Vol. 81, No. 2, 1978 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
March 30, 1978 Pages 416-421

Discrete Charge Calculations of Potentiometric Titrations for Globular Proteins: Sperm Whale Myoglobin, Hemoglobin Alpha Chain, Cytochrome c

James B. Matthew, Stephen H. Friend, Lynne H. Botelho,†
Lee D. Lehman, George I. H. Hanania, and Frank R. N. Gurd

Department of Chemistry, Indiana University,
Bloomington, Indiana 47401

Received November 16,1977

SUMMARY: The modified Tanford-Kirkwood theory of Shire et al. for intramolecular electrostatic interactions has been applied to hydrogen ion equilibria of sperm whale ferrimyoglobin, human hemoglobin α-chain and horse cytochrome c. The model employs two sets of parameters derived from the crystalline protein structures, first, the atomic coordinates of charged amino acid residues and, second, static accessibility factors to reflect their solvent exposure. In addition, a consistent set of intrinsic pK values (pK_{int}) for the individual groups is employed. The theoretical pK values at half-titration for individual groups in each protein correspond to the available observed pK values, and the theoretical titration curves compare closely with experimental potentiometric curves.

INTRODUCTION: Shire et al. (1) introduced the fractional static solvent accessibility (SA) into the discrete charge calculations of electrostatic interactions formulated by Tanford and Kirkwood (2, 3). The treatment (1, 4, 5) replaced the burial of all charged residues 0.5 to 1.0 A inside the low dielectric sphere model of the protein (2, 3) with a factor reflecting the SA parameters for the individual groups (6, 7). The modified theory has previously been shown to account for the hydrogen ion titration curve of sperm whale myoglobin (Mb) as well as the titrations of individual ionizing groups (1, 4, 5). In the present work the Mb theoretical curve is recalculated using the refined X-ray coordinates of Takano (8) and the corresponding new SA values (7). The method is also extended to hemoglobin α-chains (Hb_α) and cytochrome c (cyt c) using appropriate crystallographic data (9, 10). With the minimum

† Roche Institute of Molecular Biology, Nutley, NJ 07110.

0006-291X/78/0812-0416$01,00/0

of exceptions, it was possible to fit all titration data with a single, consistent set of pK_{int} values. These three globular proteins that differ in composition and pI value demonstrate the versatility of the method and the wealth of extractable information.

METHOD OF COMPUTATION (1, 4, 5): The Debye-Hückel theory is used to determine the electrostatic free energy for a set of discrete point charges on a sphere of radius , b, and ion-exclusion radius, a, in a solvent (usually water) with an external dielectric constant, D. The charges are placed on the surface of the equivalent sphere which is assumed to form a continuous medium with a (low) internal dielectric constant, D_i. The first stage of the calculation for a given set of parameters (b, a, D, D_i, temperature T) is to compute a table of W_{ij} values, where W_{ij} is the free energy of interaction of a pair of point charges Z_i and Z_j placed on the sphere of radius, b. An intrinsic pK (pK_{int}) is assigned to every site j and the charge Z_j of every site is calculated at any particular pH. For a given set of sites the tabulated W_{ij} values are used to compute by iterative methods a set of effective pK_i values,

$$pK_i = (pK_{int})_i - 1/(2.303\ kT) \sum_{j \neq i} (W_{ij} - SA_j W_{ij})/Zi \qquad (1)$$

Every site j for which $Z_j \neq 0$ contributes to the effective equilibrium constant of site i under the particular conditions, pK_i. Since the charged groups are buried to different degrees, the SA value of each group, j, is included to reflect the degree of exposure of each point charge to the solvent with high dielectric constant. For purposes of computation the SA values were taken equal to the nearest 0.05 increment (0.05, 0.10, 0.15, etc.) in all cases. Values of SA_j vary between 0.05 and 0.95 (7).

Application to Proteins. Structural parameters were assigned to each protein based on X-ray dimensions, so that the respective values of b were for Mb, 18.0Å (8), for Hb_α, 17.0Å (9) and for cyt c, 17.0Å (10). The average ionic exclusion radii were taken as b plus 2Å. The dielectric constant of the interior, D_i, was taken as 4 and the external dielectric constant D as that of water at 298°K, D = 78.5 (1, 11).

