

Binding Affinities for a Series of Selective Inhibitors of Gelatinase-A Using Molecular Dynamics with a Linear Interaction Energy Approach

T. J. Hou, W. Zhang, and X. J. Xu*

College of Chemistry and Molecular Engineering, Peking University,
Beijing 100871, People's Republic of China

Received: December 11, 2000; In Final Form: March 2, 2001

The binding of a series of hydroxamate inhibitors with gelatinase-A is examined to evaluate the viability of calculating free energies of binding, ΔG_b , utilizing molecular dynamics (MD) simulations with a linear interaction energy approach. In our simulations, a bonded model was used to represent the potentials of the catalytic zinc center. The electrostatic distribution of this model was derived using a two-stage electrostatic potential fitting calculations. The resulting bonded model was then used to generate the MD trajectories. Coulombic, van der Waals, and coordinate bond energy components determined from MD simulations of the bound and unbound inhibitors solvated in water were correlated with the free energies of binding for the 15 hydroxamate inhibitors. In the correlation process, several linear models consisted of different energy components were tested. We found that besides the usually used Coulombic and van der Waals energy terms, the introduction of a constant term could significantly improve the correlation. The best model yields an average error of 0.6 kcal/mol for the 15 binding affinities, which cover an observed range of 7.2 kcal/mol. The predictive ability of the best model was revealed by the high value of q^2 (0.854) from the leave-one-out cross-validation. To this series of inhibitors, the constant term can be treated as effective adjustment to the entropy contribution in the binding free energies. The MD simulations predicted the binding mode of the gelatinase-A with the studied inhibitors, and also provided insights into the interactions occurring in the active site and the origins of variations in ΔG_b . The P_1' groups of inhibitors make extensive van der Waals and hydrophobic contacts with the nonpolar side chains of four residues in the $S1'$ subsite, including Leu 197, Val 198, Leu 218, and Tyr 223, which directly influence the ligand binding. Hydrogen bonds between hydroxamates and gelatinase-A are very important to stabilize the inhibitors in the active site. The hydrogen bonds between the P_3' group and gelatinase-A can produce more favorable electrostatic interactions.

1. Introductions

Molecular docking can fit molecules together in favorable configuration to form a complex system. Molecular docking has been practically applied as a very efficient way in the studies of protein–ligand interactions.¹ The structural information from the theoretically modeled complex may help us clarify the mechanism of molecular recognition, and even can instruct us to change the structure of receptor or ligand in order to improve some biological functions or design new compounds. So molecular docking may be one of the most important techniques in structure-based drug design. The key problem in molecular docking is the estimation of binding affinities for putative receptor–ligand complexes. Many approaches have been developed to explore the energy landscape of a ligand at the binding site. Among all these approaches, thermodynamics integration (TI) and free energy perturbation (FEP) may be the most rigorous and strict techniques. But these two methods are too time-consuming and can be applied only to those relatively simple and small systems. Moreover, free energy calculations often run into sampling problems (both in conformation and configuration sampling), because free energy calculations spend most of their computation time on nonphysical intermediate states. This problem becomes even more severe when TI or FEP are applied to predict binding affinities of a set of ligands

with dissimilar structural scaffolds. So, these two rigorous methods including TI and FEP are too difficult or even impossible to be applied in ligand screening.^{2–4}

To estimate the ligand binding affinities with receptor computational efficiently and practically, some empirical scoring function methods have been developed.^{5–14} The basic idea of this kind of approach is to derive an empirical function composed of free energy components, using the known ligand–protein complex structures and their experimental binding free energies. But the application is often limited, and they are often system-sensitive (including ligand and receptor), and the obtained empirical functions are only applicable to some kinds of similar ligands. Furthermore, the scoring function is derived from the known ligand–receptor structures, where receptor binding sites have formed perfect complementarity with the corresponding ligands. As a result, the scoring function works best only for well matched ligand–receptor complex structures, but it may have poor tolerance to a slight change in binding orientation or conformation of ligand near the binding site. For a set of ligands with different sizes and heteroatoms, it is often very difficult to predict the perfectly correct binding mode of ligand in the receptor binding site using available docking programs; thus, the subsequent prediction of binding affinities using an empirical function is only questionable.

