Application of the Free Energy Perturbation Method to Human Carbonic Anhydrase II Inhibitors

Karen A. Rossi,[†,§] Kenneth M. Merz, Jr.,[*,†] Graham M. Smith,[‡] and John J. Baldwin[‡,⊥]

Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, and Merck Research Laboratories, West Point, Pennsylvania 19486

Received November 14, 1994[⊗]

An analysis of the free energy perturbation (FEP) method is presented that attempts to evaluate the efficacy of the FEP method in the drug discovery process. To accomplish this we have evaluated whether the FEP technique can accurately predict energetic and structural quantities relating to the inhibition of human carbonic anhydrase II (HCAII) by sulfonamides. Three well-characterized (both structurally and energetically) sulfonamide inhibitors of HCAII were examined in this study, **1a**, **1b**, and **1c**. Results from FEP simulations on these compounds indicate that the FEP method can predict energetic trends reasonably well; however, the FEP method was less successful in reproducing detailed structural data. In particular, an expected movement of His-64 when inhibitor **1c** was bound did not occur. We conclude that the FEP method can be used to determine relative free energies of binding but cannot be relied upon to reproduce subtle geometric changes.

Introduction

Computational chemistry is rapidly becoming an essential component of modern pharmaceutical research.[1-3] Until recently, none of the results from these methods could be directly compared to experimental energetic (*i.e.*, free energies) information. The free energy perturbation (FEP) technique has emerged as a computational method that allows for the direct comparison between experimental and calculated free energies.[4-6] This method can be used to calculate the relative free energy between two similar states, as well as other free energy quantities (*e.g.*, absolute free energies of binding, *etc.*). The relative free energies can then be compared to experimental free energies of binding, which are obtained through experimental inhibition constants (K_i's).

One strategy that can be employed when using the FEP method is to start the modeling from a known enzyme/inhibitor complex.[4-6] The bound inhibitor can then be mutated via the FEP technique into another closely related compound by the introduction of various functional groups (*e.g.*, alkyl groups, halogens, *etc.*). Upon the completion of the FEP simulations, a determination could be made whether the "candidate drug" would bind effectively to the enzyme target. Hence, the FEP method would effectively aid in the selection process of new compounds with better binding affinities and could possibly save significant laboratory time and materials.

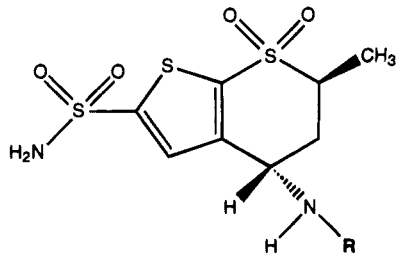
We will examine two different questions in this article: (1) Can we accurately predict relative free energies ($\Delta\Delta G$) using the FEP techniques? (2) Can we simultaneously predict associated structural changes? For the former question numerous studies have suggested that FEP methods can give accurate energetic quantities[4-6] but that care should be taken with respect to running simulations long enough to ensure that the configurations accessible to the system are sampled.[4-8] For the latter question much less effort has been focused; however, several studies have demonstrated that FEP methods may give the correct energetic quantities, while not reproducing experimental structural data.[7,9] In order to reproduce structural information, the force field used must be accurate and the available configurational space must be thoroughly sampled. Achieving both of these conditions in order to reproduce structural rearrangements is clearly quite difficult given the current level of force field sophistication as well as our ability to only carry out subnanosecond molecular dynamics simulations routinely due to algorithmic and/or computer power limitations.

Since much is known about the inhibition of HCAII, this system can serve a useful paradigm for the use of FEP methods in drug design and discovery.[10-12] Numerous crystal structures are available,[13] and among these are included the structures of the three sulfonamide inhibitors that were studied herein.[14] The function of HCAII is the rapid (10^6 s^{-1}) interconversion of carbon dioxide and bicarbonate, and it is present in the nonpigmented epithelial cells of the eye.[10] When HCAII is inhibited, the production of bicarbonate is decreased, which results in a decrease of sodium and fluid secretion to the eye.[11,12] When too much bicarbonate is produced, the intraocular pressure (IOP) in the eye is increased, and if it gets too high (above 30 mmHg) treatment is necessary, or damage to the optic nerve may occur, which can lead to blindness.[15] Elevated IOP is a symptom of glaucoma, and while the causes of glaucoma are unknown, it is possible, through early treatment, to relieve the condition.

Carbonic anhydrase inhibitors (CAIs) are used to treat glaucoma and can be administered orally or topically. Many studies have shown unpleasant side effects from oral CAIs such as acetazolamide and methazolamide.[16,17] Problems resulting from these CAIs range from altered taste and skin eruptions to anorexia and depression.[18] Additionally, the orally active drugs inhibited HCAII processes in the kidneys. Topically active inhibitors are, therefore, more desirable

[†] The Pennsylvania State University.
[‡] Merck Research Laboratories.
[§] Current address: DuPont Merck Pharmaceuticals, Wilmington, DE.
[⊥] Current address: Pharmacopia, Princeton, NJ 08540.
[⊗] Abstract published in *Advance ACS Abstracts*, May 1, 1995.

0022-2623/95/1838-2061$09.00/0



1a. R=H
1b. R=CH_3
1c. R=CH_2CH_3

Figure 1. Inhibitor structures **1a**, **1b**, and **1c**.

because they avoid these systemic side effects. The most effective topically active treatment should be readily absorbed into the eye. β-Blockers[19] are presently used as a topically applied treatment for glaucoma, but they have serious side effects that fall into three areas: pulmonary,[20] cardiovascular,[21] and central nervous system effects. This has led to an extensive search for sulfonamides that can be topically applied. Extensive research on thienothiopyrans have been performed which has demonstrated that these compounds are water soluble, active inhibitors with few side effects.[22,23] The most effective inhibitors contained a sulfonamide in the 2-position with a basic amine located elsewhere in the molecule.[22]

In the uninhibited form, HCAII has a zinc-bound hydroxide (or water at low pH) at the active site. This zinc-bound hydroxide (water) has a role in the CO_2/bicarbonate interconversion process[10]:

$$CO_2 + H_2O + \text{E-ZnOH}^- \rightleftharpoons HCO_3^- + \text{E-ZnH}_2O \quad (1)$$

The zinc-bound hydroxide ion must be regenerated, and this occurs through an intermolecular proton transfer between the zinc-bound H_2O and His-64 with the aid of intervening water molecules.[10] This completes the catalytic cycle.

$$\text{His64-E-ZnH}_2O \rightleftharpoons \text{H}^+\text{-His64-E-ZnOH}^- \overset{\text{buffer}}{\rightleftharpoons} \text{His64-E-ZnOH}^- + \text{H}^+\text{-buffer} \quad (2)$$

This cycle shows the catalytic importance of His-64, and motions of this residue are the key structural fluctuation that will be examined in this study. This residue is located about 7.5 Å away from the zinc ion[13,24] and aids the interconversion of CO_2 and HCO_3^- by translocating a proton. When a potent inhibitor like the thienothiopyran **1c** (see Figure 1) is bound to the zinc ion, the imidazole ring of His-64 is moved approximately 3 Å away from the active site (considered to be the "away" position) as shown in Figure 2.[14] In Figure 2 the His-64 in the "native" position is labeled, while the His-64 in the "away" position is unlabeled.

