Binding Preferences of Hydroxamate Inhibitors of the Matrix Metalloproteinase Human Fibroblast Collagenase

Samuel Toba, K. V. Damodaran, and Kenneth M. Merz, Jr.*

152 Davey Laboratory, Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

Received October 12, 1998

In this paper we report molecular dynamics (MD) and free energy perturbation (FEP) studies carried out on enzyme−inhibitor (two hydroxamates that only differ by a carbon−carbon double bond) complexes of human fibroblast collagenase to obtain insights into the structural and energetic preferences of these inhibitors. We have developed a bonded model for the catalytic and structural zinc centers (Hoops, S. C.; et al. *J. Am. Chem. Soc.* **1991**, *113*, 8262−8270) where the electrostatic representation for this model was derived using a novel quantum-mechanical/molecular-mechanical (QM/MM) minimization procedure followed by electrostatic potential fitting. The resulting bonded model for the zinc ions was then used to generate MD trajectories for structural analysis and FEP studies. This model has satisfactorily reproduced the structural features of the active site, and furthermore, the FEP simulations gave relative free energies of binding in good agreement with experimental results. MD simulations in conjunction with the FEP are able to provide a structural explanation regarding why one hydroxamate inhibitor is favored over the other, and we are also able to make predictions about changes in the inhibitor that would enhance protein−inhibitor interactions.

Introduction

Matrix metalloproteinases (MMPs) are a large family of enzymes that are involved in the degradation of the components of the extracellular matrix.[2,3] They are secreted as zymogens, which are activated by proteolytic cleavage of a ∼10 kDa proenzyme region.[2,3] The active enzyme can cleave one or more components of the extracellular matrix. For example, collagenases cleave specific collagens, and stromelysin cleaves gelatins, fibronectins, etc.[2,3] The activity of MMPs is regulated by what is known as tissue inhibitors of metalloproteinases (TIMPs).[2,3] However, under stoichiometric imbalance or weakly regulated conditions, MMPs can cause tissue damage or modifications.[2,3] This damage can give rise to disease conditions such as rheumatoid or osteoarthritis,[2] metastasis of cancer cells,[4] periodontal disease,[2] or corneal ulceration.[5] MMPs are also involved in many of the bodily functions associated with growth processes, such as embryogenesis, bone remodeling, and tooth eruption and in mediation of blood platelet aggregation.[6] The far-reaching impact of MMPs further includes adamalysins which are components of snake venom.[7] Hence, the molecular-level insights obtained on MMP inhibition in this study will have an impact on the future rational design of MMP inhibitors that could be useful in the treatment of the disease states described above.[8]

At the molecular level, MMPs are characterized by a zinc atom at the active site (zinc catalytic site) with a conserved zinc binding motif, HExxHxxGxxH. The proenzyme region also consists of a cystine residue which is conserved in MMPs, and in the inactive form this cystine is bonded to the catalytic zinc. This class of enzymes also contains a methionine residue below the active site zinc, which forms part of a family-wide superimposable "Met-turn".[3,9] Given that the X-ray structures of a number of MMPs have been determined recently,[2,3] it is possible to generate a generalized catalytic mechanism based on the residues that are located close to the catalytic zinc ion.[2,3] A proposed mechanism for the catalytic activity based on these crystal structures has many features in common with that of thermolysin,[10] although there are differences in the metal ion environments and substrate binding interactions.[2,3] The mechanism given in Figure 1 leaves out many of the residues within the active site since the role many of these play in MMP catalysis is uncertain at this time. In the proposed catalytic mechanism the peptide is bound into the active site (via a zinc−carbonyl interaction as well as other noncovalent peptide−active site interactions) and an active site water molecule transfers a proton to Glu 219, thereby generating a tetrahedral intermediate (see **1** → **2**). This Glu residue then transfers its proton to the amide nitrogen of the peptide (**3**) which facilitates the collapse of **3** into two peptide fragments (**4**) based on the N- and C-terminal fragments of the starting protein. The peptide fragments are then lost to generate the resting state on the enzyme (**5**). Numerous open questions remain regarding this catalytic cycle (e.g., is intermediate **3** actually stable or does it directly collapse to **4** once intermediate **2** has a proton transferred to it?), but this is a good working mechanism until further corroborating data is obtained.

Since MMPs are involved in a wide range of diseases, there has been intense interest in obtaining effective small-molecule inhibitors. Recently, a number of pseudopeptide inhibitors have been reported in the literature.[11] Generally speaking, these inhibitors are characterized by a zinc binding group on one end attached to varying numbers of small-molecule or 'peptide-like' substituents, each of which has to be optimized in order to get optimal

10.1021/jm980577f CCC: $18.00
Published on Web 03/16/1999

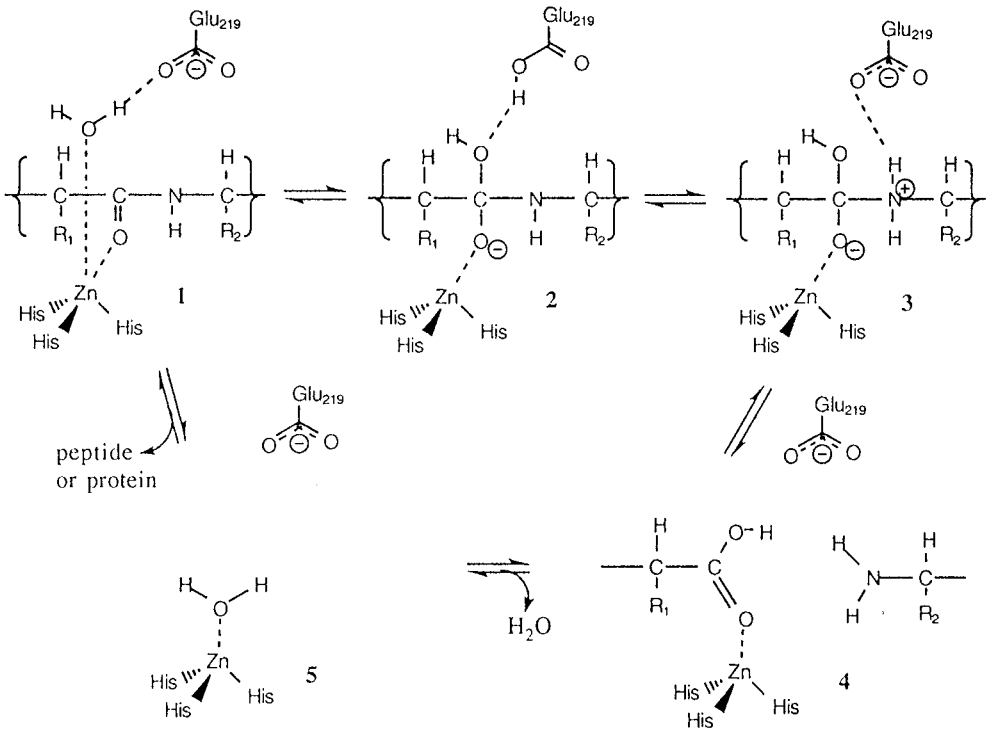

Figure 1. Postulated catalytic mechanism for MMPs.

