A New Force Field for Modeling Metalloproteins

Angelo Vedani* and David W. Huhta

Contribution from the Biographics Laboratory, Departments of Chemistry and Biochemistry, University of Kansas, Lawrence, Kansas 66045. *Received November 13, 1989*

Abstract: We present the development of a force field for the simulation of metalloproteins featuring a new potential function for modeling metal–ligand interactions. This function includes as variables the metal–ligand separations, the symmetry at the metal center, directionality of the metal–ligand bonds, ligand–metal charge transfer, and (for transition-metal ions) ligand-field stabilization. The function was developed based on the analysis of accurate small-molecule crystal structures retrieved from the Cambridge Structural Database and incorporated into the molecular mechanics program "YETI" which also includes directional terms for H-bonds and salt linkages in its force field energy expression. The program was then used to model details of metal-coordination, H-bond network formation and protein–solvent interactions in native, complexed, and Co(II)-substituted human carbonic anhydrase I.

Most molecular mechanics programs in general use[1-5] treat metal–ligand interactions by defining a covalent bond between the metal ion and the ligand atom ("bonded approach") or, alternatively, by using electrostatic and van der Waals forces instead ("nonbonded approach"). In the "bonded approach", the results are biased insofar as the optimal geometry of the metal center must be defined beforehand. As a consequence, the metal can hardly change its coordination type during the refinement. This implies that subtle effects in the ligand sphere, such as the ones proposed by Dutler for zinc enzymes,[6] cannot be simulated. By using a "nonbonded approach", on the other hand, difficulties arise mainly from the chosen electrostatic model.[7-9] Calculations based on inappropriate atomic partial charges or an unrealistic dielectric parameter can lead to atomic arrangements around metal centers that lack any resemblance to the more frequently observed types found in small-molecule crystal structures;[10] in the presence of small, highly charged ligands (e.g., OH^-) they can also fail to reproduce experimental metal–ligand bond lengths.[7]

In a previous paper[10] we have described our first attempt to derive a potential function for modeling metal centers in proteins. This function was developed based on a study of small-molecule crystal structures containing four-, five-, and six-coordinate zinc, retrieved from the Cambridge Structural Database.[11] The function included as variables the metal-ligand separations and the angles subtended at the metal. These two variables allowed specifically for distortions from and transitions between frequently occurring types of coordination geometries (e.g., tetrahedron, square plane, square pyramid, trigonal bipyramid, and octahedron). The function was then incorporated in the molecular mechanics program "YETI" and used to study various complexes of native and complexed human carbonic anhydrase I and II.[10,12]

Table I. Metal–Ligand Distances Observed in Small-Molecule Crystal Structures Retrieved from the CSD, V3.4[11]

CN[a]	ligand type	no. of observtns[b]	av M–L, Å	std dev,[c] Å	*d*(min), Å	*d*(max), Å
			A. Zinc(II)			
4	O	52	1.961	0.028	1.913	2.036
	N	136	2.041	0.047	1.917	2.214
	S	44	2.317	0.048	2.252	2.496
5	O	41	2.068	0.098	1.926	2.390
	N	88	2.126	0.086	1.995	2.389
	S	3	2.412	0.089	2.361	2.514
6	O	194	2.106	0.069	1.973	2.362
	N	60	2.147	0.063	2.051	2.287
	S	0				
			B. Co(II)			
4	O	38	1.918	0.069	1.838	2.061
	N	119	1.964	0.091	1.791	2.159
	S	29	2.282	0.050	2.118	2.342
5	O	26	1.989	0.034	1.862	2.377
	N	54	2.045	0.119	1.804	2.249
	S	4	2.327	0.034	2.300	2.374
6	O	325	2.090	0.043	1.941	2.325
	N	213	2.120	0.083	1.873	2.302
	S	29	2.452	0.120	2.250	2.585

[a] Coordination number; four-coordinate includes tetrahedral geometries only. [b] The number of observations corresponds to the number of individual metal–ligands bonds. [c] This standard deviation represents the distribution width of the CSD search; the uncertainties of the individual bond lengths are significantly smaller, typically 0.001–0.005 Å.

Although the metal-center function proved to be more versatile than a "pure bonded" or a "pure nonbonded" approach, it failed to model details of the "ligand-sphere transition" mechanism[6] and to handle metal-bound anions in a reasonable fashion.[7]

In this paper we describe results of a new analysis of the environments of Zn(II) and Co(II)[13] in small-molecule crystal structures retrieved from the Cambridge Structural Database.[11] These results are then used as the basis for an extended potential function to model metal centers in proteins. The new potential function has again been incorporated into the molecular mechanics program "YETI" (Version 4.5) and used to model native, complexed, and Co(II)-substituted human carbonic anhydrase I.

Methods

1. Analysis of Zn(II) and Co(II) Complexes. The Cambridge Structural Database (CSD), version 3.4,[11] was used to analyze the environments of zinc(II) and cobalt(II) in accurate small-molecule crystal structures. Because low-symmetry atomic ar-

(1) Singh, U. C.; Weiner, P. K.; Caldwell, J. W.; Kollman, P. A. AMBER (UCSF Version 3.0), Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, 1986.
(2) Brooks, B. R.; Bruccolieri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(3) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
(4) Tripos Associates, 1699 S. Hanley Road, Suite 303, St. Louis, MO 63144, SYBYL 5.2, 1988.
(5) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; ACS Monograph No. 177, American Chemical Society: Washington, D.C., 1982.
(6) Dutler, H.; Ambar, A. In *The Coordination Chemistry of Metalloenzymes*; Bertini, I., Drago, S. R., Lucchinat, C., Eds.; NATO Adv. Study Inst. Series; Reidel: Dordrecht, 1982; pp 135–145.
(7) Vedani, A.; Huhta, D. W.; Jacober, S. P. *J. Am. Chem. Soc.* **1989**, *111*, 4075–4081.
(8) Pettitt, B. M.; Karplus, M. *J. Am. Chem. Soc.* **1985**, *107*, 1166–1173.
(9) Berkovitch-Yellin, Z.; Leiserowitz, L. *J. Am. Chem. Soc.* **1980**, *102*, 7677–7690.
(10) Vedani, A.; Dobler, M.; Dunitz, J. D. *J. Comput. Chem.* **1986**, *7*, 701–710.
(11) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr., Sect B* **1979**, *B35*, 2331–2339.
(12) Vedani, A. *J. Comput. Chem.* **1988**, *9*, 269–280.
(13) Since 1984/5, when two previous analyses on Zn(II) have been published,[10,14] the number of structures stored in the CSD[11] increased from 45 000 to over 70 000.

