PROTEINS: Structure, Function, and Genetics 23:12–31 (1995)

Zinc Binding in Proteins and Solution: A Simple but Accurate Nonbonded Representation

Roland H. Stote and Martin Karplus
Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138 and Laboratoire de Chimie Biophysique, Institut Le Bel, 4 rue Blaise Pascal, Université Louis Pasteur, 67000 Strasbourg, France

ABSTRACT **Force field parameters that use a combination of Lennard–Jones and electrostatic interactions are developed for divalent zinc and tested in solution and protein simulations. It is shown that the parameter set gives free energies of solution in good agreement with experiment. Molecular dynamics simulations of carboxypeptidase A and carbonic anhydrase are performed with these zinc parameters and the CHARMM 22 β all-atom parameter set. The structural results are as accurate as those obtained in published simulations that use specifically bonded models for the zinc ion and the AMBER force field. The inclusion of longer-range electrostatic interactions by use of the Extended Electrostatics model is found to improve the equilibrium conformation of the active site. It is concluded that the present parameter set, which permits different coordination geometries and ligand exchange for the zinc ion, can be employed effectively for both solution and protein simulations of zinc-containing systems.** © 1995 Wiley-Liss, Inc.

Key words: molecular dynamics, electrostatics, carboxypeptidase A, carbonic anhydrase, zinc ion

INTRODUCTION

Dipositive Zn, one of the metal ions that is widely distributed in living systems, is known to be an integral component of a large number of enzymes and other proteins involved in many aspects of metabolism. A variety of roles for the zinc ion have been reviewed by Vallee and Auld.[1] Zinc is required for catalytic function in such biologically important enzymes as carbonic anhydrase,[2–5] carboxypeptidase A,[6–11] liver alcohol dehydrogenase,[12] and thermolysin.[13,14] Structural roles have been suggested for zinc in alcohol dehydrogenase and aspartate transcarbamoyltransferase.[15] The contribution to protein stability is illustrated by the nucleic acid binding protein, g32P, which is involved in DNA replication, repair, and recombination. Removal of the zinc ion results in a decreased denaturation temperature and a corresponding decrease in the enthalpy change on denaturation.[16] Zinc is an essential component of zinc finger proteins and other DNA-binding proteins that exhibit distinctive structural motifs. It is present in the complex system of gene transcription[17–21] and in metalloproteases[22] that are involved in extracellular matrix turnover. Also of interest are the enzymes in which two zinc atoms appear to be involved in catalysis.[23,24] Structures of carboxypeptidase A with a number of inhibitors and transition state analogs have demonstrated multiple binding modes around the ion at various stages of the enzyme reaction, suggesting that this flexibility in ligation may be important for enzyme function.[25]

The interactions of the zinc ion with small ligands have been studied with ab initio techniques.[3,26–28] Such calculations are difficult to extend to macroscopic systems, in which the zinc interacts with bulk solvent or a protein. In cases where the details of the electron distribution are not of primary interest, molecular mechanics and dynamics calculations offer an attractive alternative. This is particularly true for zinc in its most common (2+) oxidation state because it is an essentially spherical, closed shell (d^{10}) ion. This electron configuration allows a variety of bonding geometries of which the six coordinate octahedral and four coordinate tetrahedral are most common.[29] A survey of the Cambridge Structural Database[30] of small molecule crystal structures yielded 89 structures containing four coordinate zinc, of which 85 were classified as having a tetrahedral geometry; 28 five coordinate structures gave a distribution of trigonal bipyramids and square pyramids and the 43 structures with six ligands had an octahedral geometry.[31]

Molecular dynamics simulations of zinc containing proteins require a description of the binding site within the framework of an empirical force field. Two broad categories of models have been tried; in the present paper they are referred to as *bonded* models and *nonbonded* models. In *bonded* models explicit bond and angle terms are introduced into the potential energy function to account for interactions between the ion and its protein ligand at-

Received September 22, 1994; revision accepted March 7, 1995.

Address reprint requests to Martin Karplus, Laboratoire de Chimie Biophysique, Institut Le Bel, 4 rue Blaise Pascal, Université Louis Pasteur, 67000 Strasbourg, France.

oms.[31–35] The special attribute of such bonded models for zinc–protein interactions is that the crystal geometry is preserved by the choice of ligands and the bond length and bond angle parameters. However, although appropriate in many cases, this model introduces certain limitations. First, by imposing the crystallographic geometry through the use of constraints, the overall conformational flexibility of the metal binding site is restricted. This could, in turn, affect calculated thermodynamic quantities such as free energies and entropies, which depend on the fluctuations of the system. There is also the difficulty of using a bonded model to treat the multiplicity of zinc ion binding modes, such as those found in carboxypeptidase A.[25] The assumption that the crystal structure of the metal binding site remains unchanged in going from crystal to solution may not always be correct. Recent XAFS studies indicate that metal binding sites can have different conformations in solution from that in crystal.[36–39] Finally, such *bonded* models cannot be valid for zinc in solution, where the ligands, such as water, can exchange.

The alternative *nonbonded* models for zinc include only van der Waals and electrostatic terms. They avoid the conformational restrictions of the *bonded* models and should be appropriate for Zn^{2+}, as well as other ions such as Ca^{2+}, Mg^{2+}, and Cd^{2+}, which have closed shells and thus no crystal field splitting. Attempts to use such models have encountered problems stemming primarily from the difficulty in treating the strong electrostatic interactions of the divalent zinc ion.[40] Applications in molecular mechanics calculations of proteins resulted in rearrangements around the metal center which were in disagreement with experimental data.[34,41,42]

In this paper we show that a carefully parameterized nonbonded model can be used for zinc binding sites in proteins and for solution simulations. An important issue for any such highly charged ionic species, which has been neglected in previous models for the zinc ion, is the effect of the truncation that is commonly applied to the long-range electrostatic interactions. A number of experimental[43–47] and theoretical[48–50] studies have demonstrated the importance of long-range protein–ion interactions where significant contributions come from distances as large as 15 Å. These findings suggest that it is essential to investigate the effects of long-range electrostatic interactions on the binding of the zinc ion in proteins as part of any analysis of potential function models. In all previous studies, standard cutoffs in the range of 8 to 10 Å have been used. At these distances, the unshielded Coulomb interaction between the Zn^{2+} ion and a point charge is 83 to 66 kcal/mol, respectively, and the force on the Zn^{2+} due to a point charge is between 10.4 and 6.0 kcal/mol-Å. In some of the earliest studies, electrostatic interactions between the metal ion and its environment were neglected[51] or, as in the case of a simulation study of carboxypeptidase A carried out in vacuum, the metal ion was eliminated entirely and only the apoprotein was simulated because the system with the ion was unstable.[40]

The debate over bonded versus nonbonded models for the zinc has a history analogous to that of hydrogen bonding interactions and their treatment in empirical energy functions. In some energy functions, particularly those used in earlier studies, hydrogen bond interactions were accounted for by the introduction of special energy terms.[31,51–53] Detailed comparisons with ab initio calculations have shown that no special terms are needed and that the energetics of hydrogen bonds can be accounted for by a combination of Lennard–Jones and electrostatic terms.[54–57]

We investigate two aspects of the development of models for the zinc ion in protein systems. The first is a detailed comparison of bonded and nonbonded models for the zinc ion. For this purpose, it was necessary first to develop a reliable nonbonded model for the zinc ion. This was done by fitting the empirical potential function to the ab initio potential surface for the interaction between zinc ion and water. This initial parameter set was then refined by solution simulations and comparison to experimental radial distribution functions and free energies of solvation. The model was tested in simulations of two zinc metalloenzymes, carboxypeptidase A and carbonic anhydrase, and the results from these simulations were compared with results obtained from the use of bonded models. The second aspect is the effect of long-range electrostatic interactions on the behavior of the zinc in these systems. Long-range electrostatic interactions are included in the protein simulations via the "Extended Electrostatics" model based on a multipole expansion of distant charges.[53,58] The results obtained have implications for potential functions for any charged species in a protein environment.

In the next section we outline the methods used for this study. We then describe the development of a nonbonded model for the zinc ion and make comparisons with the more complex nonbonded model for zinc ion proposed by Åqvist and Warshel.[59] This is followed by the presentation and discussion of the protein simulation results and concluding comments.

METHODS

This section is divided into two parts. In the first, we outline the methodology used in developing the model potential for the zinc ion with emphasis on the treatment of the electrostatics and in the second, we describe the methods used in the dynamics simulations of the zinc binding sites in carboxypeptidase A and carbonic anhydrase.

The Nonbonded Model and Its Development

The form of the force field used for zinc ion interactions with its surroundings that may consist of protein, water, or both is given by

$$E_{\text{ion}} = \sum_j 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_j \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}. \quad (1)$$

The subscript i corresponds to the ion and j goes over all other atoms. The first term is the standard 6–12 form of the Lennard–Jones potential and the second term is the Coulomb potential for the electrostatic interactions. Although this is a standard expression, the description of the potential function is not complete without specifying the type of truncation; we describe different truncation schemes in Treatment of Electrostatic Interactions. The general approach taken in the development of the model was to utilize ab initio results to develop an initial model for the interaction of the zinc ion with a water molecule. The model was then refined by comparing the Zn^{2+}–O radial distribution function (rdf) and the free energy of solvation of the Zn^{2+} ion calculated from condensed phase simulations with experimental results (details are given below). This is an important element of parameter development. Similar approaches have been employed by other groups for ion parameters[60] and for protein and nucleic acids force fields.[57]

A number of ab initio studies of the interaction between the zinc ion and biologically significant ligands exist.[26,27,61–63] Given the important role played by water in the biological function of zinc-containing enzymes, potential energy surfaces have been calculated for Zn^{2+}/H_2O interactions and used in the development of empirical force field parameters for the zinc ion. The ion–water interactions have been modeled by standard 6–12 Lennard–Jones terms and a Coulombic term[62] or by a more complicated potential energy function which includes corrections for three body interactions.[63] In the present study, the ab initio potential energy surface for zinc ion–water interactions calculated by Clementi and co-workers[62] was used to develop an initial trial parameter set within the framework of the CHARMM potential energy function [see Eq. (1)]. A formal charge of 2+ was used for the zinc ion, water was represented by a modified TIP3P water model,[55] and standard combining rules were used. With these constraints, *only* the Lennard–Jones parameters (σ, ϵ) of the zinc ion can be adjusted to fit the salient features of the interaction potential, such as the well depth and the position of the energy minimum.

The refinement of the initial parameter set obtained from the analysis of the ab initio results was done by calculating the Zn^{2+}–O radial distribution function from Monte Carlo simulations of the solvated zinc ion and adjusting the Lennard–Jones parameters (σ, ϵ) of the zinc ion such that the positions of the first and second peaks of the rdf, as well as the coordination numbers, were in agreement with the experimentally determined values. The Monte Carlo simulation method has been described elsewhere.[64] The simulations were carried out with the BOSS program[65] modified to employ the standard combination rules [i.e., $\sigma_{ij} = 0.5(\sigma_{ii}+\sigma_{jj})$ and $\epsilon_{ij}=(\epsilon_{ii}\epsilon_{jj})^{1/2}$], used in the CHARMM potential energy function. Simulations were made with a cubic periodic boundary condition system; initial edge lengths of 18.34 and 24.75 Å were used for 216 and 512 water molecules, respectively. Ion–water interactions were truncated by a switching function at 8.5 Å for simulations employing the smaller box and 9.5 Å for simulations using the larger box. Water–water interactions were switched off at 8.5 Å in all the simulations. Use of a relatively short cutoff in simulations of infinitely dilute ionic solutions is justified, given the high dielectric constant of water (ε = 80) relative to that of a protein (ε = 2 to 4). The NPT ensemble at 25°C and 1 atm pressure was used with Metropolis/preferential sampling. Simulations for determining the radial distribution functions were carried out for 1.5×10^6 Monte Carlo moves.

Free Energy Simulations

The free energy of solvation was also calculated with the BOSS program[65] by transforming the Zn^{2+} into an extended-atom methane sphere for which the solvation free energy has been calculated.[65,66] The total free energy, ΔG, is given as a sum over a series of calculations for a set of λ values between zero and one,[67,68] i.e.,

$$\Delta G = -\sum k_B T \ln\langle\exp(U_{\lambda_{i\pm1}} - U_{\lambda_i})/RT\rangle_{\lambda_i} \quad (2)$$

where $U_{\lambda_i} = \lambda_i U_1 - (\lambda_i - 1)U_0$; U_0 is the potential of the initial state and U_1 is the potential of the final state.

The simulations were run at λ_i and Eq. (2) was used to compute the free energy change for $\lambda_i \rightarrow \lambda_{i+1}$ and $\lambda_i \rightarrow \lambda_{i-1}$ using double wide sampling. For all calculations, $\Delta\lambda = |\lambda_{i\pm1} - \lambda_i| = 0.1$. Equilibration at each λ point was carried out for 200K MC moves and between 1.5M and 2 M moves were done for averaging.