Site Assignments. Inspection of the crystal structures leads to three categories of proton binding sites (1). The first class comprises all groups with normal intrinsic pK values: terminal carboxyl, 3.60 (1); Asp, 4.00 (1); Glu, 4.50 (1); propionic acid, 4.00 (1); His ND1, 6.00, His NE2, 6.60 (12-14); terminal amino, 8.00 (1); Cys, 9.10 (15); Tyr, 10.00 (16); Lys, 10.40 (17); and Arg, 12.00 (1). The choice of the His value depends on the preponderant exposure of the given imidazole N according to the SA listing (12). The second class consists of sites also available to hydrogen ion equilibrium with abnormal pK values, e.g. those groups determined to be influenced by such effects as hydrogen bonding. These intrinsic pK values were estimated by adding 0.50 (Lys, Tyr) or subtracting 0.50 (Glu, Asp). The charge location was assigned to the appropriate atom, e.g. either O in a carboxyl group, according to the hydrogen bonding pattern or else the greater degree of SA for that atom. The third class represents the masked sites (1) defined according to specific chemical or structural evidence, e.g. masked histidine residues (12), or cysteine residues covalently attached to heme (18).

RESULTS: Figure 1 compares individual computed and observed pK_i values at half-titration for certain groups that have been recognized by nuclear magnetic

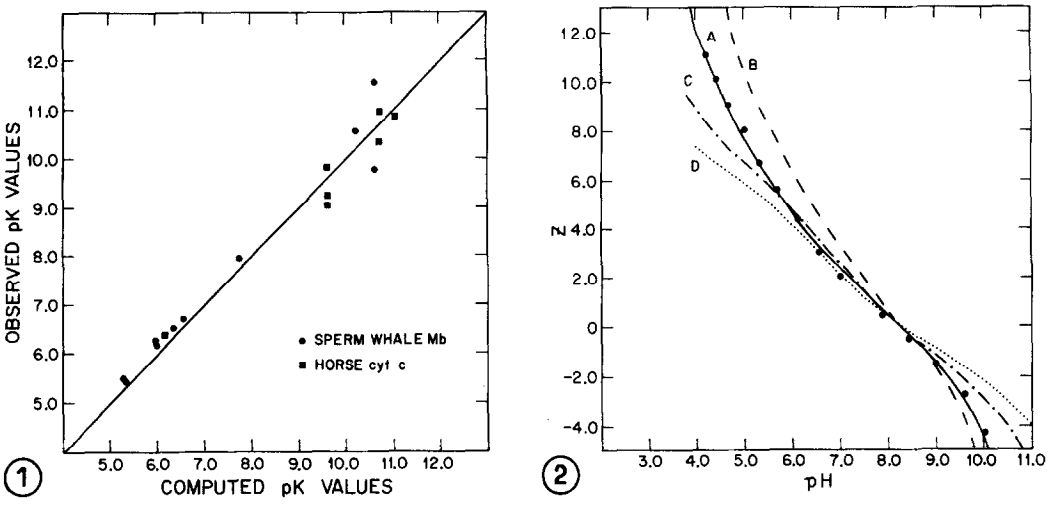

Figure 1. Comparison of observed and computed pK values.

Figure 2. Computed curves for Mb, 0.2 mM, I = 0.01M, 25°. Curve A, according to the present treatment; Curve B, with SA taken always as 1; Curve C, with SA taken always as 0; Curve D, with all groups placed 1Å inside the low dielectric medium.

resonance or optical spectroscopy (12, 14, 18). The majority of the points refer to histidine residues in Mb which were studied at 17° and ionic strength 0.1 M (12).[1] The theoretical values for 25° and ionic strength 0.01 M may be expected to range slightly higher by approximately 0.15 pK unit (1, 4, 5). Points are also included for tyrosine residues 151, 103 and 146 that were found (19) to have pK values of 10.6, 9.8 as well as 11.7[2], and 12.5, respectively. The computed values were 10.2, 10.6 and 11.9, respectively. Points in Figure 1 are also included for cyt c residues (20).

Figure 2 shows experimental points for the titration of Mb at 25° and ionic strength 0.01 M, and four theoretical titration curves for the same conditions. The continuous curve (A) is predicted from the present model and falls in close agreement with the observed points. The steepest curve (B) uses the

[1] The pK_{int} for histidine residue 36 was arbitrarily taken as 7.8 (12).

[2] The ^{13}C resonance of Tyr 103 could not be fitted with the use of a single pK value, but was consistent with two pK values (19). Our predicted value is intermediate.

