hexane to yield 28 mg of white crystals, mp 135–136° whose ir was identical with that of the starting material. Some tar was also recovered from the work-up which could not be resolved chromatographically.

c. Basic Reagent. Addition of a few drops of dilute NaOH or 5% $NaHCO_3$ to a solution of I caused rapid darkening of the solution. Work-up attempts yielded only tars. Addition of a few drops of pyridine to a solution of I also caused darkening. Addition of hydrazine hydrate gave a precipitate of a tacky off-white product which gummed and darkened in a matter of minutes.

Ozonolysis of I. Approximately 250 mg of I was dissolved in 10 ml of CH_2Cl_2 and cooled to around −5°. The solution was then subjected to a stream of ozonized oxygen from a Welsbach generator for 20 min. The solvent was removed under reduced pressure and the yellow-green ozonide gum was decomposed by H_2O and Zn dust. The suspension was steam distilled into Brady's reagent and the resultant precipitate was chromatographed over alumina and eluted with benzene to get approximately 70 mg of a 2,4-DNP derivative, mp 163.5–165°. The literature melting point for the 2,4-DNP of acetaldehyde is 165–166°.[15] The ir spectrum of the ozonolysis derivative from I was identical with that of freshly prepared acetaldehyde 2,4-dinitrophenylhydrazone.

2,4-Dinitrophenylhydrazone of I. To a solution of 0.5 g of I in 30 ml of MeOH and 6 ml of concentrated HCl approximately 0.6 g of 2,4-dinitrophenylhydrazone was added. The reaction mixture was refluxed for 4 hr and the resultant solution was concentrated to 10 ml and refrigerated overnight. The precipitated solid was recovered to yield 0.9 g of red solid which was stirred with 100 ml of ether for 15–20 min. The insoluble material was discarded and the ether solution was concentrated to an oil which was chromatographed over silica gel. A major component of 200 mg was eluted with benzene–EtOAc (95:5) and recrystallized successively from MeOH–H_2O and ether to get 140 mg of reddish crystals, mp 188–189° (*Anal.* Calcd for $C_{16}H_{14}N_4O_7Cl_2$: C, 43.14; H, 3.14; N, 12.58; Cl, 15.95. Found: C, 43.53; H, 3.50; N, 12.16; Cl, 16.13); λ_{max}^{MeOH} 385 mμ (36,500), 303 (10,460), and 269 mμ (ε 13,800); ν_{max}^{KBr} 3410, 3290, 1750, 1620, 1592, 1530, 1430, 1362, 1340, 1315, 1270, 1215, 1140, 1118, 1100, 970, 924, 876, 833, and 742 cm^{-1}; nmr in $CDCl_3$: doublet at τ 8.12 (3 H, J = 6–7 cps), singlet at 6.23 (3 H), singlet at 4.88 (1 H), broad band at 4.58 (1 H), exchangeable with CD_3OD, doublet at 3.60 (3 H, J = 2–3 cps), doublet at 1.93 (1 H, J = 7–8 cps), split doublet at 1.63 (1 H, J = 7–8, 2–3 cps), doublet at 0.97 (1 H, J = 2–3 cps), and singlet at −1.77.

(15) G. J. Karabatsos, B. L. Shapiro, F. M Vane, J. S. Fleming, and J. S. Ratka, *J. Am. Chem. Soc.*, **85**, 2784 (1963).

Acknowledgments. We express our appreciation to our colleagues Dr. H. Tresner for culture isolation and identification, to Mr. A. J. Shay and Mr. M. Dann for fermentation and large-scale concentrations, to Mr. L. M. Brancone and staff for microanalyses, to Mr. W. Fulmor and staff for spectra and optical rotations, and to Dr. J. Karliner for mass spectra. We also thank Professor Murray Goodman of the Polytechnic Institute of Brooklyn for allowing us to use the Cary 60 spectropolarimeter and Mr. N. R. Nelson for his assistance in the X-ray data collection.

Interpretation of the Titration Curve of Oxyhemoglobin. Detailed Consideration of Coulomb Interactions at Low Ionic Strength[1]

William H. Orttung

Contribution from the Department of Chemistry, University of California, Riverside, California 92502. *Received July 1, 1968*

Abstract: An extension of the Tanford–Kirkwood theory of protein titration curves has been used to fit the titration curve of horse oxyhemoglobin at low ionic strengths in the pH range 4.5–9.0. Site coordinates were taken from a model built according to Perutz' 1965 description. The effect of Coulomb interactions on the titration of the individual sites was obtained from the calculations. The sensitivity of the fit to variations of intrinsic pK's, protein dielectric constant, and site depth was investigated. The possibility that some sites might be masked was also considered, and a plausible interpretation of the Bohr effect was obtained. An attempt to correlate the results with short-range interactions was only partially successful, suggesting that longer range interactions also contribute to the variation of site environment.

