

Calculated pH-Dependent Population and Protonation of Carbon-Monoxo-Myoglobin Conformers

Björn Rabenstein and Ernst-Walter Knapp

Institut für Chemie, Fachbereich Biologie, Chemie, Pharmazie, Freie Universität Berlin, 14195 Berlin, Germany

ABSTRACT X-ray structures of carbonmonoxymyoglobin (MbCO) are available for different pH values. We used conventional electrostatic continuum methods to calculate the titration behavior of MbCO in the pH range from 3 to 7. For our calculations, we considered five different x-ray structures determined at pH values of 4, 5, and 6. We developed a Monte Carlo method to sample protonation states and conformations at the same time so that we could calculate the population of the considered MbCO structures at different pH values and the titration behavior of MbCO for an ensemble of conformers. To increase the sampling efficiency, we introduced parallel tempering in our Monte Carlo method. The calculated population probabilities show, as expected, that the x-ray structures determined at pH 4 are most populated at low pH, whereas the x-ray structure determined at pH 6 is most populated at high pH, and the population of the x-ray structures determined at pH 5 possesses a maximum at intermediate pH. The calculated titration behavior is in better agreement with experimental results compared to calculations using only a single conformation. The most striking feature of pH-dependent conformational changes in MbCO—the rotation of His-64 out of the CO binding pocket—is reproduced by our calculations and is correlated with a protonation of His-64, as proposed earlier.

INTRODUCTION

Myoglobin is one of the most widely studied proteins. Numerous x-ray structures with different myoglobin ligands and under different physical conditions were solved in the past. Here, we focus on a set of structures of carbonmonoxymyoglobin (MbCO), which are prepared at pH values of 4, 5, and 6 (Yang and Phillips, Jr., 1996). This set of structures shows that at low pH, His-64—an important residue for O₂ specificity (Springer et al., 1994)—swings out of the CO binding pocket. It is assumed that a protonation of His-64 goes along with this remarkable conformational change (Yang and Phillips, Jr., 1996; Johnson et al., 1996). In this work we study the protonation behavior of MbCO by the method of continuum electrostatics (Ullmann and Knapp, 1999), where first the electrostatic interactions are calculated by solving the Poisson-Boltzmann equation (Bashford and Karplus, 1990), and after that the large number of possible protonation states is sampled with a Monte Carlo method (Beroza et al., 1991). We have already successfully applied this method to investigate protonation and electron transfer processes in several other proteins (Rabenstein et al., 1998a,b, 2000; Vagedes et al., 2000).

The protonation behavior of myoglobin has been studied extensively by theoretical methods (Shire et al., 1975; Bashford et al., 1993; Yang and Honig, 1994; Sandberg and Edholm, 1999). However, in these studies conformational flexibility was not considered explicitly. To take into ac-

count the structural differences within the set of x-ray structures, the continuum electrostatics method must be extended to include an ensemble of multiple conformations (Ullmann and Knapp, 1999). There are numerous approaches to include multiple conformations in the calculations of protonation behavior (Buono et al., 1994; You and Bashford, 1995; Beroza and Case, 1996; Sham et al., 1997; Schaefer et al., 1997; Alexov and Gunner, 1997, 1999). However, these methods are not feasible here, because they account only for conformational changes of titratable groups, explicitly simulated water molecules, and/or hydrogens.

Here, a small number of protein structures is given with completely different sets of atomic coordinates. Hence, not only atoms of the backbone and non-titratable groups are displaced, but also the dielectric boundary is changed. In this work we present a new method to treat a given ensemble of a small number of arbitrarily different conformers of a protein. Please note that throughout this work the term “conformer” refers to a whole structure, not only to a single residue. This method is essentially a Monte Carlo version of the approach described by Rabenstein et al. (1998a) and Ullmann and Knapp (1999). Our aim is to show that our method generates consistent population probabilities of the different conformers and to calculate titration curves of the different titratable groups of MbCO—especially of His-64—in agreement with experimental data. Our results will also have implications for the assignment of spectroscopically detected taxonomic substates (Johnson et al., 1996).

MATERIALS AND METHODS

Coordinates and parameters

In our calculations we used x-ray structures of sperm whale MbCO at pH values 4, 5, and 6 (Yang and Phillips, Jr., 1996). The PDB identifiers and resolutions of these structures are 1spe and 2.0 Å for pH 4, 1vxc and 1.7

Received for publication 30 June 2000 and in final form 14 November 2000.

Address reprint requests to Dr. Ernst-Walter Knapp, Institut für Chemie, Freie Universität Berlin, Takustrasse 6, 14195 Berlin, Germany. Tel.: 49-30-83854387; Fax: 49-30-83853464; E-mail: knapp@chemie.fu-berlin.de.

© 2001 by the Biophysical Society

0006-3495/01/03/1141/10 \$2.00

Å for pH 5, and 1vxf and 1.7 Å for pH 6. For the residues 79–81, the structures 1spe and 1vxc contain two alternative conformations A and B, which we considered separately. Thus, we got five different structures called in the following 4a, 4b, 5a, 5b, and 6, with the digit corresponding to their respective pH value of preparation and the letter corresponding to the conformation of residues 79–81. We removed the sulfate ion and all water molecules from each structure. The influence of water was considered exclusively by a higher dielectric constant in cavities and outside of the protein because the orientation of the water molecules is not known, which makes their electrostatic effect uncertain (Rabenstein et al., 1998b). We considered as titratable groups the C-terminus, the N-terminus, the two propionates of the heme, and all aspartates, glutamates, arginines, lysines, tyrosines, and histidines, with the exception of His-93, which coordinates the heme iron and is therefore considered to be non-titrating. We used an all-atom model where all hydrogen atoms were treated explicitly with the exception of the hydrogen atom of protonated carboxyl groups. In this case the hydrogen atom was treated implicitly by distributing its charge symmetrically to the two oxygen atoms of the carboxyl group. The coordinates of hydrogen atoms were generated with CHARMM (Brooks et al., 1983). The positions of hydrogen atoms were energetically optimized using the CHARMM program with the CHARMM22 force field (MacKerell et al., 1992), while the heavy atom positions were fixed. For this optimization all titratable groups were in their standard protonation (i.e., aspartate, glutamate, heme-propionate, and the C-terminus unprotonated, arginine, histidine, lysine, tyrosine, and the N-terminus protonated). All atomic partial charges, including that of the heme, and the atomic radii were taken from the CHARMM22 force field (MacKerell et al., 1992). However, we needed to apply some modifications and endorsements: as mentioned above, the proton of protonated carboxyl groups is not represented explicitly, but simply by symmetrically distributing its charge to the atoms of the unprotonated carboxyl group (Table 1). We did this because we cannot know the exact binding site of the proton. We demonstrated in the past that this simplification works properly (Rabenstein et al., 1998a,b, 2000; Vagedes et al., 2000). A more detailed approach would be to introduce more than two states for a titratable group similar to the way it was done for histidines here (Bashford et al., 1993; Beroza and Case, 1996; Alexov and Gunner, 1997). For similar reasons, the deprotonation of Arg, Lys, and the N-terminus was represented by symmetrically removing a unit positive charge from the atoms of the titratable group (Table 1). For the charges of deprotonated tyrosine, which are not part of the CHARMM22 parameter set, we based our charges on quantum chemical calculations as described in

Rabenstein et al. (1998b). Atomic partial charges that are not explicitly part of the CHARMM22 parameter set are listed in Table 1.

