Metalloion−Ligand Binding Energies and Biological Function of Metalloenzymes Such as Carbonic Anhydrase. A Study Based on ab Initio Calculations and Experimental Ion−Ligand Equilibria in the Gas Phase

Michael Peschke, Arthur T. Blades, and Paul Kebarle*

Contribution from the Department of Chemistry, University of Alberta, Edmonton, AB, Canada T6G 2G2

Received July 9, 1999

Abstract: The known structure of carbonic anhydrase includes the structure of the active site, which consists of Zn^{2+}, coordinated to three histidine (His) residues and a water molecule. The ligand involved in the catalysis is the water molecule. Using imidazole (Im) to model histidine, it is found that histidine is the strongest bonding residue to Zn^{2+} of all neutral amino acid residues. On the basis of high-level ab initio calculations, the sequential bond energies for one to three imidazoles attached to Zn^{2+} were evaluated. The bond energy $Zn(Im)_3^{2+}-(H_2O)$ was also determined by ab initio calculations and (gas phase) ion equilibria measurements. From a comparison of the free energy of stabilization of Zn^{2+} in the enzyme and that in aqueous solution, we conclude that very strongly bonding residues such as histidine are essential to make the Zn^{2+} ion in the enzyme stable, relative to the aqueous environment. The strongly bonding histidine also has another more important role. Due to the very large charge transfer and polarizability energy component with such a ligand and ligand−ligand repulsion, the bonding to the fourth, i.e., the H_2O ligand, is very much weakened. It is shown that weakening of the $Zn(L)_3^{2+}-(H_2O)$ bond energy is *unfavorable*, i.e., *increases* the energy required for the first step of the accepted mechanism, $Zn(His)_3OH_2^{2+} \rightarrow Zn(His)_3OH^+ + H^+$, while the second step, $Zn(His)_3OH^+ + CO_2 \xrightarrow{H_2O} Zn(His)_3(H_2O)^{2+} + HCO_3^-$, is favorably affected, i.e., requires less energy. The choice of the ligands L therefore must be such so as to lead to a compromise between the two opposing effects. By making a semiquantitative inclusion of the solvent effect of the protein and aqueous environment on the reaction free energies for the above two ionic reactions, it was possible to show that the "choice" of the three histidine ligands is "just right" to provide this compromise. These ligands lead to reaction free energies that are close to zero for both reactions.

I. Introduction

"Structure" as determined mainly by X-ray crystallography and nuclear magnetic resonance spectroscopy, has been an extraordinarily successful guide to "function" for biocompounds and particularly enzymes. The relationship between energy and function is examined less often. However, the relationship between structure, energy, and function becomes *the* focus, in investigations of the mechanism of the reaction catalyzed by the enzyme which include an examination of the structure and energy of the transition state(s) involved in the proposed mechanism.

In the area of metalloenzymes and particularly Zn^{2+} metalloenzymes such as carbonic anhydrase, impressive advances have been made in establishing the reaction mechanism of the catalyzed reaction.[1] The experimental investigations involved mainly determinations of the rate of catalysis, such as the steady-state rate constants and determinations of the crystallographic structure of carbonic anhydrase and particularly the structure of the active site.[2] The power of these techniques was enormously increased by the use of site-directed mutagenesis, where specific amino acid residues, suspected to be participating in the stabilization of the transition states in the wild type of the enzyme, are replaced and the crystallographic structure and enzymatic activity of these variants is determined and compared with that of the wild type.[3]

Numerous purely theoretical studies of the mechanism, i.e., the structure and energy of reactants and transition states, have been made.[4] These started with very simple models that involved only the Zn^{2+} ion and a few simple ligands such as H_2O and NH_3. With the rapid development of computational techniques, the modeling grew in sophistication and relevance so that the latest work,[4d,e] using coupled quantum mechanical calculations of transition states with molecular mechanics and molecular dynamic methods, provide very relevant information and details on individual steps in the reaction mechanism.

(1) (a) Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30. (b) Coleman, J. E. *Curr. Opin. Chem. Biol.* **1998**, *2*, 222.

(2) (a) Ericksson, A. E.; James, T. A.; Liljas, A. *Proteins Struct. Funct. Genet.* **1988**, *4*, 274. (b) Häkansson, K.; Carlsson, M.; Svenson, L. A.; Liljas, A. *J. Med. Biol.* **1992**, *227*, 1192. (c) Xue, Y.; Vidgren, J.; Svensson, L. A.; Liljas, A.; Jonsson, B. H.; Lindskog, S. *Proteins* **1993**, *15*, 80.

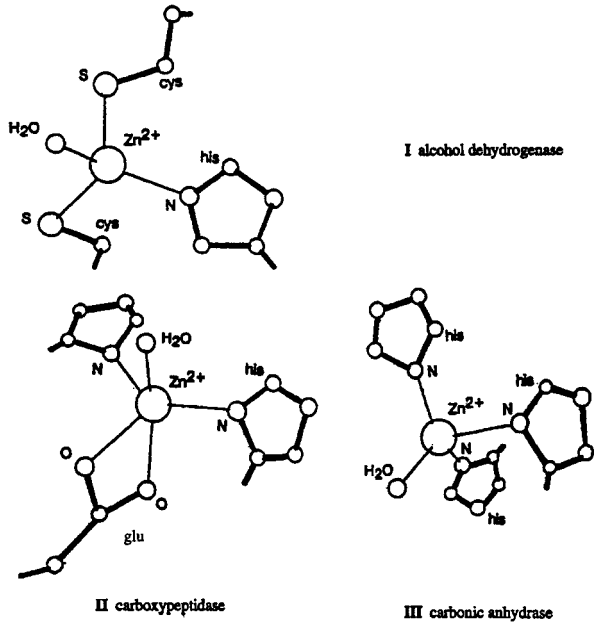
(3) (a) Kiefer, L. L.; Fierke, L. L. *Biochemistry* **1994**, *33*, 15233 and references therein. (b) Kiefer, L. L.; Palermo, S. A.; Fierke, C. A. *J. Am. Chem. Soc.* **1995**, *117*, 6831. (c) Ipolito, J. A.; Christianson, D. W. *Biochemistry* **1994**, *33*, 15241. (d) Lesburg, C. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1995**, *117*, 6838.

(4) (a) Pullman, A. *Ann. N. Y. Acad. Sci.* **1981**, *367*, 340. (b) Allen, L. C. *Ann. N. Y. Acad. Sci.* **1981**, *367*, 385. (c) Liang, J. Y., Lipscomb, W. N. *Int. J. Quantum Chem.* **1989**, *36*, 299. (d) Zheng, Y. J.; Merz, K. M. *J. Am. Chem. Soc.* **1992**, *114*, 10498. (e) Merz, K. M., Jr.; Banei, L. *J. Am. Chem. Soc.* **1997**, *119*, 863. (f) Zheng, Y. J.; Merz, K. H., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 10498.

10.1021/ja992406l CCC: $19.00
Published on Web 02/03/2000

As a result of the experimental and theoretical investigations mentioned above, the mechanism of the enzyme-catalyzed reactions involving carbonic anhydrase and specifically the isoenzyme human carbonic anhydrase (CAII) can be considered as well established.

The previous theoretical work,[4] on CAII was largely centered on the transition states involved in the catalyzed reaction and the participation of specific amino acid residues in these transition states. For several Zn^{2+} metalloenzymes (see structures **I**−**III**) one of the ligands is H_2O and it is this ligand that



participates in the catalysis, while the remaining three ligands remain bonded during the reaction. The remaining directly bonded ligands, which we will call the primary ligands, are often the amino acid residues: His, Cys, and Glu. They can be modeled by the the following compounds: imidazole, deprotonated CH_3S^-, and $C_2H_5CO_2^-$. The question can be asked, why were the specific primary ligands, **I**−**III**, selected by the evolutionary process?

The primary ligands can have two main purposes. One is to make the metal ion Zn^{2+} stable in the enzyme, relative to the aqueous environment. For that purpose, these ligands should be strongly bonding to Zn^{2+}. The other purpose should be to tune the bonding of the active ligand (H_2O) to the specific requirements of the catalytic reaction. These questions have been addressed experimentally for CAII, by site-directed mutagenesis, in which the primary histidine ligands were replaced by other residues.[3] The results from these beautiful experiments are very interesting; however, in some cases the explanation of the observations becomes very challenging because the changes observed involve not only the bonding changes at the reaction site, narrowly defined as the Zn−ligand complex, but also conformational adjustments of the protein backbone, repositioning of residues near the reaction site, and even reactive changes such as the ionization or neutralization of acidic residues such as Glu and Asp near the reactive site.[3] It appears desirable to be able to isolate the chemical changes of the narrowly defined reactive site, from the far more complex, but certainly very important, other changes. The present work is an attempt to provide an insight into this role of the primary ligands.

The results and discussion concerning the role of the primary ligands to provide stability of the Zn^{2+} ion in the enzyme, relative to the aqueous environment are given in section III. We discuss in section IV how the selection of the primary ligands can contribute to the minimization of the transition-state(s) energy. We hope to be able to show that useful insights concerning the choice of the primary ligands in the enzyme can be obtained from studies of the ion−ligand bond energies and an approximate inclusion of the ion solvation effect provided by the protein and aqueous solvent environment surrounding the active site.

The required ion−ligand bond energies were obtained from theoretical ab initio calculations and to a lesser extent from gas-phase ion−ligand equilibria determinations. The latter method has been our specialty over many years[5a,b] and has produced a wealth of data on sequential, n to $(n-1)$ ligand, dissociation enthalpies and free energies, of the ion−ligand complex ML_n^{z+}; see eq 1. Unfortunately, the ion equilibrium method, with

$$ML_n^{z+} = ML_{n-1}^{z+} + L \qquad (n, n-1) \qquad (1)$$

presently available instrumentation, can provide bond free energies, enthalpies, and entropies only for reactions in which the free energy change does not exceed ~28 kcal/mol. This range covers most ion−ligand interactions for singly charged ions. For doubly and triply charged metal ions, some difficulties are encountered. First, ion−ligand complexes of these ions are difficult to produce by the techniques used in the previous work.[5] Fortunately, the advent of electrospray, with which ions present in solution can be "transferred" to the gas phase and detected mass spectrometrically,[6a] has provided the means for production of many multiply charged metalloion complexes.[6b−d] The second difficulty is the limitation of the large free energy change. The bonding of doubly charged ions such as Zn^{2+} to most ligands is very strong for $n = 1$ to $n = 4$, and generally $(n, n-1)$ equilibria in that range cannot be determined with the present instrumentation. For a third difficulty due to side reactions, see section IIa. However, sufficiently accurate results can be obtained, particularly for smaller ligands, from ab initio calculations[7−9] in the low-n range. For higher values of n, where ab initio calculations become very difficult particularly for more complex ligands, the ion equilibrium technique can often be applied, because the $(n, n-1)$ bond free energies are lower for higher n values and often within the range of the technique.[6d,10] This complementarity of the two methods has been used also in the present work. However, because many of the bonding interactions of interest were strong, ab initio calculations have been the major source of data.

IIa. Experimental Results. Determination of the Gas-Phase Equilibria, $Zn(Im)_3^{2+} + H_2O = Zn(Im)_3H_2O^{2+}$ and $Zn(Im)_3(H_2O)^{2+} + H_2O = Zn(Im)_3(H_2O)_2^{2+}$, and Some Other

(5) (a) Dzidic, I.; Kebarle, P. *J. Phys. Chem.* **1970**, *74*, 1466. (b) Kebarle, P. Ion Thermochemistry and Solvation from Gas Phase Ion Equilibria. In *Annu. Rev. Phys. Chem.* **1977**, *28*, 445. (c) Keesse, R. G.; Castleman, A. W. *J. Phys. Chem. Ref. Data* **1986**, *15*, 1011.

(6) (a) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, *88*, 4451; **1984**, *88*, 4671. (b) Blades, A. T.; Jayaweera, P.; Ikonomou, M. G.; Kebarle, P. *J. Chem. Phys.* **1990**, *92*, 5900. (c) Blades, A. T.; Jayaweera, P.; Ikonomou, M. G.; Kebarle, P. *Int. J. Mass Spectrom. Ion Processes* **1990**, *101*, 355; **1990**, *102*, 251. (d) Peschke, M.; Blades, A. T.; Kebarle, P. *J. Phys. Chem. A* **1998**, *102*, 9978.

(7) Garmer, D. R.; Gresh, N. *J. Am. Chem. Soc.* **1994**, *116*, 3556.