The titration curve of a protein is the sum of the titration curves of all of its proton-binding groups. The total number of protons bound at each pH may be calculated from the experimental titration curve, isoionic pH, and composition of the protein. In addition, the number of protons bound to a particular type of group may be obtained from a spectrophotometric titration if the group has an observable spectral change with ionization, or from reaction rates if a catalytic effect changes with ionization. The difficulty with the latter two methods is that they are not generally applicable from the practical point of view. Experimental determination of the titration curves of the individual sites seems even more remote at this time.

The theoretical approach to the problem was initiated by Linderstrom-Lang[2] who calculated the Coulomb interactions of the charges bound to the protein by assuming all charges to be spread uniformly over the surface of a spherical protein. Scatchard[3] found that this model gave a satisfactory representation of the interactions of eight binding sites at the vertices of a cube. In a recent analysis, Tanford and Nozaki[4] have

(1) This investigation was supported in part by Public Health Service Research Grant GM 11683 from the Division of General Medical Sciences, and was presented at the 156th National Meeting of the American Chemical Society, Atlantic City, N. J., Sept 1968, Abstract BIOL-248.

(2) K. Linderstrom-Lang, *Compt. Rend. Trav. Lab. Carlsberg*, **15**, No. 7 (1924).
(3) G. Scatchard, *Ann. N. Y. Acad. Sci.*, **51**, 660 (1949).
(4) C. Tanford and Y. Nozaki, *J. Biol. Chem.*, **241**, 2832 (1966).

used this model to estimate the number of masked histidines in hemoglobin at high ionic strength.

To obtain a more detailed understanding of specific environmental effects, Laskowski and Scheraga[5] considered hydrogen-bonding effects, and Tanford and Kirkwood[6,7] developed a theory that included the detailed Coulomb interactions of the sites. The Tanford–Kirkwood theory has recently been extended to allow calculation of the titration curves of the individual sites.[8,9] Calculations of this type have now become possible because site coordinates may be estimated from three-dimensional structural models of proteins such as hemoglobin.[10]

In the present paper an attempt has been made to apply the extended Tanford–Kirkwood theory to the case of horse oxyhemoglobin at low ionic strength. Since Coulomb interactions must occur if charges are present, this approach seems most logical. It is not obvious *a priori* that the Tanford–Kirkwood model is sufficiently realistic or that the calculations will be sufficiently accurate to reliably predict the effect of Coulomb interactions. If these problems turn out to be minimal, then differences between the data and such calculations might be ascribed to specific effects such as hydrogen bonding or site masking. (A masked site is considered to be uncharged, or charged by borrowing a proton internally from an adjacent initially uncharged site.) In any case, it seems probable that an understanding of the behavior of the individual sites will be necessary to understand such subtle phenomena as the Bohr effect[11] and unmasking of sites by acid denaturation.[12]

Site Coordinates

Models of the α and β chains of horse oxyhemoglobin were constructed to a scale of 1 cm/Å. The α helices were bent from brass wire and soldered as rigid structures. The peptides in the coiled regions were represented by properly shaped Lucite sheets, and the ϕ–ψ rotations between peptides[13] could be read to $\pm 5°$ from small angle wheels. The side chains were made from brass wire and full allowance was made for bond rotations. The model was assembled by first positioning the ends of the helices to the coordinates estimated by Perutz.[10] The coiled regions were modified as little as possible from the myoglobin configuration to fit into the hemoglobin structure. Adjustments were then carried out to remove as many discrepancies as possible.[14]

The coordinates of all nonhydrogen atoms of the model were then generated mathematically using an IBM 7040 computer. The atoms of each helix were generated backward from the C end, and the coil atoms were generated forward from the C end of the preceding helix. The computer model was self-consistent to about 1 Å, and also agreed with the physical model to about 1 Å. The coordinates of all 2500 atoms, in addition to other parameters such as bond rotations, etc., were stored on magnetic tape for ready access. The coordinates of the sites in the α_1 and β_1 units[10] were then obtained from the computer model. Perutz' sequence listing[10] was used, except that Asp F3α, G4α, and GH3α were assumed to be Asn[15] and therefore uncharged. In the β chain, Glx and Asx were assigned by comparison with the human sequence, and residues A13 β, B2β, and H3β were assumed to be Asp or Glu.[15] Residue E9α may either Lys or Gln.[15] All site separations were calculated, and slight adjustments were made in a few sets of coordinates to avoid separations significantly less then 4 Å.

