Studying How Different Terminal Groups Change the Motion of $H_2NSO_2C_6H_4CONH(EG)_3R$ When Bound to the Active Site of Human Carbonic Anhydrase II

Donovan N. Chin,† Albert Y. Lau, and George M. Whitesides*

Department of Chemistry, and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received March 18, 1997

Molecular dynamics simulations have been used to explore the motions of a series of ligands containing coupled benzenesulfonamide and oligoethylene glycol moieties ($H_2NSO_2C_6H_4CONH(CH_2CH_2OCH_2CH_2OCH_2CH_2)\mathbf{R}^+$; $\mathbf{R}^+ = NH_3^+$, $NHCOCH_2NH_3^+$, $NHCOCH(CH_2Ph)NH_3^+$) bound at the active site of human carbonic anhydrase II (HCAII; E.C. 4.2.1.1). These complexes have been examined previously by X-ray crystallography; the locations of the terminal groups of these ligands were not defined in the crystal structures. These simulations, carried out in the presence of water, provide qualitative dynamic information about the motion of the bound ligand that supplements the quasistatic information from crystallography. Our results suggested that the Gly and Phe groups of these ligands interacted weakly with the protein adjacent to the active site. Quantitative estimates of energies of binding did not correlate usefully with observed free energies of binding, but in the absence of information about entropies, it is not possible to tell if the lack of correlation between calculated energies and observed free energies represents inaccuracies in the energies, or a compensation between enthalpies and entropies. When the terminal Phe group was placed near a previously identified hydrophobic patch in the active site (Phe20 and Pro202) the average conformation of the ligand inferred from this simulation was inconsistent with that from the crystal structure; this result illustrates the problems of misleading local minima in these types of simulations.

Introduction

Oligo(ethylene glycol) linking groups are useful in medicinal and biochemistry for several reasons:[1] they are biocompatible; they are not degraded by peptidases; they increase the solubility of hydrophobic species in water; they are conformationally flexible, and can conform well to the surface of a protein; and they are inexpensive and easily modified synthetically. We report here a study, using stochastic boundary molecular dynamics (SBMD) simulations, of the dynamics of interaction of several triethylene glycol-terminated *p*-benzenesulfonamide inhibitors—$H_2NSO_2C_6H_4CONH-CH_2CH_2OCH_2CH_2OCH_2CH_2\mathbf{R}^+$ ($S(EG)_3\mathbf{R}^+$; $\mathbf{R}^+ = NH_3$, $NHCOCH_2NH_3^+$, $NHCOCH(CH_2Ph)NH_3^+$)—with the active site of human carbonic anhydrase II (HCAII; E.C. 4.2.1.1) in the presence of water. This work focuses on three molecules (Table 1): the parent compound, $S(EG)_3NH_3^+$, and two derivative compounds, $S(EG)_3GlyNH_3^+$ and $S(EG)_3PheNH_3^+$ (S = benzenesulfonamide); these two latter compounds have the highest affinity for HCAII of this series of inhibitors (Table 1), but all three have similar free energies of binding.[2] The three compounds differ only (but substantially) in the terminal group on the triethylene glycol moiety.

We and others have used HCAII as a model system with which to study the physical−organic chemistry of interactions of ligands with a representative protein surface.[3−8] HCAII has 259 amino acids and a structurally well-defined, conical active site that is roughly 15 Å deep. The active site is characterized by one hydrophilic and one hydrophobic wall. A zinc ion, bound to three histidines, is located at the bottom of this active site. This zinc atom is responsible for catalyzing the hydration of

Table 1. Values of K_d for Several Oligo(ethylene glycol)-Based Inhibitors; $H_2NSO_2C_6H_4Conhch_2CH_2OCH_2CH_2OCH_2CH_2R^+$. Molecules Studied Here Are in Bold

$\mathbf{R}^+$	abbrev	K_d[a] (nM)
$-NHCOCH(CH_2Ph)NH_3^+$	**$S(EG)_3PheNH_3^+$**	**14**
$-NHCOCH(CH_2CH(CH_3))NH_3^+$	$S(EG)_3LeuNH_3^+$	16
$-NHCOCH_2NH_3^+$	**$S(EG)_3GlyNH_3^+$**	**19**
$-NHCOCH(CH_2OH)NH_3^+$	$S(EG)_3SerNH_3^+$	41
$-NH_3^+$	**$S(EG)_3NH_3^+$**	**43**
$-NHCOCH((CH_2)_4NH_3^+)NH_3^+$	$S(EG)_3LysNH_3^+$	50
$-NHCOCH(CH_2CH_2COO^-)NH_3^+$	$S(EG)_3GluNH_3^+$	100

[a] Values taken from ref 2; we estimate that they are accurate to within ±20%.

† Moldyn, Inc., 955 Massachusetts Ave., Cambridge, MA 02139.

(1) Harris, J. M. *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biochemical Applications*; Harris, J. M., Ed.; Plenum Press: New York, 1992.

(2) Boriak, A.; Christainson, D.; Kingery-Wood, J.; Whitesides, G. M. *J. Med. Chem.* **1995**, *38*, 2286−2291.

(3) Avila, L. Z.; Chu, Y. H.; Blossey, E. C.; Whitesides, G. M. *J. Med. Chem.* **1993**, *36*, 126.

(4) Jain, A.; Huang, S. G.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 5057.

(5) Jain, A.; Whitesides, G. M.; Alexander, R. S.; Christianson, D. W. *J. Med. Chem.* **1994**, *37*, 2100.

(6) Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murko, M. A.; Randall, W. C.; Schwam, H.; Sugrue, M. F.; Springer, J. P.; Gautheron, P. *J. Med. Chem.* **1989**, *32*, 2510.

(7) Greer, J.; Erickson, J. W.; Baldwin, J. J.; Varney, M. D. *J. Med. Chem.* **1994**, *37*, 1035.

(8) Alexander, R. S.; Nair, S. K.; Christianson, D. W. *Biochemistry* **1991**, *30*, 17320.

(9) Dodgson, S. J.; Tashian, R. E.; Gros, G.; Carter, N. D. *The Carbonic Anhydrase: Cellular Physiology and Molecular Genetics*; Plenum Press: New York, 1991.