Titration of single conformers

First, we calculated the protonation pattern of all five structures separately. The theoretical background of the calculation of protonation patterns by solving the Poisson-Boltzmann equation (Bashford and Karplus, 1990) and sampling the possible protonation states with a Monte Carlo (MC) method (Beroza et al., 1991) is reviewed by Ullmann and Knapp (1999). We give here only an outline of the explanations, representing the most essential parts to understand the new methods applied in this work.

The energy G^n of a protonation state n of a protein, which is characterized by the protonation state vector $\mathbf{x}_n = (x_1^n, x_2^n, \dots, x_N^n)$, is given by Eq. 1 (Bashford and Karplus, 1990, 1991; Beroza et al., 1991; Yang et al., 1993; Beroza and Fredkin, 1996; Antosiewicz et al., 1996; Ullmann and Knapp, 1999).

$$G^n = \sum_{\mu=1}^N \left((x_{\mu}^n - x_{\mu}^0) RT \ln 10 (\text{pH} - \text{pK}_{\text{intr},\mu}) + \sum_{\nu>\mu}^N W_{\mu\nu} (x_{\mu}^n - x_{\mu}^0) (x_{\nu}^n - x_{\nu}^0) \right) \quad (1)$$

Depending on the group μ , which can be protonated or unprotonated, x_{μ}^n adopts the values 1 or 0, respectively. \mathbf{x}_0 is the protonation state vector of the so-called reference state, where all titratable groups are in their uncharged protonation state, i.e., x_{μ}^0 is 1 if μ is an acid and 0 if μ is a base. The sums run over all N titratable groups. $\text{pK}_{\text{intr},\mu}$ is the so-called intrinsic pK_a value of the group μ , and $W_{\mu\nu}$ is the electrostatic interaction energy of the groups μ and ν if both are charged. pK_{intr} and the symmetrical W -matrix are both calculated by solving the Poisson-Boltzmann equation on a grid (Warwicker and Watson, 1982). For doing this, we used the program MEAD (Bashford and Gerwert, 1992; Bashford, 1997). In our calculation, the protein is represented by a cavity with the dielectric constant $\epsilon_p = 4$ containing the protein atoms with their atomic partial charges. In the remaining volume, the solvent is represented by setting the dielectric constant to $\epsilon_s = 78.5$ and the ionic strength to $I = 0.1$ mol/l. For a discussion of the value for ϵ_p , see Warshel and Russel (1984), Warshel and Åqvist (1991), Honig and Nicholls (1995), Warshel et al. (1997), Rabenstein et al. (1998a), and Ullmann and Knapp (1999). For the solution of the Poisson-Boltzmann equation, we performed grid focusing (Klapper et al., 1986) in two steps. Initially, we used an 80-Å cube with a 1.0 Å lattice spacing centered at the geometric center of the protein, followed by a 20-Å cube with a 0.25 Å lattice spacing centered at the considered titratable group. We used an ion exclusion layer of 2 Å and a solvent probe radius of 1.4 Å. The pK_a values used for the model compounds are listed in Table 2. We accounted for the δ and ϵ tautomers of histidine by considering histidines as double sites, as described by Bashford et al. (1993).

The protonation probability $\langle x_{\mu} \rangle$ of the group μ can be calculated by evaluating the thermodynamic average over all 2^N possible protonation states n given by Eq. 2,

$$\langle x_{\mu} \rangle = \frac{\sum_{n=1}^{2^N} x_{\mu}^n \exp(-G^n/RT)}{\sum_{n=1}^{2^N} \exp(-G^n/RT)} \quad (2)$$

This thermodynamic average cannot be calculated exactly, since the number of possible protonation states of myoglobin (2^N with $N = 73$) is far too large. Instead, we sampled the set of protonation states using a Metropolis MC method (Beroza et al., 1991). This MC method is implemented in our program KARLSBERG (Rabenstein, 1999), which is freely available under

TABLE 1 Atomic partial charges of titratable groups

Group	Protonated	Deprotonated
Carboxyl (except C-term.)		
C- α	-0.21	-0.28
carboxyl C	0.75	0.62
carboxyl O (2 \times)	-0.36	-0.76
C-terminus		
carboxyl C	0.34	0.34
carboxyl O (2 \times)	-0.17	-0.67
Arginine		
N- ϵ	-0.70	-0.81
C- ζ	0.64	0.71
guanidino N (2 \times)	-0.80	-0.90
guanidino H (4 \times)	0.46	0.27
Lysine/N-terminus		
amino N	-0.30	-0.97
amino H (3 \times)	0.33	0.22
Tyrosine		
C- ζ	0.11	-0.18
hydroxyl O	-0.54	-0.82
hydroxyl H	0.43	0.00

TABLE 2 Groups considered as titratable with the pK_a values of their model compounds

Titratable Group	Model Compound pK_a	Reference
Arginine	12.0	Nozaki and Tanford (1967)
Aspartate	4.0	Nozaki and Tanford (1967)
Glutamate	4.4	Nozaki and Tanford (1967)
δ -Histidine	7.0	Tanokura (1983)
ϵ -Histidine	6.6	Tanokura (1983)
Lysine	10.4	Nozaki and Tanford (1967)
Tyrosine	9.6	Nozaki and Tanford (1967)
C-terminus	3.8	Nozaki and Tanford (1967)
N-terminus	7.5	Nozaki and Tanford (1967)
Heme propionate	4.8	Moore and Pettigrew (1990)

the GNU public license from our webserver (<http://lie.chemie.fu-berlin.de/karlsberg/>). To increase sampling efficiency, pairs of titratable groups coupled by >3 pK_a units were treated by special double moves (Beroza et al., 1991), and triples of titratable groups, where at least two of the possible three pairs are coupled by >6 pK_a units, were treated by triple moves (Rabenstein et al., 1998a). The elementary MC step is the attempt to change the protonation state of a randomly chosen titratable group with subsequent evaluation of the Metropolis criterion. In the case of a double or triple move, the protonation state of all two or three titratable groups are changed before the Metropolis criterion is applied. An MC scan comprises as many MC steps as titratable single groups, doubles, and triples are available. We did one complete sampling for each 0.1 pH increment in the range from pH 3 to pH 7. For each sampling we performed 10,000 full scans, and after that 100,000 so-called reduced scans, where all titratable groups whose protonation probability during the full sampling differed from unity or zero by $<10^{-4}$ were fixed in their respective protonation state and excluded from further sampling. We calculated the statistical error of the protonation probability of a single group by evaluating its correlation function as described by Beroza et al. (1991). For each single group it was smaller than 0.003, and the statistical error of the total protonation of the whole myoglobin was always smaller than 0.05 protons. All calculations were done for $T = 300$ K.