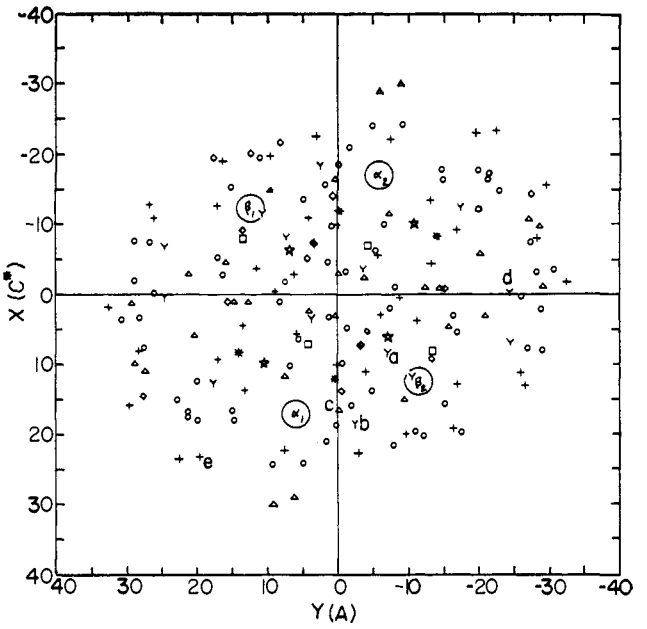

Figure 1. Projections of the site coordinates parallel to the twofold *b* axis onto the *ac** plane. The site types are labeled as follows: Y, Arg; +, Lys; ◇, Tyr; *, CysH; ☆, NTA; △, His; O, COO; □, CTC. Sites of special interest are labeled as follows: a, FG4α; b, C6β; c, FG4β; d, Gl8β; e, E9α. The labels are on the α_1 and β_2 subunits. The α–β labels are at the positions of the corresponding iron atoms. Fe_I–Fe_{IV} are correlated with β_2, β_1, α_1, and α_2, respectively.

The heme ligand and distal heme-linked histidine residues (eight per undissociated molecule) were assumed to be masked, and were left out of the titration calculation. Possible masking of tyrosine and cysteine residues did not have to be considered since these groups are all uncharged in the pH range of our calculations (4.5–9.0). Projections of the remaining site coordinates estimated from our models are shown in Figures 1 and 2.

From Perutz' discussion of his model,[10] the following sites may also be masked: Asp H9α, Arg Gl8β, Arg H24α, Arg FG4α, Arg C6β, and His FG4β. The last three are in the $\alpha_1\beta_2$ interface between the two closest heme groups, and occur in all mammalian hemoglobins studied so far.[10]

(5) M. Laskowski and H. A. Scheraga, *J. Am. Chem. Soc.*, **76**, 6305 (1954).
(6) C. Tanford and J. G. Kirkwood, *ibid.*, **79**, 5333 (1957).
(7) C. Tanford, *ibid.*, **79**, 5340 (1957).
(8) W. H. Orttung, *J. Phys. Chem.*, in press.
(9) W. H. Orttung, *ibid.*, in press.
(10) M. F. Perutz, *J. Mol. Biol.*, **13**, 646 (1965).
(11) J. Wyman, Jr., *Advan. Protein Chem.*, **4**, 458 (1948).
(12) J. Steinhardt, R. Ona, and S. Beychok, *Biochemistry*, **1**, 29 (1962).
(13) J. T. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Nemethy, G. N. Ramachandran, and H. A. Scheraga, *Biopolymers*, **4**, 121 (1966).
(14) The assistance of J. Augeri, W. Bottenberg, Y. Chen, and R. Shealy in setting up, adjusting, and analyzing the models is gratefully acknowledged.
(15) M. O. Dayhoff and R. V. Eck, Ed., "Atlas of Protein Sequence and Structure 1967–1968," The National Biomedical Research Foundation, Silver Spring, Md., 1968.