(8) (a) Bauschlicher, C. W., Jr.; Sodupe, M.; Partridge, H. *J. Chem. Phys.* **1992**, *96*, 4453. (b) Bauschlicher, C. W., Jr.; Langhoff, S. R.; Partridge, H.; Rice, J. E.; Komanicki, A. *J. Chem. Phys.* **1991**, *95*, 5142.

(9) (a) Pavlov, M.; Siegbahn, P. E. E.; Sandström, M. *J. Phys. Chem. A* **1998**, *102*, 219. (b) Glendening, E. D.; Feller, D. *J. Phys. Chem.* **1996**, *100*, 4790.

(10) Blades, A. T.; Klassen, J. S.; Kebarle, P. *J. Am. Chem. Soc.* **1995**, *117*, 10563.

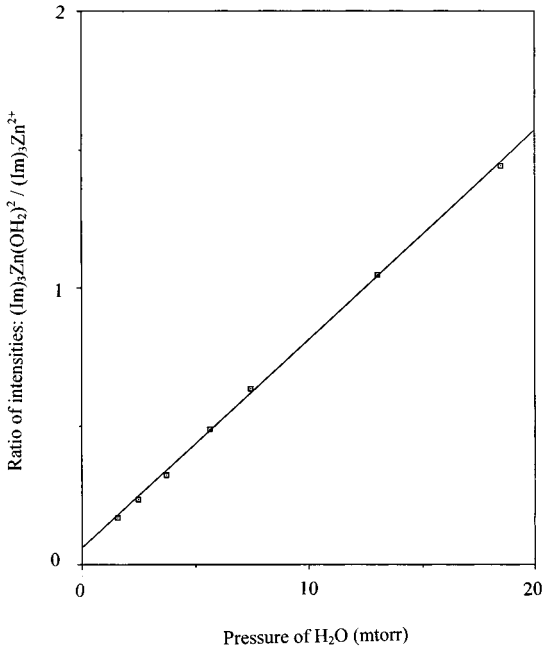


Figure 1. Experimental determination of equilibrium: $Zn(Im)_3H_2O^{2+} = Zn(Im)_3^{2+} + H_2O$. Ratio of detected ion intensity of $Zn(Im)_3H_2O^{2+}/Zn(Im)_3^{2+}$ shown on vertical axis while horizontal axis given water vapor pressure (in mTorr). Temperature 442 K. Constant slope demonstrates that equilibrium is present; see section IIa experimental results.

Reactions. The experiments were performed in an ion equilibration chamber into which ions produced by electrospray are introduced via a capillary. This apparatus has been described previously.[10] For the above equilibria, water vapor at variable pressures from 0 to 20 mTorr, carried in the nitrogen N_2 bath gas, flowed through the reaction chamber. To obtain $Zn(Im)_3^{2+}$ by electrospray, we used a solution of 3×10^{-4} M $ZnCl_2$ and 10^{-3} M imidazole (Im). Both methanol or water as solvents could be used.

The major ions observed under these conditions and an ion equilibration chamber temperature of ~440 K, were $Zn(Im)_4^{2+}$, $(Im)_2H^+$ and $Zn(Im)_3^{2+}$, $Zn(Im)_3(H_2O)_n^{2+}$. The combined intensity of the first two ions was generally ~3 times higher than the intensity of the desired ions $Zn(Im)_3^{2+}$ and $Zn(Im)_3(H_2O)_n^{2+}$.

Shown in Figure 1 is the ion intensity ratio $[Zn(Im)_3H_2O^{2+}]/[Zn(Im)_3^{2+}]$, plotted versus the partial pressures of water vapor that was present in the equilibration chamber. The straight line obtained fits the equation

$$[Zn(Im)_3(H_2O)^{2+}]/[ZnIm_3^{2+}]p_{H_2O} = K_p$$

where K_p is the equilibrium constant. The determination of K_p by such a plot was used to obtain the free energy change:

$$\Delta G^\circ_{0,1} = -RT \ln K_p \qquad -\Delta G^\circ_{0,1} = \Delta G^\circ_{1,0}$$

The free energies obtained for the first and second hydration equilibrium are given in Table 3A, section III.

The plot shown in Figure 1 was obtained at 442 K, since at 298 K, the equilibrium was strongly shifted toward the hydrates. Attempts to measure the equilibrium over a range of temperatures centered at 442 K were made. However, a wide enough range, required to obtain a reliable van't Hoff plot allowing determinations of the enthalpy and entropy change, could not be covered because of the relatively low intensities of the two reactant ions, and possible interference from other ion species present in the mass spectrum.

The free energy $\Delta G^\circ_{1,0} = 9.7$ kcal/mol at 442 K was converted to $\Delta G^\circ_{1,0} = 14.0$ kcal/mol at $T = 298$ K by using the theoretically determined entropy, $\Delta S^\circ_{1,0} = 30$ cal/(deg·mol) obtained by the theoretical calculations; see section IIb and Table 3A in section III. Dehydration reactions (1,0) most often lead to positive entropy changes between 25 and 30 cal/(deg·mol).[5,6d] Had an arbitrary value of $\Delta S^\circ_{1,0} = 25$ cal/(deg·mol) been used, the value $\Delta G^\circ_{1,0} = 13.3$ kcal/mol at $T = 298$ K would have been obtained.

Another ion observed at high intensity was $Zn(Im)_3OH(H_2O)^+$. This ion is interesting because it models an important reactant in the carbonic anhydrase mechanism;[1−4] see reaction 13 in section IV. The mode of formation in the present work could be via the side reaction

$$Zn(Im)_3(H_2O)_2^{2+} + B \rightarrow Zn(Im)_3OH(H_2O)^+ + BH^+$$

where B is a base such as imidazole. Such a reaction could be occurring already in the solution or at a later stage. Much less likely is a gas-phase reaction in which $B = H_2O$. A charge separation reaction of this type would not be expected for a Zn^{2+} ion stabilized by the three imidazole ligands.[11]

Since it is desirable to provide some experimental determinations involving Zn−ligand complexes, for comparison with the theoretical predictions, the determination of the hydration equilibrium

$$Zn(H_2O)_4^{2+} = Zn(H_2O)_3^{2+} + H_2O$$

was attempted. According to the theoretical data by Pavlov and co-workers,[9a] the $\Delta G^\circ_{4,3}$ for this process was small enough for the equilibrium to be observable within the high-temperature range of the present apparatus. However, when the temperature was increased above 310 K, the charge separation side reaction leading to $ZnOH(H_2O)_n^+$ became visible. Above 373 K, these ions became dominant over the observed hydrates: $Zn(H_2O)_n^{2+}$, $n = 6, 7, 8$. The $Zn(H_2O)_4^{2+}$ could not be obtained, because the charge-separated products were completely dominant at the higher temperatures at which it was expected. Therefore, equilibria involving $ZnOH(H_2O)_n^+$ were measured. The values obtained, $\Delta G^\circ_{4,3} = 6.8$ kcal/mol at $T = 357$ K and $\Delta G^\circ_{3,2} = 10.4$ kcal/mol at $T = 440$ K are given in section III, Table 3A. For comparison with theoretical results, see section IIb.

The $Zn(Im)_2(CH_3COO_2)^+$ ion, where the acetate anion $CH_3COO_2^- = Ac^-$, is used as a model for the glutamate residue in the ion−ligand complex of carboxypeptidase (see structure II) and could be prepared by electrospray of solutions of $Zn(Ac)_2$ and imidazole, both at 10^{-4} M concentrations in methanol. The equilibrium: $Zn(Im)_2(Ac)^+ + H_2O = Zn(Im)_2(Ac)(H_2O)^+$ could be determined at 298 K, using water vapor in the equilibration chamber in the range 10−100 mTorr. The free energy obtained $\Delta G^\circ_{1,0} = 5.0$ kcal/mol is given in section III, Table 3A.

IIb. Theoretical Calculations. All computations, unless otherwise noted, were done at the Becke3LYP/6-311++G(d,p) level as implemented in the Gaussian94 suite.[12] For the large chemical systems examined in this paper, DFT methods are a necessity because they are fast, compared to other correlated methods, and have much lower memory demands. In a recent critical evaluation, Curtiss et al.[13] compared G2 and DFT

(11) Peschke, M.; Blades, A. T.; Kebarle, P. *Int. J. Mass Spectrom. Ion Processes* **1999**, *185/186/187*, 685.

methods by computing the well-established enthalpies of formation of 148 molecules. The DFT functionals that comprise the Becke3LYP method were shown to give the smallest average absolute deviation (3.11 kcal/mol at the B3LYP/6-311+G(3df,-2p)//B3LYP/6-31G(d) level) of the seven DFT methods surveyed. By comparison, the G2 method produced an average absolute deviation of 1.58 kcal/mol. The energies that were calculated are heats of formations for each molecule, not reaction enthalpies. In heats of formation calculations, each bond is determined computationally so that there is little chance for cancellation of errors. In reaction enthalpy calculations, because many of the bonds are conserved with only small variations, some cancellation of error can occur generally reducing the error bars. In this work, a slightly smaller basis set was used, but it has been shown that DFT does not have a large basis set dependence, once a sufficiently large basis set is reached.[14]

In comparison with the MP2 method, B3LYP is of similar quality as shown in a number of studies.[14] However, the reliability of DFT does depend to some extent on the type of system studied. For multiple metal ion−ligand interactions such as $Zn(Imidazole)_3^{2+}$, it is possible that the indicated error of ∼3 kcal/mol could be considerably larger. Therefore, comparisons with experimental determinations are very useful. Using Pavlov's computational ΔH_0° values at the B3LYP/6-311+G-(2d,2p) level for the loss of one H_2O molecule from $Mg(H_2O)_6^{2+}$, $Mg(H_2O)_7^{2+}$, $Ca(H_2O)_6^{2+}$, and $Ca(H_2O)_7^{2+}$, which are 24.5, 19.0, 24.7, and 17.6 kcal/mol, respectively,[9a] and comparing them to the experimental values of 24.6, 20.3, 25.3, and 16.9 kcal/mol, respectively, obtained in this laboratory,[6d] shows that even for multiple ion−ligand interactions the error can be quite small. For these four systems, the average error is 0.7 kcal/mol. This extremely good agreement between theory and experiment is quite possibly fortuitous and other more complex ligands with more charge-transfer contributions and electronic rearrangements when coordinating to the doubly charged metal cation would increase the average error. For the transition metal ion Zn^{2+}, of special interest in the present work, comparisons between the theoretically evaluated and experimentally determined ΔG values for the loss of one water molecule were obtained for $ZnOH(H_2O)_3^+$ and for $ZnOH(H_2O)_4^+$. For the complex with three waters, the experimental value (see section IIa and Table 3A in section III) for ΔG_{442}° of 10.4 kcal/mol has been obtained, while the theoretical result for ΔG_{442}° is 9.7 kcal/mol. For the complex with four waters, the experimental ΔG_{357}° is 6.8 kcal/mol compared to the theoretical ΔG_{357}° of 6.6 kcal/mol. For the Zn−imidazole complexes, $Zn(Im)_3H_2O^{2+}$, the experimental result is $\Delta G_{440}^\circ = 9.7$ kcal/mol, and the theoretical value is $\Delta G_{442}^\circ = 7.5$ kcal/mol (Table 3A in section III). In the last case, the difference between the two results is larger, 2.2 kcal/mol, but still very satisfactory considering the complexity of the ion involved.

For the imidazole binding energy in the $Zn(Im)^{2+}$ complex, Garmer and Gresh[7] obtained an MP2 energy value that differs from our DFT result by ∼12.5 kcal/mol (approximate, because it is not clear to what extent zero-point energy correction is included in their work) (see Table 1 in section III). Redoing, in this work, a full MP2 optimization at the same basis set level as was used for the DFT calculation (6-311++G(d,p)) yields a difference of 12.4 kcal/mol. A single-point energy calculation at the MP4/6-311++G(d,p) level, using the MP2 geometry, reduces the difference to 11.5 kcal/mol. At this point it is not clear which of the methods, DFT or MP4//MP2, provides the more reliable energy value, given the good agreement of DFT with experiment for the water dissociation energy of the $Zn(Im)_3H_2O^{2+}$ complex. It is, however, possible that in the energy determination for the complex a possible zinc−imidazole interaction error is reduced due to cancellation, so that the good agreement of DFT to experiment cannot be used to imply equal accuracy for the $Zn(Im)^{2+}$ binding energy determination. Consequently, a reasonable average error for the more strongly bonded systems could be 5−10 kcal/mol. Successive binding energies are expected to be in better agreement and for a subsequent error analysis where the total ligand dissociation energy for three ligands is considered (see Table 1 in section III) an average error of 20 kcal/mol is proposed.