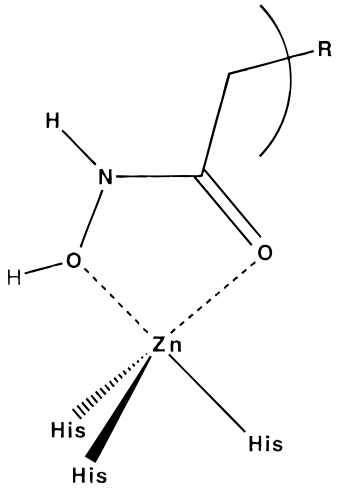

Figure 2. Both hydroxamate oxygens and the three liganding histidine nitrogens of the enzyme make up a trigonal-bipyramidal coordination sphere around the catalytic zinc.

efficacy. Of these, inhibitors with the hydroxamate zinc binding group have been found to be potent inhibitors in MMPs such as human fibroblast collagenase (HFC)[12] and human neutrophil collagenase (HNC).[13] Inhibitors with other zinc binding groups such as carboxylate, amino carboxylate, and sulfydryl have also been synthesized.[14] High-resolution X-ray structures of the enzyme−inhibitor complexes have also become recently available.[2,3] The crystal structures of HNC and HFC complexed with hydroxamate inhibitors have revealed that the catalytic zinc is pentacoordinated (3 His + 2 oxygens from the inhibitor zinc binding group; see Figure 2). [12] The hydroxyl oxygen is at a slightly larger distance from the zinc than the other ligands. It is worth noting that hydroxamate inhibitors were found to have a different mode of binding in human carbonic anhydrase-II, where the terminal nitrogen became ionized and bonded directly with the zinc.[15] The protein also contains a second zinc ion in tetrahedral coordination (zinc structural site) and a calcium ion in octahedral coordination. These metal ions are necessary for the structural stability of the enzyme.[16]

Because of the presence of zinc ions (one catalytic and one structural) in MMPs, we are faced with selecting a way in which to model these ions. There are two basic ways to model metal ions using a purely classical potential function that has been described in the literature: the bonded model[1] and the nonbonded model.[17,18] In this manuscript we have adopted the bonded approach for metal ion representation, which involves placing explicit bonds between a metal ion and its surrounding ligands.[1] In the nonbonded approach, nonbonded electrostatic and van der Waals terms are used to retain the metal−ligand coordination.[17,18] The major limitations of the bonded approach are its inability to allow for coordination changes around the metal ion during the dynamics and the limited metal ion−ligand flexibility conferred due to the presence of the metal−ligand bonds. On the other hand, the nonbonded approach is very sensitive to the electrostatic model chosen[17] and can suffer from the inability to retain low coordination number.[1] We note, however, that the nonbonded model for zinc has recently been used successfully in carbonic anhydrase and carboxypeptidase by Stote and Karplus[18] using a specially designed cutoff scheme with stochastic boundary conditions and the CHARMM force field. However, with the AMBER force field we find that the nonbonded approach generally fails to give the correct coordination number even when long-range electrostatic interactions are correctly accounted for using an infinite cutoff. Thus, given our past success with the bonded model,[1,19,20] we have decided to continue its use. In addition, with the presence of zinc in two different coordination environments in MMPs, we felt that the use of a bonded

approach would most accurately reproduce the HFC−zinc−inhibitor system in our simulations.

In this paper we report molecular dynamics (MD) and free energy perturbation (FEP) studies carried out on enzyme−inhibitor complexes of human fibroblast collagenase to obtain insights into the structural and energetic preferences of these inhibitors. We have developed a bonded model for the catalytic and structural zinc centers[1] where the electrostatic representation for this model was derived using a novel procedure in which a quantum-mechanical/molecular-mechanical (QM/MM) minimization procedure was followed by electrostatic potential fitting. This approach has the advantage in that it includes environmental polarization effects, while in the previous incarnation of this model[1] all charges were evaluated in the absence of the enzyme or aqueous environment (i.e., the charges were determined in the gas phase). The resulting bonded model for the zinc ions was then used to generate MD trajectories for structural analysis and FEP studies. This approach has satisfactorily reproduced the salient structural features of the active site. Furthermore, the FEP simulations gave relative free energies of binding in good agreement with experimental results.

Experimental Section

Crystal structures and binding information (i.e., K_i's) were obtained from experimentally published results. The crystal structure of HFC obtained by Spurlino et al.[12] contains a hydroxamate inhibitor with a Phe side chain at the P_2' position. The active enzyme, after the cleavage of the proenzyme (42 kDa), is unstable in vitro. Hence the crystal structure was determined using a "matured, truncated" form (19 kDa) of the enzyme without the hemopexin domain present at the C-terminus. This matured, truncated enzyme is also catalytically active.[12] The active site consists of three histidines (His 218, 222, 228) bound to the catalytic zinc and a glutamate (Glu 219) hydrogen bound to the hydroxyl hydrogen of the inhibitor. The structural zinc site is tetrahedral, coordinated to three histidines (His 168, 183, 196) and an aspartate (Asp 170) bound to the metal ion in a monodentate manner. The Ca^{2+} present is coordinated to the carbonyl oxygens of Gly 176, Gly 178, and Asn 180 and the side chain oxygens of Asp 175, Asp 198, Glu 201 in an octahedral manner. The latter two metal ion centers are required for the stability of the HFC enzyme.

In the present work we have investigated succinyl hydroxamate inhibitors with indole side chains at the P_2' position and have concentrated on studying the effects of mutations at the P_1' position (positions as defined in Beckett et al.[8]). These inhibitors and their experimentally observed inhibition constants are shown in Figure 3a−c.[8,21] The starting structure for the simulation given in Figure 3b was obtained by replacing the Phe side chain with an indole group with the help of the graphics program MIDAS[22] (i.e., from OH-NH-CO-Leu-**Phe**-CO-NH-CH_3 to OH-NH-CO-Leu-**Trp**-CO-NH-CH_3). This inhibitor, which we refer to as IND1, is better known by the tradename Galardin[5] developed by Glycomed Inc. Galardin is reported to be undergoing phase II/III clinical trials for treatment of patients with corneal ulcers[5] and has been extensively used as a lead molecule in the study of small-molecule MMP inhibitors.[23]

The conformation of the protein was not altered during this model-building process, and the inhibitor structure was altered only minimally. The starting structure for IND2 (see Figure 3c) was obtained from the minimized structures (see below for more details) of IND1. In all the simulations described in this work, we have retained all the water molecules observed in the X-ray structure.[12] We then added a water cap with a radius of 22 Å (from the zinc ion) to ensure that the inhibitor and the surface region near the active site were fully solvated.

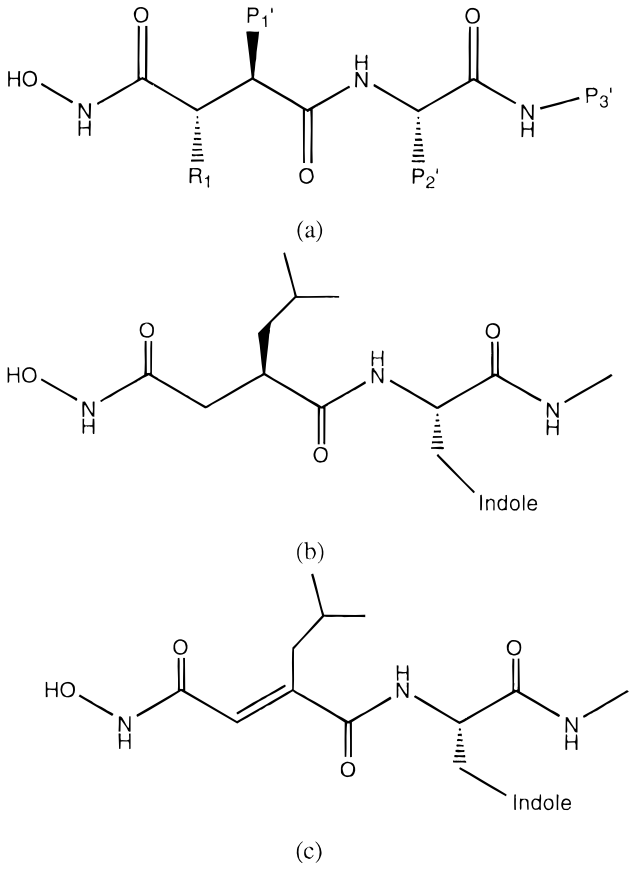

Figure 3. (a) Succinyl hydroxamate inhibitor; (b) inhibitor IND1 (K_i = 0.002 μM); (c) inhibitor IND2 (K_i = 0.13 μM).