S0022-3263(97)00493-3 CCC: $15.00
Published on Web 01/30/1998

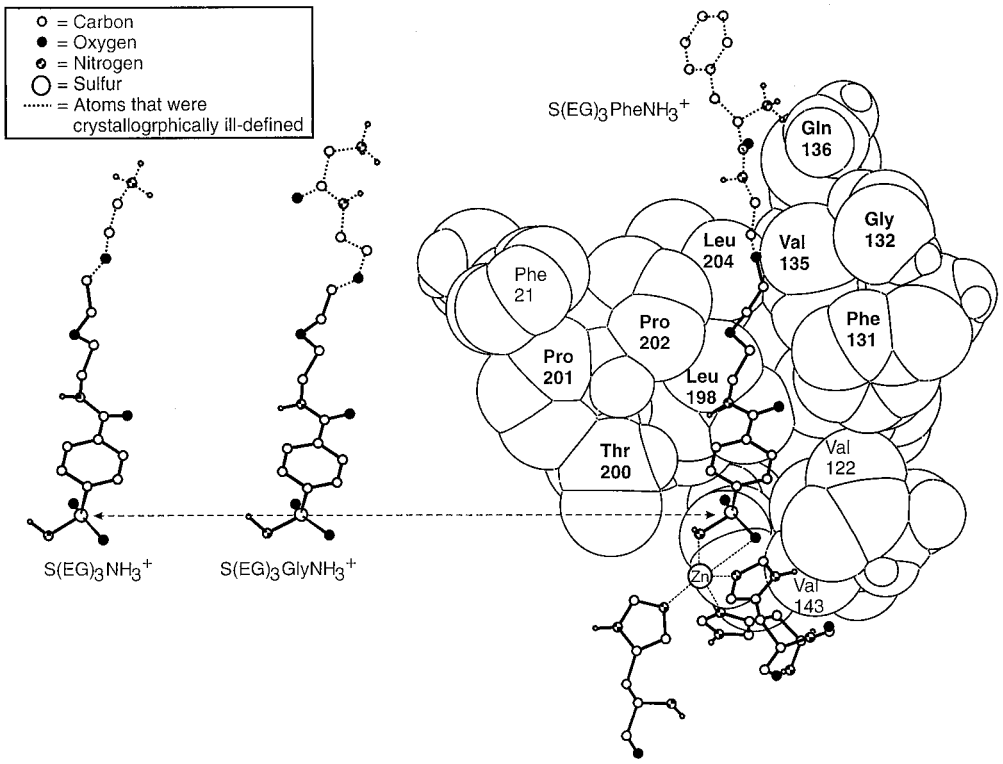

Figure 1. Partial view of $S(EG)_3PheNH_3^+$ bound to the hydrophobic wall of HCAII. The inhibitors $S(EG)_3NH_3^+$ and $S(EG)_3GlyNH_3^+$ are similarly bound and are shown without HCAII for clarity—the perspective for $S(EG)_3NH_3^+$ and $S(EG)_3GlyNH_3^+$ is the same as that for $S(EG)_3PheNH_3^+$.

carbon dioxide and is the center to which the sulfonamide group binds.[9] The location of the benzene sulfonamide group in the active site does not change with changes in the *para* substituent.[10]

Christianson and co-workers have determined the crystal structure of these three $S(EG)_3\mathbf{R}^+$ ligands bound to the active site of HCAII, and we have determined their dissociation constants (Table 1).[2] The crystallographic data establish that, for all three ligands, the $(EG)_3$ tails associate with the hydrophobic wall (Figure 1); the terminal glycol groups do not appear in the electron difference map (that is, they are crystallographically undefined) and, therefore, are less localized in their binding than other groups in the inhibitor. Analyses of the values of T_2 relaxation times from ^1H NMR spectroscopy show that the three EG groups closest to the aryl group interact sufficiently strongly with the active site of bovine carbonic anhydrase (BCA) to restrict their motion but that EG units in tails longer than $(EG)_3$ do not interact significantly when bound to the active site.[4]

In a previous study, we used molecular dynamics to suggest the role of a similarly disordered terminal group on a (benzyloxy)oligoglycine terminated *p*-benzene sulfonamide inhibitor ($SGly_3Bn$) bound to the active site of HCAII.[11] The oligoglycine tail of $SGly_3Bn$ aligns itself along the hydrophobic wall in a way similar to that of the $(EG)_3$ tails. The simulations suggested that weak hydrophobic interactions between the disordered terminal −Bn group on $SGly_3Bn$ and a hydrophobic cleft (defined by Phe20 and Pro202) located at the lip of the active site contributed to tighter binding (lower value of K_d) than that of an oligoglycine inhibitor terminating in a carboxylate group. These simulations also showed that the range of conformations of the $SGly_3Bn$ inhibitor, when bound to HCAII, was substantially greater than that inferred from the crystal structure of the $SGly_3Bn$–HCAII complex. Here, we have used similar molecular dynamics methods to study the motion of the EG groups of the inhibitors when bound to HCAII and to help understand the influence of the different terminal groups on the interaction of these ligands with the protein.

An important question emerging from our studies of the binding between CA and sulfonamide inhibitors with tails based on EG and Gly is the following. Why does the change in the length of the tail (composed of either $(EG)_n$ or $(Gly)_n$, where $n = 1-3$), add nothing to the ΔG of binding for these inhibitors to CA?[4] The answer is probably a compensation between the enthalpy and entropy of binding.[4]

Table 1 indicates that terminal amino acid groups that are hydrophobic had marginally lower values of K_d than those that are hydrophilic. We wished to understand how three terminal groups with substantial differences in hydrophobicity but similar binding constants interacted with the protein, when the interaction was too weak to localize the terminal groups crystallographically. We hoped these studies might suggest new strategies for modifying the terminal group in the "tail" of the ligand to increase binding.