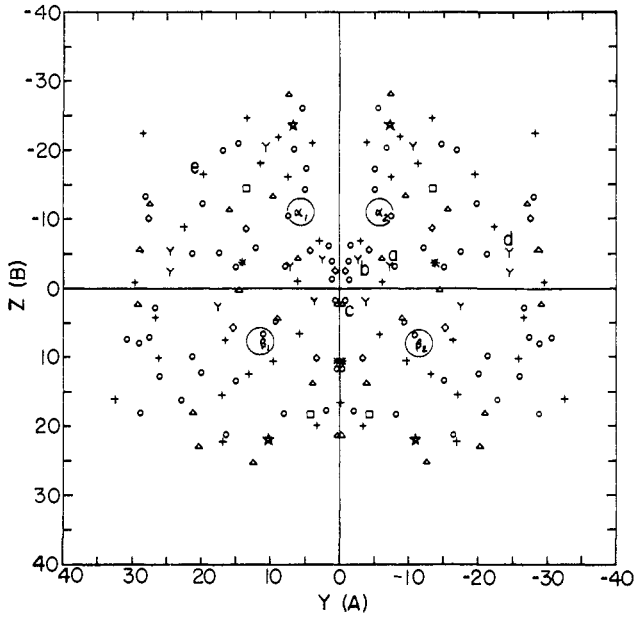

Figure 2. Projection of the site coordinates parallel to the c^* axis onto the *ab* plane. The notation is the same as in Figure 1.

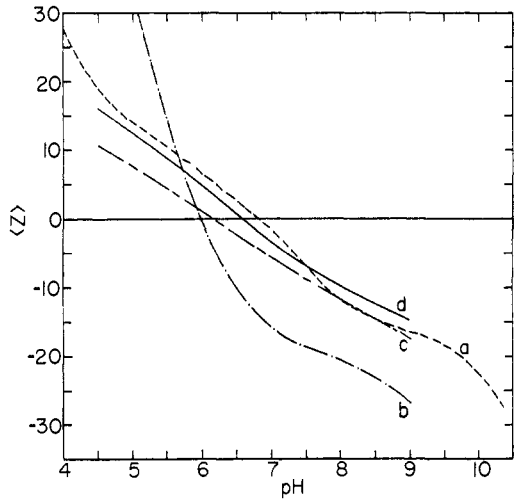

Figure 3. Fit of the titration curve of horse hemoglobin: (a) data of Cohn, *et al.*, after adjustment; (b) independent site average calculation; (c) group average calculation; (d) individual site average calculation. The parameters used in the calculations are described in the text.

Fitting the Titration Curve

The data chosen for comparison was that of Cohn, Green, and Blanchard[16] on horse carboxyhemoglobin at 25° in the absence of added salt. A molecular weight of 64,500 was used, corresponding to undissociated tetramer, and the data were shifted vertically by 3.50 charge units to correspond to zero net charge at the isoionic point (pH 6.8).[4] In the pH range 5–9, the ionic strength did not exceed 0.01. Extrapolation of data at higher ionic strength suggests that the curve shown in Figure 3 should be about one or two charge units lower at pH 5 and about one or two charge units higher at pH 9 if the ionic strength were zero. This correction is roughly indicated in Figures 4–6.

(16) E. J. Cohn, A. A. Green, and M. H. Blanchard, *J. Am. Chem. Soc.*, **59**, 509 (1937).

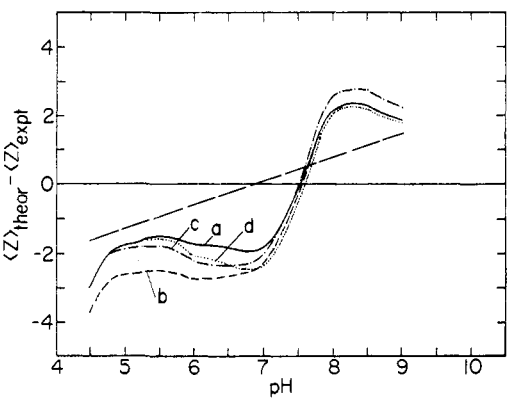

Figure 4. Differences of the individual site average calculation from the data, showing the sensitivity of the results to intrinsic pK variation. The estimated correction of the data to zero ionic strength is indicated by the dashed straight line: (a) parameters of Figure 3; (b) COO pK of 4.7; (c) HIS pK of 6.1; (d) NTA pK of 7.7.

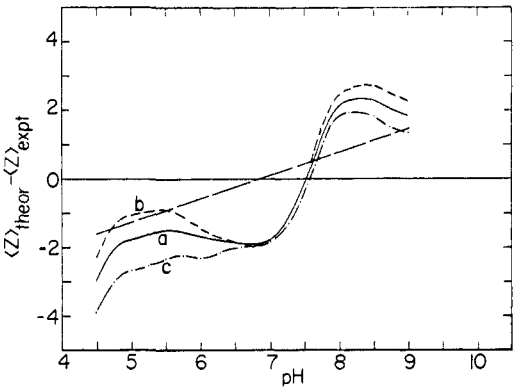

Figure 5. Differences as in Figure 4, showing the sensitivity of the results to parameter variation: (a) parameters of Figure 3; (b) site depth of 0.7 A; (c) protein dielectric constant of 6.

