glycine has a similar local structural effect in these molecules remains to be seen.

Conclusions

The conclusions of this study are that the effect of a single glycine substitution in an isolated α-helix reduces the mean helix content of the peptide as one would predict given the s value for this side chain.[9,10] However, an individual glycine exerts a strong local destabilizing effect on the helix at and close to the substitution site. The helical structure in these peptides is likely to reflect an equilibrium between helix and coil that is rapid on the NMR time scale at 25 °C.

Acknowledgment. This research was supported by a grant from the NIH (GM 40746). The peptide synthesizer was acquired with support from the NSF (DIR 87-22895). The Bruker AM 500 NMR spectrometer was purchased with funds from the NIH (RR 02497) and the NSF (DMB 84-13986).

Registry No. LLL, 131131-67-4; GGG, 132791-42-5; LGL, 131131-69-6; G, 56-40-6.

Determination of pK_as of Ionizable Groups in Proteins: The pK_a of Glu 7 and 35 in Hen Egg White Lysozyme and Glu 106 in Human Carbonic Anhydrase II

Kenneth M. Merz, Jr.

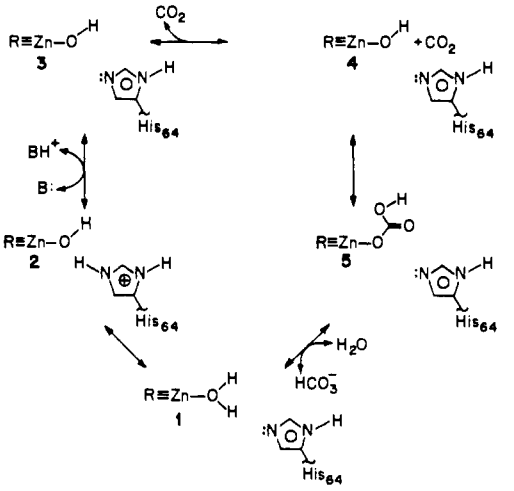
Contribution from the Department of Chemistry and the Department of Molecular and Cell Biology, The Pennsylvania State University, University Park, Pennsylvania 16802.
Received May 29, 1990

Abstract: We report a method by which molecular dynamics–free energy perturbation simulations can be used to estimate the pK_a of ionizable groups in proteins. The method has been tested to demonstrate its effectiveness in determining the pK_a of Glu 7 and 35 in hen egg white lysozyme, where the former has a pK_a of 2.6 and the latter a perturbed pK_a of 6.0–6.5 (in the presence of un-ionized Asp 52). We predict that the pK_a of Glu 7 is 3.1 ± 3.1, while we predict that Glu 35 has a pK_a of 12.0 ± 1.3. These test simulations indicate that (1) our approach is capable of predicting the perturbation of the pK_a of glutamic acid to higher values, (2) the precision of the approach depends on the conformational flexibility of the glutamic acid side chain, and (3) the accuracy of the approach, on average, is ±3 pK_a units. This approach was then applied to the problem of the pK_a of Glu 106 in human carbonic anhydrase II (HCAII). The activity of HCAII is dependent on a group whose pK_a is around 7.0. Glu 106 has been implicated as this group, but this requires the pK_a of this residue to be around 7. We predict that this group has a pK_a of 2.2 ± 2.8, which, even given the accuracy of our method, suggests that this group is not the activity-linked group. The present work demonstrates that our approach can be fruitfully applied to chemically important questions and that free energy methods can be applied to the determination of pK_a's in proteins with an accuracy to about ±3 pK_a units.

Introduction

The catalytic mechanism of human carbonic anhydrase (HCA) has been studied in detail.[1-4] The catalysis is dependent on a group whose pK_a is around 7.[1] After much debate it was decided that a zinc-bound water best satisfied this criterion, which led to the formulation of the zinc–hydroxide mechanism (see Scheme I).[1] However, one mechanism that is significantly different than the zinc–hydroxide mechanism, but cannot be ruled out is the proton shuttle mechanism of Kannan et al. (see Scheme II).[2,3] The zinc ion sill has a similar role, but the proton is not shuttled out of the active site via His 64 (**1** → **2**). Instead, this mechanism involves the active site residues Thr 199 and Glu 106 in a proton relay (forward A → B).[2-4] The one drawback of this mechanism,