0002-7863/90/1512-4759$02.50/0

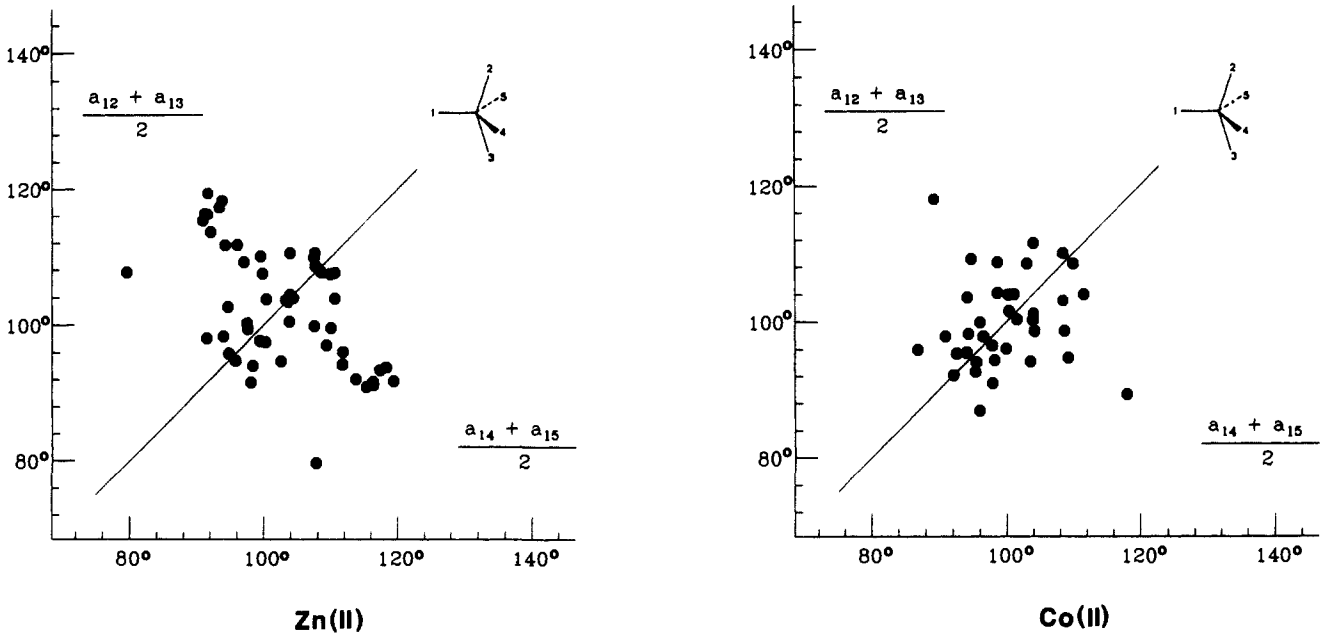

Figure 1. Distribution of $(a_{12} + a_{13})/2$ vs $(a_{14} + a_{15})/2$ angles for five-coordinate Zn(II) and Co(II) compounds with N, O, and S ligands. The Berry pseudorotation coordinate lies approximately normal to the diagonal which was taken as a line of symmetry.

rangements in condensed phases are difficult to classify in terms of discrete coordination numbers (see, for example, refs 15 and 16), we arbitrarily defined the coordination shell of Zn and Co to include all first-row atoms (N, O) closer than 2.40 Å and all second-row atoms (S, Cl) closer than 2.60 Å. The criteria for accepting an entry from the database were as follows: (1) the metal should have at least one O, S, or N ligand, (2) no crystallographic disorder or errors, (3) no strong metal–metal interactions, (4) no ligands bridging two metal centers, (5) no metal-π interactions, and (6) a crystallographic R-factor of 0.06 or less.

We found 89 structures containing four-coordinate zinc(II). On the basis of comparison of the angles subtended at the metal to the ideal value, 85 of the 89 were classified as tetrahedral. The 28 five-coordinate fragments were further analyzed to find the distribution between trigonal bipyramids (TBP) and square pyramids (SQP). As shown in Figure 1 for five-coordinate Zn(II) and Co(II) complexes, the preference for trigonal bipyramids or square pyramids can be visualized by plotting the angles subtended at the metal following Bürgi and Dunitz.[17] The distribution of structures along the "Berry pseudorotation" coordinate[18] is an indicator for the barrier height of the transition between TBP and SQP. The almost random experimental distribution suggests that, for Zn(II) systems, this barrier is not particularly high. The 43 six-coordinate structures were almost exclusively octahedral with deviations arising from influence of the ligand geometry. Mean metal–ligand distances are given in Table I.[19]

As shown in Figure 2, there is a clear preference for metal-bound ligand atoms to orient their lone pair towards the metal center like a H-bond acceptor atom prefers to orient its lone pair towards the donor-H atom (cf. refs 20–22). For modeling

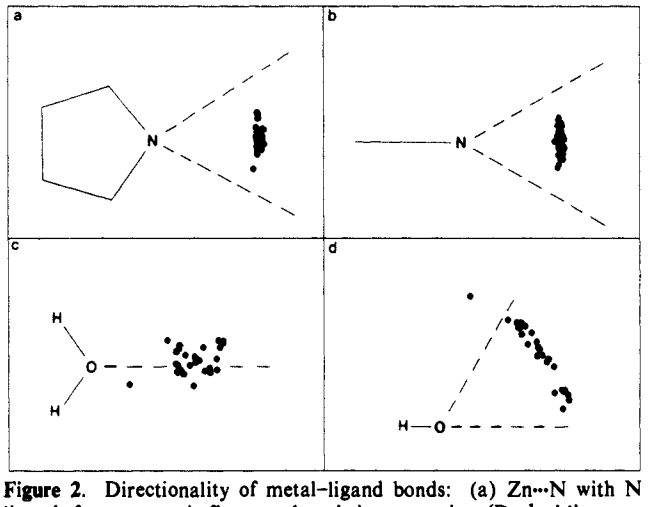

Figure 2. Directionality of metal–ligand bonds: (a) Zn···N with N ligands from aromatic five-membered rings, top view (Dashed lines are drawn at ±30° from the C–N–X (X = C, N) bisector.); (b) side view (Dashed lines are drawn at ±30° from the C–N–X (X = C, N) bisector.); (c) Zn···O with O ligands from H_2O, top view (Dashed line indicates H–O–H bisector.); and (d) side view (Dashed lines indicate sp^2 (0°) and sp^3 directions (60°).).

complexes of human carbonic anhydrase (see below), imidazole (Figure 2a,b) and water ligands (Figure 2c,d) around zinc were of particular interest.

Our search identified 70 structures containing four-coordinate cobalt(II). Tetrahedrons were formed by 51 with the balance being square planes. The distribution of the 19 five-coordinate cobalt species indicates a clear preference for SQP (15 fragments) vs TBP (1 fragment), with three intermediate arrangements (cf. Figure 1). Preference for SQP arises from the influence of ligand-field stabilization (LFS) which favors the square-pyramidal arrangement for both high- and low-spin configuration.[23] The 98 six-coordinate cobalt(II) structures are, as found for Zn(II), essentially octahedral with distortions caused by ligand constraints. Mean metal–ligand distances are given in Table I.[19]

2. The New Metal-Center Function. Most molecular mechanics programs in general use[1-5] treat metal–ligand interactions either

(14) Auf der Heyde, T. P. E.; Nassimbeni, L. R. *Acta Crystallogr. Sect B* **1984**, *B40*, 582–590.
(15) Schwarzenbach, D.; Brunner, G. O. *Z. Kristallogr.* **1971**, *133*. 127–133.
(16) O'Keeffe, M. *Acta Crystallogr.* **1979**, *A35*, 772–775.
(17) Bürgi, H. B.; Dunitz, J. D. *Acc. Chem. Res.* **1983**, *16*, 153–161.
(18) "Berry pseudorotation" refers to the conversion of a TBP to a SQP by angular deformation.
(19) The variation in metal–ligand distances is larger for axial than for equatorial ligands; indeed, in a few square pyramids, trigonal bipyramids, and octahedra the axial ligands are so distant from the metal center (but still included in the above definition) that the coordination types could also be described as (4 + 1), (3 + 2), and (4 + 2), respectively.
(20) Murray-Rust, P.; Glusker, J. P. *J. Am. Chem. Soc.* **1984**, *106*, 1018–1025.
(21) Taylor, R.; Kennard, O. *Acc. Chem. Res.* **1984**, *17*, 320–326.
(22) Vedani, A.; Dunitz, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7653–7658.
(23) Huhey, J. E. *Inorganic Chemistry—Principles of Structure and Reactivity*, Harper & Row: London, 1975; p 336, 381.

$$E_{total} = \sum_{\text{dihedrals}} \frac{V_n}{2}\left(1+\cos\left(n\phi-\gamma\right)\right) + \sum_{\text{all pairs}} \frac{1}{4\pi\varepsilon_0}\cdot\frac{q_i q_j}{D(r)\,r_{ij}} + \sum_{\text{nb pairs}} \frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}}$$

$$+ \sum_{\text{H-Bonds}} \left(\frac{A'}{r_{H\cdots Acc}^{12}} - \frac{C'}{r_{H\cdots Acc}^{10}}\right)\cdot\cos^2\left(\theta_{Don-H\cdots Acc}\right)\cdot\cos^n\left(\omega_{H\cdots Acc-LP}\right)$$

$$+ \sum_{\text{ML pairs}} \left(\frac{A''}{r_{M\cdots Lig}^{12}} - \frac{C''}{r_{M\cdots Lig}^{10}}\right)$$

$$+ \left(E^{\circ}_{MC} + E_{LFS}\right)\cdot\prod_{\substack{\text{indep.}\\ \text{angles}}}\cos^2\left(\Psi_{Lig-M-Lig'}-\Psi_0\right)\cdot\frac{1}{n}\sum_{\substack{n\\ 1^{st}\text{shell}\\ \text{ligands}}}\cos^n\left(\omega_{M\cdots Lig-LP}\right)$$

Figure 3. Force field energy expression in "YETI" V4.5.

by a "pure bonded" approach (i.e., defining the metal–ligand bond as a covalent bond and using appropriate parameters for bond stretching and angle bending) or by a "pure nonbonded" approach (i.e., by not defining a covalent bond between metal and ligand but treating the interaction by means of electrostatic and van der Waals forces instead).