Table 1. Derived Force-Field Parameters for the Zinc Ions and the Hydroxamate Group Associated with MMP Inhibitors

Bond Parameters		
bond	K_r (kcal/Å^2)	R_{eq} (Å)
Zn−NB	40.0	2.05
Zn−OH	40.0	2.20
Zn−O	40.0	2.05
N−O	539.0	1.37
Angle Parameters		
angle	K_θ (kcal/rad^2)	θ_{eq} (deg)
NB−Zn−NB	20.0	105.0
NB−Zn−O	20.0	115.0
CR−NB−Zn	20.0	126.0
N−OH−HO	83.6	105.6
OH−N−C	137.4	116.1
H−N−OH	94.8	109.1
Dihedral Parameters		
HO−OH−N−C	$v/2 = 3.0$, $\gamma = 180$, $n = 1$	

Counterions were placed near the five glutamate residues on the surface of the protein. All residues within 15 Å radius of the two zinc ions as well as all the cap water molecules were allowed to move during the dynamics.

Force Field. It was necessary to develop the force-field parameters associated with the N−O functionality for the hydroxamate inhibitors used in this investigation since these parameters were not available in the AMBER[24−26] force field. Acetohydroxamate ($CH_3C(O)NHOH$) was used as a small-molecule model system for the hydroxamate functionality of the inhibitors. Using the HF/6-31G* ab initio level, appropriate bond, angle, and dihedral parameters were determined. The resulting parameters are reported in Table 1. Bond and angle parameters associated with the zinc center were taken from

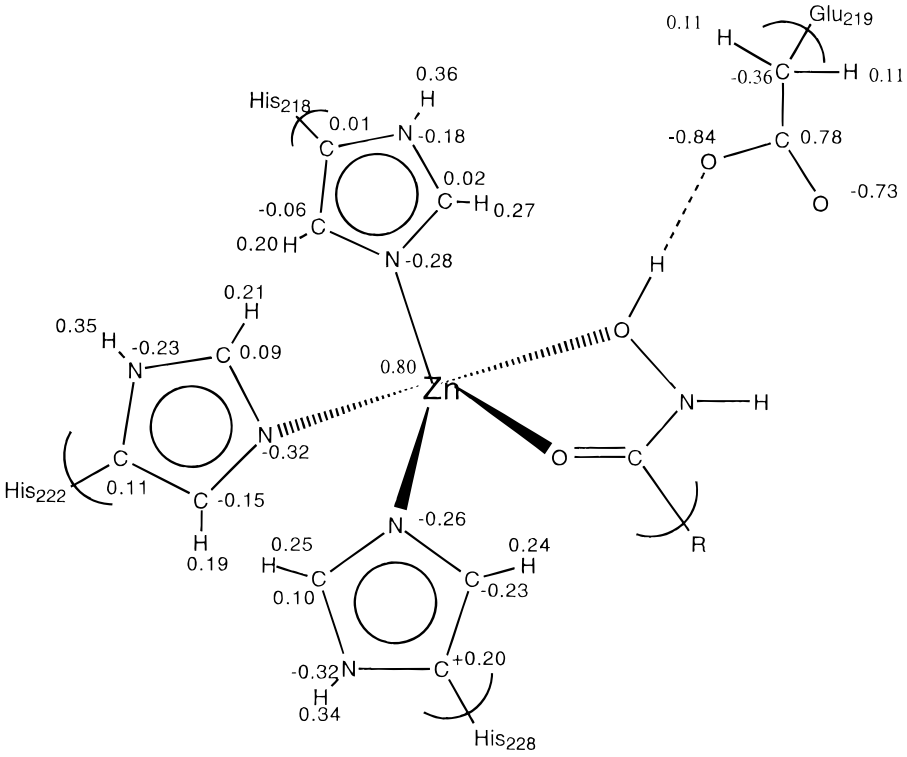

Figure 4. Partial charges for the HFC active site complexed with IND1.

Hoops et al.[1] All the torsions associated with the zinc−ligand bonds were set to zero as in Hoops *et al.*.[1]

Partial Charges. In addition to the bond, angle, and dihedral parameters, the different coordination geometries of the two zinc sites required the development of appropriate charge representations. Since a nonbonded metal−ligand representation involving a formal charge (+2) often leads to octahedral coordination for Zn, we have adopted and extended the bonded model approach designed for metalloproteins in our laboratory.[1] In our new approach, we have derived the partial charges for the zinc ions and their ligands using a coupled QM/MM[27] energy minimization followed by electrostatic potential fitting (ESP).[28] The important new feature of this approach, is the inclusion of polarization effects arising from the surrounding environment. In the previous incarnation of this approach the ESP charges were determined in the gas phase and not directly in the local environment.[1] In the QM/MM energy minimization, the zinc ions and their coordinating ligands were treated using a quantum-mechanical (the semiempirical PM3 Hamiltonian) representation while the rest of the protein−solvent environment was treated using a molecular-mechanical (AMBER) representation. For the catalytic region we included three histidine side chains, one glutamate side chain, the entire inhibitor, and the zinc ion. Similarly, the structural zinc ion along with its three histidine side chains and one aspartate side chain were treated using a QM representation. The junction between the QM and MM regions was made between C-β and C-γ of the His residues and between C-γ and C-δ for the Glu residue. When a residue was split between the QM and MM regions, the partial charges on the MM atoms of that residue were modified to maintain the charge neutrality of the MM region.[27] The QM/MM energy minimization involved 200 steps of steepest descent minimization and 800 steps of conjugate gradient minimization. The distances from the coordinating atoms to Zn (e.g., the NE2 nitrogen from the histidines and the hydroxyl and carbonyl oxygens at the hydroxamate-terminus of the inhibitor) were constrained at their crystallographic values. This was done because an unconstrained QM/MM minimization resulted in a local geometry that was in unsatisfactory agreement with the X-ray structure. In particular, the distance between the hydroxyl oxygen of the inhibitor to the catalytic zinc ion was too long. The PM3 Hamiltonian[29−31] was used during the QM/MM energy minimization, while the MNDO Hamiltonian was used for the ESP calculations. The PM3 Hamiltonian yields better structure[29−31] yet produces charges that are in poor agreement with ab initio charges.[28] MNDO, on the other hand, was found to be superior in charge determination,[28] and, therefore, was used in the ESP fitting calculations.[1] The net charges of the QM regions around the catalytic site or the structural site were both set to +1 (Zn = +2, inhibitor = 0, histidines = 0, glutamate/aspartate = −1) during the QM/MM energy minimization. The QM/MM minimized structure for the enzyme−inhibitor complex with the inhibitor IND1 was used to set up the models for IND2. Two hydrogens in the IND1 structure near the P_1' position were removed to generate the carbon−carbon double bond for IND2. The partial charges for the active site QM atoms were then calculated for this system using the same method as described above. The partial charges used in the enzyme MD simulations are summarized in Figures 4−7.

