The pH-Induced Release of Iron from Transferrin Investigated with a Continuum Electrostatic Model

David A. Lee and Julia M. Goodfellow
Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, United Kingdom

ABSTRACT A reduction in pH induces the release of iron from transferrin in a process that involves a conformational change in the protein from a closed to an open form. Experimental evidence suggests that there must be changes in the protonation states of certain, as yet not clearly identified, residues in the protein accompanying this conformational change. Such changes in protonation states of residues and the consequent changes in electrostatic interactions are assumed to play a large part in the mechanism of release of iron from transferrin. Using the x-ray crystal structures of human ferri- and apo-lactoferrin, we calculated the pK_a values of the titratable residues in both the closed (iron-loaded) and open (iron-free) conformations with a continuum electrostatic model. With the knowledge of a residue's pK_a value, its most probable protonation state at any specified pH may be determined. The preliminary results presented here are in good agreement with the experimental observation that the binding of ferric iron and the synergistic anion bicarbonate/carbonate results in the release of approximately three H^+ ions. It is suggested that the release of these three H^+ ions may be accounted for, in most part, by the deprotonation of the bicarbonate and residues Tyr-92, Lys-243, Lys-282, and Lys-285 together with the protonation of residues Asp-217 and Lys-277.

INTRODUCTION

Computer simulation methods, when carefully allied with experimental methods, provide tools for investigating electrostatic interactions in proteins. When protein structural stability and biological function are strongly correlated with pH it is necessary to define the protonation states of titratable residues as these may have a significant effect upon the electrostatic interactions. As the identification of such a state for a given residue is not always amenable to biochemical or spectroscopic investigation and hydrogen atoms cannot be located in a typical protein crystallographic study, there is a role for computational calculations designed to yield this information. As environmental factors affect the pK_a of a residue, a number of factors have to be taken into account in their calculation, including the degree of solvent exposure of a residue, the electrostatic potential of its environment in the protein, and the protonation states (and thus net charges) of all other titrating residues in the protein. Large shifts in the pK_a values of residues have been reported experimentally in a number of proteins and seem quite common in enzyme active sites (e.g., Bartik et al., 1994).

The most common approach to the calculation of pK_a values in protein molecules is through the use of the computationally efficient continuum electrostatic model in which solvent is represented as a region with a high dielectric constant and the protein is represented as a cavity within this region with a low dielectric constant. The electrostatic potential around the protein may then be calculated using a numerical solution of the linearized Poisson-Boltzmann equation. This approach has been used to study a number of proteins, including bovine erythrocyte Cu-Zn superoxide dismutase (Klapper et al., 1986), hen egg-white lysozyme (Bashford and Karplus, 1990), λ repressor (Zacharias et al., 1992), and yeast iso-1-ferricytochrome c (Zhou and Vijayakumar, 1997).

We have chosen the continuum electrostatic approach to study the transferrins, a family of glycoproteins that includes serum transferrin, ovotransferrin, and lactoferrin (Baker, 1993). They consist of a single polypeptide chain of 680–700 amino acid residues. These show a high degree of similarity, approximately 70% identity between lactoferrins and 50–60% between lactoferrins and transferrins. Each of the proteins also shows a striking twofold internal-sequence homology, with typically ~40% sequence identity between the N-terminal and C-terminal halves of each molecule. Given that stable half-molecule fragments can be prepared and that each possesses an iron site with very similar spectroscopic characteristics, the clear implication is that the two-sited transferrins have developed by gene duplication from a one-iron, 40-kDa precursor molecule. Each Fe(III) ion is bound together with a bicarbonate (HCO_3^-) or, more likely, carbonate (CO_3^{2-}) ion; neither is bound strongly, if at all, in the absence of the other, but when taken together the binding is extremely strong, with effective stability constants of the order 10^{20} (Aasa et al., 1963). Nevertheless, binding is reversible, both in vitro and in vivo, so mechanisms must exist for the release of this tightly bound iron.

X-ray crystal structures of transferrins, including that of human lactoferrin (Anderson et al., 1987), show that the polypeptide chain is folded into two globular lobes, representing the N-terminal and C-terminal halves of the mole-

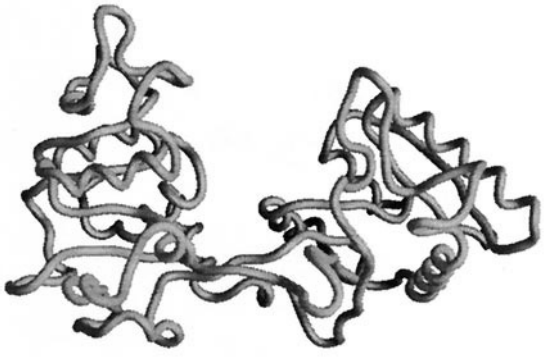
Received for publication 17 October 1997 and in final form 12 March 1998.

Address reprint requests to Dr. Julia Goodfellow, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, UK. Tel.: 44-171-631-6833; Fax: 44-171-631-6833; E-mail: j.goodfellow@mail.cryst.bbk.ac.uk.

0006-3495/98/06/2747/13 $2.00

cule (residues 1–333 and 345–691, respectively, in human lactoferrin). These N- and C-lobes are joined by a three-turn α-helix (residues 334–344 in human lactoferrin). Each lobe is then further subdivided into two domains (NI and NII and CI and CII) and contains a single iron-binding site, located in a deep cleft between the domains. Apart from the connecting helix, interactions between the two lobes involve primarily the domains NI and CI, which interact via a number of hydrophobic side chains. These give a hydrophobic cushion between the lobes, which may both explain the strong tendency of separated half-molecules to associate and also allow some relative movement of the two lobes.

The metal ion in both lobes is coordinated by four protein side chains. In the N-lobe of human lactoferrin (Figs. 1 and 2 *A*), this is via the phenolate oxygens of Tyr-92 and Tyr-192, the imidazole nitrogen of His-253, and the carboxylate oxygen of Asp-60 (Fig. 2 *B*). The metal coordination is completed by the carbonate ion, which is bound in a bidentate fashion to give a six-coordinate, distorted octahedral geometry around the metal. Although the iron-binding site in each lobe is buried deep within the protein, some



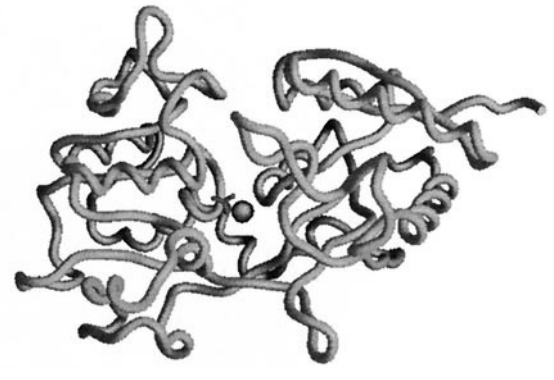


FIGURE 1 Main chain worm representations produced using GRASP (Nicholls et al., 1993) for the open iron-free and closed iron-loaded forms of the N-terminal half-molecule of human lactoferrin. The positions of the iron and carbonate in the closed form are indicated.

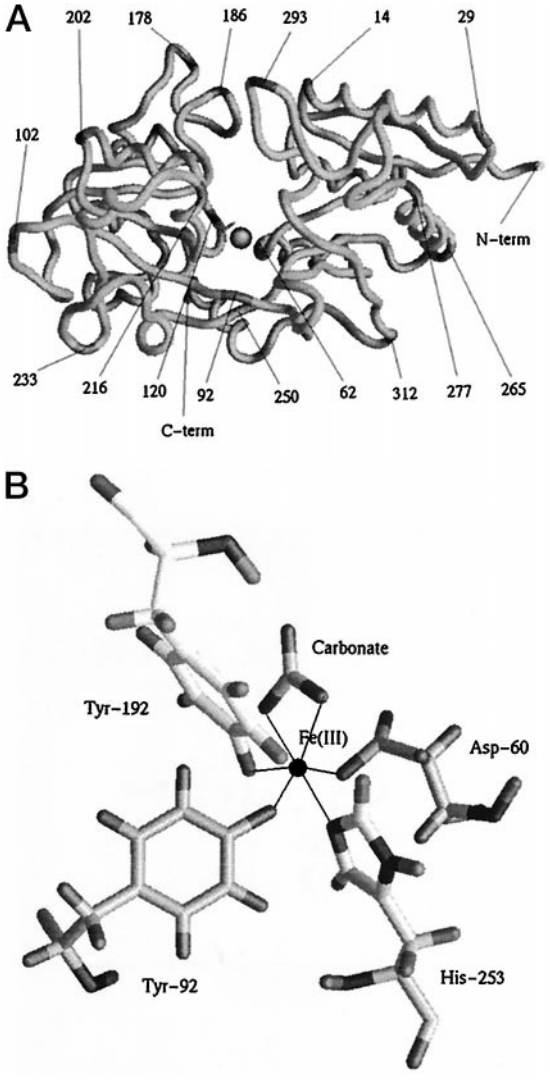


FIGURE 2 (*A*) Main chain worm representation of the N-terminal half-molecule of human ferri-lactoferrin with some α-carbon positions indicated to assist in locating the positions of residues discussed in the text. (*B*) The iron-binding site of the N-lobe of human lactoferrin. The geometry of the complex is distorted octahedral.