The basis set used is the 6-311++G** set as supplied by Gaussian94. Use of such a relatively large basis set allows for a better description of the often longer range ion−ligand interactions. In addition, the basis set superposition error (BSSE) is reduced for large basis sets, especially in conjunction with DFT methods.[14b] The expected BSSE would be close to the error limit due to other assumptions inherent in the method, and consequently, no BSSE corrections were done. While the basis set does not include as many polarization functions as were used for the energy determinations by Curtiss et al.,[13] it does include an additional diffuse function on the hydrogens. In addition, all geometries were optimized at the higher level rather than the B3LYP/6-31G(d) level.

For the larger systems, generally only one conformer was considered and it is possible that alternate conformers exist that would be lower in energy. Attention was given to avoid obviously unfavorable ligand interactions, but to definitely determine the most energetically favorable conformer was computationally too demanding. However, most of these conformers that are important for the model would be just rotation of the ligand around the metal ion−ligand bond. The difference in energy between such conformers would be expected to be no more than 2−3 kcal/mol in most cases. For some systems, there is also the possibility of water or OH^- interacting with the N−H hydrogen of the imidazole to form hydrogen bonds rather than direct interaction with the metal cation center. While these conformers can be of comparable energy, they are irrelevant to the model studied and therefore outside the focus of this paper.

Finally, a note of explanation must be added for the $Zn(Im)_3(H_2O)^{2+}$ and the $Zn(Im)_3(OH)^+$ systems. These are very large systems for the theoretical level chosen, but because long-range interactions play such an important role in ion−multiple ligand interactions, it was felt that a smaller basis set might not provide the necessary degree of accuracy. However, a consequence of the large basis set, in addition to the normal computational expense, is the tendency that the potential energy surface (PES) for ion−multiple ligand systems becomes "grainy", with many small minimums and maximums often relatively close together, depending on the total combination of the many long-range

(12) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision D.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

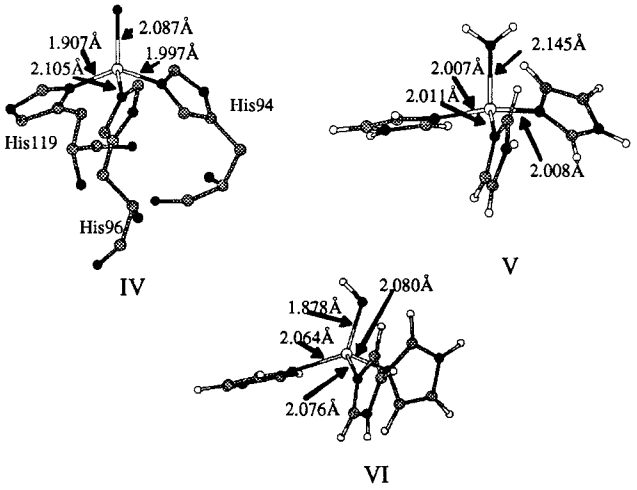
(13) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063.

(14) (a) Wiberg, K. B.; Ochterski, J. W. *J. Comput. Chem.* **1997**, *18*, 108. (b) Smith, B. J.; Radom, L. *Chem. Phys. Lett.* **1994**, *231*, 345. (c) Hobza, P.; Šponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315. (d) González, L.; Mó, O.; Yáñez, M. *J. Comput. Chem.* **1997**, *18*, 1125. (e) Kafafi, S. A.; Krauss, M. *Int. J. Quantum Chem.* **1999**, *75*, 289.

interaction forces.[15] This grainy property of the PES makes optimizations quite difficult since the actual forces are very small, but the optimization algorithm is often incapable of deciding the right step size to obtain new values for the various coordinates without the help of very expensive frequency calculations. Once the optimization procedure has reached this stage, the energy differences between new optimization guesses are generally on the order of 0.1 kcal/mol. The electronic energy that is obtained from such a structure will therefore be within 0.1−0.2 kcal/mol of an eventual minimum. Moreover, in any experimental system, the entire closeby portion of the PES surface with similar energy values will be sampled at ambient temperatures.[16] Consequently, energy values for systems that have reached this stage can be regarded as reliable, even if not all standard optimization criteria have been fulfilled. While the $Zn(Im)_3(H_2O)^{2+}$ system fortuitously fulfilled all standard optimization criteria, the $Zn(Im)_3(OH)^+$ system did not. All forces were extremely small—approaching or exceeding the tight optimization criteria, but the proposed displacement values did not meet the standard optimization criteria. Nevertheless, on the basis of observations of many other ion−multiple ligand systems (see Tables 1 and 3 in section III), the electronic energy for $Zn(Im)_3(OH)^+$ is expected to be within 0.1−0.2 kcal/mol of an eventual minimum for that portion of the hyper-PES.

Computational results are diplayed in Tables 1 and 3 in section III along with comparison to available literature values by Garmer and Gresh,[7] Pavlov et al.,[9a] and Glendening and Feller.[9b]

IIc. Comparison of Native Carbonic Anhydrase Structure of Zn−Histidine Complex with Theoretically Predicted Structure of Zn(Imidazole)$_3$ Complexes. The structure in human native carbonic anhydrase of the Zn−histidine complex (Expasy database structure 1CA2 of Swiss 3D collection)[17] (see structure **IV**) can be compared with the theoretically calculated



structures (see section IIb) for $Zn(Im)_3OH_2^{2+}$ (structure **V**) and $Zn(Im)_3OH^+$ (structure **VI**). The X-ray structure is probably that of the hydrate since the Zn−O distance in **IV** is 2.087 Å, while that in the theoretical hydrate, **V**, is 2.145 Å, much longer than the Zn−O 1.878 Å in the hydroxide of **VI**.

(15) Instead of being a real property of the PES, a reviewer suggested that the "grainy" property might be due to the particular grid used in the DFT calculation.

(16) Stöckigt, D. *Chem. Phys. Lett.* **1996**, *250*, 387.

(17) (a) Appel, R. D.; Bairoch, A.; Hochstrasser, D. F. *Trends BioChem. Sci.* **1994**, *19*, 253. (b) Peitsch, M. C.; Stampf, D. R.; Wells, T. N. C.; Sussman, J. L. *Trends BioChem. Sci.* **1995**, *20*, 82. (c) Peitsch, M. C. Protein Modelling by E-mail *Bio/Technology* **1995**, *13*, 658. (d) Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins, Struct., Funct.* **1988**, *4*, 274.

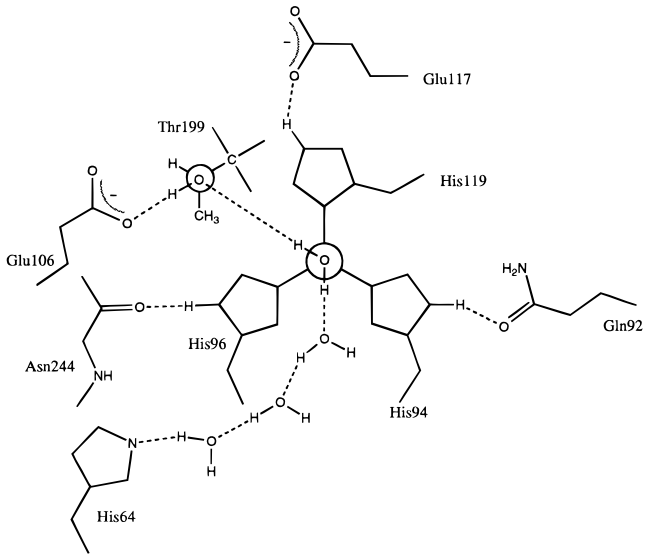


Figure 2. Amino acid residues and hydrogen bonds in the active center of human carbonic anhydrase. His-96, His-94, and His-119 are the three ligands of Zn^{2+}. The residues Asn-244, Glu-117, and Gln-92 form strong hydrogen bonds with hydrogens of the three imidazole (histidine) ligands. Four water molecules participate in a proton shuttle which transfers a proton from the water ligand to His-64 (adapted from Coleman[1,28]).

The Zn−N bonds to the imidazole nitrogens in the native **IV** are shorter than those in the theoretical hydrate **V**. The shorter native bonds can be attributed to strengthening of the bonds by the well-known[1−3] H-bonding interactions of the histidines with nearby residues (see Figure 2), which are as follows: His-119 with Glu-117, His-96 with Asn-244, and His-94 with Gln-92. By far the strongest H bond is expected to be with the negative carboxylic group of Glu-117, and this is reflected also in the corresponding Zn−N bond with the His-119 nitrogen, which is the shortest 1.907 Å of the three Zn−N bonds in the native **IV**.

The N−Zn−N bond angles in **IV** are smaller than 120° so that the structure is close to tetrahedral. For the calculated structure **V**, the bond angles are closer to 120° so that a more planar structure results. This flattening out is probably due to repulsions between the three partially positive imidazoles. The repulsion in the gas phase is expected to be larger than is the case in the environment provided by the enzyme.

The Zn−N bonds in the theoretical structure for the hydrate **V,** which are in the average ∼2.009 Å, are considerably shorter than those in the theoretical hydroxide (**VI**), which are on average ∼2.075 Å. Also the N−Zn−N bond angles in the hydroxide are closer to tetrahedral. These two effects reflect the weaker Zn−N bonding in the hydroxide and the lesser repulsion between the imidazoles, both effects being due to the lower positive Zn charge in the hydroxide, due to electron donation from the hydroxide group.

One other important feature in **VI** is the distorted position of the hydroxide itself. The two electron pairs of the hydroxide oxygen seem to undergo a long-range interaction with a slightly positive hydrogen on two imidazole rings (interaction distances between the oxygen and hydrogen are 2.340 and 2.689 Å—longer than traditional hydrogen bonds but close enough to provide some additional charge stabilization). Due to this interaction, the O−Zn−N angle to the imidazole ring with the closer hydrogen is 98.0° and to the second imidazole ring is 106.1°. The O−Zn−N angle to the third imidazole ring is 124.8°. The hydroxide interaction with two of the imidazoles

Table 1. Theoretical Sequential and Total Binding Energies for Zn^{2+}-Imidazole Complexes[a]

	ΔE_{elec} (kcal/mol)	ΔH°_{298} (kcal/mol)	ΔS°_{298} (cal/mol·K)	ΔG°_{298} (kcal/mol)
1,0 $ZnIm^{2+} = Zn^{2+} + Im$	181.8	180.2(169.3)[b]	28.0	171.8
2,1 $Zn(Im)_2^{2+} = Zn(Im)^{2+} + Im$	128.2	126.6	36.3	115.7
3,2 $Zn(Im)_3^{2+} = Zn(Im)_2^{2+} + Im$	64.9	63.0	34.5	52.7
3,0 $Zn(Im)_3^{2+} = Zn^{2+} + 3Im$	374.9	369.9	98.8	340.2
$Zn(Im)_3^{2+}(H_2O) = Zn(Im)_2^{2+}(H_2O) + Im$	55.2	52.8	35.3	42.3

[a] Present work at the B3LYP/6311++G(d,p) level, unless otherwise noted. [b] Garmer and Gresh.[7]

draws those three groups closer together leaving more space, i.e., less ligand−ligand repulsion, for the third imidazole, which in turn can interact more strongly with the Zn cation center, indicated by the shorter bond distance (2.064 Å vs 2.076 and 2.080 Å). This tendency of the hydroxide group to undergo strong enough long-range interactions can be also of some significance for the enzyme active site where the hydroxide group could be able to pull in other nearby groups. Due to the rigid structure of the histidines (His-94, His-96, His-119), who even in the absence of zinc barely change their structural positions (indicated by the apoenzyme X-ray crystal structure[17b]), possible interactions of the hydroxide group would likely occur with other residues or solvent molecules and eventually with the CO_2 molecule to form HCO_3^-, according to the accepted mechanism (vide infra).

