TABLE III

INTENSITIES OF PROPYL IONS FOR *n*-BUTANE AND FOUR DEUTERATED BUTANES

Butane-d_0 Ion	I^a	Butane-d_1 Ion	I	Butane-d_3 Ion	I	Butane-d_5 Ion	I	Butane-d_7 Ion	I
$CH_3CH_2CH_2^+$	34	$CH_3CDHCH_2^+$ $CH_3CH_2CDH^+$	36	$CH_3CH_2CH_2^+$ $CD_3CH_2CH_2^+$	18 24	$CH_3CD_2CH_2^+$ $CD_3CH_2CD_2^+$	15 23	$CH_3CD_2CD_2^+$ $CD_3CD_2CD_2^+$	19 23

[a] Intensities or relative abundances as percentage of total ion current, corrected for exchange contribution indicated in Table II.

of the heavier and lighter ions. However, the sum of these was less than the intensity of the ethyl ion from butane-d_0.

One further point of interest should be noted regarding the effect of isotopic substitution. In Table IV the total relative intensities of the C_2, C_3 and C_4 groups for *n*-butane and four isotopic butanes are given. It is noted that the relative intensity of the C_2 group for butane-d_5 shows a considerable increase over that expected with a corresponding decrease in the C_3 and C_4 groups. This may be attributed to the close similarity of the two possible ethyl ions resulting from a simple bond break or, oppositely, to the greater dissimilarity of the two groups immediately adjacent to the bond. Future work on symmetrically substituted molecules such as $CD_3CH_2CH_2CD_3$ and $CH_3CD_2CD_2CH_3$ is indicated.

TABLE IV

TOTAL RELATIVE INTENSITIES FOR THE C_2, C_3 AND C_4 GROUPS FOR *n*-BUTANE AND FOUR DEUTERATED BUTANES

Group	Butane-d_0	Butane-d_1	Butane-d_3	Butane-d_5	Butane-d
C_2	30.6	28.9	23.8	28.8	22.9
C_3	56.3	59.0	63.4	60.9	65.8
C_4	12.2	11.9	12.0	9.5	10.3

SALT LAKE CITY, UTAH

[CONTRIBUTION FROM THE NAVAL MEDICAL RESEARCH INSTITUTE]

Charge Distribution in Protein Molecules. I

BY TERRELL L. HILL

RECEIVED OCTOBER 6, 1955

Assuming that ion (*e.g.*, H^+) binding sites are distributed throughout a spherical protein molecule and that electrolyte solution can penetrate somewhat into the protein molecule, it is possible to calculate the non-uniform radial distribution of bound ions at equilibrium. In the numerical cases worked out, there is some tendency, but not an overwhelming one, for the bound ions to concentrate near the surface. The titration curve (or binding isotherm) is computed and compared with more conventional titration curves. The net charge (bound charge plus electrolyte) radial distribution is also calculated. The total net charge turns out to be about half of the total bound charge, in the examples chosen.

I. Introduction

When charge interactions are taken into account in work on ion binding, titration curves, etc., of proteins, it is generally assumed that the charges on the protein molecule are distributed over the surface. This assumption is difficult to test experimentally, so we examine the question theoretically here and in a second paper to be published later. In the present paper we restrict ourselves to the case of one type of dissociable group (*e.g.*, $-NH_3^+ \rightarrow H^+ + -NH_2$). The second paper will consider a mixed case (*e.g.*, $-NH_3^+$ and $-COOH$).

A protein surface charge distribution is strictly possible only if all dissociable groups (or binding sites) are on the surface. A more reasonable assumption (see below) would appear to be that the binding sites (*e.g.*, $-NH_2$) are distributed throughout the protein but that the sites nearest the surface bind preferentially (to give, for example, $-NH_3^+$) because this results in a lower electrostatic free energy than a more or less uniform charge distribution throughout the molecule.

There might be some tendency for the binding sites themselves to be built in predominantly near the surface of the protein on synthesis (polymerization of amino acids), since this would decrease the free energy of synthesis. However, this suggestion seems rather unlikely since (1) the protein presumably folds up, forming a surface, *after* the polymerization is accomplished and (2) restricting, on synthesis, the order of amino acids along the chain (or chains), so that the binding sites will appear at or near the surface after folding, places a very severe limitation (statistically) on the possible arrangements of the amino acids along the chain (or chains).