Titration of a conformational ensemble

To account for conformational variability in computations of electrostatic energies, Eq. 1 must be modified, yielding Eq. 3, which is the energy $G^{n,l}$ of the protonation state n in a certain conformation l (Ullmann and Knapp, 1999).

$$G^{n,l} = \sum_{\mu=1}^N \left((x_{\mu}^n - x_{\mu}^0) RT \ln(10(pH - pK_{\text{intr},\mu}^l)) + \sum_{\nu>\mu}^N W_{\mu\nu}^l (x_{\mu}^n - x_{\mu}^0)(x_{\nu}^n - x_{\nu}^0) \right) + \Delta G_{\text{conf}}^l \quad (3)$$

The interaction parameters $pK_{\text{intr},\mu}^l$ and $W_{\mu\nu}^l$ depend now on the actual conformation l . The conformational reference energy $\Delta G_{\text{conf}}^l = G_{\text{conf}}^l - G_{\text{conf}}^r$ is the energy difference between an arbitrarily chosen, albeit fixed, reference conformation r and the actual conformation l . For the computation of ΔG_{conf}^l , the protein must be in its reference protonation state for both conformations r and l , i.e., all titratable groups are in their uncharged protonation state. If there are L conformations in the conformational ensemble, the total number of combined protonation and conformational states is $L \cdot 2^N$ (with N being the number of titratable groups). The

thermodynamic average in Eq. 2 must now be evaluated for all $L \cdot 2^N$ states, yielding Eq. 4 (Rabenstein et al., 1998a).

$$\langle x_{\mu} \rangle = \frac{\sum_{l=1}^L \sum_{n=1}^{2^N} x_{\mu}^n \exp(-G^{n,l}/RT)}{\sum_{l=1}^L \sum_{n=1}^{2^N} \exp(-G^{n,l}/RT)} \quad (4)$$

The occupation probability α_m of a certain conformation m within the conformational ensemble can also be calculated by a thermodynamic average:

$$\alpha_m = \frac{\sum_{n=1}^{2^N} \exp(-G^{n,m}/RT)}{\sum_{l=1}^L \sum_{n=1}^{2^N} \exp(-G^{n,l}/RT)} \quad (5)$$

Again, the thermodynamic average can only be calculated analytically for a small number of titratable groups and conformations as it was done by Rabenstein et al. (1998a). However, in this study the number of titratable groups is $N = 73$ and the number of conformations is $L = 5$, so that the number of states is $5 \cdot 2^{73}$. We also sampled this increased number of states by an MC method. We did this by introducing conformational MC moves in addition to the titration moves. However, in doing so, the conformational sampling proved to be not efficient enough. To overcome this problem, we applied parallel tempering (Geyer, 1991).

In parallel tempering the MC simulation is not performed for a single system, but for I noninteracting systems in parallel. Each system is a copy of the original single system with the only difference in temperature T . The systems are numbered from 1 to I in a way that $T_{i+1} > T_i$ (with $0 \leq i < I$). Titration and conformational moves are applied as before to each of the parallel systems independently. However, a third kind of move is introduced, a so-called tempering move. This move exchanges the protonation state \mathbf{x}_i and the conformation l_i of a randomly selected system i ($i < I$) with the protonation state \mathbf{x}_i and the conformation l_{i+1} of system $i + 1$. A tempering move is accepted with the probability

$$p = \min[1, \exp(R^{-1}(T_i^{-1} - T_{i+1}^{-1})(G_i - G_{i+1}))] \quad (6)$$

where G_i is the energy $G^{n,l}$ according to Eq. 3 for the current protonation state n and conformation l of system i . The parallel tempering procedure yields a canonical ensemble at the corresponding temperature for each parallel copy. Although the simulation of several ensembles is more costly, the tempering moves decrease the correlation within each ensemble, and hence decrease the statistical error.

In our MC simulations for the complete set of ensembles for all five myoglobin structures we applied parallel tempering at the temperatures 300 K, 400 K, 533 K, 711 K, and 948 K. For the evaluation of our results we considered only the simulation data at 300 K. We added one tempering move and 10 conformational moves per MC scan to the conventional titration moves. The simulations at higher temperatures are coupled to the 300 K simulation by the tempering moves such that the sampling efficiency of the 300 K simulation is increased as described before. For each sampling we performed 2000 full scans and 100,000 reduced scans, leading to a statistical error (at 300 K) of <0.02 for the protonation probability of single groups, of <0.1 protons for the total protonation of myoglobin, and of <0.02 for the occupancy of single conformations.

For the MC sampling of the conformational ensemble, reasonable values for ΔG_{conf}^l are required. These values can in principle be obtained from a molecular dynamics force field like CHARMM (Rabenstein et al., 1998a; Ullmann and Knapp, 1999). However, it is unlikely that this method will give reasonable results in our case. The structures used here are derived from crystallographic data and therefore inevitably involve small uncertainties in the atomic coordinates. Small structural changes within these uncertainties can dramatically change the energy value obtained from a force field. One could try to solve this problem by adjusting the structures according to the used force field, e.g., with a kind of energy minimization. It is, however, questionable in which way this should exactly be done

without introducing an uncontrolled bias of the results. Therefore, we decided to use another method, where a modification of the given x-ray structures was not necessary.

First, we determined the $\Delta G_{\text{conf}}^{\text{l}}$ values only within the two alternative conformations in the x-ray structure from pH 4 (4a and 4b) and in the x-ray structure from pH 5 (5a and 5b), i.e., we simulated two-member ensembles consisting of 4a and 4b or of 5a and 5b. The occupancy of the two conformations was determined by the crystallographers to be 0.78 for 4a, 0.22 for 4b, 0.76 for 5a, and 0.24 for 5b. Our aim was to find $\Delta G_{\text{conf}}^{\text{l}}$ values that reproduce these occupancies, denoted as $\alpha_{\text{target}}^{\text{l}}$, in the multiconformational MC sampling. We achieved this by an iterative method. We first did MC simulations (one for 4a/4b, one for 5a/5b) where the initial value of $\Delta G_{\text{conf}}^{\text{l}}$ was set to zero. The pH values for these simulations were the same as those for the preparation of the x-ray structures, i.e., pH 4 for 4a/4b, pH 5 for 5a/5b. From the calculated occupancies $\alpha_{\text{calc}}^{\text{l}}$ of the conformations, a correction energy $\Delta G_{\text{corr}}^{\text{l}}$ was derived by Eq. 7, which was added to $\Delta G_{\text{conf}}^{\text{l}}$.

$$\Delta G_{\text{corr}}^{\text{l}} = RT \ln \frac{\alpha_{\text{calc}}^{\text{l}}}{\alpha_{\text{target}}^{\text{l}}} \quad (7)$$

If our initial MC sampling were not afflicted with a statistical error, the corrected values for $\Delta G_{\text{conf}}^{\text{l}}$ would be those we were looking for. However, there is a statistical error of $\alpha_{\text{calc}}^{\text{l}}$, and the resulting error of $\Delta G_{\text{corr}}^{\text{l}}$ is especially large if $\alpha_{\text{calc}}^{\text{l}}$ is near to zero or unity, so we iteratively repeated the MC simulations with the corrected $\Delta G_{\text{conf}}^{\text{l}}$ values to get new values for $\Delta G_{\text{corr}}^{\text{l}}$. We did this until the root-mean-square deviation (rmsd) of the calculated occupancies $\alpha_{\text{calc}}^{\text{l}}$ compared to the target occupancies $\alpha_{\text{target}}^{\text{l}}$ was smaller than 0.01, which was fulfilled after one to three iterations, depending on the calculation. For convenience, we shifted the $\Delta G_{\text{conf}}^{\text{l}}$ values additively after each iteration so that the values for 4a and 5a were always zero. As $\Delta G_{\text{conf}}^{\text{l}}$ values, we obtained 6.78 kJ/mol for 4b and -0.11 kJ/mol for 5b.