Scheme I



though, is that it requires that the activity-linked group whose pK_a is around 7 be Glu 106 (C → D). The pK_a of a glutamic acid is normally around 4.0, but highly perturbed pK_a's for glutamic acids have been observed in proteins.[5] Experimentally it

(1) Merz, K. M., Jr.; Hoffmann, R.; Dewar, M. J. S. *J. Am. Chem. Soc.* **1989**, *111*, 5636. For a selection of reviews, see: Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30. Silverman, D. N.; Vincent, S. H. *CRC Crit. Rev. Biochem.* **1983**, *14*, 207. Lipscomb, W. N. *Annu. Rev. Biochem.* **1983**, *52*, 17. Lindskog, S. In *Zinc Enzymes*; Spiro, T. G., Ed.; John Wiley & Sons: New York, 1983. *Metal Ions in Biological Systems*; Sigel, H., Ed.; Marcel Dekker: New York, 1983; Vol. 15, p 77. *Biophysics and Physiology of Carbon Dioxide*; Bauer, C., Gros, G., Bartels, H., Eds.; Springer-Verlag: New York, 1980. Prince, R. H. *Adv. Inorg. Chem. Radiochem.* **1979**, *22*, 349. Chlebowski, J. F.; Coleman, J. B. In *Metal Ions in Biological Systems*; Sigel, H., Ed.; Marcel Dekker: New York, 1976; Vol. 6. Pocker, Y.; Sarkanen, S. *Adv. Enzymol.* **1978**, *47*, 149. Bertini, I.; Luchinat, C.; Scozzafava, A. *Struct. Bonding (Berlin)* **1981**, *48*, 45.

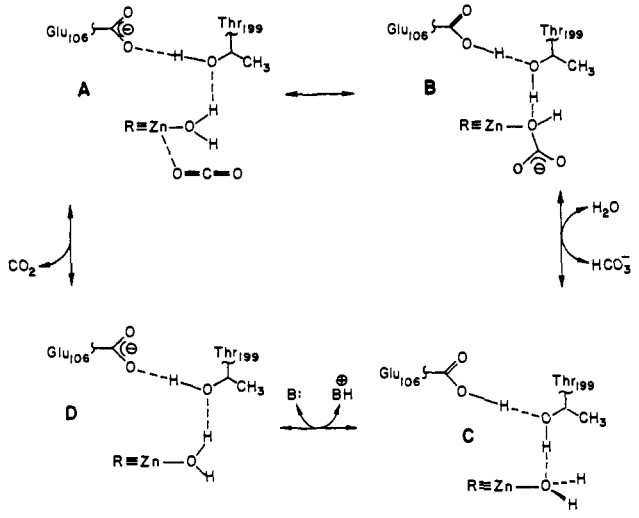
(2) Kannan, K. K.; Ramanadham, M.; Jones, T. A. *Ann. N.Y. Acad. Sci.* **1984**, *429*, 49.

(3) Kannan, K. K.; Liljas, A.; Waara, I.; Bergsten, P.-C.; Lövgren, S.; Strandberg, B.; Bengtsson, U.; Carlbom, U.; Fridborg, K.; Jarup, L.; Petef, M. *Cold Spring Harbor Symp. Quant. Biol.* **1971**, *36*, 221.

(4) Eriksson, E. A.; Jones, T. A.; Liljas, A. In *Zinc Enzymes*; Bertini, I., Luchinat, C., Maret, W., Zeppezauer, M., Eds.; Birkhäuser: Boston, 1986; p 317. Eriksson, E. A.; Jones, A. T.; Liljas, A. *Proteins* **1989**, *4*, 274. Eriksson, E. A.; Kylsten, P. M.; Jones, T. A.; Liljas, A. *Proteins* **1989**, *4*, 283.