As a model, we take a rigid, slightly expanded (hydrated), spherical protein molecule with a radially symmetrical distribution of binding sites. For a given total number of bound ions (*e.g.*, H^+), we find that distribution of bound ions among sites which minimizes the total (electrostatic plus binding) free energy. An alternative but equivalent statement is that the required distribution of bound ions among sites is such as to make the electrochemical potential of the bound ions the same at all points in the protein molecule. Having found the charge distribution, we then note the extent to which the charges are actually concentrated near the surface, according to this model.

In an earlier paper,[1] we attempted to treat the apparent swelling of bovine serum albumin at low *p*H. In the electrostatic calculation, a uniform degree of ionization throughout the protein was assumed. Here we remove this assumption, but do not allow (variable) swelling. Both swelling and a

(1) T. L. Hill, *J. Phys. Chem.*, **60**, 358 (1956).

non-uniform degree of ionization could be included in a single treatment but the numerical work would be much more formidable than in the present paper. Also, the mechanism of swelling is still uncertain (the α-β transition is one possibility[1]).

II. The Model and Equations

Consider a spherical protein molecule immersed in an electrolyte solution. The protein is rigid but swollen (to a fixed extent) by electrolyte[2] solution to an extent such that the volume fraction of protein is $1 - \alpha$ and the radius is b. Let D_i^0 be the dielectric constant of unswollen protein and D the dielectric constant of the solvent. Then we take for the dielectric constant D_i inside the sphere

$$D_i = \alpha D + (1 - \alpha)D_i^0 \cong \alpha D \quad (1)$$

since $D_i^0 = 2$ or 3.

We assume that there are sufficient sites so that continuous distributions can be used below as an adequate approximation. Let $\rho(r)$ be the density of sites (no./volume) at r and $\theta(r)$ the fraction of these sites occupied by bound ions. The charge on a vacant site is taken as zero and on a filled site $Z\epsilon$, where ϵ = protonic charge. All sites have the same intrinsic equilibrium dissociation (or binding) constant.

If j is the partition function of a bound ion on a site, then in the absence of electrostatic effects the chemical potential μ of bound ions is given by[3]

$$\frac{\mu}{kT} = \ln \frac{\theta}{(1 - \theta)j} \text{ (bound on protein)} = \frac{\mu^0}{kT} + \ln c \text{ (in soln.)} \quad (2)$$

where c is the equilibrium concentration (or activity) of ions in solution and μ^0 is a standard ($c = 1$) chemical potential. Let K^0 be the intrinsic dissociation constant of the bound ions. Then, from eq. 2

$$K^0 \equiv \frac{c(1 - \theta)}{\theta} = \frac{1}{je^{\mu^0/kT}} \quad (3)$$

In the general case, with $\psi(r)$ the potential at r, and μ the electrochemical potential

$$\frac{\mu}{kT} = \ln \frac{\theta(r)}{[1 - \theta(r)]j} + \frac{Z\epsilon\psi(r)}{kT} = \frac{\mu^0}{kT} + \ln c \quad (4)$$

or

$$\frac{Z\epsilon\psi(r)}{kT} = \ln \frac{1 - \theta(r)}{\theta(r)} + \ln \frac{c}{K^0} \quad (5)$$

This relation between $\psi(r)$ and $\theta(r)$ will be used below.

Inside the sphere, $\psi(r)$ must satisfy

$$\nabla^2\psi(r) = -\frac{4\pi}{\alpha D}\left[\rho(r)\theta(r)Z\epsilon + \sum_i z_i c_i \epsilon \alpha e^{-z_i\epsilon\psi(r)/kT}\right] \quad (6)$$

where Σ_i is the charge density due to the electrolyte (considered made up of point ions) which has penetrated the sphere. The i-th ionic species of electrolyte has a concentration c_i far from the sphere and a valence z_i. The correction factor α in Σ_i takes care of the fact that electrolyte is "diluted" by protein inside the sphere. On linearizing eq. 6

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi}{dr}\right) = -\frac{4\pi\rho(r)\theta(r)Z\epsilon}{\alpha D} + \kappa^2\psi(r) \quad (7)$$

where κ is the usual Debye–Hückel parameter (with dielectric constant D).