In the next step, we determined $\Delta G_{\text{conf}}^{\text{l}}$ values for the sampling of all conformations together. Again we applied an iterative method as before. The problem is that values for the target occupancies are not directly available from experiment. Therefore, we assumed that, averaged over the pH range from 3 to 7, all three groups of structures, 4a/4b, 5a/5b, and 6, represent the complete ensemble of the infinite number of possible myoglobin conformations equally well. Hence, their occupancy integrated over this pH range should be equal (1/3). With this assumption, we could apply the iterative method as before with only a few modifications: the occupancy of a conformation $\alpha_{\text{calc}}^{\text{l}}$ is now calculated by integrating the calculated occupancy at a certain pH $\alpha_{\text{pH}}^{\text{l}}$ over the pH range from 3 to 7:

$$\alpha_{\text{calc}}^{\text{l}} = \int_{\text{pH}3}^{\text{pH}7} \alpha_{\text{pH}}^{\text{l}} d\text{pH} \quad (8)$$

For this calculation, the occupancies of 4a and 4b (or 5a and 5b, respectively) were always added together. Their relative $\Delta G_{\text{conf}}^{\text{l}}$ values were fixed to the values given above. The initial guess for the $\Delta G_{\text{conf}}^{\text{l}}$ values was -100 kJ/mol for 4a and -5 kJ/mol for 5a. The value for structure 6 was fixed to be zero. The resulting $\Delta G_{\text{conf}}^{\text{l}}$ values were -98.29 kJ/mol for 4a, -91.51 kJ/mol for 4b, -2.05 kJ/mol for 5a, -2.16 kJ/mol for 5b, and 0.00 kJ/mol for 6.

Please note that the physical meaning of these energy values is rather limited. They result merely from the postulation that all three groups of structures, 4a/4b, 5a/5b, and 6, are populated equally averaged over the considered pH range, and that the population ratio of the structures 4a/4b or 5a/5b, respectively, is given by the occupancies from the x-ray refinement data.

RESULTS AND DISCUSSION

Population of the conformers

The calculated population probability of the conformers 4a, 4b, 5a, 5b, and 6 is plotted over the pH range from 3 to 7 in Fig. 2. Indeed, conformer 6 has its maximum population at the highest pH value and conformers 4a and 4b are most populated at the lowest pH value. Conformer 5a and 5b show a maximum of their population probability at intermediate pH values. For 5b, this maximum is located as expected at about pH 5, whereas for 5a it is shifted to a higher pH of ~6. This difference is remarkable because the two conformers differ only at the residues Lys-79, Gly-80, and His-81, and the rmsd is low (see Table 3). However, the shapes of the population curves of the corresponding conformer pair 4a and 4b differ less. Also, the differences between 4a/4b and the other conformers is not surprising because there is a larger rmsd of ~1.1 Å and the extreme displacement of the His-64 side chain. The conformational differences between 5a/5b and 6 are less pronounced, but nevertheless the shapes of their population curves are quite different.

In summary, we conclude that our method was able to calculate population probabilities of the MbCO conformers that are qualitatively correct and consistent with the experimental conditions of the x-ray structure determination. These population probabilities are sensitive to even small conformational changes.

Titration behavior of individual sites

Depending on the titratable site, the titration curves are more or less similar for the different conformations. For some sites, however, there are extreme differences. This is the case for His-64, which features the swing out of the CO binding pocket at low pH values. Fig. 3 shows the protonation probability of His-64 for different conformers. For the conformers 5a, 5b, and 6, His-64 is nearly unprotonated at all pH values. For the conformers 4a and 4b, where His-64 has swung out of the CO binding pocket, it is partially protonated, up to 80% at low pH values, but even at high pH values it is nearly 50%. If all conformers are considered together, the protonation of His-64 is closely correlated to the population probability of conformers 4a and 4b (Fig. 2) and thus with the rotation of His-64. This is in agreement with experimental findings (Yang and Phillips, Jr., 1996). The δ -proton of the protonated His-64 in the 4a/4b conformation is in hydrogen bond distance (1.8 Å) to the backbone oxygen of Asp-60. The distance to the carboxy group of Asp-60, which is according to our results negatively charged at all pH values, is 3.7 Å. The hydrogen bond and the electrostatic attraction between the positively charged His-64 and the negatively charged Asp-60 may be responsible for pulling the protonated His-64 out of the

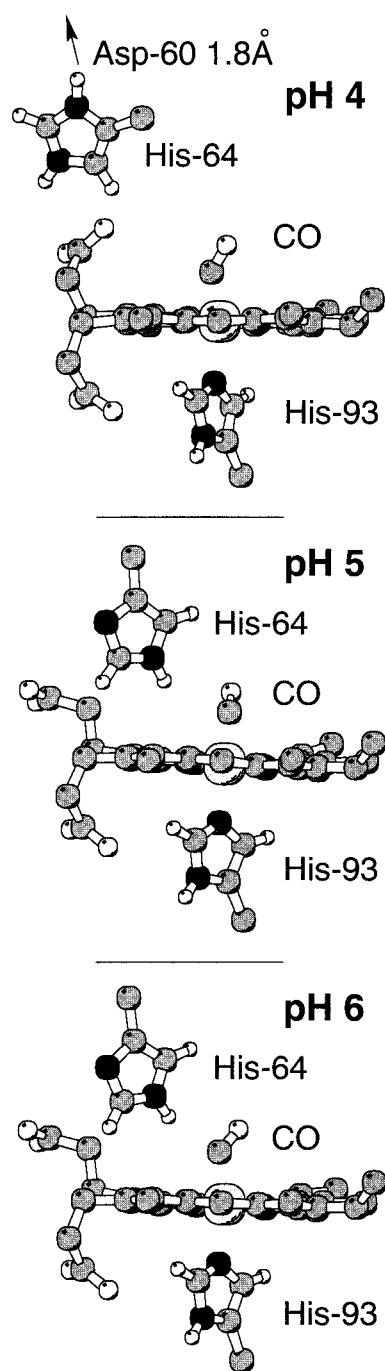


FIGURE 1 Heme, CO ligand, and side chains of His-64 and His-93 from the three x-ray structures at pH 4, 5, and 6. Hydrogen atoms of the heme are not drawn for the sake of clarity. The main feature of the structural changes is the rotation of His-64 out of the CO binding pocket at low pH. According to our results, protonation of His-64 goes along with this rotation. The unprotonated His-64 prefers the ϵ tautomeric form (drawn with the MOLSCRIPT program (Kraulis, 1991)).

binding pocket. In the conformers 5a, 5b, and 6, His-64 is not involved in any strong hydrogen bond with the protein moiety or the heme. (However, hydrogen-bonding to water

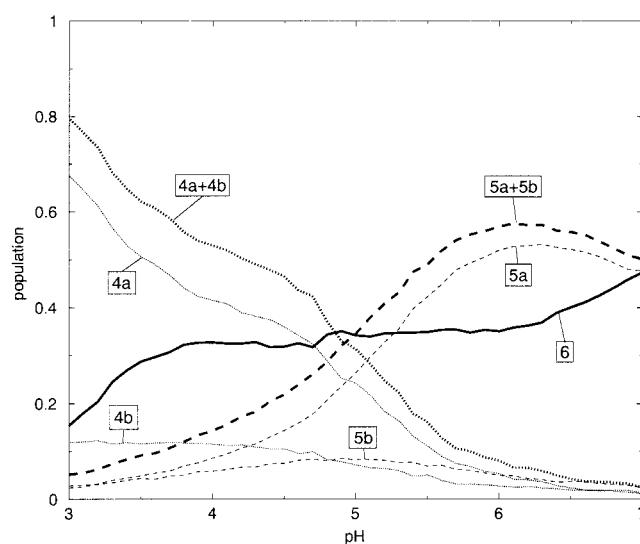


FIGURE 2 Calculated population probabilities of the different conformers in the pH range from 3 to 7. The number of each conformer corresponds to the pH value at which the respective x-ray structure was prepared.

molecules in cavities of the protein is still possible. In the x-ray structure, water molecule 35 (structure 6) or water molecule 36 (structure 5a/b) is in hydrogen bond distance (3.1 Å or 2.7 Å, respectively) to N_{δ} of His-64). In our calculations, we also determined the population probabilities for the two tautomers of histidine. Our result for His-64 is shown in Fig. 4. The unprotonated His-64 prefers the ϵ -tautomer, which is also reflected in Fig. 1.