By using a "pure bonded approach", most coordination types around metals can be reproduced from a geometrical point of view. Disadvantages of this approach include the harmonicity of the bond-stretching potential which does not seem appropriate to model the experimentally observed asymmetric distribution of metal–ligand bond lengths. Most critically, the metal can hardly change its coordination type but certainly not its coordination number during such a simulation, since covalent bonds can neither be formed nor cleaved during molecular mechanics optimizations. This implies that the preferred coordination geometry of the metal ion cannot be determined in an unbiased fashion; therefore this approach should only be used where there is clear experimental evidence for the type of metal-ion coordination.

By using a "pure nonbonded approach", on the other hand, problems arise mainly from the chosen electrostatic model.[7-9] Calculations based on inappropriate atomic charges (e.g., assigning formal charges to the metal ion) or using an unrealistic dielectric parameter will lead to atomic arrangements around metal centers that lack any resemblance to the more frequently observed types found in small-molecule crystal structures.[10] To overcome the strong electrostatic attraction between metal and ligands (leading to unreasonably short metal–ligand bonds), certain force fields assign significantly higher well-depths for the 6/12 function (cf. ref 24 for the terminology) and larger van der Waals radii for the metal ion.[25]

To achieve a compromise between these two fundamental approaches, we have developed an empirical potential function for modeling metal centers in macromolecules: The new function (Figure 3) includes two major terms, one describing the radial behavior of the metal–ligand interactions, the other analyzing the first ligand sphere at the metal.

Apart from the metal-ion and ligand-atom type, the radial term of this function depends solely on the metal–ligand distance. The summation extends over all potential metal–ligand pairs (the term "ligand" presently applies to all O, N, and S atoms/ions capable of metal-coordination, i.e., having at least one available lone pair). This nonbonded type approach allows for mobility of all ligands between various shells, i.e., the metal can change both number and arrangement of its proximal ligands during a refinement.

An exchange of ligands between the first and second coordination shell has been postulated by Dutler in the "ligand-sphere transition" model for the catalytic mechanism in zinc enzymes with particular reference to liver alcohol dehydrogenase.[6]

The coefficients A″ and C″ depends on the equilibrium distance (r_0) and on the well-depth (E_0) for each M···L interaction, both of which depend on the coordination type as well as on the nature of the atoms involved. Values for r_0 were obtained from the CSD[11] (cf. Table I), those assigned to E_0 were estimated from semi-empirical calculations on model systems (cf. below).

One disadvantage of this approach is that axial and equatorial ligands are not treated differently. The analysis of the data retrieved from the CSD shows that (4+1), (3+2), and (4+2) distorted polyhedra are fairly common. Our previous metal-center function[10] treated equatorial ligands (narrow experimental distribution of metal–ligand bond lengths) with a 10/12 function, whereas axial ligands (broader distribution of metal–ligand bond lengths) were treated with a 6/12 function. We feel, however, that for macromolecular applications an unbiased determination of coordination number and type is more important than a slightly different treatment of axial and equatorial ligands. Moreover, the calibration of such subtleties would require more highly resolved structural data than is presently available for metalloproteins.

The directional term of the metal-center function analyzes the first ligand shell at the metal: its energy depends on the symmetry at the metal center, the directionality of the metal–ligand bonds, and (for transition-metal ions) the ligand-field stabilization. A weighting factor determines the absolute value of E°_{MC} and allows variation of the ratio of radial/directional energy. Best results were obtained by assigning a weight from 0.75–0.667 to the radial term and a corresponding weight of 0.25–0.333 to the directional term. By assigning a weight of 1.00, the directional term can be totally disabled.

Optimal values for Ψ_0 used in the symmetry term are given by symmetry or were obtained by analysis of structural data retrieved from the CSD[11] (cf. above). Although all Lig-Met-Lig′ angles are explicitly evaluated for this term, the summation extends only over the independent angles (tetrahedron and square plane 5 out of 6 angles; square pyramid and trigonal bipyramid 7/10; octahedron 9/15, respectively). Values for the coefficient "n" in the directional term are identical with the ones used in the H-bond function (see below). Values for the ligand-field stabilization can be obtained from theoretical considerations or experimental data.[23,26]

Probably the most important parameter for modeling metal–ligand interactions is the charge distribution between metal ion and ligand atoms. It is obvious that the use of formal charges on the metal ion (e.g., +2.0 for Zn(II)) will overestimate the electrostatic interactions. The weight of the electrostatic term (corresponding to the ratio of ionic/covalent character for a particular metal–ligand bond) is critical for modeling transition-metal elements where the metal–ligand bond has a finite covalent character. Ligand–metal charge transfer in zinc enzymes has previously been discussed by Pullman in model studies for carbonic anhydrase,[27] Hayes and Kollman for carboxypeptidase,[28,29] and Giessner-Prettre and Jacob for thermolysin.[30]

Unfortunately, presently available molecular mechanics programs do not allow modification of the charge distribution between metal and ligand atoms during the refinement. We have, therefore, attempted to develop an empirical function for ligand–metal charge transfer, which, based on theoretical or experimental evidence, allows dynamic adjustment during the refinement. The main idea behind this function is that a reasonable estimate for the charge distribution between metal and ligands can be obtained for small, well-characterized systems. As a first approximation we have used the difference in electronegativity to estimate the extent of charge transfer. This yields atomic charges in our model systems that are in qualitative agreement with those calculated by semiempirical methods (MNDO as implemented in AMPAC[31] using zinc param-

(24) Margenau, H.; Kestner, N. *The Theory of Intermolecular Forces*; Pergammon Press: Oxford, 1970.
(25) Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. TRIPOS Technical Newsletter, 1988, 1/5; *J. Comput. Chem.* Submitted for publication.
(26) Cotton, F. A.; Wilkinson, G. *Inorganic Chemistry*, 4th ed.; Wiley & Sons: New York and London, 1980; p 646.
(27) Pullman, A. *Ann. New York Acad. Sci.* **1981**, *367*, pp 340–355.
(28) Hayes, D. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1976**, *98*, 3335–3345.
(29) Hayes, D. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1976**, *98*, 7811–7816.
(30) Giessner-Prettre, C.; Jacob, O. *J. Computer-Aided Molecular Design*; **1989**, *3*, 23–37.
(31) Distributed by *QCPE*, University of Indiana, Bloomington, IN. Original reference on MNDO method: Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.

eters obtained from the literature[30]). Further extensions of this work will include an extensive calibration of the function using input from MO calculations. For the system $Zn(H_2O)_n$ (n = 1–4), energies of the metal–ligand bonds were also estimated from PRDDO calculations. For this purpose, four water molecules were sequentially added to the zinc ion to form a tetrahedron.[32]

On the basis of a theoretical or experimental value for a model system, our function controls the amount of charge transfer depending on the actual metal–ligand distance by using two reference points: At a distance equal to the reference distance, the amount of charge transfer corresponds to the theoretical or experimental value derived for the model compound; at infinite separation charge transfer vanishes (i.e., pure electrostatic/van der Waals interactions prevail). In between, the amount of charge transfer falls off exponentially. The steepness of this exponential decay can also be used to control the amount of charge transfer among four-, five-, and six-coordinate species, for which different mean metal–ligand separations are observed (cf. Table I).