III. Function of Strongly Bonding Ligands To Make the Metal Ion Stable in the Enzyme Relative to the Aqueous Environment

In a recent study based on gas-phase ion−ligand equilibria determinations of CuL_2^+ complexes with a variety of ligands L, we found[18] that imidazole, which can serve as a model of the histidine residue, leads to the strongest bonding with Cu^+ out of some 22 different ligands including all ligands used to model the remaining neutral amino acid residues. We considered this finding as very significant because the histidine residue is so often present in metallo complexes of biological origin. Imidazole leads also to the strongest bonding with the doubly charged Zn^{2+} ion. This is not surprising because these two ions are isoelectronic. High-level ab initio calculations by Garmer and Gresh,[7] undertaken in a study of the applicability of the hard and soft acid and base (HSAB) principle, provide the bond energies for Zn^{2+}−L, with 11 ligands L, many of which model amino acid residues. The highest binding energy for all uncharged (neutral) ligands was observed for L = imidazole.

The strong bonding achieved with imidazole suggests that one purpose for the use of this ligand in metalloenzymes is to make the metal ion in the enzyme (En′) stable relative to the cytoplasm, i.e., the aqueous solution. To examine more closely this proposition, CA II was chosen, for which all three primary ligands are histidine residues (see **III**). Ab initio calculations using the density functional theory (B3LYP/6-311++G**) performed in this work provide the three successive bond energies for L = Im shown in Table 1. For more details see section IIb. These results lead to

$$\Delta H_{1,0}\,(Zn^{2+}\text{−imidazole}) \approx 180\ \text{kcal/mol} \qquad (2)$$

The sum of the three sequential dissociations provides the total dissociation energy $\Delta H_{3,0}$ of $Zn(Im)_3^{2+}$, of 370 kcal/mol. The sequential and total free energy changes were also obtained. The total free energy change is given below.

Table 2. Model for Interaction of Zn^{2+} with Ligands, Protein, and Aqueous Solution for CAII

type of interaction	free energy change $\Delta G^\circ_{3,0}$ (kcal/mol)[a]	
(1) $Zn(Im)_3^{2+} = Zn^{2+} + 3Im$	340	Table 1
(2) $Zn(Im)_3H_2O^{2+} = Zn(Im)_3^{2+} + H_2O$	13	Table 3A
(3) noncovalent interactions, i.e., strong hydrogen bonds induced by Zn^{2+} charge: His-94 to Gln-92; His-96 to Asn-244, His-119 to Glu-117, and H_2O ligand to Thr-199	55	see Appendix I
(4) solvation by protein	34	see Appendix I
(5) solvation by aqueous environment	26	see Appendix I
(6) solvation by conical cleft in caii	9	see Appendix I
total	477	

[a] For error estimates, see Appendix IB.

$$\Delta G^\circ_{3,0}\,(ZnIm_3^{2+}) \approx 340\ \text{kcal/mol} \qquad (3)$$

This value can be compared with the hydration free energy of the free ion Zn^{2+} in aqueous solution, $\Delta G^\circ_h\,(Zn^{2+})$. Shown below are hydration data for Zn^{2+} and several other ions, given for comparison.

$$\Delta G^\circ_h = Cu^+\,(-128),\ Cu^{2+}\,(-482),\ Mg^{2+}\,(-440),\ Ca^{2+}\,(-362),\ Zn^{2+}\,(-470)\ (\text{kcal/mol})^{19} \qquad (4)$$

The values given in parentheses represent the free energy released when 1 mol of ions is transferred from the gas phase to the aqueous phase.[19,20] To make the ion thermodynamically stable in the biocomplex, the free energy released by the total interactions of the ion with ligands and remaining environment, i.e., the protein and the aqueous medium surrounding the protein, should be similar or larger than the Zn^{2+} interaction with the aqueous environment, i.e., $-\Delta G^\circ_h\,(Zn^{2+})$.

The estimated values of the interactions providing additional stabilization of the $Zn(Im)_3^{2+}$ complex in the enzyme in aqueous solution are listed in Table 2. Justification for the chosen energy values is given in Appendix IA. Here we provide a brief description of the energy terms that were considered. Additional stabilization of the complex is provided by the fourth ligand, H_2O, for which a bonding free energy of 13 kcal/mol was obtained in the present work (see experimental section and Table 3A). Hydrogen bonding between hydrogen atoms of the histidine ligands and nearby polar groups of the protein (Figure 2), such as Gln-92 with the His-94 hydrogen and Glu-117 with the hydrogen of His-119, and Asn-244 with the hydrogen of His-96 and the fourth ligand H_2O with Thr-199 (see Figure 2), adds further stability to the complex. The hydrogen bonds are

(18) Deng, H.; Kebarle, P. *J. Am. Chem. Soc.* **1998**, *120*, 2925.

(19) Marcus, Y. *Ion Solvation*; John Wiley & Sons: New York, 1985; Table 5.10, p 107.

(20) (a) The absolute free energies of hydration $\Delta G^\circ_h(M^{z+})$ and for negative ions $\Delta G^\circ_h(X^{z-})$ are obtained from conventional free energies of hydration for these ions based on measurements of galvanic cell standard potentials and supporting data such as ionization energies of M. electron affinities of X, and free energies of sublimation and evaporation leading to the free energies of the ions in the gas phase. The conventional free energies of hydration are available in standard compilations.[13b] These experimental data lead to values for the sum of the absolute hydration energies for an ion pair such as $\Delta G^\circ_h(Na^+)$ and $\Delta G^\circ_h(Cl^-)$ for NaCl or $\Delta G^\circ_h(H^+)$ and $\Delta G^\circ_h(Cl^-)$ for HCl. All single-ion absolute values are then obtained[19] from the above sum and an extra thermodynamic assumption which selects the absolute value of one ion (such as H^+). (b) U.S. National Bureau of Standards NBS Tables of Chemical Thermodynamic Properties *J. Phys. Chem. Ref. Data* **1984**, *11* (Suppl. 2).

Table 3. Reaction Energies

A. Bond Energies[a] for Dehydration Reactions: $M(L)_x(H_2O)^{2+} \rightarrow M(L)_x^{2+} + H_2O$

	ΔE_{elec}	ΔH_0	ΔH°_{298}	ΔS°_{298}	ΔG°_{298}
$Ca(H_2O)^{2+}$	57.3	55.4(56.9)[b]	56.5(53.0)[c]	23.4	49.5
$Mg(H_2O)^{2+}$	82.7	80.6(81.5)[b]	81.8(78.2)[c]	24.2	74.6
$Mg(H_2O)_2^{2+}$	73.9	71.5(70.9)[b]	72.0(68.9)[c]	28.9	63.4
$Mg(H_2O)_3^{2+}$	58.6	55.9(55.1)[b]	56.6(56.6)[c]	26.6	48.7
$Zn(H_2O)^{2+}$	103.6	102.0 (101.9)[b]	103.1	22.6	96.3
$Zn(H_2O)_2^{2+}$	90.0	87.3(86.6)[b]	88.0	33.4	78.1
$Zn(H_2O)_3^{2+}$	57.7	55.3(53.6)[b]	55.8	25.2	48.3
$Zn(H_2O)_4^{2+}$	44.7	42.6(41.2)[b]	42.9	28.9	34.3
$ZnOH(H_2O)_3^{+}$	24.7	22.6	22.4	28.9	13.8
$ZnOH(H_2O)_3^{+}$					9.7,[e] 10.4[f]
$ZnOH(H_2O)_4^{+}$	22.1	19.3	19.9	37.3	8.8
$ZnOH(H_2O)_4^{+}$					6.6,[i] 6.8[j]
$Zn(Im)(H_2O^{2+}$	75.3	73.5	74.0	30.5	64.9
$Zn(Im)_2(H_2O)^{2+}$	32.2	30.1	30.5	29.5	21.7
$Zn(Im)_3(H_2O)^{2+}$	22.5	20.6	20.9	30.3	11.9, 14.0[d]
$Zn(Im)_3(H_2O)^{2+}$					7.5,[g] 9.7[h]
$Zn(Im)_3(H_2O)_2^{2+}$					6.3[d]
$Zn(Im)_2(CH_3CO_2)(H_2O)^{+}$					5.0[d]

B. Energies[k] for Reactions: (14a) $M(L)_x(OH)^{+} \rightarrow M(L)_x^{2+} + OH^{-}$ and (14b) $M(L)_x(OH)^{+} + CO_2 + H_2O \rightarrow M(L)_xOH_2^{2+} + HCO_3^{-}$

	ΔE_{elec}(14a)	ΔH_o(14a)	ΔH°_{298}(14a)	ΔS°_{298}(14a)	ΔG°_{298}(14a)	ΔG°_{298}(14b)
$Ca(OH)^{+}$	338.9	336.3	337.4	23.6	330.3	243.3
$Mg(OH)^{+}$	384.3	382.4	383.3	20.8	337.1	265.0
$Mg(H_2O)(OH)^{+}$	354.7	352.2	352.5	25.6	344.9	244.0
$Mg(H_2O)_2(OH)^{+}$	320.7	318.0	318.5	25.0	311.0	224.8
$ZnOH^{+}$	436.5	434.3	435.4	21.1	429.1	295.3
$Zn(H_2O)OH^{+}$	391.0	387.9	388.5	28.2	380.1	264.5
$Zn(H_2O)_2(OH)^{+}$	331.9	329.1	329.5	22.9	322.7	236.9
$Zn(H_2O)_3(OH)^{+}$	298.8	296.4	296.7	26.7	288.8	217.0
$Zn(Im)OH^{+}$	351.2	348.9	349.4	28.3	341.0	238.6
$Zn(Im)_2(OH)^{+}$	266.6	264.2	264.7	27.9	256.4	197.2
$Zn(Im)_3(OH)^{+}$	230.6	*228.2*[l]	*228.7*	*27.9*	*220.4*	*171.0*

C. Energies[k] for Deprotonation Reactions: (13a) $M(L)_x(H_2O)^{2+} \rightarrow M(L)_x(OH)^{+} + H^{+}$ and (13b) $M(L)_x(H_2O)^{2+} + 2\ H_2O \rightarrow M(L)_x(OH)^{+} + (H_2OHOH_2)^{+}$

	ΔE_{elec}(13a)	ΔH_o(13a)	ΔH°_{298}(13a)	ΔS°_{298}(13a)	ΔG°_{298}(13a)	ΔG°_{298}(13b)
$Ca(H_2O)^{2+}$	114.4	107.1	107.4	21.8	100.9	−61.3
$Mg(H_2O)^{2+}$	94.4	86.3	877	25.4	80.1	−82.1
$Mg(H_2O)_2^{2+}$	115.2	107.3	108.7	25.3	101.2	−61.0
$Mg(H_2O)_3^{2+}$	133.9	125.9	127.3	23.7	120.3	−41.9
$Zn(H_2O)^{2+}$	63.1	55.7	56.9	23.5	49.9	−112.3
$Zn(H_2O)_2^{2+}$	95.0	87.4	88.7	27.3	80.6	−81.6
$Zn(H_2O)_3^{2+}$	121.9	114.2	115.5	24.3	108.2	−54.0
$Zn(H_2O)_4^{2+}$	141.9	134.2	135.4	24.3	128.1	−34.1
$Zn(Im)(H_2O)^{2+}$	120.1	112.6	113.8	24.3	106.5	−55.7
$Zn(Im)_2(H_2O)^{2+}$	161.6	153.9	155.0	23.8	148.0	−14.2
$Zn(Im)_3(H_2O)^{2+}$	187.9	*180.2*[m]	*181.3*	*23.8*	*174.3*	*12.1*

[a] Present results from theoretical calculations (see section IIb), except when otherwise noted. All energy values in kcal/mol, Entropy in cal/(deg·mol). [b] Literature values[9] in parentheses B3LYP/tvz//B3LYP/Lan12dz. [c] Literature values[16] in parentheses MP2/dvz/HF/dvz. [d] Experimental values at 298 K; see section IIa. [e] Theoretical value at 440 K. [f] Experimental value at 440 K. [g] Theoretical value at 442 K. [h] Experimental value at 442 K. [i] Theoretical value at 357 K. [j] Experimental value at 357 K. [k] All values based on theoretical calculations; see section IIb, present work. [l] Italics in this row indicate that the thermal corrections were taken from the reactions for $Zn(Im)_2(OH)^{+}$. Examination of the thermal corrections shows that for similar systems they are almost identical (within 0.1−0.3 kcal/mol) and therefore can be used as a very good approximation for a similar system for which frequency information is not available (see section IIb). [m] Italics in this row indicate that the thermal corrections were taken from the reactions for $Zn(Im)_2(H_2O)^{2+}$. Examination of the thermal corrections shows that for similar systems they are almost identical (within 0.1−0.3 kcal/mol) and therefore can be used as a very good approximation for a similar system for which frequency information is not available (see section IIb).

expected to be strengthened by the protic character of the histidine hydrogens induced by the electron transfer from the histidines to the Zn^{2+} core ion. For the sum of these hydrogen-bonding interactions, a total free energy of 55 kcal/mol was used (see Appendix IA).