Use of intermediate values of λ defines a particular path for the free energy simulation. As the free energy is a thermodynamic state function, the calculated value is, in principle, independent of the path. However, in practice, one generally observes some degree of path dependence, so calculations along two different paths between the zinc ion and the methane sphere were done. The first path (Path 1) involved concurrent changes in the charge and the Lennard–Jones parameters of the zinc ion going directly from the zinc ion to the methane sphere. In

the second path (Path 2), this transformation was done in two stages; first, the charged ion was mutated into an uncharged sphere with the same Lennard–Jones parameters, and the neutral zinc sphere was then changed to the methane sphere.[69] Calculations along the reverse direction were also done to determine the magnitude of the hysteresis. The Lennard–Jones parameters, σ_{Zn} and ε_{Zn} as well as the charge q_{Zn} of the zinc were scaled linearly with λ; e.g., $\sigma_i = \lambda_i \sigma_{Zn}$, $\varepsilon_i = \lambda_i \varepsilon_{Zn}$, and $q_i = \lambda_i q_{Zn}$ for λ_i values between zero and one. The individual free energies were then averaged to obtain the final result. The total free energy of solvation, $\Delta G_{Solvation}$, was calculated by summing the terms

$$\Delta G_{Solvation} = \Delta G_{(ion\text{–}methane\ sphere)} + \Delta G_{(methane\ sphere)} + \Delta G_{Born} \quad (3a)$$

or

$$\Delta G_{Solvation} = \Delta G_{(ion\text{–}uncharged\ ion)} + \Delta G_{(uncharged\ ion\text{–}methane\ sphere)} + \Delta G_{(methane\ sphere)} + \Delta G_{Born} \quad (3b)$$

for the two paths. The first term in Eq. (3a) corresponds to the free energy difference between the charged ion and the standard united atom model for methane in solution while the first term in Eq. (3b) corresponds to that between the charged ion and an uncharged van der Waals sphere with the zinc Lennard–Jones parameters. The second term in Eq. (3b) corresponds to the free energy of forming the uncharged van der Waals sphere relative to the standard united atom model for methane in solution, for which the free energy of solvation, $\Delta G_{(methane\ sphere)}$, has already been calculated.[66,70] The last term in both equations is a correction term which accounts, in part, for the error introduced by use of a finite truncation for the electrostatic interaction. Effectively, one views the system as a spherical hydrated "ion" with a radius equivalent to the cutoff distance in water. The hydration of this macroion is estimated by the Born formula,

$$\Delta G_{Born} = -(q^2/2R)(1 - 1/D) \quad (4)$$

where q is the net charge of the ion, R is the cutoff distance, and D is the dielectric constant of the solvent, which is taken to be the experimental value for water ($D = 78$). Water models similar to the TIP3P water model have a smaller dielectric constant ($\varepsilon \approx 50$ for TIP4P and $\varepsilon \approx 68$ for SPC at 293 and 300K, respectively[71]), but use of such a value would yield essentially the same result.

Protein Simulations

Version 22 of the CHARMM program,[53] with a preliminary version of the CHARMM22 all-atom potential function,[57] was used for all calculations. Crystal structures of carboxypeptidase A (5CPA)[8] and carbonic anhydrase II (1CA2)[5] were obtained from the Brookhaven Protein Data Bank[72,73] and used as the starting configuration for all calculations. Initial hydrogen atom positions were determined using the HBUILD facility[74] in CHARMM. To prepare the system, an energy minimization was done to remove short van der Waals contacts and strained bond angles. Harmonic constraints defined as $k(r-r_o)^2$ with force constants of 50 and 100 kcal/mol-Å^2 were placed on the protein backbone atoms and the zinc ion, respectively. Weaker constraints with a force constant of 25 kcal/mol-Å^2 were placed on the protein sidechain atoms. The nonbond list was constructed using a 13 Å cutoff with pair interactions terminated at 12 Å. Van der Waals interactions were truncated over a range of 10–12 Å with a switching function, while electrostatic interactions were truncated using a shift function coupled with an r-dependent dielectric constant and an ϵ value equal to four.[75,76] This form of ϵ ($4r$) was chosen after several trial minimizations of carboxypeptidase A were done in which the value of ϵ was varied. When the value of ε was less than four, the side chain of Glu-270 was found to move away from its crystallographic position toward the zinc ion during minimization; this resulted in the ion being pulled away from its crystallographic position when the constraints were removed and the protein relaxed. With ε equal to four, the displacement of the Glu side chain did not occur and structural integrity was maintained. This form of ε ($4r$) was used only during the preliminary energy minimization in vacuum. For all dynamics calculations ε was set equal to 1 (see below). The entire enzyme was subjected to 1000 steps of steepest descent minimization. Every 100 steps the constraints were scaled by 0.65, progressively diminishing their value. The final value of the force constraints were 1.04 for backbone atoms, 0.52 for the side chain atoms, and 2.07 for the zinc ion. The rms coordinate differences from the crystal structure after energy minimization for carboxypeptidase A are 0.18, 0.27, and 0.26 Å for the backbone, side chains, and zinc ion, respectively; for carbonic anhydrase, the corresponding deviations are 0.13 Å for the backbone atoms, 0.25 Å for side chain atoms, and 0.25 Å for the zinc ion.

The stochastic boundary method for molecular dynamics (SBMD) was employed in the simulations of the zinc binding sites in carboxypeptidase A and carbonic anhydrase. This methodology has been described in detail elsewhere.[67,77,78] For the present study we defined a 16 Å reaction zone centered on the zinc ion of energy minimized structures of carboxypeptidase A and carbonic anhydrase. If any part of a residue was within a distance of 16 Å from the origin, the entire residue was treated as part of the reaction zone. The remainder of the molecule was designated as the reservoir zone. The reaction zone was further partitioned into a 12 Å reaction region

and a 4 Å buffer region. Main chain atoms were partitioned on an atom by atom basis, i.e., if any main chain atom was within 12 Å of the origin it was labeled a reaction region atom; if a main chain atom was between 12 and 16 Å it was labeled a buffer region atom. If any atom of the side chain was within the 12 Å sphere centered on the origin, the entire side chain was included in the reaction region. The protein connectivity of the reaction region was {12, 15, 18, 47, 63–80, 82, 107–120, 123–130, 139–147, 149, 151, 155–157, 162–167, 173, 175–176, 179, 192–208, 230, 236, 238–244, 246–258, 266–282, 286, 289–290, 292–293, 297, 300, 308(Zn)} for carboxypeptidase A and {5–8, 11–14, 16, 20, 22–23, 27–33, 47, 51, 58–70, 79, 84, 88–99, 103–109, 112–124, 131–132, 134–135, 140–149, 157, 160, 164, 170, 179, 181, 184–186, 191–212, 215–216, 218, 223, 226–227, 229–231, 241–249, 254, 256, 262(Zn)} for carbonic anhydrase. The reaction zone was solvated by overlaying a 30 Å box of water and deleting all water molecules which were either outside of the 16 Å radius, or overlapped a protein atom, or overlapped a crystallographic water molecule. To equilibrate the water around the protein, a Langevin dynamics simulation was done with the protein fixed at its minimized configuration. The friction coefficient for the water oxygen was assigned a value derived from the atomic friction coefficient computed from the velocity autocorrelation function of water, $\beta = 62\ ps^{-1}$.[77] No friction coefficient was placed on the hydrogen atoms. During the simulation, the list of water molecules being treated as Langevin particles was updated every 5 steps to account for diffusion of water across the boundary between the reaction region and buffer region. The equilibration of the water in the field of the fixed protein was carried out in a number of stages. First a high temperature simulation (T equal to 1000K) was done for 2.5 ps to facilitate barrier crossing as the water moved toward more favorable interactions. This was followed by 5 ps of simulation at $T = 300$K. A second overlay of water was followed by an additional 3 ps of dynamics at 300K. For carboxypeptidase A, the first overlay of water added 79 water molecules to the already present crystallographic waters, the second overlay added another 3 waters. In carbonic anhydrase, the first overlay added 27 waters and the second added 7 waters. At this point, protein atoms were allowed to move; weak harmonic constraints were placed on the reaction zone protein atoms with force constraints of 1.0, 0.5, and 1.0 kcal/mol-Å^2 for the reaction region backbone atoms, side chain atoms, and the zinc ion, respectively; values of 2.0 and 1.0 were used for the buffer region backbone atoms and side-chain atoms, respectively and the reservoir region was fixed at its energy minimized configuration. The system was equilibrated under these conditions for 2.5 ps at 100K. This was followed by equilibration at 300 K for another 2.5 ps with weaker harmonic constraints on the reaction region protein atoms with force constants of 0.2, 0.1 and 0.1 kcal/mol-Å^2 for backbone atoms, side chain atoms, and the zinc ion, respectively, and harmonic constraints with force constants derived from the crystallographic B factors[77] were placed on the buffer region protein atoms. A friction coefficient of 250 ps^{-1}, determined from the velocity autocorrelation function of several protein atom types,[77] was applied to the buffer region atoms; the value of 62 ps^{-1} was maintained for the water oxygens. In the simulations that used a standard shift model[53] for electrostatic truncation, the reservoir atoms were fixed at their minimized configuration. In the simulations that used the extended electrostatics model,[53,58] harmonic constraints with the force constants derived from crystallographic B factors were placed on the reservoir atoms to keep them close to their energy minimized positions while allowing them to fluctuate during the simulation. The range of values for the force constants derived from B factors is from 0.6 for the $O_{\gamma 1}$ of Thr-246 to 7.85 for the N of Met-301. By letting the reservoir atoms fluctuate during the extended electrostatics simulation one generates a *fluctuating* long-range electrostatic field at the solvated metal binding site. Previous studies have shown that a *static* long-range electrostatic field leads to increased damping of atomic fluctuations.[79] The results from the present studies indicate that atomic fluctuations tend to increase when including a fluctuating long-range electrostatic field. The treatment of electrostatic interactions is discussed in more detail below. After 2.5 ps of equilibration dynamics under these conditions, the harmonic constraints on the reaction region atoms were lifted, marking the point at which time t was set equal to zero. For the simulation of carboxypeptidase A using the standard shift function for electrostatic truncation, the molecular dynamics continued for another 68 ps and all averages were calculated over the final 42.9 ps; for the extended electrostatics simulation of carboxypeptidase A, the simulation continued for another 41 ps and the final 30.9 ps were used in averaging. For carbonic anhydrase, another 62 ps of molecular dynamics followed with the final 36.9 ps used for averaging. One simulation of carbonic anhydrase was run for a longer time (101 ps) and the final 74.9 ps were used for averaging. The time step used in all simulations was 1 fs and the SHAKE algorithm[80] was used to constrain the bonds between hydrogen atoms and heavy atoms. In the present application of the SBMD method with extended electrostatics, the reservoir atoms interact with the other reservoir atoms in the conventional way, i.e., via bonding and nonbonding interactions. This treatment represents an extension of the SBMD method that is very simple to implement. In the original SBMD calculation, the reservoir region

was neglected.[77] In a refinement of the present approach, the interactions between reservoir atoms are neglected and only isotropic harmonic fluctuations of reservoir atoms are allowed; this should yield a more efficient algorithm.

Treatment of Electrostatic Interactions

In the protein simulations, two different models for the electrostatic interactions were employed. One set of simulations used a standard shift model[53] in which the electrostatic interactions are smoothly reduced to zero at a cutoff distance of 12 Å. We refer to this as the SHIFT model. In the second set of simulations, an extended electrostatics model was used. This is referred to as the EXTENDED model, which approximates the full electrostatic interaction without truncation; the EXTENDED electrostatics model is described in detail elsewhere.[53,58] The EXTENDED electrostatics model was applied in the simulation study of binding interactions in the RNase A/3′-UMP enzyme–product complex.[81] This study demonstrated that inclusion of the extended electrostatics contribution was necessary for the uridine phosphate to remain correctly positioned in the active site. The present simulations show, particularly in the case of carbonic anhydrase, that long-range electrostatic interactions are important for maintaining the correct geometry of the zinc binding site when using the *nonbonded* model for zinc–protein interactions.

A constant dielectric with ϵ equal to one was used throughout the dynamics simulations. This is appropriate in the presence of an explicit solvent model.

RESULTS AND DISCUSSION

Solution Simulation Results

In Figure 1 the initial potential energy function for the interaction of Zn^{2+} with a water molecule is shown and compared with that obtained with the parameters given by Clementi et al.[62] The Lennard–Jones parameters for the CHARMM force field obtained from this initial fit of the ab inito derived surface are σ (Å) = 1.70 and ϵ (kcal/mol) = 0.67. The curves in Figure 1 were calculated for the Zn^{2+} ion–water interaction as a function of the distance between the zinc ion and the oxygen atom. The water model used in this initial fit was an early version of the water model currently used in the CHARMM22 all atom potential function[57]; the final stages of the ion parameter refinement used the current water model. The orientation of the water is such that the ion lies on the bisector of the H–O–H angle. The potential energy function obtained from the final parameter set is compared to potential energy function of Clementi et al.[62] (Fig. 1). The initial parameter set was refined by solution simulations and comparison to experimental radial distribution functions and free energies of solvation. The final interaction parameters are σ (Å) = 1.95 and ϵ (kcal/mol) = 0.25. The refined interaction surface yields a distance between the Zn^{2+} and water oxygen of 2.0 Å; the interaction energy at the minimum is approximately −69 kcal/mol. This result differs from that obtained from the purely ab initio calculations of Clementi et al.[62] In their calculation, using a minimal basis set, they determined a well depth of −75.6 kcal/mol at a minimum distance of 1.9Å; their empirical potential function has a slightly less deep well of −71.2 kcal/mol at a minimum distance of 1.9 Å (see Fig. 1).

