although other information, such as the internuclear separation and absolute sign of J, can be used to limit the possible values for ΔJ.

In light of this work and the results available in the literature, from both experimental and theoretical investigations, it is clear that there are significant anisotropies in the indirect spin–spin coupling involving heavier elements of the periodic table. Such observations conclusively show that these couplings are not determined solely by contact terms and that interpretations based upon this assumption need to be reevaluated. It may be revealed, after further effort, that the trends in J couplings previously proposed from solution studies are indeed primarily due to variations in the contact terms. However, at present, there is no evidence to support this, due in part to the paucity of experimental data for ΔJ. We believe that the examples presented here provide unequivocal evidence for the magnitude of ΔJ in these compounds. It is hoped that investigations such as that performed here will continue and increase our knowledge of the origins and factors determining the nature of the indirect spin–spin coupling between nuclear spins.

Acknowledgment. We thank the Natural Sciences and Engineering Research Council (NSERC) of Canada for financial support in the form of operating and equipment grants (R.E.W.) and scholarship assistance (W.P.P.). The NSERC also supports the Atlantic Region Magnetic Resonance Centre, where all spectra were obtained. We also appreciate the assistance of Dr. A. Berger and Dr. A. Kentgens in providing software to aid in the fitting of slow-spinning MAS spectra and Prof. G. Penner for providing some of the preliminary experimental spectra.

Force Field Design for Metalloproteins

Stephen C. Hoops,[†] Kenneth W. Anderson,[‡] and Kenneth M. Merz, Jr.*

Contribution from the Department of Chemistry and Molecular and Cell Biology, The Pennsylvania State University, University Park, Pennsylvania 16802. *Received April 1, 1991*

Abstract: Herein we describe a general approach to systematic derivation of force field parameters for metalloproteins. In particular, we have extended the AMBER force field to model the zinc ion in human carbonic anhydrase II (HCAII) in both the high (zinc–hydroxide) and low (zinc–water) pH forms. Key elements of this approach are the systematic evaluation of molecular orbital methods for modeling the metallic center (zinc in HCAII) with its ligands in the active site and the force field representation of electrostatic interactions with point charges fitted to the electrostatic potential. The resulting greater accuracy in modeling electrostatic interactions for the active site embedded in the protein environment increases the capacity for studying the static and dynamic structure of human carbonic anhydrase II. This approach should be transferable to computational studies of other metalloproteins at fixed coordination numbers.

Introduction

Several force fields have been developed in order to conduct computer simulations of the structure and function of proteins.[1-6] In general these force fields have undergone continuous improvement, and they are increasingly useful in simulating the three-dimensional structures and dynamic properties of proteins. Herein we describe the application of a systematic method for accurately incorporating a metal ion and its ligands into a classical force field for subsequent modeling of the static and dynamic structure of a metalloprotein.

The present study utilizes and builds upon the AMBER force field.[1] In this force field the total energy of the system is modeled by the potential function.

$$E_{\text{Total}} = \sum_{\text{bonds}} \frac{K_r}{2}(r - r_{\text{eq}})^2 + \sum_{\text{angles}} \frac{K_\theta}{2}(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \sum_{n} \frac{V_n}{2}[1 + \cos(n\theta - \gamma)] + \sum_{i<j} \epsilon_{ij}\left[\left(\frac{R_{ij}^*}{R_{ij}}\right)^{12} - \left(\frac{R_{ij}^*}{R_{ij}}\right)^{6}\right] + \frac{1}{\text{VDW}_{\text{scale}}} \sum_{\substack{i<j \\ 1\text{-}4 \text{ terms}}} \epsilon_{ij}\left[\left(\frac{R_{ij}^*}{R_{ij}}\right)^{12} - \left(\frac{R_{ij}^*}{R_{ij}}\right)^{6}\right] + \sum_{\text{H-bonds}} \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}}\right] + \sum_{i<j} \frac{q_i q_j}{\epsilon R_{ij}} + \frac{1}{EE_{\text{scale}}} \sum_{\substack{i<j \\ 1,4 \text{ terms}}} \frac{q_i q_j}{\epsilon R_{ij}}$$

The first three terms present energy contributions for the internal motions, bond stretching, angle bending, and dihedral torsions, present in the molecule. The fourth and fifth terms (Lennard-Jones terms) represent the nonbonded interactions between atoms separated by distance R, and the sixth term represents hydrogen-bonding interactions. The last two terms shown represent the contributions to potential energy due to electrostatic interactions between point charges q_i and q_j separated by distances R_{ij} in a medium of dielectric constant ϵ. The assignment of point charges can be a critical problem in modeling proteins. In metalloproteins the presence of the metallic ion produces particularly strong electrostatic interactions which can be a major influence in determining the three-dimensional structure, modes of binding between small molecule and ion substrates, and mechanisms for catalysis.[7] The accurate modeling of these last two electrostatic terms for metalloenzymes will be an important focus of this article.

Force field calculations can be coupled with other computational techniques, such as molecular dynamics (MD),[8,9] quantum me-

[†] Present address: The Pennsylvania State University, New Kensington Campus, New Kensington, PA 15068.

[‡] Participant in the National Science Foundation Research Experience for Undergraduates Program.

(1) (a) Weiner, S. J.; Kollman, P. A. *J. Comput. Chem.* **1986**, *6*, 718. (b) Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1986**, *7*, 230.

(2) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J. *J. Comput. Chem.* **1983**, *4*, 187.

(3) Hermans, J.; Berendson, J. C.; van Gunsteren, W. F.; Potsma, J. P. M. *Biopolymers* **1984**, *23*, 1513.

(4) Momany, F.; McGuire, R.; Burgess, A.; Scheraga, J. *J. Phys. Chem.* **1975**, *79*, 2361.

(5) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(6) Merz, K. M., Jr.; Kollman, P. Unpublished results.

(7) See for example: Burley, S. K.; Petsko, G. A. *Adv. Protein Chem.* **1988**, *39*, 125. Harvey, S. C. *Proteins: Struct. Func. Genet.* **1989**, *5*, 78. Davis, M.; McCammon, J. A. *Chem. Rev.* **1990**, *90*, 509 and references cited therein.

(8) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Advances in Chemical Physics. Volume LXXI: Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; John Wiley and Sons: New York, 1988.

(9) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: New York, 1987.

0002-7863/91/1513-8262$02.50/0

chanical/molecular mechanical coupling,[10-12] and free energy perturbation (FEP),[13-16] to expand the scope of modeling applications. Among force fields the most useful provide the capacity to perform molecular dynamics (MD) simulations.[1-6] Applied alone energy minimization algorithms are subject to the general problem of yielding structures at local minima. The addition of MD techniques allows one to sample more phase space and makes it less likely to become trapped in local minima.[8,9] FEP techniques provide the capacity to analyze dynamic processes for selected substructures. Thus, one of our goals is to develop a force field model of the metal ion coordination that will allow us to routinely use MD and FEP techniques.

The molecular modeling of the structure and dynamics for a metalloprotein is a more complex problem than modeling a protein with no incorporation of metal ions. Force fields have been successfully utilized, simulating structures and dynamic processes for hemoglobin and myoglobin.[17-19] Among the metalloproteins these iron systems can be regarded as most amenable to molecular modeling. The polar iron atom tends to be coordinatively saturated and embedded in a heme ring system, and as a result, there are weak electrostatic interactions between the active site and the protein.