For distances significantly shorter than the reference distance (as they might occur at the beginning of a molecular mechanics refinement), this function would lead to an unrealistic charge distribution by transferring too much charge from the ligand atom to the metal ion. In such cases, the charge transfer is disabled and, instead, electrostatic interactions involving the metal ion are damped until a reasonable geometry is obtained (usually after one iteration cycle).

For modeling complexes of human carbonic anhydrase, the metal-center and charge-transfer functions were calibrated by using the model systems ML_n: M = Zn(II), Co(II); L = H_2O, NH_3, H_2S, OH^-, SH^-; n = 4, 5, 6 (except for L = H_2S, SH^- and M = Zn(II) where n = 4, 5, since no corresponding six-coordinate structures were found in the database), and mean M–L distances obtained from the search of the CSD.[11]

3. Treatment of H-Bonds. Directionality of H-bonds has first been discussed by Kroon et al. in 1975. Their analysis of 45 small-molecule crystal structures (containing 195 O–H···O H-bonds) showed a preference for the donor H atom to cluster at the acceptor O atom around the bisector of the R–O–H angle (see Figure 6, ref 33). In 1984, Murray-Rust and Glusker analyzed the spatial geometry of H-bonds in small-molecule crystal structures with particular reference to preferred directions at O acceptor fragments.[20] By using data retrieved from the CSD,[11] they showed that H-bond donors are concentrated in directions commonly ascribed to the lone-pair orbitals of the O acceptor atom. Similarly, Kennard and Taylor studied H-bonds involving carbonyl and hydroxyl O atoms as H-bond acceptors;[21] Dunitz and Vedani studied hydroxyl and sulfonamide O as well as N acceptor atoms in aromatic five- and six-membered rings.[22] In 1984, Baker and Hubbard published a very detailed study on the geometry of H-bonds in high-resolution protein structures.[34]

In 1985, we proposed an extended potential function to allow for directionality of H-bonds in molecular mechanics calculations.[22] This function includes as an additional variable (besides the H···Acc distance and the linearity of the H-bond, i.e., the angle subtended at the H atom) the deviation of the H-bond from the closest lone-pair direction at the acceptor atom (angle H···Acc-LP; see Figure 3 and refs 7 and 22). A critical parameter for the evaluation of this "explicit" H-bond term (electrostatic contributions are calculated separately) is the exponent "n" weighting the penalty for the deviation of the actual H-bond from the closest lone-pair direction at the acceptor fragment. The exponent "n" was calibrated for each individual H-bond acceptor to give the best possible agreement with the experimental distribution.[20-22] Details are given in Table II.

The H-bond function in "YETI" was calibrated by using small-molecule systems for which structural and, preferably, energetic data were available. If not otherwise cited, structural information was obtained from the CSD.[11] O–H···O H-bonds were calibrated by using the water dimer[35,36] and a homodromic water pentagon.[37] Energies of these systems were calibrated by using experimental data[38] as well as ab initio calculations with a 6-31G* basis set (as implemented in GAUSSIAN 86[39]). N–H···O H-bonds were calibrated by using the *N*-methylacetamide dimer and geometric data for >N–H···O=C< H-bonds found in accurate small-molecule crystal structures.[21] O–H···N H-bonds were calibrated by using the water–imidazole system, whereas for N–H···N H-bonds we made use of the *N*-methylacetamide–imidazole pair. H-bonds involving S atoms as H-bond donors or acceptors were calibrated by using the systems cysteine–water

Table II. Penalty Coefficients "n" for the H-Bond and Metal-Center Potential Functions in "YETI" V4.5 (cf. Figure 3)

H-bond/Met Acc type	number/type of lone pairs[a]	"n"	ref
carbonyl O	2 sp^2	3	20, 21
carboxyl O	2 sp^2	4	66
hydroxyl O	3 sp^2/sp^3	3	21, 22
water O[b]	3 sp^2/sp^3	3	22, this work
ether O[c]	2 pseudo	4	20
arom. N five-membered ring	1 sp^2	4	22, this work
arom. N six-membered ring	1 sp^2	4	22
sulfhydryl S	2 intermediate	3	67–70
thioether S	2 intermediate	3	67–70

[a] The term "lone pair" refers to the experimental (i.e., crystallographically) observed maximum of the distribution of donor-H atoms around H-bond acceptors (cf. refs 20–22, 68 as well as metal ions around ligand atoms (cf. Figure 2). [b] For special applications, the "YETI" force field recognizes the residues WAT3 (sp^3-type) and WAT2 (sp^2-type), whereas the default type WAT as well as all R–OH acceptors include both sp^2 and sp^3 directions. [c] Two pseudo lone pairs to reproduce the broad experimental distribution (cf. ref 20).

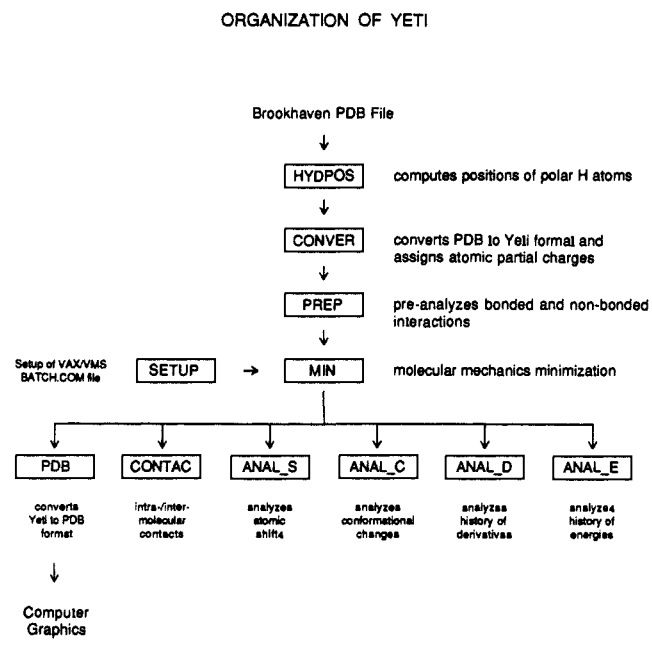

Figure 4. Modules of the molecular mechanics program "YETI" V4.5.

(32) Snyder, J. P., G. D. Searle & Co., Research & Development, Skokie, IL; 1989. Personal communication.
(33) Kroon, J.; Kanters, J. A.; Van Duijneveldt-Van de Rijdt, J. G. C. M.; Van Duijneveldt, F. B.; Vliegenthart, J. A. *J. Mol. Struct.* **1975**, *24*, 109–129.
(34) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Molec. Biol.* **1984**, *44*, 97–179.
(35) Dyke, T.; Muenter, J.; *J. Chem. Phys.* **1974**, *60*, 2929–2930.
(36) Curtiss, L.; Frurip, D.; Blander, M. *J. Chem. Phys.* **1979**, *71*, 2703–2711.
(37) Saenger, W. *Principles of Nucleic Acid Structure*, Springer: New York, 1984; p 382.
(38) Curtiss, L.; Frurip, D.; Blander, M. *J. Am. Chem. Soc.* **1978**, *100*, 79–86.
(39) Frisch, M. J.; Binkley, J. S.; Schlegel, H. B.; Raghavachari, K.; Melius, C. F.; Martin, R. L.; Stewart, J. J. P.; Bobrowicz, F. W.; Rohlfing, C. M.; Kahn, L. R.; Defrees, D. J.; Seeger, R.; Whiteside, R. A.; Fox, D. J.; Fleuder, E. M.; Pople, J. A. Carnegie-Mellon Quantum Chemistry Publishing Unit: Pittsburgh, PA, 1984.

(S-H···O), cysteine–imidazole (S-H···N), water–cysteine (O-H···S), *N*-methylacetamide–cysteine (N-H···S), and cysteine–methionine (S-H···S).