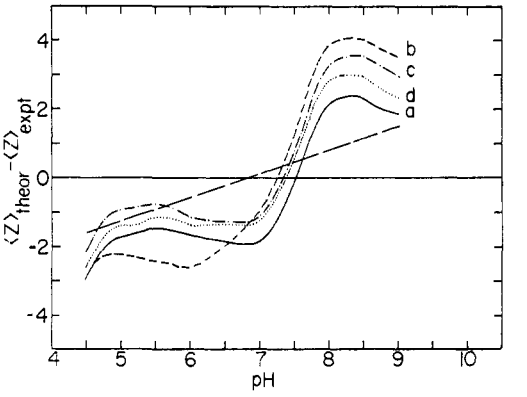

Figure 6. Differences as in Figure 4, showing the effect of unmasking the following sites: (a) none; (b) FG4α, C6β, FG4β; (c) E9α; (d) G18β.

In the calculations, the radius of the protein was taken as 28 Å and the ionic strength was set equal to zero. Pairwise interactions were calculated using the site separations obtained from the model and assuming that all sites were the same depth within the protein. Five of the eight types of site were either fully titrated or untitrated in the pH range 4.5–9.0. The intrinsic pK's used for these sites were as follows: Arg, 12.5;

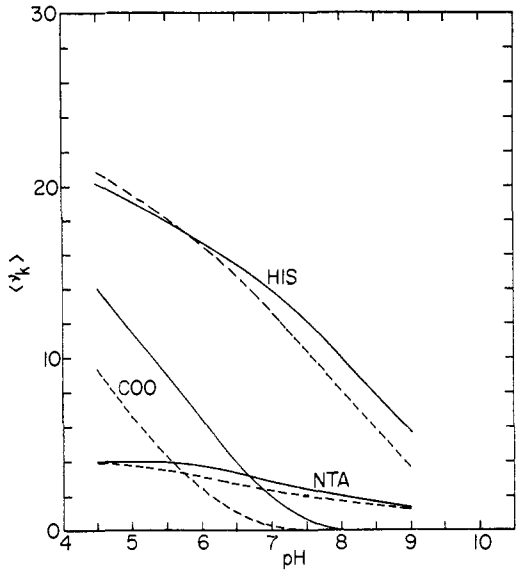

Figure 7. Theoretical number of protons bound by COO, HIS, and NTA according to the individual site average calculation of Figure 3. The dashed curves are obtained when FG4α, C6β, and FG4β are unmasked.

Lys, 10.2; Tyr, 9.6; CysH, 9.1; C-terminal carboxyl, 3.75.

Preliminary calculations were made using the group average approximation[8] without the second term of the W expansion.[8] Each calculation required 20 min on an IBM 7040 computer. Using intrinsic pK's of 7.6, 6.7, and 4.6 for N-terminal amino (NTA), imidazole (HIS), and carboxyl (COO), respectively, the results were found to be insensitive to variations of the protein dielectric constant and site depth. It was also necessary to allow all possible singly excited distributions[8] to avoid anomalous results. However, it was not possible to fit the data except by large shifts of intrinsic pK's.

Successful fitting of the data required the more detailed individual site approximation.[8] Each calculation of this type required 7 hr on an IBM 7040 computer. The ν summation[8] was limited to the range 69–110 (zero charge corresponds to ν = 88). Four and ten singly excited distributions were allowed from the NTA and HIS residues. The results were not sensitive to the artificial limitation of HIS–COO excitations. Intrinsic pK's of 7.8, 6.0, and 4.8 for NTA, HIS, and COO; a site depth of 0.6 Å; and a protein dielectric constant of 4 gave a fit which apparently could not be improved, except by unlikely variations of several parameters simultaneously. The possibility of site masking was then considered. Residue E9α was taken to be Gln (fast component) rather than Lys (slow component)[15] and Arg G18β was assumed to be masked. Removal of each site caused a uniform lowering of the curve over the calculated range of pH. Sites FG4α, C6β, and FG4β were then assumed to be masked. Removal of these sites raised the curve on the acid side and lowered it on the basic side, considerably improving the fit. The calculated result is compared with the data in Figure 3. Curves for independent site and group average calculations, using the same parameters as in the final individual site average calculation, are also shown in Figure 3 for comparison.