0002-7863/91/1513-3572$02.50/0

Scheme II



has been difficult to evaluate the pK_a of Glu 106 in HCAII to test the viability of the Kannan mechanism. Hence, the development of theoretical techniques that could routinely and accurately determine the pK_a of ionizable amino acids would significantly enhance our ability to understand protein function among other things. In the present work we present an approach to estimate the pK_a of ionizable groups in proteins using molecular dynamics–free energy perturbation (MD–FEP) simulations.[6]

Several simple microscopic models have been utilized to determine the pK_a's of various ionizable groups. Warshel's PDLD and SCSSD methods were used to evaluate pK_a's in solution and in proteins and gave quantitative results.[7] Recently, a continuum model of the surrounding solvent has been employed by Rashin et al. to determine hydration enthalpies and pK_a's for small molecules in solution.[8] These authors found that the hydration enthalpies determined with continuum model correlated well with free energies of hydration evaluated by using free energy perturbation techniques.[9] These approaches have the advantage that they are computationally inexpensive and can provide useful insights into environmental effects on pK_a's. While these approaches have been generally successful, it is still important to employ more sophisticated models because they give a more detailed and realistic representation of the system and the factors affecting pK_a's of ionizable groups.

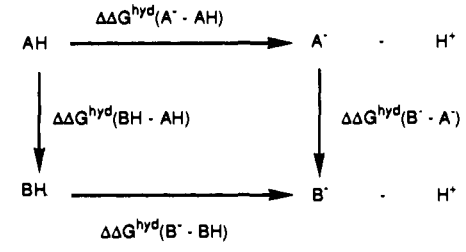
Some theoretical work has been directed toward the determination of pK_a values using more realistic molecular representations. Warshel et al.[10] described efforts using MD–FEP to determine the free energies of hydration of two carboxyl groups in bovine pancreatic trypsin inhibitor (BPTI). They found that they could get reasonable estimates for this quantity with very high precision (±1 kcal/mol). For small organic molecules Jorgensen et al.[11] described an approach to determine pK_a's using ab initio and Monte Carlo (MC) simulations using FEP techniques. Guissani et al.,[12] using MD simulations and particle insertion techniques to determine the free energy of hydration of H_2O, OH^-, and H_3O^+, have determined the pH of water. Continuum models have recently been employed with success to determine pK_a shifts in active site residues upon the alteration of the amino acid composition of a protein via site-directed mutagenesis.[13] Recently, Bashford and Karplus[14] have used these techniques to determine the pK_a of all of the ionizable groups in hen egg white lysozyme and found that they were able to predict pK_a's to an accuracy from 0 to 3.3 pK_a units. Our current approach is an extension of the seminal work done by both Warshel et al.[10] and Jorgensen et al.[11]

The "master equation" for dealing with small organic solutes is[11]

$$2.3RT[pK_a(BH) - pK_a(AH)] = \Delta\Delta G^{hyd}(B^- - A^-) - \Delta\Delta G^{hyd}(BH - AH) + \Delta\Delta G_{gas}(BH - AH) \quad (1)$$

where the first two terms on the right-hand side of eq 1 are the difference in the free energy of hydration of the neutrals (AH and BH) and the corresponding anions (A^- and B^-), which can be determined from FEP simulations, while the third term is the difference in the gas-phase free energy of ionization of AH and BH and can be evaluated by ab initio techniques.[11] This equation is well suited to study the pK_a differences between small molecules, but it is not suited to the study of pK_a differences in proteins. The equation above requires that we perturb BH into AH, which implies that in order to study a protein we must perturb our reference molecule BH (a dipeptide, for example) into our protein AH. Furthermore, the determination of $\Delta\Delta G_{gas}(BH - AH)$ by ab initio techniques is impossible for studies involving proteins.

In order to solve this problem we have adopted the following approach, which has been used previously by Russell and Warshel[15] and Bashford and Karplus:[14] (1) Since we are interested in the ionization of a glutamic acid embedded in a protein matrix we can eliminate the need to determine $\Delta\Delta G_{gas}(BH - AH)$ by using a free glutamic acid dipeptide in solution as our reference molecule of known pK_a. The implicit assumption being that the differences in the gas-phase quantum mechanical contribution to the free energies of ionization of AH and BH will be zero since the ionizing groups are essentially identical. This approximation is justified because the group we are dealing with is always a glutamic acid, which is embedded in a protein matrix. In cases where one is looking for pK_a differences in, for example, para-substituted benzoic acids this approximation is less justified because resonance and inductive effects in these cases will have a profound affect on the pK_a's of these compounds. Thus, $\Delta\Delta G_{gas}(BH - AH) \approx 0.0$. (2) The following free energy cycle offers us a useful relationship:



$$\Delta\Delta G^{hyd}(B^- - A^-) - \Delta\Delta G^{hyd}(BH - AH) = \Delta\Delta G^{hyd}(B^- - BH) - \Delta\Delta G^{hyd}(A^- - AH) \quad (2)$$