We multiply eq. 7 by $r^2\,dr$ and integrate from r to b

$$r^2\psi'(r) = b^2\psi'(b) + \frac{4\pi Z\epsilon}{\alpha D}\int_r^b \rho(y)\theta(y)y^2dy - \kappa^2\int_r^b \psi(y)y^2dy \quad (8)$$

To obtain an expression for $b^2\psi'(b)$, we put $r = 0$ in eq. 8

$$b^2\psi'(b) = -\frac{4\pi Z\epsilon}{\alpha D}\int_0^b \rho(y)\theta(y)y^2\,dy + \kappa^2\int_0^b \psi(y)y^2\,dy \quad (9)$$

Next, we multiply eq. 8 by dr/r^2 and integrate again from r to b

$$\psi(r) = \psi(b) - b^2\psi'(b)\left(\frac{1}{r} - \frac{1}{b}\right) - \frac{4\pi Z\epsilon}{\alpha D}\int_r^b \frac{du}{u^2}\int_u^b \rho(y)\theta(y)y^2dy + \kappa^2\int_r^b \frac{du}{u^2}\int_u^b \psi(y)y^2dy \quad (10)$$

We now consider the region outside the sphere, where

$$\nabla^2\psi(r) = \kappa^2\psi(r)$$

and hence

$$\psi(r) = \frac{Ce^{-\kappa r}}{r} \quad (11)$$

To find C, we set $D\psi'(b)$ from eq. 11 equal to $\alpha D\psi'(b)$ from eq. 9, and obtain

$$C = \frac{e^{\kappa b}}{1 + \kappa b}\left[\frac{4\pi Z\epsilon}{D}\int_0^b \rho(y)\theta(y)y^2\,dy - \kappa^2\alpha\int_0^b \psi(y)y^2\,dy\right] \quad (12)$$

From eq. 11 and 12

$$\psi(b) = \frac{1}{1 + \kappa b}\left[\frac{4\pi Z\epsilon}{Db}\int_0^b \rho(y)\theta(y)y^2\,dy - \frac{\kappa^2\alpha}{b}\int_0^b \psi(y)y^2\,dy\right] \quad (13)$$

This expression can be used for ψ(b) in eq. 10 since ψ is continuous at $r = b$.

Returning to eq. 10, we substitute eq. 13 for $\psi(b)$, eq. 9 for $b^2\psi'(b)$, and simplify the last two integrals by an integration by parts. The result is

$$f(s) = A\left(\frac{\alpha}{1 + \kappa b} - 1\right)\int_0^1 \rho(x)\theta(x)x^2\,dx - \kappa^2b^2\left(\frac{\alpha}{1 + \kappa b} - 1\right)\int_0^1 f(x)x^2\,dx + \frac{A}{s}\int_0^s \rho(x)\theta(x)x^2\,dx - \frac{\kappa^2b^2}{s}\int_0^s f(x)x^2\,dx + A\int_s^1 \rho(x)\theta(x)x\,dx - \kappa^2b^2\int_s^1 f(x)x\,dx \quad (14)$$

where

$$f(s) = \frac{Z\epsilon\psi(s)}{kT}$$

$$s = r/b$$

$$A = \frac{4\pi Z^2\epsilon^2b^2}{\alpha DkT}$$

(2) If bound ions are assumed able to penetrate the protein to reach interior binding sites, for consistency we should assume in general that electrolyte ions can also penetrate the protein.

(3) See, for example, T. L. Hill, "Statistical Mechanics," McGraw-Hill Book Co., New York, N. Y., 1956, Appendix 4.

Finally, if we replace $f(s)$ (and similarly $f(x)$) in eq. 14 by

$$f(s) = \ln \frac{1 - \theta(s)}{\theta(s)} + \ln \frac{c}{K^0} \qquad (15)$$

we obtain a non-linear integral equation for $\theta(s)$, which can be solved numerically. The parameters on which the solution depends are c/K^0, A, α, κb, and the assumed site density distribution $\rho(s)$.