Model and Computational Methodology

Each crystal structure served as the starting geometry for its simulation. The crystal structures of the complexes of HCAII with the three ligands were kindly provided to us by Professor David Christianson.[2] The terminal glycol and functional units of the ligands not located in the crystal

(10) Cappalonga, A. M.; Alexander, R. S.; Christianson, D. W.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 5063.
(11) Chin, D. N.; Whitesides, G. M. *J. Am. Chem. Soc.* **1995**, *117*, 6153.

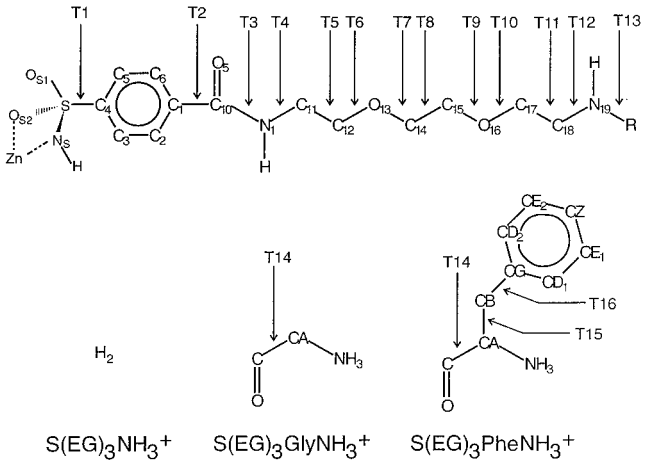

Figure 2. Nomenclature used for the atomic names, and the torsions (T#) for the molecules of $S(EG)_3\mathbf{R}^+$.

structure were added using standard valence geometries from the QUANTA modeling program.[12] All computations involved the CHARMm 22 program and the QUANTA 3.3 parameter set.[12,13] Modifications to this parameter set involved a new charge distribution, and harmonic constraints, at the $-SO_2$-$NHZn(His)_3-$ binding site.[11] These modifications were necessary to maintain the correct pentacoordination at the Zn II ion. Nonbonded interactions were cut off at 15 Å in the computations, and a switching potential between 11 and 14 Å was used for van der Waals and electrostatic terms (with the dielectric constant equal to 1).[13] Hydrogens on polar or charged atoms were added to the crystal structure using HBUILD,[14] and an extended atom representation was used for all other nonpolar hydrogenated groups.[13] The TIP3P model of water described the solvent.[15]

Solvating and Relaxing the Active Site Energetically. To make the simulation more realistic, we overlaid a pre-equilibrated sphere (20 Å radius) of bulk TIP3P waters centered on the C_{12} carbon atom of each inhibitor (Figures 2 and 3). Molecules of added water that were within 2.8 Å of any atoms of the protein, ligand, or crystallographically defined waters were deleted. The geometries of all atoms in the protein and ligand were fixed, and a brief Langevin dynamics simulation (0.1 ps) at 300 K was carried out to readjust the molecules of water to the presence of the enzyme and ligand.[16] To avoid creating any artificial "holes" in this initial shell of solvating water molecules, we repeated the process of overlaying and deleting molecules of bulk water three times, where each process was followed by the short Langevin dynamics "shake-up." The geometry of the system after the final short Langevin dynamics run was used for further computational procedures.

Stochastic Boundary Molecular Dynamics. An efficient way of reducing the amount of computation associated with *solvated* biological systems is to use stochastic boundary molecular dynamics (SBMD).[17−19] We have previously detailed the use of SBMD for similar oligoglycine-terminated inhibitors

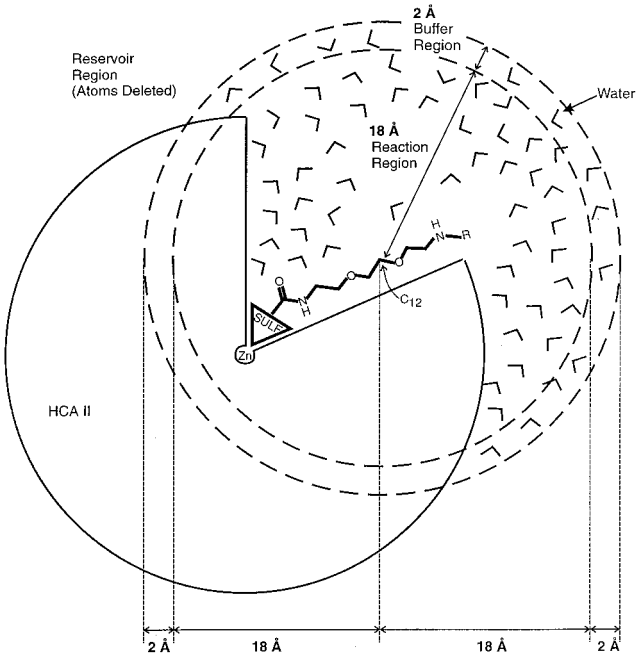

Figure 3. Description of the solvated system and the partitioning used in SBMD. The center of the SBMD method was located on the atom labeled C_{12}.

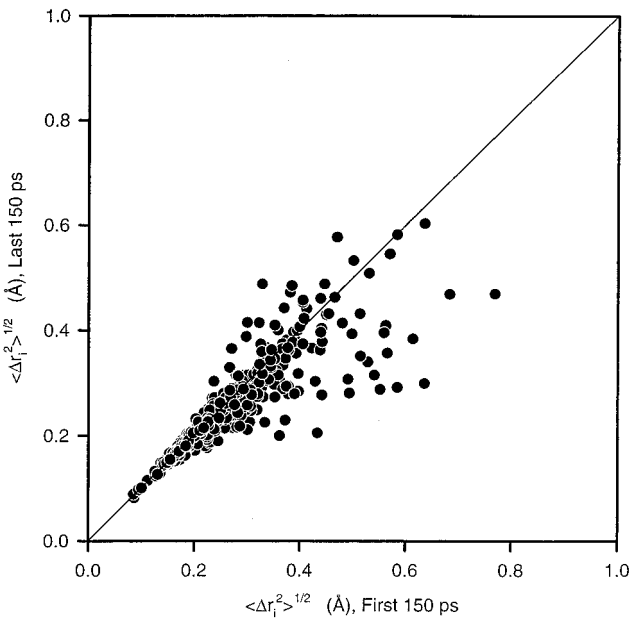

Figure 4. Comparison of the rms fluctuations of amino acids from the first 150 ps and the last 150 ps simulations.

bound to HCAII[11] and will, therefore, only highlight the pertinent details here. In SBMD, atoms that did not directly affect the spherical region of interest were deleted,[20] and the remaining atoms were partitioned into two concentric spherical regions: an outer region referred to as the buffer region and an inner region referred to as the reaction region (Figure 3). In the buffer region, the influences of the deleted atoms were simulated by mean and stochastic forces and were added to the forces caused by atoms explicitly present in this region. In the inner region, the atoms were propagated only by the forces acting on them due to the other atoms present.