Sulfonamides inhibit HCAII in their anionic form by binding the amide portion of the molecule to the zinc ion in the active site.[11,12] When the inhibitor is bound to the zinc ion, the resulting coordination environment is a distorted tetrahedron with three histidine ligands (94, 96, and 119) and the fourth coordination site occupied by the RNH^- group of the sulfonamide. The inhibitor prevents the interconversion process that HCAII is responsible for by occupying the fourth coordination site where a hydroxide ion resides in the native HCAII active site. Numerous sulfonamides have been synthesized,[11,12] and ample data is available,[22,23] including inhibition constants (K_i's), which can be directly converted to ΔG_{bind} values. Three closely related sulfonamide inhibitors were chosen for this study: **1a**, **1b**, and **1c** (see Figure 1). Structures **1a** and **1b** have very similar inhibition constants (1.52 and 1.88 nM), while **1c** is roughly 5 times more effective (0.37 nM).[14] The only difference between the three inhibitors is the length of the amino side chain. Inhibitor **1c**, which is the tightest binding inhibitor, forces His-64 to move into the "away" position (see Figure 2).[14] This is believed to occur because the ethylamine side chain is long enough to create unfavorable steric interactions with His-64. The methylamine (**1b**) and amine (**1a**) side chains are not long enough to cause interactions of this kind, with the end result being no movement of His-64.

These three inhibitors will be examined in order to test the FEP method's ability to predict correct energetic values as well as structural changes. These structures have different binding and structural properties and, therefore, will be effective indicators of the ability of the FEP method and its efficacy in drug design applications.

Computational Approach

Free energy perturbation (FEP) is a method which can be used to obtain a relative free energy between two similar structures.[4-6] In this method, one structure is "perturbed" into another by slowly modifying the parameters describing the molecule. This method determines the free energy difference between the starting state *versus* the final state. Hence, FEP is a useful tool in the simulation of many biologically and chemically relevant processes.[4,25-27]

Not only does the FEP method determine thermodynamic information on biomolecular systems, but it also gives a description of their structural evolution over time. The FEP method is useful for many types of applications, such as mutagenesis studies, to determine the effects of certain residues,[9,28,29] or it can be used as a preliminary device in planning for experimental studies. Additionally, the FEP method can be applied to structure/function relationships and the study of the function of biomolecules.[30-32] Finally, the FEP method can be used in examining specific ligands and inhibitors.[33,34] This research will focus on this final aspect.

FEP techniques can be used to determine the relative free energy between two states. Although an absolute free energy (horizontal directions in the thermodynamic cycle below) is too computationally intensive to be determined on large systems like proteins, the relative free energy is useful for drug design applications. The relative free energy of substrate 1 (S_1) and substrate 2 (S_2) can be computed directly by using the following thermodynamic cycle.

In this method, S_1 is perturbed into S_2 in the enzyme, yielding ΔG_{bind}, and a perturbation of S_1 to S_2 is performed in aqueous solution to represent the sub-

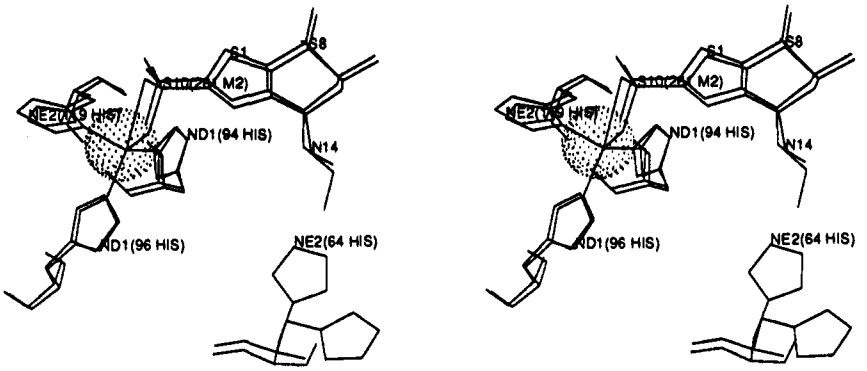


Figure 2. Movement of His-64 from the "native" (labeled) conformation to the "away" (unlabeled) conformation.

strate in solution (ΔG_{sol}). Then, by employing the thermodynamic cycle and the fact that free energy is a state function, eq 1 and 2 can be derived which give a relative free energy of binding as a difference between ΔG_{sol} and ΔG_{bind}.[4-6]

$$\Delta G_1 + \Delta G_{bind} - \Delta G_2 - \Delta G_{sol} = 0 \quad (1)$$

$$\Delta\Delta G_{bind} = \Delta G_2 - \Delta G_1 = \Delta G_{bind} - \Delta G_{sol} \quad (2)$$

Hence, through the application of this thermodynamic cycle–FEP (TC–FEP)[4-6] approach one can determine the binding ability of one compound relative to another.

Parameterization

It was necessary to develop a force field for the sulfonamide inhibitors used in this investigation, since many of the parameters were not included in the AMBER 4.0[35] force field. Other parameter sets for sulfonamides have been reported,[36] but these sets are not compatible with the AMBER force field (*e.g.*, MM2, *etc.*), so we carried out our own parameterization effort. Recall that the TC–FEP method requires two sets of simulations; hence, two sets of parameters need to be developed.

The first set of parameters determined were atomic point charges for the inhibitor in aqueous simulations. Since the inhibitors bind to HCAII in their anionic forms, the formal charge on all inhibitors was set to −1. There is some uncertainty regarding whether these inhibitors bind as the anion or as the neutral molecule (see refs 10–13 and references cited therein), but since we are not creating or destroying charge during the simulation, either a neutral or charged model will give us an accurate free energy of solvation difference between the inhibitors. We chose the anionic model for our condensed phase simulations since it is compatible with the enzyme simulations. The inhibitors were minimized in the gas phase using MOPAC 5.0[37] and the AM1[38] Hamiltonian. Figure 3 lists atomic point charges that were calculated for each of the inhibitors in the gas phase using electrostatic potential (ESP)[39] fitting in conjunction with the MNDO[40] Hamiltonian. The AM1 Hamiltonian yields better structures yet produces charges that are in poor agreement with *ab initio* charges.[39] MNDO, on the other hand, was found to be superior in charge determination and, therefore, was used in the studies presented herein. It has been shown that scaled MNDO charges match 6-31G* charges more closely than unscaled charges;[39] however, this was not