III. Calculations

Because of the complexity of eq. 14 and the fact that the more interesting case has two types of sites (see Part II), we have computed[4] curves for only one choice of A, α, κb and $\rho(s)$, and three choices of c/K^0.

An iteration scheme was used in which, for given A, α, κb and $\rho(s)$, each successive approximation $\theta_i(s)$ was forced to pass through θ (0.5) = constant, with an initial approximation $\theta_1(s) \equiv$ same constant.

The choices of parameters were the following: T = 310.1°K., D = 74.31, α = $^1/_3$, $^1/_\kappa$ = 7.79 Å. (0.15 M NaCl), b = 31.16 Å. (so that κb = 4), Z = 1, and ρ = constant, with a value corresponding to a total of 25 binding sites in the molecule.[5]

The choices of the constant in $\theta(0.5)$ = constant, referred to above, were 0.1, 0.5 and 0.8. The corresponding values to which $\ln(c/K^0)$ converged were −1.936, 1.350 and 3.492, respectively.

The solutions $\theta(s)$ are shown in Fig. 1. In the $\theta(0.5)$ = 0.1 case, the concentration of charge near the surface is negligible, because of the low charge density. There is a much bigger effect in the other two cases, especially $\theta(0.5)$ = 0.5. Of course, when c/K^0 becomes very large, $\theta(s) \to 1$ for all s. Hence the surface concentration effect will be small for $c/K^0 \to \infty$ and $c/K^0 \to 0$, and will be largest for intermediate c/K^0.

Also of interest are the $s^2\theta(s)$ curves included in Fig. 1, since the total number of charges between s and $s + ds$ is proportional to $s^2\theta(s)ds$. These curves show the usual assumption that the charge resides only on the surface to be a fair but not a good approximation, for the model and case worked out here.

As already pointed out (eq. 6), when electrolyte is allowed to penetrate the protein, the charge density inside the protein is not just that due to the bound ions. Thus, from eq. 7, if $P(r)$ is the net charge density at r

$$P(r) = \rho(r)\theta(r)Z\epsilon - \frac{\kappa^2 \alpha D \psi(r)}{4\pi} \qquad (16)$$

When ρ = constant, a convenient reduced form is

$$\frac{P(s)}{\rho Z\epsilon} = \theta(s) - \frac{\kappa^2 b^2}{\rho A}\left[\ln \frac{1 - \theta(s)}{\theta(s)} + \ln \frac{c}{K^0}\right] \qquad (17)$$

where eq. 15 has been used to eliminate $\psi(r)$. Equation 17 gives the net charge density in the form of a "corrected" or "net" $\theta(s)$. Figure 2 compares

(4) Most of the numerical work was carried out by the National Bureau of Standards Computation Laboratory.

(5) When ρ = constant, only the combination ρA occurs, where

$$\rho A = \frac{3zZ^2\epsilon^2}{\alpha b D k T}$$

and z = total number of binding sites per protein molecule (here z = 25).

