction of sites of type } s$$

and

$$\sum_{s} a_s = z \tag{1}$$

Thus the number of sites of type i having a_1, a_2, . . . nearest neighbor sites of the different types is

$$B_i(\mathbf{a}) = B_i \frac{z!}{\prod_{s} a_s!} \prod_{s} f_s^{a_s} \tag{2}$$

where $\mathbf{a}$ represents a set of numbers a_1, a_2, . . . a_n satisfying eq. 1.

When an ion is bound on one of the $B_i(\mathbf{a})$ sites just described, there is an additional binding energy (beyond the interaction (a) referred to above), owing to interactions (c), of magnitude $\Sigma_s\, a_s\, W_s$.

We can now take over directly the earlier results[2] if we consider each group of $B_i(\mathbf{a})$ sites as a different "type." From eq. 30,[2] of all the B sites, the fraction θ filled is

$$\theta = \sum_{i} f_i \sum_{\mathbf{a}} \frac{\dfrac{z!}{\prod_{s} a_s!} \prod_{s} f_s^{a_s}}{\dfrac{e^{\alpha}}{j_i e^{-\Sigma_s a_s W_s/kT}} + 1} \tag{3}$$

where the inner sum is over all sets $\mathbf{a}$ consistent with eq. 1 and j_i is the partition function (including the binding or interaction energy (a)) for an ion bound at a site of type i and

$$\alpha = \frac{1}{kT}\left[-\mu + W\left(\frac{\partial \eta'}{\partial N}\right)_{B,T}\right] \tag{4}$$

In eq. 4, μ is the chemical potential of the bound ions, N is the total number of bound ions (*i.e.*, $\theta = N/B$) and $\eta' W$ is the contribution of nearest neighbor bound ion interactions to the Helmholtz free energy for a system with N bound ions on one type of site.

At equilibrium the chemical potential of the ions in solution is also μ, so that

$$\mu = \mu^* + kT \ln c \tag{5}$$

where c is the concentration (or activity) of ions in solution and μ^* is a standard or reference chemical potential. The intrinsic dissociation constant K_i for an ion bound to a site of type i is then[5]

$$K_i = \frac{e^{-\mu^*/kT}}{j_i} \tag{6}$$

Equation 3 can now be written as

$$\theta = \sum_{i} f_i \sum_{\mathbf{a}} \frac{\dfrac{z!}{\prod_{s} a_s!} \prod_{s} f_s^{a_s}}{(K_i/c)e^{\Sigma_s a_s W_s/kT}\, F(\theta) + 1} \tag{7}$$

where

$$F(\theta) = \exp\left[\frac{W}{kT}\left(\frac{\partial \eta'}{\partial N}\right)_{B,T}\right] \tag{8}$$

The above treatment of a random distribution of sites introduces no approximations not already included in the nearest neighbor problem with one type of site (*i.e.*, in $\partial\eta'/\partial N$). The two most useful approximations for $F(\theta)$ are[2,6]

(1) Quasi-chemical approximation

$$\ln F(\theta) = \frac{z}{2}\left[\frac{W}{kT} + \ln \frac{(\beta - 1 + 2\theta)(1-\theta)}{(\beta + 1 - 2\theta)\theta}\right] \tag{9}$$

$$\beta^2 = 1 - 4\theta(1-\theta)(1 - e^{-W/kT}) \tag{10}$$

(2) Bragg–Williams approximation

$$\ln F(\theta) = \frac{Wz\theta}{kT} \tag{11}$$

For computational purposes, z, W, the f_i, the K_i and the W_s are assumed known. A value is then assigned to $F(\theta)/c$ and θ calculated from eq. 7. This value of θ then gives $F(\theta)$ from eq. 9 or 11 and hence finally the value of c associated with this θ from the originally assigned value of $F(\theta)/c$.

Special Cases.—(1) If all the W's are zero, eq. 7 becomes

$$\theta = \sum_{i} \frac{f_i c}{K_i + c} \tag{12}$$

(2) If there is only one type of site, and we use the Bragg–Williams approximation

$$\theta = \frac{1}{(K_1'/c)e^{Wz\theta/kT} + 1} \tag{13}$$

or

$$\frac{c}{K_1'} = \frac{\theta}{1-\theta} e^{Wz\theta/kT} \tag{14}$$

where

$$K_1' = K_1 e^{zW_1/kT}$$

As has been pointed out before,[7] eq. 14 has the same formal appearance as in the usual[4] treatment of electrostatic effects in terms of the net charge (rather than in terms of nearest neighbor interactions).

(3) Suppose there are two types of sites for the binding of, say, H^+, (a) uncharged, *e.g.*, $-NH_2$ and (b) negatively charged, *e.g.*, $-COO^-$, and that $z = 4$. Then $W > 0$, $W_1 = 0$ and $W_2 = -W$. Equation 7 becomes

$$\theta = \sum_{i=1}^{2} f_i \left[\frac{f_1^4}{(K_i/c)F(\theta) + 1} + \frac{4f_1^3 f_2}{(K_i/c)\delta F(\theta) + 1} + \frac{6f_1^2 f_2^2}{(K_i/c)\delta^2 F(\theta) + 1} + \frac{4f_1 f_2^3}{(K_i/c)\delta^3 F(\theta) + 1} + \frac{f_2^4}{(K_i/c)\delta^4 F(\theta) + 1}\right] \tag{15}$$

where $\delta = e^{-W/kT}$.