4. The Molecular Mechanics Software "YETI".[40] The molecular mechanics program "YETI" was developed over the past 6 years at the ETH Zürich and the University of Kansas.[7,10,12] It was designed for modeling proteins and small-molecule protein complexes with particular emphasis on metalloproteins. The program consists of 11 modules (Figure 4) and is written in VAX FORTRAN. "YETI" reads/writes standard Brookhaven Protein Data Bank[41] format and can therefore be interfaced with most computer graphics software, e.g., the program FRODO,[42] which was used for our studies.

Special features of the program include the directional potential functions for hydrogen bonds, salt linkages and metal–ligand interactions, a dynamic charge-transfer function for modeling metal centers, the optimization in a mixed internal/cartesian coordinate space, a conjugate-gradient minimizer with step-length optimizer (implicit second derivative determination), full compatibility with the molecular mechanics software AMBER (Version 3.0),[1] and the interactivity of all modules.

5. Molecular Mechanics Refinements. Energy calculations on the various complexes of human carbonic anhydrase I were performed by using the molecular mechanics programs "YETI" V4.5 and AMBER V3.0[1] on μ-VAX II computers of the KU Biographics Laboratory and the VAX 8650 computer of the Academic Computing Services of the University of Kansas. The refinements included all 2008 heavy atoms (C, N, O, S) and 450 polar H atoms (i.e., O–H, N–H, S–H) of the protein, the metal-ion cofactor, the two zinc-bound water molecules, and the substrate or inhibitor molecule as well as 501 water molecules representing an explicit solvent structure. For the protein portion, aliphatic and aromatic H atoms were included within united atoms.

In "YETI", optimizations are performed in a mixed internal/cartesian coordinate space by using a conjugate-gradient minimizer with step-length optimizer. Degrees of freedom include the conformation of all protein side chains; position, orientation, and conformation of the substrate or inhibitor molecule; position of any metals and anions as well as position and orientation of all water molecules. For our study, this yielded 3564 degrees of freedom for both the native and the active enzyme, 3565 for the complex with bicarbonate (502 water molecules), and 3533 for the complex with acetazolamide (496 water molecules), respectively. In the "YETI" force field, the total energy of the system is represented in terms of torsional, electrostatic, van der Waals, H-bond, and metal–ligand energies. Bond lengths and bond angles are not varied during the refinement (those have been optimized using AMBER; see below); therefore, no bond-stretching and angle-bending terms are present in the "YETI" force field energy expression (cf. Figure 3). Force field parameters are given in Table III.

By using smooth cutoff criteria (switch-on, switch-off; cf. ref 2) of 9.5/10.0 Å for electrostatic interactions, 6.5/7.0 Å for van der Waals interactions, and 5.5/6.0 Å for H-bonds, the initial list of nonbonded interactions numbered about 615 000.[43] Convergence criteria for the energy refinement were set at 0.025 kcal/(mol·deg) for torsional, at 0.025 kcal/(mol·deg) for rotational, and at 0.250 kcal/(mol·Å) for translational RMS first derivatives, respectively.

Full relaxations (including also the protein backbone) were performed by using the molecular mechanics software AMBER,[1] version 3.0. AMBER uses a combined steepest-descent/conjugate-gradient minimizer in cartesian coordinate space. Apart from the treatment of H-bonds and metal–ligand interactions, the AMBER force field for nonbonded interactions is identical with the "YETI" force field. In addition, AMBER allows bond distances and bond angles to vary and therefore includes terms for bond stretching and angle bending.[44] Convergence criteria for all AMBER refinements were set at 0.050 kcal/(mol·Å) for first de-

Table III. Force Field Parameters for Nonbonded Interactions Used in "YETI" V4.5[d]

1. Van der Waals Parameters

atom type(s)		VdW radius, Å	well-depth, kcal/mol[a]
H:	polar (i.e., H–O, H–N, H–S)	1.00	−0.020
	apolar (i.e., H–C)	1.375	−0.038
C:	>CH– united[2], aliphatic	1.85	−0.090
	$-CH_2-$ united[2], aliphatic	1.925	−0.120
	$-CH_3$ united[2], aliphatic	2.00	−0.150
	=CH— united[2] aromatic	1.85	−0.120
	sp^2 >C= aromatic/aliphatic	1.85	−0.120
	sp^3 >C< aliphatic	1.80	−0.060
N:	formal positive charge	1.85	−0.080
	formal neutral	1.75	−0.160
O:	carbonyl O	1.60	−0.200
	carboxyl, phosphate, nitrate O	1.60	−0.200
	hydroxyl, sugar, ether, ester O	1.65	−0.150
	water O	1.768	−0.152
P		2.10	−0.200
S:	formal neutral	2.00	−0.200
	anionic, i.e., $R-S^-$	2.00	−0.250
Co(II)		0.69	−0.014
Zn(II)		0.69	−0.014

2. H-bond Parameters

H-bond	equil. distance, Å	r_o, Å	E_o, kcal/mol
O–H···O	1.79	1.746	−4.946
O–H···N	1.89	1.878	−4.655
O–H···S	2.54	2.535	−1.746
N–H···O	1.87	1.877	−4.073
N–H···N	1.99	2.003	−3.491
N–H···S	2.64	2.667	−1.455
S–H···O	2.09	2.099	−2.328
S–H···N	2.19	2.088	−2.037
S–H···S	2.84	3.009	−1.164

3. Metal Center Parameters

metal	ligand	CN	equil. distance, Å	r_o, Å	E_o, kcal/mol
Zn(II)	O	4	1.961	1.969	−31.329
Zn(II)	N	4	2.041	2.043	−35.143
Zn(II)	S	4	2.317	2.327	−24.357
Zn(II)	O	5	2.068	2.075	−22.262
Zn(II)	N	5	2.126	2.126	−28.597
Zn(II)	S	5	2.412	2.409	−17.792
Co(II)	O	4	1.918	1.921	−37.368
Co(II)	N	4	1.964	1.963	−40.838
Co(II)	S	4	2.282	2.289	−29.096

[a] $\epsilon_{ij} = (\epsilon_i \cdot \epsilon_j)^{1/2}$. [b] United atoms include one C atom and one to three H atoms; compared with "explicit" atoms, they feature a larger van der Waals radius and an increased well-depth. [c] Total well-depth except electrostatic and LFS energy. [d] Torsional parameters are identical with those used in AMBER V3.0, cf. ref 44.

(40) Nonprofit organizations may obtain a free copy of the VAX/VMS object codes. Details for the distribution of object codes for other operating systems as well as for the source codes should be requested from the author of the program (A.V.).

(41) Bernstein, F.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. J. *J. Mol. Biol.* **1977**, *112*, 535.

(42) Pflugrath, J. W.; Saper, M. A.; Quiocho, F. A. In *Methods and Applications in Crystallographic Computing*; Hall, S., Ashiaka, T., Eds.; Clarendon Press: Oxford, 1984; p 407.

(43) Compared with earlier refinements,[7] the number of interactions in the nonbonded list increased by about 6%. This increase is due to a slight contraction (approximately 1.8% in volume) of the protein during the AMBER refinements. The cause for this shrinkage is presently unclear since the refinements included the whole protein and a reasonable solvent shell of 503 water molecules. Possible reasons might include the use of a distance-dependent dielectric parameter (overweighting of the van der Waals interactions) or the use of "united atoms" for H atoms attached to C atoms. We therefore plan to perform control refinements including all H atoms and using different dielectric parameters, i.e., $D(r) = 1, 2, 4; r$. Corresponding calculations without relaxing the protein backbone conformation and thus disabling the shrinking were carried out simultaneously.

(44) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Chandra Singh, U.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.