Defining the System for SBMD. The reaction region, which had a radius of 18 Å, was centered on the C_{12} atom in

(12) *QUANTA 3.3 Parameter Handbook*; MSI, USA: Burlington, MA, 1992.

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

(14) Brunger, A. T.; Karplus, M. *Proteins: Struct. Funct. Genet.* **1988**, *4*, 148.

(15) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 927.

(16) The collisional frequency applied to the oxygen atoms in the molecules of water was 62−1. For a discussion of Langevin dynamics: Brooks, C. L. Karplus, M. *Biopolymers.*

(17) Brooks, C. L.; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843.

(18) Brooks, C. L.; Karplus, M. *J. Mol. Biol.* **1993**, *16*, 172.

(19) Nakagawa, S.; Yu, H.; Karplus, M. *Proteins: Struct. Funct. Genet.* **1985**, *24*, 843.

(20) That is, if the peptide group was outside the buffer region, then the entire amino acid was deleted.

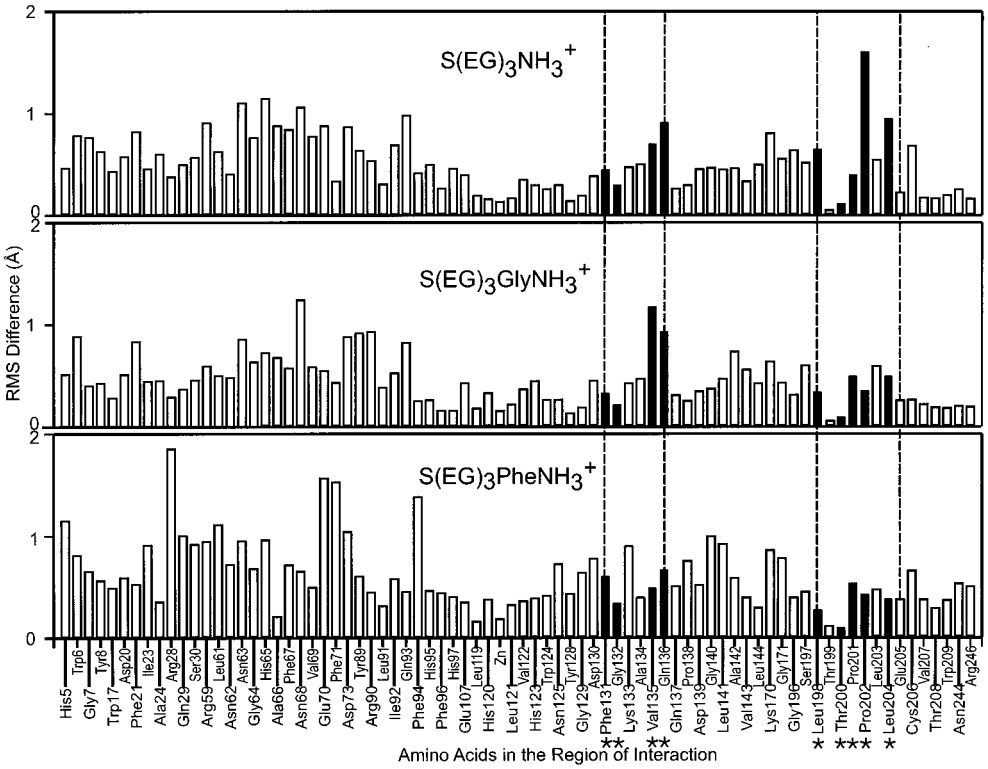

Figure 5. Rms difference between amino acids in the reaction region for the average structure from the simulation and the structure from the crystal. Amino acids highlighted by filled columns are in labeled in bold in Figure 1.

the inhibitors. A buffer region of 2 Å thick surrounded the reaction region (Figure 3). The Verlet algorithm, using a time step of 1 fs, was used to propagate atoms in the reaction region; mean and stochastic boundary forces were applied to non-hydrogen atoms in the buffer region. The mean boundary forces were of two types: those acting on the protein and those acting on the solvent.

Boundary Forces on the Protein. The geometries of the amino acids in the buffer region required harmonic restoring forces to be consistent with those in the crystal structure. We applied these harmonic forces to each non-hydrogen atom in the buffer region. The force constants used were calculated from the average isotropic Debye−Waller temperature factor (B_{iso}) for the various amino acids ($B_{iso} = 5$ Å^2 for sulfur atoms and 9 Å^2 for all other side and main-chain atoms) in the crystal structure.[11]

Boundary Forces on the Solvent. We applied a radially distributed potential to oxygen atoms in the molecules of water.[21] This potential helps to keep the structure of the molecules of water in the simulation close to that of molecules of water in the bulk. While only those molecules of water in the buffer region were directly affected by this potential, molecules of water were allowed to freely diffuse across the boundary between the buffer and reaction regions.