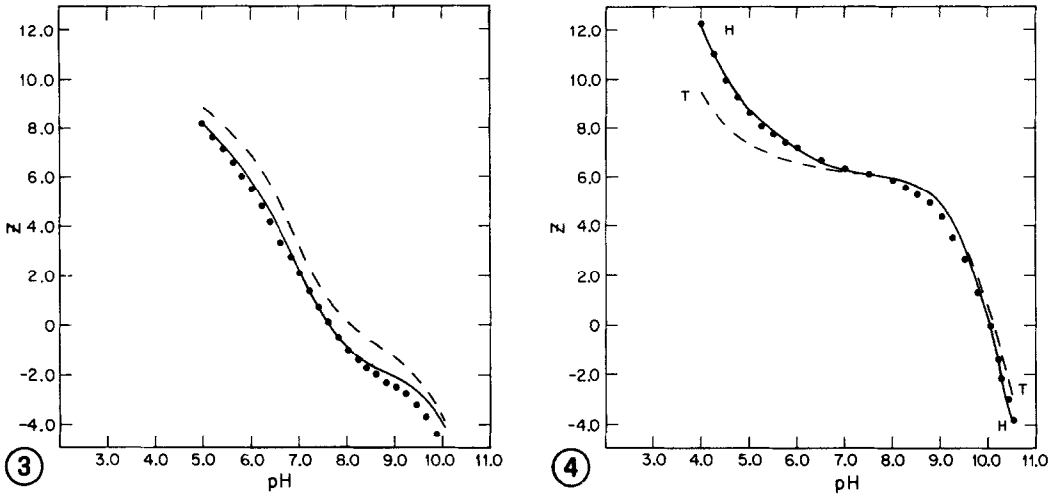

Figure 3. Application to human hemoglobin α-chain, 0.2 mM, I = 0.2M, 25°. Solid curve, computed for p-mercuribenzoate derivative of cysteine 104; broken curve, computed with mercurial omitted.

Figure 4. Computed curve, broken, for tuna ferricytochrome c; solid curve, computed for horse cyt c based on sequence differences, with observed points for 0.4 mM protein, I = 0.1M, 25°.

assumptions of the present model except that SA is taken as 1 in all cases, which has the effect of obliterating the electrostatic contributions. The curve (C) that falls first on the other side of the experimental points is computed with the assumption that SA is zero, in effect treating all charged groups as inaccessible to solvent. The least shapely curve (D) involves the same assumption of zero SA with the depth parameter d of Tanford and Kirkwood set equal to 1.0Å, in keeping with the usual practice according to the original treatment (2, 3).

Figure 3 shows the experimental titration points at ionic strength 0.2 M for isolated Hb_{α} bearing one p-mercuribenzoate molecule per molecule of the protein (21). The continuous curve includes a contribution from all titratable groups excluding the proximal and distal histidines and with all other cases assigned to the first or second classes described above (9). The coordinates of the negatively charged p-mercuribenzoate, pK_{int} 4.00, were estimated by

inspection of a Hb_{α} model (9). The broken curve in Figure 3 represents the computed form when the mercurial is omitted from consideration and the free cysteine 104 is now included. The computations are made for ionic strength 0.1 M according to an extended Debye-Hückel form (4).

Figure 4 shows the experimental titration points at ionic strength 0.1 M for horse heart cyt c titrated from the isoelectric region down to near pH 3 (20). The theoretical curves exclude from consideration the two heme linked cysteines, the iron-coordinated histidine, and the acetylated amino terminus (18). The broken curve represents that computed for tuna cyt c (10). The continuous curve represents that computed for horse cyt c with substitution of the appropriate amino acid sidechains into the crystallographic data for the tuna molecule which has been determined with higher resolution (10, 18).

CONCLUSIONS: The treatment presented here follows a consistent procedure for identifying charge loci and SA values (7), uses a standard set of pK_{int} values, and reduces *ad hoc* assumptions to a minimum. Charges have been located on given atomic coordinates as implied by the crystallographic structures rather than being arbitrarily placed. For example, in a carboxyl group the charge is placed on the more accessible O atom unless it is hydrogen bonded to a positively charged amino group. Solvent exposure is also used to guide the assignment of imidazole N atom in a given histidine residue, with the concomittant choice of pK_{int} (12). For sperm whale myoglobin the identification of residue 122 as Asp (8, 22) rather than Asn (1, 23, 24) has been confirmed for the preparation used for the present measurements of titration curve and proton NMR analyses (12).[3]