This thermodynamic cycle tells us that we can relate the free energy differences between the anions and neutrals to that of the free energy difference between the un-ionized and ionized forms of the protein or small molecule. This indicates that all we have

(5) Fersht, A. R. *Enzyme Structure and Mechanism*; W. H. Freeman and Co.: New York, 1985.

(6) van Gunsteren, W. F. *Protein Eng.* **1988**, *2*, 5. Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1. Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184. Kollman, P. A.; Merz, K. M., Jr. *Acc. Chem. Res.* **1990**, *23*, 246. Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Advances in Chemical Physics. Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; John Wiley and Sons: New York, 1988; Vol. 71. McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: New York, 1987.

(7) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. Warshel, A. *J. Phys. Chem.* **1979**, *83*, 1640. Warshel, A. *Biochemistry* **1981**, *20*, 3167. Warshel, A.; Russell, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283.

(8) Rashin, A. A.; Namboodiri, K. J. *J. Phys. Chem.* **1987**, *91*, 6003. Rashin, A. A. *Int. J. Quantum Chem.: Quant. Biol. Symp.* **1988**, *15*, 103.

(9) Rashin, A. A.; Rabinowitz, J. R.; Banfelder, J. R. *J. Am. Chem. Soc.* **1990**, *112*, 4133.

(10) Warshel, A.; Sussman, F.; King, G. *Biochemistry* **1986**, *25*, 8368.

(11) Jorgensen, W. L.; Briggs, J. M. *J. Am. Chem. Soc.* **1987**, *109*, 6857. Jorgensen, W. L.; Briggs, J. M. *J. Am. Chem. Soc.* **1989**, *111*, 4190.

(12) Guissani, Y.; Guillot, B.; Bratos, S. *J. Chem. Phys.* **1988**, *88*, 5850.

(13) Gilson, M. K.; Honig, B. H. *Nature* **1987**, *330*, 84. Sternberg, M. J. E.; Hayes, F. R. F.; Russell, A. J.; Thomas, P. G.; Fersht, A. R. *Nature* **1987**, *330*, 86. Bashford, D.; Karplus, M.; Canters, G. W. *J. Mol. Biol.* **1988**, *203*, 507. Gilson, M. K. *Proteins* **1988**, *3*, 32.

(14) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219.

(15) Russell, S. T.; Warshel, A. *J. Mol. Biol.* **1985**, *185*, 389.

to do is perturb the acid between its ionized and un-ionized form (i.e., $A^- \rightarrow AH$) and not between BH and AH (or $B^- - A^-$). Note that what we are determining here will be a relative free energy of ionization, which can be related to a relative pK_a change that occurs in response to the protein environment versus an aqueous environment. The new master equation becomes:

$$2.3RT[pK_a(BH) - pK_a(AH)] = \Delta\Delta G^{hyd}(B^- - BH) - \Delta\Delta G^{hyd}(A^- - AH) \quad (3)$$

This equation allows us to take a reference molecule, like a glutamic acid dipeptide in the present case, of known pK_a and use it to determine the unknown pK_a of a glutamic acid in a protein.[11]

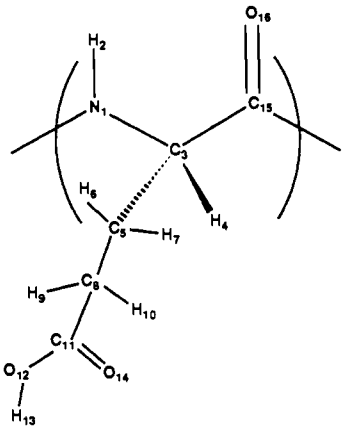
In order to test this approach we will first determine if we can make predictions about the pK_a's of two glutamic acids in hen egg white lysozyme (HEWL). In HEWL the pK_a of Glu 35 is known to be between 6.0 and 6.5[16], while a normal pK_a for glutamic acid is about 4.0.[5] Thus Glu 35 is, at most, perturbed by 2 pK_a units or 2.7 kcal/mol. As a further check on our calculations we will determine the pK_a of Glu 7 in HEWL. The pK_a of Glu 7, which is exposed to the surrounding solvent, has been determined experimentally to be 2.6.[17] Following these tests we will make a prediction about the pK_a of Glu 106 in HCAII.