The titratable residues Lys-79 and His-81 are directly involved in the conformational difference between 5a and 5b, or 4a and 4b, respectively. Their titration behavior may be a key to understanding the different shapes of the 5a and 5b population curves in Fig. 2. Lys-79 is completely protonated for all considered pH values and conformers. The titration behavior of His-81 is shown in Fig. 5. Although the titration curves for the conformers 4a and 4b are similar, the curves for 5a and 5b are much more different: the 5b curve resembles more the curves for 4a and 4b, whereas the 5a curve is similar to the curve for conformer 6. There are extreme differences between the a and b conformations in

TABLE 3 Root-mean-square deviations (rmsd) in angstroms for the five x-ray structures of MbCO considered in this work. Each pair of compared structures was aligned using the algorithm of Kabsch (1976)

	4a	4b	5a	5b	6
4a	0.00	0.20	1.09	1.07	1.07
4b	0.20	0.00	1.09	1.08	1.06
5a	1.09	1.09	0.00	0.16	0.63
5b	1.07	1.08	0.16	0.00	0.65
6	1.07	1.06	0.63	0.65	0.00

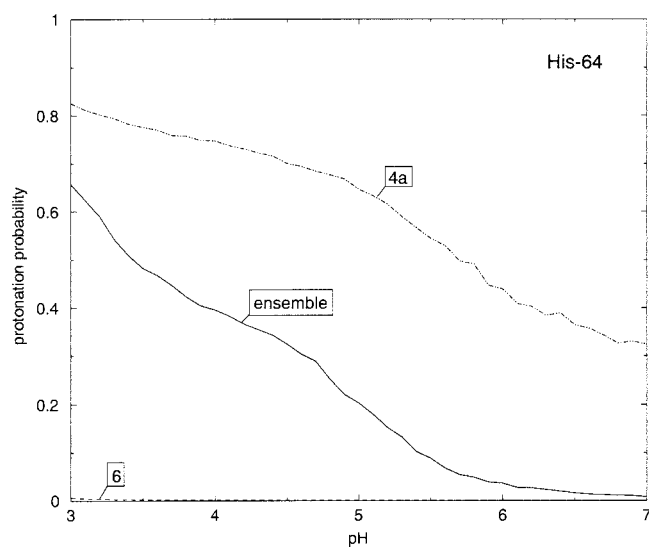


FIGURE 3 Protonation probability of His-64 for the single conformers 4a and 6, and for the complete ensemble of conformers ("ensemble"). The protonation probabilities for the other conformers are not shown because they are very similar to those of conformer 4a (4b) or 6 (5a and 5b). His-64 is protonated to a significant amount only in conformers 4a and 4b. The protonation of His-64 for the whole conformational ensemble reflects the increasing population probability of the conformers 4a and 4b with decreasing pH (see also Fig. 2).

both cases (4a/4b and 5a/5b) for the neighboring residue His-82 (Fig. 6). The titration curve of 4a is similar to that of 5b, and the titration curve of 4b is similar to that of 5a, so the curves are swapped between the more populated a conformations and the less populated b conformations.

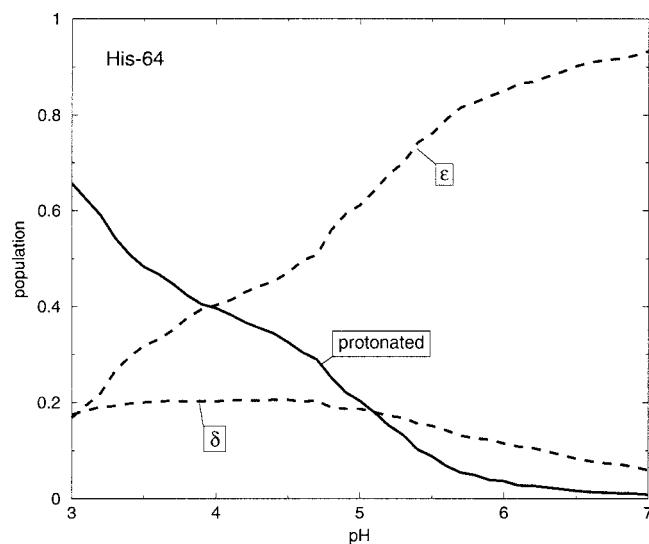


FIGURE 4 Population probability of the tautomers of His-64, considering the complete ensemble of conformers. The curve for the protonation probability is identical to the "ensemble" curve in Fig. 3.

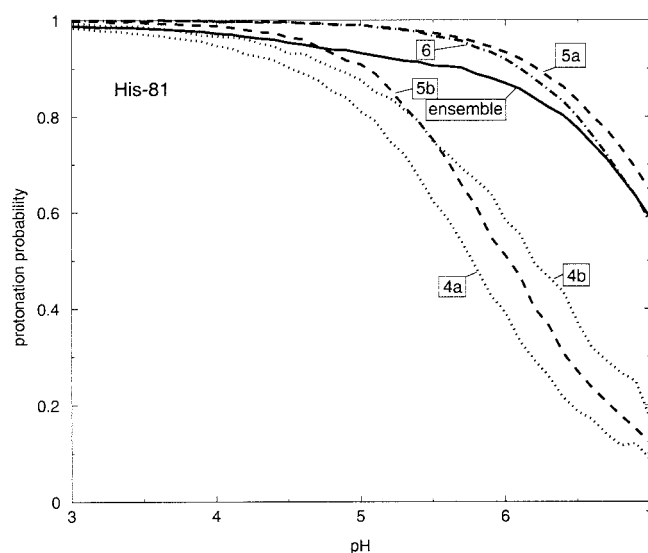


FIGURE 5 Protonation probability of His-81 for the different conformers and for the complete ensemble of conformers ("ensemble").

The strongly coupled pair His-24/His-119

As discussed by Bashford et al. (1993), the ϵ nitrogen atoms of His-24 and His-119 are in hydrogen-bond distance and the two residues constitute a strongly coupled pair of titratable groups. During titration, they effectively share their ϵ -proton. Protonation of both residues at the same time nearly never occurs. In the unprotonated state, His-24 nearly completely adopts the δ -tautomeric form. The ϵ -tautomer of His-24 is nearly unpopulated. If and only if His-24 is protonated, His-119 adopts the δ tautomeric form, so only three states are practically occurring: $p\delta$, δp , and $\delta\epsilon$ (Fig. 7).