10–12 Å beneath the molecular surface, its environment is distinctly hydrophilic. The two-domain structure, with the binding site in the cleft between the domains, provides a mechanism for binding and release; a hinge in the backbone strands behind the iron site would allow the cleft to be opened or closed over the bound substrate by relative movement of the domains (Gerstein et al., 1993). An important part of this design is that the metal takes its ligands from four different parts of the structure, the Asp residue from the first domain, one Tyr residue from the second domain, and the other two ligands from the two backbone strands.

The x-ray crystal structure of human apo-lactoferrin (Anderson et al., 1990) shows that iron and carbonate removal produces a very large conformational change in the N-terminal half of the molecule (see Fig. 1). The NII domain has rotated by 54°, relative to the NI domain, about an

axis near the back of the iron site, such that the N-lobe binding cleft is opened wide. This results in the exposure of a number of residues previously buried in the interdomain cleft, including several basic side chains. It also results in the separation of the potential iron ligands, such that the two tyrosine residues remain with domain NII (which also carries the anion site) whereas the histidine and aspartate ligands remain with domain NI. The conformational change is essentially a rigid-body movement of the whole NII domain and is achieved by flexing of the hinge region. Various physical studies, including hydrodynamic measurements (Rossenau-Motreff et al., 1971) and small-angle x-ray and neutron scattering studies (Kilar and Simon, 1985; Vigh et al., 1989; Grossman et al., 1992, 1993) support such a conformational change and show that the molecule is markedly more compact when iron is bound.

Iron release can be triggered by a reduced pH, with release from serum transferrin beginning at a pH of 6.0 and being complete at a pH of 4.0, whereas for lactoferrin, release occurs in the pH range from 4.0 to 2.5 (Mazurier and Spik, 1980). The conventional view has been that, in vivo, iron-loaded transferrin binds to its cellular receptor, is internalized, releases its iron at the lower intravesicular pH (5.5), and is then recycled to the outside of the cell as apoprotein (for recent reviews see Qian and Tang, 1995; De Silva et al., 1996). The only experimentally determined pK_a values that have been reported for transferrins, however, are for some of the histidine residues in half-molecules of ovotransferrin (Woodworth et al., 1987) and for some of the histidine residues in human serum apo-transferrin (Kubal et al., 1994). No similar results are available for human lactoferrin, which also has a somewhat different distribution of histidine residues. In the absence of experimentally determined pK_a values for most of the titratable residues in transferrins, we have used computational methods.

A mechanism for iron release has been proposed, involving a pH-sensitive dilysine trigger in the N-lobe of hen ovotransferrin (Dewan et al., 1993) based on the crystal structure in which an unusual interdomain interaction is formed between the NZ atoms of Lys-209 and Lys-301, which are 2.3 Å apart and which are also involved in hydrogen-bonding interactions with the aromatic ring of a tyrosine residue. The protein crystals were grown at pH 5.9 (i.e., well below the usual pK_a ~10 for a lysine side chain), and it was suggested that the pK_a of either one or both of these residues lies below the pH of the structure determination and is, therefore, not positively charged and that uptake of the Fe(III)-transferrin complex into an acidic endosome via receptor-mediated endocytosis would result in the protonation of both lysine residues. The close proximity of the two resulting positive charges, and their location on opposite domains of the N-lobe, might well be the driving force that opens the two domains of the protein, exposing the Fe(III) ion and facilitating its release.

Examination of the amino acid sequences of other transferrins indicated that similar pH-sensitive dilysine triggers may be possible in the N-lobe, but not the C-lobe, of most serum transferrins. Dilysine triggers are not possible in the C-lobe of hen ovotransferrin or in either lobe of most lactoferrins. Although Lys-301 is maintained in the human lactoferrin structure, Arg-210 occupies a position similar to that of Lys-209 of ovotransferrin. Lys-301 also forms a salt bridge with Glu-216 and is ~5 Å from Arg-210. As Glu-216 and Lys-301 are on opposite domains, this salt bridge would tend to hold the domains together. Clearly, if this release mechanism is correct, then different release mechanisms must exist in different transferrins.

Gelb and Harris (1980) demonstrated that binding of ferric iron and bicarbonate to human serum transferrin at pH 7.6 and 20 mM ionic strength resulted in the release of 3.1 H^+ ions per bound metal ion, and binding of ferric iron and bicarbonate to hen ovotransferrin over the pH range 7.5–9.5 and 20 mM ionic strength resulted in the release of 2.9 H^+ ions per bound metal. The conformational change illustrated in Fig. 1 is assumed to accompany this process. Addition of oxalate ($C_2O_4^{2-}$) as the anion instead of bicarbonate (HCO_3^-) resulted in the similar release of approximately three H^+ ions per bound metal ion, and this was interpreted as indicating that the bicarbonate was not deprotonated upon binding. A growing number of x-ray crystal structures of transferrins suggest that this conclusion was incorrect and that the bicarbonate is most likely deprotonated to carbonate (CO_3^{2-}) upon binding. It is possible that the binding of oxalate (which is a larger anion than carbonate) promotes a different pattern of deprotonation in the proteins. If one of the H^+ ions is released from the bicarbonate, then this implies that there is a net release of approximately two H^+ ions from the transferrins. It is doubted that one can simply attribute the release of the other two H^+ ions to deprotonation of the two binding-site tyrosine residues, which are assumed to be in a deprotonated state when the iron is bound.

We have used continuum electrostatic calculations to generate titration curves for the two forms of human lactoferrin. Comparison of the average protonation states of the two forms of the protein at the same pH in the pH range 7.5–9.5 should yield a difference of approximately two protons, the open iron-free form of the lactoferrin being more highly protonated than the closed iron-loaded form. Analysis of the titration curves of individual residues in the two forms of the protein should then allow identification of the residues responsible for this difference.

METHODS

Model building

Initial models were constructed using the suite of programs AMBER 4.1 (Pearlman et al., 1995). Atomic coordinates were obtained from the Brookhaven Protein Databank for human ferri-lactoferrin (1lfg) and for human apo-lactoferrin (1lfh). Given the size of the protein, the degree of homology between the N-lobe and C-lobe and the fact that both have iron-binding sites, it was decided to use half-molecules in these calculations. Residues 1–333 were used to represent the N-lobe half-molecule of the protein. As coordinates for the first two residues of the protein are

missing, it was considered preferable to place an acetyl blocking group at the N terminus of the protein rather than have a titratable amino group in an incorrect position. Similarly, the C terminus of the half-molecule was artificially created, and it was thus considered preferable to add an N-methyl blocking group at the C terminus rather than have a titratable carboxylic acid group. Coordinates for the side-chain atoms of Arg-86 in the ferri-lactoferrin structure were also missing, and these were built in by the EDIT program of AMBER 4.1. An all-hydrogen-atom model was employed and partial atomic charges were taken from the AMBER 4.1 1994 parameter set where possible.

The programs used in this work to calculate pK_a values require that coordinates are given for all titratable hydrogen atoms, so all titratable residues in the protein had to be built into the models in their protonated states. The coordinates for the non-hydrogen atoms of the C-terminal blocking group were derived from the coordinates of residue 334. The coordinates of the N-terminal blocking group, the Arg-86 side chain atoms in ferri-lactoferrin and all of the hydrogen atoms, which were determined by EDIT, were subjected to simultaneous energy minimization using the belly option, which means that the positions of all other atoms were kept frozen. 10 cycles of steepest descent were followed by conjugate gradient minimization until an energy gradient tolerance of 0.01 kcal/mol Å was reached. 1–4 electrostatic interactions were divided by 1.2 as recommended for the 1994 parameters.

The pK_a values have been experimentally determined for all of the titratable residues of hen egg-white lysozyme, and so this protein provides a good system to test the predictive power of the calculation scheme presented herein. An all-atom model of hen egg-white lysozyme using the AMBER 1994 parameters was therefore also built. Atomic coordinates were obtained from the Brookhaven Protein Databank (2lzt). As no atomic partial charges appropriate to a deprotonated N terminus or a protonated C terminus have yet been calculated (see below), the termini of the protein were maintained as nontitrating groups in their charged states. The positions of all hydrogen atoms were energy minimized as described for lactoferrin.

Electrostatic calculations

Electrostatic calculations were performed using version 1.1.3 of MEAD, macroscopic electrostatics with atomic detail (Bashford, 1993). The electrostatic potential in and around a molecule in solvent is assumed to be governed by the linearized Poisson-Boltzmann equation:

$$\nabla[\epsilon(r)\nabla\phi(r)] - \kappa^2\epsilon(r)\phi(r) = -4\pi\rho(r) \qquad (1)$$

where $\phi(r)$ is the electrostatic potential at position r due to the charge distribution ρ, which has the form of point charges on atom centers; ϵ is the dielectric constant, which took the value $\epsilon_s = 80$ in the solvent-accessible region and $\epsilon_m = 4$ in the molecular interior; and κ is a modified Debye-Hückel parameter, which accounts for counterion screening in the solvent. The solvent-accessible region was defined as the volume that could be swept out by a spherical probe of radius 1.4 Å without overlapping any atomic radii of the molecule; what remained was defined as the molecular interior. A finite difference approach was employed to find the electrostatic potential ϕ in the model system. A two-step focusing method was employed using an initial 81^3 grid with a 1-Å spacing centered on the geometrical center of the protein, followed by a 81^3 grid with a 0.25-Å spacing centered on the titrating residues. Calculations for the model compounds adopted the same approach but with a grid dimension of 61^3.