Stochastic Forces. A Langevin dynamics algorithm was used to control the rate of exchange of thermal energy in the system.[22] This algorithm was only applied to the non-hydrogen atoms in the buffer region. The essential parameter in Langevin dynamics is the collisional frequency. Values of the collisional frequencies used were 100 and 62 ps^{-1} for non-hydrogen atoms in the protein and molecules of water, respectively.[17]

The Length of the Simulations, 300 ps, Was Adequate for These Studies. Each simulation of the protein−ligand−water complex was equilibrated for 50 ps at 298 K and then analyzed over a period of 300 ps (the coordinates were saved every 0.1 ps). We used two metrics to check for convergence

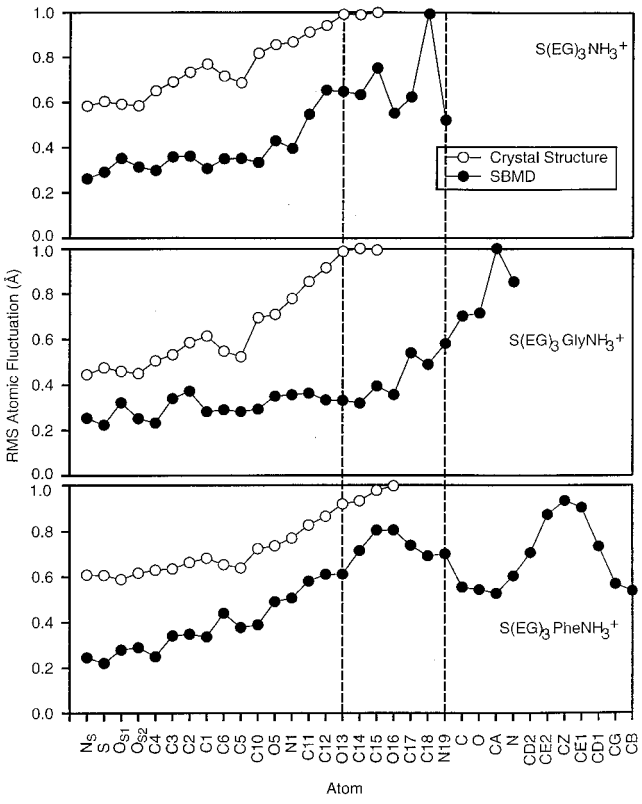

Figure 6. Rms atomic fluctuations for atoms in the inhibitors (nomenclature from Figure 2). The values were normalized to the largest value from each series.

(21) Brooks, C. L.; Karplus, M. J. *J. Chem. Phys.* **1983**, *79*, 6312.
(22) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: a theoretical perspective of dynamics, structure, and thermodynamics*; Wiley: New York, 1988; Vol. 71.

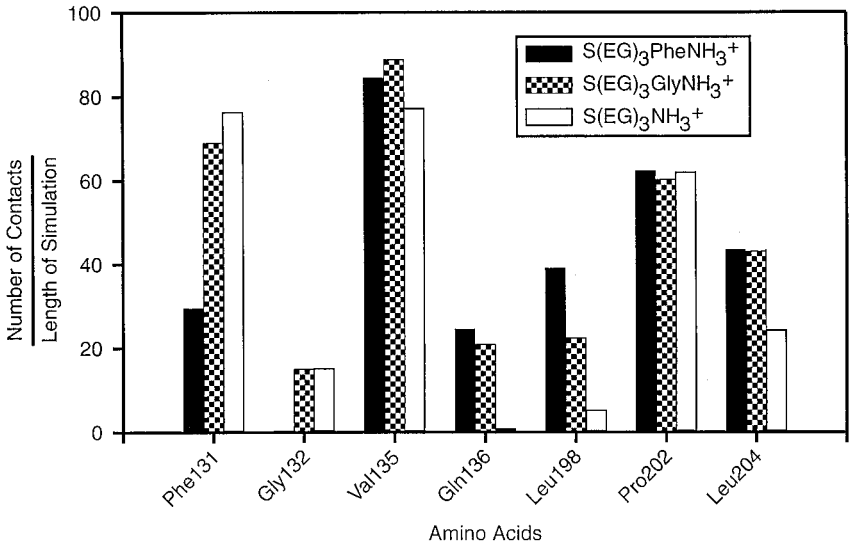

Figure 7. Average number of contacts for atoms in the $-(EG)_3R$ tails of the inhibitors with the amino acids in the active site of HCAII.

of structural properties in the simulation.[23,24] First, we compared the values of the root-mean-square (rms) fluctuations of the *amino acids* from the first 150 ps to those from the last 150 ps (a representative plot from the simulation of $S(EG)_3NH_3^+$·HCAII is shown in Figure 4). The comparison in Figure 4 shows that most of the values of the rms for amino acids were similar in the two halves of the simulations, and some differed by only 0.2 Å. Second, the isomerization rates for torsions in the inhibitors (other than those involving peptide bonds) all converged within 50 ps.[25]

Results and Discussion

The structure of the amino acids in the reaction region of HCAII during the simulation was similar to that from the crystal structure. As a measure of the accuracy of the simulation, we compared the rms *difference* between positions of the amino acids in the reaction region of the average structure of HCAII during the simulation and those from the crystal structure (Figure 5).[26] Amino acids that compose the hydrophobic wall have filled columns in Figure 5. These differences were, for the most part, small. The largest, but not unreasonable, differences in Figure 5 were for amino acids on the surface of HCAII; these differences resulted from a simulation that represented a protein molecule in solution and not in a crystalline array. In other words, the proximity of other proteins in the crystal may have dampened the fluctuations of amino acids on their surfaces.

Mobility of Atoms in the Inhibitors. Figure 6 shows the rms fluctuations for non-hydrogen atoms in the inhibitors. The values of the rms fluctuations for the crystal were computed from their isotropic temperature factors[27] and were normalized to the largest value for each inhibitor. The values of the rms fluctuations from the simulation were also normalized by the largest respective value in each inhibitor. These data showed that fluctuations of the atoms in the terminal groups were consistent with the increased disorder observed for these groups in the crystal: values of the fluctuations were greatest for atoms whose electron densities were not well defined in the Fourier difference map. The terminal $-CONH-$ group of $S(EG)_3PheNH_3^+$ was less mobile than adjacent atoms; the simulations suggested that weak electrostatic interactions between this peptide group and the underlying Gln136 amino acid were responsible for this decreased mobility (as we will show). The simulation also suggests that the $-CH_2C_6H_5$ group of $S(EG)_3PheNH_3^+$ conformationally restricts the terminal amide group and directs it toward Gln136; $S(EG)_3GlyNH_3^+$ lacks this restriction in conformation, and therefore, its terminal peptide group is less oriented toward the underlying amino acids in the active site than those of the other inhibitors.[28]