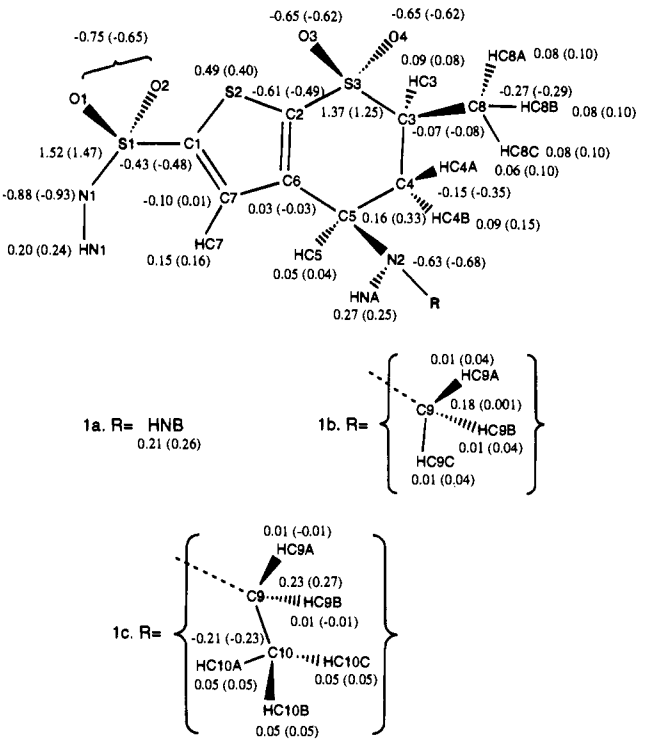


Figure 3. Atom numbering and atomic point charges for **1a**, **1b**, and **1c**. The first atomic point charge given corresponds to the value used in the solution phase simulations, while the charges in parentheses were those used in the enzyme simulations.

done since scaled charges would have resulted in noninteger net charges on the charged inhibitors.

The same procedure was used in the determination of the atomic point charges for the inhibitor while bound to the enzyme. These charges are also given in Figure 3 (see values given in parentheses) for each inhibitor. The computational model employed to determine the atomic point charges included the zinc ion as well as three imidazoles and the sulfonamide inhibitor. The zinc and imidazole active site model was included in the determination of the charges to allow the inhibitor charges to be influenced by the species surrounding it (see Figure 4). In these models the net charge was +1 (Zn = +2, inhibitor = −1, imidazoles = 0). Again, the structure was minimized with MOPAC 5.0.[37] using the AM1[38] Hamiltonian, and charges were determined using MNDO.[40] Since we used truncated histidine residues in the ESP calculations, we needed to determine charges for the remaining atoms of the residue. The backbone charges for the histidines were obtained directly from the AMBER database while the C_β and HC_β AMBER

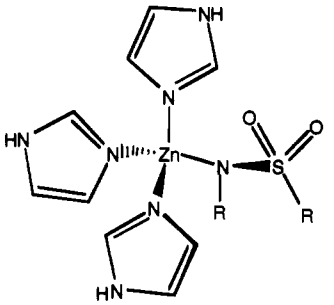


Figure 4. Model of the HCAII active site where imidazoles are substituted for histidines.

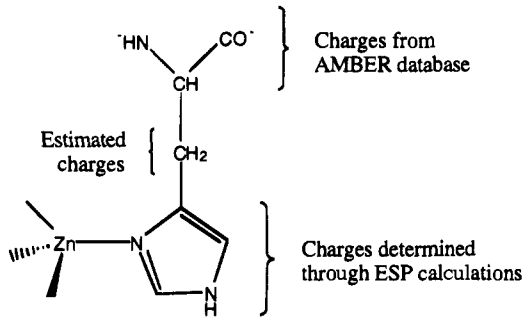


Figure 5. Procedure used for charge determination of the zinc-bound histidines.

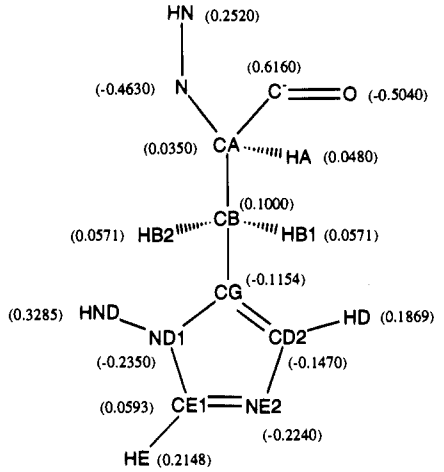


Figure 6. Atomic point charges for histidine residue, HDL.

charges were altered to set the overall charge of the inhibitor and active site model to +1. Figure 5 summarizes the procedure used. The resulting charges for the two types of histidines are given in Figures 6 and 7, and the charge on zinc was +0.76 (these charge sets were averaged over the three ESP calculations) throughout. The inhibitor charge was also averaged over the three ESP calculations except for the charges of the amine, methylamine, and ethylamine charges. This averaging procedure resulted in charge sets that only needed to be altered slightly in order to retain charge neutrality. The resulting charge model had a constant charge for the zinc ion as well as the surrounding histidine ligands and for most of the inhibitor except for the side chains. This procedure was used since it isolated the changes in the charges in these systems to a small region of the inhibitor molecule. However, this approach is not the only one possible, but it offered us the most flexibility, while retaining a simple representation. In general, however, it should be stressed that due to the modest structural changes in these inhibitors

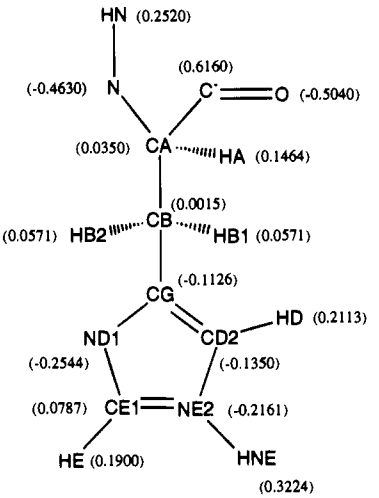


Figure 7. Atomic point charges for histidine residue, HEL.

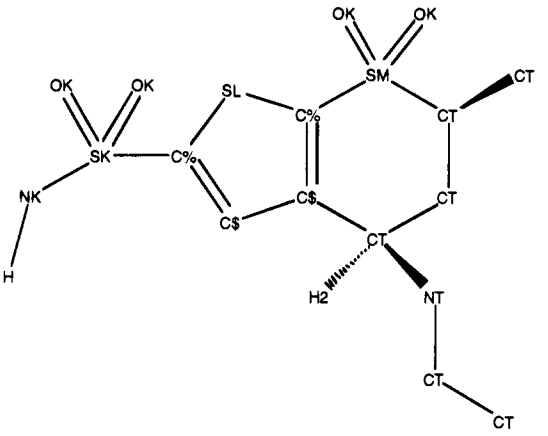


Figure 8. Atom labels used in the force field representation.

only small charge changes were observed on going from **1a** to **1b** to **1c**. This is as opposed to a perturbation that involves, for example, the creation or annihilation of a net charge.