The surrounding protein provides also a nonspecific stabilization of the charge of the ion−ligand complex. This solvation-type energy can be estimated with the Born charging equation, as shown in Appendix IA. We estimate (Appendix I) the ion−ligand complex to have a radius of $R = 7$ Å and assume the dielectric constant for the protein to be $D = 2$. The protein extends from $R = 7$ Å to an outer radius of $R = 25$ Å. With these values one obtains (Table 2) the result

$$-\Delta G^\circ_{sol}\ (\text{by protein}) = 34\ \text{kcal/mol} \qquad (5)$$

The aqueous medium outside the protein sphere of 25 Å provides an additional stabilization which can be evaluated with the Born equation, $R = 25$ Å, and the dielectric constant of water $D(H_2O) = 78$. This leads to

$$-\Delta G^\circ_{hyd}\ (\text{by } H_2O) = 26\ \text{kcal/mol} \qquad (6)$$

Carbonic anhydrase is not perfectly approximated by a sphere. In particular there is one approximately conical cleft ~15 Å deep, which allows access of the aqueous medium to the active site. The solvation contribution to this aqueous cleft is difficult to estimate. We propose (Appendix IA) a value of

$$-\Delta G^\circ_{\text{hyd (cleft)}} = 9 \text{ kcal/mol} \tag{7}$$

The total stabilization is

$$-\Delta G^\circ \text{ (stab)} = 477 \pm 27 \text{ kcal/mol} \tag{8}$$

For the origin of estimated error, see Appendix IB.

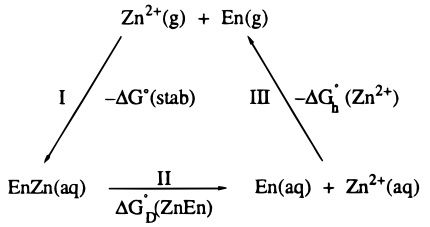
The total stabilization energy ΔG°(stab) is seen to be somewhat larger than the free energy due to the hydration of Zn^{2+}, i.e. $-\Delta G^\circ_h(Zn^{2+}) = 470$ kcal/mol. This is in qualitative agreement with the known stability of Zn^{2+} in the enzyme in aqueous solution. Kiefer and Fierke[3a] used the wild type of human carbonic anhydrase as well as variants in which one of the histidine ligands was replaced with different residue. For the wild type, the Zn^{2+} dissociation constant was found to be

$$K_D(\text{En'Zn}) = 4 \times 10^{-12} \text{ M} \qquad (\text{pH } 7,\ T = 303 \text{ K})$$

which corresponds to

$$\Delta G^\circ_D(\text{ZnEn'}) = -RT \ln K_D = 15.8 \text{ kcal/mol}$$

The ΔG°_D(ZnEn′) can be related to $\Delta G^\circ_h(Zn^{2+})$ and ΔG°_h (stab), via the thermodynamic cycle



(9)

It will be noted that step I, ΔG°(stab), does not include any interactions of the En′Zn with the aqueous medium that are not due to the +2 charge; i.e., the interactions of En′ with the water were not considered (Table 2). Similarly, in step III, interactions of En′ with the water are again not included. This procedure is justified because the two omitted terms are approximately equal and are expected to cancel in the result from the cycle

$$\Delta G^\circ_D(\text{ZnEn'}) = \Delta G^\circ_h(Zn^{2+}) - \Delta G^\circ(\text{stab})$$
$$= -470 + 477 = 7 \text{ kcal/mol}$$

The relatively close agreement with the Kiefer and Fierke[3a] experimental result ΔG°_D(ZnEn) = 15.8 kcal/mol, should not be given much significance because of the large uncertainty in the estimates in the present work used to evaluate ΔG°(stab). The significance of the result is that with plausible assumptions a qualitative agreement with experiment[3a] can be achieved with the crude model used. The results in Table 2 provide some indication of the relative importance of the various stabilization energies. The most important term, some 70% of the total, is the stabilization due to the bonding of Zn^{2+} with the three imidazole ligands. The strong hydrogen bonds of imidazole hydrate hydrogens with nearby protein residues lead to some 12% of the stabilization while the nonspecific solvation by the protein and water provides the remaining 18%.

It is interesting to compare the above results with data obtained by Kiefer and Fierke[3a] and Ipolito and Christianson,[3c] which dealt with changes of Zn affinity[3a] and structural changes[3c] on site-directed mutagenesis when one histidine of CAII is replaced by another residue. The replacement with the essentially non-Zn-bonding methyl residue of alanine leads[3a] to a large ~10^5 fold decrease of Zn affinity which corresponds to a free energy bonding decrease of ~7 kcal/mol. Assuming that the bonding to methyl is negligible, the gas-phase reaction,

$$Zn(Im)_3(H_2O)^{2+} = Zn(Im)_2(H_2O)^{2+} + Im \qquad \Delta G^\circ_{10a} = 42 \text{ kcal/mol} \tag{10a}$$

whose energy can be evaluated from the present theoretical calculations, provides an upper limit for the decrease in Zn affinity. Since there is no change in the charge, changes due to the nonspecific solvation by the protein and aqueous medium cannot be expected to provide much stabilization. However, if there are no suitable residues that could lead to tetracoordination with the Zn^{2+} in the mutant, then at least one water molecule can be expected to get involved. Therefore, the reaction energy of the exchange reaction

$$ZnIm_3(H_2O)^{2+} + H_2O = ZnIm_2(H_2O)_2{}^{2+} + Im \qquad \Delta G^\circ_{10b} \approx 24 \text{ kcal/mol} \tag{10b}$$

should be set as upper limit. The ΔG°_{10b} value was estimated on the basis of the known values for the dehydration reactions for $ZnIm_2(H_2O)^{2+}$ and $Zn(H_2O)_4{}^{2+}$, which are 22 and 34 kcal/mol (see Table 3A), and suggest a value of 18 kcal/mol for the dehydration of $ZnIm_2(H_2O)_2{}^{2+}$. The latter combined with eq 10a leads to ΔG°_{10b}. The ΔG°_{10b} obtained is still too high; however, the enzyme structures exhibit a certain conformational plasticity[3c] which can lead to further specific stabilization. Unfortunately an X-ray structure for the mutant His, Ala could not be obtained,[3c] because the low Zn^{2+} affinity led on crystallization to a metal free active site,[3c] i.e., to the apoenzyme of the mutant.

IV. Tuning of the Bond Energy of the Ligand Involved in the Catalytic Process by Selection of the Remaining Ligands

We assume that the bonding of the metal to the ligand involved in the catalyzed reaction must be of a given magnitude for optimal catalytic effect. For example, we assume that for the three Zn-containing enzymes **I**−**III**, where H_2O is the ligand at the site of catalysis, the remaining three directly bonding ligands (the primary ligands in our nomenclature) are selected to control the bonding of the H_2O and the energies of subsequent transition states involving the H_2O site, so as to achieve lowest transition-state(s) energies. We will call the control obtained by selection of the primary ligands the "primary tuning". This is to distinguish it from the well-known and very important additional stabilization of transition states by the participation of nearby residues that might hydrogen bond to the active ligand site, which is known as "fine tuning".[1,3]

Sequential bond energies for ion−ligand bonding in the gas phase obtained from ab initio calculations and gas-phase ion−ligand equilibria, summarized in Table 3A, provide a good illustration of the changes in bonding for an H_2O ligand as the number of other ligands is increased and as these ligands are changed.

For Zn^{2+} and H_2O as the only ligand type, the bond energy is observed to decrease from $\Delta H^\circ_{1,0} = 103$ kcal/mol to $\Delta H_{4,3} =$

43 kcal/mol. The decrease of bond energy with increasing n has been attributed[8,9a] to charge transfer from the ligand to the core ion, ligand−ligand repulsion due to dipole−dipole and Pauli repulsion, and less effective ion core polarization. For the tetracoordinated $ZnL_3H_2O^{2+}$, both charge transfer from the ligands to the core ion and ligand−ligand repulsion should be very important in lowering the bond energy of the H_2O ligand. Imidazole, which has a large charge transfer and polarizability term (see energy decomposition analysis by Garmer and Gresh[7]), is expected to lead to a large bond weakening in $ZnIm_3^{2+}$−H_2O. Ligand−ligand repulsion should also be important. Pavlov and co-workers[9a] have shown that, even for the hydrates $Zn(H_2O)_4^{2+}$, ligand repulsion must be significant. Thus for the pentahydrate, the $Zn(H_2O)_4(H_2O)^{2+}$ species in which one water molecule was in an outer shell was found to be of lower energy than the $Zn(H_2O)_5^{2+}$ species, in which all water molecules were directly interacting with the ion.[9a] Since the imidazole molecules are bigger, ligand−ligand repulsions in the $ZnIm_3(H_2O)^{2+}$ should be also very important, as discussed in section IIc. These considerations are supported by the ΔH°_{298} changes (from Table 3A), which provide a comparison between the bond enthalpies to Zn^{2+}, when H_2O molecules or imidazoles are the primary ligands:

$$Zn(H_2O)(H_2O)^{2+} \rightarrow Zn(H_2O)^{2+} + H_2O \quad 88 \text{ kcal/mol}$$
$$Zn(Im)(H_2O)^{2+} \rightarrow Zn(Im)^{2+} + H_2O \quad 74 \text{ kcal/mol}$$

$$Zn(H_2O)_3(H_2O)^{2+} \rightarrow Zn(H_2O)_3^{2+} + H_2O \quad 43 \text{ kcal/mol}$$
$$Zn(Im)_3H_2O^{2+} \rightarrow Zn(Im)_3^{2+} + H_2O \quad 21 \text{ kcal/mol}$$

The result for the three primary ligands is particularly striking, showing that the replacement of water with imidazole decreases the bond energy for the H_2O molecule from 43 to 21 kcal/mol. These data demonstrate that the choice of the primary ligands has a profound effect on the bonding of the active ligand.

We will consider below what the carbonic anhydrase catalysis requirements are and how they may affect the choice of the primary ligands. Carbonic anhydrase catalyzes the reaction

$$CO_2 + H_2O = H_2CO_3 \quad (11a)$$

$$H_2CO_3 + H_2O(aq) = H_3O^+ + HCO_3^- \quad (11b)$$

in an aqueous medium. The reaction requires catalysis because the first step (11a) is associated with a large energy barrier.[21−23] The second step of the reaction in aqueous solution (11b) is very fast in both the forward and reverse directions, compared to (11a), and therefore the reactions producing H_2CO_3 or H_3O^+ and HCO_3^- in aqueous solution from aqueous CO_2 are functionally equivalent steps from the standpoint of the catalysis.

The free energy for the reaction

$$OH^- + CO_2 = HCO_3^- \qquad \Delta H^\circ_{12} = -46 \text{ kcal/mol}$$
$$\Delta G^\circ_{12} = -37.5 \text{ kcal/mol} \quad (12)$$

will be needed for the subsequent analysis (vide infra). The above values were obtained with density functional theory in the present work; see section IIb.

The reaction catalyzed by CAII is extremely fast, having a CO_2 turnover rate[1−4] of $k = 10^6$ s^{-1}. The generally accepted mechanism[1−4] involves two steps. The deprotonation of the H_2O ligand

$$(En'ZnH_2O)^{2+}(aq) \rightarrow (En'ZnOH)^+(aq) + H^+(aq) \quad (13)$$

followed by the formation of the bicarbonate anion by a nucleophilic attack of the hydroxide group in $(En'ZnOH)^+$ on CO_2.

$$(En'ZnOH)^+(aq) + CO_2 \rightarrow (En'ZnHCO_3)^+(aq) \xrightarrow{H_2O} (En'ZnH_2O)^{2+}(aq) + HCO_3^-(aq) \quad (14)$$

The sum of reactions 13 and 14 corresponds to the formation in the aqueous phase of H_3O^+ and HCO_3^- from H_2O and CO_2. The deprotonation of $En'ZnH_2O$, eq 13, is facilitated by a "proton shuttle" mechanism[1−3] involving a chain of water molecules connecting the hydroxide to the nearby His-64 residue; see Figure 2. The assistance of the His-64 group is an example of "fine tuning". The occurrence of the bicarbonate-forming reactions is also assisted by fine-tuning. In that case, several residues are believed to be involved (vide infra). The participation of such residues is well documented, particularly so with mutants in which the residue in question is replaced with another residue and the effect of the replacement is established by determination of the changes in the rates of the catalyzed reaction.[2,3]

The effect of the primary histidine ligands relative to other possible ligands, on the rates of the catalyzed reaction and a demonstration that the histidine ligands lead to an optimal result has received far less attention in the theoretical work.[4] Experimental work based on site-directed mutagenesis in which the directly bonding His ligands are replaced by other ligands is available.[3] Some of the findings will be considered below in conjunction with the present results.