Computational Procedure

The computations on HEWL and HCAII were carried out as follows: The coordinate set for HEWL (triclinic form) was taken from the Brookhaven Protein Databank,[18] and the HCAII (2.0-Å resolution) coordinates were supplied to us by E. A. Eriksson.[4] The total charge of the protein was first neutralized by the addition of counterions and all charged residues that did not form a salt bridge with another charged residue also had a counterion placed in its proximity by use of the CION option in the EDIT program contained in the AMBER suite of programs.[19] The net charge of the protein was −1 in all cases, where the glutamic acid that was being studied carried the net negative charge. For the glutamic acid residue 6-31G* ESP derived point charges[20] and the AMBER all-atom model was used.[21] The remainder of the protein was described by the AMBER united atom model.[21] Table I lists the charges used for glutamic acid in both the ionized and neutral forms. Table I also gives STO-3G charges that were determined for neutral aspartic acid. The details concerning the incorporation of the zinc ion into HCAII are given elsewhere.[22] The starting structures for the MD simulations were first minimized fully to remove any bad intermolecular contacts. These structures were then solvated with a 15-Å sphere of TIP3P[23] water molecules, which was centered at the C_β carbon of the glutamic acid. For HCAII, the crystallographically observed water molecules were incorporated into the model. This resulted in about 300 water molecules solvating the glutamic acid in both HEWL and HCAII. The water molecules were kept in this sphere with the use of harmonic restraining forces (0.5 kcal/mol Å). The FEP simulations were kept at 298 K by coupling to a temperature bath.[24] All residues within 15 Å of the glutamic acid were permitted to move during the course of the simulation as were all water molecules, while residues lying outside this region were held fixed. SHAKE[25] was used to constrain bond lengths at their equilibrium value and a time step of 1.5 fs was employed. The nonbonded pair list had a cutoff of 10 Å and was updated every 50 time steps; a constant dielectric of 1 was used throughout. The equilibration period was 18 ps. The FEP simulations were done by the slow growth procedure.[6] For each case, two separate simulations started from the equilibrated structure and another structure was generated from a further 18 ps of equilibration. The simulation time period was 90 ps and they were carried out in the forward ($\lambda = 1 \rightarrow 0$) direction only. In the case of Glu

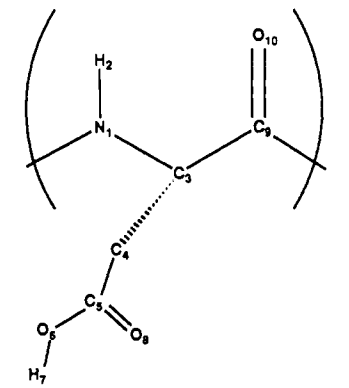
Table I. Glu Residue Charges

Neutral and Ionized Glutamic Acid
6-31G* ESP Charges (All-Atom Model)[a]



atom label	ionized Glu	neutral Glu	atom label	ionized Glu	neutral Glu
N-1	−0.566	−0.304	H-9	0.104	0.106
H-2	0.340	0.296	H-10	0.104	0.106
C-3	−0.126	−0.680	C-11	0.987	0.878
H-4	0.056	0.225	O-12	−0.872	−0.706
C-5	0.223	0.225	H-13	-	0.480
H-6	−0.004	0.036	O-14	−0.872	−0.670
H-7	−0.004	0.036	C-15	0.729	0.929
C-8	−0.444	−0.301	O-16	−0.655	−0.656

Neutral Aspartic Acid STO-3G Charges (United-Atom Model)



atom label	neutral Asp	atom label	neutral Asp	atom label	neutral Asp
N-1	−0.520	C-5	0.599	O-8	−0.433
H-2	0.248	O-6	−0.169	C-9	0.526
C-3	0.229	H-7	0.138	O-10	−0.500
C-4	−0.118				

[a] AMBER Glu residue charges are given in parentheses.