Calculation of additional atomic partial charges

As the AMBER 1994 atomic partial charge set for amino acids in proteins does not include values for the deprotonated forms of tyrosine, lysine, or arginine residues, it was necessary to calculate appropriate values, which was achieved by following as closely as possible the AMBER methodology. Models of these three residues in their deprotonated forms with acetyl and N-methyl blocking groups were built using QUANTA (Molecular Simulations Inc., 1986, 1992). Ab initio geometry optimizations were then performed by the unrestricted Hartree-Fock self-consistent field method at the 6–31G* level using the suite of quantum chemistry programs CADPAC 6 (Amos et al., 1995). A distributed multipole analysis (DMA) representation to rank 5 of the potential surrounding each geometry-optimized structure was then generated and imported into the program ORIENT (Stone, 1990). ORIENT was used to generate potential surfaces at 1.4, 1.6, 1.8, and 2.0 times the Van der Waals radii. Atomic partial charges were fitted to these surfaces using the RESP program from AMBER 4.1 in a two-stage fitting procedure as described for the derivation of the AMBER 1994 atomic partial charges (Pearlman et al., 1995). When entering these newly calculated charge parameters (Fig. 3, *A* and *B*) into the AMBER database, standard atomic partial charges were assumed for the main-chain atoms whereas small adjustments were made to the calculated charges to maintain the correct net charge.

Ab initio calculations were also undertaken on a pseudo-molecule representing the active site as a whole using the atomic coordinates of the iron, carbonate, and four binding-site residues. The protein residues were truncated by methyl groups at their α-carbon atoms using the molecular modeling package QUANTA. An unrestricted Hartree-Fock self-consistent field calculation was then performed for this pseudo-molecule at the 6–31G* level (STO3G for the iron atom) using CADPAC 6. The pseudo-molecule was given a net charge of $-2e$ and the iron was assumed to be in the high spin state and so the multiplicity was set to 6. A distributed multipole analysis (DMA) representation to rank 5 of the potential surrounding the pseudo-molecule was generated and imported into the program ORIENT. RESP charges were then calculated as described above. The two tyrosine residues were constrained to be equivalent, and standard atomic partial charges were assumed for the main-chain atoms with small adjustments to the calculated charges being made to maintain the correct net charge (Fig. 3 *A*).

pK_a calculations

The fractional degree of protonation, f_i, of any ionizable group, i, is given by

$$\ln[f_i/(1 - f_i)] = 2.3(pK_{ai} - pH) \qquad (2)$$

The relationship between the pK_a of an isolated amino acid in solution, pK_{ai}^0, and the pK_a of the same group in a protein with a single titratable site, pK_{ai}^{int} (the intrinsic pK_a), is given by

$$pK_{ai}^{int} = pK_{ai}^0 - \gamma(i)\Delta\Delta G_i^{env}/2.3k_BT \qquad (3)$$

where $\Delta\Delta G_i^{env}$ is the change in electrostatic energy of ionizing the group in the protein environment relative to solution. $\gamma(i) = -1$ or 1 for an acidic or basic group, respectively, and K_B and T are the Boltzmann constant and absolute temperature, respectively. The $\Delta\Delta G_i^{env}$ can be calculated by taking the difference between the difference in free energy of the group in the unprotonated state with its protein environment relative to the reference state in solution, $\Delta G_i^{env}(A)$ and that of the corresponding term for the same group in the protonated state, $\Delta G_i^{env}(AH)$.

Each of these terms can be considered as being due to two contributions, namely, the difference in reaction field energy in the isolated state and in the protein and the contribution of the permanent dipoles and any other pH-independent electric field in the protein. The latter would include, for example, the effects of bound ions and water molecules and the distribution of electrostatic potential on the other residues in the protein in their uncharged states. The reaction field term is just the contribution of those parts of the system treated as a dielectric continuum. These calculations were implemented using the Multiflex program of MEAD with the AMBER 4.1 1994 charge and atomic radii parameter sets.

When the protein has more than one titratable group, the mutual charge-charge interactions at the titrating site become pH dependent. pK_{ai} is now obtained by determining the pH at which the group i is 50% protonated. This is done by calculating a titration curve for the entire

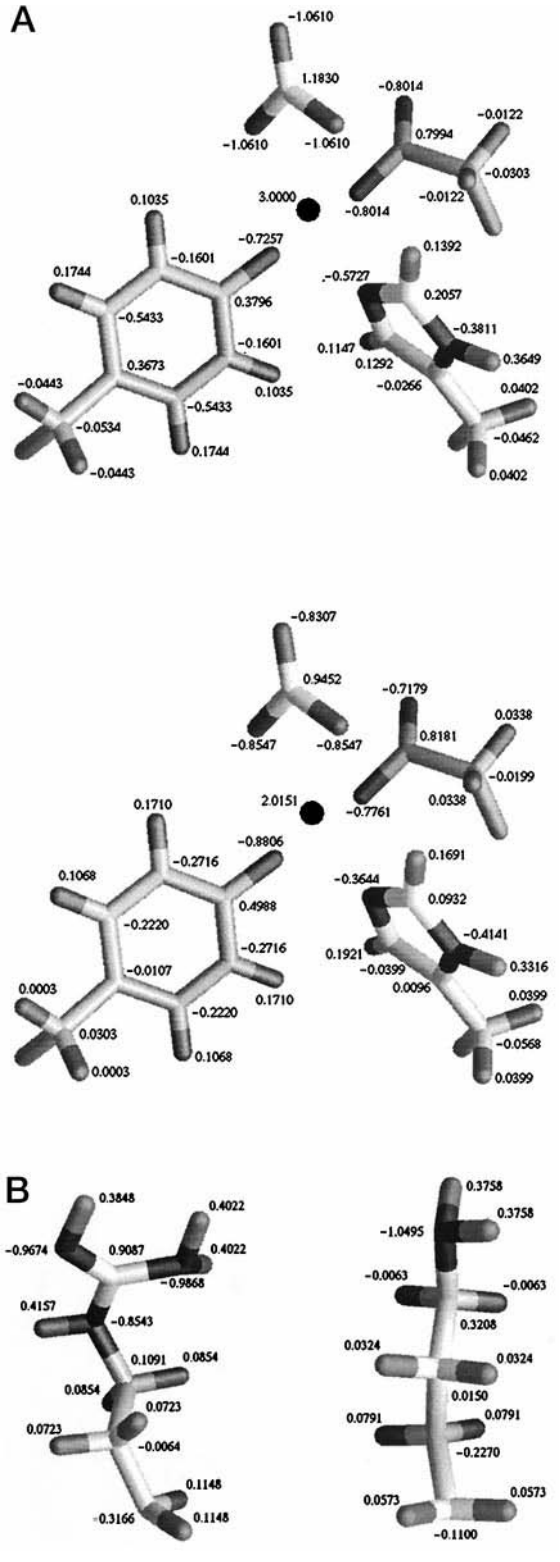


FIGURE 3 (*A*) Atomic partial charges used in modeling the iron-binding site, before and after adjustment to account for interactions between the iron and its ligands. The atomic partial charges for the deprotonated form of a tyrosine residue used in the pK_a determinations are also shown. All main chain atoms are assigned standard AMBER 4.1 atomic partial charges and so are not shown. The two tyrosine residues in the binding site were constrained to be equivalent and so only one is shown. (*B*) Atomic partial charges assigned to the side chain atoms of the neutral forms of arginine and lysine residues.

system. In most cases, the pK_a of group i can now be written in the form

$$pK_{ai} = pK_{ai}^{int} + \Delta pK_{ai}^{ttr} \quad (4)$$

The last term in Eq. 4 is simply due to charge-charge interactions between titratable groups.

In principle, the titration curve of the system can be obtained from a statistical mechanical average over the 2^N states that arise from the N ionizable sites; however, this becomes impractical as N reaches several tens of titratable groups. An alternative approach uses a Monte Carlo sampling technique, which has been implemented in the program Monti (Beroza et al., 1991). At each step of 0.1 pH units, 1000 full steps of Monte Carlo were performed followed by 5000 reduced steps in which the residues most likely to remain fully protonated or deprotonated were omitted from the sampling (tolerance $= 1 \times 10^{-6}$ proton). Increasing the number of reduced steps from 4,000 to 10,000 in blocks of 1,000 indicated that the calculation converged after 5,000 steps to within 0.1 pK unit with most values being within 0.02 pK unit (results not shown).

Lysozyme

These calculations were performed at 293 K with an ionic strength of 150 mM to allow direct comparison with the results of Antosiewicz et al. (1996). For the histidine residue, atom NE2 was assumed to be deprotonatable. Antosiewicz et al. (1996) reported that the computed pK_a of residue Glu-35 was sensitive to the details of proton placement. When they allowed this residue to be protonated on the alternative carboxylic oxygen atom (OE1) rather than the default choice (OE2), they found a significant improvement in the calculated pK_a value for this residue compared with the experimental value. Therefore, we also repeated our calculations with both protonation states for this residue. The results of our test calculations on this system (Table 1) show that we obtain similar overall agreement with experimental pK_a values as Antosiewicz et al., i.e., a root mean squared deviation (rmsd) of 2.2 pK units, and we also see the effect of the alternative protonation sites for Glu-35. If, however, the problematic residue Tyr-53 is excluded from the analysis, then our rmsd becomes 1.7 pK units compared with 1.0 pK units for Antosiewicz et al., suggesting that our parameters and methods may give a somewhat poorer agreement with experiment.