The Type of Interaction between the $-(EG)_3R$ Tails and Active Site of HCAII Depends on the Type of Terminal R Group. To assess the hydrophobic effect of the inhibitors in the simulations, we calculated the number of contacts between hydrophobic atoms in the $-(EG)_3R$ tails and hydrophobic atoms in the amino acids exposed on the surface of the active site of HCAII over the length of the simulations (300 ps). We defined a "contact" between hydrophobic carbons in the following way: two carbon atoms were in contact if one atom in the $-(EG)_3R$ tail was within 5.0 Å of one atom in the proximate surface of HCAII. The value of 5.0 Å was chosen based on the following: first, it is roughly the distance around a hydrophobic atom that represents the maximum probability of finding a water oxygen;[29] second, it represents the maximum distance that gives a significant favorable free energy of interaction between two molecules of methane.[30,31] The total number of contacts

(23) Clarage, J. B.; Romo, T.; Andrews, B. K.; Pettitt, B. M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3288−3292.
(24) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, U.K., 1987.
(25) We calculated the isomerization rate by counting the number of times a torsion crossed a barrier into a neighboring well and then divided this number by the length of the simulation.
(26) We did not include hydrogen atoms in the calculation of the rms difference in atomic positions.
(27) Willis, B. T. M.; Pryor, A. W. *Thermal Vibrations in Crystallography*; Cambridge University Press: Cambridge, 1975.
(28) We showed that, for analogous compounds based on oligoglycines, the plane of the $-CONH-$ group in the glycines preferred to pack against the hydrophobic wall in the active site of HCAII.
(29) Jorgensen, W. L.; Gao, J.; Ravimohan, C. *J. Phys. Chem.* **1985**, *89*, 3470.
(30) Jorgensen, W. L.; Gao, J.; Ravimohan, C. *Chemtracts: Org. Chem.* **1991**, *4*, 91.
(31) Blokzijl, W.; Engberts, J. B. F. N. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1545.

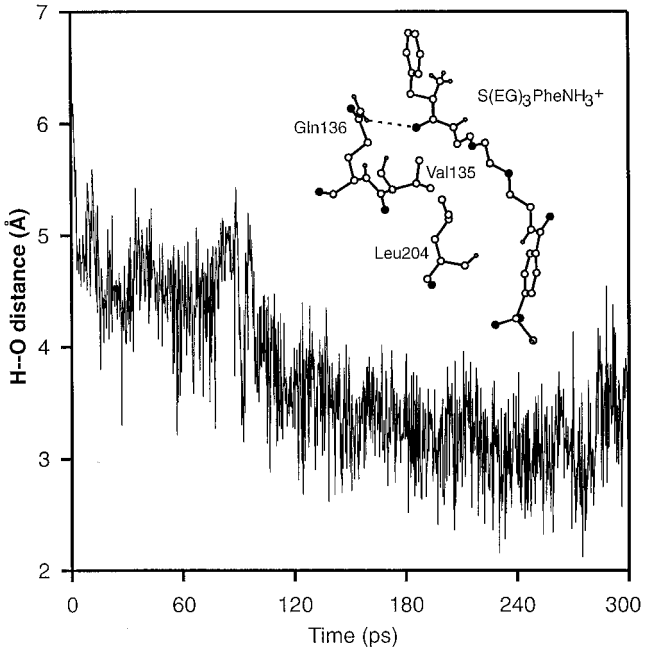

Figure 8. Interatomic distance between the carbonyl oxygen of the Phe group of $S(EG)_3PheNH_3^+$ and a hydrogen on Nϵ2 of Gln136. The inset is the average structure for this interaction from the last 50 ps of the 300 ps simulation.

for each amino acid were averaged over the length of the simulation (Figure 7). These data suggest that the $-(EG)_3R$ tails of $S(EG)_3GlyNH_3^+$ and $S(EG)_3PheNH_3^+$ have similar values of their average contacts with the hydrophobic wall of HCAII, with both values being slightly greater than that of $S(EG)_3NH_3^+$. Figure 7 suggests that these contacts are distributed over different amino acids for each inhibitor.

Where the hydrophobic atoms in the $-(EG)_3R$ tails contacted the amino acids in the active site depended on the type of terminal group on the inhibitor. For example, Gln136, which resides at the lip of the active site, is in contact with an inhibitor terminated with either Phe or Gly. The carbonyl oxygen of the terminal Phe group in $S(EG)_3PheNH_3^+$ had a significant electrostatic interaction with an Nϵ2 hydrogen of Gln136 (Figure 8). The initial orientation of the Nϵ2 group of Gln136 in the crystal structure was away from the carbonyl oxygen of the terminal Phe group in $S(EG)_3PheNH_3^+$. The initial torsional value of the Nϵ2 group in Gln136, defined about the C_γ and C_δ atoms, was 20°; this value changed to an average value of 109° over the last 50 ps (insert in Figure 8). The terminal amine of $S(EG)_3GlyNH_3^+$ interacted moderately with the carbonyl oxygen of Gly132, but this interaction was less pronounced than the electrostatic interaction involving the carbonyl oxygen of the terminal Phe group in $S(EG)_3PheNH_3^+$ and Gln136. In contrast, the terminal NH_3^+ of $S(EG)_3NH_3^+$ was conformationally more flexible than inhibitors terminated with either Gly or Phe (as will be shown) and did not make any significant electrostatic contacts with the amino acids in the active site of HCAII. These data suggest that weak electrostatic interactions between the terminal groups of $S(EG)_3GlyNH_3^+$ or $S(EG)_3PheNH_3^+$ and the proximate amino acids in the active site might be a contributing factor to the lower values of K_d than that of $S(EG)_3NH_3^+$.

(32) Holloway et al. *J. Med. Chem.* **1995**, *38*, 236−240.