The partial charges for the inhibitors in an aqueous solvent environment were also determined using the coupled potential approach. Atomic point charges for the inhibitors could have been determined in the gas phase as is typical practice;[24−26] however, this would lead to an unbalanced model[32] between the enzyme and aqueous-phase simulations. In the enzyme case we are utilizing charges that include polarization arising from the surrounding environment; thus, it is also important to include these effects in the aqueous case. Moreover, inclusion of conformational sampling effects (via MD simulations) was also regarded as being more important in solution than in the relatively rigid enzyme active site. The IND1 and IND2 inhibitor structures were initially optimized in the gas phase using PM3 semiempirical calculations (both inhibitors are neutral). Gas-phase partial charges were then determined using ab initio single-point calculations at the HF/6-31G* level. These calculated charges were then used for MD simulations of the inhibitors solvated in a 22 Å cap of water molecules, for 150 ps. Partial charges were calculated using the QM/MM ESP[28] fitting procedure with the MNDO[33,34] Hamiltonian on configurations saved every 15 ps, and the resulting charges were averaged. The observed fluctuations in the charges calculated in this manner were observed to be relatively small, with the largest rms fluctuation being only ~0.1e. The result-

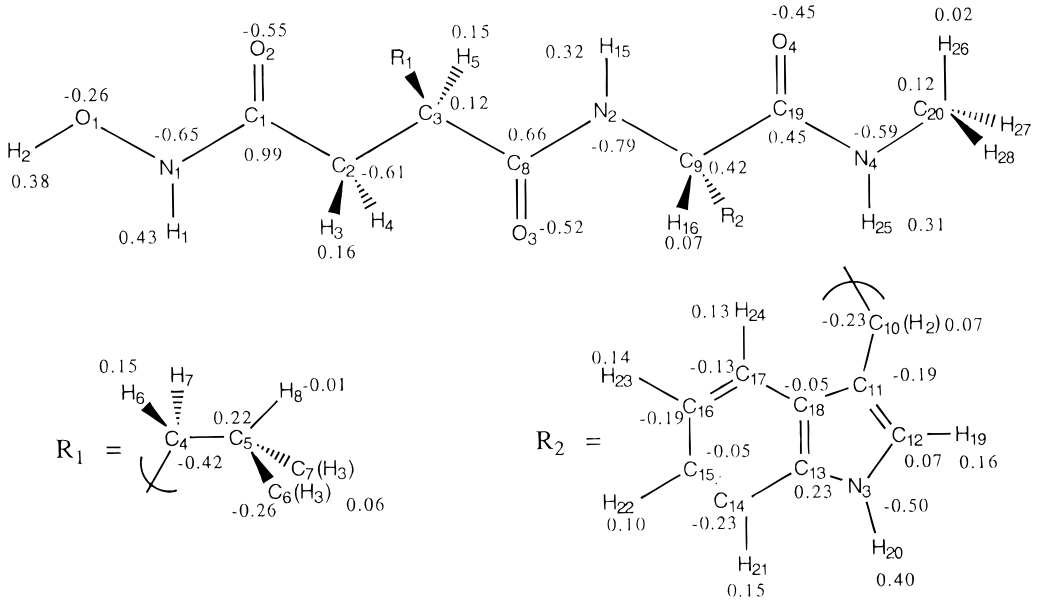

Figure 5. Partial charges for the hydroxamate inhibitor IND1. In cases where symmetry-related atoms (e.g., hydrogens in a methyl group) are present, only one atomic charge is given. The numbering scheme given here is also used for IND2 except H4 and H5 are deleted.

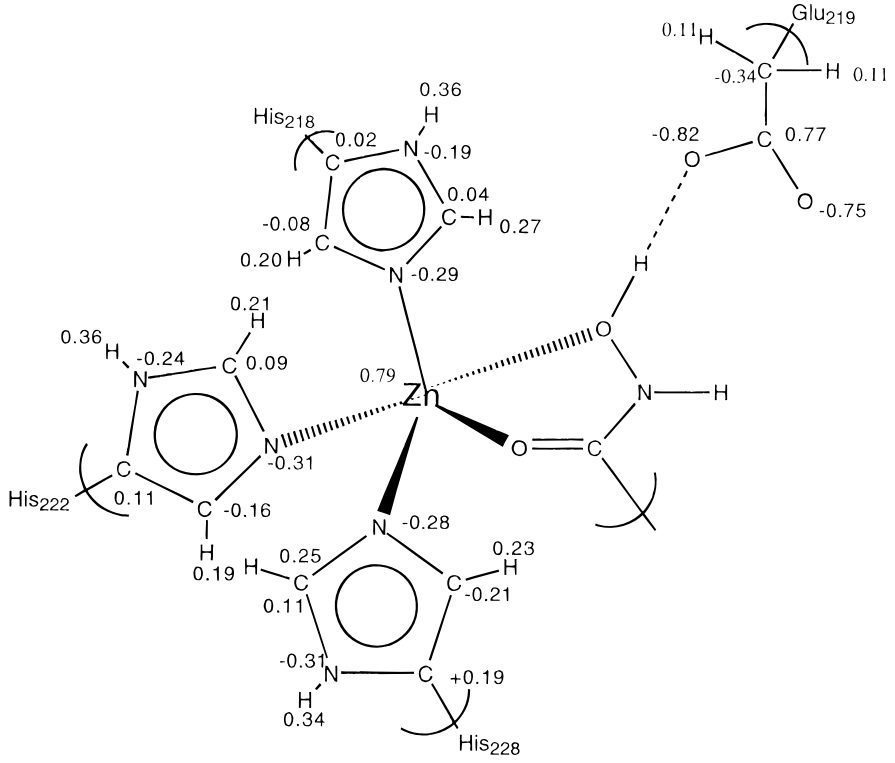

Figure 6. Partial charges for the HFC active site complexed with IND2.

ing partial charges used in the aqueous MD simulations are summarized in Figures 8 and 9.

Dynamics. Using the force-field parameters and the partial charges derived above, we obtained a minimized structure that was in good agreement with the X-ray structure with an rms deviation of only 0.35 Å. Hence, we decided that our force-field parameter set was able to represent the protein system well enough to begin our MD and FEP simulations. To set up the starting structure for the simulations described herein, we utilized the minimized structure of the HFC complexed with the IND1 inhibitor as our reference. MD simulations were carried out using a residue-based cutoff of 10 Å and a time step of 1 fs. All the bonds involving hydrogens were constrained using SHAKE,[35] while the bonds involving heavy atoms were included in the dynamics. The model was equilibrated by gradually increasing the temperature in stages (5 ps at 100 K, 5 ps at 200 K) to the target temperature of 300 K by coupling to a temperature bath.[36] After 190 ps of MD simulations, the rms deviation from the initial structure had stabilized with a value of ~0.9 Å. Atomic coordinates were collected every 0.05 ps for 300 ps for the two systems.

FEP simulations were carried out using the GIBBS module of the AMBER (version 4.0)[37] suite of programs. We have employed the "slow growth" approach for the determination of the free energy. This method has been used successfully in several previous application studies.[20,38] The relative free energy between IND1 and IND2 was obtained by mutating the CH_2R-CHR_2 linkage to $CHR=CR_2$, in which two hydrogens disappeared and the carbons were mutated from CT (sp^3) to CM (sp^2) types. We have carried out the FEP simulations

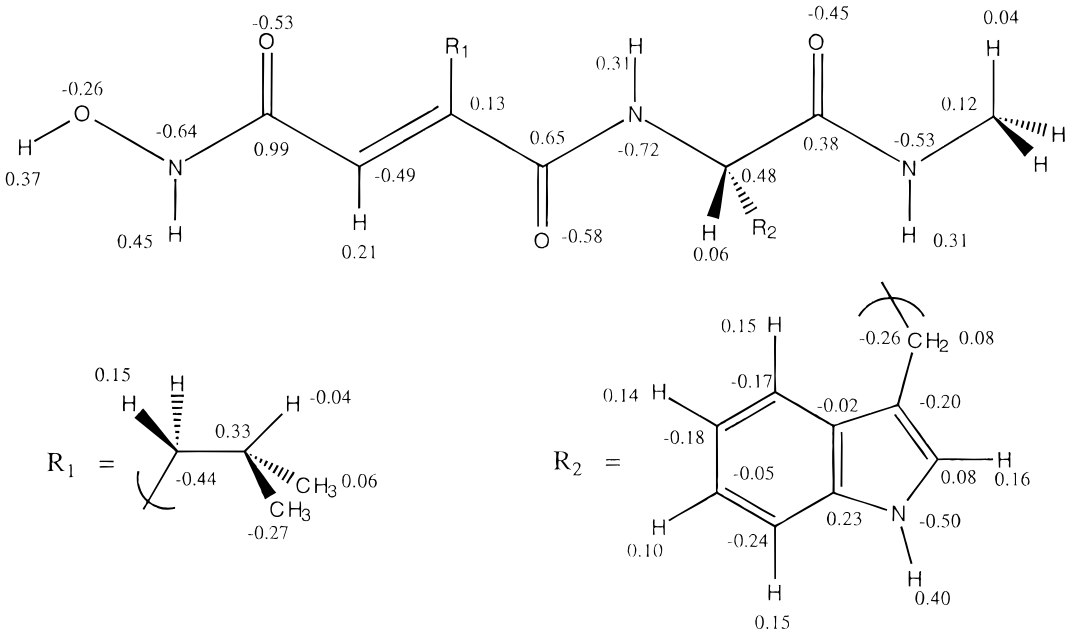

Figure 7. Partial charges for the hydroxamate inhibitor IND2. In cases where symmetry-related atoms (e.g., hydrogens in a methyl group) are present, only one atomic charge is given.