Lactoferrin

We adopted the same overall procedure as used successfully for lysozyme. For the four histidine residues, atom NE2 was assumed to be the deprotonatable atom by inspection of the environments of these residues in the x-ray crystal structures. In simulating the experimental conditions of Gelb and Harris (1980), the calculations were carried out at 294 K and with an ionic strength of 20 mM. Calculations were also repeated at an ionic strength of 145 mM and at 310 K to simulate physiological conditions. The four active-site binding residues were found to be fully deprotonated at a pH below which any of the other residues in the protein started to titrate, due to the proximity of the 3+ background charge associated with the iron. Therefore, these residues were regarded as nontitrating and contributing only to the background charges. The pK_a values were recalculated for ferri-lactoferrin with the four binding-site residues held in their deprotonated states. A further refinement was then made where the atomic partial charges on the deprotonated binding site residues were adjusted to attempt to create a more realistic model of the chemical nature of iron binding. The pK_a values for the ferri-lactoferrin were then calculated again with this modification.

RESULTS

The pK_a values for titratable residues in the N-terminal half-molecule of apo-lactoferrin have been calculated. The calculation has been repeated for the iron-bound form with

TABLE 1 Comparison of PK_a values for hen egg-white lysozyme

Residue	PROTA*	PROTA#	PROTB*	PROTB#	Experimental
N terminus			5.6	5.6	
Lys-1	9.1	9.1	10.1	10.1	10.6
Arg-5	10.5	10.5	12.8	12.8	
Glu-7	2.8	2.8	2.9	2.8	2.9
Lys-13	10.3	10.3	10.7	10.7	10.3
Arg-14	12.5	12.4	12.4	12.4	
His-15	5.3	5.3	5.1	5.1	5.4
Asp-18	1.9	1.9	2.7	2.7	2.7
Tyr-20	13.2	13.1	12.4	12.4	10.3
Arg-21	12.3	12.3	12.7	12.7	
Tyr-23	11.6	11.6	9.7	9.7	9.8
Lys-33	12.8	12.9	10.9	10.9	10.4
Glu-35	3.9	4.8	3.5	5.6	6.2
Arg-45	11.8	11.8	12.2	12.2	
Asp-48	0.6	0.7	1.5	1.5	1.6
Asp-52	2.0	1.8	4.3	3.6	3.7
Tyr-53	18.0	18.0	20.5	20.5	12.1
Arg-61	12.1	12.1	13.8	13.7	
Asp-66	5.8	5.8	1.9	1.9	0.9
Arg-68	15.3	15.3	13.0	13.0	
Arg-73	12.7	12.7	12.5	12.5	
Asp-87	2.2	2.2	2.4	2.4	2.1
Lys-96	11.6	11.6	13.6	13.6	10.7
Lys-97	11.4	11.4	11.2	11.2	10.1
Asp-101	4.4	4.3	4.4	4.4	4.1
Arg-112	11.7	11.6	11.2	11.2	
Arg-114	12.9	12.9	13.3	13.3	
Lys-116	9.3	9.3	9.3	9.3	10.2
Asp-119	3.0	3.0	3.7	3.7	3.2
Arg-125	13.3	13.3	12.9	12.9	
Arg-128	12.1	12.1	12.2	12.2	
C terminus			3.1	3.1	2.8
rmsd	2.2	2.2	2.2	2.1	

Comparison of the pK_a values calculated for hen egg-white lysozyme (PROTA) with those determined by experiment and with the computed values reported by Antosiewicz et al. (1996) (PROTB). The final line of the table provides the cumulative rmsd of the computed values from the available experimental values.

*Results when all default protonation sites were used.

#Results when OE1 of Glu-35 was treated as the protonatable atom.

TABLE 2 pK_a values calculated for apo and ferri forms of the half-molecule of lactoferrin

		Ferri form		
Residue	Apo form	a	b	c
Glu-15	0.4	−2.7	−3.0	−4.1
Arg-39	10.2	13.1	13.0	13.1
Asp-55	−4.4	−6.7	−6.8	−6.9
Asp-60*	−2.2	−34.2		
Tyr-65	13.9	14.1	12.0	12.4
Glu-66	4.2	1.2	1.3	1.6
Arg-75	15.1	18.3	18.2	18.2
Glu-80	−4.9	−7.2	−6.2	−1.7
Tyr-82	31.2	38.5	39.2	32.8
Arg-89	22.2	17.3	17.8	15.4
His-91	3.4	0.9	0.7	−0.1
Tyr-92*	18.7	−16.0		
Tyr-93	15.7	16.2	16.0	17.9
Lys-100	11.3	14.3	14.3	14.3
His-116	−8.5	−1.4	−1.3	−0.5
Arg-121	8.9	27.7	27.0	28.1
Arg-151	15.2	17.6	17.6	17.6
Tyr-189	23.5	16.9	16.9	16.9
Tyr-192*	1.4	−31.0		
Lys-197	13.3	23.2	23.2	23.2
Asp-201	−0.1	2.4	2.3	2.4
Arg-210	16.7	23.1	22.1	42.0
Glu-216	2.7	−2.8	−1.7	−2.4
Asp-217	6.2	10.5	10.0	10.0
Lys-237	9.3	14.9	14.9	14.9
Lys-243	10.9	10.2	10.1	7.7
His-246	2.1	4.2	4.2	4.4
His-253*	6.3	−52.9		
Lys-277	6.6	11.0	11.0	11.0
Lys-280	9.8	12.1	12.1	12.1
Lys-282	13.9	7.4	7.4	7.5
Lys-285	10.0	0.3	0.3	0.3
Lys-296	11.9	19.9	19.9	19.9
Asp-297	3.3	−10.6	−11.6	−10.4
Lys-301	14.6	34.6	34.9	34.8
Asp-302	−1.3	5.4	5.6	5.8
Tyr-319	20.6	25.3	24.9	24.7
Tyr-324	15.0	7.2	16.7	20.6
(Lys-28)	9.3	8.9	8.9	8.9

Only those residues are listed for which there is a >2 pK unit difference between any two models, except for residue Lys-28, which is also referred to in the text. The results of using three models of the ferri form of the binding-site residues are shown: a, with binding site residues titrating; b, with the binding-site residues nontitrating and held in their deprotonated states; and c, same as b but with atomic partial charges adjusted to reflect the chemistry of the metal ion complex. The experimental conditions of Gelb and Harris (1980) were used (temperature = 294 K; ionic strength = 20 mM).

*Binding-site residues.

three different models of the binding site. The data are presented for individual residues in Table 2, and the accumulated effect for the whole N-lobe can be seen in Fig. 4 *A* for both apo and iron-bound (model ferri(c)) forms. In Table 2 we have listed only those residues for which there is a >2 pK unit difference between any two models. We have also highlighted the residues involved directly in iron binding, i.e., Asp-60, Tyr-92, Tyr-192, and His-253.

In apo-lactoferrin, with the exception of Tyr-192, all of the tyrosine residues have substantially elevated pK_a values. The tyrosine residues Tyr-82, Tyr-189, and Tyr-319 start with elevated pK_a^{int} values due in most part to their low solvent exposure. Charges on anionic titratable residues contribute further to the elevation of their pK_a values, e.g., Glu-80, which is very close to Tyr-82. As the pH increases, the negatively charged, deprotonated form of the tyrosine residues is further disfavored by the increasing net negative charge on the protein. The intrinsic pK_a of His-116 is also substantially lowered by the buried position of the residue in the protein, and the pK_a is shifted further downwards by the proximity of Arg-121.

There appears to be a competition for protonation between the two iron-binding ligands Tyr-92 and Tyr-192. Both residues start with fairly low pK_a^{int} values (8.1 and 7.0), due in most part to the environment of background charges. Each residue contributes substantially to the background charges experienced by the other, and there may be

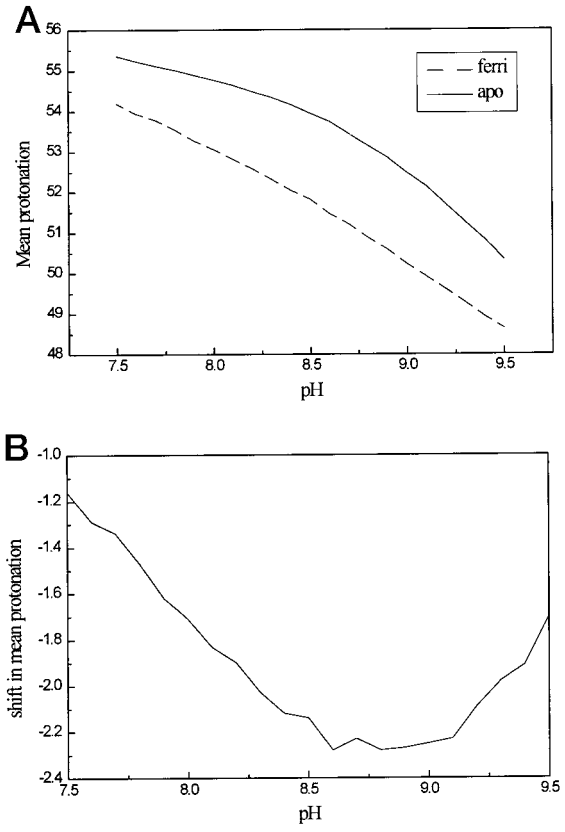


FIGURE 4 (*A*) The calculated titration curves of the N-terminal half-molecules of human apo- and ferri-lactoferrin over the pH range 7.5–9.5. The mean protonation is the mean number of titratable hydrogen atoms that are predicted to remain on the protein at a given pH. (*B*) The predicted mean shift in protonation of the N-terminal half-molecule of human lactoferrin over the pH range 7.5–9.5 as a result of the transition from the apo to the ferri form of the protein.

overcrowding of atoms when both residues are protonated. The proximity of Glu-212, however, tends to elevate the pK_a of Tyr-92 whereas the proximity of two arginine residues, Arg-121 and Arg-210, tends to keep the pK_a of Tyr-192 at a low level. These interactions tip the balance in favor of Tyr-192 being more readily deprotonated, and the resulting net negative charge on this residue then drives the pK_a of Tyr-92 even higher. The pK_a^{int} of Lys-277 is low (7.6) due to its buried position in the protein. Electrostatic interactions with other charged residues are weak, but the pK_a is kept low due to the residue being surrounded by more cationic residues (Arg-272, Lys-282, Arg-30, and Lys-18) than anionic residues (Glu-276).