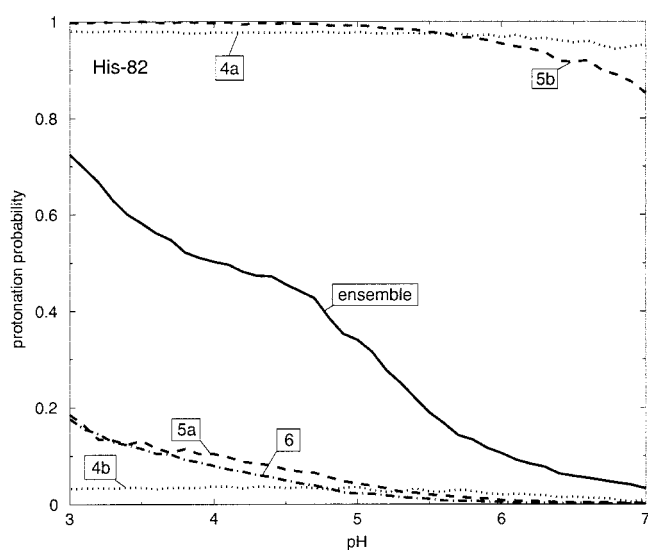


FIGURE 6 Protonation probability of His-82 for the different conformers and for the complete ensemble of conformers ("ensemble").

A “p” means protonated, “ δ ” represents the δ -tautomer, and “ ϵ ” the ϵ -tautomer. The first letter denotes the protonation state of His-24, the second that of His-119. The states p δ and δ p are equivalent if the ϵ -proton is considered as shared. This behavior is illustrated in Fig. 8.

Comparison of calculated pK_a values with experimental values and other theoretical studies

There are experimentally determined pK_a values for most of the histidines and tyrosines in MbCO. The values for Tyr-103 and Tyr-151 were determined by Wilbur and Allerhand (1976) to be 10.3 and 10.5, respectively. These values are out of the pH range we considered here (pH 3–7). According to our calculations, both tyrosine residues are protonated over the whole considered pH range. This is in agreement with the experimental findings.

More interesting is the comparison of the histidine values. They are listed in Table 4. Our calculated values are compared to experimental values and calculated values from other models. In the null model, the pK_a values are simply set to the pK_a values of the model compounds in aqueous solution. Here, the δ or ϵ tautomer is chosen according to experimental results. In the uniform-80 model (Sandberg and Edholm, 1999), an electrostatic calculation is done with the dielectric constant uniformly set to $\epsilon = 80$ everywhere. SCPB denotes a single-conformer Poisson-Boltzmann calculation (Bashford et al., 1993). DaPDS (distance and position dependent screening) is a new method to calculate protonation states in proteins (Sandberg and Edholm, 1999). It is based on empirical screening functions (Warshel et al., 1984) and is much simpler and thus faster than methods where the Poisson-Boltzmann equation must be solved. For each model we calculated the rmsd compared to the experimental values. In the calculation of the rmsd we included only histidine residues for which an experimental pK_a value in the range from 3 to 7 could be determined. In addition, we omitted the strongly coupled pair His-24 and His-119 due to their special titration behavior (s.a.), which does not allow assigning a well-defined pK_a value for each of these residues. That is also reflected in experimental findings (Bashford et al., 1993). According to our results, His-12 and His-81 have pK_a values >7 , which is out of the range of our study. However, the point of half protonation is nearly reached at pH 7, so that we could extrapolate a pK_a value for these residues and include them in the rmsd calculation. The lowest rmsd value for the non-trivial models is reached with our method (Table 4). The low rmsd value of the uniform-80 model is not surprising because all the experimental pK_a values are not far from the model compound pK_a values, and the uniform-80 model trivially provides pK_a values close to the model compound values. Especially interesting are the pK_a values for His-64, which performs the swing out of the binding pocket. The SCPB calculation (Bashford et al., 1993), which used an x-ray

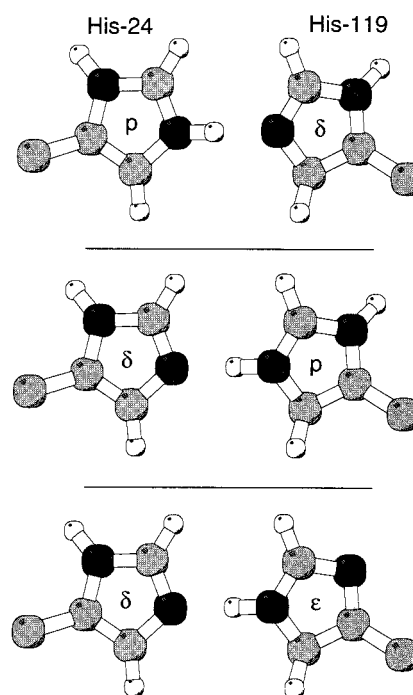


FIGURE 7 Structural representation of the strongly coupled pair His-24/His-119 with the three occurring protonation states p δ , δ p, and $\delta\epsilon$ explained in text (drawn with the MOLSCRIPT program (Kraulis, 1991)).

structure prepared at intermediate pH (Kuriyan et al., 1986), where His-64 is situated inside the CO binding pocket, shows a very low pK_a value, as this is also implied by our calculation if only the conformers 6, 5a, or 5b are considered (Fig. 3). If we consider the ensemble of all five conformers, the calculated pK_a value of His-64 is in reasonable

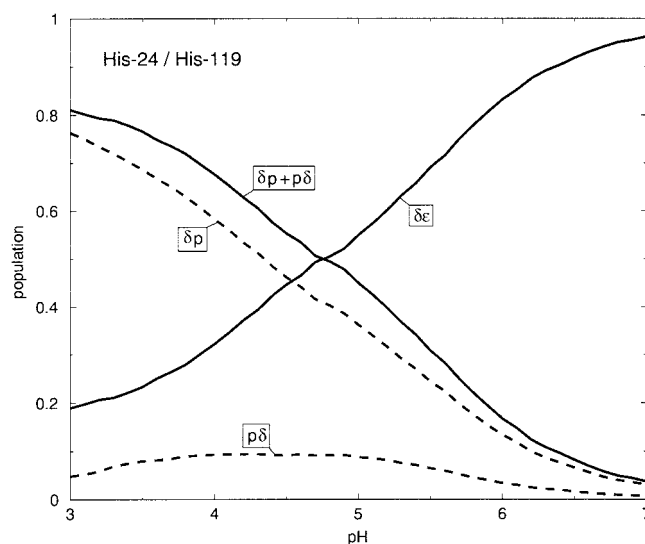


FIGURE 8 Population probabilities of the three possible states of the strongly coupled pair His-24/His-119 (see text). Because p δ and δ p are considered to be equivalent, their sum is also shown.

TABLE 4 Experimental and calculated pK_a values of histidines

Residue	Experimental	Calculated				
		Null*	Uniform 80**	SCPB††	DaPDS‡‡	This Work
<i>His-12</i>	6.3 [‡]	6.6	5.78	5.57	6.54	>7 (≈ 7.2) ^{§§}
<i>His-24</i> [†]	—	6.6	6.05	—	7.46	—
<i>His-36</i>	8.0 [‡]	7.0	6.89	6.24	7.27	>7
<i>His-48</i>	5.3 [‡]	6.6	6.16	4.79	6.74	5.09
<i>His-64</i>	<5 [‡] /4.6 [§] /4.5 [¶]	6.6	5.63	-2.90	6.30	3.43
<i>His-81</i>	6.6 [‡]	7.0	6.81	6.87	7.21	>7 (≈ 7.1) ^{§§}
<i>His-82</i>	<5 [‡]	6.6	5.82	-2.29	6.60	4.04
<i>His-97</i>	5.6 [‡] /5.9 [¶]	6.6	6.18	6.73	7.05	6.32
<i>His-113</i>	5.4 [‡]	6.6	4.96	4.26	6.08	4.36
<i>His-116</i>	6.5 [‡]	6.6	6.36	6.22	6.89	6.28
<i>His-119</i> [†]	6.1 [‡]	6.6	5.39	3.34	5.74	4.35
rmsd	0.0	1.09	0.61	2.87	1.05	0.71

* Unmodified pK_a values of model compound, the tautomer (δ or ϵ) was chosen according to experimental results.