The sensitivity of the results to shifts of the intrinsic pK's is shown in Figure 4, and to other parameters in Figure 5. The effect of including the masked sites is shown in Figure 6.

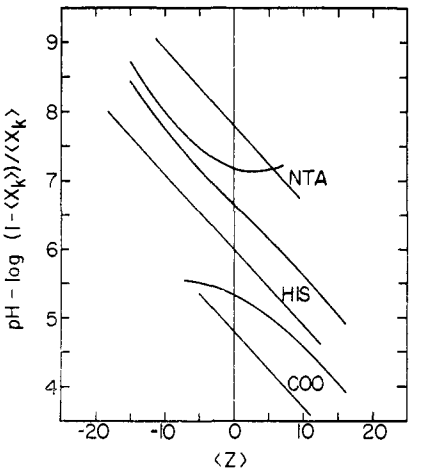

Figure 8. Theoretical logarithmic plots according to the individual site average calculation of Figure 3. The straight lines are plots of eq 1 for the parameters discussed in the text.

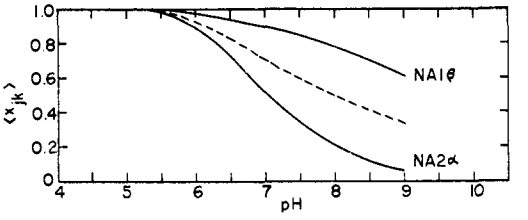

Figure 9. Theoretical titration curves of the individual NTA sites corresponding to the fit of Figure 3. The dashed curve shows the mean value.

Details of the Theoretical Fit

Only the N-terminal amino, histidine, and carboxyl groups are titrated in the pH range 4.5–9.0. The number of protons bound by each type of site is shown in Figure 7 as a function of pH. According to the smeared charge model of electrostatic interactions

$$\mathrm{pH} - \log\left[\frac{1 - \langle X_k\rangle}{\langle X_k\rangle}\right] = \mathrm{p}K_k - 0.868w\langle z\rangle \quad (1)$$

where pK_k represents the intrinsic pK of the kth type of site, $\langle z\rangle$ is the mean total charge on the protein, and

$$w = \frac{e^2}{2D_s bkT}\left(1 - \frac{\kappa b}{1 + \kappa a}\right) \quad (2)$$

where e is the electron charge, D_s is the solvent dielectric constant, b is the radius of a spherical protein, a is the radius of closest approach of counterions, and κ is the reciprocal ionic atmosphere radius of the Debye–Hückel theory. Figure 8 shows the well-known logarithmic plot suggested by eq 1. For the parameters used in these calculations, w is 0.128 at zero ionic strength. The slopes in Figure 8 are consistent with this value, although the intercepts differ. The titration curves of the individual sites are shown in Figures 9–11 for the best fit, and in Figures 12–14 assuming the three sites, FG4α, C6β, and FG4β, to be unmasked.

Comparison of Figures 10 and 13 shows that unmasking of the three sites has a small general lowering

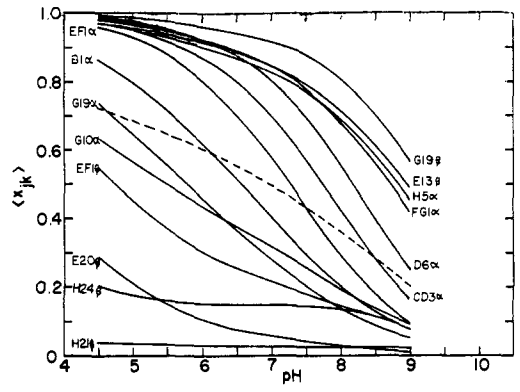

Figure 10. Theoretical titration curves of the individual HIS sites corresponding to the fit of Figure 3. The dashed curve shows the mean value.

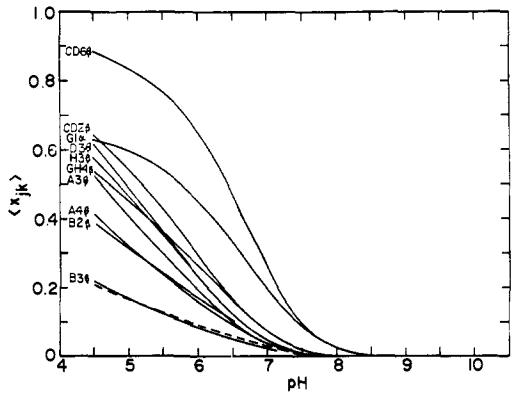

Figure 11. Theoretical titration curves of the individual COO sites corresponding to the fit of Figure 3. The dashed curve shows the mean value. The curves for the sites that fall below the mean value are not shown.