Recently, it has been reported that fairly accurate binding free energy estimates can be obtained from molecular dynamics

* Corresponding author. E-mail: xiaojxu@chemms.chem.pku.edu.cn.

simulation energy averages, which is proposed by Åqvist et al. as a semiempirical linear interaction energy (LIE) approach.^{15–19} In this method, ΔG_b can be computed on the basis of ligand interaction energies in the bound and free states. The main trick, deriving from linear response consideration, is that only averages of the interaction energies between the ligand and its surroundings need to be evaluated. The binding free energies are broken down into electrostatic (Coulombic) and van der Waals contributions.

$$\Delta G_{\text{bind}} \approx \Delta G_{\text{bind}}^{\text{el}} + \Delta G_{\text{bind}}^{\text{vdw}} \quad (1)$$

$$= \alpha(\langle U \rangle_{\text{bound}}^{\text{el}} - \langle U \rangle_{\text{free}}^{\text{el}}) + \beta(\langle U \rangle_{\text{bound}}^{\text{vdw}} - \langle U \rangle_{\text{free}}^{\text{vdw}})$$

where U^{el} and U^{vdw} are the electrostatic and van der Waals interaction energies between ligand and its surroundings in protein (*bound* form) or in aqueous solutions (*free* form), respectively; $\langle \rangle$ denotes the ensemble average over molecular dynamics or Monte Carlo simulations; and α and β are empirical parameters. MD simulations were used to determine the required energy components.

In the initial work of Åqvist et al., $\beta \approx 0.16$ was derived to give the best fit to experimental binding data, and the electrostatic scaling factor, $\alpha = 0.5$, follows from the quadratic dependence of free energy on solute charge, as embodied in the Born model for ion solvation. The calculations of Paulsen and Ornstein show that $\alpha \approx 0.5$ and $\beta \approx 1.043$ are best to correlate the calculated binding free energies with the respective experimental values.²⁰ In Wang's work, $\alpha \approx 0.5$ and $\beta \approx 1.0$ can give the best predicted binding affinities to a group of avidin ligands using AMBER force field.²¹ While Jones-Herzog and Jorgensen observed that an addition of another term concerned with molecular solvent-accessible surface as well as $\alpha = 0.131$ and $\beta = 0.131$ could produce the best correlation.²²

It seems that the variations of the parameters in the LIE models are mainly affected by two factors, receptor–ligand systems and force fields. Regardless of the transferability of these parameters, there are several appealing features of the LIE approach. First of all, LIE is faster than FEP or TI calculations because LIE requires only simulations at the endpoints of mutations. Second, LIE can take into account the protein flexibility just as FEP or TI does, which is often not included in scoring function approaches. Finally, LIE uses an explicit solvent model in its simulations; thus, the contribution of enthalpy in desolvation free energy can be reasonably handled. These advantages make LIE a potential useful tool in structure-based ligand design.

The aim of this paper is to apply the LIE method to calculate ligand binding free energies for a set of representative inhibitors interacting with gelatinase-A (MMP-2). Moreover, we want to study their binding modes in the active site.

Matrix metalloproteinases (MMPs) are key enzymes involved in connective tissue turnover in normal and pathological conditions.^{23–25} These zinc- and calcium-dependent enzymes are synthesized as zymogens and under physiological conditions the proteins are selectively regulated by inhibitors called tissue inhibitors of metalloproteinases (TIMPs).^{26,27} The extracellular matrix functions as a medium of migration, attachment, and structural support in various cell types and tissues. Therefore, MMPs play a crucial role in matrix remodeling events of connective tissues during embryonic growth and wound healing. On the basis of substrate specificity and primary sequence similarities, the members of the matrixin family can be grouped into five subfamilies: gelatinases (MMP-2, -9), which cleave

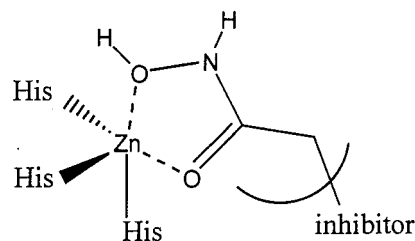


Figure 1. The catalytic zinc atom coordinated by three liganding histidine nitrogens of the enzyme and both hydroxamate oxygens.

denatured collagen, elastin, and type IV and V collagens; collagenases (MMP-1, -8, -13), which cleave triple-helical interstitial collagen; stromelysins (MMP-3, -10, -11), which mainly cleave proteoglycans; membrane-type MMPs (MMP-14, -15, -16, -17), which are associated with activation of pro-MMPs; and others (MMP-7, etc).^{28,29} At the molecular level, MMPs are characterized by a zinc atom at the active site (zinc catalytic site) with a conserved zinc binding motif, HExx-HxxGxxH. The proenzyme region also consists of a cystine residue which is conserved in MMPs, and in the inactive form which is conserved in MMPs, and in the inactive form this cystine is bonded to the catalytic zinc.