[†] His-24 and His-119 are strongly coupled, see text.

[‡] Bashford et al. (1993).

[§] Fuchsman and Appleby (1979).

[¶] Müller et al. (1999).

^{||} Compared to most recent experimental values and including only residues printed in italics.

** Continuum electrostatics with $\epsilon = 80$ everywhere (Sandberg and Edholm, 1999).

†† Poisson-Boltzmann calculation similar to this work but with a single conformer (Bashford et al., 1993).

‡‡ Distance and Position Dependent Screening, a simple electrostatic model (Sandberg and Edholm, 1999).

§§ Extrapolated.

agreement with the experimental result. The DaPDS calculation (Sandberg and Edholm, 1999) yields a high pK_a value for His-64 of 6.30, although it uses the same x-ray structure as the SCPB calculation, where His-64 is situated inside the CO binding pocket and should be mostly unprotonated. This constitutes an example where simple methods like the uniform-80 model or empirical models like DaPDS yield, on the average, results in reasonable agreement with experimental findings, but do not provide an understanding of the underlying molecular mechanism. The calculation of pK_a values by solving the Poisson-Boltzmann equation may sometimes run into difficulties and yield deviating results, but is capable of unveiling what is going on in the more special situations, which may be crucial to understand protein function.

Bashford et al. (1993) reported a strong dependency of the Poisson-Boltzmann results on the used parameter set, i.e., atomic partial charges and atomic radii. They showed that choosing an appropriate parameter set can significantly improve the reliability of the results. The CHARMM22 parameter set, which was applied here, was not specifically designed for continuum electrostatics calculation, but rather for molecular dynamics at $\epsilon = 1$ and with an explicit solvent (if any). Possibly our results could be improved even further if the parameter set were tuned for continuum electrostatics calculations. However, our results presented here and also in another work (Vagedes et al., 2000) suggest that the CHARMM22 parameter set is already quite well suited for continuum electrostatics calculations.

Implications for the assignment of taxonomic substates

Three taxonomic substates of MbCO can be distinguished by infrared spectroscopy: A_0 , A_1 , and A_3 . A_0 is more populated at low pH, whereas A_1 and A_3 are present at higher pH (Johnson et al., 1996). In consistency with our own results, the A_0 substate is usually assigned to the protonated His-64 rotated out of the CO binding pocket (for further discussion see Johnson et al. (1996), Müller et al. (1999), and references cited therein).

On the basis of this assignment, Müller et al. (1999) recently derived the pK_a values of His-64 and His-97 from their FTIR measurements of the taxonomic substates. These pK_a values are in good agreement with our calculated values (see Table 4).

In general, A_1 is the dominant substate at pH 5.7 in solution (Johnson et al., 1996). Ray et al. (1994) assigned A_1 to the ϵ -tautomer of His-64, and A_3 to the δ -tautomer. Jewsbury and Kitagawa (1994) criticized this assignment based on their MD simulations and suggested it is the other way around, i.e., A_1 is the δ -tautomer and A_3 is the ϵ -tautomer. Our results, however, support the model of Ray et al. (1994) because the ϵ -tautomer is preferred at higher pH (Fig. 4). However, it is also possible that the difference between A_1 and A_3 is not the protonation state of His-64, but a hydrogen bond to a solvent molecule. This is corroborated by the fact that the interconversion between A_1 and A_3 is very fast at room temperature, which cannot be explained by

a simple tautomeric conversion. For more details see Müller et al. (1999).

CONCLUSIONS

We calculated the titration behavior of MbCO in the pH range from 3 to 7 by conventional electrostatic continuum methods with subsequent MC sampling of the protonation states. However, we considered not only one conformer, but an ensemble of five conformers derived from x-ray structures. These structures were solved for pH values 4, 5, and 6, so that the conformational ensemble used in our calculations represents the conformational changes of MbCO induced by pH changes. We extended the MC method by introducing conformational moves to sample the conformational ensemble in addition to the ensemble of protonation states. In our extended MC method the conformational ensemble may consist of a limited number of arbitrarily different conformations. The efficiency of the sampling was successfully increased by the application of parallel tempering (Geyer, 1991). As a result, we obtained population probabilities for the five conformers as a function of pH value. According to our results, the conformers corresponding to the x-ray data determined at pH 4 are most populated at low pH, whereas the conformer corresponding to pH 6 is most populated at high pH. The population probabilities of the conformers corresponding to pH 5 reach their maximum values at intermediate pH. Hence, our method was able to yield pH-dependent population probabilities of the conformers that are in qualitative agreement with experimental findings. The pK_a values calculated by sampling the conformational ensemble are in better agreement with experimental values than values calculated by a single-conformer Poisson-Boltzmann calculation (Bashford et al., 1993) and by the recently published DaPDS model (Sandberg and Edholm, 1999). We calculated the titration behavior of His-64, which features a remarkable rotation out of the CO binding pocket at low pH value, in agreement with experimental findings. Our results support the assignment of the taxonomic substates A_0 to the protonated rotated His-64, A_1 to the ϵ -tautomer, and A_3 to the δ -tautomer.

We are grateful to Dr. Donald Bashford and Dr. Martin Karplus for providing the programs MEAD and CHARMM, respectively. We also thank Dr. Matthias Ullmann and Dr. Ulrich Nienhaus for helpful discussions. This work was supported by the Deutsche Forschungsgemeinschaft Sfb 498, Project A5.