The modeling of zinc metalloproteins, such as human carbonic anhydrase (HCA),[20-27] carboxypeptidase (CPA),[28] liver alcohol dehydrogenase (LADH),[29] and thermolysin,[30] has developed less rapidly and proved to be a more difficult and challenging problem. This difficulty is almost certainly due to modeling interactions related to stronger electrostatic interactions in a more dynamic environment.[7] Generally the zinc at an active site is more polar due to bonding to oxygen, nitrogen, and sulfur atoms. It is part of a dynamic system because it tends to have four separate ligands without chelation and may vary its coordination number. It is often located near polar side chains and exposed to highly polar water molecules. With such a complex system it is wise to note that the purpose for developing a force field is to simulate complex quantum mechanical structures with classical mechanical equations. Thus, the history of modeling all chemical systems and, particularly, zinc proteins reflects attempts to simplify structures and simulate the most essential interactions with computational efficiency.[20,21,24-30] In the earliest studies force field computations neglected electrostatic interactions between the metal ion and its environment[20] or omitted the metal ion and simulated the apoprotein.[28] The simplest electrostatic models assumed a formal 2+ charge on the zinc ion in the active site[27,30] and either produced unsatisfactory results or required constraints in order to avoid six- or five-coordination in either energy minimization[23,30] or MD simulations following energy minimization.[23,30]

Most recently, a potential function for modeling metal ligand bonding has been reported.[22] This function is based on experimental data from crystal structures for small molecules and is parametrized for predicting coordination number changes and charge transfer at Co(II) and Zn(II) centers. With inclusion in the AMBER force field it has been successfully applied to small-molecule crystal structures. It has been used to predict structures for Zn(II)- and Co(II)-substituted HCAI.[22] This force field and the one described here have been designed for different purposes and, thus, reflect different conceptual approaches and different methods. In our method the purpose is to improve the parametrization process for modeling the metal at fixed coordination or limited change in coordination with the result of improving both energy minimization and molecular dynamics calculations. Our bonded approach necessarily introduces an inherent limitation in that the coordination at the metal ion can only change in a limited manner during the course of a minimization or simulation. However, in our opinion there is such a dearth of experimental information correlating both *energy* changes and *structural* changes involved in altering the coordination around a zinc ion that this is necessary. The alternative approach of trying to dynamically model changes in coordination around a metal ion in the absence of energetic information is more problematic in our estimation.

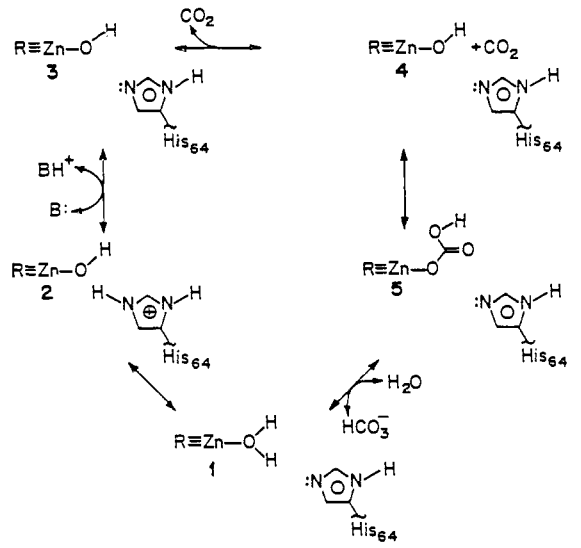
For this study we adopted the assumptions that a useful force field would simulate explicit bonding between zinc and selected ligands and model all electrostatic interactions between the active site and the protein as accurately as possible. We will call this a "bonded" approach because certain bonds are identified as conserved for the chemical system to be modeled. The energy changes related to structural changes for conserved zinc to ligand bonds are modeled by the bond, angle, and dihedral terms in AMBER. Possible additional bonds can be modeled by the electrostatics term in AMBER, producing a model that is dynamic in a limited sense in that it simulates processes with no changes or increases in coordination number. It is anticipated that the major advantage will be an improved representation of electrostatics. We have decided to simplify this bonded approach and not model charge transfer interactions between metal and ligands until it has not been demonstrated that this is necessary for improving accuracy.[22]

The adoption of these assumptions is largely based on a study of sulfonamide inhibitors of carbonic anhydrase (HCAII) in which these assumptions were made.[23] Energy minimization in the AMBER force field produced tetrahedral geometry at zinc and, most importantly, this geometry was maintained during MD simulation.[23] In unreported results we find that using an alternative "nonbonded" or more "dynamic" approach may produce correct structures in energy minimizations, but these structures are not maintained during MD simulations. A problem associated with using the bonded approach arises from treating the metal–ligand bonding as harmonic, which may lead to symmetric bond distributions around the metal center (i.e. all metal–ligand bonds have the same bond length). However, we find that with a reasonable choice of zinc–ligand bond parameters asymmetric bonding patterns can be reproduced (vide post). With these considerations we have opted to use the "bonded" approach.

The approach we have adopted requires that we obtain bond, angle, and torsion parameters for the zinc ion and its coordination sphere, and we have done this by using, where possible, experimental data or ab initio data.[31] We have determined the nonbonded parameters (i.e. Lennard-Jones and electrostatic terms) using molecular orbital techniques. The charges for the atoms are obtained via fitting the molecular electrostatic potential.[32,33]

(10) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *6*, 718.
(11) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
(12) Warshel, A.; Karplus, M. *J. Am. Chem. Soc.* **1972**, *94*, 5612. Field, M. J.; Bash, P. A.; Karplus, M. *J. Comp. Chem.* **1990**, *11*, 700.
(13) van Gunsteren, W. F. *Protein Engineering* **1988**, *2*, 5.
(14) Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1.
(15) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184.
(16) Kollman, P. A.; Merz, K. M., Jr. *Acc. Chem. Res.* **1985**, *18*, 105.
(17) Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 7690.
(18) Case, D. A.; McCammon, J. A. *Proc. N.Y. Acad. Sci.* **1986**, *482*, 222.
(19) Case, D. A.; Karplus, M. *J. Mol. Biol.* **1979**, *132*, 343.
(20) Vedani, A.; Dobler, M.; Dunitz, J. D. *J. Comput. Chem.* **1986**, *7*, 701–710. Vedani, A.; Dunitz, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7653.
(21) Vedani, A.; Huhta, D. W.; Jacober, S. P. *J. Am. Chem. Soc.* **1989**, *111*, 4075–4081.
(22) Vedani, A.; Huhta, D. W. *J. Am. Chem. Soc.* **1990**, *112*, 4759–4767.
(23) Merz, K. M., Jr.; Murcko, M. A.; Kollman, P. A. *J. Am. Chem. Soc.* Accepted for publication.
(24) Merz, K. M., Jr. *J. Mol. Biol.* **1990**, *214*, 799.
(25) Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 4060.
(26) Merz, K. M., Jr. *J. Am. Chem. Soc.* Accepted for publication.
(27) Liang, J.-Y.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. (U.S.A.)* **1990**, *87*, 3675.
(28) Makinen, M. W.; Troyer, J. M.; van der Werff, H.; Berendsen, H. J. C.; van Gunsteren, W. F. *J. Mol. Biol.* **1989**, *207*, 201.
(29) de Kok, P. M. T.; Beijer, N. A.; Buck, H. M.; Sluyterman, L. A. A. E.; Meijer, E. M. *Recl. Trav. Chim. Pays-Bas* **1988**, *107*, 355.
(30) Merz, K. M., Jr.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 5649.
(31) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; MacMillan: New York, 1982. Hehre, W. J.; Radom, L.; v. R. Schleyer; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley and Sons: New York, 1986.
(32) Cox, S. R.; Williams, D. E. *J. Comp. Chem.* **1981**, *2*, 304. D. E. Williams, J. M. Yan *Adv. Atomic Mol. Phys.* **1988**, *23*, 87. Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129. Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894. Woods, R. J.; Khalil, M.; Pell, W.; Moffat, S. H.; Smith, V. H., Jr. *J. Comp. Chem.* **1990**, *11*, 297.

Scheme I

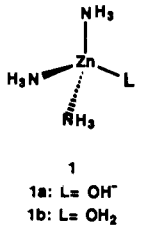


This approach has been shown to accurately predict experimentally observed electrostatic properties of molecules as well as electrostatic interactions between molecules.[32] Because the systematic approach described here has this firm theoretical basis, it should be applicable to other metalloproteins.