With the relatively high level basis sets used in the present calculations, it is not feasible to obtain the structures and energies of the transition states even in the absence of the protein environment of the enzyme. Therefore, the arguments will be based on evaluated energies for reactions 13 and 14 where the enzyme is replaced just by the three primary ligands. The stabilizing effect of the environment, i.e., the enzyme protein and the aqueous solution surrounding the reactions of eqs 13 and 14, is introduced subsequently. The use of the free energy change of the reactions, rather than the transition-state energies, is still relevant because, when the reaction free energies are positive, they set a lower limit for the sum of the activation free energies of the consecutive steps that lead to the reaction products.

The calculated energies for the reactions in the gas phase

$$L_xZnOH_2^{2+} = L_xZnOH^+ + H^+ \quad (13a)$$

$$L_xZnOH^+ = L_xZn^{2+} + OH^- \quad (14a)$$

are given in sections B and C of Table III. Also given in this table are the energies for the gas-phase reactions b,

$$L_xZnOH_2^{2+} + 2H_2O = L_xZnOH^+ + (H_2OHOH_2)^+ \quad (13b)$$

$$L_xZnOH^+ + CO_2 + H_2O = L_xZnOH_2^{2+} + HCO_3^- \quad (14b)$$

which model somewhat better the reactions 13 and 14. The

(21) Jönsson, B.; Karström, G.; Wennerström, H.; Roos, B. *Chem. Phys. Lett.* **1976**, *41*, 317. Jönsson, B.; Karström, G.; Wennerström, H.; Forsen, S.; Roos, B.; Almlöf, J. *J. Am. Chem. Soc.* **1977**, *99*, 4628.
(22) Nguyen, M. T.; Ha, T. K. *J. Am. Chem. Soc.* **1984**, *106*, 599.
(23) Wright, C. A.; Boldyrev, A. I. *J. Phys. Chem.* **1995**, *99*, 12125.

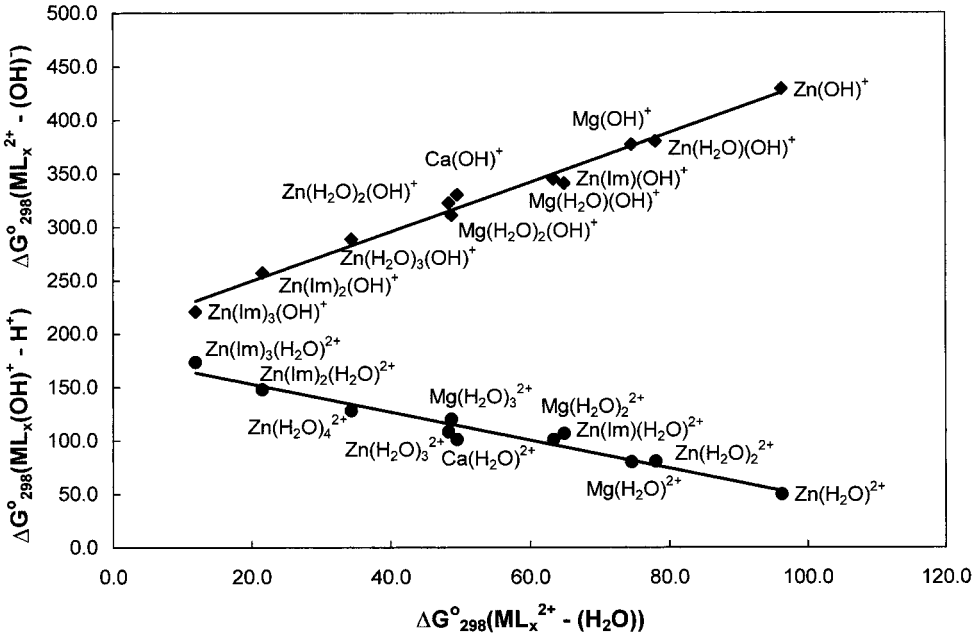


Figure 3. (Upper plot) Free energies for the dehydroxylation reaction, $L_xZnOH^+ = L_xZn^{2+} + OH^-$, plotted versus the free energies for the dehydration reaction, $L_xZnOH_2^{2+} = L_xZn^{2+} + OH_2$. A linear correlation is observed. (Lower plot:) Free energies for deprotonation reaction $L_xZnOH_2^{2+} = L_xZnOH^+ + H^+$, versus the free energy of the dehydration reaction. An inverse proportionality is observed for the deprotonation. The ligands L are identified in the figure. Data for some metal ions such as Mg^{2+} and Ca^{2+} are also included and are seen to fit in the plots.

energies of reactions 13a and 13b differ by a constant term which for the ΔG°_{13} values corresponds to the gas-phase basicity (= the deprotonation free energy) of H_2O of 137.2 kcal/mol[24] plus the dehydration free energy of $H_2OHOH_2^+$, which is 25 kcal/mol,[25] for a total of 162.2 kcal/mol. This partial inclusion of solvation effects on the ions illustrates the large effect of ion solvation on the reaction energies and the absolute necessity to include the effect of the environment. The free energies of reactions 14a and 14b differ by the following term: The free energy of the monohydration of $Zn(L)_x^{2+}$ from Table 3A and the free energy of reaction 12, $\Delta G^\circ_{12} = -37.5$ kcal/mol.

The changes of the free energies for reactions 13a and 14a as a function of the ligands used, are illustrated in Figure 3. The plot of the dehydroxylation energy $ZnL_x^{2+}-OH^-$ versus the dehydration free energy

$$L_xZnOH_2^{2+} = L_xZn^{2+} + H_2O$$

leads to an approximately linear correlation. This is expected because, in both cases, a Zn−O bond is cleaved. The energies decrease as the charge transfer from the other ligands increases. Again, this is expected because the stabilization provided by the OH_2 or the OH^- becomes less important as the stabilization provided by charge transfer from the L_x ligands increases.

The energies for the deprotonation reaction, $ZnL_xOH^+-H^+$, versus those for the dehydration reaction, $ZnL_x^{2+}-OH_2$ (Figure 3) lead to an approximate *inverse* proportionality. The deprotonation energy increases as the dehydration energy decreases. An electron-donating ligand reduces the electron transfer from the water molecule to the core ion and thus decreases the protic character of the hydrogens of the water molecule.

Since both reactions 13 and 14 must be fast, the choice of the ligands, i.e., the primary tuning, must lead to a compromise between the conflicting requirements of reactions 13 and 14. The proper balance would be found when the free energies for both reactions are negative or, if positive, then within less than 10 kcal/mol above zero.

Assuming that the Im_3 primary ligand composition is a best choice for this compromise, the free energies of the reactions (Table 3), in the gas phase,

$$Im_3ZnOH_2^{2+} = Im_3ZnOH^+ + H^+ \qquad \Delta G^\circ = 174.3 \text{ kcal/mol} \quad (13c)$$

$$Im_3ZnOH^+ + CO_2 + H_2O = Im_3ZnOH_2^{2+} + HCO_3^- \qquad \Delta G^\circ = 171.0 \text{ kcal/mol} \quad (14c)$$

when corrected for interactions with the protein environment and the aqueous solution, should each be equal to less than 10 kcal/mol.

The values obtained with these environment corrections are

$$En'(Im)_3ZnOH_2^{2+}(aq) = En'(Im)_3ZnOH^+(aq) + H^+(aq) \qquad \Delta G^\circ_{13d} = +14 \text{ kcal/mol} \quad (13d)$$

$$En'(Im)_3ZnOH^+(aq) + CO_2(aq) + H_2O(aq) = En'(Im)_3ZnOH_2^{2+}(aq) + HCO_3^-(aq) \qquad \Delta G^\circ_{14d} = -5 \text{ kcal/mol} \quad (14d)$$

The small product ions $H^+(aq)$ and $HCO_3^-(aq)$ are assumed to have left the active site and entered the aqueous solution.

The thermochemical equations and procedures used to obtain the free energies 13d and 14d from the gas-phase reactions 13c and 14c and the required thermochemical data[19,26,27] are detailed in Appendix II. The evaluation of the solvation effect of the protein and aqueous environment on $Im_3ZnOH_2^{2+}$ and Im_3-

(24) Hunter, E. P. L.; Lias, S. G. Evaluated Gas-Phase Basicities and Proton Affinities of Molecules: An Update. *J. Phys. Chem. Ref. Data* **1998**, *27*.

(25) Lau, Y. K.; Ikuta, S.; Kebarle, P. *J. Am. Chem. Soc.* **1982**, *104*, 1462.

(26) The hydration free energy $\Delta G^\circ_h(HCO_3^-)$ is not available;[19] however, $\Delta H^\circ_h(HCO_3^-) = -94$ kcal/mol and $S^\circ(HCO_3^-, aq) = 27$ cal/(deg·mol) are available.[19] The theoretical calculation (present work) of ΔG°_{12} for reaction eq 12, included an evaluation of $S^\circ(HCO_3^-, g) = 63.8$ cal/

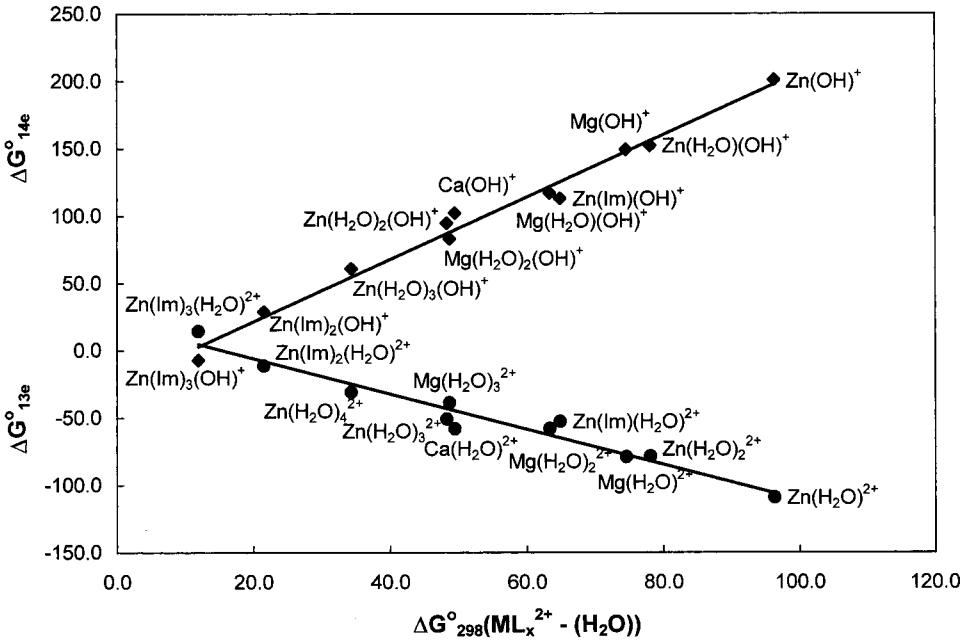


Figure 4. Plot of free energy for the reactions, $En'L_xZnOH_2^{2+}(aq) = En'L_xZnOH^+ + H^+(aq)$, ΔG°_{13e} and $En'L_xZnOH^+(aq) + CO_2 + H_2O = En'L_xZnOH_2^{2+}(aq) + HCO_3^-(aq)$, ΔG°_{14e}. The two data points, where $L_x = (Im)_3$, correspond to reactions 13d and 14d, whose energies model approximately the reactions of the CAII enzyme in aqueous solution. Free energies for other ligands are also given. The corrections for the effect of the environment for the other ligands were taken to be the same as those for (13d) and (14d). The plot shows that only the three imidazole ligands lead to free energies which are close to zero for *both* reactions.