Table IV. Three Potential Proton-Relay Networks in Native Human Carbonic Anhydrase I

donor acceptor	Don···Acc, Å	H···Acc, Å	Don–H···Acc, deg	H···Acc–LP, deg
		Relay I[a]		
Wat 262 O–H···OH Thr 199	2.67	1.73	164.7	11.7
Thr 199 O–H···O2 Glu 106	2.79	1.86	159.8	22.9
		Relay II[a,b]		
		(a) His 64 Branch		
Wat 624 O–H···OH Wat 452	2.83	1.88	170.4	11.4
Wat 452 O–H···OH Wat 387	2.80	1.84	174.2	6.3
Wat 387 O–H···OH Wat 339	2.77	1.81	176.2	4.4
Wat 339 O–H···OH Wat 400	2.78	1.83	167.5	4.6
Wat 400 O–H···OH Wat 422	2.82	1.86	171.4	16.5
Wat 422 O–H···OH Wat 483	2.82	1.86	177.9	12.9
Wat 483 O–H···OH Wat 435	2.93	1.99	163.1	18.7
Wat 435 O–H···NB His 64	2.88	1.94	166.5	3.7
		(b) His 67 Branch		
Wat 483 O–H···OH Wat 610	2.86	1.93	161.5	27.1
Wat 610 O–H···OH Wat 313	2.72	1.80	159.3	7.6
Wat 313 O–H···NB His 67	2.85	1.89	177.0	24.9
		Relay III[c]		
Wat 262 O–H···OH Wat 381	2.73	1.81	160.3	19.0
Wat 381 O–H···OH Wat 452	2.78	1.86	158.7	9.0

[a] Deprotonation of the proximal zinc-bound water molecule, i.e., Thr 199–Glu 106 pathway. [b] Deprotonation of the distal zinc-bound water molecule: (a) protonation of His 64 and (b) protonation of His 67. [c] Deprotonation of the proximal zinc-bound water molecule (via Wat 381; joining proton-relay II at Wat 452).

rivative RMS values; the cutoff value for nonbonded interactions was set at 10.0 Å. In both the "YETI" and AMBER refinements, a weight of 0.5 was assigned to nonbonded 1–4 interactions (electrostatic and van der Waals terms).

Atomic partial charges for amino acid residues of the protein were taken from ref 44; those for the substrate and the inhibitor molecules were calculated by MNDO (program AMPAC[32]) or ab initio methods (program GAUSSIAN 86[39]). Electrostatic energies are evaluated by using a distance-dependent dielectric parameter of $D(r) = 2 \cdot r$. Earlier studies justify this value for modeling metalloenzymes.[10]

Results and Discussion

Carbonic anhydrase, a zinc-containing enzyme, is an extremely efficient catalyst of the reversible hydration of carbon dioxide.[45] The crystal structure of the native enzyme has been determined to a resolution of 2.0 Å and refined to a R-factor of 0.19 by Kannan and co-workers.[46] The catalytic zinc is located at the bottom of a conical cavity, coordinated to the N atoms of His 94, His 96, His 119 and to the O atom of a water molecule.

Various mechanistic models for the catalytic reaction have been put forward based on structural, spectroscopic, and kinetic data.[47-53] It is almost universally assumed that a zinc-bound hydroxide ion is the nucleophile in the catalytic reaction.[53] However, it is still uncertain which pathway generates the nucleophile OH^- from a water molecule, and what the details of metal coordination during catalysis might be.

Table V. Details of Metal Coordination in Native and Complexed Human Carbonic Anhydrase I

	native (Zn)[a]	active (Zn)[a] (a)H64/(b)H67	Bic (Zn)	AAA (Zn)	native (Co)[a]	Bic (Co)
N: His 94	2.044	2.045/2.045	2.043	2.052	1.995	1.986
N: His 96	2.044	2.052/2.051	2.080	2.053	1.990	2.021
N: His 199	2.039	2.037/2.035	2.048	2.045	1.992	2.007
O: Wat 262	1.991	1.991/1.990	2.006		2.031	2.042
O: Wat 624	2.918				3.024	
O: OH^-		2.938/2.933				
O: OH Bic			2.066			2.134
C: Bic[b]			3.075			3.156
N: AAA[b]				2.045		
O: AAA[b]				2.757		
q(M)[c]:	0.833	0.838/0.842	0.747	0.751	0.667	0.564

[a] Native: zinc-bound water molecule. Active: zinc-bound OH^-: (a) His 64 and (b) His 67 pathways, respectively (cf. text). [b] Bic: bicarbonate; AAA: actetazolamide (2-acetamido-1,3,4-thiadiazole-5-sulfonamide). [c] Effective partial charge on metal ion after ligand–metal charge transfer.

In an ab initio study, Cook and Allen suggested a five-coordinated metal center where a second zinc-bound water molecule would act as the nucleophile.[54] Most recently, Merz, Hoffmann, and Dewar published a very detailed AM1 study addressing the generation of the OH^- ion, the formation of the bicarbonate ion from CO_2, and the relay of the bicarbonate ion from the active-site zinc to the surface of the protein.[55] They propose a four-coordinate zinc with only one metal-bound water molecule. Since both studies are based on very small model systems, molecular mechanics calculations including the whole protein, and parts of the solvent shell should provide additional insight into the details of the catalytic mechanism. Of course, they cannot be used to simulate properties of transition states and high-energy intermediates.

Of the two structures of human carbonic anhydrase (HCA I: low-activity form; HCA II: high-activity form) deposited in the Brookhaven Protein Data Bank (BPDB,[41]), HCA I has been refined to a higher degree and therefore was chosen as the starting structure for our model studies. Differences in the primary sequence of the two isoenzymes include residues 200 (HCA I: His, HCA II: Thr), 67 (His,Asn), 69 (Asn,Glu); 91 (Phe,Ile), 121

(45) Notstrand, B.; Vaara, I.; Kannan, K. K. In *The Isozymes*; Markers, C. L., Ed.; Academic Press: New York, NY, 1975; pp 575–599.
(46) Kannan, K. K.; Ramanadham, M.; Jones, T. A. In Biology, and Chemistry of the Carbonic Anhydrases. Tashian, R. E., Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 49–60.
(47) Keilin, D.; Mann, T. *Nature (London)* **1940**, *146*, 164.
(48) Pocker, Y.; Sarkanen, S. *Adv. Enzymol.* **1978**, *47*, 149–274.
(49) Silverman, D. N.; Vincent, S. H. *CRC Crit. Rev. Biochem.* **1983**, *14*, 207–255.
(50) Lindskog, S.; Engberg, P.; Forsman, C.; Ibrahim, S. A.; Jonsson, B.-H.; Simonsson, I.; Tibell, L. In Biology and Chemistry of the Carbonic Anhydrases. Tashian, R. E., Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 61–75.
(51) Pocker, Y.; Deits, T. L. In Biology and Chemistry of the Carbonic Anhydrases. Tashian, R. E., Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 76–83.
(52) Maren, T. H. In Biology and Chemistry of the Carbonic Anhydrases. Tashian, R. E., Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 10–17.
(53) Coleman, E. In Biology and Chemistry of the Carbonic Anhydrases. Tashian, R. E.; Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 26–48.
(54) Cook, C. M.; Allen, L. C. In Biology and Chemistry of the Carbonic Anhydrases. Tashian, R. E., Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 84–88.
(55) Merz, K. M.; Hoffmann, R.; Dewar, M. J. S. *J. Am. Chem. Soc.* **1989**, *111*, 5636–5649.

Figure 5. Stereoscopic view of the active site cleft of native human carbonic anhydrase I. The zinc is represented by a sphere, zinc-bound ligands have been drawn enhanced.

(Ala,Val), and 131 (Leu,Phe). Residues 67 and 200 would seem to be most critical for the observed differences in activity as well as for the arrangement of the solvent within the active-site cleft.

Since the structure of HCA I as deposited in the BPDB does not include coordinates of the solvent,[56] we have used the program SOLVGEN, developed at our laboratory, to generate an explicit solvent structure. The algorithm in SOLVGEN is based on length, linearity, and directionality of H-bonds.[57] It searches the protein for unsaturated H-bond donors and acceptors and can generate both internal and surface water molecules. For native HCA I, SOLVGEN generated two zinc-bound, 16 internal and 485 surface water molecules.[7] Some of these water molecules are associated with potential H-bond relay networks; details are given in Table IV.