(4) If $W \neq 0$ but $W_1 = W_2 = \cdots = W_n = 0$, then

$$\theta = \sum_{i=1}^{n} \frac{f_i}{(K_i/c)F(\theta) + 1} \tag{16}$$

---

(5) T. L. Hill, THIS JOURNAL, **78**, 3330 (1956).

(6) We mention again that eq. 9 is exact for $z = 2$ (linear system).

(7) T. L. Hill, *J. Phys. Chem.*, **57**, 324 (1953).

TABLE I

EFFECTIVE DIELECTRIC CONSTANTS, D_E

ξ	r/b	$x = 0$	0.5	1.0	2.0	4.0	6.0	8.0
0.50	1.000	86.5	148	254	75×10^1			
.60	0.894	82.1	130	207	50×10^1			
.70	.775	77.1	113	167	37×10^1			
.80	.632	71.0	95.1	130	244	90×10^1		
.90	.447	62.9	75.9	93.1	143	35×10^1	89×10^1	
.95	.316	56.9	64.4	73.9	99.3	18×10^1	35×10^1	
.98	.200	51.3	55.1	59.9	71.7	11×10^1	15×10^1	
.99	.141	48.3	50.7	53.6	60.8	79.3	10×10^1	14×10^1

An example, with $n = 2$ and $z = 2$, would be the back titration curve of DNA from low pH between pH 3 and 9 where presumably only the $-NH_2$ groups of adenine and cytosine are involved.

III. Estimation of W for Charges on Surface of Sphere

In this section we show how W may be estimated if, for a model, we can use a sphere with charges located on the surface. These results may also be of interest in the theory of enzyme action; from this point of view they supplement tables published in an earlier paper.[8]

If we express W in the form

$$W = \frac{e_1 e_2}{D_E r} \tag{17}$$

where e_1 and e_2 are charges, then the problem becomes one of calculating D_E, the effective dielectric constant. As before,[8] we use the model and basic theory of Kirkwood.[9]

Consider a sphere of low dielectric constant D_i ($\cong 2$) immersed in a solvent of dielectric constant D. Electrolyte (point) ions at low concentration are present in the solvent and fixed (point) charges are imbedded at the surface of the sphere (Fig. 2). Electrolyte does not penetrate the sphere. If, in eq. 8,[8] we put

$$\left(\frac{r}{b}\right)^2 = 2(1 - \xi)$$
$$\xi = \cos\theta$$
$$y = 1$$

then

$$\frac{1}{D_E} = \frac{1}{D}\left\{2 + \sqrt{2(1 - \xi)}\ln\frac{1 - \xi}{1 - \xi + \sqrt{2(1 - \xi)}} - \sqrt{2(1 - \xi)}\sum_{n=0}^{\infty}\frac{(2n + 1)xq_n(x)P_n(\xi)}{(n + 1)[n + 1 + xq_n(x)]}\right\} \tag{18}$$

where $x = \kappa b$ (κ is the Debye–Hückel parameter). the $P_n(\xi)$ are Legendre polynomials and

$$q_n(x) = 1 - \frac{K_n'(x)}{K_n(x)} \tag{19}$$

$$K_n(x) = \sum_{s=0}^{n}\frac{2^s n!\,(2n - s)!\,x^s}{s!\,(2n)!(n - s)!} \tag{20}$$

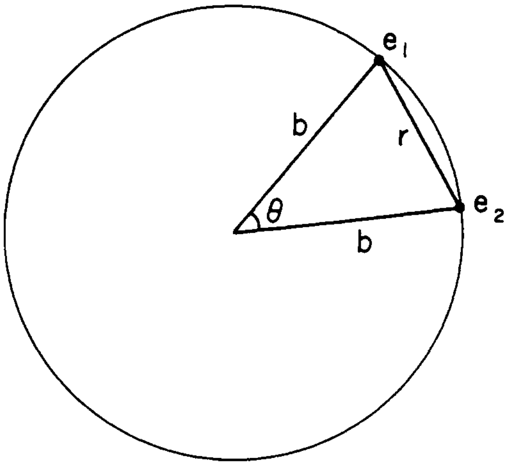

Fig. 2.—Geometry for charges on surface of sphere.

Table I gives the calculated values[10] of D_E for various choices of x and ξ, using $D_i = 2$ and $D = 78$.

BETHESDA 14, MD.

---

(8) T. L. Hill, *J. Phys. Chem.*, **60**, 253 (1956).

(9) J. G. Kirkwood, *J. Chem. Phys.*. **2**, 351 (1934).

(10) Calculated by the Computation Laboratory of the National Bureau of Standards.