It has been thought that the N–S–C–S (see Figure 1) torsion profile can be related to the efficacy of a HCAII inhibitor.[22] *Ab initio* calculations on 2-sulfonamide–thiophene at the 3-21G* level give a minimum at a ~75–90° torsion angle and the maximum at 180° (~3.5 kcal/mol above the minimum).[22,41] However, we found in test runs that by placing the torsion profile minimum and maximum in these locations we were unable to reproduce the active site geometry. According to the X-ray crystal structures, this torsion angle appeared to be ~140°, rather than 75°. For example, experimental N–S–C–S dihedral angles in inhibitors **1a**, **1b**, and **1c** bound to HCAII were 151.7°, 154.1°, and 139.1°, respectively.[14] When a 140° minimum energy dihedral was implemented into the force field, the structural characteristics of the inhibitor/HCAII complex were well-reproduced (see results below). Therefore, a minimum energy torsion angle of 140° was used in the force field. Unique atoms labels, shown in Figure 8, were assigned to the inhibitor for use in the force field. The additional parameters added to the force field are listed in Tables 1–4.

Procedure

Once the parameterization was completed, the simulations were performed. Perturbations were completed in the follow-

Table 1. Derived Bond Parameters

bond	K_r (kcal/Å^2)	r_{eq} (Å)
ZN–NB	40.	2.050
ZN–OK	20.	3.050
C%–SK	222.	1.710
C%–SL	222.	1.740
C%–SM	222.	1.72
ZN–NK	100	1.910
SK–NK	230.	1.72
CT–SM	222.	1.78
H–NK	434.	1.01
OK–SM	683.	1.435
OK–SK	683.	1.435
DC–HC	331.	1.09
HC–DH	434.	1.01
H–DH	434.	1.01
DH–DC	331.	1.09
H–DC	337.	1.45
DC–DC	310.	1.40
NT–CT	367.	1.471
C%–C$	469.	1.4
C$–C$	469.	1.4
C$–CT	317.	1.51
C$–HC	340.	1.08
H2–DH	331.	1.09

Table 2. Derived Angle Parameters

atoms	K_θ (kcal/rad^2)	θ_{eq} (deg)
CR–NB–ZN	20.	126.0
CV–NB–ZN	20.	126.0
CC–NB–ZN	20.	126.0
CR–NB–ZN	20.	126.0
NB–ZN–NB	20.	109.5
NB–ZN–NK	80.	109.5
NB–ZN–OK	0.	85.0
NK–ZN–N	20.	109.5
H–NK–H	35.	120.
ZN–NK–H	100.	137.6
ZN–NK–SK	100.	110.3
ZN–OK–SK	35.	110.3
NK–ZN–OK	20.	60.0
NK–SK–OK	100.	107.0
NK–SK–C%	100.	108.5
H–NK–SK	35.	108.9
SK–C%–SL	100.	121.4
SK–C%–C$	100.	125.8
OK–SK–OK	104.	116.0
OK–SK–C%	74.	108.9
C%–SL–C%	58.	91.8
SL–C%–SM	80.	122.0
SL–C%–C$	70.	112.7
C%–SM–OK	74.	110.1
C%–SM–CT	58.	100.6
SM–CT–HC	51.	104.0
SM–CT–CT	69.	106.7
SM–C%–C$	69.	124.2
OK–SM–OK	104.	118.0
OK–SM–CT	74.	108.2
N–CT–C$	80.	109.5
DH–DC–DH	35.	109.5
DC–DC–DH	35.	109.5
HC–DC–HC	35.	109.5
HC–DC–DH	35.	109.5
HC–DC–DC	80.	109.5
DC–DC–DC	40.	109.5
CT–HC–DH	35.	109.4
CT–HC–DC	80.	109.4
DH–HC–DC	35.	109.4
N–H–DH	35.0	109.5
DH–H2–DH	35.0	109.5
DH–HC–DH	35.0	109.5
CT–NT–H2	35.0	109.5
CT–NT–H2	35.0	109.5
CT–CT–NT	80.0	109.7
DH–H2–NT	80.0	109.7
CT–NT–CT	80.0	109.7
HC–CT–NT	35.0	109.7
CT–NT–CT	50.0	108.0
CT–NT–H2	35.0	107.0
C$–CT–NT	63.0	109.0
C%–C$–C$	85.0	111.4
C%–C$–HC	35.0	120.0
C$–C$–HC	35.0	120.0
C%–C$–CT	70.0	120.0
C$–C$–CT	70.0	124.0
C$–CT–HC	35.0	109.5
C$–CT–CT	63.0	114.0

ing manner: **1a** → **1b**, **1b** → **1c**, **1c** → **1b**, and **1b** → **1a** in both solution and in the enzyme active site. In this notation the first number indicated (*e.g.*, **1a**) identifies the crystal structure that we started from. Note that we have run four separate aqueous phase and enzyme sets of FEP simulations where in each case we do the forward (180 ps) and reverse (180 ps) simulation (see below) starting from the indicated crystal structure. In this way perturbations were performed in both directions starting with all possible crystal structures.[14]

Each aqueous phase simulation used the anionic form of the inhibitor represented by the all atom model. Each inhibitor was placed in a computational cell with approximately 550 TIP3P[42] water molecules. Periodic boundary conditions were used during these simulations. Each structure was minimized for 2000 steps using the steepest descent for the first 500 steps and conjugate gradient for the remaining steps. Molecular dynamics equilibration was done for 72 ps, which was followed by a 180 ps slow growth FEP simulation. The endpoint of the "forward" run was then equilibrated for another 72 ps, and a "reverse" simulation of 180 ps was then carried out.

All enzyme simulations were begun with the crystal structure of the relevant enzyme and inhibitor complex. The active site residues (His-94, -96, and -119) as well as the inhibitor used the all atom model,[43] while the rest of the protein used the united atom model.[44] The enzyme structure was solvated by placing a water "cap" with a 15 Å radius from the Cγ on residue His-64. A 15 Å "belly" was created which centered around the zinc ion. Within this "belly" all residues within 15 Å of the zinc ion as well as all cap water molecules were included in the MD and FEP runs while residues outside these regions remained fixed. Minimization was done for 100 steps without SHAKE,[45] using steepest descent. The SHAKE[45] algorithm was then activated, and an additional 500 steps of steepest descent minimization was carried out. The resulting structure was equilibrated for 72 ps, and a 180 ps FEP simulation was then carried out. The end point of the FEP simulation was then equilibrated for 72 ps, and the "reverse" run was done in the same manner.

All MD and FEP simulations were done at constant temperature (298 K).[46] The aqueous phase simulations were carried out at a constant pressure of 1 atm.[46] SHAKE[45] was employed during the FEP simulations to restrain the bond lengths to their equilibrium distances in order to remove high-frequency motions, thus, allowing a time step of 1.5 fs.