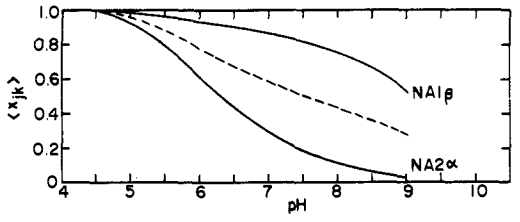

Figure 12. The analog of Figure 9 when sites FG4α, C6β, and FG4β are unmasked.

effect on the strongly binding histidine sites, particularly at alkaline pH. Figures 11 and 14 illustrate the greater reduction for the carboxyls when the three sites are unmasked.

Comparison of Binding Strengths and Local Environments

To determine how closely variations of binding strength correlate with short-range Coulomb interactions, site separations of less than 10 Å have been collected in Table I. Each integer in the table represents the distance from the indicated site to a neighboring site of a given type. Decimals have been truncated, so that a separation such as 7 indicates a distance of 7–8 Å. The sites are listed in order of decreasing binding strength. The last column of the table is an estimate of the relative Coulomb potentials

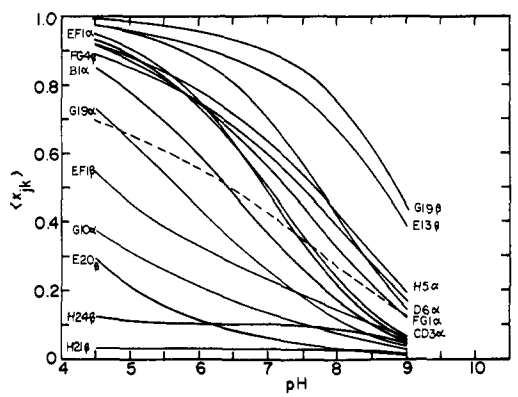

Figure 13. The analog of Figure 10 when sites FG4α, C6β, and FG4β are unmasked.

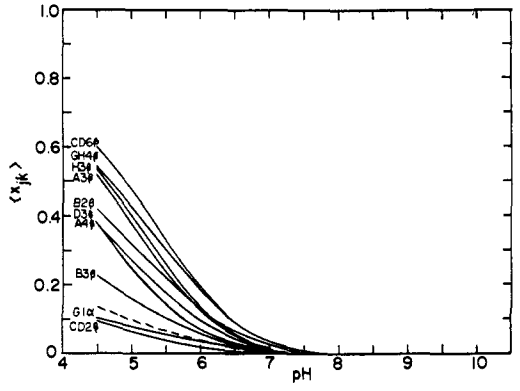

Figure 14. The analog of Figure 11 when sites FG4α, C6β, and FG4β are unmasked.

at the sites due to neighbors within 10 Å. The site charges were taken as $z_i = 1, 0, -1$ for Arg–Lys, NTA–H*IS*, and COO–CTC in this calculation.

If all sites were uniformly spread over the surface of the molecule, each site would have about five neighbors within 10 Å. Examination of the table shows that most sites have approximately five neighbors, although some carboxyls have more, and some histidines have less. There is a significant correlation of binding strength with the Coulomb potential of near neighbors for NTA and strongly binding HIS and COO sites, although the correlation is not very good for the weakly binding HIS sites. Since the weakly binding HIS sites have more HIS neighbors than the other sites, the use of $z_i = 0$ rather than $z_i = 1$ for HIS may lead to insufficiently positive results in the last column of the table. However, most of the HIS neighbors of these sites are also weakly binding.

It must be concluded, then, that longer range interactions make a significant contribution to the variation of binding strength among sites of a given type. It should also be pointed out that variation of site depth of the individual sites may occur, but is not considered in our calculation. The latter effect could also cause considerable variation.

Discussion

The calculations described in this paper demonstrate that a moderately good fit of the titration curve of oxyhemoglobin at low ionic strength is obtained by considering only Coulomb interactions and using

Table I. Local Environments of NTA, HIS, and Strongly Binding COO Sites

Site		r_i[a] Arg + Lys	r_i[a] NTA + HIS	r_i[a] COO + CTC	$\sum_i z_i/r_i \times 10$
NTA	NA1β	6	9	89	−2
	NA2α	6677	9		+5
HIS	G19β			589	−4
	FG1α	8		689	−3
	E13β			577	−5
	H5α	7		588	−3
	D6α		9	5	−2
	CD3α		9	9	−1
	EF1α	9		58	−2
	B1α		6	7	−1
	G10α	9			+1
	G19α	7	6		+1
	EF1β	9	99	68	−2
	H24β	88	89	66	−1
	E20β	4	9	9	+1
	H21β	58	58	689	−1
COO	CD6β	99	9	55	−2
	G1α	56		47899	−4
	CD2β	66		59	0
	D3β			59	−3
	GH4β		8	59	−3
	H3β	99		56	−2
	A3β			56	−4
	B2β	9	5	58	−2
	A4β	6		55	−1
	B3β	7	7	6899	−4