Since MMPs are involved in a wide range of diseases, there has been interest in obtaining effective small-molecule inhibitors. Many MMP inhibitors have been reported in the previous paper.³¹ Generally speaking, the requirement for a molecule to be an effective inhibitor of the MMPs is a functional group (e.g., hydroxamic acid, carboxylic acid, and sulfhydryl, etc.) capable of attaching to the catalytic zinc atom, at least one functional group which provides a hydrogen bond interaction with the enzyme back-bond, and one or more side chains which undergo effective van der Waals interactions with the enzyme subsites. Moreover, some inhibitors demonstrate binding affinities to several MMPs, which can be naturally deduced that the binding modes between different inhibitor and different MMP are somewhat similar. Of all these inhibitors, compounds with the hydroxamate zinc binding group may be the most popular ones in MMPs such as MMP-1, MMP-2, MMP-3, MMP-9, etc. The crystal structure of hydroxamate inhibitors complexed with MMPs have revealed that the catalytic zinc was pentacoordinated with three histidine nitrogens in MMPs and two hydroxamate oxygens in the inhibitor (see Figure 1).

Among the subfamilies of MMPs, gelatinases have been considered very promising for use in drug development. Since gelatinases are thought to play an important role in triggering the processes of tumor growth, invasion, and metastasis by cleaving the vascular basement membrane which consists of type IV collagen, gelatinase inhibitors have been studies extensively in the search of a new type of anticancer drug.^{32,33} The recent resolution of the crystal structure of MMP-2 opens the opportunity to develop new drugs by the structure-based approach.³⁴ However, few research studies have been reported concerned with the theoretical estimation of the binding free energy of MMP-2 with its inhibitors.

Because of the presence of zinc ions (one catalytic and one structural) in MMPs, we are faced with selection of a way in which to model these ions. There are two ways to model the potentials of these ions: a nonbonded model and a bonded model. In the nonbonded approach, nonbonded electrostatic and van der Waals terms are used to model the metal–ligand/enzyme interactions. While in the nonbonded approach, the coordinate bonds between ions and ligand/enzyme are described by the usually used bonded terms including bond stretching,

angle bending, and torsional terms, etc. The nonbonded approach is very simple, but its limitation is obvious. The nonbonded method is very sensitive to the electrostatic model chosen and can suffer from the inability to retain low coordination number. In Toba's work, it has been confirmed that using AMBER force field, 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.³⁵ The bonded model can describe the coordination bonds more precise, but the difficulties on the parametrizations of the bonded terms concerned with these ions limit the applications of the bonded approach. In previous work of Hoops et al., the authors have extended the AMBER force field to model the zinc ion in human carbonic anhydrase II (HCAII) in both the high (zinc-hydroxide) and low (zinc-water) pH forms. The resulting energy minimized structures of the high pH (hydroxide) form of HCAII based on the "large" and "small" active site models are very similar with the experimental results.³⁶ In Toba's work of MD simulations and free energy perturbation (FEP) studies of human fibroblast collagenase (HFC) with two hydroxamate inhibitors, the authors also used a bonded model to describe the zinc ions. The bonded 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. So, in this paper we have adopted the bonded approach for zinc ions representation, and we think that the use of the bonded approach will accurately predict the MMP-2-hydroxamate systems.

In this paper, a set of representative MMP-2 inhibitors, hydroxamates, with wide range of binding affinities, have been selected for free energy calculations. A bonded model for the catalytic zinc center where the electrostatic representation for this model was derived using electrostatic potential (ESP) fitting was proposed to generate the MD trajectories for analyses.

2. Methods and Computational Details

Construction of the Initial Structures for MD Simulations.

A recent report described the structure of the full-length proform of MMP-2 (proMMP-2, PDB code: 1QIB).³⁴ The catalytic domain of the proMMP-2 is similar to that of proMMP-3. The same residues form the substrate binding pockets and coordination of the catalytic Zn^{2+} ion is quite similar. Also, the binding site is identical to a well-conserved motif found in all known MMP structures. But the complex structure of proMMP-2 and its inhibitors are unavailable, so, it is a challenge for us to construct the initial structures for MD simulations. According to the references, it can be found that the basic structures of MMP inhibitors are quite similar; moreover, considerable insight into MMP ligand interaction has revealed that the inhibitors of different MMPs generally adopt similar binding modes with their receptors. So we believe that the constructed complexes of proMMP-2 with hydroxamates from the crystal structure of a proMMP-3 complex are precise enough to be the initial structures for MD simulations. In this paper, an X-ray crystal of proMMP-3 with a hydroxamate inhibitor (PDB code: 1BIW) obtained by Natchus et al. was used as the template molecule.³⁷ The hydroxamate inhibitor in 1BIW is shown in Figure 2.