Results

We first consider the energetic information resulting from the FEP simulations which is summarized in Table 5. The calculated free energies of solvation are consistent with the changes in the characteristics of the inhibitors. The free energy of solvation for **1a** → **1b** (*i.e.*, H → Me) yielded a ΔG_{sol} of 0.82 ± 0.08 kcal/mol (the error bars represent ±1σ), and the perturbation from **1b** → **1a** (*i.e.*, Me → H) was −0.73 ± 0.10 kcal/mol. This suggests that **1a** is better solvated than **1b** by about −0.75 kcal/mol. This agrees with intuition, in that the amino side chain on **1a** has two protons which can hydrogen bond to solvent, in addition to the nitrogen. On the other hand, **1b** is not as well solvated because a nonpolar methyl group replaces an hydrogen and, thus, will not allow for an additional hydrogen bond at that point. We can estimate an experimental value for this free energy conversion by considering the conversion of

Table 3. Derived Dihedral Angles

dihedral	division	V_n (kcal/rad^2)	γ (deg)	n
X-ZN-NK-X	9	0.0	0.	3.
X-ZN-NK-S	3	5.5	0.	3.
X-NK-SK-X	6	8.0	180.	3.
OK-SK-C%-SL	2	0.	0.	2.
OK-SK-C%-C$	2	0.	0.	2.
NK-SK-C%-SL	1	5.00	180.	-2.
NK-SK-C%-SL	1	7.00	0.	4.
NK-SK-C%-C$	1	0.0	0.	2.
X-C%-SL-X	6	2.0	180.	2.
C$-C%-SL-C%	1	4.0	180.	2.
SK-C%-SL-C%	1	4.0	180.	2.
X-SM-C%-X	6	0.5	0.	2.
X-SM-CT-X	9	3.2	0.	3.
HC-CT-CT-SM	1	0.5	180.	3.
CT-CT-CT-SM	1	0.5	180.	3.
X-DC-DC-X	9	1.3	0.	3.
X-HC-DC-X	6	1.0	0.	3.
X-CT-HC-X	9	1.0	0.	3.
X-ZN-NB-X	6	1.0	0.	3.
X-ZN-OK-X	9	0.0	0.	3.
X-SK-OK-X	9	0.0	0.	3.
X-CT-NT-X	6	4.0	0.	3.
X-C%-C$-X	4	6.3	180.	2.
SM-C%-C$-X	4	3.0	180.	2.
CT-C%-C$-X	4	3.0	180.	2.
X-CT-C$-X	4	3.0	180.	2.
X-C$-C$-X	4	5.3	180.	2.
X-NT-H2-X	6	1.3	180.	3.

Table 4. Derived Parameters for Nonbonded Interactions

nonbond atoms	R^* (Å)	ϵ (kcal/mol)
OK	1.6	0.20
NK	1.75	0.16
SK	2.00	0.20
SL	2.00	0.20
SM	2.00	0.20
ZN	1.1	0.0125
C%	2.2385	0.049
C$	2.2385	0.049
DC	0.0	0.0
DH	0.0	0.0

Table 5. Results from the FEP Simulations[a]

inhibitor (-R)	ΔG_{bind}	ΔG_{sol}	$\Delta\Delta G_{bind}$	expt $\Delta\Delta G_{bind}$
1a → **1b**	1.04 ± 0.31	0.82 ± 0.08	0.22 ± 0.32	0.13
1b → **1a**	−1.56 ± 0.32	−0.73 ± 0.10	−0.83 ± 0.33	−0.13
1b → **1c**	0.18 ± 0.16	0.56 ± 0.04	−0.38 ± 0.16	−0.96
1c → **1b**	0.71 ± 0.19	−0.63 ± 0.12	1.34 ± 0.23	0.96

[a] All values listed in kcal/mol.

$MeNH_2 \rightarrow Me_2NH$ which gives a solvation free energy difference of 0.27 kcal/mol, which is smaller but in the same direction as the calculated value of ~0.75 kcal/mol for **1a** → **1b**.[47]

The solvation free energy difference between inhibitors **1b** and **1c** were also consistent with intuition. The FEP simulations predict a value of 0.56 ± 0.04 kcal/mol for **1b** → **1c** (*i.e.*, Me → Et) and −0.63 ± 0.12 kcal/mol for **1c** → **1b**. These results suggest that **1b** is better solvated than **1c** by ~0.60 kcal/mol. Although **1b** is not as well solvated as **1a**, it is reasonable to expect **1b** to be better solvated than **1c** since inhibitor **1c** has an ethylamine side chain, while **1b** only has a methylamine. A suitable experimental model system would be the conversion of $Me_2NH \rightarrow MeEtNH$, but we are unaware of any existing experimental free energy data for this conversion. However, we can estimate the full conversion (*i.e.*, **1a** → **1c**) from $EtNH_2 \rightarrow Et_2NH$, which has a solvation free energy difference of 0.43 kcal/mol. This can be compared to the calculated value of ~1.4 kcal/mol. Hence, the overall sign of the calculated free energy matches experiment reasonably well, but the magnitude of the calculated free energy for this conversion may be overestimated.

The simulation of **1a** → **1b** within the HCAII active site resulted in a relative free energy of 1.04 kcal/mol, and the reverse simulation of **1b** → **1a** gave a similar value of −1.56 kcal/mol. These simulations both favor inhibitor **1a**, which has the hydrogen on the amino side chain, versus the methyl group on inhibitor **1b**. The simulation of **1b** → **1c** yielded a ΔG_{bind} of 0.18 kcal/mol, which favors **1b**, while the reverse simulation of **1c** → **1b** yielded a ΔG_{bind} of 0.71 kcal/mol, which favors **1c**. Hence, the difference in the calculated ΔG_{bind} for the latter sets of simulations (*i.e.*, **1b** → **1c** and **1b** → **1c**) is much larger than for the former. This suggests that the FEP simulations are having difficulty placing the ethyl side chain in the active site and perhaps the side chain is becoming trapped in long-lived (relative to our simulation time scale) local minimum.

By applying the thermodynamic cycle to these results we find that the calculated $\Delta\Delta G_{bind}$ between all structures is in reasonable agreement with experiment. Perturbation from **1a** → **1b** resulted in a relative binding free energy of −0.22 ± 0.32 kcal/mol, which is in excellent agreement with the experimental value of −0.13 kcal/mol. When the perturbations were started from the crystal structure of **1b**, the calculated relative free energy for **1b** → **1a** was 0.83 ± 0.33 kcal/mol, which is in reasonable accord with experiment. Although the runs yield somewhat different results when using the different crystal structures as starting points, they both correctly predict the better inhibitor.

Applying eq 2 to the FEP simulations for **1b** → **1c** and **1c** → **1b** gave values of 0.40 ± 0.16 kcal/mol and −1.33 ± 0.23 kcal/mol, respectively. The experimental relative free energy for this change is 0.96 kcal/mol (**1b** → **1c**). Hence, these runs were again in reasonable accord with experiment, but the difference in the calculated $\Delta\Delta G_{bind}$ are more substantial in this case than in the former (*i.e.*, **1a** → **1b** and **1b** → **1a**). It has been suggested that the origin of the preference for **1c** over **1a** and **1b** has to do with the gain in entropy (due to the loss of a water molecule) upon the introduction of the ethyl side chain. From our FEP analysis we find that this is only part of the story. Indeed, there is a gain in free energy when the RNHMe group is perturbed into RNHEt within the enzyme (the average ΔG_{bind} (**1b** → **1c**) = −0.27 ± 0.25 kcal/mol), but this does not account for the total difference. The solvation component of 0.60 kcal/mol (the average ΔG_{sol} (**1b** → **1c**) = 0.60 ± 0.13 kcal/mol) in this case also makes a very large contribution. Hence, from our calculations it appears that the binding affinity of **1c** over **1a** and **1b** arises mostly from solvation effects, which drive **1c** into the enzyme active site where the loss of a water molecule or favorable inhibitor side chain/protein contacts favor **1c**.