We have selected the metalloprotein human carbonic anhydrase II (HCAII) as the test metalloprotein system for our general approach. Human carbonic anhydrase II (carbonate dehydrase, EC 4.2.1.1) is an enzymatic protein and one of seven isozymes with designations HCAI–HCAVII. It is predominantly found in red blood cells and catalyzes the interconversion of carbon dioxide and the bicarbonate ion.

$$CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+$$

This protein with 260 amino acid residues and a molecular mass of 29 300 Da is a metalloenzyme with a single zinc atom in the active site for catalysis. The structure of the high pH (hydroxide) form has been determined by X-ray crystallography[34] and refined at 2.0-Å resolution.[35] The active site possesses several unique features and lies in a conical cavity which is unusually wide at the entrance (ca. 15 Å) and deep (ca. 15 Å). This cavity is divided into a hydrophobic region and a hydrophilic region with a large number of water molecules interconnecting the active site and external solvent. The zinc is located at the bottom of the cavity and has a nearly tetrahedral geometry with coordination to three histidine nitrogens (His-94, -96, and -119) and hydroxide oxygen. It is deduced that the oxygen is part of a hydroxide group because the pH of the crystallization medium was 8.5[34] and the pK_a of zinc-bound water has been estimated as 7.0.[36] Recent unpublished

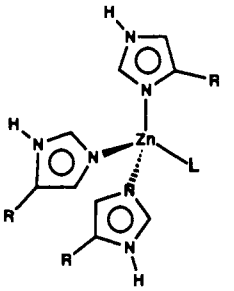


1

1a: L= OH^-

1b: L= OH_2

Figure 1. Models of active site in HCAII used in the MO calculations.



2

2a: L= OH^-; R= $CH_2CH(COCH_3)(NHCOCH_3)$

2b: L=OH^-; R= H

2c: L= OH_2 ; R= $CH_2CH(COCH_3)(NHCOCH_3)$

2d: L=OH_2 ; R= H

Figure 2. Small and large models of active site in HCAII used in the MO calculations.

X-ray crystallographic results for the low pH 5.7 (water) form of HCAII show the zinc has a coordination number of four.[37]

The catalytic mechanism of human carbonic anhydrase (HCA) has been studied in detail but has not been completely elucidated.[36] The catalysis is dependent on a group whose pK_a is around 7.[36] After much debate it was decided that a zinc-bound water best satisfied this criteria, which led to the formulation of the zinc-hydroxide mechanism (see Scheme I).[36] A considerable body of evidence indicates the step converting **1** into **2** is rate limiting proton transfer, which is kinetically separated from the sequence of steps converting **2** into **1** via **3**, **4**, and **5**. The separation is necessary to explain the ping-pong kinetics observed.[36] Although experiments have not completely elucidated the detailed structural changes in the mechanism for catalysis, there is considerable evidence that certain residues are probably catalytically important. These include His-64, Glu-106, Thr-199, and water molecules near the active site.[35,36] Thr-199 is positioned with Thr-200 on the opposite side of the cavity from the zinc ion (see e.g. Figure 4a).[35] These threonine residues, His-64 (located at the entrance of the cavity) and Glu-106 combine with other polar residues to constitute the hydrophilic half of the cavity. Thr-199 is a key residue which is centered between this hydrophilic half and the more hydrophobic half. It is locked in position as part of an important hydrogen-bonded subsystem. The hydrogen on oxygen bonded to zinc is the donor bond for hydrogen bonding to Oγ1 of Thr-199 and the hydrogen bound to Oγ1 at Thr-199 is a donor bond for hydrogen bonding to Oϵ1 of Glu-106.[24,35] The proximity of Thr-199 to zinc and the rigidity of this hydrogen-bonding subsystem are considered to be crucial for inhibitor binding,[35] and it may be a crucial structural feature for CO_2 binding and catalysis.[24,25] In general small molecules or ions with donor bonds for hydrogen bonding can hydrogen bond at the acceptor oxygen at Thr-199, but other species, e.g. SCN^-, must bind at zinc from the side in the hydrophobic region.[35] In has been suggested that the zinc hydroxide–Thr-199–Glu-106 subsystem may be a crucial structural feature which opens a hydrophobic pocket for CO_2 binding and orients the CO_2 molecule for nucleophilic addition.[24] Bordering on this hydrophilic region is a hydrogen-bonding network of eight water molecules which extends toward bulk water.[35] During catalysis some of these waters may participate in proton relays

(33) Besler, B. H.; Merz, K. M., Jrr.; Kollman, P. *A. J. Comp. Chem.* **1990**, *11*, 431.

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

(35) 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.

(36) 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. *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.

(37) Christianson, D. W. Personal communication.

Table I. Calculated Electrostatic Potential Charges for **1a** and **1b** by Using Molecular Orbital Methods

atom in ligand[a]	ligand[b] at zinc	MNDO	ab initio 6-31G*	ab initio STO-3G*	ab initio MINI-4	ab initio MIDI-4
O	H_2O	−0.7176	−1.0863	−0.5779	−0.9765	−0.9956
O	HO	−0.8570	−1.2250	−0.7070	−1.2182	−1.3792
Zn	H_2O	1.1133	1.5536	0.6767	1.6417	1.2118
Zn	HO	0.8288	1.3750	0.4572	1.7183	1.3233
N	H_2O	−0.7567	−1.2179	−0.7182	−0.8977	−0.7996
N	H_2O	−0.7450	−1.1831	−0.6991	−0.8777	−0.7760
N	H_2O	−0.7446	−1.1822	−0.7405	−0.9766	−0.8587
N	HO	−0.5677	−1.0828	−0.6448	−0.9469	−0.8142
N	HO	−0.5494	−1.0686	−0.6368	−0.9337	−0.8000
N	HO	−0.4267	−0.8842	−0.5345	−0.9362	−0.8711
H (O)	H_2O	0.4440	0.5682	0.4278	0.5038	0.5475
H (O)	H_2O	0.4439	0.5680	0.4299	0.5088	0.5513
H (O)	HO	0.3348	0.4532	0.3106	0.2751	0.5130
H (N)	HO	0.2590	0.4030	0.3156	0.3451	0.3379
H (N)	HO	0.2586	0.3929	0.3092	0.3228	0.3118
H (N)	HO	0.2606	0.4019	0.3185	0.3510	0.3464
H (N)	HO	0.2582	0.3997	0.3184	0.3477	0.3423
H (N)	HO	0.2523	0.3862	0.3056	0.3169	0.3059
H (N)	HO	0.2548	0.3989	0.3134	0.3440	0.3361
H (N)	HO	0.2394	0.3635	0.2980	0.3497	0.3622
H (N)	HO	0.2128	0.3191	0.2766	0.3134	0.3216
H (N)	HO	0.2414	0.3671	0.2997	0.3510	0.3640
H (N)	H_2O	0.3336	0.4510	0.3551	0.3373	0.3460
H (N)	H_2O	0.3256	0.4400	0.3521	0.3306	0.3320
H (N)	H_2O	0.3329	0.4500	0.3561	0.3396	0.3476
H (N)	H_2O	0.3304	0.4391	0.3518	0.3338	0.3362
H (N)	H_2O	0.3235	0.4328	0.3464	0.3247	0.3291
H (N)	H_2O	0.3315	0.4478	0.3539	0.3346	0.3478
H (N)	H_2O	0.3316	0.4486	0.3656	0.3627	0.3725
H (N)	H_2O	0.3234	0.4325	0.3582	0.3519	0.3516
H (N)	H_2O	0.3301	0.4379	0.3621	0.3590	0.3566

[a] Attached atom in parentheses. [b] H_2O in $[Zn(OH_2)(NH_3)_3]^{2+}$ and HO in $[Zn(OH)(NH_3)_3]^{+}$.

to buffers in external water via His-64 or to buffer molecules in the aqueous network.[36] The position of His-64 is of particular interest because it is generally accepted that it is an important basic site in rate-limiting proton relays.[36]

This study will focus on modeling the structures of two of the intermediates shown in Scheme I. The first is the zinc–water form (**1**) and the second is the zinc–hydroxide form (**2**). These were chosen because they allow us to develop a force field for both the 2+ and 1+ formal charge states of the zinc ion with its associated ligands. The accuracy of the force field will be assessed by comparison of energy-minimized structures produced with the X-ray crystallographic structures. It is of particular interest to assess the ability to correctly position the ligands and the catalytically important residues near zinc.