In the current work, we selected 15 hydroxamate inhibitors from reference, and some of them show strong binding affinity to proMMP-2.^{38–40} The studied compounds are shown in Table 1. The construction of the complex structure of proMMP-2 with the studied molecules could be divided into three stages. First, the structural alignment was used to superimpose proMMP-2 and proMMP-3 together. Then, the hydroxamate inhibitor was

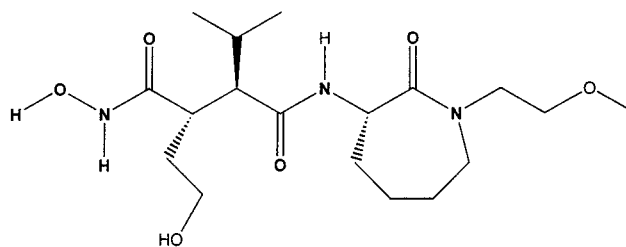


Figure 2. The structure of hydroxamate inhibitor in 1BIW.

extracted from 1BIW and merged into 1QIB. Finally, the structure of hydroxamate inhibitor was modified to obtain the studied molecules. During this model-building process, the conformation of the protein was not altered, and the structures of inhibitors were altered minimally to avoid unacceptable atom bumps. All the constructing processes were completed on InsightII molecular simulation package on a SGI Octane 2-CPU workstation.⁴¹

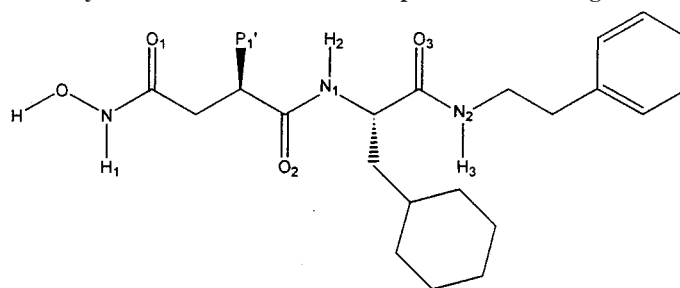
Force Field. Most bond and angle parameters associated with the zinc center were taken from Hoops et al.³⁶ All the torsions associated with the zinc–ligand bonds were set to zero as in Hoops et al. Three angle parameters concerned with the zinc center, absent from the work of Hoops et al., were taken from the studies of Ryde.⁴³ These three angles are C–O–Zn, C–V–NB–Zn, and HO–OH–Zn. The force field parameters associated with N–O functionality were taken from Taba et al.³⁵ In Taba's work, the acetohydroxamate ($\text{CH}_3\text{C}(\text{O})\text{NHOH}$) was used as a small molecule model system to determine the bond, angle, and dihedral parameters by HF/6-31G* *ab initio* level calculations. Some missing parameters concerned with inhibitors were obtained from the newest AMBER force field (parm99) revised by Wang et al.^{43,44} (Some parameters especially for small molecules, absent in the previous AMBER force field, have been developed in the newest AMBER force field.) All force field parameters for the zinc ion and the hydroxamate group associated with inhibitors were shown in Table 2 (the parameters in the AMBER force field are not listed). In this paper, all atoms are represented with the atom types defined in the AMBER force field.

Partial Charges. In our MD simulations, the structural and catalytic zinc centers were represented with the nonbonded and bonded approaches, respectively. We believe that the structural zinc does not directly influence the ligand binding, and the conformation near the structural zinc can be maintained using a nonbonded model combined with a structural restrain. Three calcium ions in 1QIB were also represented with nonbonded approach. So formal charges of $+2|e|$ were assigned to the structural zinc and the three calcium ions. But to the catalytic zinc ion, in addition to the bond, angle, and dihedral parameters, the coordination geometries of this zinc site required the development of appropriate charge representations.