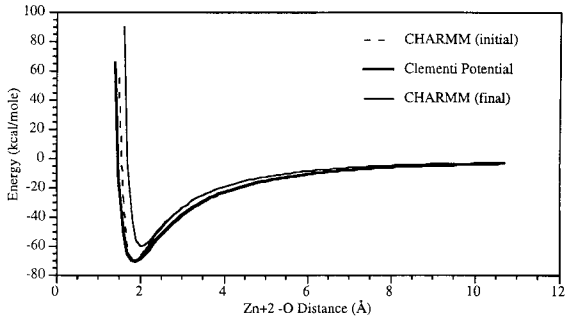

Fig. 1. Zinc ion–water interaction potentials for the initial potential energy function, the Clementi potential energy function[62] and the final potential energy function. Lennard–Jones parameters for the zinc ion within the CHARMM force field are $\sigma = 1.70$ Å and $\epsilon = 0.67$ kcal/mol for the initial potential energy function and $\sigma = 1.95$ Å and $\epsilon = 0.25$ kcal/mol for the final potential energy function.

The Zn^{2+}–O and Zn^{2+}–H radial distribution functions were calculated from the Monte Carlo simulations described in the Methods section with both the initial, {σ(Å) = 1.70 and ϵ(kcal/mol) = 0.76} and final, {σ(Å) = 1.95 and ϵ(kcal/mol) = 0.25}, parameter sets. The Zn^{2+}–O peak of the radial distribution function calculated from the simulation with the parameter set fit to the ab initio results was at 1.96 Å; this distance is shorter than those found in X-ray diffraction experiments.[82,83] The radial distribution function resulting from the final parameter set shows a well defined first solvation peak at 2.12 Å, which is within the range of experimental results (Fig. 2A); they place the first water solvation shell between 2.05 and 2.17 Å.[82,83] Integration over the first solvation shell yields a coordination number of 6, in agreement with the experimentally determined values. The second shell peak between Zn–O is at about 4.1 Å, in agreement with the experimental results.[83] The small sharp peaks at around 8 Å result from the truncation of the electrostatic term.[84]

The computed free energies of solvation are summarized in Table I. Entries 1 and 2 correspond to single step path (Path 1), i.e., the calculation changes the charged zinc ion to the methane sphere following a single path (see Methods); entries 3 and 4 correspond to the two-step path (Path 2) where the transformation of the zinc ion to the methane sphere

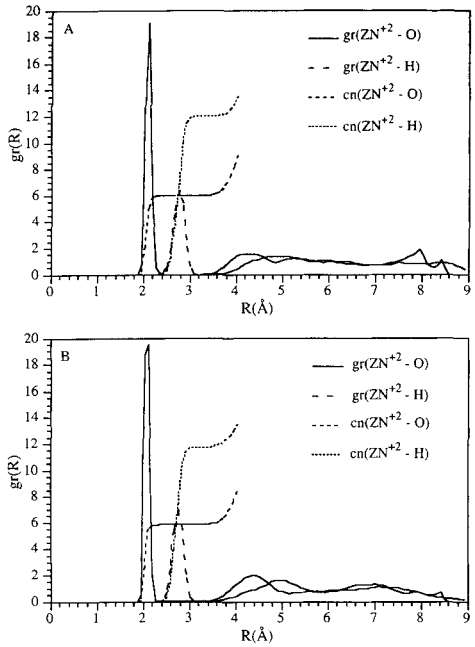

Fig. 2. Radial distribution functions for the zinc ion in water calculated from Monte Carlo simulations. (**A**) Lennard–Jones parameters of zinc corresponding to the final parameter set. (**B**) Åqvist–Warshel model for the zinc ion.

occurs in two steps, first the ion is decharged then the Lennard–Jones parameters are changed from those of the zinc ion to those of the methane sphere. Entries 5 and 6 correspond to the calculation of the two-step pathway with the larger water box and the longer cutoff, as indicated in Table I. The values in the column headed ΔG_{MC} are the results obtained from the MC simulations and the column headed $\Delta G_{solvation}$ gives the values obtained from them by use of Eq. (3). The free energy of solvation for the methane sphere in a box of TIP3P water is calculated to be 2.9 kcal/mol.[66] The Born correction was obtained from Eq. (4); it equals −77.14 kcal/mol at a cutoff distance of 8.5 Å and −69.02 kcal/mol at a cutoff distance of 9.5 Å. The value of ΔG determined by averaging the results from the individual simulations is −472.7 kcal/mol, 3% smaller than the experimental estimates which range from −480 to −485 kcal/mol.[85,86] The spread of values obtained from the different simulations is relatively small, indicating that errors due to system size, cutoff and nonconvergence are not important. The magnitude of the solvation free energy along the pathways where an ion was being discharged was greater than along the reverse pathway, i.e., the charging pathway. Differences between individual simulations can provide a measure of the statistical uncertainty of the calculated value. The rms difference is 18.15 kcal/mol or 3.84% of the mean calculated value. The hysteresis is less for the simulation with the two-step path than that using a single step. Slightly better agreement with the experimental values was obtained when employing a larger water box and cutoff as indicated by the results presented by entries 5 and 6. For comparison, the solvation free energy calculated using the initial potential function is −512 kcal/mol. These results show that the simple spherical model for the zinc ion in solution can reproduce the structural and energetic properties of the solvated ion within a few percent.

It is of interest to compare the present results with those obtained with the more complicated nonbonded model for zinc ion–water/protein interactions introduced by Åqvist and Warshel.[59] Their model corresponds to adding six fractional positive charges +0.5 in an octahedral geometry at a distance r_δ = 0.9 Å from center of the ion and a corresponding charge of −1.0 in the center, giving the ion a net charge of +2. The ion interacts as a single charged Lennard–Jones sphere at the center surrounded by 6 electrostatic centers. The model parameters used for zinc correspond to a Lennard–Jones particle with values of 1.49 Å and −38.19 kcal/mol for σ and ε, respectively; such a large ϵ value is apparently required to make their model fit the experimental results. The geometry of the six peripheral fractional charges is rigid, but overall rotation of the six-center frame about the nucleus is allowed and no internal forces are associated with such rotations, i.e., it rotates freely in response to interactions with the surrounding solvent and exerts no force on the central ion. This is done by placing small masses on the hydrogens and using the SHAKE algorithm to maintain the geometry.[87] In justifying their model, Åqvist and Warshel argue that ligand field effects in transition metal ions lead to a complex pattern of solvation energies. However, as mentioned earlier, the divalent zinc ion is a closed shell (d^{10}) ion and is not subject to ligand field effects. Berg and Merkle[88] have proposed that the unique conformational properties of the zinc ion are due, at least in part, to the *lack* of ligand field effects. The potential energy surface for the interaction of a zinc ion and a water molecule was calculated for the Åqvist/Warshel model with two different orientations of the water relative to the charges. In one orientation, the zinc ion lies on the line δ–Zn–δ and in the second, the zinc ion lies on the bisector of the δ–Zn–δ in the x–y plane. A small degree of anisotropy in the interaction energy is manifested by the orientation dependence of the interaction energy surface. However, the Åqvist/Warshel model and the present model give very similar ion–water interaction energies as a function of distance.

To compare the solution properties of the two models, Monte Carlo calculations employing the Åqvist model were done here to obtain the ion–water radial distribution function as well as the free energy of solvation. For these calculations, the SPC water

TABLE I. Solvation Free Energies for Zn^{2+} Calculated From Monte Carlo Simulations*

No.	Pathway	Cutoff/No. H_2O	ΔG_{MC}		$\Delta G_{Solvation}$
1	$Zn^{2+} \rightarrow CH_4$	8.5/216	403.62		−477.83
2	$CH_4 \rightarrow Zn^{2+}$	8.5/216	−388.87		−463.11
3	$Zn^{2+} \rightarrow Zn \rightarrow CH_4$	8.5/216	394.55	1.26	−470.05
4	$CH_4 \rightarrow Zn \rightarrow Zn^{2+}$	8.5/216	−1.25	−389.71	−465.20
5	$Zn^{2+} \rightarrow Zn \rightarrow CH_4$	9.5/512	416.73	1.71	−484.56
6	$CH_4 \rightarrow Zn \rightarrow Zn^{2+}$	9.5/512	−1.66	−407.57	−475.35

*Two pathways and their reverse are examined. System dependencies are examined by carrying out the simulations using two different cutoff lengths and two different box sizes. ΔG_{Born} is calculated using the experimental dielectric constant of 78 and it equals −77.14 and −69.02 kcal/mol at the cutoff distance of 8.5 and 9.5 Å, respectively. The free energy of solvation of a methane sphere in TIP3P water equals 2.902 kcal/mol (J. Gao, M. Karplus, unpublished calculations). The calculated free energies for each step of the pathway are given in the column labeled ΔG_{MC}; for the two-step pathway, the values of both are given. The total free energy, which is the sum of the entries in ΔG_{MC}, the solvation free energy of methane, and of the Born correction is given in the last column, $\Delta G_{Solvation}$.

model and the combination rules $\sigma_{ij} = (\sigma_i \sigma_j)^{1/2}$ and $\varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2}$ were used to be consistent with the conditions under which the ion parameters were developed. The simulation protocol of the previously described free energy perturbation calculations was employed. In Figure 2B, the ion to oxygen and ion to hydrogen radial distribution functions and the first peak integrations are shown. The results are very similar to those obtained with the simple spherical model (Fig. 2A). The calculated free energy of −488 kcal/mol is slightly larger than the experimental values, while the values obtained from the simple spherical model are slightly smaller. Thus, the more complex Åqvist/Warshel model does not give decidedly better or different results than a simple spherical model for the zinc ion in solution. It is possible that a model of the type suggested by Åqvist and Warshel would be more useful for open-shell ions. Of course, the bonded models make no sense for extensive solution simulations where exchange of waters can play a role. In the case of Zn^{2+}, the rate of water exchange has been estimated to be between 0.3 and 0.6×10^8 s^{-1}.[89] The solution simulations performed here were not sufficiently long to observe any water exchange.

Protein Simulation Results

A comparison of the simulation results obtained with the present model to those from simulations employing other models of the zinc binding site are used to examine the performance of the different types of models. In addition, we examine the effects of long-range electrostatic interactions on the structure and dynamics of a zinc-containing system.

Carboxypeptidase A

The protein studied first is the native form of carboxypeptidase A, the structure of which was solved by Rees and Lipscomb to 1.5 Å resolution[8] and for which XAFS structural data in solution are available.[39] There are 307 protein residues and 1 zinc ion in the active site. It is ligated to 2 histidines (His-69 and -196), 1 glutamic acid (Glu-72), and 1 water molecule. All histidine residues except His-13 and His-166 were assigned to the His($N^{\epsilon 1}$–H) tautomeric form. Earlier studies of carboxypeptidase A found that the His-166 was best assigned as the His($N^{\delta 1}$–H) and the His-13 was best assigned as a protonated histidine.[40] Other important residues present in the active site are Glu-270, Arg-127, Arg-145, and Tyr-248. The energy-minimized structure, prepared following the approach outlined in the Methods section, was used as the starting configuration in the SBMD calculations. The protocol for system equilibration given in the Methods section was followed. The production phase of the trajectory was started when the total energy plotted as a function of time reached a relatively steady value. For the shift simulation this required 25 ps of equilibration and only 10 ps for the extended electrostatics simulation. Data analysis was carried out on the remainder of the trajectory, or between 25 and 68 ps for the SHIFT simulation and between 10 and 41 psec for the EXTENDED simulation. We refer to this as the "production phase."

The average dynamics structure for carboxypeptidase A was calculated from the atomic coordinates of all atoms averaged over the production phase of the trajectory. In Table IIa, the rms coordinate difference between the time-averaged dynamics structure and the crystal structure is compared for the two simulations for all atoms and for the subset of reaction region atoms. The results are generally very similar. In all groupings of protein atoms, except for the zinc ion, the EXTENDED calculation yields slightly larger rms differences than the SHIFT simulation. This is probably due to the fact that in the EXTENDED calculation the reservoir region is not held rigidly in place as it is in the SHIFT simulation. This results in greater flexibility, which is confirmed by the slightly larger rms coordinate fluctuations observed in the EXTENDED simulation (see below).