Computational Procedure

Figures 1 and 2 show structural formulas for ions used as models for molecular orbital calculations in this study.

Comparative Study of Molecular Orbital Methods for Bonding at Zinc. We have performed semiempirical MNDO[38] and ab initio[31] calculations for two ions, $[Zn(OH_2)(NH_3)_3]^{2+}$ (**1a**) and $[Zn(OH)(NH_3)_3]^{+}$ (**1b**), in order to compare methods for molecular orbital calculations in modeling bonding for zinc in the active site of the enzyme.

All MNDO semiempirical molecular orbital calculations included full geometry optimizations obtained by using MOPAC 5.0 ESP (Unix version)[39] implemented on a Silicon Graphics Iris 4D/220 GTX system. MNDO derived electrostatic potential (ESP) charges for atomic centers were calculated by using the method of Besler et al.[33] The ab initio calculations were single determinant (Hartree–Fock) calculations using Gaussian 88[40] implemented on a Multiflow 14/300 and using the previously calculated MNDO optimized geometries. In the ab initio calculations four different basis sets, STO-3G,[31] MINI-4,[41] MIDI-4,[41] and a basis set we will designate as "6-31G*", were investigated. For the 6-31G* calculation a 5333/531/5 basis set split to 53321/531/5 was employed for zinc,[42] while the rest of the ligand atoms (H, C, N, and O) used the standard 6-31G* basis set.[31] Ab initio electrostatic potential charges were calculated for each atomic center with use of the method of Besler et al.,[33] which has been incorporated into Gaussian 88.[43]

Table II. Coefficients of Correlation (R) for Linear Regression Analyses of Electrostatic Potential Charges Calculated for **1a** and **1b** by Using Molecular Orbital Methods

MO method	semiemp MNDO	ab initio STO-3G*	ab initio MINI 4	ab initio MIDI 4	ab initio 6-31G*
semiemp MNDO		0.9712	0.9756	0.9816	0.9925
ab initio STO-3G*	0.9712		0.9280	0.9556	0.9720
ab initio MINI 4	0.9756	0.9280		0.9872	0.9854
ab initio MIDI 4	0.9816	0.9556	0.9872		0.9858
ab initio 6-31G*	0.9925	0.9720	0.9854	0.9858	

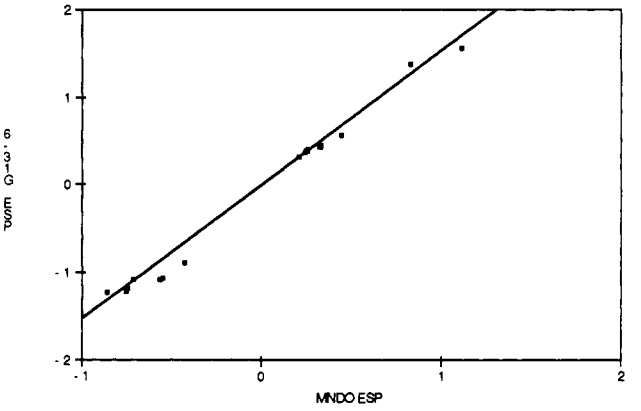


Figure 3. Linear regression analysis of electrostatic potential derived point charges.

(38) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899, 4907.
(39) Merz, K. M., Jr.; Besler, B. H. *QCPE Bull* **1990**, *10*, 15.
(40) Gaussian 88: Frisch, M. J.; Head-Gordon, M.; Schlegel, H. B.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 1988.
(41) Tatewaki, H.; Sakai, Y.; Huzinaga, S. *J. Comp. Chem.* **1981**, *2*, 278.
(42) Huzinaga, S.; Andzelm, J.; Klobukowski, M.; Radzio-Andzelm, E.; Sakai, Y.; Tatewaki, H.; *Gaussian Basis Sets for Molecular Calculations*; Elsevier: Amsterdam, 1984.

Table III. Electrostatic Potential Charges Calculated by Using MNDO for Small Model Structures **2b** and **2d**[a]

atom[b]	ESP charge: small model OH ligand **2b**	ESP charge: small model H_2O ligand **2d**	atom[b]	ESP charge: small model OH ligand **2b**	ESP charge: small model H_2O ligand **2d**
H (O)	+0.2975	+0.3995	C2	−0.0336	+0.0105
H (O)		+0.4013	N1 (H)	−0.2241	−0.2107
O (Zn)	−0.8242	−0.6197	C5	−0.1346	−0.0493
Zn	+0.6886	+0.8663	C4	−0.1596	−0.1703
N3 (Zn)	−0.1853	−0.2633	H (C2)	+0.2427	+0.2263
C2	−0.0048	−0.0075	H (N1)	+0.3215	+0.3432
N1 (H)	−0.2252	−0.1725	H (C5)	+0.2097	+0.2229
C5	−0.1308	−0.0900	H (C4)	+0.1848	+0.2016
C4	−0.1190	−0.1230	N3 (Zn)	−0.1576	−0.3138
H (C2)	+0.2215	+0.2341	C2	+0.0288	+0.0485
H (N1)	+0.3202	+0.3330	N1 (H)	−0.2689	−0.2215
H (C5)	+0.2092	+0.2300	C5	−0.0752	−0.0421
H (C4)	+0.1730	+0.1857	C4	−0.2177	−0.1455
N3 (Zn)	−0.1082	−0.2445	H (C2)	+0.2396	+0.2207
			H (N1)	+0.3276	+0.3446
			H (C5)	+0.1976	+0.2173
			H (C4)	+0.2073	+0.1885

[a] From MNDO. [b] Attached atom in parentheses.

Table IV. Partial Atomic Charges Assigned to Zinc and Ligand Atoms in Force Field Models

atom	hydroxide ligand: large model[a]	hydroxide ligand: small model[b]	hydroxide ligand: Zn^{2+} model	water ligand: large model[c]	water ligand: small model[d]	water ligand: Zn^{2+} model
Zn	+1.016	+0.688	+2.000	+1.033	+0.866	+2.000
O	−0.811	−0.824	−1.3170[e]	−0.629	−0.620	−0.834[f]
H	+0.253	+0.298	+0.317[e]	+0.378	+0.400	+0.417[f]
H				+0.378	+0.400	+0.417[f]

[a] Calculated by MNDO/ESP for structure **2a**. [b] Calculated by MNDO/ESP for structure **2b**. [c] Calculated by MNDO/ESP for structure **2c**. [d] Calculated by MNDO/ESP for structure **2d**. [e] Reference 25. [f] Reference 47.

Results for all molecular orbital calculations for ions **1a** and **1b** are shown in Table I. Table I clearly indicates that modeling zinc with a formal 2+ point charge is unrealistic. The variation of the computed ESP charge on zinc as a function of method is of interest. Thus, STO-3G* [31] gives the most covalent picture of bonding at zinc, while MINI-4[41] gives the most ionic picture. The larger MIDI-4 and "6-31G*" basis sets, as well as MNDO, give an intermediate picture. For **1a** the charge on zinc with hydroxide ligand varies from +0.83 to +1.72, while for **1b** with the water ligand the charge on zinc ranges from +1.11 to +1.64.

The sets of electrostatic potential charges calculated by different methods were analyzed for correlations with each other by linear regression analyses and the coefficients of correlation are shown in Table II. The highest correlation was observed between electrostatic potential charges derived from MNDO and ab initio 6-31G* methods, and the graph of the linear regression line is shown in Figure 3. These results demonstrate that electrostatic potential charges for zinc and ligand atoms computed from semiempirical MNDO wave functions are highly correlated to charges computed by more rigorous ab initio methods with use of the 6-31G* basis set. For this study it is important to note that the highly efficient MNDO method produces accurate results for Zn as well as H, C, N, and O. These correlations are consistent with correlation results reported in other studies which did not include zinc.[33,44] Given that the MNDO method is much faster and that it correlates reasonably well with the large basis set ab initio method, we decided to use the MNDO method in subsequent development of the force field.