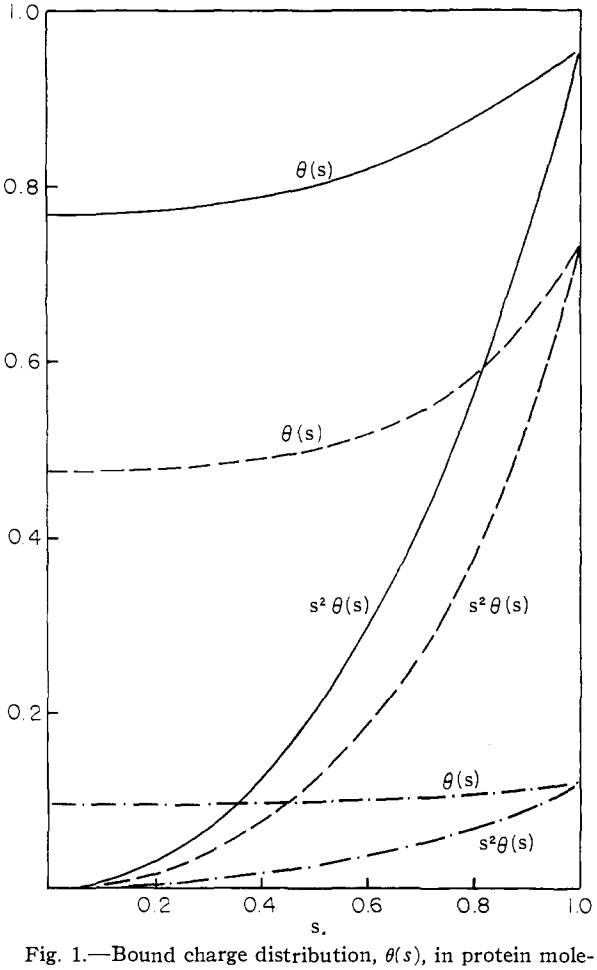

Fig. 1.—Bound charge distribution, $\theta(s)$, in protein molecule, together with $s^2(\theta)(s)$: (a) θ (0.5) = 0.8, ——; (b) θ (0.5) = 0.5, ------; (c) θ (0.5) = 0.1, -·-·-·.

$\theta(s)$ with $P(s)/\rho\epsilon$ in the numerical case of Fig. 1. The net charge density is seen to be very small near the center of the sphere (since $b >> 1/\kappa$).

In general, the total number of bound ions is

$$N = 4\pi b^3 \int_0^1 \rho(s)\theta(s)s^2\, ds \qquad (18)$$

which can be calculated from the solution $\theta(s)$ of eq. 14. In the special case ρ = constant, $N = z\bar{\theta}$, where $\bar{\theta}$ is the average value of θ

$$\bar{\theta} = 3\int_0^1 s^2\theta(s)\, ds \qquad (19)$$

Similarly (ρ = constant), the total net charge on the protein, in the same units, is $z\bar{P}/\rho Z\epsilon$, where

$$\frac{\bar{P}}{\rho Z\epsilon} = 3\int_0^1 s^2\left[\frac{P(s)}{\rho Z\epsilon}\right] ds \qquad (20)$$

These quantities have been calculated using the data of Fig. 2 and numerical integration. For $\theta(0.5)$ = 0.1, 0.5 and 0.8, we find $\bar{\theta}$ = 0.108, 0.589 and 0.872, respectively, and $\bar{P}/\rho\epsilon$ = 0.052, 0.292 and 0.419, respectively. In each case the total net charge on the protein is about half of the charge due to the bound ions alone. These considerations have obvious implications in electrophoresis theory, etc., for protein molecules with penetrating electrolyte.