The isotropic rms fluctuations of the heavy atoms

TABLE II. Rms Coordinate Differences Between the Crystal Structure of Carboxypeptidase A and the Average Simulation Structure (A) and rms Coordinate Fluctuations of the Reaction Region Atoms Calculated From the Molecular Dynamics Simulations (B)

Simulation	Backbone	Side chain	Zinc ion
	A*		
SHIFT	0.23, 0.28	0.44, 0.58	0.69, 0.69
EXTENDED	0.34, 0.35	0.68, 0.82	0.53, 0.53
	B†		
Crystal	0.38	0.48	0.45
SHIFT	0.24	0.35	0.25
EXTENDED	0.29	0.50	0.32

*Rms coordinate differences between the crystal structure of carboxypeptidase A and the average structure calculated from the molecular dynamics simulations (in Å). For each entry, the first value is that for the entire protein and the second value is that for the subset of atoms comprising the reaction region.
†Rms coordinate fluctuations of the reaction region atoms (in Å) calculated from the molecular dynamics simulations of carboxypeptidase A. The crystal structure values were calculated from the crystallographic B factors where $B_j = 8/3\ \pi^2 <\Delta r_j>$.

in the reaction region are compared with the rms fluctuations calculated from the experimental temperature factors in Table IIb. The theoretical values for the rms fluctuations calculated from the SHIFT simulations for the backbone atoms, the side chain atoms, and the zinc ion are smaller than the experimental estimates. This is consistent with what has been observed in other molecular dynamics simulations.[90] The rms fluctuations calculated from the EXTENDED simulations for the backbone atoms, side chain atoms, and the zinc ion are larger than in the SHIFT simulation and, except for the side chain atoms, are smaller than the experimental estimates. The rms fluctuations for the side chain atoms are slightly larger than the experimental estimates. The fact that the reservoir region is constrained either by being fixed to its energy-minimized position in the case of the SHIFT simulation or by the imposition of harmonic constraints in the case of the EXTENDED simulation may contribute to restraining the atomic fluctuations in the reaction region. However, permitting the reservoir region to fluctuate and including long-range electrostatic interactions seem to increase the overall flexibility of the reaction region. This pattern is observed in the simulations of carbonic anhydrase, discussed below. It should be noted that the simulations are too short[91,92] for full convergence of the rms fluctuations and, in addition, there are contributions to the experimental B factors from static or lattice disorder in the crystal.[93]

Of primary interest in this work is the comparison of the average structure of the zinc binding site obtained in the simulations described here, which employ a nonbonded model for the zinc ion, with the experimental X-ray crystal structure and with the average structures obtained from simulations employing bonded models for the zinc ion. Such a comparison is meaningful only if the protein structure is well preserved, as it is (see Table II). We consider the simulation results of Banci et al.[94,95] where a bonded model for the zinc binding site was used with the AMBER force field in a dynamics study of zinc-containing carboxypeptidase A. Their model is derived from the model developed by Merz and co-workers for simulation studies of carbonic anhydrase.[35] In developing the model for carboxypeptidase A, Banci and co-workers[94,95] used MNDO calculations to obtain the charge distribution for the zinc ion and active site residues His-69, Glu-72, and His-196. The resultant charge on the zinc ion was 0.71.[94] Explicit bonds with a force constant of 100 kcal/mol were introduced between the Zn^{2+} and the ligating nitrogen atoms of His-69 and -196. Explicit angle terms for N–Zn–N with a force constant of 50 kcal/mol-rad^2 and for Zn–N–C with a force constant of 10 kcal/mol-rad^2 were also introduced. Equilibrium bond lengths and angles, required for the model, were obtained from the X-ray structure of carboxypeptidase A. Initially, there was no direct bond between the zinc ion and the carboxylate group of the Glu-72. However, during the simulations the bidentate binding of Glu-72 to the zinc was not preserved and additional bonding terms with a force constant of 20 kcal/mol were introduced between the ion and the carboxylate oxygens.[94] Consistent with the bonded models proposed by Merz et al.,[35] specific bonds were introduced between the zinc ion and the bound water oxygen. Parameters for the van der Waals interaction, σ and ε of the zinc ion, were not listed in the publications of Banci et al.[94,95] If they used the same values as in the reference of Merz et al.,[35] σ equals 1.25 Å and ε equals 0.1 kcal/mol. However, these parameters are expected to play a small role in the bonded models. The simulation described by Banci et al.[94] used an 8 Å residue-based truncation model, while the study described by Banci et al.[95] used a slightly longer (9 Å) residue based truncation. *All* bond lengths were constrained to their equilibrium values using the SHAKE algorithm, including the zinc ion–water oxygen bond distance. This would make the choice of force constants for protein–zinc bonds immaterial. In the CHARMM calculations, only bonds between heavy atoms and hydrogen atoms were constrained using the SHAKE algorithm. Thus there is a much higher degree of conformation freedom. In particular, there are no constraints (either via SHAKE or via bonding terms) between the zinc and any of its ligands. Consequently the structural results obtained from the simulations are a stringent test of the zinc nonbonded interaction model. Coordinates of the average structure of the zinc binding site, which includes the zinc ion, His-69 and -196,

Glu-72 and -270, and the zinc bound water molecule, as determined from the molecule dynamics simulation described by Banci et al.,[95] were made available by Dr. L. Banci and Dr. G. LaPenna for the present comparison.

Figure 3 shows stereo views of the average structures of the zinc binding sites superimposed on the crystal structure. One sees that all three models give approximately the same structure in the zinc binding site. The largest deviation from the crystal structure and, correspondingly, difference among the models involves Glu-270. To obtain a more detailed understanding of the results for the zinc coordination from the different models, we present bond lengths, bond angles, and dihedral angle values in the subsequent tables. In Table III, we list certain distances from the different average structures and the crystal structure of the zinc binding site which permit a comparison of the different models. The first entries of Table III focus on the five ligands coordinated directly with the zinc in the X-ray structure. The interatomic distances between the zinc ion and the ligating atoms ND1(His-69), OE1(Glu-72), OE2(Glu-72), ND1(His-196), and the oxygen of the zinc bound water are listed. The average structure from the dynamics simulations of Banci et al. agrees very well with the crystal structure; the deviations are 0.1 Å or less. This is not surprising since all the bond distances between the zinc ion and its ligands are introduced as restraints in the model. The average structures from the SHIFT simulations and from the EXTENDED simulations, in which no explicit restraints are used, are also in good agreement with both the crystal structure and the average structure from the simulations described by Banci et al. All distances are within 0.1 Å or less from the crystal structure, except that to ND1 of His-196, which is 0.2 Å longer in the SHIFT simulation, and that to OE2 of Glu-72, which is 0.3 Å shorter in the EXTENDED simulation. In both CHARMM calculations, the zinc ion to water distance is in slightly better agreement with the crystal structure than in the calculations of Banci et al. The remaining entries in Table III are distances between the zinc or the zinc bound water and other atoms of the binding site. For these cases, the Banci et al. model included only nonbonded interactions, as in the CHARMM calculations. Glu-270 is believed to be involved in the activation of the water molecule during the enzymatic reaction. In the crystal structure the distance between the zinc ion and OE1 of this residue is 4.2 Å. In the simulations of Banci et al. the distance is 5.8 Å (see Table I of Banci et al.[94]) and 4.9 Å from Banci et al.[95] The first calculation used an 8 Å residue based cutoff while the second used a longer cutoff distance of 9 Å. The SHIFT simulation in this work, which uses a 12 Å cutoff, yields an improved value of 4.1 Å and inclusion of the long-range electrostatic interactions via the EXTENDED model gives a value of 5.2 Å. For OE2 of Glu-270, the crystal structure gives a distance of 4.7 Å from the zinc ion and the Banci simulations give distances of 5.8 and 6.2 Å for the two simulations. The SHIFT simulation gives a value of 4.4 Å while inclusion of long-range electrostatic interactions improves this value to 4.7 Å. Corresponding results (i.e., that the EXTENDED value is closer to the crystal than the SHIFT) are obtained for most of the distances given in Table III. The only exceptions are the Zn–OE1 and the OH2–OE1 of Glu-270 which show a larger deviation in the EXTENDED simulation than in the SHIFT, but it is still an improvement over the model of Banci et al. The results are shown in Figure 4 which gives a scatter diagram for the calculated and experimental distances; along the diagonal, the simulation distances equal the crystal distances. It is clear from this diagram that the inclusion of long-range electrostatic interaction, either by increasing the cutoff distance or by using the extended electrostatics model, improves the agreement between the average simulation structure and the crystal structure.

In Tables IV and V angle and dihedral measures are given to further compare the geometry of the zinc ion binding site obtained using the various models. In Table IV angles defined by the positions of three atoms including the zinc ion are given. The first five entries in Table IV report angle values defined by atoms which ligate the zinc ion. For the first two entries, the bonded model specifies explicit energy terms in the potential energy function, while in the case of the SHIFT and EXTENDED simulations, no explicit angle terms are included. The agreement between the three models is quite good, though the value of Banci et al. is closer to the X-ray value than the SHIFT and EXTENDED results, as expected. For the third and fourth entries no explicit *angle* terms are included in the bonded model, but there are bonding terms between the zinc ion and the oxygens of the glutamic acid which constrain the angle. The results of the two models are comparable. For the angle ZN–OE1–CD, the bonded model and the EXTENDED model give essentially the same result of 90° and 91°, respectively, while the SHIFT model gives a result in better agreement with the crystal structure, 97° vs 95°. For the angle ZN–72OE2–72CD, the bonded model is in better agreement with the crystal structure value than either of the two nonbonded calculations. For the angle ZN–196ND1–196CG, one finds the deviation of the bonded model from the crystal value is less than that of the EXTENDED result and both are greater than that of the SHIFT result. The bonded model includes an angle energy term defined by the atoms ZN–ND1–C; in the paper of Merz et al.[35] it is not specified which carbon angle this last angle term uses. The final entry involves Glu-270, the zinc bound water oxygen, and the zinc ion. Both the

TABLE III. Interatomic Distances in Carboxypeptidase A Calculated From the Crystal Structure (XTAL), the Present Molecular Dynamics Simulations (SHIFT and EXTENDED), and the Molecular Dynamics Simulations Employing the Bonding Model (BONDED) Proposed by Bianci et al.[93,94]

Residue/atom	Residue/atom	R_{ij} (XTAL)	R_{ij} (SHIFT)	R_{ij} (EXTENDED)	R_{ij} (BONDED)
Zn^{2+}	69 ND1	2.1	2.2	2.2	2.1*
Zn^{2+}	72 CD	2.6	2.5	2.5	2.5*
Zn^{2+}	72 OE1	2.2	2.1	2.1	2.1*
Zn^{2+}	72 OE2	2.3	2.2	2.0	2.2*
Zn^{2+}	196 ND1	2.1	2.3	2.2	2.1*
Zn^{2+}	OH2 (zinc bound)	2.1	2.1	2.1	2.0*
Zn^{2+}	270 OE1	4.2	4.1	5.2	4.9*, 5.9†
Zn^{2+}	270 OE2	4.7	4.4	4.7	6.2*, 5.8†
OH2 (zinc bound)	270 OE1	2.5	3.3	3.4	4.0†
OH2 (zinc bound)	270 OE2	3.2	2.6	2.8	3.9†
Zn^{2+}	248 OH	18.7	17.7	18.1	18.0†
OH2 (zinc bound)	248 OH	17.3	17.5	17.2	16.4†

*Banci, L., Bertini, I., La Penna, G., Inorg. Chem. 32:2207, 1993.
†Banci, L., Bertini, I., Carloni, P., Luchinat, C., Orioli, P.L. J. Am. Chem. Soc. 114:10683, 1992.

bonded model and the EXTENDED model result in deviations from the crystal structure greater than 20°, while the SHIFT model shows a deviation of about 12°.

A comparison of dihedral angles χ_1 and χ_2 for the zinc binding site residues is given in Table V. For the residues directly bonded to zinc, the various models give similarly good agreement with deviations on the order of 20° or less. The exception is the χ_2 of His-69 where the EXTENDED model shows a deviation of 30°. In some cases (e.g., χ_1 of residue 69 and χ_2 of residue 72, EXTENDED) the nonbonded model is better, while for others (e.g., χ_2 of residue 69 and χ_1 of residue 196) the bonded model is better; in the latter there are no specific restraints on the dihedral angles involved. For χ_1 of residue 270, the nonbond shift model gives results close to the X-ray structure, while for χ_2 of residue 270 the EXTENDED model and the bonded model have dihedral angles in an alternative minimum. It is not clear whether this difference is significant, e.g., the change in both χ_1 and χ_2 found in the EXTENDED simulation leaves CB and CG in alternate minima and the remainder of the residue 270 in a similar position. Such "anticorrelated" changes, i.e., when two adjacent dihedrals rotate in opposite directions, have been observed in other simulations.[90]

Additional dihedral measures are given (Table V) for angles defined between binding site residues and the zinc ion; an example is the dihedral defined by the four atoms {72CG, 72CD, 72OE1, ZN} or {72CD, 72OE1, ZN, 69ND1}. These dihedrals reflect the relative orientation between one group and another. It is clear from this result that the bonded model does not give results that are significantly better than either of the two nonbonded models and the results from the models are all within 10° of the other. The only exception is the fourth dihedral defined by the His-196 residue and the zinc ion (CB–CG–ND1–ZN) where the bonded model results are noticeably closer to the crystal structure than the nonbonded models. Overall, however, these results indicate that for the simulations described here, the deviations from the crystal structure of the bonded model are not significantly smaller than either the SHIFT or the EXTENDED simulations that use a nonbonded model for the zinc ion.