The transition from apo-lactoferrin to the model of ferri-lactoferrin with titrating binding site residues (ferri(a)) is accompanied by some dramatic shifts in pK_a values (Table 2). The pK_a values of the binding-site residues are all obviously predicted to be massively lowered by the proximity of the 3+ charge on the iron atom. Lys-277 becomes more readily protonated because the solvent exposure of the residue is increased by the conformational change. Asp-217 has a fairly high pK_a^{int} (5.9) in the apo-lactoferrin structure due to both partial burial in the protein and adverse interactions with the background electrostatic potential. This situation is exacerbated in the ferri-lactoferrin structure by increases in both factors such that the pK_a^{int} becomes 9.5. Lys-285 is predicted to be deprotonated at pH 7.4 due to the partial burial of the residue in the protein and its movement toward the positively charged residue Lys-28. The lower pK_a^{int} of Lys-285 compared with that of Lys-28 ensures that Lys-285 is the one that is deprotonated more readily.

The pK_a of Tyr-324 is predicted to be substantially lowered in the model with titrating binding-site residues but not in the other models. This leads to the difference between models ferri(a) and ferri(b) in the predicted release of H^+ ions. The effect of this stage in the refinement of the model building is to worsen the apparent agreement with experiment.

In model ferri(a), Tyr-324 has a pK_a^{int} of 0.7 due to the background charge of 3+ associated with the iron atom. In model ferri(b) Tyr-324 has a pK_a^{int} of 13.9 as the net negative charges on three of the binding-site residues are now also background charges and Asp-60 and Tyr-92 are closer to Tyr-324 than the iron atom is. In model ferri(a), the negative charges on the binding-site residues have their effect as titrating residues and result in a pK_a of 7.2 for Tyr-324. The much higher pK_a of 16.7 when model ferri(b) is assumed results from the fact that the phenyl hydrogen atom in Tyr-324 adopts a position such that it donates to a hydrogen bond between this residue and the hydroxyl oxygen atom of Thr-122. In model ferri(a), this does not happen as in this case the binding-site residue Asp-60 is protonated and this additional hydrogen atom adopts a position such that it now donates to a hydrogen bond with the hydroxyl oxygen atom of Thr-122 and the hydroxyl hydrogen atom of Thr-122 now becomes the hydrogen donor in a hydrogen bond with the phenyl oxygen atom Tyr-324. The result is that the phenyl hydrogen atom in Tyr-324 is forced to adopt a much less favorable position in model ferri(a) due to an extra hydrogen atom on residue Asp-60, which is very unlikely to be there in any realistic protonation state of ferri-lactoferrin. Model ferri(b) is thus assumed to be a better model as its fully protonated state is closer to what is assumed to be the protonation state of the protein over the experimental range 7.5–9.5 and at the physiological pH of 7.4.

A few quite large pK_a shifts are seen in residues close to the binding site with the transition from model ferri(b) to model ferri(c). These are due to the changes in the background charges in the binding-site residues contributing to shifts in the pK_a^{int} values of residues. The downward shift in the pK_a of Lys-243 is the main reason for the difference between models ferri(b) and ferri(c) in the predicted release of H^+ ions and for the restoration of a good agreement with experimental results comparable with that of model ferri(a). Using model ferri(b), the pK_a^{int} of Lys-243 is 10.4, but using model ferri(c), it is 8.2. Lys-243 is in fact quite distant from the iron-binding site (Fe(III) − NZ distance = 16.41 Å), and the difference in pK_a^{int} is due to a difference in the position adopted by the phenyl hydrogen atom on Tyr-93. In

model ferri(c), the phenyl hydrogen atom on Tyr-93 is much closer to the NZ atom of Lys-243 (1.972 Å compared with 3.062 Å) and thus destabilizes a positive charge on this residue. The difference in the optimal position adopted by the phenyl hydrogen atom of Tyr-93 must be a long-range effect of the redistribution of charge at the iron-binding site.

The massive pK_a shift for Arg-210 in the transition from model ferri(b) to model ferri(c) is due to the redistribution of electrostatic potential in the binding-site residues; i.e., the positive charge on Arg-210 helps stabilize the redistribution of electrostatic potential that occurs when iron binds, and its presence in that position in the protein is thus predicted to favor iron binding. Model ferri(c) is assumed to be the best model for ferri-lactoferrin because the four unnecessary hydrogen atoms on the binding-site residues are omitted from the calculation scheme, thus leaving more space for other titratable hydrogen atoms to find their optimal position and because the chemical interaction between the iron and its ligands is fully accounted for. Polarizability of the protein is modeled by the dielectric constant of 4, but this may be inadequate to account for the polarization of its ligands by the ferric iron atom.

Titration curves were prepared for the half-molecule of apo-lactoferrin and for the three model forms of the half-molecule of ferri-lactoferrin to allow comparison of the calculated results with the experimental results of Gelb and Harris (1980). Comparison of the titration curves for the apo and ferri forms of the protein revealed the predicted number of H^+ ions released in the experimental pH range 7.5–9.5 (Table 3). The titration curves for apo-lactoferrin and the most accurate model of ferri-lactoferrin (ferri(c)) are shown in Fig. 4 *A*. The difference between them is highlighted in Fig. 4 *B*. This difference can be decomposed into the contributions made by individual residues by preparing titration curves for each residue in the two forms of the protein and calculating the difference between them as for the full protein. An example of the titration curves calculated for one residue (Lys-285) is shown in Fig. 5 *A*, and the difference between these two curves and the differences for the other twelve residues that make the major contribution to the overall difference for the protein in this pH range are shown in Fig. 5 *B*. The characteristic sigmoidal shape of the titration curves for Lys-285 is typical of those calculated for all of the residues in this study. Note that titration of the residue extends over the fairly wide pH range of approximately 4 units.

TABLE 3 Predicted number H^+ ions released for the three model forms of ferri-lactoferrin

	Ferri		
pH	a	b	c
7.6	1.5 (2.5)	0.9 (1.9)	1.3 (2.3)
7.5–9.5	2.0 (3.0)	1.3 (2.3)	1.9 (2.9)

The pH range shown allows comparison with the experimental results of Gelb and Harris (1980), which showed the release of approximately three H^+ ions per bound metal ion. The values in parentheses include deprotonation of bicarbonate.

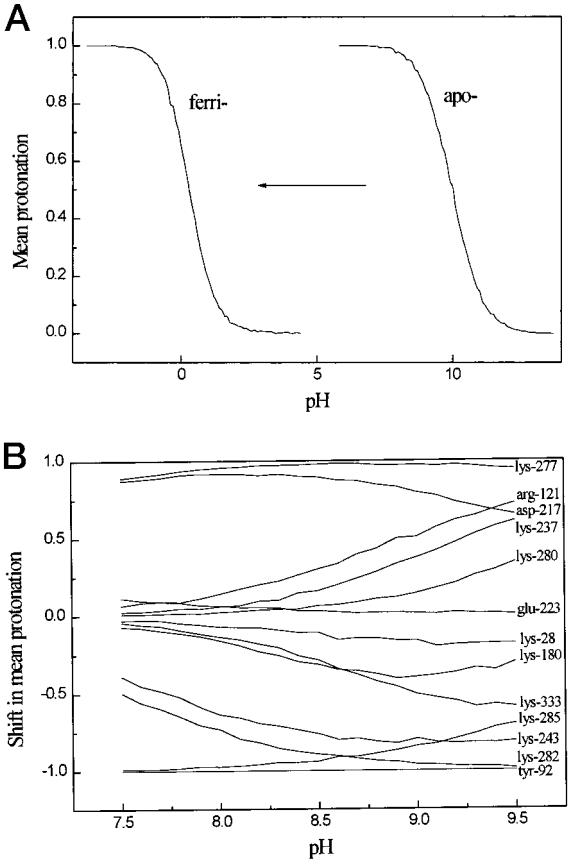


FIGURE 5 a. Calculated titration curves for Lys-285 in the N-terminal half-molecules of human apo- and ferri-lactoferrin. The arrow indicates the shift that is predicted to occur when the protein makes the transition from the apo to the ferri form. (*B*) The shift in the mean protonation states of the thirteen residues that are predicted to change in the pH range 7.5–9.5 when the N-terminal half-molecule of human lactoferrin makes the transition from the apo to the ferri form.