The experimental relative free energy for **1a** → **1b** is small (−0.13 kcal/mol), and therefore, it is more difficult to accurately predict this difference. Additionally, the FEP method can have reasonably large statistical or convergence problems, making it harder to predict

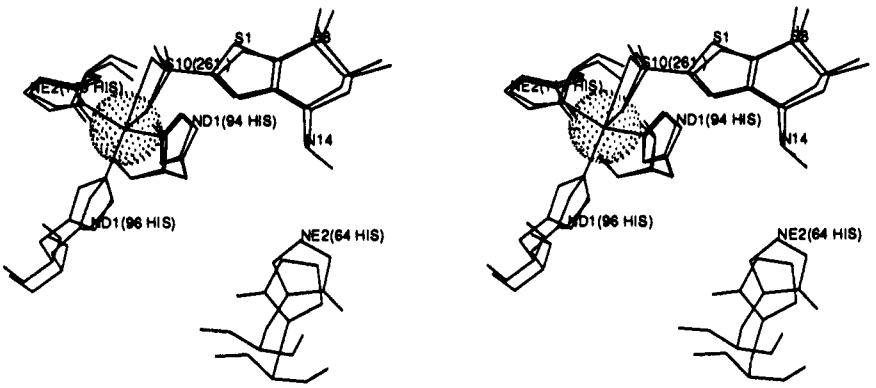


Figure 9. Stereoview of the HCAII active site with the crystal structure of **1b** bound to HCAII (labeled) and after (unlabeled) the FEP conversion of **1a** → **1b**. The unlabeled structure represents the final structure obtained at the end of the FEP simulation and corresponds to **1b**.

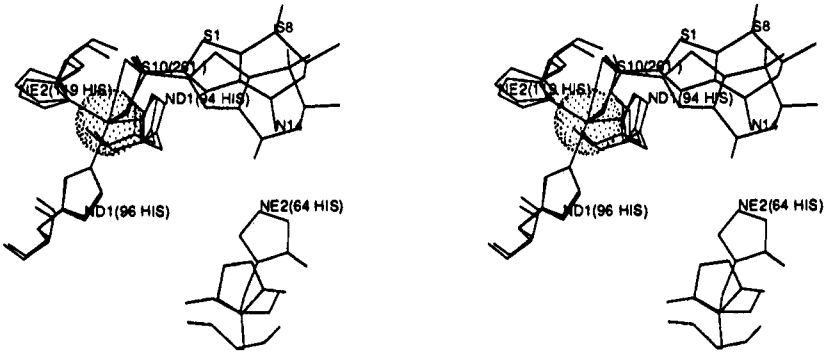


Figure 10. Stereoview of the HCAII active site with the crystal structure of **1a** bound to HCAII (labeled) and after (unlabeled) the FEP conversion of **1b** → **1a**. The unlabeled structure represents the final structure obtained at the end of the FEP simulation and corresponds to **1a**.

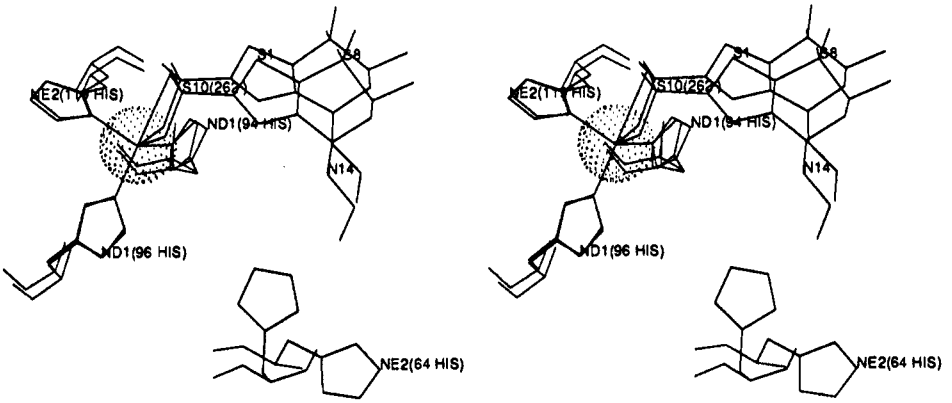


Figure 11. Stereoview of the HCAII active site with the crystal structure of **1c** bound to HCAII (labeled) and after (unlabeled) the FEP conversion of **1b** → **1c**. The unlabeled structure represents the final structure obtained at the end of the FEP simulation and corresponds to **1c**.

accurate free energies. However, what is important to note is that all runs predicted the more favorable structure in every case, including runs which started with opposing crystal structures (*i.e.*, **1a** → **1b** vs **1b** → **1a**). Hence, we find that the FEP method can reproduce experimental relative free energies. This has been observed by others previously and is not that surprising overall.[4-6,33,34]

Another important consideration is the analysis of the structures generated by the FEP simulations. Figures 9–12 describe the active site of the enzyme and shows the His-64 position when **1a**, **1b**, and **1c** are bound. No structural changes were expected to occur within the protein for the perturbation of **1a** → **1b**.[14] Figures 9 and 10 show the results of the FEP simulations for **1a** → **1b** and **1b** → **1a**, respectively. In the figures, the experimental crystal structure of the final state is labeled (*e.g.*, for the conversion of **1a** → **1b** the final state will be **1b**), while the same structure derived from the FEP simulations is unlabeled. One part of the enzyme that was not expected to change in this case was His-64 (see Figures 9 and 10). In this set of FEP simulations, this residue did move around significantly. The imidazole ring flips over and back during the equilibration phase of the simulation, and the backbone appears to shift downward, possibly as a result of the inhibitor moving in the binding pocket. This shift occurs early in the equilibration phase and remains throughout the simulation. Overall, the FEP simulations for **1a** → **1b** and **1b** → **1a** retain the active site structure quite well.