Molecular mechanics refinements of the solvated native enzyme yielded a practically undistorted tetrahedron at the zinc. In agreement with the ab initio study (but not in contradiction to the X-ray structure[46]), an ideal position for a second zinc-bound water molecule has been identified at a distal, fifth coordination site of the metal (d = 2.918 Å), opposite the imidazole N atom of His 96. (For the proximal/distal terminology of zinc ligands in proteins, see ref 6). Details of the zinc coordination are given in Figure 5 and Table V.

Several possible proton-relay networks for the interconversion of water and OH^- are presently under discussion: One is associated with the proximal water molecule and shuttles the proton via Thr 199 to Glu 106. This pathway would require an anomalously high pK_a of Glu 106 and is therefore not considered to be very likely.[55] Another possibility involves the relay of a proton via water molecules bound within the active-site cleft. Both the proximal and the distal water molecule are engaged in a complex H-bond network with that solvent. This relay could lead to an external base (e.g., buffer) or to a His residue located at the entrance of the active site. For HCA I, two His residues (His 64 and His 67) are located in solvent-accessible positions; for HCA II (the high-activity isozyme), only His 64 (but not Asn 67) could act as a suitable base. Both the ab initio and the AM1 study have been carried out for HCA II and favor the water-relayed proton transfer where His 64 acts as the terminal base. The role of His 64 in the catalytic mechanism of HCA II has also been studied with a site-specific mutant.[58]

In HCA I, no short proton-relay analogous to the one identified in HCA II (i.e., Zn(II)···Wat···Wat···His 64) seems possible. This

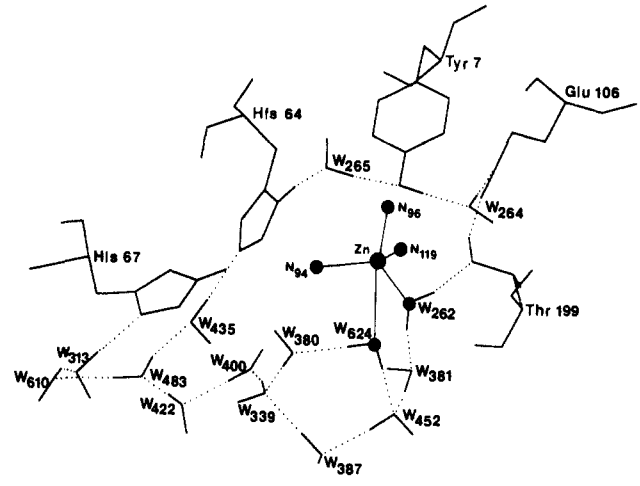

Figure 6. Schematic representation of possible proton-relay networks in native human carbonic anhydrase I is shown. The zinc and the zinc-bound ligands are represented by spheres. Hydrogen bonds are indicated by dotted lines.

is mainly due to the presence of residue His 200 (HCA II: Thr 200) which has been found engaged in a water-mediated H-bond with His 67.[46] However, our studies identified a longer connection from the catalytic zinc to His 64 (relay via seven water molecules) and to His 67 (relay via eight water molecules). Details are illustrated in Figure 6; geometries are listed in Table IV.

In agreement with the ab initio study[54] and the mechanism proposed by Coleman,[53] our molecular mechanics calculations favor the deprotonation of the distal water molecule via protein-bound water molecules to either His 64 or His 67. Based on H-bond geometries and energies, both His 64 and His 67 seem equally likely candidates for the terminal base in the proton relay (see Figure 6 and Table IV). The apparent disagreement with the AM1 calculation[55] might be due to the very limited size of model system used in that study; for example, residue Thr 199 and several experimentally determined water molecules were not included.

A ^{13}C NMR study on the Co(II)-substituted enzyme by Henkens, Merrill, and Williams gave experimental evidence that the natural substrate bicarbonate binds within the first coordination shell of the metal and that the average Co···C distance is 3.20 Å.[59] In an earlier molecular mechanics study[7] we showed that a monodentate binding mode would be in best agreement with the NMR result. Such a monodentate binding mode would also

(56) In ref 46, the authors specifically discuss protein-bound water molecules; Figure 3 shows six of them in a stereographic representation. These six waters were reproduced by SOLVGEN but will be referred to as "experimental" water molecules.

(57) Jacober, S. P. SOLVGEN: *An Approach to Protein Hydration*; M.S. Thesis, Department of Computer Sciences, University of Kansas, 1988.

(58) Tu, C.; Silverman, D. N.; Forsman, C.; Jonsson; B.-H.; Lindskog, S. *Biochemistry* **1989**, *28*, 7913–7918.

(59) Henkens, R. W.; Merrill, S. P.; Williams, T. J. In Biology and Chemistry of the Carbonic Anhydrases. Tashian, R. E., Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 143–145.

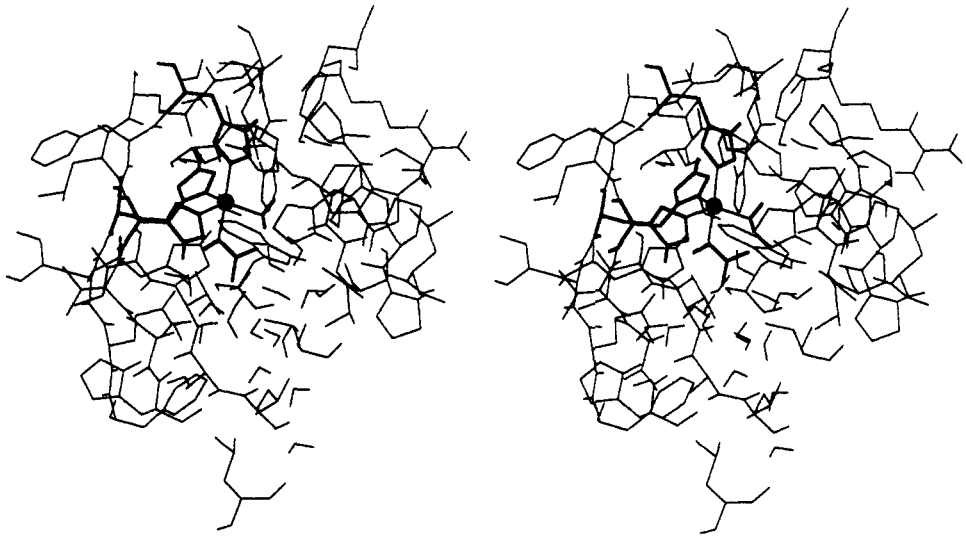

Figure 7. Stereoscopic view of the complex of HCA I and the natural substrate bicarbonate along with details of the active-site region are shown. The zinc is represented by a sphere, zinc-bound ligands have been drawn enhanced.

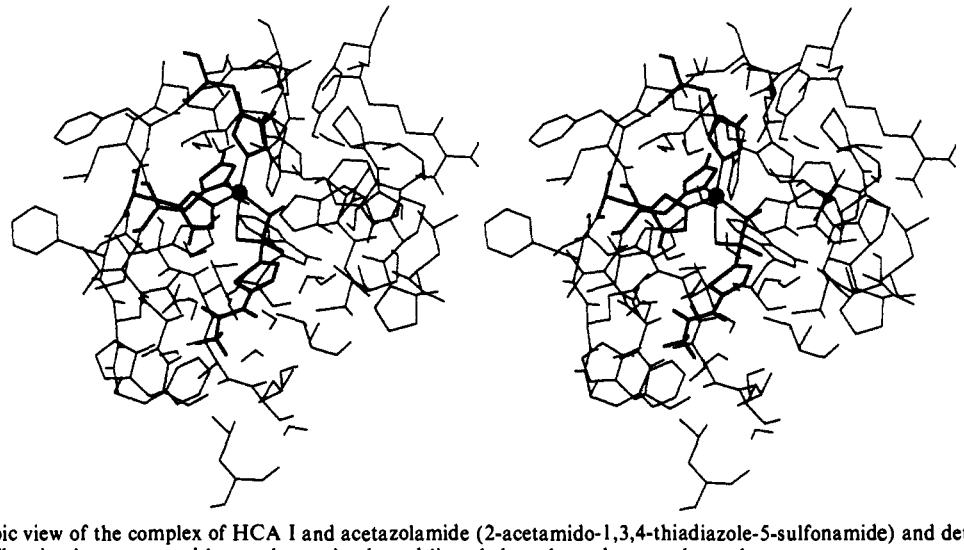

Figure 8. Stereoscopic view of the complex of HCA I and acetazolamide (2-acetamido-1,3,4-thiadiazole-5-sulfonamide) and details of the active-site region are shown. The zinc is represented by a sphere, zinc-bound ligands have been drawn enhanced.

be in agreement with the ab initio study[54] and the mechanism proposed by Coleman.[53] In the AM1 study, Merz, Hoffmann, and Dewar propose a bidentate binding mode.[55] There, the Zn···C distance in an early transition state is 3.10 Å (cf. ref 55, Figure 7); distances for the zinc–bicarbonate complex (cf. ref 55, structures **5** and **29**) are not reported.