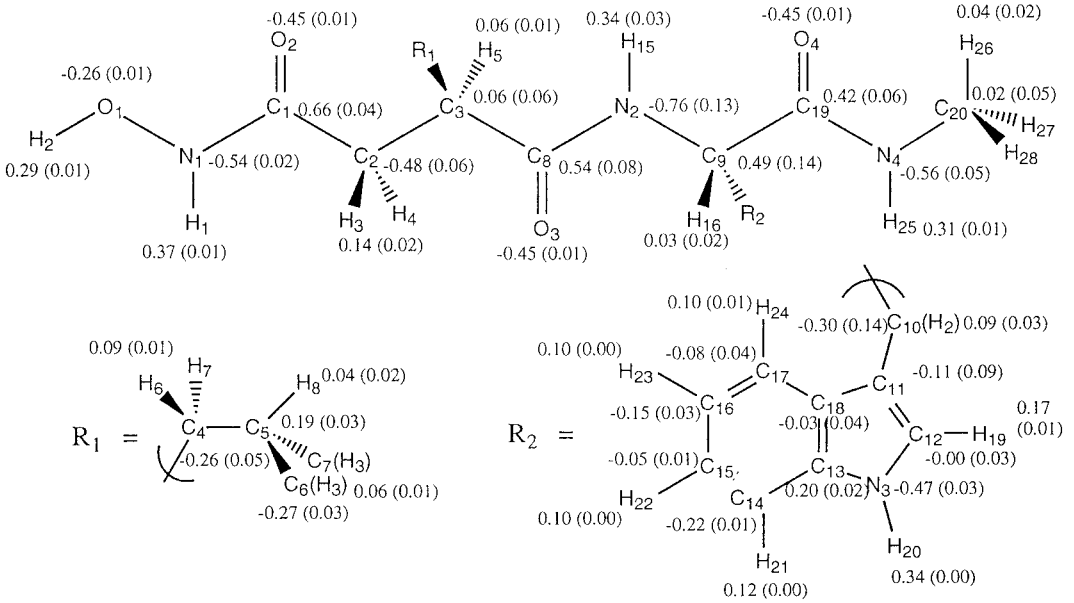

Figure 8. Partial charges for the hydroxamate inhibitor IND1 in solvent. The rms fluctuations in the charges are given in parentheses. In cases where symmetry-related atoms (e.g., hydrogens in a methyl group) are present, only one atomic charge is given.

using a time scale of 500 ps. To investigate the effect of sampling on the computed free energy values, we first carried out 190 ps equilibration runs followed by 500 ps FEP simulations in the enzyme and in solution. For the enzyme case we ran a total of five simulations that started from temporally separated (150 ps each after the initial 190 ps equilibration phase) starting configurations. For the solution runs the variation among the first four temporally separated FEP simulations was so low (±0.13 kcal/mol) that a fifth run was not carried out.

Results and Discussions

The results of the MD and FEP simulations will be divided into two categories: energetic and structural. First, we shall discuss the energetic portion of our simulations. The FEP runs were performed on inhibitors IND1 and IND2 with and without (i.e., in aqueous solution) the protein HFC. The calculations were run from IND1 → IND2 only starting with the equilibrated structures of the enzyme−substrate complex. All runs were done in the forward direction five times (for the enzyme, only four runs were done in the aqueous case) to obtain insight into the statistical errors present in the calculation. The ΔG vs λ plot for the IND1 → IND2 mutation was linear and showed very good convergence both in the solvent and in the enzyme−inhibitor complex. The results of the FEP simulations are listed in Table 2. The calculated average relative free energy of binding ($\Delta\Delta G$) was 3.9 kcal/mol, which is in reasonable agreement with the experimental value of 2.5 kcal/mol. Thus, both experiment and theory predict that IND1 is a better inhibitor of HFC than is IND2.

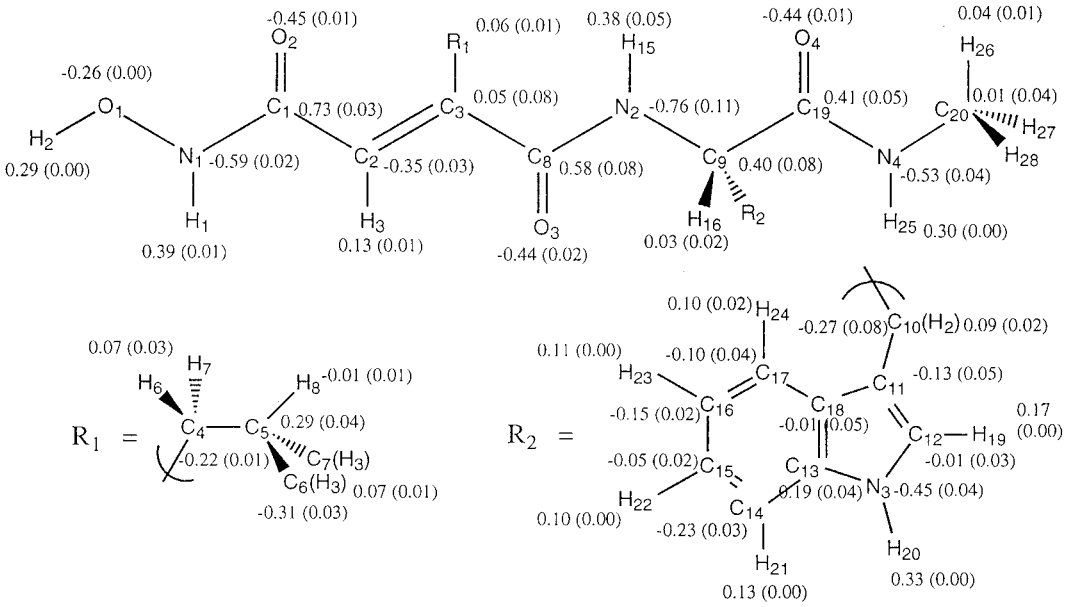

Figure 9. Partial charges for the hydroxamate inhibitor IND2 in solvent. The rms fluctuations in the charges are given in parentheses. In cases where symmetry-related atoms (e.g., hydrogens in a methyl group) are present, only one atomic charge is given.

Table 2. Results Obtained from the FEP Simulations for IND1 → IND2

starting structure	ΔG_{enz} (kcal/mol)	ΔG_{sol} (kcal/mol)	$\Delta\Delta G_{bind}$[a] (kcal/mol)
190 ps	97.0	102.0	
340 ps	98.8	102.2	
490 ps	96.9	102.2	
640 ps	99.7	101.9	
790 ps	98.7		
averages	98.2 ± 1.1	102.1 ± 0.13	3.9

[a] Experimental value: 2.5 kcal/mol.