One measure of dynamic behavior in protein systems is provided by the rms fluctuations. In Figure 5, heavy atom rms coordinate fluctuations for the zinc ion, the zinc ion binding residues His-69, Glu-72, His-196, the zinc bound water, and Glu-270 are compared. Values from the crystal structure were calculated from the temperture factors in the manner described earlier (without any disorder corrections). These are compared to the results from the SHIFT and EXTENDED electrostatics simulation as well as from the simulation results of Banci et al.[95] One can be seen that in general there is good agreement between the simulations and experimental estimates. The SHIFT and the EXTENDED simulations for the ion-bound ligands are very similar, except for Glu-72, which shows larger fluctuations in the EXTENDED simulations than the SHIFT simulations or the experimental estimates; for the more distant residue Glu-270, the rms fluctuations of the SHIFT simulations are in better agreement with the crystal values. It is interesting to note that XAFS studies of carboxypeptidase A indicate that in solution one oxygen of the Glu-27 is displaced from its crystallographic position[39]; the larger fluctuations seen for the Glu-72 residue in the EXTENDED simulations may reflect this conformational change. The results from the Banci model show qualitatively the same behavior as the CHARMM simulations, although some of the maxima correspond to larger

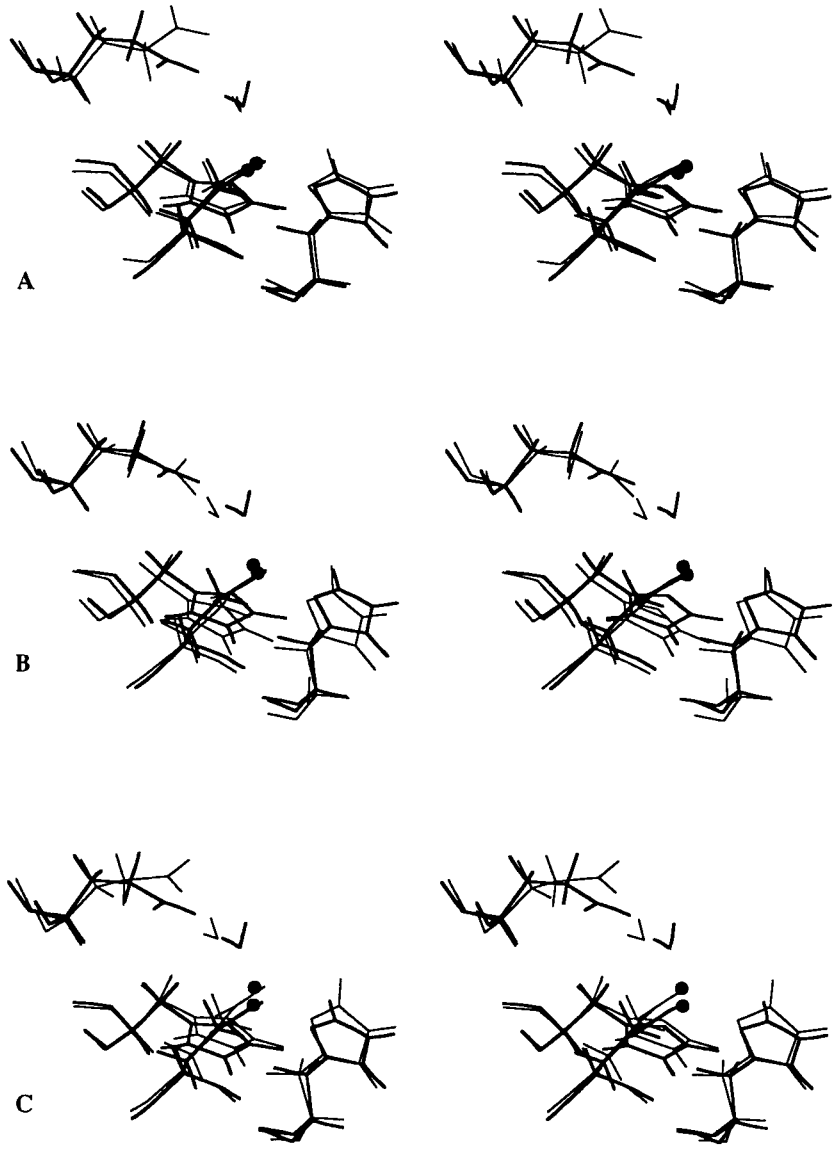

Fig. 3. Stereo views of the zinc binding site. Average dynamics structures (thin lines) are superimposed over the X-ray crystal structure (thick lines). (**A**) Average structure from dynamics simulation employing the model of Banci et al.[95] (**B**) same as (**A**) except that the average structure was calculated from the dynamics simulation employing the SHIFT model for electrostatics: (**C**) As in (**B**) but for the EXTENDED electrostatics model.

fluctuations, particularly for the carbonyl oxygen of the Glu-72 residue.

Carbonic anhydrase

The second protein studied is the native form of carbonic anhydrase II (1CA2), the structure of which was solved by Eriksson et al. to 2.0 Å resolution.[5] This zinc metalloenzyme catalyzes the reversible hydration of CO_2 to produce bicarbonate, HCO_3^-. There are 260 protein residues and 1 zinc ion in the active site ligated to 3 histidines (His-94, -96, and -119), and 1 oxygen from either a water molecule or a hydroxide ion. Except for His-94 and His-96, all histidines were assigned the His($N^{\epsilon 1}$–H) tautomeric form. The former two were assigned the His($N^{\delta 1}$–H) form after examination of the crystal structure around the zinc ion. The $N^{\epsilon 1}$ of these two His residues ligate the zinc ion, while in His-119 the $N^{\delta 1}$ ligates the zinc ion. Other important residues present in the active site are His-64, Glu-106, and Thr-199.

The energy minimization followed the procedure outlined in the Methods section, with the one exception that during the energy minimization, distance

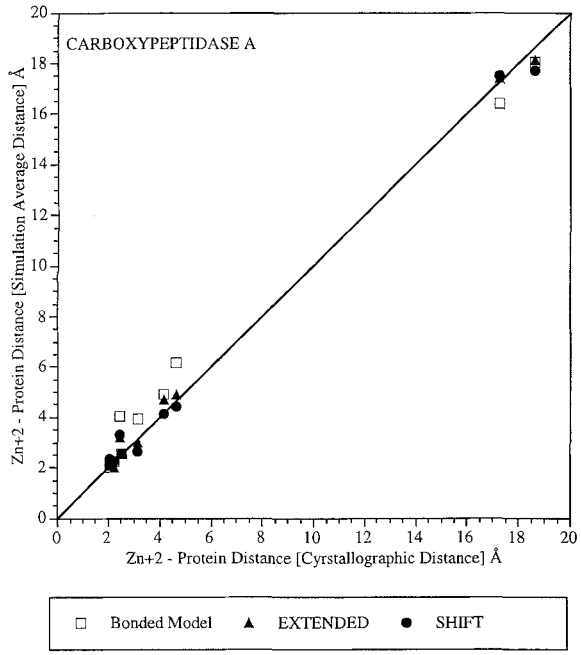

Fig. 4. Scatter diagram of the average distances from the present SHIFT (●) and EXTENDED (▲) simulations of carboxypeptidase A and the average distances from simulations of Banci et al.[95] (□) as a function of the crystal distances (see Table III). Along the diagonal, the simulation distances equal the crystal distances.

constraints were introduced between the zinc ion and the ligating nitrogen atoms of the zinc bound histidine ligands as well as to the OE1 of Glu-106. This was found to be necessary to minimize the deviation of the zinc binding site from the crystal structure, in accord with the results for carboxypeptidase A. In both cases energy minimization is done under vacuum conditions with electrostatic shielding introduced only via a distance-dependent dielectric constant. Although a useful approximation, such a simple distance-dependent dielectric is insufficient to reproduce the subtle effects of explicit solvation.[96] All constraints were removed in the molecular dynamics simulations. The molecular dynamics procedure outlined in the Methods section was used. Four simulations of 61 ps were done using the EXTENDED model and two simulations using the shift truncation model were done, one for 61 ps and the other for 101 ps. The initial conditions were changed by assigning a different initial velocity distribution. Results from the two SHIFT simulations and from the four EXTENDED simulations are reported here. These simulations employed the nonbonding model for the zinc ion as used in carboxypeptidase A.

An average structure of carbonic anhydrase from each simulation was obtained by averaging the atomic coordinates over the production phase of each simulation. The structural results obtained are compared to the experimental X-ray crystal structure as well as to those reported by Hoops et al.[34] Hoops et al. compared simulation results obtained using both a nonbonded model and two forms of a bonded model, referred to as "large" and "small,"[34] in vacuum simulations with the AMBER potential function. The terms "large" and "small" refer to the size of the model systems used in the determination of the partial charges, which differ between the two models. The simulation time was 12 ps. A 10 Å cutoff distance was used with simple truncation and a constant dielectric of ε equal to 1. During their energy minimization and dynamics calculations, *all* bond lengths, including the distances between the zinc ion and its ligating atoms in the nonbonded model, were constrained to their equilibrium values by use of the SHAKE algorithm. Thus even the "nonbonded" model has specific terms that constrain the distances of all atoms coordinated to the zinc. There are no such constraints in the CHARMM calculations. In the calculations of Hoop et al., only crystallographic waters were included while in the CHARMM calculation, a 16 Å sphere of water solvated the reaction region centered on the zinc ion. In the crystal structure, there are 40 water molecules within a 16 Å sphere around the zinc ion, while the water sphere in the present simulations was made up of 68 waters. These differences in the simulation systems complicate the comparison of the results, i.e., observed differences are not necessarily due to the model for the zinc. Nevertheless, they are of interest as an additional test of the nonbonded model. Rms coordinate differences between the two SHIFT and the four EXTENDED simulations and the X-ray structure are similar to those obtained for carboxypeptidase A; the same hold true for the rms fluctuations.

In Table VI, average distances between the zinc ion and a number of protein atoms calculated from the present molecular dynamics simulation are listed, as well some results obtained by Hoops et al.[34] The distances from the present simulations are averages over the two SHIFT and over the four EXTENDED simulations. These data are shown in a distance correlation diagram in Figure 6. The results of Hoops et al.[34] were obtained from molecular dynamics simulations in vacuum where the zinc–protein interactions were treated by either a nonbonded model (column 4; NB) and or one of two bonded models (columns 5 and 6; B small and B large, respectively). Their results were based on the single structure obtained at the end of their simulation protocol; however, the results published are representative of results obtained at times earlier in the simulations.[97] The results from the present simulations are obtained by averaging the distances from each time frame of the simulations; these will be compared to the distances obtained by Hoops et

TABLE IV. Angles in the Zinc Ion Binding Site of Carboxypeptidase A Calculated From the Crystal Structure (XTAL), the Present Molecular Dynamics Simulations (SHIFT and EXTENDED), and the Molecular Dynamics Simulation Employing the Bonded Model Proposed by Banci et al.[95] (BONDED)

	θ (IJK)					
I	J	K	θ (XTAL)	θ (SHIFT)	θ (EXTENDED)	θ (BONDED)
69 ND1	Zn^{2+}	196 ND1	98.8	83.8	89.2	100.1
Zn^{2+}	69 ND1	69 CG	124.5	132.7	129.7	127.9
Zn^{2+}	72 OE1	72 CD	95.1	97.4	89.8	90.8
Zn^{2+}	72 OE2	72 CD	87.6	89.0	98.7	87.3
Zn^{2+}	196 ND1	196 CG	127.7	125.8	133.2	124.7
Zn^{2+}	OH2	270 OE2	125.9	137.5	163.1	149.4

TABLE V. Dihedral Angles in the Zinc Binding Site of Carboxypeptidase A Calculated From the Crystal Structure (XTAL), the Present Molecular Dynamics Simulations (SHIFT and EXTENDED), and the Molecular Dynamics Simulations Employing the Bonded Model Proposed by Banci et al.[94,95]

	ϕ (IJKL)							
	I	J	K	L	ϕ (XTAL)	ϕ (SHIFT)	ϕ (EXTENDED)	ϕ (BONDED)
χ_1	69 N	69 CA	69 CB	69 CG	−66.82	−67.80	−72.35	−57.00
χ_2	69 CA	69 CB	69 CG	69 ND1	141.44	162.79	171.52	130.72
χ_1	72 N	72 CA	72 CB	72 CG	−61.33	−59.29	−87.11	−65.09
χ_2	72 CA	72 CB	72 CG	72 CD	165.03	−178.53	167.96	171.29
χ_1	196 N	196 CA	196 CB	196 CG	−71.31	−82.03	−83.77	−80.50
χ_2	196 CA	196 CB	196 CG	196 ND1	−108.79	−89.52	−96.48	−96.68
χ_1	270 N	270 CA	270 CB	270 CG	−67.44	−56.56	−162.66	−69.91
χ_2	270 CA	270 CB	270 CG	270 CD	−81.43	−98.60	144.03	−155.39
	69 CB	69 CG	69 ND1	ZN	−6.0	−11.9	−15.32	−4.3
	72 CG	72 CD	72 OE1	ZN	−177.9	−178.9	177.25	173.2
	72 CG	72 CD	72 OE2	ZN	178.2	179.1	−177.02	−173.2
	196 CB	196 CG	196 ND1	ZN	8.5	−31.1	−11.86	4.2
	72 OE1	ZN	OH2	270 OE1	−143.9	−179.1	177.90	−103.2
	72 CD	72 OE1	ZN	69 ND1	−128.2	−117.5	−121.63	−134.2
	72 CD	72 OE1	ZN	196 ND1	−14.1	−34.3	−27.53	2.4

al.[34] This is different from carboxypeptidase A where *average* structures from different simulations were compared. The distinction drawn here between the average structure and average distances is analogous to the distinction between the *actual* structure of a protein and its crystal structure. In X-ray crystallography an optimized average structure is found where the average is over a large number of molecules present in the crystal and optimized through the use of any one of several refinement procedures.[98] Using molecular dynamics simulation of bovine pancreatic trypsin inhibitor (BPTI) it was shown that bond lengths obtained from an average structure are generally shorter than explicitly averaged distances, which are evaluated by determining the bond lengths for each individual coordinate set in the trajectory and then averaging.[98,99] In the present simulations, corresponding dynamic averaging is observed for longer distance atom pairs. In all cases, the distances calculated from the average structures are equal to or slightly less than the explicitly averaged distances. A striking example is found in one of the EXTENDED simulations where the zinc–water oxygen distance calculated from the average structure is 1.62 Å, but the explicitly averaged value is 2.11 Å.