For the **1b** → **1c** FEP simulations, His-64 was expected to move away from the position it adopts in

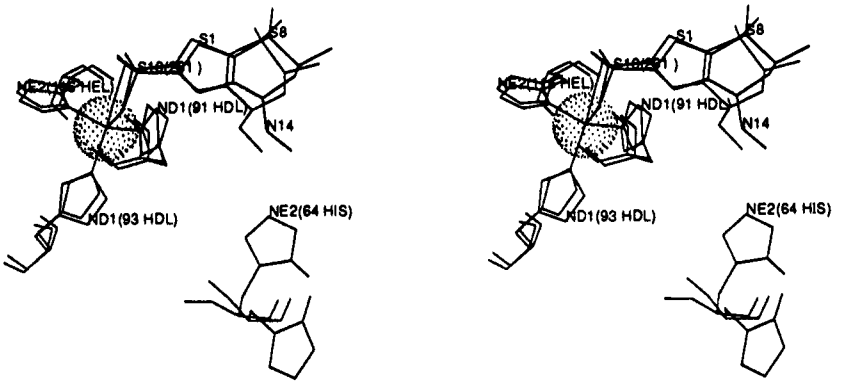


Figure 12. Stereoview of the HCAII active site with the crystal structure of **1b** bound to HCAII (labeled) and after (unlabeled) the FEP conversion of **1c** → **1b**. The unlabeled structure represents the final structure obtained at the end of the FEP simulation and corresponds to **1b**.

the native structure. Inhibitor **1c** binds more tightly than **1b**, and the origin of its binding strength is thought to be partially entropic due to the shifting of His-64 approximately 3 Å away from the active site which releases a tightly bound water molecule.[14] Figure 2 displays this movement, where the "native" position of His-64 is labeled, and the "away" position is unlabeled. In the FEP simulation of **1b** → **1c**, the imidazole ring moved in the incorrect direction (see Figure 11) and essentially adopts a structure similar to that seen for **1a** or **1b**. Upon reversing the direction of this simulation, the imidazole ring did not move back into the native position for inhibitor **1b** (data not shown). Alternatively, when starting with the crystal structure of **1c** (*i.e.*, **1c** → **1b**), which has His-64 in the "away" position, the histidine shifted away from the active site even more and remained there throughout the MD and FEP simulations (see Figure 12).

In all cases, there was a lot of movement of the inhibitor in the active site during the simulations. As noted earlier, there is a large pocket (approximately 15 Å wide and deep) that is divided into hydrophobic and hydrophilic parts. The hydrophobic portion of the binding pocket does not get extensively solvated in our simulations, and this leaves a lot of room for the inhibitor to move. When it moves, it shifts portions of the backbone of the enzyme and, thus, forces important residues, such as His-64 and neighboring residues, out of place.

When **1c** is bound to HCAII, it is presumed that the steric signature of the ethyl side chain forces the movement of His-64 into the "away" position. Hence, when **1c** is perturbed into **1b**, it is reasonable to assume that the imidazole ring would move into the native conformation, but we do not find this to be the case. The dynamics of the inhibitor and His-64 in the active site is too great and the length of our simulations are too short to allow for a thorough sampling of phase space. In this case, we have found it to be difficult to obtain accurate structural pictures of the inhibitor/enzyme complexes. We expect that the reason for this has to do with two factors: The time scale of our simulations, while longer than many FEP simulations,[5,6] is probably not long enough to sample all of the available conformational space. The second factor has to do with the force field. In the present set of simulations the changes we expected are very subtle and are, therefore, very sensitive to the choice of force field parameters.

Conclusions

The free energy perturbation (FEP) method has been examined for its potential use in drug design. Three sulfonamide inhibitors of HCAII were used as models to test this technique. Energetic and structural changes were monitored and compared to experimental data. This research shows that the FEP method can correctly predict the better inhibitor, but it is not able to predict geometric changes as effectively. The expected movement of the imidazole ring of His-64 is caused by steric influences of the amino side chain of the inhibitor and is directly related to the length of the chain. Since this structural change is simply due to steric interactions and not to more specific charge effects (*i.e.*, directional hydrogen bonding, *etc.*), it was difficult to reproduce. Again, if a different system had been studied, perhaps better results could have been obtained. In addition, longer simulation times may need to be run in order to sample conformational space thoroughly. Another factor that has to be considered is the fact that our simulations are more representative of the solution phase than of the crystalline phase, and it may be possible that the positioning of His-64 is different in these two situations. However, we do not feel that this is the case in the present example since the motion of His-64 as a function of the inhibitor appears to be governed solely by bad van der Waals contacts.

From this study it can be concluded that the free energy perturbation method can be problematic for use in drug design and should be employed with care. Too many discrete structural changes were expected in the HCAII/inhibitor complexes that did not occur with this method. These changes are especially important in drug design where adding different substituents can have small but significant effects on the efficacy of the drug candidate. Indeed, small geometric changes may differentiate between a good and poor inhibitor, and the absence of these changes during the course of a FEP simulation could result in the rejection of an improved inhibitor. Furthermore, the amount of time required to set up an effective model as well as the time required to carry out the simulations required months of effort. This kind of time requirement may make the use of the FEP method in the drug discovery process too time consuming at the present time. However, with the development of better force fields and the availability of longer FEP simulations (through the use of massively

parallel processing), the FEP method could become an effective method in the drug discovery process.

Acknowledgment. We thank the NIH (GM-29072) and Merck for supporting this research. Generous allocations of supercomputer time from the Pittsburgh Supercomputer Center, the San Diego Supercomputer Center, and the Cornell Theory Center are gratefully acknowledged.