The titration curve or adsorption isotherm of the

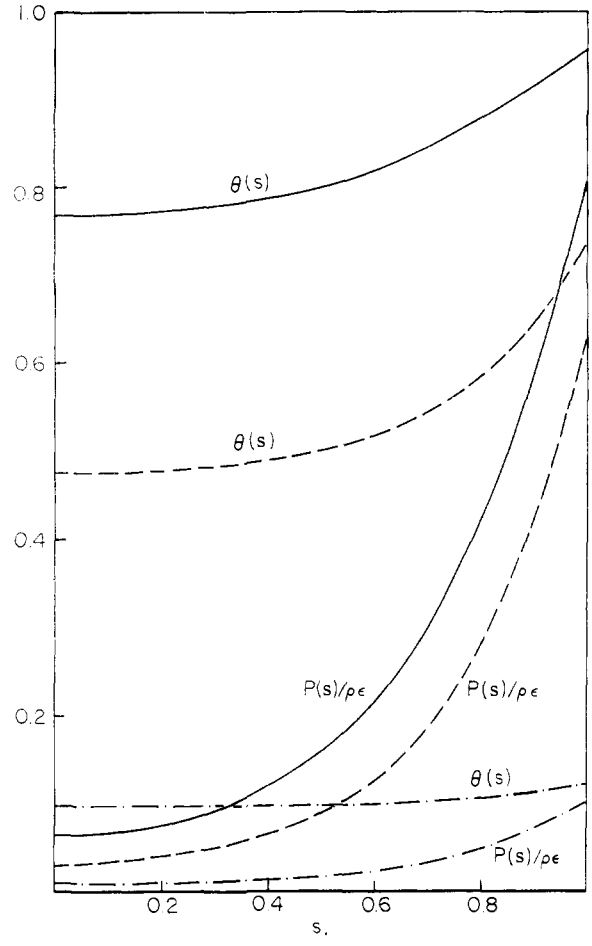

Fig. 2.—Bound charge distribution, $\theta(s)$, and net charge distribution, $P(s)/\rho\epsilon$, in protein molecule. (a) θ (0.5) = 0.8, ——; (b) θ (0.5) = 0.5, ------; (c) θ (0.5) = 0.1, -·-·-·.

bound ions is obtained by plotting N *versus* ln (c/K^0). For the numerical case under discussion here, we plot, in Fig. 3, $\bar{\theta} = N/z$ *versus* ln (c/K^0). For purposes of comparison, three other curves are also included in Fig. 3.

(1) The "unperturbed" or "Langmuir" isotherm, eq. 3, which holds in the absence of electrostatic effects.

(b) The well-known isotherm

$$\ln \frac{c}{K^0} = \ln \frac{\theta}{1-\theta} + \frac{zZ^2\epsilon^2\theta}{DbkT(1+\kappa b)} \qquad (21)$$

taking $Z = 1$, $z = 25$, $\kappa b = 4$, etc., as above. This equation applies to a protein molecule immersed in non-penetrating electrolyte (point ions) and with z binding sites distributed uniformly over the surface of the sphere.

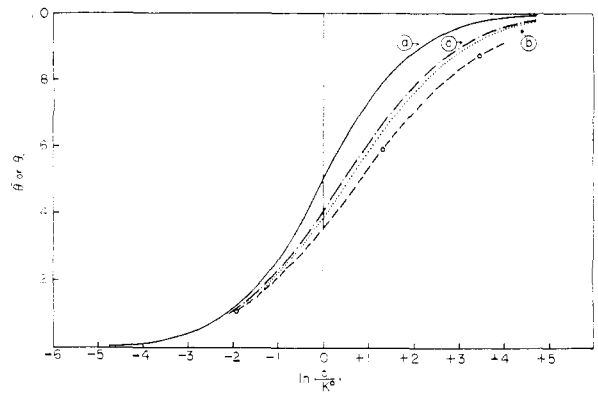

Fig. 3.—Titration curves according to various models: curve (a), no electrostatic effects; curve (b), charge distributed on surface and no electrolyte in protein molecule; curve (c), charge distributed on surface, with electrolyte penetrating protein molecule. The dashed curve (------, based on three points only, as shown) corresponds to Figs. 1 and 2; that is, charge distributed (non-uniformly) throughout protein and electrolyte penetrating the protein.

(c) Let the z sites be on the surface, but let electrolyte (point ions) penetrate as before. The potential is then

$$\psi_{in} = C_1 \frac{(e^{\kappa r} - e^{-\kappa r})}{r}$$

$$\psi_{out} = \frac{C_2 e^{-\kappa r}}{r}$$

with boundary conditions.

$$\psi_{in}(b) = \psi_{out}(b)$$

$$D\psi'_{out}(b) - \alpha D\psi'_{in}(b) = -\frac{zZ\epsilon\theta}{b^2}$$

This gives

$$\ln \frac{c}{K^0} = \ln \frac{\theta}{1-\theta} + \frac{zZ^2\epsilon^2\theta}{DbkT[1+\kappa b + \alpha(\kappa b \coth \kappa b - 1)]} \qquad (22)$$

All the curves in Fig. 3 appear to be part of a family. It would therefore presumably be difficult to distinguish between the corresponding models by fitting experimental titration curves, except possibly through the ionic strength dependence.

Attention should be called to a paper in the same general field by Tanford,[6] which appeared while the present manuscript was in preparation. Tanford does not discuss the main points considered here: non-uniform radial bound and net charge distributions.

The author is indebted to Drs. Frank Westheimer and Manuel Morales for stimulating discussions.

BETHESDA, MARYLAND

(6) C. Tanford, *J. Phys. Chem.*, **59**, 788 (1955).