[a] The distances are shown as truncated integers (Å) for all sites of each type within 10 Å of the specified site. Sites are listed in order of decreasing binding strength.

normal values for pK's and protein dielectric constant. Assumed masking of certain sites greatly improves the fit. The site depth required in the fit is somewhat less than the 1 Å value found to be most suitable for small molecules.[17] However, the fact that a protein surface is not sharply defined may lead to this difference.

It has been assumed that all sites of a given type have the same intrinsic pK. In reality variations may occur because of variations of site depth or because of hydrogen-bonding interactions. In the absence of more precise structural data, it is difficult to consider these possibilities in any detail. It is also difficult to decide whether or not additional sites are masked or involved in salt bridges without better structural information.

The changes in the titration curve upon unmasking Arg FG4α, Arg C6β, and His FG4β are of considerable interest. Perutz has already focussed attention on these sites,[10] since they occur in all mammalian hemoglobins studied so far, but not in myoglobins. The sites appear to be masked in the oxy form, but are disrupted by the relative motions in the $\alpha_1\beta_2$ interface upon deoxygenation.[18] The difference between the

(17) C. Tanford, *J. Am. Chem. Soc.*, **79**, 5348 (1957).
(18) H. Muirhead, J. M. Cox, L. Mazzarella, and M. F. Perutz, *J. Mol. Biol.*, **28**, 117 (1967).

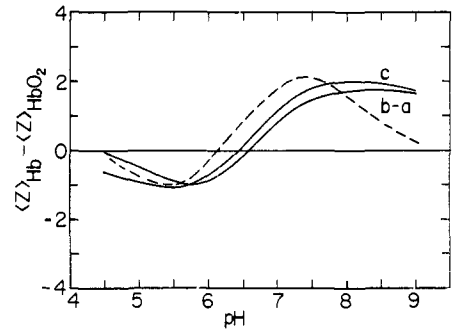

Figure 15. Comparison of the difference between curves a and b of Figure 6 (solid line) with the difference between the titration curves of oxy- and deoxyhemoglobin (dashed line). Curve c shows the difference between curve a and the modification of curve b obtained by rotating the subunits to the deoxy positions.[18]

titration curves calculated with these three sites masked and unmasked (Figures 6 and 15) is almost identical with the difference between the titration curves of oxy- and deoxyhemoglobin.[19,20] The latter difference is responsible for the well-known Bohr effect,[11,21] in which acidification of the blood by carbon dioxide (at the muscle tissue) causes additional release of oxygen by the hemoglobin. The present calculations thus provide a reasonable molecular interpretation of the Bohr effect.[22] The alkaline Bohr effect is a direct consequence of adding protons to the unmasked sites, while the acid Bohr effect is the reduction of proton binding by a larger number of carboxyls because of the addition of nearby positive charges. This effect is illustrated in Figures 8, 11, and 14. The residual discrepancy between the data and the fit in Figure 3 may be due to additional masking of sites of the same charge type which are near carboxyls.

The present calculations suggest that electrostatic repulsions prevent about eight histidines from titrating (Figure 7). If we add to this number the ten histidines and four to six arginines assumed to be masked, we obtain a total of 22–24 groups untitrated as low as pH 4.5, in agreement with Steinhardt, *et al.*[12] However, their data refer to ionic strengths of 0.25, while the present calculations were carried out at zero ionic strength. We could not ascribe the weak binding of the histidines to short-range effects in a very convincing manner in Table I, so that it will not be clear whether the present results provide an adequate interpretation of unmasking by acid denaturation until the shielding effect of ionic strength on long-range interactions is included in the calculations.

(19) B. German and J. Wyman, Jr., *J. Biol. Chem.*, **117**, 533 (1937).
(20) E. Antonini, J. Wyman, M. Brunori, C. Fronticelli, E. Bucci, and A. Rossi-Fanelli, *ibid.*, **240**, 1096 (1965).
(21) C. Bohr, K. Hasselbalch, and A. Krogh, *Scand. Arch. Physiol.*, **16**, 402 (1904).
(22) W. H. Orttung, *Nature*, in press.