The first entries of Table VI give the distance between the zinc ion and its 4 ligating atoms. The interatomic distances between the zinc ion and the ligating atoms NE1 (His-94), NE1 (His-96), ND1 (His-116), and the oxygen of the zinc bound water are listed; note that the definitions of these nitrogens are inverted in Table X of Hoops et al.,[34] i.e., NE1 (His-94) is reported as ND1 (His-94), NE1 (His-96) is reported as ND1 (His-96), and ND1 (His-116) is reported as NE1 (His-116). The four distances from the structure of Hoops et al. agree well with the crystal structure; the deviations are 0.15 Å or less. This is not surprising as explicit constraints on all of these distances are present in both the nonbonded and bonded models of Hoops et al.[34] The dynamics averages from the SHIFT and EXTENDED simulations tend to be slightly longer (0.1 to .2 Å). For distances between atoms that are not bonded, the nonbonded model of Hoops et al.[34] has many large errors; for example, the distance between the zinc and the OG1 of Thr-199 is 3.83 Å in the crystal structure and the nonbonded model of Hoops et al.[34]

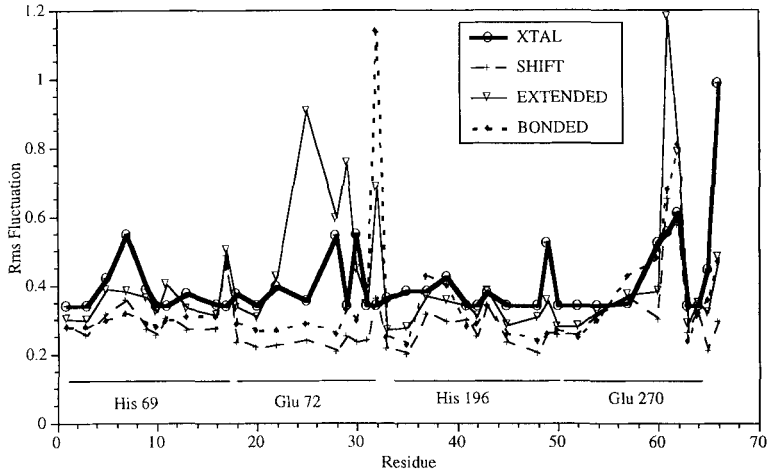

Fig. 5. Rms coordinate fluctuations for the zinc ion binding site atoms including His-69, Glu-72, His-196, Glu-270, the zinc ion (#65), and the zinc bound water oxygen (#66).

TABLE VI. Interatomic Distances in Carbonic Anhydrase Calculated From the Crystal Structure (XTAL), the Present Molecular Dynamics Simulations (SHIFT and EXTENDED), and the Molecular Dynamics Simulations Employing the Nonbonded Model (NB) and the Two Bonded Models (B), Small and Large, Proposed by Hoops et al.[34,*]

Residue	Atom	XTAL	NB	B small	B large	SHIFT	EXTENDED
94 His	NE2	2.00	2.05	2.05	2.05	2.26	2.17
96 His	NE2	2.10	2.05	2.05	2.05	2.15	2.22
119 His	ND1	1.91	2.05	2.05	2.05	2.21	2.26
H2O	OH2	2.00	2.05	2.05	2.05	2.11	2.11
199 Thr	OG1	3.83	6.07	4.40	4.44	3.51	4.27
106 Glu	OE1	4.06	1.79	3.95	4.08	1.99	4.66
106 Glu	OE2	5.46	3.63	5.44	4.08	3.79	5.81
200 Thr	OG1	6.42	4.69	4.43	4.12	6.48	4.88
117 Glu	OE1	6.45	6.68	6.78	6.70	5.05	8.68
117 Glu	OE2	7.72	6.80	8.13	7.90	6.60	6.65
64 His	NE2	7.92	10.25	7.96	8.04	8.47	6.77

*The values given for the present simulations are average values from the two SHIFT simulations and the four EXTENDED simulations.

gives a value of 6.07 Å. The bonded models of Hoops et al.[34] are significantly better (4.4 and 4.44 Å). For this distance, both the SHIFT and the EXTENDED nonbonded model calculations give values in better agreement with the crystal result than any of the models of Hoops et al.[34] The extended electrostatics model has some errors that are similar to those of Hoops et al.,[34] e.g., the distance between zinc and OG1 of Thr-200, where the crystal structure distance is 6.42, the bonded models of Hoops et al.[34] give distances of 4.43 and 4.12 Å; the EXTENDED simulations give values between 4.19 and 5.40 Å. Of particular interest are the residues Glu-106 and Thr-199 which are believed to be important in catalysis, i.e., Glu-106 is part of a hydrogen bonding network along with the zinc bound water, or hydroxide ion, and Thr-199. Although the function of this network is not known, it is believed to be important because it is present in all the carbonic anhydrase structures[2,5] and Thr-199 and Glu-106 are conserved in all sequences of carbonic anhydrase.[100] Results from molecular dynamics simulations[101] and from crystallographic analysis[102] have led to the proposal that the function of this hydrogen bonding network is to position the zinc bound water (or hydroxide ion) in an optimal orientation for nucleophilic attack on the CO_2 substrate.

In Table VII, several distances relevant to this hydrogen bonding network are given. All the distances, except for the crystal structure values, are obtained by averaging over the trajectory; the rms fluctuations are given in parentheses. Hydrogen positions determined using the HBUILD facility in the CHARMM program were used in calculating the crystal values. Referring to Table VII, all of the values for the distance between the zinc and the zinc

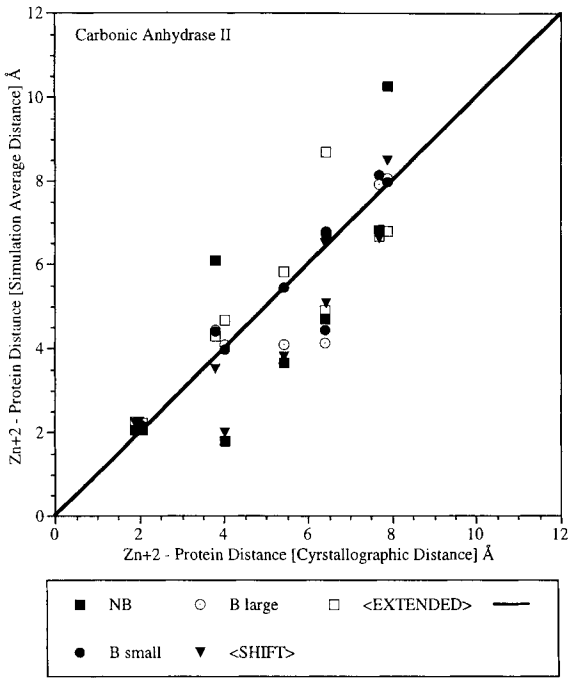

Fig. 6. Scatter diagram of the average distances (shown in Table VI) from the two SHIFT (▼) and the four EXTENDED (□) simulations of carbonic anhydrase and the distances from the simulations of Hoops et al.[45] as a function of the crystal distances. For the results of Hoops et al., the nonbonded model (NB) (■); the bonded model labeled small (●); the bonded model labeled large (○). Along the diagonal, the simulation distances equal the crystal distances.

bound water oxygens (Zn–OH2) are in good agreement with the crystal structure values. This was discussed in detail above. In all simulations but one, the water molecule originally present in the crystal structure remained in its crystallographic position; however, in one simulation which used the extended electrostatics model, this crystal water was displaced by another water molecule which then occupied the crystallographic position. This type of interchange is expected to occur in reality, but would be prevented by use of a bonded model. Such an interchange of a catalytically significant water has also been observed in the solvated active site of acylchymotrypsin.[103] In the columns labeled H1(H2O)–OG1 and H2(H2O)–OG1 are the distances between the hydrogens of the zinc bound water (H1 and H2) and the OG1 oxygen of Thr-19 (see Fig. 7). In both the SHIFT and EXTENDED simulations one of the two water hydrogens is involved in a hydrogen bond to the OG1 of Thr. For the distances between the hydrogen bound to OG1 of Thr-199 (HG1) and the OE1 oxygen of Glu-106, the EXTENDED values are shorter than those from the SHIFT simulations. In the SHIFT simulations, the distance HG1–OE2 is the shorter of the two; this is a result of the conformational change of the Glu-106. In the crystal structure, the distance between the zinc ion and OE1 of this residue is 4.1 Å. In a molecular dynamics simulation which used their nonbonding model, Hoops et al.[34] obtained a distance of 1.79 Å. The OE1 oxygen moved into the first solvation sphere of the zinc ion. The distance obtained when using the bonded models is in better agreement with the crystal structure results, i.e., 3.95 and 4.08 Å. In the present calculations, we observe the same behavior in the SHIFT simulation as Hoops et al.[34] observed in their nonbonded model simulations, i.e., the OE1 of Glu-106 moves into the first solvation sphere of the zinc ion to a relative separation of 2.0 Å. when using the EXTENDED electrostatics model, however, one finds in all four simulations, Glu-106 remains close to its crystallographic position and that OE1 does not move into the first solvation sphere of the zinc ion. The EXTENDED electrostatics model gives very good results for both OE1 and OE2 of Glu-106. The nonbonded model of Hoops et al.[34] used a 2+ charge on the zinc and a reduced charge of either 0.866 or 1.033 for the "small" and "large" bonded models, respectively. It was found necessary in their calculations to reduce the charge on the zinc ion to obtain agreement with the crystal structure. The SHIFT calculations use a 12 Å cutoff and a solvated active site but they still show similar behavior. However, the EXTENDED calculations are significant in this respect since the OE1 of Glu-106 remains close to its crystallographic position. This indicates that the difficulties concerning the displacement of Glu-106 in the simulations of Hoops et al.[34] had more to do with their treatment of electrostatic interactions and solvation than with the charge of the zinc ion. For Glu-117, the EXTENDED model shows greater deviations from the crystal structure than the bonded models of Hoops et al.[34]

CONCLUSIONS

Molecular dynamics and Monte Carlo simulations have been used to investigate the utility of a simple nonbonded model for the zinc ion in aqueous solution and in protein simulations. The nonbonded model introduced here treats the zinc as a charged (2+) sphere with nondirectional Coulombic and Lennard–Jones interactions with the rest of the system; other nonbonded models have used reduced charges. The results from these simulations are compared to those obtained with specific bonded models in proteins and more complicated nonbonded models in solution. In addition, comparisons with experiment are made for the corresponding systems.

The studies described here provide a detailed comparison between the different treatments to assess the strengths and weaknesses of each. Such an analysis is essential for the continuing development of computer models for the study of a wide range of zinc-containing systems. The results demonstrate that the use of a more complicated model (either

TABLE VII. Distances in the Zn–Thr-199-Glu-106 Hydrogen Bond Network (see Fig. 7)*

	Zn–OH2	H1 (H2O)–OG1	H2 (H2O)–OG1	HG1–OE1	HG1–OE2
XTAL	2.09	3.38	2.76	2.82	3.22
SHIFT	2.13 (0.06)	2.09 (0.27)	2.77 (0.21)	2.10 (0.19)	1.94 (0.20)
EXTENDED 1	2.11 (0.37)	2.48 (0.57)	2.62 (0.57)	1.72 (0.13)	2.96 (0.30)

*The distances are given in Å and are average values calculated and averaged over each time frame. The numbers in parentheses are the average rms fluctuation in the distance. Column 1 contains the distances between the zinc and the zinc bound water and the second and third columns contain the distances between the two hydrogens of this water and the OG1 oxygen of Thr-199. The final two columns contain the distances between the HG hydrogen, the hydrogen bound to the OG1 oxygen of Thr-199, and the two carboxyl oxygens of Glu-106.