Table 2. Average Interaction Energies between the Ligands and HCAII (<...> Denotes the Average)

ligand	⟨IE⟩ (kcal/mol)	⟨vdW⟩ (kcal/mol)	⟨Elec⟩ (kcal/mol)	K_d (nM)	$\Delta\Delta G_b$[a] (kcal/mol)
$S(EG)_3PheNH_3^+$	−95.8	−30.4	−65.4	14	−0.66
$S(EG)_3GlyNH_3^+$	−110.1	−29.3	−80.8	19	−0.48
$S(EG)_3NH_3^+$	−112.3	−24.9	−87.3	43	

[a] $\Delta\Delta G_b = \Delta G_b(NH_3^+) - \Delta G_b(RNH_3^+) = RT \ln(K_d/K_d)$; $R = 1.98$ cal·mol^{-1}·K^{-1}, $T = 300$ K.

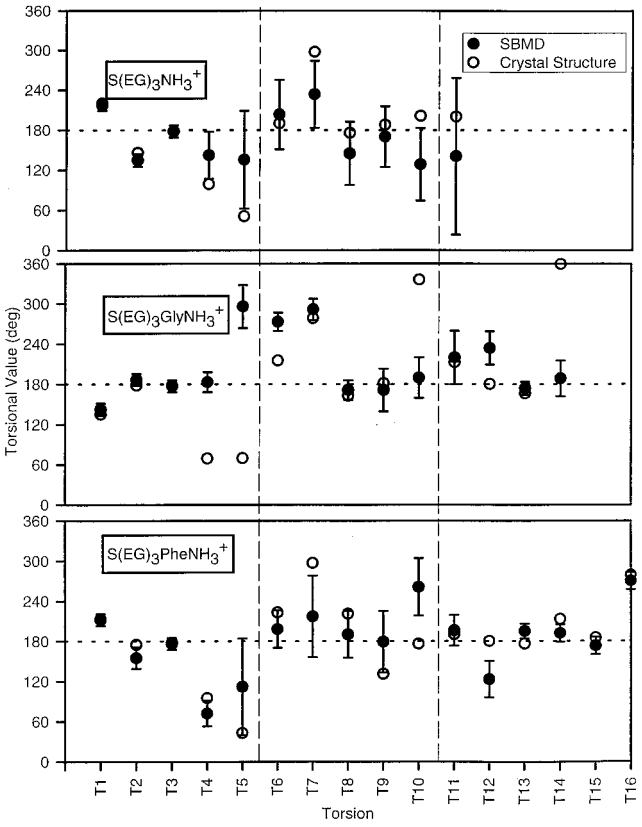

Figure 9. Torsional ranges for the inhibitors during the simulation (nomenclature from Figure 2).

We calculated the average interaction energy (⟨IE⟩, a value that approximates an enthalpy) between the ligands and the protein and the average van der Waals (⟨vdW⟩) and electrostatic ⟨Elec⟩ components of this energy from the simulations (Table 2). The purpose of this calculation was to see if calculated energies and observed free energies of binding (obtained from binding constants) showed either qualitative or quantitative correlation.[32] The results indicate that there is no interpretable correlation between these calculated energies and the observed free energies. The three inhibitors studied here have almost the same values of ΔG_b: the difference is ∼0.7 kcal/mol between tightest and weakest binding. By contrast, the spread in calculated energies is more than 16 kcal/mol, or a factor of $\sim 10^{12}$ in binding constants.

There are two possible inferences from these data: (i) either the calculated energies are meaningless or (ii) the calculated energies are meaningful (to some extent), but there are large entropic contributions to ΔG of binding that largely cancel the enthalpic terms. Since the simulations seem to reproduce the conformation observed crystallographically, and also to correlate in general terms with the conformational mobilities inferred from structure factors (see below; Figure 9), we tend to favor the second inference.