The treatment assumes, over the pH range of interest, that the crystal structure of a particular protein adequately represents the solution structure. The model predicts effective pK values (Figure 1) and ionic strength dependence (1, 4, 5) for each titratable group. Deviations of theoretical curves from experimental data may be due to pH-dependent conformational changes. Incorpor-

[3] Unpublished work of R. A. Bauman, R. D. DiMarchi, S. H. Friend, B. N. Jones, L. D. Lehman, and C. C. Wang.

ation of chemical evidence for solution structure variances, e.g. anion binding sites or difference in solution versus crystal side chain reactivities (12, 25), will allow refinement of the theoretical model for a particular protein. However, the good agreement with experiment observed for Mb, Hb_{α} and cyt c shows the versatility of the method. The method is now being applied to myoglobin from other species, and to the oxy and deoxy quaternary structures of hemoglobin.

ACKNOWLEDGMENTS: We wish to thank Dr. S. J. Shire for advice and Prof. C. R. Hartzell for making available the detailed cytochrome c titration data. The structural coordinates were made available through the cooperation of Ms. F. C. Bernstein of the Brookhaven Laboratory Protein Data Bank. This work was supported by United States Public Health Service Research Grant HL-05556 and Biochemistry Training Grant T01-GM-1046 (L.H.B., L.D.L. and J.B.M.). It is the 91st paper in a series dealing with coordination complexes and catalytic properties of proteins and related substances.

References

1. Shire, S. J., Hanania, G. I. H. and Gurd, F. R. N. (1974) Biochemistry 13, 2967-2974.
2. Tanford, C. and Kirkwood, J. G. (1957) J. Amer. Chem. Soc. 79, 5333-5339.
3. Tanford, C. and Roxby, R. (1972) Biochemistry 11, 2192-2198.
4. Shire, S. J., Hanania, G. I. H. and Gurd, F. R. N. (1974) Biochemistry 13, 2974-2979.
5. Shire, S. J., Hanania, G. I. H. and Gurd, F. R. N. (1975) Biochemistry 14, 1352-1357.
6. Lee, B. and Richards, F. M. (1971) J. Mol. Biol. 55, 379-400.
7. Matthew, J. B., Hanania, G. I. H. and Gurd, F. R. N. (1978) Biochem. Biophys. Res. Commun. (preceding paper).
8. Takano, T. (1977) J. Mol. Biol. 110, 537-568.
9. Fermi, G. (1975) J. Mol. Biol. 97, 237-256.
10. Mandel, N., Mandel, G., Trus, B. L., Rosenburg, J., Carlson, G. and Dickerson, R. E. (1977) J. Biol. Chem. 252, 4619-4636.
11. Ortung, W. H. (1970) Biochemistry 9, 2394-2402.
12. Botelho, L. H. (1975) Ph.D. Thesis, Indiana University.
13. Reynolds, W. F., Peat, I. R., Freedman, M. H. and Lyerla, J. R. (1973) J. Amer. Chem. Soc. 95, 328-331.
14. Wilbur, D. J. and Allerhand, A. (1977) J. Biol. Chem. 252, 4962-4975.
15. Tanford, C. (1962) Adv. Prot. Chem. 17, 69-165.
16. Gurd, F. R. N., Keim, P., Glushko, V. G., Lawson, P. J., Marshall, R. C., Nigen, A. M., and Vigna, R. A. (1972) in Chemistry and Biology of Peptides (Meienhofer, J. ed) p. 45-49, Ann Arbor Science Publishers, Inc., Ann Arbor, Mich.
17. Keim, P., Vigna, R. A., Nigen, A., Morrow, J. S. and Gurd, F. R. N. (1974) J. Biol. Chem. 249, 4149-4156.
18. Margoliash, E. and Schejter, A. (1966) Adv. Prot. Chem. 21, 113-286.
19. Wilbur, D. J. and Allerhand, A. (1976) J. Biol. Chem. 251, 5187-5194.
20. Shaw, R. W. and Hartzell, C. R. (1976) Biochemistry 15, 1909-1914.
21. Bucci, E., Fronticelli, C. and Ragatz, B. (1968) J. Biol. Chem. 243, 241-249.
22. Romero Herrera, A. E., and Lehmann, H. (1974) Biochim. Biophys. Acta 336, 318-323.
23. Edmundson, A. B. (1965) Nature 205, 883-887.
24. Watson, H. C. (1969) Progr. Stereochem. 4, 299-333.