Molecular Orbital Modeling of the Active Site of HCA. Structures **2a** and **2b** were adopted as "large" (84 atoms) and "small" (30 atoms) models for the active site of the hydroxide form of the enzyme. For **2b** MNDO calculations were performed with full geometry optimization and electrostatic potential (ESP) charges were calculated as described above. Calculated ESP charges are shown in Table III. For **2a** the coordinates for all non-hydrogen atoms in His-94, Gly-95, His-96, and His-119, zinc, and oxygen (bonded at zinc) for HCAII were obtained from the Brookhaven Protein Data Bank[45] and AMBER was used to add hydrogens in order to convert the PDB united atom representation to the all-atom representation. Methanoyl groups were substituted at N-terminal residues and *N*-methylamino groups were substituted at carbonyl groups at C-terminal residues. The resulting geometry was fully optimized to remove unfavorable contacts by using AM1[46] before performing a single-point MNDO calculation with no further geometry optimization. AM1 was used here in order to retain the hydrogen-bonding pattern present in the fragment structure of the enzyme active site model. MNDO derived electrostatic potential charges for **2a** and **2c** are provided as supplementary material in the form of an AMBER database file. In a similar manner ESP charges were computed for **2c** and **2d**, the "large" and "small" models of the water form of HCAII. Table III gives the charges for **2d**. MNDO calculations for the two active site models for the hydroxide form, **2a** and **2b**, produce quite different sets of ESP charges, and the same is true for models of the water form, **2c** and **2d**. These differences necessitated separate electrostatic models for the enzyme (see Figure 2). The ESP charge on zinc ranges from a minimum +0.69 with hydroxide ligand in the small model to +1.03 with water ligand in the large model, again demonstrating that the assumption of a +2 charge overestimates charge at the zinc (see Table IV).

Incorporation of Electrostatic Models. The three bonds to imidazole rings and the one bond to oxygen in hydroxide or water ligand were considered to be conserved and the bond, angle, and dihedral parameters are shown in Table V.

The charges resulting from electrostatic potential calculations were incorporated into the AMBER[1] force field. We will describe the procedure applied to the hydroxide case as an example; a similar procedure was followed for the water case. Enzyme models were developed as follows. The enzyme model labeled "large" was derived from active site model **2a**. The charges on Zn, O, H, all atoms in the imidazole rings, and C_β of histidines 94, 95, and 119 were extracted from the MNDO ESP calculations, while the remaining charge centers used the standard AMBER values. In order to ensure charge neutrality the charge at C_α was adjusted. Note that His-94 and His-96, which are coordinated to zinc in a similar manner, were not equivalent, in contrast to the small model (**2b**) described below.

The enzyme model labeled "small" was derived from active site model **2b**. The charges on Zn, O, H and the C, N, and H atoms in the imidazole rings were extracted and incorporated into the AMBER force field. Histidine residues 94 and 96 were made equivalent by averaging. AMBER charges for main chain atoms in the protein were preserved, and the charge at C_β was adjusted to preserve total charge.

(43) Besler, B. H.; Merz, K. M., Jr. Unpublished results.
(44) Gruschus, J. M.; Kuki, A. *J. Comp. Chem.* **1990**, *11*, 978. Orozco, M.; Luque, F. J. *J. Comp. Chem.* **1990**, *11*, 909.
(45) Bernstein, F.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, J. J. *J. Mol. Biol.* **1977**, *112*, 535.
(46) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

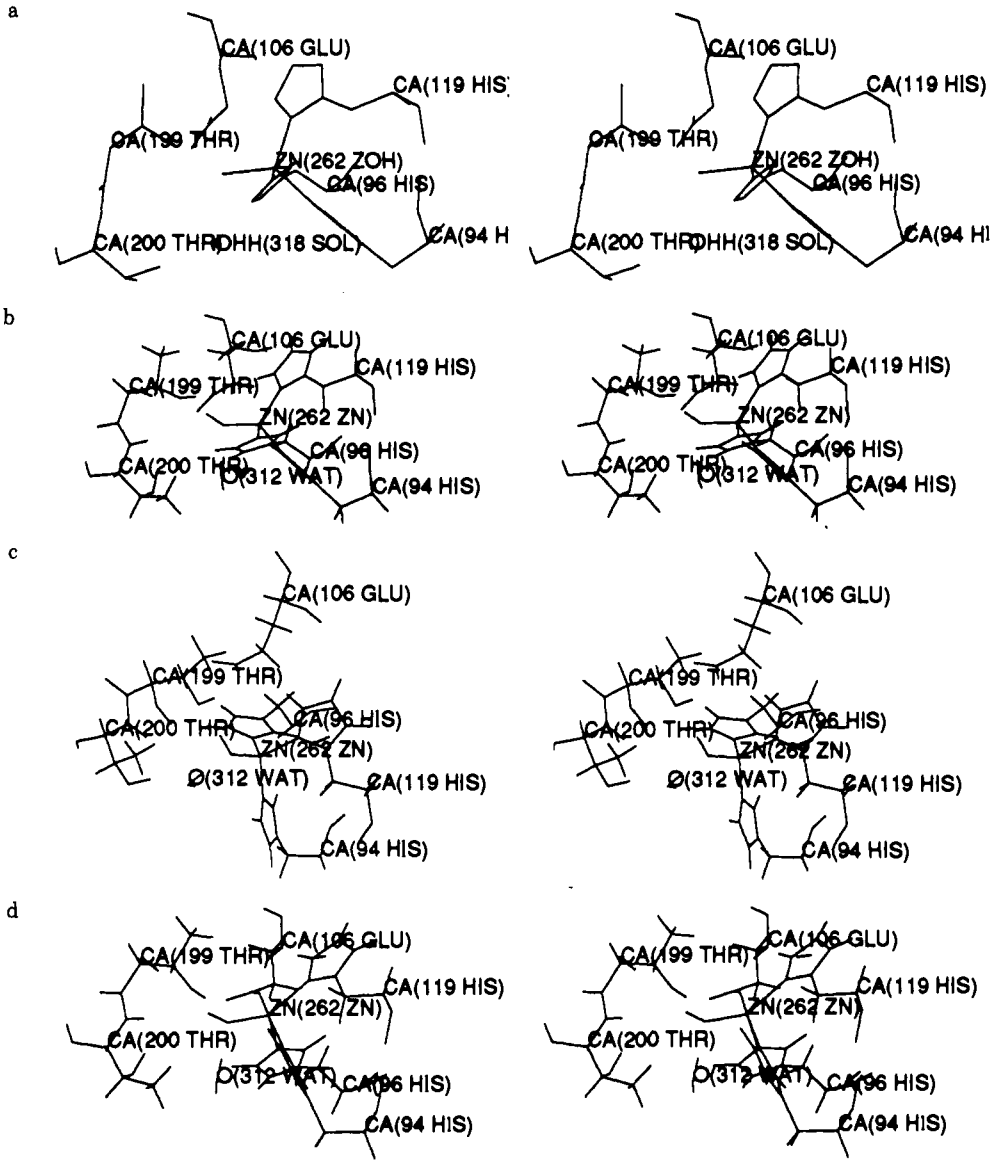


Figure 4. Stereodrawings of the active site of HCAII in the zinc–hydroxide form: (a) experimental X-ray structure; (b) HCAII large model; (c) HCAII small model; (d) HCAII Zn^{2+} model.

In the enzyme model labeled as "Zn^{2+}", the charge on Zn was assigned 2+ and the charges on O and H in hydroxide ion were based on ab initio 6-31G* ESP calculations.[31,33] The AMBER charges on all atoms in the amino acid residues bound to the zinc ion were not changed.

All assigned partial charges for AMBER database files for zinc and the hydroxide or water ligand are shown in Table IV. Assigned partial charges for AMBER database files for atoms in the histidine ligands are available as Supplementary Material.