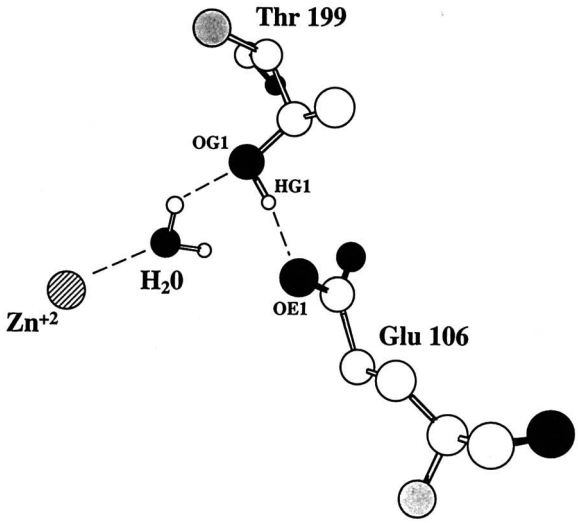

Fig. 7. Schematic of the hydrogen bond network involving the zinc ion, the zinc bound water molecule, Thr-199 and Glu-106.

bonded or nonbonded) does not lead to an improved representation of the system. In the solution simulations, experimental quantities such as the radial distribution function and the free energy of solvation can be obtained to a corresponding degree of accuracy with either the simple spherical model developed here or the more complicated distributed charge model of Åqvist and Warshel.[59] In the simulations of carboxypeptidase A and carbonic anhydrase there are no significant differences between simulation with the nonbonded spherical model and those employing a bonded model for the zinc binding site.

It has been shown that parameters determined from ab initio calculations, when used directly in condensed phase simulations, tend not to reproduce experimental values of solvation energies or free energies.[57,104,105] This is confirmed for the zinc ion in solution by the present free energy of solvation calculations. The solvation free energy calculated using the potential energy function parameters obtained by fitting to the ab initio result overestimates the solvation free energy. This may be due to the fact that a simple functional form like Eq. (1) is used to represent the interactions. Possible improvements to the nonbonded model could include a softer van der Waals potential and local polarization, such as that introduced by Roux and Karplus.[106,107]

The present simulations provide no evidence that bonded models are unsatisfactory for a well-defined fixed site. Both the structural and dynamics results appear to be satisfactory. However, the simpler, nonbonded model with proper treatment of truncation (see Fig. 7) gives equally good results and includes the flexibility required for a change in the coordination. This is essential for solution studies (i.e., solvent molecules move in and out of the first coordination shell) and is likely to be of importance for studying catalysis by zinc enzymes where changes in the coordination play an important role, e.g., in carboxypeptidase A. Thus, use of a bonded model does not appear justified except in cases where the exact geometry of an essentially fixed zinc coordination site must be preserved. Otherwise, the nonbonded model, which is simpler and physically more valid, is the preferred choice.

In a nonbonded model, and in a bonded model that includes nonbonded terms, complete specification of the model involves the nature of the truncation scheme for the interactions. This is particularly important for the electrostatic interactions of charged species like the zinc ion. Results have been given for two types of truncation. One that is simple and widely used is a shift truncation which smoothly reduces the electrostatic interactions to zero at the cutoff distance. The other is an extended electrostatics model,[53,58] which approximates the full electrostatic interaction by using a combination of the conventional pairwise additive sum for close interactions and a multipole approximation for longer range interactions. In the case of carboxypeptidase A there was only a small effect when including the long range electrostatics, given that a sufficiently long cutoff distance was used (13 Å). However, in the case of carbonic anhydrase, exclusion of the long-range contribution leads to a striking conformational change of the Glu-106 around the zinc ion in disagreement with the experimental structure, i.e., in the crystal structure the distance between the zinc ion and OE1 of this residue is 4.1 Å. In the structures obtained from simulations which excluded long-range electrostatic interactions, the OE1 oxy-

gen moved into the first solvation sphere of the zinc ion. When long-range electrostatic interactions were included via the extended electrostatics model, the distances obtained were in better agreement with the crystal structure. It should be noted that the overall charge of the zinc coordination sphere in carboxypeptidase A is 1+ (Zn^{2+}, His^0, Glu^{1-}, His^0) while in carbonic anhydrase the zinc coordination sphere has a total charge of +2 (Zn^{2+}, His^0, His^0, His^0). This is likely to be the reason why longer-range electrostatic interactions are more important in carbonic anhydrase than in carboxypeptidase A.

The present studies confirm that the Extended Electrostatic method provides an accurate treatment of the long-range forces with manageable computational requirements and show that it can be readily applied to the study of moderately sized proteins containing several thousand atoms. We expect this treatment to be particularly useful in the study of the thermodynamics of zinc binding in proteins (e.g., carbonic anhydrase) where the long-range electrostatic interactions are important. Such a study is underway.

ACKNOWLEDGMENTS

The authors thank Dr. Lucia Banci and Dr. Giovanni La Penna for providing coordinates from their simulations of carboxypeptidase A. We would like to acknowledge helpful discussions with Dr. Johan Åqvist, Dr. Lucia Banci, Dr. Carmay Lim, and Dr. Kenneth Merz, Jr. This work was supported in part by the National Science Foundation and an equipment grant from Hewlett-Packard Corporation. The National Center for Supercomputing Applications provided an allocation of CRAY Y-MP computer time which was used for this work. Additional calculations were done on SUN, Stardent, and SGI workstations. The final stages of this work were carried out in the laboratory of Prof. Jean-François Lefèvre at the Ecole Supérieure de Biotechnologie de Strasbourg (ESBS), France. His hospitality is gratefully acknowledged. Supported in part by a grant from the National Science Foundation and a gift from Molecular Simulations, Inc. R.H.S. was supported by a fellowship from the Direction de la Recherche et des Etudes Doctorales from the Ministére de l'Education National in France.

REFERENCES

1. Vallee, B.L., Auld, D.S. Functional zinc-binding motifs in enzymes and DNA-binding proteins. Faraday Discuss. 93:47–65, 1992.
2. Kannan, K.K., Ramanadham, M., Jones, T.A. Structure, refinement, and function of carbonic anhydrase isozymes: Refinement of human carbonic anhydrase I. Ann. N.Y. Acad. Sci. 429:49–60, 1984.
3. Liang, J.-Y., Lipscomb, W.N. Hydration of carbon dioxide by carbonic anhydrase: Internal proton transfer of Zn^{2+}-bound HCO_3^-. Biochemistry 26:5293–5301, 1987.
4. Liang, J.-Y., Lipscomb, W.N. Hydration of CO_2 by carbonic anhydrase: Intramolecular proton transfer between Zn^{2+}- bound H_2O and histidine 64 in human carbonic anhydrase II. Biochemistry 27:8676–8682, 1988.
5. Eriksson, E.A., Jones, T.A., Liljas, A. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. Proteins 4:274–282, 1988.
6. Quiocho, F.A., Lipscomb, W.N. Carboxypeptidase A: A protein and an enzyme. Adv. Prot. Chem. 25:1–49, 1971.
7. Lipscomb, W.N. Carboxypeptidase A mechanisms. Proc. Natl. Acad. Sci. U.S.A. 77:3875–3878, 1980.
8. Rees, D.C., Lewis, M., Lipscomb, W.N. Refined crystal structure of carboxypeptidase A at 1.54 Å resolution. J. Mol. Biol. 168:367–387, 1983.
9. Christianson, D.W., Lipscomb, W.N. Binding of a possible transition state analog to the active site of carboxypeptidase A. Proc. Natl. Acad. Sci. U.S.A. 82:6840–6844, 1985.
10. Christianson, D.W., Lipscomb, W.N. Mechanism of carboxypeptidase A: Hydration of a ketonic substrate analogue. Proc. Natl. Acad. Sci. U.S.A. 84:1512–1515, 1987.
11. Christianson, D.W., Lipscomb, W.N. Carboxypeptidase A. Acc. Chem. Res. 22:62–69, 1989.
12. Branden, C.I., Eklund, H., Nordstrom, B., Boiwe, T., Soderlund, G., Zeppezauer, E., Ohlsson, I., Åkeson, Å. Structure of liver alcohol dehydrogenase at 2.9-Ang. resolution. Proc. Natl. Acad. Sci. U.S.A. 70:2439–2442, 1973.
13. Holmes, M.A., Matthews, B.W. Structure of thermolysin refined at 1.6 Å resolution. J. Mol. Biol. 160:623–639, 1982.
14. Matthews, B.W. Structural basis of the action of thermolysin and related zinc peptidases. Acc. Chem. Res. 21:333–340, 1988.
15. Honzatko, R.B., Crawford, J.L., Monaco, H.L., Ladner, J.E., Edwards, B.F.P., Evans, D.R., Warren, S.G., Wiley, D.C., Ladner, R.C., Lipscomb, W.N. Crystal and molecular structures of native and CTP-liganded aspartate carbamoyltransferase from Escherichia coli. J. Mol. Biol. 160:219–263, 1982.
16. Keating, K.M., Ghosaini, L.R., Giedroc, D.P., Williams, K.R., Coleman, J.E., Sturtevant, J.M. Thermal denaturation of T4 gene 32 protein: Effects of zinc removal and substitution. Biochemistry 27:5240–5245, 1988.
17. Hanas, J.S., Hazuda, D.J., Bogenhagen, D.F., Wu, F.Y., Wu, C. Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene. J. Biol. Chem. 258:14120–14125, 1983.
18. Berg, J.M. Zinc finger domains, hypotheses and current knowledge. Annu. Rev. Biophys. Biophys. Chem. 19:405–421, 1990.
19. O'Halloran, T.V. Transition metals in control of gene expression. Science 261:715–725, 1993.
20. Meyers, L.C., Cushing, T.D., Wagner, G., Verdine, G.L. Metal-coordination sphere in the methylated Ada protein-DNA co-complex. Chem. Biol. 1:91–97, 1994.
21. Schwabe, J.W.R., Klug, A. Zinc mining for protein domains. Struct. Bio. 1:345–349, 1994.
22. Kleiner, D.E., Jr., Stetler-Stevenson, W.G. Structural biochemistry and activation of matrix metalloproteases. Curr. Opinion Cell Biol. 5:891–897, 1993.
23. Hough, E., Hansen, L.K., Birknes, K., Jynge, K., Hansen, S., Horvik, A., Little, C., Dodson, E., Derewenda, Z. High-resolution (1.5 Å) crystal structure of phospholipase C from Bacillus cereus. Nature (London) 338:357–360, 1989.
24. Kim, E.E., Wyckoff, H.W. Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis. J. Mol. Biol. 218:449–464, 1991.
25. Christianson, D.W. Structural biology of zinc. Adv. Protein Chem. 42:281–355, 1991.
26. Demoulin, D., Pullman, A. An ab initio theoretical study of the binding of ZnII with biologically significant ligands CO_2, H_2O, OH^-, imidazole, and imidazolate. Theoret. Chim. Acta 49:161–181, 1978.
27. Nakagawa, S., Umeyama, H. Ab initio molecular orbital study on the effects of zinc ion, its ligands and ionic amino acid residues for proton transfer energetics between Glu 270 and Zn coordinated water molecules in carboxypeptidase A. J. Theor. Biol. 96:473–493, 1982.
28. Bertini, I., Luchinat, C., Rosi, M., Sgamellotti, A., Tarantelli, F. pKa of zinc-bound water and nucleophilicity of hydroxo-containing species. Ab initio calculations on