Apart from the Zn–OH bond (d = 2.066 Å), the bicarbonate ion engages in H-bonds with the proximal zinc-bound water molecule as well as with two other water molecules bound to the active-site cleft. The Zn···C distance of 3.075 Å is clearly shorter than the one obtained in the NMR experiment, which, however, represents the Co(II)-substituted enzyme. As we will discuss below, corresponding molecular mechanics simulations with Co^{II}–HCA I lead to a Co^{II}–C distance of 3.156 Å. Details of the Zn(II) complex are shown in Figure 7; geometries are listed in Table V.

Calculations have also been made for the complex of HCA I with acetazolamide (2-acetamido-1,3,4-thiadiazole-5-sulfonamide), a potent inhibitor of the enzyme. In agreement with experimental data (a 3.0-Å resolution X-ray structure[60,61] and a ^{15}N NMR study[62]), the deprotonated sulfonamide N atom and one O atom of the sulfonamide moiety coordinate to the zinc and displace both metal-bound water molecules from the native enzyme, thus interrupting all possible H-bond relay networks. In contrast to the complex with the natural substrate bicarbonate, the O atom of the inhibitor molecule binds only very weakly to the metal (d = 2.757 Å) which therefore retains an almost undistorted tetrahedral symmetry. Details of the complex with Zn(II)–HCA I are illustrated in Figure 8; geometries are listed in Table V. The role of sulfonamide binding has recently been elucidated in a theoretical study by Liang and Lipscomb.[63]

Since most spectroscopic studies were performed by using the Co(II)-substituted enzyme,[62,64] we have carried out corresponding molecular mechanics calculations with Co(II) carbonic anhydrase. Differences between Zn(II) and Co(II) include different average metal–ligand bond lengths (see Table I), a different amount of charge transfer (Co–ligand bonds are slightly less electrostatic in character than Zn–ligand bonds), and the presence of ligand-field stabilization in the case of Co(II). The LFSE has been estimated using g- and f-factors to obtain a value for the ligand-field splitting 10 Dq (cf. Table 9.1, ref 23) for Co(II) in

(60) Kannan, K. K.; Vaara, I.; Notstrand, B.; Lovgren, S.; Borell, A.; Fridborg, K.; Petef, M. In *Proceedings on Drug Action at the Molecular Level*; Roberts, G. C. K., Ed.; McMillan: London, 1977; pp 73–91.

(61) Kannan, K. K. In *Biophysics and Physiology of Carbon Dioxide*; Gros, H., Bartels, H., Eds.; Springer: Berlin, 1979; pp 184–205.

(62) Mukherjee, J.; Rogers, J. I.; Khalifah, R. G.; Everett, G. W., Jr.; *J. Am. Chem. Soc.* **1987**, *109*, 7232–7233.

(63) Liang, J.-Y.; Lipscomb, W. N. *Biochemistry* **1989**, *28*, 9724–9733.

(64) Bertini, I.; Luchinat, C. In Biology and Chemistry of the Carbonic Anhydrases. Tashian, R. E., Hewett-Emmett, D., Eds.; *Ann. New York Acad. Sci.* **1984**, *429*, 89–98.

high-spin configuration and the various ligands. The *f*-factor for the imidazole N ligand was approximated by the available data for NH_3.

Both native Co(II)-substituted carbonic anhydrase as well as the complex with the natural substrate bicarbonate were studied. Results from the complex with bicarbonate can be compared to the ^{13}C NMR data.[59] The Zn(II)- and the Co(II)-enzyme show no principal structural differences. The distance from the bicarbonate C atom to the Co(II), as obtained by our calculations for the bicarbonate complex, of 3.157 Å, is in good agreement with the ^{13}C NMR study.[65] Geometries for the Co(II) complexes are given in Table V.

Conclusions

Structural data retrieved from the Cambridge Structural Database demonstrates that directionality is an important factor for the strength of H-bonds and metal–ligand interactions. Since these interactions play a key role for both structure and function of proteins, we have developed and applied a new force field which accounts for such directional preferences as well as for ligand–metal charge transfer and ligand-field stabilization. With the incorporation of this force field into the molecular mechanics program "YETI", we have then tried to demonstrate that results from such calculations on macromolecular systems are complementary to information obtained by ab initio and semiempirical studies on small model systems.

Acknowledgment. We express our gratitude to Drs. Jim P. Snyder and Dayle Spangler (Searle, Research & Development, Skokie, IL) for challenging discussions. We gratefully acknowledge financial support from the University of Kansas and the Swiss "Foundation for the Replacement of Animals in Medical Research" (FFVFF, Zürich). D.W.H. is a NIH predoctoral trainee under PHS Grant GM-07775.

Note Added in Proof. The most recent version of the molecular mechanics program "YETI", version V4.8, allows the treatment of up to eight different protein-bound metal ions. In addition, a function calculates an adjusted amount of charge transfer for ligand atoms bridging two or more metal centers. Version V4.8 has been tested and recalibrated by using the X-ray crystal structures of the enzymes thermolysin (1.6-Å resolution),[71] trypsin (1.7-Å resolution),[72] carboxipeptidase (1.54-Å resolution),[73] and tonin (1.8-Å resolution),[74] all of which were retrieved from the Brookhaven Protein Data Bank.[41]

Registry No. Zinc(II), 7440-66-6; bicarbonate, 71-52-3; acetazolamide, 59-66-5; cobalt(II), 7440-48-4.

(65) A better agreement of the Co–C distance could be obtained by using a weighting factor for the radial/directional term (cf. above) of 0.667 (standard value: 0.750). Here, the resulting Co···C distance of 3.195 Å is identical with the one obtained by the ^{13}C NMR study.[59] However, since we do not presently have a solid theoretical basis for determining an optimal value, this agreement should not be overinterpreted.

(66) Vedani, A.; Dunitz, J. D. Unpublished results.
(67) Guru Row, T. N.; Parthasarathy, R. *J. Am. Chem. Soc.* **1981**, *103*, 477–479.
(68) Ramasubbu, N.; Parthasarathy, R.; Murray-Rust, P. *J. Am. Chem. Soc.* **1985**, *108*, 4308–4314.
(69) Jones, D. D.; Bernal, I.; Frey, M. N.; Koetzle, T. F. *Acta Crystallogr.* **1974**, *B30*, 1220–1227.
(70) Kerr, A. K.; Ashmore, J. P.; Koetzle, T. F. *Acta Crystallogr.* **1975**, *B31*, 2022–2026.
(71) Holmes, M. A.; Matthews, B. W. *J. Mol. Biol.* **1982**, *160*, 623–639.
(72) Walter, J.; Steigemann, W.; Singh, T. P.; Bartunik, H.; Bode, W.; Huber, R. *Acta Crystallogr., Sect. B* **1982**, *38*, 1462–1472.
(73) Rees, D. C.; Lewis, M.; Lipscomb, W. N. *J. Mol. Biol.* **1983**, *168*, 367–387.
(74) Fujinaga, M.; James, M. N. G. *J. Mol. Biol.* **1987**, *195*, 373–396.