Energy Minimization and Molecular Dynamics Simulations. The AMBER[47] suite of programs was used in the energy minimizations and MD simulations. Parameters derived here for bonds, angles, dihedrals, and Lennard-Jones potentials utilized in the AMBER force field are shown in Table V. Explicit details about how these values were determined are given elsewhere.[25] The structures included the crystallographically observed water molecules,[35] and these were represented with use of the TIP3P water model.[48] The initial structures, in an all-atom representation, were subjected to energy minimization with no constraints with use of the MIN module *in* AMBER.[47] Minimization was successfully achieved by using initial steepest descent minimization followed by conjugate gradient minimization with a convergence criterion of 10^{-6} kcal/mol. A nonbonded cutoff of 10 Å was used and was updated every 100 steps of minimization. SHAKE was used to constrain all bond lengths to their equilibrium values.[49] The MD simulations were done in the gas phase, and all residues were allowed to move. The duration of the simulation was 12 ps with a time step of 1 fs. The temperature was increased from 0 to 300 K and maintained at 300 K by coupling to the temperature bath.[50] The nonbonded cutoff was 10 Å and the non-

Table V. AMBER Force Field Parameters

atom	R^*	ϵ
Zn	1.1	0.0125
O	1.768	0.152
bond	r, Å	K_r, kcal/(mol·Å^2)
O–H	0.957	553
Zn–N	2.05	40
Zn–OH	1.80	94
Zn–OH_2	2.05	40
angle	θ, deg	K_θ, kcal/(mol·rad^2)
C–N–Zn	126	20
H–O–Zn	126	100
N–Zn–N	109.5	20
N–Zn–O	109.5	20
torsion	V_n, kcal/mol	γ, deg
X–Zn–N–X	0	0
X–Zn–O–X	0	0

(47) AMBER 3.0 UCSF: Singh, U. C.; Weiner, P. K.; Caldwell, J.; Kollman, P.

(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(49) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.

Table VI. Calculated Angles at Zinc in Models of HCAII with Hydroxide Ligand

		X–Zn–Y angle (deg) from Zn											
			energy minimization						MD equilibration					
X	Y	exp[a]	calc large **2a**	dev large **2a**	calc small **2b**	dev small **2b**	calc ionic **2+**	dev ionic **2+**	calc large **2a**	dev large **2a**	calc small **2b**	dev small **2b**	calc ionic **2+**	dev ionic **2+**
N94	O262	107.9	104.9	3.0	103.7	4.2	78.3	29.6	120.4	7.8	112.6	4.7	62.5	57.9
N96	O262	113.5	109.4	4.1	115.7	2.2	112.8	0.7	80.5	28.1	108.6	4.9	63.0	17.5
N116	O262	105.6	121.9	16.3	115.7	10.1	128.9	23.3	108.2	6.9	115.1	9.5	137.9	29.7
N94	N96	107.5	108.4	0.9	105.2	2.3	86.9	20.6	86.7	21.1	107.8	0.3	96.6	9.9
N94	N119	113.5	110.7	2.8	114.5	1.0	93.4	20.1	128.0	16.1	111.9	1.6	87.6	40.4
N96	N119	101.6	101.0	0.6	101.9	0.3	117.0	16.0	84.2	15.7	99.9	1.7	91.5	7.3

[a] Reference 35.

Table VII. Calculated Distances from Zinc with Hydroxide Ligand in Models of HCAII

		distance (Å) from Zn												
			energy minimization						MD equilibration					
residue	atom	exp[a]	calc large **2c**	dev large **2c**	calc small **2d**	dev small **2d**	calc ionic **2+**	dev ionic **2+**	calc large **2c**	dev large **2c**	calc small **2d**	dev small **2d**	calc ionic **2+**	dev ionic **2+**
His-119	Nε	1.91	1.97	0.06	1.97	0.06	2.22	0.31	2.05[b]	0.14	2.05[b]	0.14	2.05[b]	0.14
His-94	Nδ	1.99	1.91	0.08	1.94	0.05	2.20	0.21	2.05[b]	0.06	2.05[b]	0.06	2.05[b]	0.06
OH-262	O	2.09	1.77	0.32	1.76	0.33	1.94	0.15	1.80[b]	0.29	1.80[b]	0.29	1.80[b]	0.29
His-96	Nδ	2.10	1.99	0.11	1.95	0.15	2.01	0.09	2.05[b]	0.05	2.05[b]	0.05	2.05[b]	0.05
Thr-199	Oγ	3.83	3.90	0.07	3.82	0.01	3.51	0.32	5.43	1.60	4.18	0.35	3.81	0.02
H_2O-318	O	3.91	3.35	0.56	3.50	0.41	3.83	0.08	2.08	1.83	3.74	0.17	1.93	1.98
Glu-106	Oε	4.06	3.85	0.21	3.72	0.34	1.71	2.35	4.24	0.18	4.11	0.05	1.74	2.32
H_2O-338	O	4.09	4.47	0.38	4.45	0.36	4.53	0.44	3.58	0.51	3.89	0.20	4.03	0.06
H_2O-292	O	4.14	4.12	0.02	3.98	0.16	4.83	0.69	4.08	0.06	5.46	1.32	4.29	0.15
Glu-106	Oε	5.46	5.25	0.21	5.27	0.19	4.04	1.42	5.55	0.09	5.37	0.09	3.80	1.66
H_2O-264	O	5.84	5.89	0.05	5.59	0.25	6.25	0.41	4.66	1.18	5.55	0.29	5.10	0.74
H_2O-389	O	5.93	4.74	1.19	4.73	1.20	5.51	0.42	4.85	1.08	5.80	0.13	6.04	0.11
H_2O-265	O	5.51	5.54	0.47	5.15	0.34	5.16	0.35	5.45	0.47	6.56	0.64	6.46	0.54
H_2O-369	O	5.92	5.18	0.74	6.07	0.15	6.17	0.25	5.79	0.22	6.88	0.87	6.79	0.78
Thr-200	Oγ	6.42	5.92	0.50	6.02	0.40	6.25	0.17	6.26	0.16	5.02	1.40	6.88	0.46
Glu-117	Oε	6.45	6.55	0.10	6.59	0.14	6.57	0.12	6.53	0.08	6.68	0.23	6.31	0.14
H_2O-385	O	7.54	6.85	0.69	6.87	0.67	6.09	1.45	6.68	0.86	7.46	0.08	8.65	1.11
Glu-117	Oε	7.72	7.71	0.01	7.62	0.10	6.68	1.04	7.42	0.30	7.95	0.23	7.06	0.66
His-64	Nε	7.92	7.72	0.20	7.95	0.03	8.83	0.91	11.02	3.10	7.99	0.07	8.05	0.13
H_2O-295	O	7.93	7.80	0.13	9.17	1.24	8.32	0.39	7.48	0.45	8.18	0.25	9.80	1.87

[a] Reference 35. [b] Constrained by SHAKE.

Table VIII. Summary for Comparisons of Geometry: RMS Differences between X-ray Structure of HCAII[a] and Minimized and MD Equilibrated Models of HCAII with Hydroxide Ligand

	root mean square deviations					
	energy minimization			MD equilibration		
	large **2c**	small **2d**	ionic **2+**	large **2c**	small **2d**	ionic **2+**
rms bond angles at zinc, deg	±6	±4	±12	±16	±5	±30
rms distances from zinc, Å						
selected side chain N,O atoms[d] (0–5.0 Å)	±0.1	±0.2	±1.0	±0.7	±0.2	±1.1
selected side chain N,O atoms[d] (0–8.0 Å)	±0.2	±0.2	±0.7	±1.0	±0.4	±0.8
main chain N,C,C atoms (0.0–5.0 Å)	±0.3	±0.3	±0.4	±0.7	±0.6	±0.7
main chain N,C,C atoms (0.0–8.0 Å)	±0.3	±0.3	±0.4	±0.9	±0.6	±0.9
all atoms (0.0–8.0 Å)						
all atoms						

[a] Reference 35. [d] O,N atoms for side chains in residues listed in Table VII.

bonded pairlist was updated every 20 time steps.