model for zinc enzymes. Inorg. Chem. 29:1460–1463, 1990.
29. Purcell, K.F., Kotz, J.C. "Inorganic Chemistry." Philadelphia: W.B. Saunders, 1977.
30. 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., Podgers, J.R., Watson, D.G. Cambridge structural database. Acta Crystallogr. Sect. B B35:2331–2339, 1979.
31. Vendani, A., Huhta, D.W. A new force field for modeling *metallo*proteins. J. Am. Chem. Soc. 112:4759–4767, 1990.
32. Vedani, A., Dobler, M., Dunitz, J.D. An empirical potential function for metal centers: Application to molecular mechanics calculations on metalloproteins. J. Comput. Chem. 7:701–710, 1986.
33. Vedani, A., Huhta, D.W., Jacober, S.P. Metal coordination, H-bond network formation, and protein-solvent interactions in native and complexed human carbonic anhydrase I: A molecular mechanics study. J. Am. Chem. Soc. 111:4075–4081, 1989.
34. Hoops, S.C., Anderson, K.W., Merz, K.M. J. Force field design for metalloproteins. J. Am. Chem. Soc. 113:8262–8270, 1991.
35. Merz, K.M.J., Murcko, M.A., Kollman, P.A. Inhibition of carbonic anhydrase. J. Am. Chem. Soc. 113:4484–4490, 1991.
36. Lin, S., Stern, E.A., Kalb, A.J., Zhang, Y. Evidence from x-ray absorption fine structure spectroscopy for significant differences in the structure of concanavalin A in solution and in the crystal. Biochemistry 29:3599–3603, 1990.
37. Lin, S., Stern, E.A., Kalb, A.J., Zhang, Y. X-ray absorption fine structure investigation of the zinc transition metal binding site of zinc concanavalin A in solution and in the crystal. Biochemistry 30:2323–2332, 1991.
38. Zhang, K., Chance, B., Ready, K.S., Ayene, I., Stern, E.A., Bunker, G. Structural differences in solution and crystalline forms of metmyoglobin. Biochemistry 30: 9116–9120, 1991.
39. Zhang, K., Chance, B., Auld, D.S., Larsen, K.S., Vallee, B.L. X-ray absorption fine structure study of the active site of zinc and cobalt carboxypeptidase A in their solution and crystalline forms. Biochemistry 31:1159–1168, 1992.
40. Makinen, M.W., Troyer, J.N., van der Werff, H., Berendsen, J.C., van Gunsteren, W.F. Dynamical structure of carboxypeptidase A. J. Mol. Biol. 207:210–216, 1989.
41. Liang, J.-H., Lipscomb, W.N. Binding of substrate CO_2 to the active site of human carbonic anhydrase II: A molecular dynamics study. Proc. Natl. Acad. Sci. U.S.A. 87: 3675–3679, 1990.
42. Liang, J.-Y., Lipscomb, W.N. Correction to Proc. Natl. Acad. Sci. U.S.A. 87 3675–3679. Proc. Natl. Acad. Sci. U.S.A. 87:6475, 1990.
43. Thomas, P.G., Russel, A.J., Fersht, A.R. Tailoring the pH dependence of enzyme catalysis using protein engineering. Nature (London) 318:375–376, 1985.
44. Russell, A.J., Fersht, A.R. Rational modification of enzyme catalysis by engineering surface charge. Nature (London) 328:496–500, 1987.
45. Russell, A., Thomas, P.G., Fersht, A.R. Electrostatic effects on the modification of charged groups in the active site cleft of subtilisin by protein engineering. J. Mol. Biol. 193:803–813, 1987.
46. Pantoliano, M.W., Whitlow, M., Wood, J.F., Rollence, M.L., Finzel, B.C., Gilliland, G.L., Poulos, T.L., Bryan, P.N. The engineering of binding affinity at metal ion binding sites for the stabilization of proteins: Subtilisin as a test case. Biochemistry 27:8311–8317, 1988.
47. Groeneveld, C.M., Ouwerling, M.C., Erkelens, C., Canters, G.W. ^1H nuclear magnetic resonance study of the portonation behaviour of the histidine residues and the electron self-exchange reactions of azurin from Alcaligenes denitrificans. J. Mol. Biol. 200:189–199, 1988.
48. Sternberg, M.J.E., Hayes, F.R.F., Russell, A.J., Thomas, P.G., Fersht, A.R. Prediction of electrostatic effects of engineering of protein charges. Nature (London) 330:86–88, 1987.
49. Gilson, M.K., Honig, B.H. Calculation of electrostatic potentials in an enzyme active site. Nature (London) 330: 84–86, 1987.
50. Bashford, D., Karplus, M., Canters, G.W. Electrostatic effects of charge perturbations introduced by metal oxidation in proteins. A theoretical analysis. J. Mol. Biol. 203:507–510, 1988.
51. Vedani, A., Dunitz, J.D. Lone-pair directionality in hydrogen bond potential functions for molecular mechanics calculations: The inhibition of human carbonic anhydrase II by sulfonamides. J. Am. Chem. Soc. 107:7653–7658, 1985.
52. Dunfield, L.G., Burgess, A.W., Scheraga, H.A. Energy parameters in polypeptides. 8. Empirical potential energy algorithm for the conformational analysis of large molecules. J. Phys. Chem. 82:2609–2616, 1978.
53. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy minimization and dynamics calculations. J. Comp. Chem. 4:187–217, 1983.
54. Hagler, A.T., Huler, E., Lifson, S. Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals. J. Am. Chem. Soc. 96:5319–5327, 1974.
55. Reiher, W.E., III. "Theoretical Studies of Hydrogen Bonding." Cambridge, MA: Harvard University, 1985.
56. Jorgensen, W.L., Tirado-Rives, J. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. J. Am. Chem. Soc. 110: 1657–1666, 1988.
57. MacKerell, A., Karplus, M. To be published. 1995.
58. Stote, R.H., States, D.J., Karplus, M. On the treatment of electrostatic interactions in biomolecular simulation. J. Chim. Phys. 88:2419–2433, 1991.
59. Åqvist, J., Warshel, A. Computer simulation of the initial proton transfer step in human carbonic anhydrase I. J. Mol. Biol. 224:7–14, 1992.
60. Åqvist, J. Ion-water interaction potentials derived from free energy perturbation simulations. J. Phys. Chem. 94: 8021–8024, 1990.
61. Demoulin, D., Pullman, A. Theoretical study of binding and proton-labilizing properties of zinc(2+) ion. J. Am. Chem. Soc. 99:8498–8500, 1977.
62. Clementi, E., Corongiu, G., Jonsson, B., Romano, S. Monte Carlo simulations of water clusters around Zn^{++} and a linear Zn^{++} CO_2 complex. J. Chem. Phys. 72:260–263, 1980.
63. Yongyai, Y.P., Kokpol, S., Rode, B.M. Zinc ion in water: intermolecular potential with approximate three-body correction and Monte Carlo simulation. Chem. Phys. 156: 403–412, 1991.
64. Allen, M.P., Tildesley, D.J. "Computer Simulation of Liquids." Oxford: Clarendon Press, 1989.
65. Jorgensen, W. BOSS V. 3, Yale University, New Haven, CT.
66. Gao, J., Karplus, M. Unpublished calculations.
67. Brooks, C.L. III, Karplus, M., Pettitt, B.M. "Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics." Adv. Chem. Phys. New York: John Wiley, 1988.
68. Beveridge, D.L., DiCapua, F.M. Free energy via molecular simulation: Applications to chemical and biomolecular systems. Annu. Rev. Biophys. Biophys. Chem. 18: 431–492, 1989.
69. Bash, P.A., Singh, U.C., Langridge, R., Kollman, P.A. Free energy calculations by computer simulation. Science 236:564–568, 1987.
70. Jorgensen, W.L., Madura, J.D., Swenson, C. Optimized intermolecular potential functions for liquid hydrocarbons. J. Am. Chem. Soc. 106:6638–6646, 1984.
71. Alper, H.E., Levy, R.M. Computer simulations of the dielectric properties of water: Studies of the simple point charge and transferrable intermolecular potential models. J. Chem. Phys. 91:1242–1251, 1989.
72. Bernstein, F.C., Koetzle, T.F., Willams, G.J.B., Meyer, E.F.J., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., Tasumi, M. The Protein Data Bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
73. Abola, E., Berstein, F.C., Bryant, S.H., Koetzle, T.F., Weng, J. Protein Data Bank. In: "Crystallographic Data-

bases—Information Content, Software Systems, Scientific Applications." Bonn: Data Commission of the International Union of Crystallography, 1987: 107–132.
74. Brunger, A.T., Karplus, M. Polar hydrogen positions in proteins: Empirical energy placement and neutron diffraction comparison. Proteins 4:148–156, 1988.
75. Gelin, B., Karplus, M. Mechanism of tertiary structural change in hemoglobin. Proc. Natl. Acad. Sci. U.S.A. 74: 801–805, 1977.
76. Whitlow, M., Teeter, M.M. An empirical examination of potential-energy minimization using the well determined structure of the protein crambin. J. Am. Chem. Soc. 108: 7163–7172, 1986.
77. Brooks, C.L. III, Brünger, A.T., Karplus, M. Active site dynamics in protein molecules: A stochastic boundary molecular-dynamics approach. Biopolymers 24:843–865, 1985.
78. Brooks, C.L. III, Karplus, M. Solvent effects on protein motion and protein effects on solvent motion. J. Mol. Biol. 208:159–181, 1989.
79. Berendsen, H.J.C., Van Gunsteren, W.F., Zwinderman, H.R.J., Guertsen, R.G. Simulations of proteins in water. Ann. N.Y. Acad. Sci. 482:269–286, 1986.
80. Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. J. Comp. Phys. 23:327–341, 1977.
81. Straub, J.E., Lim, C., Karplus, M. Simulation analysis of the binding interactions in the RNase A/3′-UMP enzyme-product complex as a function of pH. J. Am. Chem. Soc. 116:2591–2599, 1994.
82. Ohtaki, H., Yamaguchi, T., Maeda, M. X-ray diffraction studies of the structures of hydrated divalent transition-metal ions in aqueous solution. Bull. Chem. Soc. Jpn. 49:701–708, 1976.
83. Johansson, G. Structure of complexes in solution derived from X-ray diffraction measurements. Adv. Inorg. Chem. 39:159–232, 1992.
84. Brooks, C.L. III, Pettitt, B.M., Karplus, M. Structural and energetic effects of truncating long ranged interactions in ionic and polar fluids. J. Chem. Phys. 83:5897–5908, 1985.
85. Gmelin Institute, Handbuch der Anorganischen Chemie 8 Auflage. Berlin: Springer, Ger. 720, 1956.
86. Pogue, R.F., Atkinson, G. Solution thermodynamics of first row transition elements. 3. Apparent molal volumes of aqueous $ZnCl_2$ and $Zn(ClO_4)_2$ from 15 to 55 C and an examination of solute-solute and solute-solvent interactions. J. Soln. Chem. 18:249–264, 1989.
87. Åqvist, J. Personal communication.
88. Berg, J.M., Merkle, D.L. On the metal ion specificity of zinc finger proteins. J. Am. Chem. Soc. 111:3759–3761, 1989.
89. Laurenczy, G., Ducommun, Y., Merbach, A.E. Variable-pressure kinetic and equilibrium study of monocomplex formation of copper(II) and zinc (II) with 2-chloro-1,10-phenanthroline in aqueous solution. Inorg. Chem. 28: 3024–3028, 1989.
90. Post, C.B., Brooks, B.R., Dobson, C.M., Artymiuk, P.J.C., Phillips, D.C., Karplus, M. Molecular dynamics simulations of native and substrate-bound lysozyme. A study of the average structures and atomic fluctuations. J. Mol. Biol. 190:455–479, 1986.
91. Kuczera, K., Kuriyan, J., Karplus, M. Temperature dependence of the structure and dynamics of myoglobin. A simulation approach. J. Mol. Biol. 213:351–373, 1990.
92. Loncharich, R.J., Brooks, B.R. Temperature dependence of dynamics of hydrated myoglobin. Comparison of force field calculations with neutron scattering data. J. Mol. Biol. 215:439–455, 1990.
93. Hartmann, H., Parak, F., Steigemann, W., Petsko, G.A., Ringe Ponzi, D., Fraunfelder, H. Conformation substates in a protein: Structure and dynamics of metmyoglobin at 80 K. Proc. Natl. Acad. Sci. U.S.A. 79:4967–4971, 1982.
94. Banci, L., Schroder, S., Kollman, P.A. Molecular dynamics characterization of the active cavity of carboxypeptidase A and some of its inhibitor adducts. Proteins 13:288–305, 1992.
95. Banci, L., Bertini, I., La Penna, G. A molecular dynamics study of carboxypeptidase A: Effect of protonation of Glu 270. Inorg. Chem. 32:2207–2211, 1993.
96. Schaefer, M., Karplus, M. To be published.
97. Merz, K.M., Jr. Personal communication.
98. Yu, H.-A., Karplus, M., Hendrickson, W.A. Restraints in temperature-factor refinement for macromolecules: An evaluation by molecular dynamics. Acta Cryst. B 41:191–201, 1985.
99. Karplus, M. Dynamic aspects of protein structure. Ann. N.Y. Acad. Sci. 439:107–123, 1985.
100. Tashian, R.E. The carbonic anhydrases: Widening perspectives on their evolution, expression and function. Bioessays 10:186–192, 1989.
101. Merz, K.M., Jr. Insights into the function of the zinc hydroxide-Thr199-Glu106 hydrogen bonding network in carbonic anhydrases. J. Mol. Biol. 214:799–802, 1990.
102. Xue, Y., Liljas, A., Jonsson, B.-H., Lindskog, S. Structural analysis of the zinc hydroxide-Thr-199-Glu-106 hydrogen-bond network in human carbonic anhydrase II. Proteins 17:93–106, 1993.
103. Nakagawa, S., Yu, H.-A., Karplus, M., Umeyama, H. Active site dynamics of acyl-chymotrypsin. Proteins 16: 172–194, 1993.
104. Migliore, M., Corongiu, G., Clementi, E., Lie, G.C. Monte Carlo study of free energy of hydration for lithium(1^{+1}), sodium(1^{+}), potassium(1^{+}), fluoride, and chloride ions with ab initio potentials. J. Chem. Phys. 88:7766–7771, 1988.
105. Straatsma, T.P., Berendsen, H.J.C. Free energy of ionic hydration: Analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics simulations. J. Chem. Phys. 89:5876–5886, 1988.
106. Roux, B., Karplus, M. Ion transport in a gramicidin channel: Structure and thermodynamics. Biophys. J. 59:961–981, 1991.
107. Roux, B. Nonadditivity in cation-peptide interactions: A molecular dynamics and ab initio study of Na^+ in the gramicidin channel. Chem. Phys. Lett. 212:231–240, 1993.