$ZnOH^+$, which is the most difficult step, was obtained by a procedure closely analogous to that used for the evaluation of the stability of the Zn^{2+} ion in $En'ZnOH_2^{2+}$ relative to aqueous solution; see section III. The success of the procedure to provide a result in agreement with the experimental Zn stability constant (see section III) provides some confidence for its use also in the present evaluation (Appendix II).

The free energies for (13d) and (14d) are close enough to zero to be consistent with the assumed mechanism[1−3] for CAII. Shown in Figure 4 are the calculated values for reactions 13 and 14, which include also the other ligands L (Table 3). The reactions have been corrected to include solvation energy terms. For simplicity, the solvation energy terms were assumed to be the same for all ligands. The plot in Figure 4 illustrates that the "choice" of the three ligands to be $(Im)_3$ is "just right". It is these ligands that bring the reaction free energies for both reactions, 13 *and* 14, closest to zero.

The dehydration energy for $Zn(Im)_2(CH_3CO_2)H_2O^+$, which models the ligand composition of the reactive center of carboxypeptidase (see structure **II** in the Introduction), was determined experimentally in the present work (Table 3A), $\Delta G^\circ_{298} = 5.0$ kcal/mol. It is much lower than the energy for $Zn(Im)_3(H_2O)^{2+}$, which is 13 kcal/mol (Table 3A). Assuming that the $Zn(Im)_2(CH_3CO_2)^+$ hydrate could be included in the plot given in Figure 4, one would come to the conclusion that the carboxypeptidase ligand composition is less well suited to be used for the carbonic anhydrase catalysis because the Glu ligand, while destabilizing further the Zn−OH^- bond, raises the deprotonation energy of the Zn hydrate into the positive free energy range, which should slow the overall reaction. This conclusion is in line with the assumption of the primary tuning hypothesis that the choice of three His ligands is best. However, the energetics for the reactions 13c and 14c when one of the imidazole ligands is replaced with the acetate anion may be significantly different. An actual theoretical calculation of these reactions with the acetate ligand would be required to check whether the energies will fall on the straight lines (Figure 4). Such a calculation is more difficult than for the $Zn(Im)_3OH_2^{2+}$ system and has not been carried out yet.

Kinetic measurements[3c] with CAII variants produced by site-directed mutagenesis, in which one of the primary His ligands was replaced by a negatively charged residue such as glutamate or aspartate, have shown that these mutants have very much lower catalytic activity than the wild type. It was also deduced[3a] that these mutant ligands destabilize the Zn−OH^- bonding and increase the deprotonation energy of the hydrate in agreement with the predictions in Figure 4; see discussion above. However, it was also deduced that the negative residues increase the transition-state energy of the HCO_3^--forming reaction (equivalent to reaction 14). The present results deal only with the reaction energies and therefore provide no direct information on the transition states. Therefore, assuming that the deduction[3a] concerning the activation free energy for the HCO_3^- formation is correct, the observed decreased catalytic activity of the mutants[3a] will not be only due to the increased deprotonation energy of the Zn−hydrate, as the present results indicate.

In the following discussion we examine how the present results for the reaction energies (eqs 13d and 14d) can be correlated with results available in the literature, which deal with the transition-state energies.

The deprotonation reaction 13 has been modeled by Liang and Lipscomb[28] with the quantum mechanical PRDDO com-

(deg·mol), and this value was used to obtain $\Delta G^\circ_{hi}(HCO_3^-) = -83$ kcal/mol. The theoretical value for the entropy is consistent with the experimental values for similar molecules in the gas phase[27] such as $S^\circ(CO_2, g) = 51$ cal/(deg·mol), $S^\circ(HCO) = 53.7$ cal/(deg·mol), and $S^\circ(HCHO, g) = 52.3$ cal/(deg·mol).

(27) Wagman, D. D.; Evans, W. H.; Parker, V. B.; Hallow, I.; Bailey, S. M.; Schrumm, R. H. *Selected Values of Chemical Thermodynamic Properties*.; NBS Technical Note 270-3; U.S. Government Printing Office: Washington, DC, 1968.

(28) Liang, J. Y.; Lipscomb, W. N. *Biochemistry* **1988**, *27*, 8676.

putational method.[29] A schematic representation of the proton shuttle involving three water molecules bridging the gap from $(His)_3ZnOH_2^{2+}$ to the residue His-64 is shown in Figure 2. In the calculations,[28] the histidines were replaced by ammonia molecules. The authors[28] found an energy barrier only for the first proton transfer, i.e., from $(NH_3)_3ZnOH_2^{2+}$ to the first water molecule. The barrier obtained with PRDDO was quite big, 34 kcal/mol. The authors stated that a very much smaller barrier would be expected with an extended basis set and inclusion of electron correlation. The assistance of the residue Thr-199 in which the oxygen of the methyl ether forms a hydrogen bond with the hydrogen of the hydroxy group bonded to Zn (see Figure 2) can be expected to lead to additional lowering of the barrier. Liang and Lipscomb[28] also examined how the barrier changes with the nature of the ligands used on Zn. From the data given (Table 1),[28] one can deduce that the barrier increases as the deprotonation energy (see reaction 13a) increases. Since the deprotonation energy with imidazole rather than ammonia as ligands is predicted to be higher,[7] a higher barrier for the proton transfer to the first water molecule would be expected with the imidazole ligands.

Even though the height of the proton transfer barrier is not well known, the considerations given above indicate that the barrier is sufficiently high so that selection of ligands that are more strongly electron donating than the three imidazoles, by the substitution of, say, one glutamate for one imidazole, will make the barrier too high and slow the proton-transfer rate below the rate required for the catalysis. Therefore, one can assume that the choice of the three imidazole ligands is close to the choice that keeps the proton-transfer reaction sufficiently fast *and* provides the maximum assistance for the carboxylation reaction 14 whose rate is expected to increase as electron donation from the ligands increases; see Figures 3 and 4.

Some of the intermediates of reaction 14 proposed in recent theoretical modeling studies by Merz et al.[4e,f] are shown in Figure 5. The calculated energies of the intermediates included modeling of the effect of the environment (PM3/MM). The CO_2 molecule, which has entered via the hydrophobic pocket formed by the three Val residues, is attacked by the nucleophilic hydroxy ligand. This leads through the transition-state structure **A** to the intermediate **B**. Intramolecular proton transfer in the HCO_3^- from one C−O oxygen to the other (structures **B**−**D**, Figure 5) is mediated by the participation of residues Glu-106, Thr-199, and Thr-200. In the final step, structures **D** and **E**, the bicarbonate anion HCO_3^- is displaced by a water molecule. The authors made calculations only up to structure **D**, Figure 5. The activation energies required up to that point were relatively small, less than 10 kcal/mol.[4e,f] They assumed that the displacement of HCO_3^- by H_2O, step D to E, occurs readily.

The relatively low activation energy for the preceding steps A to D are approximately in line with a fast turnover rate observed experimentally for CAII. However, it is unfortunate that data for the final steps D to E are not available. This step probably is very difficult to model with a proper inclusion of the effects of the aqueous environment. If one accepts the assumption of the authors[4a,f] that the final step, D to E, occurs readily, and combines the low activation energy obtained[4f] with the low reaction free energy $\Delta G^\circ_{14d} = -5$ kcal/mol (present work), one comes to the conclusion that the proposed mechanism is energetically possible and that the three imidazole ligands are just right to provide the low free energy required for the fast catalytic process.

(29) Partial retention of differential overlap (PRDDO), which uses an osthogonalized basis set, is a close approximation to SFC MO calculations at the minimum basis set level.[4e]

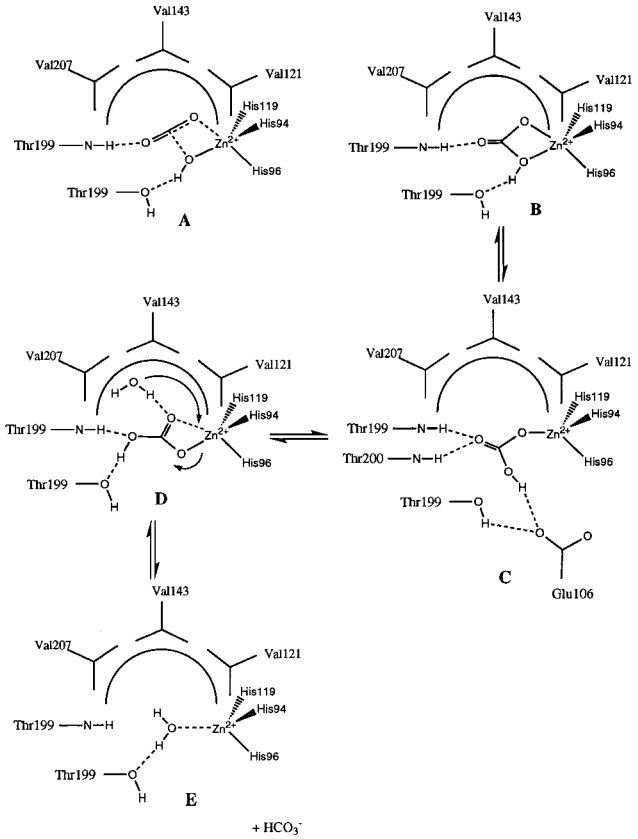


Figure 5. Intermediates for the formation of the bicarbonate anion HCO_3^- proposed by Merz and co-workers.[4e,f] The CO_2 molecule, which enters the reaction center via the hydrophobic pocket formed by the three Val residues, is attacked by the nucleophilic hydroxy ligand. This leads via transition state **A** to intermediate **B**. Intramolecular proton transfer in the HCO_3^-, structures **A**, **C**, and **D**, is mediated by residues Glu-106, Thr-199, and Th-200. In the final step, structures **D** and **E**, the bicarbonate anion is displaced by a water molecule. (Reproduction of Figure 4 from Merz, K. M., Jr.; Banei, L. *J. Am. Chem. Soc.* **1997**, *119*, 863).

The value obtained for ΔG°_{14d} was based on HCO_3^- being aqueous. A considerably higher value would be expected for HCO_3^- located near the reaction site, because of the less favorable ion solvation in the partially hydrophobic environment of the enzyme. The transport of the HCO_3^- from structure **D** to the aqueous medium could be possible but should require the participation of suitable residues.

Conclusions

(a) Ligands such as histidine (imidazole), which bond very strongly to the metal ion, provide the needed stability for the metal ion in the enzyme, thus preventing the ion from escaping into the strongly solvating aqueous environment outside the enzyme. The present results, involving theoretical calculations of the binding energies of the directly bonded ligands in the $Zn(Im)_3(H_2O)^{2+}$ complex and a semiquantitative inclusion of the stabilization due to the protein and aqueous environment, provide a semiquantitative breakdown of the total stabilization of Zn^{2+} in the enzyme, which shows that by far the most important stabilization is due to the three imidazole ligands.

(b) Due to the very substantial stabilization of the ion by the polarizability and electron transfer from the three strongly bonding ligands, the bonding to the fourth ligand, H_2O (or OH^-), is much weakened. Such a weakening is required for the catalytic process. The generally accepted mechanism for CAII

involves a deprotonation reaction 13 and an HCO_3^--forming reaction 14. The two reactions have conflicting requirements concerning the nature of the direct-bonding ligands. Reaction 14 is facilitated by strongly charge-donating ligands while reaction 13 is slowed by such ligands. To satisfy these conflicting requirements, the three primary ligands must be so chosen to lead to free energy changes ΔG°_{13} and ΔG°_{14}, which are both close to zero. The three imidazole ligands fulfill this requirement as shown in Figure 4. Thus, the "choice" of the three histidine ligands in CAII appears to be the optimal choice—the choice that is just right.

(c) The semiquantitative method for the inclusion of the stabilization due to the enzyme and aqueous environment, as evaluated in the present work, greatly increases the uncertainty in the values of reactions 13d and 14d, which "model" the reactions in solution. Much better modeling of the solvation effects are desirable. The authors hope that the present results may stimulate such better modeling by research groups which specialize in such work.

Appendix I

(A) Some Details Concerning Values Used in Table 2. (1) Stabilization of Zn^{2+} by Three Imidazole Ligands. This is a theoretical result; see Table 1.

(2) Stabilization of $Zn(Im)_3^{2+}$ by H_2O. Theoretical and experimental results are available; see Table 3A.