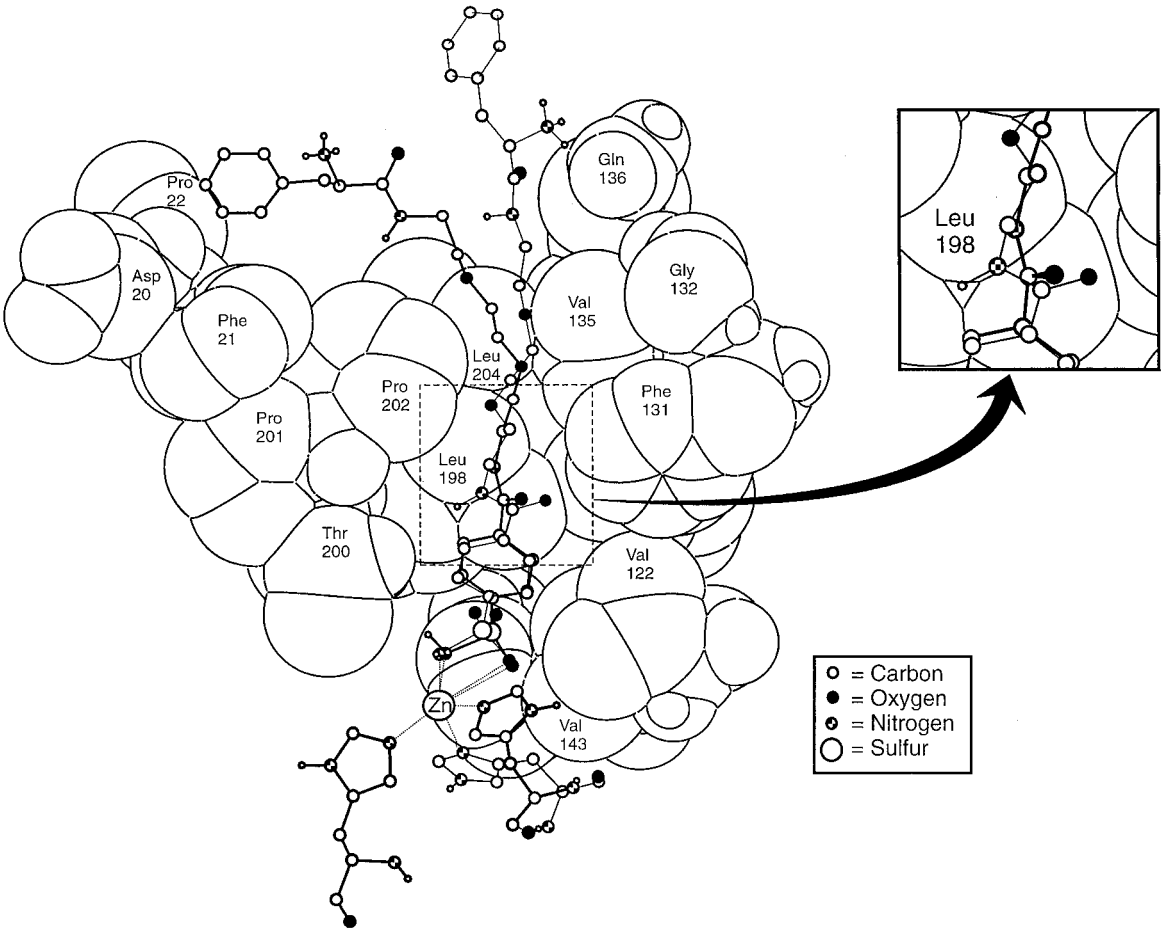

Figure 10. Comparison between the average conformation of $S(EG)_3PheNH_3^+$ from a simulation where the Phe group was initially placed near Phe20 and Pro202 on HCAII (thick line) and that of the crystal structure (thin line).

The Average Torsional Values of the Inhibitors from the Simulations Were Similar to Those from the Crystal Structure, and Their Ranges Were Consistent with the Observed Values of K_d. The ranges of torsional values for each inhibitor in the simulations were compared to those from the crystal structure (Figure 9). These data show that the range of values from the simulation mostly contained those from the crystal structure. Inhibitors terminating in Gly and Phe were less conformationally flexible than the parent inhibitor, and their torsional ranges were consistent with the mode of binding of these inhibitors. That is, groups that interacted most strongly with the active site of HCAII—atoms toward the sulfonamide group and toward the Phe and Gly terminal group—were conformationally more restricted than those that did not bind tightly.

Simulating Possible Hydrophobic Interactions between the Phe Group in $S(EG)_3PheNH_3^+$ and a Known Hydrophobic Patch on the Active Site Caused the Ligand to Adopt a Conformation That Was Not Consistent with That in the Crystal Structure. We wanted to see if simulating a hydrophobic interaction between the terminal Phe group of $S(EG)_3PheNH_3^+$ and the lip of the active site would result in a plausible conformation in the ligand. That is, could the terminal Phe group of $S(EG)_3PheNH_3^+$ bind to a hydrophobic patch located nearby on the lip of the active site and result in a conformation of the inhibitor consistent with that from the crystal structure. To test this hypothesis, we carried out a simulation in which the conformation of the terminal Phe group of $S(EG)_3PheNH_3^+$ was adjusted so that it was near a hydrophobic binding cleft defined by Phe20 and Pro202. This cleft was suggested to be a plausible binding site for the −Bn group of the SG_3Bn ligand in a previous molecular dynamics simulation.[11] The simulation of this hypothetical conformation of the $S(EG)_3PheNH_3^+$ ligand was set up similarly to the other simulations in this paper and was carried out for 61 ps. Within the first 3 ps, this ligand adopted a conformation that was different from that of the crystal structure and retained this conformation for the remainder of the simulation. The average structure of the ligand from the simulation, significantly different than that of the crystal structure, is shown in Figure 10. Consequently, the simulations suggest that interactions between the −Phe group and the hydrophobic patch on the active site do not contribute to a plausible conformation for $S(EG)_3PheNH_3^+$.

Conclusions

We have performed several molecular dynamics simulations to determine the motion of three oligoethylene glycol-based ligands when bound to the active site of HCAII. These simulations roughly reproduce conformations observed for the crystal structure but do not reproduce even qualitatively the observed free energies of binding. The fact that conformations are reproduced (albeit in simulations starting with the crystallographic conformation), but free energies are not, is compatible

with the hypothesis that entropy terms contribute importantly to the free energy of binding; in the absence of calorimetric data, we cannot test the hypothesis of compensating contributions from entropy and enthalpy to free energy. The observation that the "tails" on the ligands that we are examining are mobile, even when associated with the protein, is not incompatible with the hypothesis of an entropy−enthalpy compensation or other data derived from trends in experimental binding constants that suggest this kind of compensation.[4,5]

Overall, what value are these types of simulations? We believe that their greatest value may be to provide a general sense of energetically accessible conformations that bring molecular surfaces of protein and ligand in close proximity and, thus, to suggest opportunities for modifying the structure of the ligand to increase its molecular surface in contact with protein. They thus supplement qualitatively the quasistatic picture of the ligand−protein interaction that emerges from the crystal structure to characterize different parts of the ligand as "mobile" or "immobile," they focus attention on parts of the ligand that may contribute to enthalpies and entropies.

We suspect that these types of simulations would help to identify energetically very unfavorable modifications to structures. We note, however, the problem of local minima. The hypothesis of a favorable hydrophobic interaction between the phenyl group of the ligand terminating in Phe and a hydrophobic patch on the surface of the protein resulted in a long-lived (by standards of these simulation) interaction, but whose conformation does not agree with the crystal structure.

The increasing atomic fluctuation of the atoms in the inhibitors from the sulfonamide end to the terminal end parallels the increasing disorder observed for the ligands when bound to the active site of HCAII in the crystal structure. The small differences between the geometries of the amino acids in simulations and the crystal structures suggest that the parameters used in the SBMD method are able to reproduce at least some features important in these types of studies. As a result, the SBMD method remains a possible way of studying the dynamics of protein−ligand interactions in water, while significantly reducing the computational expense of similar studies that include all solvated atoms in a protein−ligand system. Together, principles of physical−organic chemistry used in conjunction with molecular dynamics are useful tools in the study of protein−ligand interactions: they provide a way of exploring conformations and some aspects of protein−ligand interactions in the presence of solvent, and they augment information from crystallographic studies. Their inability to estimate entropic terms, and the unknown accuracy of their potential functions, remain serious weaknesses.

Acknowledgment. This study was supported by the NIH (NIH GM30367). We would like to thank George Sigal, Mathai Mammen, and Jinming Gao for critical reading of this manuscript.

Supporting Information Available: Listings of parameters used (6 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO970493Z