Table II. Computed Free Energies of Ionization for a Glu Dipeptide, Glu 35 and 7 in HEWL and Glu 106 in HCAII[a]

simulation	18 ps	36 ps	54 ps	ΔG_{av}
dipeptide[b]	−70.2 ± 2.1	−71.2 ± 0.6		−70.7 ± 1.6
Glu 7	−76.7	−68.3	−71.2	−72.1 ± 3.8
Glu 35[c]	−81.3	−82.0		−81.7 ± 0.4
Glu 35[d]	−59.0	−60.8		−59.9 ± 0.9
Glu 35[e]	−34.9	−33.3		−34.1 ± 0.8
Glu 106	−75.2	−71.2		−73.2 ± 2.3

[a] All free energies are given in kilocaries per mole. The error bars are $\pm 1\sigma$. [b] The free energies for the dipeptide model are averages for the forward ($\lambda = 1 \rightarrow 0$) and reverse ($\lambda = 0 \rightarrow 1$) simulations. For the rest they are only for the forward. [c] Asp 52 had a counterion associated with it. [d] Asp 52 is un-ionized. [e] Asp 52 is ionized.

(16) Parsons, S. M.; Raftery, M. A. *Biochemistry* **1972**, *11*, 1623, 1630, 1633.
(17) Kuramitsu, S.; Hamaguchi, K. *J. Biochem.* **1980**, *87*, 1215.
(18) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.
(19) Singh, U. C.; Weiner, P. K.; Caldwell, J.; Kollman, P., AMBER 3.0 UCSF.
(20) Williams, D. E.; Yan, J.-M. *Adv. At. Mol. Phys.* **1988**, *23*, 87. Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894. Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129. Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
(21) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Am. Chem. Soc.* **1986**, *7*, 230. For the neutral glutamic/aspartic acid parameters, see: Merz, K. M., Jr.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 5649.
(22) Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 406.
(23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(24) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. *J. Chem. Phys.* **1984**, *84*, 3684.
(25) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.

7 in HEWL, we ran another 18 ps of equilibration (for a total of 54 ps) and ran another free energy evaluation due to the observation of a significant amount of hysteresis between the first two simulations.

The determination of the free energy of ionization of a dipeptide model of glutamic acid was done in an analogous manner, with the exception that the dipeptide was solvated in a cube of 822 TIP3P water molecules and the simulations were carried out in the forward and reverse directions for 45 ps each way. Periodic boundary conditions were employed with constant temperature (298 K) and pressure (1 atm).[24] All enzyme FEP simulations used a variable-cutoff approach,[26] where the glutamic acid had a 300-Å cutoff while the remaining residues had the 10-Å cutoff. The MD–FEP simulations were all carried out with the GIBBS module in the AMBER suite of programs.[19]

Results and Discussion

The results are given in Table II. The difference in the free energy of hydration of the dipeptide model is calculated to be −70.7 ± 1.6 kcal/mol, which is in reasonable accord with the value of 70 kcal/mol determined by Warshel for an aspartic acid in solution and for a surface glutamic acid in BPTI.[7] The differences are probably due to the different force fields used to model the protein and the surrounding water. Assuming that the value for the pK_a of the dipeptide is 4.0 and using the computed free energies in Table II, we arrive at a value of 3.1 ± 3.1 for the pK_a of Glu 7 in HEWL. Analogously, we predict that the pK_a of Glu 35 in HEWL should have a pK_a of 12.0 ± 1.3 when Asp 52 is un-ionized. When we place a counterion around Asp 52 to neutralize its charge, the predicted pK_a is −4.0 ± 1.7, which is much too negative. This arises from the overestimation of the stability of the ionized form of Glu 35 relative to the neutral from. Thus, it appears that a counterion is probably not present in the vicinity of Asp 52 and Glu 35. We have also determined the pK_a for Glu 35 in the presence of ionized Asp 52 (net protein charge of −2) and found a very positive pK_a value of 30.9 ± 1.8. Here we are destabilizing the ionized form of Glu 35 relative to the un-ionized form. The experimental value for the pK_a of Glu 35 in the presence of ionized Asp 52 is about 8, but it is clear that our current model is unable to reproduce this value. Similar results were obtained by Warshel et al.[7]