Another important component of this research is the structural analysis which allows us to garner chemical insights into why IND1 is a better inhibitor of HFC that IND2. The inhibitors IND1 and IND2 both contain the hydrophobic side chains of Leu and Trp interacting with the S_1' and S_2' pockets of the active site, respectively. The Leu side chain of the inhibitor is in the protein interior, whereas the Trp side chain of the inhibitor is solvent-exposed (as was the Phe side chain in the original X-ray structure of Spurlino et al.[12]). There was a significant amount of movement of the inhibitors in the active site during the simulations, but in both cases it was confined primarily to the Trp side chain of the inhibitors. Most of the effective MMP inhibitors have a hydrophobic group in the P_1' position (e.g., the isobutyl group in IND1 and IND2) and take advantage of the favorable hydrophobic interactions provided by Val 215 in the enzyme active site. Hydrophilic interactions such as hydrogen bonds between the inhibitor amide protons and the enzyme have been reported to be extremely important in stabilizing the inhibitor in the active site also.[21,23] Both IND1 and IND2 form approximately 5−6 stable hydrogen bonds to the protein, and their distances are given in Table 3. Overall, we find that all the protein−inhibitor backbone hydrogen bond interactions observed in the crystal structure have been maintained throughout in the simulations (see Table 3). Besides the protein−inhibitor interactions, we also find that our model retains the metal (Zn and Ca) distances quite well (note that these distances were not constrained by SHAKE).

Table 3. Selected Metal−Ligand and Inhibitor−Protein Hydrogen-Bonding Distances Calculated from the MD Trajectories[a,b]

ligand	IND1	IND2	X-ray[c]
Catalytic Zn Site			
NE2 His 218	1.97(0.08)	2.00(0.08)	2.00
NE2 His 222	2.04(0.07)	2.02(0.07)	2.13
NE2 His 228	2.04(0.07)	2.04(0.08)	2.07
O1 INH	2.21(0.07)	2.22(0.08)	2.17
O2 INH	2.09(0.08)	2.08(0.08)	2.08
Structural Zn Site			
NE2 His 168	2.05(0.09)	2.06(0.09)	2.09
OD2 Asp 170	1.80(0.07)	1.83(0.06)	1.88
NE2 His 183	2.04(0.09)	2.07(0.08)	2.03
NE2 His 196	2.13(0.09)	2.11(0.08)	2.08
Calcium Site			
OD1 Asp 175	2.28(0.05)	2.29(0.05)	2.30
O Gly 176	2.35(0.06)	2.37(0.06)	2.33
O Gly 178	2.35(0.06)	2.33(0.06)	2.44
O Asn 180	2.35(0.06)	2.37(0.07)	2.39
OD2 Asp 198	2.26(0.04)	2.25(0.04)	2.32
OE2 Glu 201	2.27(0.05)	2.27(0.05)	2.22
Inhibitor−Protein H-Bond Interactions			
N1−O Ala 182	3.02(0.18)	3.23(0.22)	2.83
N1−OE2 Glu 219	3.05(0.24)	2.93(0.18)	3.00
O1−OE1 Glu 219	2.59(0.08)	2.62(0.09)	2.50
O3−N Leu 181	2.88(0.11)	2.86(0.10)	2.81
N2−O Pro 238	3.00(0.14)	3.04(0.18)	3.10
O4−N Tyr 240	2.95(0.16)	2.97(0.17)	2.84
N4−O Gly 179	2.91(0.14)	3.02(0.20)	2.80

[a] Results averaged over 300 ps of MD trajectory. [b] The rms fluctuations are given in parentheses. [c] Values from Spurlino et al.[12]

The hydroxamate zinc binding group acts as a bidentate ligand with each oxygen (O1 and O2) situated at an optimum distance (1.9−2.3 Å) from the active site zinc ion. Our simulations have shown that both the oxygens at O1 and O2 positions retain their stable ligand interaction with the catalytic zinc ion (the distances are shown in Table 4) close to the crystal structure average value of 2.09 Å. Moreover, the zinc

Table 4. Select Bond and Angle Values between IND1 and IND2 and the Enzyme Active Site

atoms	IND1	IND2
O Ala 182−H1 IND	2.1 ± 0.2 Å	2.6 ± 0.3 Å
OE2 Glu 219−H1 IND	2.5 ± 0.4 Å	2.2 ± 0.2 Å
OE2 Glu 219−H1 IND−N1 IND	117.0 ± 15.0°	132.1 ± 11.0°
O Ala 182−H1 IND−N1 IND	146.8 ± 15.0°	128.2 ± 12.2°

ion remained pentacoordinated with the two oxygens and the three histidines in the active site. The hydroxamate nitrogen is thought to be protonated (H1), and it forms a hydrogen bond with the backbone carbonyl oxygen of Ala 182.[12] From our simulations we have noted a change in the mode of interaction of the zinc binding group of IND1 and IND2 with the active site of the enzyme. In the IND1 simulations, H1 has a short hydrogen bond with the carbonyl oxygen at Ala 182 (2.1 ± 0.2 Å) and a longer one with the carboxylate oxygen (OE2) of Glu 219 (2.5 ± 0.4 Å) as is observed in the crystal structure.[12] However, the opposite is observed in the IND2 simulations: The hydroxamate H1 in IND2 favors a short hydrogen bond with the carboxylate oxygen (OE2) of Glu 219 (2.2 ± 0.2 Å) and a longer one with the carbonyl oxygen of Ala 182 (2.6 ± 0.3 Å). This observation is a direct result of mutating IND1 → IND2 where the backbone atoms shift to such an extent to allow the hydroxamate hydrogen to favor interaction with the carboxylate oxygen of Glu 219. To further examine these two hydrogen bond interactions, we examined the angle between the hydrogen bond donors and acceptors (recall that the optimum value would be 180°). The angle values (Table 4) for these hydrogen bonds further support the presence of a "strong" hydrogen bond between Ala 182 and IND1 (hydrogen bond angle of ∼147°) and a "weak" hydrogen bond between Glu 219 and IND1 (hydrogen bond angle of ∼117°). In IND2, both these hydrogen bonds were retained, but both of the angles are far away from the optimal value (∼132° and ∼128°, respectively). From this analysis we propose that the observed differences in these two hydrogen bonds partly account for the tighter binding of IND1. Glu 219 and Ala 182 have also been proposed as the residues responsible in promoting the nucleophilic attack of a water molecule on the scissile peptide bond and the residue in stabilizing the protonated nitrogen of the scissile bond, respectively.[12]

The FEP trajectories further support the observed changes in the protein−inhibitor hydrogen bonds as revealed in the independent MD trajectories described above. Figure 10a,b clearly shows the crossover, at a particular λ (i.e., the FEP coupling parameter) value, for the two hydrogen bonds as the inhibitor is mutated from IND1 to IND2 (1 → 0). Although this is not the only contribution to the tighter binding between IND1 and the protein, we feel that this interaction plays a crucial role in making IND1 the tighter binding inhibitor.

Another major determinant of selectivity is the P_1' substituent of the inhibitor which is the isobutyl group in the IND1 and IND2. Nonpolar alkyl groups have been shown to be the preferred substituents for the P_1' position in HFC inhibitors.[8] This group is capped by the S_1' pocket consisting of the side chains of Val 215 and His 218 along with the backbones of Ser 239 and Tyr

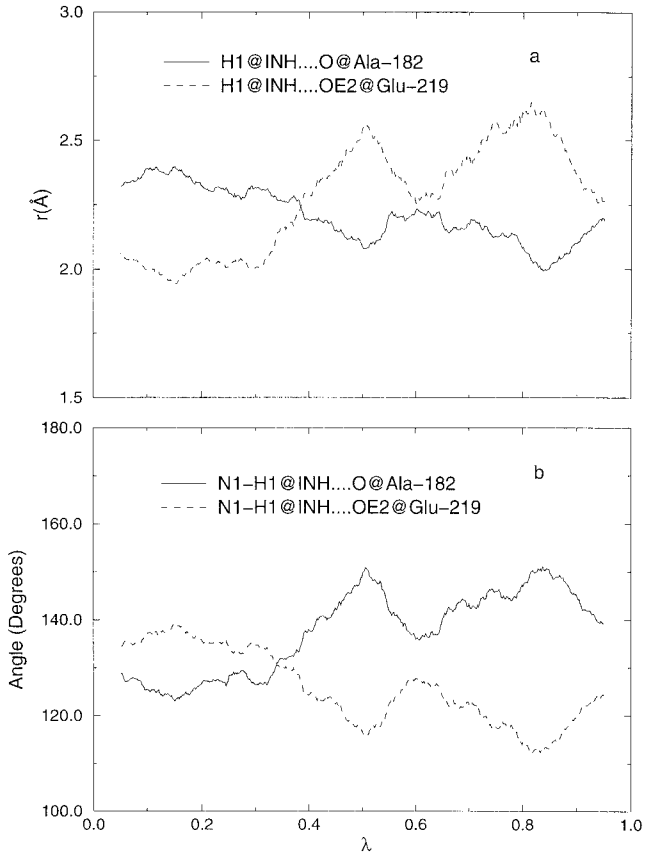

Figure 10. (a) Running averages of the H1 IND−O Ala 182 and the H1 IND−OE2 Glu 219 distances; (b) running averages of the N1 IND−H1 IND−O Ala 182 and N1 IND−H1 IND−OE1 Glu 219 angles.