References

(1) Fruhbeis, H.; Klein, R.; Wallmeier, H. Computer-Assisted Molecular Design (CAMD) – An Overview. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 403–418.
(2) Hopfinger, A. J. Computer-Assisted Drug Design. *J. Med. Chem.* **1985**, *28*, 1133–1139.
(3) Cohen, N. C.; Blaney, J. M.; Humblet, C.; Gund, P.; Barry, D. C. Molecular Modeling Software and Methods for Medicinal Chemistry. *J. Med. Chem.* **1990**, *33*, 883–894.
(4) Kollman, P. A.; Merz, K. M. J. Computer Modeling of the Interactions of Complex Molecules. *Acc. Chem. Res.* **1990**, *23*, 246–252.
(5) Straatsma, T. P.; McCammon, J. A. Computational Alchemy. *Annu. Rev. Phys. Chem.* **1992**, *34*, 407–435.
(6) Kollman, P. A. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.
(7) van Gunsteren, W. F. Molecular Dynamics Studies of Proteins. *Curr. Opin. Struct. Biol.* **1993**, *3*, 277–281.
(8) Mitchell, M. J.; McCammon, J. A. Free Energy Difference Calculations by Thermodynamic Integration: Difficulties in Obtaining a Precise Value. *J. Comput. Chem.* **1991**, *12*, 271–275.
(9) Dang, L. X.; Merz, J., K. M.; Kollman, P. A. Free Energy Calculations on Protein Stability: Thr-15 → Val-157 Mutation of T4 Lysozyme. *J. Am. Chem. Soc.* **1989**, *111*, 8505–8508.
(10) Silverman, D. N.; Lindskog, S. The Catalytic Mechanism of Carbonic Anhydrase: Implications of a Rate-Limiting Protolysis of Water. *Acc. Chem. Res.* **1988**, *21*, 30–36.
(11) Maren, T. H. Carbonic Anhydrase: Chemistry, Physiology and Inhibition. *Physiol. Rev.* **1967**, *47*, 595–781.
(12) Maren, T. H. Carbonic Anhydrase – General Perspectives and Advances in Glaucoma Research. *Drug Dev. Res.* **1987**, *10*, 255.
(13) Liljas, A.; Håkansson, K.; Jonsson, B. H.; Xue, Y. Inhibition and Catalysis of Carbonic Anhydrase. Recent Crystallographic Analysis. *Eur. J. Biochem.* **1994**, *219*, 1–10.
(14) Smith, G. M.; Alexander, R. S.; Christianson, D. W.; McKeever, B. M.; Ponticello, G. S.; Springer, J. P.; Randall, W. C.; Baldwin, J. J.; Habecker, C. N. Positions of His-64 and a Bound Water in Human Carbonic Anhydrase II Upon Binding Three Structurally Related Inhibitors. *Protein Sci.* **1994**, *3*, 118–125.
(15) Everitt, D. E.; Avorn, J. Systemic Effects of Medications Used to Treat Glaucoma. *Ann. Intern. Med.* **1990**, *112*, 120–125.
(16) Lichter, P. R.; Newman, L. P.; Wheeler, N. C.; Beall, O. V. Patient Tolerance to Carbonic Anhydrase Inhibitors. *Am. J. Ophthalmol.* **1978**, *85*, 495–502.
(17) Arrigg, C. A.; Epstein, D. L.; Giovanoni, R.; Grant, W. M. The Influence of Supplemental Sodium-Acetate on Carbonic Anhydrase Inhibitor-Induced Side Effects. *Arch. Ophthalmol.* **1981**, *99*, 1969–72.
(18) Wallace, T. R.; Fraunfelder, F. T.; Petursson, G. J.; Epstein, D. L. Decreased Libido – A Side Effect of Carbonic Anhydrase Inhibitor. *Ann. Ophthalmol.* **1979**, *11*, 1563–6.
(19) Lewis, R. V.; Lennard, M. S.; Jackson, P. R.; Tucker, G. T.; Ramsay, L. E.; Woods, H. F. Side Effects of Beta-blockers Assessed Using Visual Analog Scales. *Eur. J. Clin. Pharmacol.* **1985**, *19*, 826–8.
(20) Schoene, R. B.; Abuan, T.; Ward, R. L.; Beasley, C. H. Effects of Topical Betaxolol, Timolol and Placebo on Pulmonary Function in Asthmatic Bronchitis. *Am. J. Ophthalmol.* **1984**, *97*, 86–92.
(21) McMahon, C. D.; Shaffer, R. N.; Hoskins, H. D. Jr.; Hetherington, J., Jr. Adverse Effects Experienced by Patients Taking Timolol. *Am. J. Ophthamol.* **1979**, *88*, 736–8.
(22) Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; Randall, W. C.; Schwan, H.; Sugrue, M. F.; Springer, J. P.; Gautheron, P.; Grove, J.; Mallorga, P.; Viader, M. P.; McKeever, B. M.; Navia, M. A. Thienothiopyran-2-sulfonamides: Novel Topically Active Carbonic Anhydrase Inhibitors for the Treatment of Glaucoma. *J. Med. Chem.* **1989**, *32*, 2510.
(23) Prugh, J. D.; Hartman, G. D.; Mallorga, P. J.; McKeever, B. M.; Michaelson, S. R.; Murcko, M. A.; Schwam, H.; Smith, R. L.; Sondey, J. M.; Springer, J. P.; Sugrue, M. F. New Isomeric Classes of Topically Active Ocular Hypotensive Carbonic Anhydrase Inhibitors. *J. Med. Chem.* **1991**, *34*, 1805.
(24) Eriksson, A. E.; Jones, A. T.; Liljas, A. Refined Structures of Human Carbonic Anhydrase II at 2.0Å Resolution. *Proteins* **1988**, *4*, 274–282.
(25) Rao, B. G.; Singh, U. C. Studies on the Binding of Pepstatin and Its Derivatives to Rhizopus Pepsin by Quantum Mechanics, Molecular Mechanics, and Free Energy Perturbation Methods. *J. Am. Chem. Soc.* **1991**, *113*, 6735–6750.
(26) Cummins, P. L.; Ramnarayan, K.; Singh, U. C.; Gready, J. E. Molecular Dynamics/Free Energy Perturbation Study on the Relative Affinities of the Binding of Reduced and Oxidized NADP to Dihydrofolate Reductase. *J. Am. Chem. Soc.* **1991**, *113*, 8247–8256.
(27) Amisaki, T.; Hakashima, T.; Tomita, K.-i.; Nishikawa, S.; Uesugi, S.-i.; Ohtsuka, E.; Ikehara, M.; Yoneda, S.; Kitamura, K. Molecular Dynamics and Free Energy Perturbation Calculations on the Mutation of Tyrosine 45 to Tryptophan in Ribonuclease T_1. *Chem. Pharm. Bull.* **1992**, *40*, 1303–1308.
(28) Prevost, M.; Wodak, S. J.; Tidor, B.; Karplus, M. Contribution of the Hydrophobic Effect to Protein Stability: Analysis Based on Simulations of the Ile-96 → Ala Mutation in Barnase. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10880–10884.
(29) Tidor, B.; Karplus, M. Simulation Analysis of the Stability Mutant R96H of t4 Lysozyme. *Biochemistry* **1991**, *30*, 3217–3228.
(30) Warshel, A. *Computer Modelling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons, Inc.: New York, 1991; p 236.
(31) Merz, K. M., Jr. Computer Simulation of Enzymatic Reactions. *Curr. Opin. Struct. Biol.* **1993**, *3*, 234–240.
(32) Zheng, Y. J.; Merz, K. M., Jr. Mechanism of the Human Carbonic Anhydrase II Catalyzed Hydration of Carbon Dioxide. *J. Am. Chem. Soc.* **1992**, *114*, 10498–10507.
(33) Merz, K. M., Jr.; Kollman, P. A. Free Energy Perturbation Simulation of the Inhibition of Thermolysin: Prediction of the Free Energy of Binding of a New Inhibitor. *J. Am. Chem. Soc.* **1989**, *111*, 5649–5658.
(34) Merz, K. M., Jr.; Murcko, M. A.; Kollman, P. A. Inhibition of Carbonic Anhydrase. *J. Am. Chem. Soc.* **1991**, *113*, 4484–4490.
(35) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. In University of California: San Francisco, 1991.
(36) Nicholas, J. B.; Vance, R.; Martin, E.; Burke, B. J.; Hopfinger, A. J. A Molecular Mechanics Valence Force Field for Sulfonamides Derived by *ab initio* Methods. *J. Phys. Chem.* **1991**, *95*, 9803–9811.
(37) Stewart, J. J. P. MOPAC 5.0. *Quantum Chemistry Program Exchange 455* **1986**, *6*.
(38) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
(39) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431–439.
(40) Dewar, M. J. S.; Thiel, W. Ground States of Molecules. 38. The MNDO Method, Approximations and Parameters. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
(41) Smith, G. M. Unpublished results.
(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for the Simulation of Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926935.
(43) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230–252.
(44) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A New Force Field of Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
(45) 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. Comput. Phys.* **1977**, *23*, 327–341.
(46) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(47) Arnett, E. M.; Jones, F. M.; Taagepera, M.; Henderson, W. G.; Beauchamp, J. L.; Holtz, D.; Taft, R. W. A Complete Thermodynamic Analysis of the "Anomalous Order" of Amine Basicities in Solution. *J. Am. Chem. Soc.* **1972**, *94*, 4724–4726.

JM940757Z