The experimental value for the pK_a of Glu 7 is 2.6, while that for Glu 35 (with neutral Asp 52) is 6.0–6.5. For Glu 7 the agreement with experiment is reasonable and the accuracy is similar to that found by Jorgensen for the pK_a's he determined for small organic solutes. The calculated pK_a of Glu 35 is larger than the experimental value, but it is in agreement with experiment in suggesting that the pK_a of this group is perturbed to significantly higher pK_a values. The reason why the pK_a of Glu 35 is so high is due to its location in a hydrophobic region of the active site region in HEWL,[16] and our results suggest that this effect is overestimated by the AMBER[21] force field. From these test calculations we estimate that the average accuracy of our theoretical approach is about ±3 pK_a units. The precision of these simulations (as indicated by the error bars) varies quite dramatically and is probably due, in part, to the conformational flexibility of the glutamic acid side chain. Thus, the free energy determined for Glu 35 is more precise than that determined for Glu 7, presumably due to the greater conformational freedom for the latter. This same conclusion was arrived at by Bashford and Karplus in their work on HEWL.[14] This suggests that in order to increase the precision simulations longer than 90 ps are needed. Finally, these results demonstrate that our theoretical model is capable of clearly indicating that a particular pK_a is perturbed relative to its normal value. Finally, we note that determination of pK_a's in proteins is a powerful test of the molecular mechanical representation of nonbonded interactions in proteins. Specifically, this is a powerful test of the electrostatic interactions. Thus, in the future these techniques could be used to assist in the verification process of protein force fields.

The pK_a of Glu 106 in HCAII can be evaluated by using the data in Table II. Again using eq 3 and assuming that the pK_a of the dipeptide model is 4.0, we arrive at a prediction 2.2 ± 2.8 for the pK_a of Glu 106. This clearly indicates that the pK_a of Glu 106 is probably not perturbed even in light of our ±3.0 pK_a accuracy, which at most would give a pK_a of 5.2 for this residue. These results suggest that the Kannan mechanism is not operative in HCAII. Given our previous MD results, which suggested that the hydrogen-bonding network of Glu 106–Thr 199–zinc-hydroxide plays a role in orienting the substrate CO_2 for facile reaction, and the present results we suspect that the only role that Glu 106 is playing is as an anchor for the hydroxyl hydrogen of Thr 199.[27]

Conclusions

We have demonstrated that the MD–FEP method can be used to determine the pK_a of amino acid residues in proteins. The accuracy of the approach is at the ±3.0 pK_a unit level, but with improved potential functions and more realistic representations of the protein environment it is expected that this number will improve. The precision of the present simulations can be enhanced by using longer simulation time scales. Nevertheless, the method should be qualitatively useful in making predictions about the pK_a's of functionally important residues, as was demonstrated with HCAII.

It is also appropriate to compare our approach with that of Bashford and Karplus.[14] The continuum approach used by these authors employs a static picture of HEWL, while the MD–FEP approach uses a dynamic approach. This should allow the MD–FEP approach to be more sensitive to conformational changes that may be important in governing the pK_a of an ionizable group in a protein. The clear advantage of the continuum approach is the inexpensiveness of the calculations. The MD–FEP approach is very computationally intensive compared to the continuum approaches. The continuum models also have the advantage in the inclusion of ionic strength effects, but the MD–FEP approach the disadvantage that it treats the protein and the surrounding solvent as a constant dielectric medium ($e = 4$ and $e = 80$, respectively). Clearly, the present approach and the continuum models are both powerful methods and should be useful in addressing a number of issues concerning potential function quality and the pK_a's of ionizable groups in proteins.

Acknowledgment. We thank the Center for Academic Computing at the Pennsylvania State University for generous allocations of IBM 3090-600S computer time. We also thank Jim Caldwell for supplying us with the variable cutoff version of the GIBBS program. The Office of Naval Research is acknowledged for support of this work via an Office of Naval Research-Young Investigator Award.

Registry No. Glu, 56-86-0; lysozyme, 9001-63-2; carbonic anhydrase, 9001-03-0; Glu dipeptide, 3929-61-1.

(26) Caldwell, J., unpublished results.

(27) Merz, K. M., Jr. *J. Mol. Biol.* **1990**, *214*, 799.