When protonation state calculations were performed with the simulation of physiological conditions (temperature = 310 K; ionic strength = 145 mM) and model ferri(c) of ferri-lactoferrin, the release of 1.1 H^+ ions was predicted upon the transition from apo- to ferri-lactoferrin (2.1 if the bicarbonate is assumed to be deprotonated also). The calculated pK_a values varied to a small degree from those shown in Table 2, which were calculated assuming the experimental conditions of Gelb and Harris (1980). Table 4 shows the values calculated for those residues that are assumed to be predominantly in different protonation states at pH 7.4.

The work of Demchuk and Wade (1996) suggests that computed pK_a values are sensitive to the value chosen for the protein dielectric constant (ϵ_p). Higher values can, through scaling electrostatic interactions, compensate for features that are not explicitly present in the model, such as dynamic motions. All calculations have been repeated with ϵ_p values of 10 and 20. For the model of lysozyme with

TABLE 4 Comparison of the calculated pK_a values for the N-terminal half-molecules of human apo and ferri-lactoferrin under simulated physiological conditions

Residue	Apo	Ferri(c)
Tyr-92	17.6	<7.4
Asp-217	6.3	9.3
Lys-277	7.1	11.0
Lys-285	10.1	1.1

The model form of the ferri-lactoferrin is the same as model c in Table 2. Only those residues for which a shift in the pK_a value from one side of pH 7.4 to the other is predicted are shown. Temperature = 310 K; ionic strength = 145 mM.

default protonation sites, the rmsd of computed from experimental pK_a values is improved, to give a rmsd of 1.0 pK units for both values of ϵ_p, even when the result for the problematic residue Tyr-53 is included. In fact, when an ϵ_p of 10 is assumed, a very realistic pK_a of 11.7 is computed for this residue (experimental pK_a = 12.1). The rmsd values of 1.0 pK unit are significantly better than that of 1.4 pK unit attained when the null hypothesis is assumed; i.e., the pK_a values of amino acids in aqueous solution are assumed, with the protein environment having no effect.

When ϵ_p values of 10 and 20 are employed in the calculation of pK_a values for the N-terminal half-molecule of human lactoferrin, it is striking that the computed titration curves, in the relevant pH range of 4–9.5, for the apo and ferri (model ferri(c)) forms of the protein are not altered significantly. Simulation of the experimental conditions of Gelb and Harris (1980) results in the average net release of H^+ ions, in the pH range of 7.5–9.5, of 1.8 with ϵ_p = 10 and 1.6 with ϵ_p = 20 (i.e., 2.8 and 2.6, respectively, assuming deprotonation of bicarbonate upon binding). At pH 7.6, the net release of H^+ ions is 3.0 with ϵ_p = 10 and 2.9 with ϵ_p = 20 (including deprotonation of bicarbonate). A very similar degree of protonation in the transition under physiological conditions from ferri-lactoferrin at pH 7.4 to apo-lactoferrin at pH 4 is also predicted. A gain of 12.8 and 12.6 protons is predicted using ϵ_p = 10 and ϵ_p = 20, respectively, compared with our original calculation of 12.0 protons with ϵ_p = 4. The net release of H^+ ions under physiological conditions at pH 7.4 due to iron and carbonate binding is a little more sensitive, being 2.2 and 1.8 protons with ϵ_p = 10 and ϵ_p = 20, respectively, compared with 1.1 using ϵ_p = 4.

As expected, some sensitivity to the value chosen for ϵ_p is seen for individual residues. However, we find, when simulating the experimental conditions of Gelb and Harris (1980), that 69% of residues contributing to the release of approximately two H^+ ions with ϵ_p = 4 are the same as those that contribute with ϵ_p = 10 and 62% with ϵ_p = 20. The degree to which these remaining residues contribute, of course, differs somewhat. The prediction of unusual protonation states in the apo and ferri forms of the protein under physiological conditions at pH 7.4 are also somewhat sensitive to the choice of ϵ_p. In the apo form, the pK_a values of Tyr-192 and Lys-277 are 7.0 and 9.0 with ϵ_p = 10 and 7.9 and 9.7 with ϵ_p = 20, compared with 1.4 and 6.6 with ϵ_p = 4. In the ferri form, the pK_a values of Asp-217 and Lys-285 are 5.3 and 6.4 with ϵ_p = 10 and 3.2 and 8.0 with ϵ_p = 20, compared with 10.0 and 0.3 with ϵ_p = 4. The predictions of the deprotonated states of Tyr-192 and Lys-285 are clearly more robust.

DISCUSSION

Lactoferrin has been used here as a model for the transferrin family as x-ray crystal structures are available for both the open and closed forms of the protein. The fact that both crystal forms of human lactoferrin were grown at pH 7.8 indicate that their protonation states should be reasonably appropriate to the physiologically relevant pH of 7.4 and to the experimental pH range (7.5–9.5) used by Gelb and Harris (1980). Experimental evidence is available to support the use of lactoferrin as a model for serum transferrin. Chung and Raymond (1993) made a comparison of the relative rates of ferric iron removal at pH 7.4 and 37°C from human milk lactoferrin and human serum transferrin by a synthetic triatecholate sequestering agent. Results indicated that iron removal from lactoferrin proceeds by a mechanism essentially the same as that for removal of iron from serum transferrin. Day et al. (1992) have expressed an N-terminal half-molecule (residues 1–333) of human lactoferrin from cloned cDNA in baby hamster kidney cells. Crystals have been grown, and a structure determination is awaited. Iron release from the half-molecule occurred in the pH range 6.0–4.0 compared with 4.0–2.5 for native lactoferrin and 6.2–4.0 for transferrin. These results suggest that the more facile release of iron from the half-molecule compared with the native lactoferrin resulted from the absence of stabilizing contacts between the N- and C-terminal halves and that the characteristic difference in pH stability between lactoferrins and transferrins is due primarily to differences in these interactions. This observation is interpreted as strengthening the case for the use of the N-lobe of human lactoferrin in this study. Not only are the contacts between the N- and C-lobes limited, but moreover, these are composed of mostly nontitratable residues.

El Hage Chahine and Pakdaman (1995) have studied the in vitro kinetics of iron release from human serum holotransferrin and have invoked a mechanism that is independent of the nature and concentration of competing ligands and of receptor binding. It is controlled by a slow proton transfer attributed to a rate-limiting slow proton gain by a protein ligand subsequent to a fast decarbonation of the N-site. Such a mechanism offers hope that simulation of iron release can be achieved with a relatively simple model, without the need to explicitly represent a competing ligand or receptor binding.

In reproducing the experimental results of Gelb and Harris (1980), the calculations were intended to identify residues that are likely to have different protonation states in the two conformations of the half-molecule of human lactoferrin at the same pH. A few quite extreme pK_a shifts in the

two structures of lactoferrin were predicted, and these are most likely artifacts of the method of solution of the multiple titration site problem for this large protein. Many of these are irrelevant to this investigation as the same protonation state in the pH range 7.5–9.5 is predicted in both structures.

Atypical predominant protonation states predicted to be found in apo-lactoferrin in the pH range 7.5–9.5 are for the deprotonated residues Tyr-192 and Lys-277, and for ferri-lactoferrin, these are for the deprotonated residues Tyr-92, Tyr-192, Lys 243, Lys-282, and Lys-285 together with a protonated Asp-217. Although Fig. 5 *B* shows that 13 residues are predicted to contribute to the observed net release of approximately two H^+ ions from the protein upon binding of iron and carbonate and the conformational change, the major contributions can clearly be seen to be due to the deprotonation of residues Tyr-92, Lys-243, Lys-282, and Lys-285 together with the protonation of residues Asp-217 and Lys-277. The positions of these residues in the protein structure are shown in Fig. 6. The apparent agreement with the experimentally observed release of H^+ ions is reassuring. However, with the exception of Tyr-92, the other five residues all have computed pK_a values in the apo and/or ferri forms, which are close to the experimental pH range, and given the estimated accuracy of the calculations employed here (rmsd = 2.2 pK units for lysozyme), some caution should be exercised in the interpretation of these results. The method does not allow precise quantification of the relative contributions of individual residues.

Calculations were also undertaken with physiologically relevant conditions. If one assumes that the most probable protonation state corresponds to the biologically significant form of the protein, then, at pH 7.4 it is necessary only to ask for which residues the pK_a has shifted from one side of 7.4 to the other as a consequence of iron and carbonate binding and the conformational change. In Table 3 it can be seen that these residues are Tyr-92, Asp-217, Lys-277, and Lys-285. Lys-243 and Lys-282 are predicted to be much less likely to deprotonate if iron binding occurs under physiological conditions at pH 7.4. In response to iron and carbonate binding and the conformational change under physiological conditions, Tyr-92 and Lys-285 are predicted to become deprotonated and Asp-217 and Lys-277 to become protonated. If Tyr-192 were already deprotonated before iron binding, this could contribute to priming the binding site.

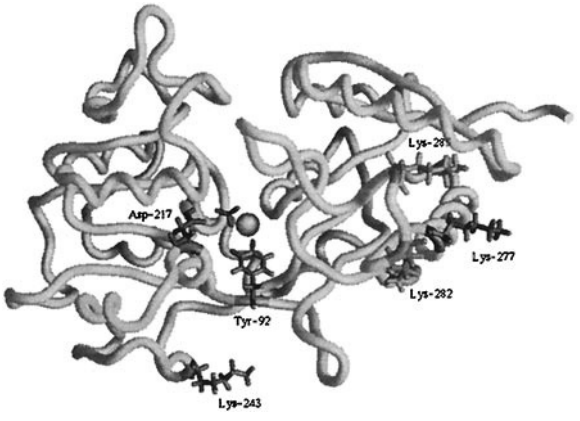


FIGURE 6 The locations in the N-terminal half-molecule of human ferri-lactoferrin of carbonate and the six residues that are predicted to make a major contribution to the net release of three H^+ ions upon metal and carbonate binding. Residues Tyr-92, Lys-243, Lys-282, and Lys-285 and bicarbonate are deprotonated whereas Asp-217 and Lys-277 are protonated. The Fe(III) ion is represented by a sphere.