240, and at the bottom of this pocket (∼7 Å away from the alkyl side chain of the inhibitors) is the side chain of Arg 214. A methyl group (C7) of the Leu is well-capped by Val 215 forming van der Waals interactions. The positions of these capping residues, particularly Val 215, make long alkyl or phenyl alkyl chains an unfavorable choice for the P_1' inhibitor side chain. The other methyl group in the Leu side chain (C6) makes van der Waals contacts with His 218 and with the carbonyl of Pro 238, but these hardly appear to be optimal. Although the S_1' pocket does not tolerate charged and polar groups well in the methyl C7 position, we feel that the C6 methyl group can be replaced with a small polar group without adversely affecting the inhibition. Rather, a small polar group replacement for C6 may result in favorable electrostatic interactions with the other polar residues found on this side of the S_1' pocket (e.g., His 218 and Pro 238). In particular, a hydroxyl group could form a hydroxyl oxygen stacking interaction with the π-face of His 218 (note that the imidazole ring in this case is partially positively charged since it is bound to the zinc ion), while the hydroxyl hydrogen could hydrogen bond to the carbonyl oxygen of Pro 238. However, it is not clear whether this would also attract water molecules closer to the active site or if the enhanced solvation free energy of the inhibitor would be overcome by these new enzyme−inhibitor interactions. A radial distribution function analysis showed that there are ∼2 water molecules "solvating" the C6 methyl 5.4 Å away, whereas the C7 methyl is "solvated" by 2 water molecules at 6.6 Å away. From this analysis it is apparent

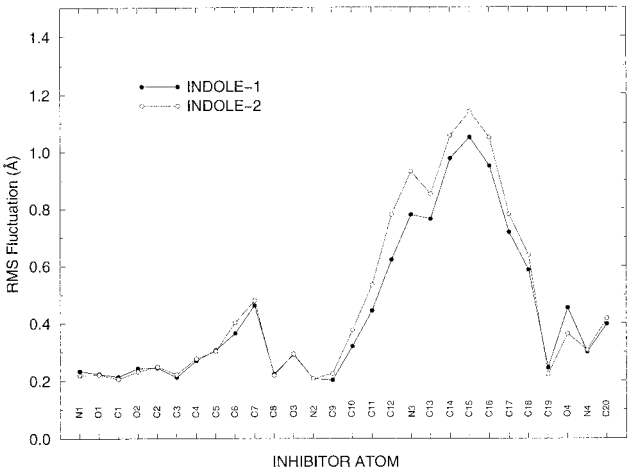

Figure 11. rms fluctuations of IND1 and IND2 atoms obtained from MD simulations.

that the Leu side chain in the P_1' site acts as a "solvent screen" to effectively shield the inhibitor from potential water molecules entering the active site and affecting, for example, the binding of the zinc binding group.

The P_2' and P_3' have been shown to be able to tolerate various substituents which results in the suggestion that they do not play a dominant role in enzyme binding. However, bulky side chains in the P_2' site have been shown to have an increased affinity for HFC.[8] The formation of a cyclic lactam between P2′ and P3′ has also been shown to result in increased inhibition.[8] In our simulations, both the IND1 and IND2 inhibitors have the bulky Trp side chain in the P_2' site and a methyl group in the P_3' site. The Trp side chain in IND2 has undergone a "ring flip" (i.e., it is oriented away from the P_3' methyl group) allowing it to be more solvent-exposed. Thus, the distance from the methyl carbon in the P_3' position to the C-γ in the Trp side chain has increased from 4.3 ± 0.2 Å for IND1 to 5.8 ± 0.1 Å in IND2. The dihedral angle between the Trp side chain and the backbone atoms of the inhibitor (χ_1 (C19−C9−C10−C11); see Figure 5 for the numbering) shows that the indole ring is pointed "away" from the backbone atoms in IND2 (180°) compared to the position in IND1 (70°). This closer interaction between P_2' and P_3' substituents in IND1 is also evident in the water radial distribution function of the amide hydrogen (H25, located between P_2' and P_3'). In IND1 the first water density peak is approximately 4.2 Å away because the indole ring of the Trp side chain blocks the close approach of water, while in IND2 the first water density peak appears at ∼2.2 Å. The P_2' and P_3' sites in IND1 remained close throughout the trajectory, forming a "cleftlike" shape, which favors van der Waals contacts between the two sites. Additionally, we have observed that the close contact between P_2' and P_3' in IND1 renders the indole ring slightly less flexible than in IND2. Figure 11 gives the rms fluctuation of the Trp side chain for both IND1 and IND2 when bound to the protein which supports the notion that the fluctuations of the Trp side chain are greater in IND2. Further support of the Trp ring flip comes from the dihedral formed by the Trp side chain atoms (C9−C10−C11−C18) which for IND1 is +70° and for IND2 is −70°. As noted previously cyclization between the P_2' and P_3' sites has been shown to result in increased activity of HFC inhibitors.[8] The rigidity and the orientation of the Trp side chain in IND1 clearly show that cyclization between the P_2' and P_3' sites is possible, while the orientation observed in IND2 is not conducive to cyclization.

There is also another zinc site in the HFC protein designated as the structural zinc site (ZNS) due to its inaccessibility to the protein surface. The ZNS is located ∼12.3 Å away from the catalytic zinc. The presence of this second zinc ion differentiates the collagenases from thermolysin and astacin families of enzymes.[9] The structural zinc is tetrahedrally coordinated by imidazole nitrogens from His 168, 183, and 196 and one carboxylate oxygen of Asp 170. The other Asp 170 carboxylate oxygen is hydrogen-bound to the hydroxyl hydrogen of Ser 172. However, this hydrogen bond was not very stable and was occasionally replaced by a hydrogen bond to a water molecule. The calcium ion, which is essential for enzymatic activity,[9] remained octahedrally coordinated and stabilizes the glycine residues in the loop (Gly 176, 178, and 179) (see Table 4). The structural zinc, along with the calcium site, maintains the structural integrity of the 170−184 loop in the active site region which in turn helps maintain a number of the protein−inhibitor contacts described above.

Conclusion

The FEP method has been successfully applied to investigate the relative binding affinities of the two succinyl hydroxamate inhibitors IND1 and IND2 of HFC. In these simulations, have adopted the bonded approach for the metal ions[1] and we have calculated the partial atomic charges for the metal ions and the coordinating ligands using a novel QM/MM ESP fitting procedure.[27] In this procedure we include the effect of the enzyme or solution-phase environment into the computed charges as opposed to previous approaches that utilized identical gas-phase charges throughout a set of simulations representing very different environments (i.e., enzyme versus solution). Thus, we are including effects arising from the polarization of the inhibitors by the local environment[32] as well as charge transfer[1,39] between the metal ions and their associated ligands. We find that the molecular mechanics parameters developed for HCAII[1] can also be successfully extended to other zinc metalloenzymes such as MMPs, with missing parameters being added using ab initio calculations (e.g., for the hydroxamate group). Importantly, our active site models were able to retain the experimentally observed structure of HFC. Furthermore, the computed free energy for conversion of IND1 into IND2 was in satisfactory agreement with the experiment.