Results and Discussion

The resulting energy minimized structures of the high pH (hydroxide) form of HCAII based on the "large" and "small" active site models are very similar with a coordination number of four and tetrahedral geometry at zinc (see Figure 4, b and c). These geometries compare favorably with the X-ray crystallographic structure (see Figure 4a).[35] In contrast, the minimized structure based on the ionic (Zn^{2+}) active site model has a coordination number of five for zinc (see Figure 4d). In a distorted trigonal-bipyramidal geometry Glu-106 is drawn into the coordination sphere at an axial position and His-94 occupies the other axial position. For detailed analysis we have tabulated the calculated angles and absolute deviations in angles from experiment[35] (X-ray crystallography) in Table VI. Similarly, distances from zinc to its ligands and neighboring polar atoms and absolute deviations in these distances are tabulated in Table VII. Table VIII shows rms deviations computed from the angle and distance results in Tables VI and VIII.

The results in Tables VI–VIII indicate that the two models using ESP-derived point charges produce energy minimized structures which more accurately predict the geometry (angles and distances) of the crystal structure than does a Zn^{2+} model. The ESP-based models position the catalytically important His-64 and the Thr-199 and Glu-106 subsystem within ±0.2 Å from zinc. However, the Zn^{2+} model which overestimates positive charge at the ion produces much larger deviations in bond angles and large displacements in the distances from zinc to nitrogen in His-64 (0.91 Å) and oxygens in Glu-106 (1.42 and 2.35 Å) and water-385 (1.45 Å) (see Table VII). Table VIII shows that the small and large models

(50) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

Table IX. Calculated Angles at Zinc in Models of HCAII with Water Ligand

		X–Zn–Y angle (deg) from Zn												
			energy minimization						MD equilibration					
X	Y	exp[a]	calc large **2a**	dev large **2a**	calc small **2b**	dev small **2b**	calc ionic **2+**	dev ionic **2+**	calc large **2a**	dev large **2a**	calc small **2b**	dev small **2b**	calc ionic **2+**	dev ionic **2+**
N94	O262	107.9	128.0	0.1	107.3	0.6	107.3	0.6	128.1	20.2	123.9	16.0	105.6	2.3
N96	O262	113.5	123.9	10.4	112.3	1.2	151.6	38.1	112.8	0.7	84.3	29.2	146.9	33.4
N116	O262	105.6	92.1	13.5	115.8	10.2	95.0	10.6	105.3	0.3	119.6	14.0	73.9	31.7
N94	N96	107.5	108.0	0.5	87.5	20.0	77.2	30.3	110.2	2.7	95.7	11.8	95.2	12.3
N94	N119	113.5	93.8	19.7	113.9	0.4	89.0	24.5	98.6	14.9	116.5	3.0	91.3	22.2
N96	N119	101.6	88.1	13.5	98.0	3.6	99.1	2.5	94.9	6.7	90.4	11.2	80.5	21.1

[a] With hydroxide ligand from ref 35.

Table X. Calculated Distances from Zinc with Water Ligand in Models of HCAII

		distance (Å) from Zn												
			energy minimization						MD equilibration					
residue	atom	exp[a]	calc large **2c**	dev large **2c**	calc small **2d**	dev small **2d**	calc ionic **2+**	dev ionic **2+**	calc large **2c**	dev large **2c**	calc small **2d**	dev small **2d**	calc ionic **2+**	dev ionic **2+**
His-119	Nε	1.91	2.10	0.19	2.03	0.12	2.26	0.35	2.05[b]	0.14	2.05[b]	0.14	2.05[b]	0.14
His-94	Nδ	1.99	2.05	0.06	1.96	0.03	2.33	0.34	2.05[b]	0.06	2.05[b]	0.06	2.05[b]	0.06
OH-262	O	2.09	1.98	0.11	1.82	0.27	2.25	0.16	2.05[b]	0.04	2.05[b]	0.04	2.05[b]	0.04
His-96	Nδ	2.10	2.07	0.03	1.94	0.16	2.29	0.19	2.05[b]	0.05	2.05[b]	0.05	2.05[b]	0.05
Thr-199	Oγ	3.83	3.80	0.03	3.82	0.01	5.98	2.15	4.44	0.61	4.40	0.57	6.07	2.24
H_2O-318	O	3.91	2.06	1.85	3.65	0.26	3.89	0.02	2.93	0.98	4.31	0.41	1.78	2.13
Glu-106	Oε	4.06	4.08	0.02	3.97	0.09	1.73	2.33	4.08	0.02	3.95	0.11	1.79	2.27
H_2O-338	O	4.09	4.01	0.08	4.56	0.47	4.29	0.20	4.49	0.40	2.31	1.78	3.77	0.32
H_2O-292	O	4.14	3.82	0.32	4.15	0.01	1.82	2.32	5.29	1.15	4.75	0.61	3.86	0.28
Glu-106	Oε	5.46	5.27	0.19	5.31	0.15	3.72	1.74	4.08	1.38	5.44	0.02	3.63	1.83
H_2O-264	O	5.84	5.80	0.04	5.75	0.09	4.34	1.50	5.34	0.50	5.16	0.68	4.68	1.16
H_2O-389	O	5.93	3.80	2.13	4.39	1.54	5.64	0.29	5.97	0.04	6.10	0.17	5.49	0.44
H_2O-265	O	5.51	5.61	0.31	5.59	0.33	6.94	1.02	6.81	0.89	6.48	0.56	5.84	0.08
H_2O-369	O	5.92	5.50	0.51	5.50	0.51	5.14	0.87	6.93	0.92	6.57	0.56	6.99	0.98
Thr-200	Oγ	6.42	4.50	1.92	5.95	0.47	6.07	0.35	4.12	2.30	4.43	1.99	4.69	1.73
Glu-117	Oε	6.45	4.08	2.37	6.66	0.21	6.70	0.25	6.70	0.25	6.78	0.33	6.68	0.23
H_2O-385	O	7.54	6.56	0.98	7.47	0.07	7.68	0.14	7.51	0.03	7.06	0.48	8.29	0.75
Glu-117	Oε	7.72	5.27	2.45	7.77	0.05	6.98	0.74	7.90	0.18	8.13	0.41	6.80	0.92
His-64	Nε	7.92	7.72	0.20	7.77	0.16	8.17	0.24	8.04	0.12	7.96	0.04	10.25	2.33
H_2O-295	O	7.93	8.82	0.90	8.59	0.67	9.02	1.10	8.00	0.07	7.36	0.57	10.12	2.19

[a] Reference 35. [b] Constrained by SHAKE.

produce bond angles with rms deviations of ±6° and ±4° compared to ±12° for the Zn^{2+} model for the zinc coordination sphere.

For further analysis we have analyzed rms deviations for distances between zinc and polar heteroatoms in side chains and between zinc and main chain atoms (N, Ca, and carbonyl C) in spheres of 5 and 8 Å centered at zinc. We have separated the effects of zinc representation on polar side chains and effects on secondary structure. The 8-Å sphere contains the structural features that have been postulated or identified as crucial for catalysis. Furthermore, it is expected that oxygen and nitrogen atoms in these polar side chains will be subjected to the strongest electrostatic interactions with the zinc ion and inaccurate charge representation can be assessed most directly. The results of this analysis are shown in Table VIII. The small and large models position side chain N and O atoms more accurately (±0.1 to ±0.2 Å) than the Zn^{2+} model (±0.7 to 1.0 Å) for both the 5- and 8-Å spheres. The three models position main chain atoms with nearly equal accuracy (±0.3 to ±0.4 Å); however, the Zn^{2+} model does give the largest variation (±0.4 Å). Thus, varying the representation of the positive charge at zinc produces larger effects on the placement of conformationally more mobile polar side chains than main chain atoms.