The suggestion that the proximity of Lys-28 to Lys-285 in the ferri-lactoferrin structure is the primary reason for the downward shift in the pK_a value for Lys-285 is reminiscent of the dilysine trigger mechanism proposed by Dewan et al. (1993) and discussed in the Introduction. In this case, however, the two lysine residues are not located on opposing domains but are both on the same domain and quite distant from the binding-site cleft. With 4 Å between the two NZ atoms in the crystal structure, these lysine residues also do not appear to be hydrogen bonded to each other.

The crystal structures indicate that, upon iron binding, the loop where Lys-285 is located moves slightly toward the α-helix where Lys-28 is located. The major difference, however, is in the moving together of the side chains of these residues. It is expected that during the pH-induced release of iron from transferrins there is some mechanism in operation to communicate changes in the pH of the solvent to the interior of the protein. The position of Lys-28 and Lys-285 near the surface of the protein make these residues candidates for involvement in such a mechanism. The structural changes in this region do, however, appear to be very localized.

The position of Lys-285 on a surface exposed loop also promotes speculation that it would be very suitable to form part of a receptor-binding site. Not only would there be a difference in the shape of this region of the protein surface depending upon whether iron was bound or not but there would also be a difference in the electrostatic potential at the surface (Fig. 7). It is clear from analysis of crystallographic B-values that this is a highly mobile region in lactoferrin. In apo-lactoferrin the average B-factor for Lys-285 is 44.59 $Å^2$ for main-chain atoms and 64.258 $Å^2$ for side-chain atoms, and these values are even higher in ferri-lactoferrin.

A reduction in pH from pH 7.4 to ~4 is required to fully induce the release of iron from the N-lobe of human lactoferrin (Day et al., 1992). Inspecting the pK_a values calculated for apo-lactoferrin under physiological conditions indicates that, in addition to the protonation of Tyr-92, Lys-285, and presumably the carbonate, a number of other residues become protonated, and these are listed in Table 5 (note that Lys-277 is predicted to be protonated at pH 7.4 in ferri-lactoferrin). We have an indication of the protonation states of ferri-lactoferrin at pH 7.4 and apo-lactoferrin at pH 4.0 that suggests that the protein gains 12.0 protons (and presumably the carbonate gains one proton) during the process of releasing the Fe(III) ion. Knowledge of these

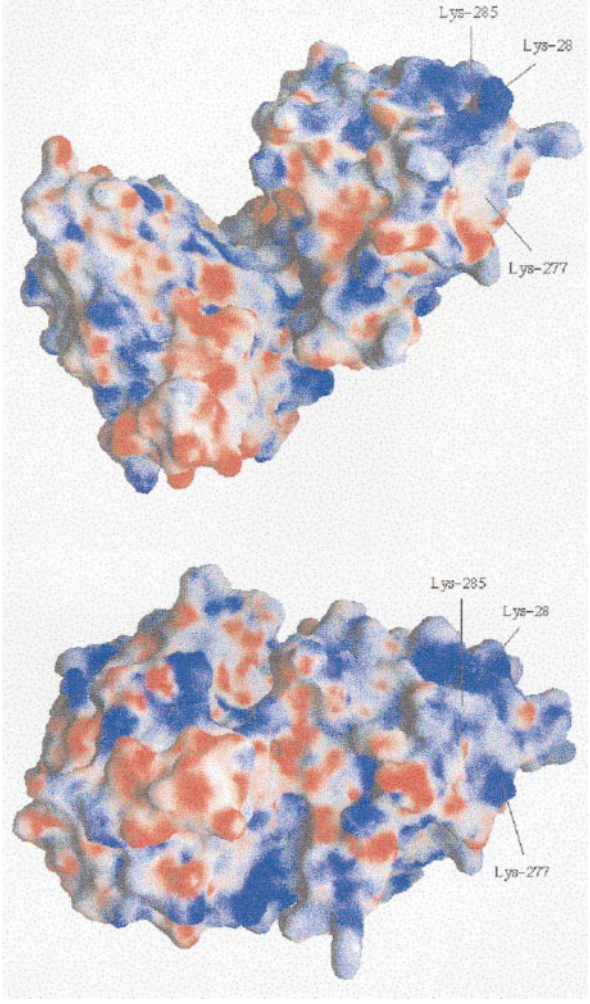


FIGURE 7 The change in surface shape and surface electrostatic potential in the region of Lys-285 associated with the conformational change accompanying iron and carbonate binding. Electrostatic potential is color coded using a sliding scale (≤ -10KT to $\geq +10$KT). Red represents negative electrostatic potential, blue represents positive electrostatic potential, and white is neutral.

states at each end of the conformational change does not, however, give a clear indication of the order of protonation that occurs along the path of the conformational change. The pK_a values are conformation dependent, and the conformational change is, of course, a dynamic process.

The calculation of pK_a values in proteins is a very challenging problem. The results of our test calculations on lysozyme show that the performance of the programs and parameters used here is comparable with that of those used by Antosiewicz et al. (1996) who used a similar method although with a different parameter set. Exclusion of the problematic residue Tyr-53 does suggest, however, that our results may be a little poorer. More accurate predictions have recently been achieved for lysozyme by Sham et al. (1997) and Demchuk and Wade (1996) using somewhat different approaches. Much of the work presented here represents preparation for subsequent molecular dynamics (MD) simulations involving the N-lobe of human lactoferrin, and as such, a compromise has been struck between accuracy and consistency between the continuum electrostatic and MD approaches.

TABLE 5 Residues in the N-terminal half-molecule of human apo-lactoferrin that are predicted to have pK_a values in the range 7.4-4.0 under physiological conditions

Residue	pK_a
Glu-51	4.0
Glu-66	4.4
His-91	4.0
Glu-108	4.7
Asp-217	6.3
Glu-221	4.6
Glu-223	6.6
Glu-226	4.5
His-253	6.7
Lys-277	7.1
Asp-297	4.1

Temperature = 310 K; ionic strength = 145 mM.

There are a number of potential sources of error in the calculation scheme presented here. Many extreme pK_a shifts are associated with the close proximity of charged residues, which raises the possibility that inaccuracies in the solution of the Poisson-Boltzmann equation at close range may contribute to the size of the shifts.

A significant problem with the calculation scheme used here is that a model structure of a protein is required that has all the titratable hydrogen atoms in place. At the pH of the crystal structure, such a degree of protonation is inappropriate (indeed, the low pH required to fully protonate all of the titratable residues in a protein would most likely cause chemical damage to and denaturation of the protein), and thus there will always be some difficulty fitting in all of the hydrogen atoms, which will be reflected in the calculated pK_a shifts. Demchuk and Wade (1996) have developed a method that circumvents this problem and suggests a future development of the work presented here. A consequent but smaller source of error in the method used here is that the surface of the protein is determined and fixed by the fully protonated model of the protein and does not change in shape when a surface residue is deprotonated.

Factors that are presently thought to be among the most important in improving the accuracy of predicted pK_a values in proteins are the inclusion of a detailed representation of the redistribution of electrostatic potential that occurs in a residue when its protonation state changes, the introduction of conformational flexibility into the model of the protein (You and Bashford, 1995; Zhou and Vijayakumar, 1997), and the inclusion of a term that estimates the entropy associated with first hydration shell solvent ordering (Warwicker, 1997). A detailed charge model is used in the work presented here. Antosiewicz et al. (1996) have shown that the use of the commonly assumed value of $\epsilon_p = 4$ works well with a detailed charge model, and this value was assumed in the main part of the work presented here.

Using higher values of ϵ_p improved the results of our test calculations on lysozyme. Different profiles of sensitivity of

calculations to ϵ_p are found for different proteins (Demchuk and Wade, 1996), and the N-terminal half-molecule of human lactoferrin is considerably larger than lysozyme, which makes it difficult to draw any conclusions about the most appropriate ϵ_p to use with this protein. Demchuk and Wade (1996) have also devised a method for predicting what level of ϵ_p is most suitable for each individual residue in a protein, and this approach may prove useful in future investigations of transferrins where experimentally determined pK_a values are not available.

The calculated net effects of pH in the range 4–9.5 on the N-terminal half-molecule of human lactoferrin, i.e., the titration curves and net gain or loss of protons, was insensitive to ϵ_p in the range 4–20. The protonation states of the individual residues that contribute to these net effects were, however, sensitive to ϵ_p, and so we cannot at this stage make very reliable predictions about individual shifts in pK_a values. Although it may be tempting to assume a higher value for ϵ_p, calculations will be less likely to identify those key residues in proteins that are found experimentally to have significantly shifted pK_a values, and so unique and interesting results may be overlooked in the drive to achieve an overall improvement in one's predictions.