The simulations reported here have also provided us with structural insights into why IND1 is a better inhibitor than IND2. Hydroxamate inhibitors of MMPs have been generally shown to be more potent than the alternative zinc binding groups such as carboxylic acids, thiols, or phosphinic acids. However, the answer to why the hydroxamate group is the preferred zinc binding ligand is still not clear. Some of the proposed explanations include (a) the presence of short zinc−hydroxamate oxygen distances, (b) the longer distance between the two zinc-binding oxygens in hydroxamate which readily favors the trigonal-bipyramidal binding geom-

etry, and (c) the contribution to the free energy of binding arising from desolvation of the inhibitor.[13] In this study, secondary hydrogen bond interactions have emerged as another possible factor favoring hydroxamate as the zinc binding group. For example, the presence and correct orientation of the amide hydrogen in hydroxamate allow for the formation of a hydrogen bond between the carbonyl oxygen in Ala 182. The same is true for the presence of the hydroxyl hydrogen of hydroxamate which can form a hydrogen bond with Glu 219. The importance of Ala 182 and Glu 219 can be further extended to MMP inhibitors with different zinc binding groups. Replacement of the hydroxamate zinc binding group in IND1 with a carboxylic acid moiety reduces the binding constant by 2 orders of magnitude.[23] In addition, the hydrogen bonding of Ala 182 is no longer directly with the zinc binding group (it is shifted down along the inhibitor backbone), while Glu 219 forms no hydrogen bonds with the inhibitor.[40] Other factors are likely to be important, but the role of secondary hydrogen-bonding interactions cannot be underestimated.

We have further shed insight into the role other inhibitor substituents play in protein−inhibitor interactions. Our simulations showed that all the necessary hydrogen bonds are satisfied and retained for both of the inhibitors. The P_1' substituent complements the S_1' binding site through favorable van der Waals contacts (between Val 215 and C7) while simultaneously removing solvent ("solvent screen") from the vicinity of the catalytic metal ion. We also proposed that a small polar group (e.g., hydroxyl) replacing the existing nonpolar P_1' substituent (C6) may better complement the other side of the S_1' binding pocket and, thereby, possibly resulting in tighter binding. The ability to form a more rigid P_2' and P_3' grouping further appeared to enhance the binding ability of IND1 over IND2, which is consistent with the enhancement of binding that is observed when the P_2' and P_3' substituents are cyclized.

Acknowledgment. We would like to thank the NIH for supporting this research through Grant GM44974. We also thank the Pittsburgh Supercomputer Center, National Center for Supercomputer Applications, San Diego Supercomputer Center, and Cornell Theory Center for generous allocations of supercomputer time.

References

(1) Hoops, S. C.; Anderson, K. W.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 8262−8270.
(2) Birkedal-Hansen, H.; Moore, W. G. I.; Bodden, M. K.; Windsor, L. J.; Birkedal-Hansen, B.; DeCarlo, A.; Engler, J. A. *Crit. Rev. Oral Biol. Med.* **1993**, *4*, 197−250.
(3) Bode, W.; Gomis-Rüth, F. X.; Stöcker, W. *FEBS Lett.* **1993**, *331*, 134−140.
(4) Pyke, C.; Ralfkiaer, E.; Huhtala, P.; Hurskainen, T.; Dano, K.; Tryggvason, K. *Cancer Res.* **1992**, *52*, 1336−1341.
(5) Galardy, R. E. *Drugs Future* **1993**, *18*, 1109−1111.
(6) Sawicki, G.; Salas, E.; Murat, J.; Miszta-Lane, H.; Radomski, M. *Nature* **1997**, *386*, 616−619.
(7) Gomis-Ruth, F. X.; Meyer, E. F.; Kress, L. F.; Politi, V. *Protein Sci.* **1998**, *7*, 283−292.
(8) Beckett, R. P.; Davidson, A. H.; Drummond, A. H.; Huxley, P.; Whittaker, M. *Drug Discov. Today* **1996**, *1*, 16−26.
(9) Stöcker, W.; Grams, F.; Baumann, U.; Reinemer, P.; Gomis-Rüth, F. X.; McKay, D. B.; Bode, W. *Protein Sci.* **1995**, *4*, 823−840.
(10) Matthews, B. W. *Acc. Chem. Res.* **1988**, *21*, 333−340.
(11) Rockwell, A.; Melden, M.; Copeland, R. A.; Hardman, K.; Decicco, C. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1996**, *118*, 10337.
(12) Spurlino, J. C.; Smallwood, A. M.; Carlton, D. D.; Banks, T. M.; Vavra, K. J.; Johnson, J. S.; Cook, E. R.; Falvo, J.; Wahl, R. C.; Pulvino, T. A.; Wendoloski, J. J.; Smith, D. L. *Proteins: Struct., Funct., and Genet.* **1994**, *19*, 98−109.
(13) Grams, F.; Crimmin, M.; Hinnes, L.; Huxley, P.; Pieper, M.; Tschesche, H.; Bode, W. *Biochemistry* **1995**, *34*, 14012−14020.
(14) Lovejoy, B.; Cleasby, A.; Hassell, A. M.; Longley, K.; Luther, M. A.; Weigl, D.; McGeehan, G.; McElroy, A. B.; Drewry, D.; Lambert, M. H.; Jordan, S. R. *Science* **1994**, *263*, 375−377.
(15) Scolnick, L. R.; Clements, A. M.; Liao, J.; Crenshaw, L.; Hellberg, M.; May, J.; Dean, T. R.; Christianson, D. W. *J. Am. Chem. Soc.* **1997**, *119*, 850−851.
(16) Woessner, J. F., Jr. *FASEB J.* **1991**, *5*, 2145−2154.
(17) Vedani, A.; Huhta, D. W. *J. Am. Chem. Soc.* **1990**, *112*, 4759−4767.
(18) Stote, R. H.; Karplus, M. *Proteins* **1995**, *23*, 12−31.
(19) Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 406−411.
(20) Merz, K. M., Jr.; Murcko, M. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 4484−4490.
(21) Johnson, W. H.; Roberts, N. A.; Borkakoti, N. *J. Enzyme Inhib.* **1987**, *2*, 1−22.
(22) Ferrin, T. E. *J. Mol. Graph.* **1988**, *6*, 13−27.
(23) Levy, D. E.; Lapierre, F.; Liang, W.; Ye, W.; Lange, C. W.; Li, X.; Grobelny, D.; Casabonne, M.; Tyrrell, D.; Holme, K.; Nadzan, A.; Galardy, R. E. *J. Med. Chem.* **1998**, *41*, 199−223.
(24) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.
(25) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.
(26) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(27) Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1995**, *99*, 11266−11275.
(28) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431−439.
(29) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221−264.
(30) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220.
(31) Stewart, J. J. P. *J. Comput. Chem.* **1991**, *12*, 320−341.
(32) Marrone, T. J.; Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1994**, *98*, 1341−1343.
(33) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4907−4917.
(34) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.
(35) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.
(36) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(37) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. AMBER 4.0; University of California: San Francisco, CA, 1991.
(38) Berendsen, H. J. C. In *Molecular Dynamics and Protein Structure*; Berendsen, H. J. C., Ed.; Polycrystal Book Service: Western Spring, IL, 1985; p 18.
(39) Nadig, G.; Van Zant, L. C.; Dixon, S. L.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1998**, *120*, 5593−5594.
(40) Becker, J. W.; Marcy, A. I.; Rokosz, L.; Axel, M. G.; Burbaun, J. J.; Fitzgerald, P. M. D.; Cameron, P. M.; Esser, C. K.; Hagmann, W. K.; Hermes, J. D.; Springer, J. P. *Protein Sci.* **1995**, *4*, 1966−1976.

JM980577F