REFERENCES

- Alexov, E. G., and M. R. Gunner. 1997. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys. J.* 74:2075–2093.
- Alexov, E. G., and M. R. Gunner. 1999. Calculated protein and proton motions coupled to electron transfer: electron transfer from Q_A^- to Q_B in bacterial photosynthetic reaction centers. *Biochemistry*. 38:8253–8270.
- 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. Comput. Chem.* 17:1633–1644.
- Bashford, D. 1997. An object-oriented programming suite for electrostatic effects in biological molecules. In *Scientific Computing in Object-Oriented Parallel Environments*, Vol. 1343 of Lecture Notes in Computer Science. Y. Ishikawa, R. R. Oldehoeft, J. V. W. Reynders, and M. Tholburn, editors. Springer, Berlin. 233–240.
- Bashford, D., D. A. Case, C. Dalvit, L. Tennant, and P. E. Wright. 1993. Electrostatic calculations of side-chain pK_a values in myoglobin and comparison with NMR data for histidines. *Biochemistry*. 32:8045–8056.
- Bashford, D., and K. Gerwert. 1992. Electrostatic calculations of the pK_a values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* 224:473–486.
- Bashford, D., and M. Karplus. 1990. pK_a s of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*. 29:10219–10225.
- Bashford, D., and M. Karplus. 1991. Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculations. *J. Phys. Chem.* 95:9557–9561.
- Beroza, P., and D. A. Case. 1996. Including side chain flexibility in continuum calculations of protein titration. *J. Phys. Chem.* 100:20156–20163.
- Beroza, P., and D. R. Fredkin. 1996. Calculation of amino acid pK_a s in a protein from a continuum electrostatic model: method and sensitivity analysis. *J. Comput. Chem.* 17:1229–1244.
- 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 center. *Proc. Natl. Acad. Sci. U.S.A.* 88:5804–5808.
- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculation. *J. Comput. Chem.* 4:187–217.
- Buono, G. S. D., F. E. Figueirisco, and R. Levy. 1994. Intrinsic pK_a s of ionizable residues in proteins: an explicit solvent calculation for lysozyme. *Proteins: Struct., Funct., Genet.* 20:85–97.
- Fuchsman, W. H., and C. A. Appleby. 1979. CO and O₂ complexes of soybean leghemoglobins: pH effects upon infrared and visible spectra. Comparison with CO and O₂ complexes of myoglobin and hemoglobin. *Biochemistry*. 18:1309–1321.
- Geyer, C. J. 1991. *Computing Science and Statistics*. American Statistical Association, New York. 156–163.
- Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science*. 268:1144–1149.
- Jewsbury, P., and T. Kitagawa. 1994. The distal residue-CO interaction in carbonmonoxy myoglobins: a molecular dynamics study of the two distal histidine tautomers. *Biophys. J.* 67:2236–2250.
- Johnson, J. B., D. C. Lamb, H. Frauenfelder, J. D. Müller, B. McMahon, G. U. Nienhaus, and R. D. Young. 1996. Ligand binding to heme proteins. VI. Interconversion of taxonomic substates in carbonmonoxy-myoglobin. *Biophys. J.* 71:1563–1573.
- Kabsch, W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A*. 32:922–923.
- Klapper, I., R. Fine, K. A. Sharp, and B. H. 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.
- Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.
- Kuriyan, J., S. Wilz, M. Karplus, and G. A. Petsko. 1986. X-ray structure and refinement of carbon-monooxy(Fe II)-myoglobin at 1.5 Å resolutions. *J. Mol. Biol.* 192:133–154.
- MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnik, T. Ngo, D. T. Nguyen, B. Prodhom, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, J. Wiorkiewicz-Kuczera, and M. Karplus. 1992. Self-consistent parameterization

- of biomolecules for molecular modeling and condensed phase simulation. *FASEB J.* 6:143a. (Abstr.).
- Moore, G. R., and G. W. Pettigrew. 1990. Cytochromes c: Evolutionary, Structural and Physicochemical Aspects. Springer, Berlin.
- Müller, J. D., B. H. McMahon, E. Y. T. Chien, S. G. Sligar, and G. U. Nienhaus. 1999. Connection between the taxonomic substates and protonation of histidines 64 and 97 in carbonmonoxy myoglobin. *Biophys. J.* 77:1036–1051.
- Nozaki, Y., and C. Tanford. 1967. Examination of titration behavior. *Methods Enzymol.* 11:715–734.
- Rabenstein, B. 1999. Karlsberg Online Manual. <http://lie.chemie.fu-berlin.de/karlsberg/>.
- Rabenstein, B., G. M. Ullmann, and E. W. Knapp. 1998a. Calculation of protonation patterns in proteins with structural relaxation and molecular ensembles—application to the photosynthetic reaction center. *Eur. Biophys. J.* 27:626–637.
- Rabenstein, B., G. M. Ullmann, and E. W. Knapp. 1998b. Energetics of electron transfer and protonation reactions of the quinones in the photosynthetic reaction center of *Rhodospseudomonas viridis*. *Biochemistry*. 37:2488–2495.
- Rabenstein, B., G. M. Ullmann, and E. W. Knapp. 2000. Electron transfer between the quinones in the photosynthetic reaction center and its coupling to conformational changes. *Biochemistry*. 39:10487–10496.
- Ray, G. B., X.-Y. Li, J. A. Ibers, J. L. Sessler, and T. G. Spiro. 1994. How far can proteins bend the FeCO unit? Distal polar effects in heme proteins and models. *J. Am. Chem. Soc.* 116:162–176.
- Sandberg, L., and O. Edholm. 1999. A fast and simple method to calculate protonation states in proteins. *Proteins: Struct., Funct., Genet.* 36:474–483.
- Schaefer, M., M. Sommer, and M. Karplus. 1997. pH-Dependence of protein stability: absolute electrostatic free energy differences between conformations. *J. Phys. Chem. B.* 101:1663–1683.
- Sham, Y. Y., Z. T. Chu, and A. Warshel. 1997. Consistent calculation of pK_a 's of ionizable residues in proteins: semi-microscopic and microscopic approaches. *J. Phys. Chem. B.* 101:4458–4472.
- Shire, S. J., G. I. H. Hanania, and F. R. N. Gurd. 1975. Electrostatic effects in myoglobin. Application of the modified Tanford-Kirkwood theory to myoglobins from horse, California grey whale, harbor seal, and California sea lion. *Biochemistry*. 14:1352–1358.
- Springer, B. A., S. G. Sliger, J. S. Olson, and G. N. Phillips, Jr. 1994. Mechanisms of ligand recognition in myoglobin. *Chem. Rev.* 94:699–714.
- Tanokura, M. 1983. ^1H -NMR study on the tautomerism of the imidazole ring of histidine residues. I. Microscopic pK values and molar ratios of tautomers in histidine-contacting peptides. *Biochim. Biophys. Acta.* 742:576–585.
- Ullmann, G. M., and E. W. Knapp. 1999. Electrostatic models for computing protonation and redox equilibria in proteins. *Eur. Biophys. J.* 28:533–551.
- Vagedes, P., B. Rabenstein, J. Åqvist, J. Marelus, and E. W. Knapp. 2000. The deacylation step of acetylcholinesterase: computer simulation studies. *J. Am. Chem. Soc.* 122:12254–12262.
- Warshel, A., and J. Åqvist. 1991. Electrostatic energy and macromolecular function. *Annu. Rev. Biophys. Biophys. Chem.* 20:267–298.
- Warshel, A., A. Papazyan, and I. Muegge. 1997. Microscopic and semi-macroscopic redox calculations: what can and cannot be learned from continuum models. *JBIC*. 2:143–152.
- Warshel, A., and S. T. Russel. 1984. Calculations of electrostatic interaction in biological systems and in solution. *Q. Rev. Biophys.* 17:283–422.
- Warshel, A., S. T. Russell, and A. K. Churg. 1984. Macroscopic models for studies of electrostatic interactions in proteins: Limitations and applicability. *Proc. Natl. Acad. Sci. U.S.A.*, 81:4785–4789.
- Warwicker, J., and H. C. Watson. 1982. Calculation of the electric potential in the active site cleft due to α -helix dipoles. *J. Mol. Biol.* 157:671–679.
- Wilbur, D. J., and A. Allerhand. 1976. Titration behavior of individual tyrosine residues of myoglobins from sperm whale, horse, and red kangaroo: application of natural abundance carbon 13 nuclear magnetic resonance spectroscopy. *J. Biol. Chem.* 251:5187–5194.
- Yang, A.-S., M. R. Gunner, R. Sompogna, and B. Honig. 1993. On the calculation of pK_a 's in proteins. *Proteins: Struct., Funct., Genet.* 15:252–265.
- Yang, A.-S., and B. Honig. 1994. Structural origins of pH and ionic strength effects on protein stability. *J. Mol. Biol.* 237:602–614.
- Yang, F., and G. N. Phillips, Jr. 1996. Crystal structure of CO-, deoxy- and met-myoglobins at various pH values. *J. Mol. Biol.* 256:762–774.
- You, T., and D. B. Bashford. 1995. Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. *Biophys. J.* 69:1721–1733.