Progress has been reported in the use of explicit solvent models for the prediction of pK_a shifts. Figuerido et al. (1996) calculated pK_a shifts for succinic acid by an explicit solvent MD simulation that were very close to experimental values. Accurate modeling of long-range effects using the Ewald summation method was found to be necessary to properly model the electrostatic shielding effect of the aqueous solvent. Although this approach would be highly computationally expensive for calculating pK_a shifts in proteins, Figuerido et al. suggest that comparisons between explicit and continuum solvent models can reveal differences that have their true physical origin in the inherent molecularity of the surrounding medium.

Baptista et al. (1997) have developed a method for incorporating pH effects into MD simulations with explicit solvent that involves periodically updating the mean protonation states of titrating residues according to a continuum electrostatic calculation. This has been implemented in the constant pH MD simulation of the small protein BPTI. It is recognized that such an approach would be very valuable in the study of transferrins, although application of the method to such a large protein may, at present, be prohibitively expensive computationally. As computers become faster it may be possible to further develop this method such that the mean protonation states are also calculated using the explicit solvent model.

We thank Professor P.F. Lindley (ESRF) for useful discussions.

The BBSRC provided an earmarked studentship to D.A. Lee, and computer resources were provided by The Wellcome Trust and by the BBSRC (EIIO5501).

REFERENCES

Aasa, R., B. G. Malmström, P. Saltman, and T. Vänngard. 1963. The specific binding of iron(III) and copper(II) to transferrin and conalbumin. *Biochim. Biophys. Acta.* 75:203–222.

Amos, R. D., I. L. Alberts, J. S. Andrews, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, K. E. Laidig, G. Laming, A. M. Lee, P. E. Maslen, C. W. Murray, J. E. Rice, E. D. Simandiras, A. J. Stone, M-D. Su, and D. J. Tozer. 1995. CADPAC: The Cambridge Analytical Derivatives Package Issue 6. Cambridge, UK.

Anderson, B. F., H. M. Baker, E. J. Dodson, G. E. Norris, S. V. Rumball, J. M. Waters, and E. N. Baker. 1987. Structure of human lactoferrin at 3.2-Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* 84:1769–1773.

Anderson, B. F., H. M. Baker, G. E. Norris, S. V. Rumball, and E. N. Baker. 1990. Apolactoferrin structure demonstrates ligand-induced conformational change in transferrins. *Nature (London).* 344:784–787.

Antosiewicz, J., J. M. Briggs, A. H. Elcock, M. K. Gilson, and J. A. McCammon. 1996. Computing ionization states of proteins with a detailed charge model. *J. Comp. Chem.* 17:1633–1644.

Baker, E. N. 1993. Transferrins from a structural perspective. *In* Perspectives in Bioinorganic Chemistry, Vol. 2. JAI Press, Greenwich, CT. 161–203.

Baptista, A. M., P. J. Martel, and S. B. Petersen. 1997. Simulation of protein conformational freedom as a function of pH: constant-pH molecular dynamics using implicit titration. *Proteins Struct. Funct. Genet.* 27:523–544.

Bartik, K., C. Redfield, and C. M. Dobson. 1994. Measurement of the individual pK_a values of acidic residues of hen and turkey lysozymes by two-dimensional ^{1}H NMR. *Biophys. J.* 66:1180–1184.

Bashford D. 1993. MEAD Version 1.1.3. The Scripps Research Institute, La Jolla, CA.

Bashford, D., and M. Karplus. 1990. pK_as of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry.* 29: 10219–10225.

Beroza, P., D. R. Fredkin, M. Y. Okamura, and G. Feher. 1991. Protonation of interacting residues in a protein by a Monte Carlo method: application to lysozyme and the photosynthetic reaction centre of *Rhodobacter sphaeroides*. *Proc. Natl. Acad. Sci. U.S.A.* 88:5804–5808.

Chung, T. D. Y., and K. N. Raymond. 1993. The role of conformational changes in iron binding and release. *J. Am. Chem. Soc.* 115:6765–6768.

Day, C. L., K. M. Stowell, E. N. Baker, and J. W. Tweedie. 1992. Studies of the N-terminal half of human lactoferrin produced from the cloned cDNA demonstrate that interlobe interactions modulate iron release. *J. Biol. Chem.* 267:13857–13862.

Demchuk, E., and R. C. Wade. 1996. Improving the continuum approach to calculating pK_as of ionizable groups in proteins. *J. Phys. Chem.* 100:17373–17387.

De Silva, D. M., C. C. Askwith, and J. Kaplan. 1996. Molecular mechanisms of iron uptake in eukaryotes. *Physiol. Rev.* 76:31–47.

Dewan, J. C., B. Mikami, M. Hirose, and J. C. Sacchettini. 1993. Structural evidence for a pH sensitive dilysine trigger in the hen ovotransferrin N-lobe: implications for transferrin iron release. *Biochemistry.* 32: 11963–11968.

El Hage Chahine, J-M., and R. Pakdaman. 1995. Transferrin, a mechanism for iron release. *Eur. J. Biochem.* 230:1102–1110.

Figuerido, F., G. S. Del Buono, and R. M. Levy. 1996. Prediction of pK_a shifts without truncation of electrostatic interactions: an explicit solvent calculation for succinic acid. *J. Phys. Chem.* 100:6389–6392.

Gelb, M. H., and D. C. Harris. 1980. Correlation of proton release and ultraviolet difference spectra associated with metal binding by transferrin. *Arch. Biochem. Biophys.* 200:93–98.

Gerstein, M., B. F. Anderson, G. E. Norris, E. N. Baker, A. M. Lesk, and C. Chothia. 1993. Domain closure in lactoferrin. Two hinges produce a see-saw motion between alternative close-packed interfaces. *J. Mol. Biol.* 234:357–372.

Grossman, J. G., M. Neu, R. W. Evans, P. F. Lindley, H. Appel, and S. S. Hasnain. 1993. Metal-induced conformational changes in transferrins. *J. Mol. Biol.* 229:585–590.

Grossman, J. G., M. Neu, E. Pantos, F. J. Schwab, R. W. Evans, E. Townes-Andrews, P. F. Lindley, H. Appel, W.-G. Thies, and S. S.

Hasnain. 1992. X-ray solution scattering reveals conformational changes upon iron uptake in lactoferrin, serum and ovo-transferrins. *J. Mol. Biol.* 225:811–819.

Kilar, F., and I. Simon. 1985. The effect of iron binding on the conformation of transferrin. *Biophys. J.* 48:799–802.

Klapper, I., R. Hagstrom, R. Fine, K. Sharp, and B. Honig. 1986. Focusing of electric fields in the active site of Cu-Zn, superoxide dismutase: effects of ionic strength and amino-acid modification. *Proteins Struct. Funct. Genet.* 1:47–59.

Kubal, G., P. J. Sadler, and A. Tucker. 1994. pH-induced structural change in human serum apotransferrin, pK_a values of histidine residues and N-terminal amino group determined by ^{1}H-NMR spectroscopy. *Eur. J. Biochem.* 220:781–787.

Mazurier, J., and G. Spik. 1980. Comparative study of the iron-binding properties of human transferrins. I. Complete and sequential iron saturation and desaturation of the lactotransferrin. *Biochim. Biophys. Acta.* 629:399–408.

Molecular Simulations Inc. 1986, 1992. The QUANTA Program, Waltham, MA. Quanta Release. 4.1.1. Version 95.0320.

Nicholls, A., R. Bharadwaj, and B. Honig. 1993. GRASP: graphical representation and analysis of surface properties. *Biophys. J.* 64:166.

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham III, D. M. Ferguson, G. L. Seibel, U. Chandra Singh, P. K. Weiner, and P. A. Kollman. 1995. AMBER 4.1. University of California, San Francisco, San Francisco.

Qian, Z. M., and P. L. Tang. 1995. Mechanisms of iron uptake by mammalian cells. *Biochim. Biophys. Acta.* 1269:205–214.

Rossenau-Motreff, M. Y. F., R. Soetewey, R. Lamote, and H. Peeters. 1971. Size and shape determination of apotransferrin and transferrin monomers. *Biopolymers.* 10:1039–1048.

Sham, Y. Y., T. C. Zhen, and A. Warshel. 1997. Consistent calculations of pK_a's of ionizable residues in proteins: semi-microscopic and microscopic approaches. *J. Phys. Chem. B.* 101:4458–4472.

Stone, A. J. 1990. ORIENT: A Program for Calculating Electrostatic Interactions between Molecules, Version 2. University of Cambridge, Cambridge, UK.

Vigh, R., L. Cser, F. Kilar, and I. Simon. 1989. Different segmental flexibility of human serum transferrin and lactoferrin. *Arch. Biochem. Biophys.* 275:181–184.

Warwicker, J. 1997. Improving pK_a calculations with consideration of hydration entropy. *Protein Eng.* 10:809–814.

Woodworth, R. C., N. D. Butcher, S. A. Brown, and A. Brown-Mason. 1987. ^{1}H NMR study of effects of synergistic anion and metal ion binding on pH titration of the histidinyl side-chain residues of the half-molecules of ovotransferrin. *Biochemistry.* 26:3115–3120.

You, T. J., and D. Bashford. 1995. Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. *Biophys. J.* 69:1721–1733.

Zacharias, M., B. A. Luty, M. E. Davis, and J. A. McCammon. 1992. Poisson-Boltzmann analysis of the λ repressor-operator interaction. *Biophys. J.* 63:1280–1285.

Zhou, H-X., and M. Vijayakumar. 1997. Modelling of protein conformational fluctuations in pK_a predictions. *J. Mol. Biol.* 267:1002–1011.