(3) Strong Hydrogen Bonds Induced by Zn^{2+} Charge. Three strong hydrogen bonds are formed between hydrogen of the three histidines and basic residues that are nearby. The histidine hydrogens are made more protic by the Zn^{2+} due to charge transfer from imidazole to Zn^{2+}. Alvarez-Santos et al.[30] have made calculations with the semiempirical AM1 for the hydrogen-bonding interactions of His-94 to Glu-92, His-96 to Asn-244, the H_2O ligand to Thr-199, and His-119 to Glu-117, using the model $Zn(Im)_3OH_2^{2+}$. The residues other than His were modeled by using the corresponding side-chain groups attached to a methyl group. For the first three interactions, which involve uncharged groups, they obtained 27.5 kcal/mol total binding energy, i.e., roughly 9 kcal/mol per hydrogen bond. For the interaction of the charged residue Glu-117 with His-119, a very strong bond of 118 kcal/mol was predicted[30] and can be expected due to the strong Coulombic interaction with the anion. However, the value to be used in the present estimate should be much smaller. The deprotonated Glu-117 before its interaction with His-119 was probably stabilized by interaction with water molecules and neighboring groups at the reactive center of the enzyme. This stabilization can be roughly modeled by replacing Glu with the acetate anion $CH_3CO_2^-$. The interactive free energy of the acetate anion with liquid water is $\Delta G^\circ_h(HCO_3^-) \approx -90$ kcal/mol.[31] If we assume the same stabilization being present in the enzyme plus aqueous environment, the stabilization by the gas-phase hydrogen bond formation between His-119 and Glu-117 of 118 kcal/mol will be reduced to 28 kcal/mol. Therefore, we assume a total free energy value of -55 kcal/mol.

(4) Solvation by Protein. To evaluate the solvation we use the Born equation:[32,33]

$$-\Delta G^\circ_{sol} = \frac{N_A Z^2 e^2}{8\pi\epsilon_o R}\left(1 - \frac{1}{D}\right)$$

where, e is the charge of the electron, Z is the charge of the ion ($Z = 2$ in the present case), N_A is Avogadro's number, ϵ_o is the permittivity of vacuum, R is the radius of the ion, and D is the dielectric constant of the solvating medium. Substituting the numerical values of the constants one obtains

$$-\Delta G^\circ_{sol} = \frac{165.2Z^2}{R}\left(1 - \frac{1}{D}\right) \quad \text{(kcal/mol)}$$

where the numerical constant gives ΔG°_{soli}in kilocalories per mole, when R is substituted in angstroms.

The radius R of $Zn(Im)_3^{2+}$ was based on the distance between zinc and the outer imidazole hydrogen(s) of 5 Å, available from the ab initio calculations. A distance of 2 Å was added to obtain $R = 7$ Å for the "cavity" of the protein around $Zn(Im)_3^{2+}$.

A dielectric constant of $D = 2$ was used for the protein.[34] The result obtained with these parameters and the Born equation was $\Delta G^\circ_{sol} = -47.2$ kcal/mol. To correct for the fact that the protein extends only to a radius of $R \approx 25$ Å (Protein Data Base, Swiss Institute of Bioinformatics),[17] we evaluated a solvation energy with $R = 25$ Å and $D = 2$, ΔG°_{sol}(protein) $= -13.2$ kcal/mol, and subtracted it, obtaining ΔG°_{sol}(protein) $= -47.2 + 13.2 = -34$ kcal/mol.

(5) Solvation by the Aqueous Medium. Using a radius for the protein (CAII), $R = 25$ Å and D = 78 for water, one obtains with the Born equation, ΔG°_{sol}(aqua) $= -26$ kcal/mol.

(6) Solvation by the Conical Cleft. CAII was approximated as a sphere with Zn^{2+} at the center. The angle of the cleft was approximated by using the X-ray crystallographic structure of the enzyme from the Expasy database Structure 1CA2 of the Swiss 3D collection.[17] Using the standard atom parameters of the space-filling model, the enzyme was lined up so that a view down one side of the cleft left only a small part of the Zn ion unobstructed. The structure was then rotated along the x-axes until the other side of the cleft was reached. The angle of rotation was ~60°. The angle of rotation along the y-axes across the cleft was also ~60°. Using the value of 60° for the angle and the standard surface area equation, the surface area of the cleft was determined to be $0.268\pi a^2$ where the radius of the sphere is a. Since the total surface of the sphere is $4\pi a^2$, the surface of the cleft was found to be equal to a fraction of 0.0676 of the total area of the sphere.

Using the Born equation for water $D = 78$ and a radius $R =$ 5A assumed to be the bottom of the cleft, a total hydration energy of -93 kcal/mol is obtained. Taking a fraction of 0.067 of this total value, one obtains -9 kcal/mol, which approximates the solvation due to the cleft.

(B) Estimates of Errors for Energies Given in Table 2. (1) Binding Energy of $(Im)_3Zn^{2+}$. An estimated error of 20 kcal/mol is proposed. See section IIb for a discussion.

(2) Monohydration free energy of $Zn(Im)_3^{2+}$ was confirmed by experiment within 2 kcal/mol; see Table 2A in section III. We have taken the average of the two values. Assumed uncertainty: ± 1 kcal/mol.

(3) Strong hydrogen bonds between imidazoles and nearby residues were induced by charge on Zn^{2+}. The calculated values were obtained[30] with the semiempirical AM1 method. Further-

(30) Santos, S. A.; Gonzales-Lafont, A.; Lluch, J. M. *Can. J. Chem.* **1998**, *76*, 1027.

(31) The free energy of hydration of the acetate anion ΔG°_h ($CH_3CO_2^-$) is not available;[19] however, $\Delta H^\circ_h = -101.1$ kcal/mol and $S^\circ_{aq} = 26$ cal/(deg·mol) are available.[19] S°_g can be approximated from S°_g (CH_3NO_2) = 65 cal/(deg·mol).[27] This value leads to $\Delta S^\circ_h = -39$ cal/(deg·mol) and $\Delta G^\circ_h = -90$ kcal/mol.

(32) Conway, B. E. *Ionic Hydration in Chemistry and Biophysics*; Elsevier: New York, 1981; p 313.

(33) King, G.; Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1991**, *95*, 4366.

(34) For a justification of the use of the dielectric constant value $D = 2$, see Discussion of Errors at the end of Appendix I.

more, the result[30] for Glu-117 had to be modified. The error is very difficult to estimate. We arbitrarily assume ±15 kcal/mol.

(4) Solvation by Protein. The Born equation,[32,33] while not reliable when used for small ions, is expected to provide reliable results for large ions,[32] such as the $R = 7$ Å used in the present case. The dielectric constant D used for proteins, $D = 2$, can be questioned. Calculated values[35] between 2 and 4 have been reported in the absence of solvent. Even much higher values such as $D = 10$ have been reported.[33] However the high values[33] were chosen in order to include the stabilizing effect of specific polar residues located very near to the ion. In the present case, we account for such stabilization separately; see part 3 in this appendix.

A change of dielectric constant from $D = 2$ to $D = 3$ leads to an increase of the calculated energy from ΔG°_{sol}(protein) = −34 ($D = 2$) to −45 kcal/mol ($D = 3$). On this basis, we estimate an error of 10 kcal/mol.

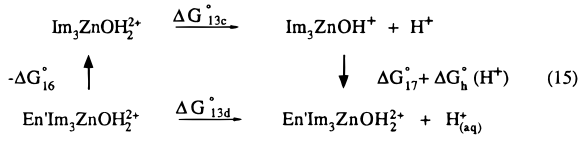
(5) Solvation by the Aqueous Medium. Because the radius of the protein is big, the Born equation should provide a good account of the solvation. Assuming an uncertainty of the protein radius by 3 Å, the estimated error with the Born equation is 2 kcal/mol.

(6) Solvation of Conical Cleft. The procedure used for the evaluation of the solvation energy is not strictly valid. However, because this energy is small, we arbitrarily assume an error of 4 kcal/mol.

Total estimated error obtained by the chain rule is ±27 kcal/mol.

Appendix II. Thermochemical Equations Used To Obtain the Free Energies for Reactions 13d and 14d

The solution-phase free energy ΔG°_{13d} was obtained via the thermodynamic cycle



(15)

which leads to

$$\Delta G^\circ_{13d} = \Delta G^\circ_{13c} - \Delta G^\circ_{16} + \Delta G^\circ_{17} + \Delta G^\circ_h(H^+)$$

ΔG°_{13c} = 174.3 kcal/mol, corresponding to the energy of the gas-phase reaction was obtained in section IV; see Table 3C. The hydration energy of the hydrogen atom is available in the literature:

$$\Delta G^\circ_h(H^+) = -253 \text{ kcal/mol}^{19}$$

The solvation interactions of $Im_3ZnOH_2^{2+}$ and Im_3ZnOH^+ with En′ and the aqueous environment,

$$Im_3ZnOH_2^{2+} \xrightarrow[H_2O(aq)]{En'} En'Im_3ZnOH_2^{2+}(aq) \qquad \Delta G^\circ_{16} = -124 \text{ kcal/mol} \quad (16)$$

$$Im_3ZnOH^{+} \xrightarrow[H_2O(aq)]{En'} En'Im_3ZnOH^{+}(aq) \qquad \Delta G^\circ_{17} = \Delta G^\circ_{16}/4 = -31 \text{ kcal/mol} \quad (17)$$

and their difference,

$$\Delta G^\circ_{18} = \Delta G^\circ_{16} - \Delta G^\circ_{17} = -93 \pm 14 \text{ kcal/mol} \quad (18)$$

were estimated as follows.

Because only the difference, $\Delta G^\circ_{17} - \Delta G^\circ_{16}$, enters the equation for ΔG°_{13d}, only the effects of the changes of charge from doubly charged to singly charged ion were taken into account. The interactions with the environment for the doubly charged species $En'Im_3Zn(OH_2)^{2+}$ were evaluated in Appendix IA; see Table 2. From the terms listed in Table 2, we take the charge-dependent terms: (3) strong hydrogen bonds induced by Zn^{2+}, $\Delta G^\circ_3 = -55$ kcal/mol; (4) solvation by protein $\Delta G^\circ_4 = -34$ kcal/mol; (5) solvation by aqueous environment, $\Delta G^\circ_5 = -26$ kcal/mol; and (6) solvation by conical cleft, $\Delta G^\circ_6 = -9$ kcal/mol. These add up to a total of $\Delta G^\circ_{16} = -124$ kcal/mol. The value for ΔG°_{17} was obtained from the following considerations. The free energy changes for (4), (5), and (6) for the doubly charged species were evaluated with the Born equation; see Appendix IA. In this equation, the dependence on the charge is given by Z^2. For the doubly charged ion $En'Im_3Zn(OH_2)^{2+}$, this leads to a factor of 4, while for the singly charged ion $En'ImZnOH^+$, the factor is 1. The dependence of (3), i.e., the strong hydrogen bonds of the imidazole ligands, on the charge Z is not exactly known. We assume that this term (ΔG°_3) also depends on Z^2. With this assumption, the estimate of ΔG°_{17} is one-quarter of ΔG°_{16}; see eq 17.

The estimated error of 14 kcal/mol is based on the error estimates (Appendix IA) for the ΔG°_3 to ΔG°_6 terms, which lead to the present ΔG°_{16}.

With these available data, ΔG°_{13d} can now be evaluated:

$$\Delta G^\circ_{13d} = \Delta G^\circ_{13c} + \Delta G^\circ_h(H^+) - \Delta G^\circ_{18} = +14 \text{ kcal/mol}$$

ΔG°_{14d} was obtained from an analgous thermodynamic cycle which leads to $\Delta G^\circ_{14d} = \Delta G^\circ_{14c} + \Delta G_h(HCO_3^-) + \Delta G^\circ_{18} - \Delta G^\circ_h(H_2O) - \Delta G^\circ_h(CO_2) = -5$ kcal/mol.

The numerical result is obtained with the available free energies of hydration of the reactants: $\Delta G^\circ_h(HCO_3^-) \approx -83$ kcal/mol,[26,27] $\Delta G^\circ_h(H_2O) = -2$ kcal/mol,[27] $\Delta G^\circ_h(CO_2) = 2$ kcal/mol.[27]

JA992406L

(35) (a) Gilson, M. K.; Rashin, A.; Fine, R.; Honig, B. *J. Mol. Biol.* **1985**, *183*, 503. (b) Gilson, M. K.; Honig, B. *Biopolymers* **1986**, *25*, 2097. (c) Pethig, R. *Dielectric and Electronic Properties of Biological Materials*; John Wiley and Sons: New York, 1979. (d) Krishtalik, L. I. *J. Theor. Biol.* **1989**, *139*, 143.