The results of short molecular dynamics (MD) simulations for each of the models are also shown in Tables VI–VIII. Only the small model retained a coordination number of four and tetrahedral geometry at zinc. This model produced an MD equilibrated structure with the smallest rms deviation in angles (±5°) and smallest rms deviations in distances (±0.4 Å for side chain atoms and ±0.6 Å for main chain atoms in the 8-Å sphere). Although the large model produced a relatively accurate minimized structure, MD equilibration produces a structure with five-coordinate zinc with large rms deviations in angles (±16°) and larger rms deviations for distances in the 8-Å sphere for both side chains (±1.0 Å) and main chain atoms (±0.9 Å). The Zn^{2+} model, which had produced a minimized structure with poor correlation with the X-ray structure, produced a more severely distorted MD equilibrated structure with six-coordinate zinc and the largest rms deviations in angles of ±30° and a large rms deviation in distances (±0.8 Å for side chain atoms and ±0.9 Å for main chain atoms in the 8-Å sphere). These results emphasize the importance of the local minimum problem in energy minimization and the importance of MD simulation as a tool for full assessment of force field models.

In conclusion, it is evident that the small model based on ESP fitted point charges produces the most accurate minimized and MD equilibrated structure. Both the large and small models yield superior structures compared to a Zn^{2+} model. Combined with the results of ESP fitting of the active site using structures **2a** and **2b** these simulation results demonstrate that a 2+ charge on zinc represents a large overestimation of charge and produces significant error in energy minimization and MD equilibration. Thus, results from force field based simulations based on this approach should be treated with caution.[27,30] Furthermore, we find that the small model performs better than the large active site model. The reasons for this are not clear, but it potentially has to do with the larger positive charge on the zinc ion in the large active site model.

This study included modeling of the form of HCAII with the water ligand bonded at zinc (see Scheme I). For the sake of completeness our results for minimizations and MD simulations with all three models (**1b**, **2c**, and **2d**) are shown in Tables IX–XI. After the completion of this research, Christianson obtained an X-ray crystallographic structure of the zinc–water form of this enzyme. The zinc ion is four coordinate possessing tetrahedral

Table XI. Summary for Comparisons of Geometry: RMS Differences between X-ray Structure of HCAII[a] and Energy Minimized and MD Equilibrated Models of HCAII with Water Ligand

	root mean square deviations					
	energy minimization			MD equilibration		
	large **2c**	small **2d**	ionic **2+**	large **2c**	small **2d**	ionic **2+**
rms bond angles at zinc, deg	±8	±9	±12	±11	±16	±21
rms distances from zinc, Å						
selected side chain N,O atoms[d] (0–5.0 Å)	±0.0	±1.0	±0.1	±0.3	±0.2	±0.0
selected side chain N,O atoms[d] (0–8.0 Å)	±1.4	±0.6	±0.8	±0.7	±0.5	±1.0
main chain N,C,C atoms (0–5.0 Å)	±0.4	±0.2	±0.5	±0.7	±0.7	±0.8
main chain N,C,C atoms (0–8.0 Å)	±1.1	±0.3	±0.4	±1.4	±0.8	±0.9
main chain N,C,C atoms (0.0–10.0 Å)	±1.37	±0.26	±0.37	±1.4	±1.0	±1.1

[a] Reference 35. [d] O,N atoms for side chains in residues listed in Table X.

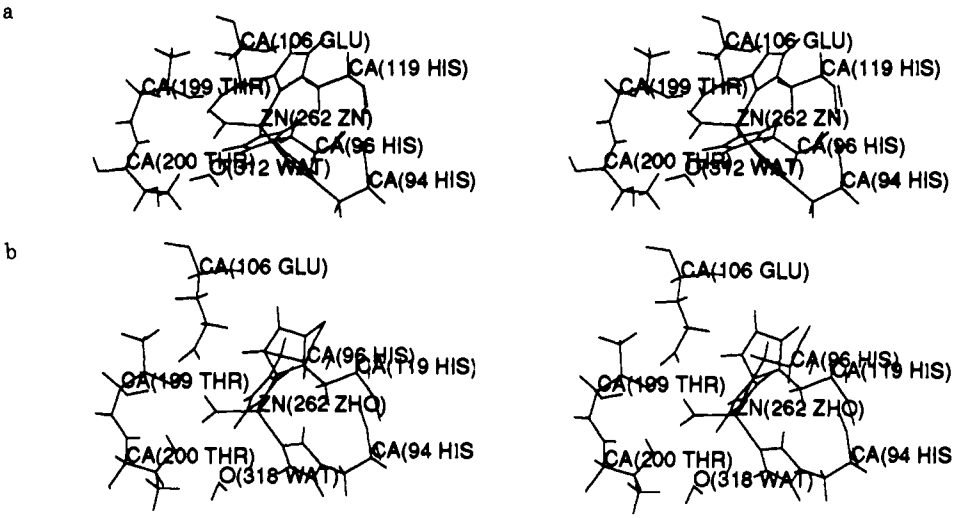


Figure 5. Stereodrawings of the active site of HCAII in the zinc–water form (small model only): (a) energy minimization only; (b) energy minimization and 12-ps MD simulation.

geometry, and research on placement of water molecules continues.[37] On the basis of our results in the hydroxide form of the enzyme, we expect the small model on ESP-derived point charges to provide the most accurate predictions. In energy minimization the small model produces a structure for the water form that has the smallest structural changes for formation from the hydroxide form. The minimized and MD equilibrated structures based on **2d** contain four-coordinate, tetrahedral zinc. This is consistent with experimental observation.[37]

The results of energy minimizations and MD simulations have been analyzed in more detail relative to the X-ray structure of the hydroxide form. The calculations reflect displacements during the process of protonating hydroxide ligand. In general, the small model predicts small displacements for polar groups within 8 Å of the zinc in the active site (see Table X). The hydrogen-bonded subsystem consisting of Glu-106, Thr-199, and one HO bond on zinc in the hydroxide form is retained in the water structure. In the energy minimized structure Oγ of Thr-200 is drawn into a position such that a water molecule serves as a hydrogen-bonding bridge between the hydroxyl hydrogen and Oγ at Thr-200 (see Figure 5). In the structure produced by MD simulation Oγ is drawn closer to zinc and the intervening water molecule is displaced. A hydrogen on the zinc-bound water ligand interacts with Oγ of Thr-200 and Oε of Glu-106 in a bifurcated hydrogen-bonding interaction. This may have mechanistic implications. The zinc–hydroxide mechanism for hydration of CO_2 requires abstraction of proton from the water ligand before nucleophilic attack. Perhaps the role of Thr-200 is to "lock" a water molecule into an appropriate position for accepting this proton during the first proton transfer in proton relays. The MD equilibrated structure suggests the first proton transfer might involve Oγ of Thr-200. It should also be pointed out that the structures do have appropriate geometries for initial proton transfer to Thr-199 and subsequent proton transfer to Glu-106 as suggested by Kannan.[34] However, this relay mechanism is considered to be unlikely because computational studies suggest the pK_a of Glu-106 is not sufficiently perturbed to make Glu-106 the group critical for catalysis.[26]

Conclusions

We have described and tested a systematic approach by which a metal ion can be incorporated into a simple force field. The key feature of our approach is in development of a viable electrostatic model for the metal ion and its coordination sphere. This was done by using an ESP-fitting procedure based on MNDO calculations. Force constants and Lennard-Jones parameters were taken from experimental and ab initio data, respectively.

A key issue remains unresolved. First, we found that ESP fitting worked for one active site model (small) but not another (large). Hence, the best way in which to consistently obtain accurate partial charges still remains an open issue. Furthermore, it might be best to abandon the simple point charge model and represent metal ions and their coordination sphere by using a multipolar representation. The observed distortion of the large model could be potentially mitigated by adjusting the force constants for the bond and angle terms linking the zinc ion to its ligands. This is artificial since our force constants were taken from experimental data. However, reexamination of this issue by determining ab initio force constants would be worthwhile.

Supplementary Material Available: Tables of partial atomic charges for histidine 91, 93, and 119 (2 pages). Ordering information is given on any current masthead page.