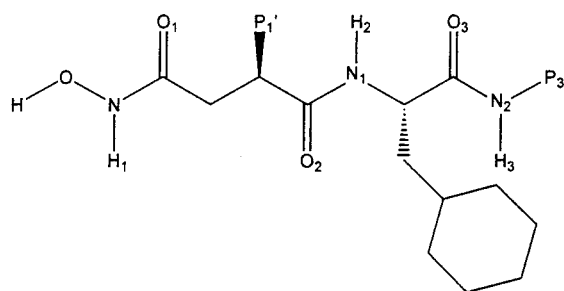
As well-known, when an inhibitor binds to its receptor, the distributions of charge are quite different from those on its free form in a vacuum or in solution. To the studied system here, the crystal complexes of enzyme and inhibitors are unavailable, so it seems that it is very difficult for us to determine the partial charge of the catalytic zinc ion and the inhibitors. In the current work, to solve this problem, we adopted a novel two-stage ESP fitting calculation.

In the first stage, a simple model was used to determine the ESP charges of the catalytic zinc ion, three histidines, and two hydroxamate oxygens coordinated. The structure shown in Figure 3 was adopted as the simple theoretic model for the

TABLE 1: Structures of the Studied Hydroxamate Inhibitors and Experimental Binding Free Energies for MMP-2



| No. | P ₁ ' | ΔG_b (kcal/mol) |
|-----|--|-------------------------|
| a1 | CH ₂ CHMe ₂ | -12.93 |
| a2 | CH ₂ Ph | -7.80 |
| a3 | (CH ₂) ₂ Ph | -11.80 |
| a4 | (CH ₂) ₃ Ph | -13.92 |
| a5 | (CH ₂) ₄ Ph | -13.09 |
| a6 | (CH ₂) ₅ C ₆ H ₁₁ | -11.13 |
| a7 | (CH ₂) ₃ Ph-4Me | -13.77 |
| a8 | (CH ₂) ₃ Ph-4OMe | -13.93 |
| a9 | (CH ₂) ₃ Ph-4F | -13.93 |
| a10 | (CH ₂) ₃ Ph-4Et | -13.28 |



| No. | P ₁ ' | P ₃ ' | ΔG_b (kcal/mol) |
|-----|--|--|-------------------------|
| a11 | (CH ₂) ₃ Ph-4Cl | (CH ₂) ₂ CO ₂ H | -15.01 |
| a12 | (CH ₂) ₃ Ph-4Cl | (CH ₂) ₂ CO ₂ Me | -15.01 |
| a13 | (CH ₂) ₃ Ph-Cl | (CH ₂) ₄ N | -14.35 |

complex structure of proMMP-2 with inhibitors in the active site. The simplified model includes only the catalytic zinc ion, three histidines, and an inhibitor coordinated. In this model, acetohydroxamate (CH₃C(O)NHOH) was used as the simplified inhibitor. First, we constructed the complex structure of proMMP-2 with acetohydroxamate from 1BIW. Then we extracted the theoretic model from the constructed structure, and AMBER 6.0 molecular simulation package was used to add hydrogens in order to convert the PDB united atom representation to the all-atom representation. Hydrogen atoms were substituted at N-terminal residues and hydroxy groups were substituted at carbonyl groups at C-terminal residues. The resulting geometry was fully optimized by using the AM1 Hamiltonian in MOPAC 7.0 before performing a single-point ab initio calculation with no further geometry optimization.⁴⁵ The Hartree–Fock method with the 6-31G* basis set applied in the Gaussian 98 program was used to determine electrostatic potentials at a grid of points around the theoretic model.⁴⁶ Finally, the partial charges were derived using the Restrained Electrostatic Potential (RESP) fitting method provided by AMBER.⁴⁷ In this stage, the partial charges for the catalytic

zinc, three liganding histidines, and two hydroxamate oxygens were determined. Figure 3 shows the partial charges on some atoms for the simplified active site model.

In the second stage, we want to obtain the partial charges of the studied inhibitors. All the molecules shown in Table 1 modified from the hydroxamate in 1BIW were fully minimized by the AM1 Hamiltonian in MOPAC 7.0. The Hartree–Fock method with the 6-31G* level applied in the Gaussian 98 program was used to determine electrostatic potentials. RESP fitting technique in AMBER was applied to determine the partial charges. Because of the absence of the crystal structures of the complexes, we could not explicitly take the consideration of the influence of environment to the inhibitors during the determination of partial charges. We believe that when the hydroxamates in Table 1 are recognized by proMMP-2, the partial charges on two hydroxamate oxygens should be quite similar to those on two hydroxamate oxygens in acetohydroxamate in Figure 3. So for the bound inhibitors with proMMP-2, during the ESP calculations, two Lagrange constrains in RESP were used to fix the partial charges on the O1 and O2 atoms to those values from the first stage, while the partial charges on

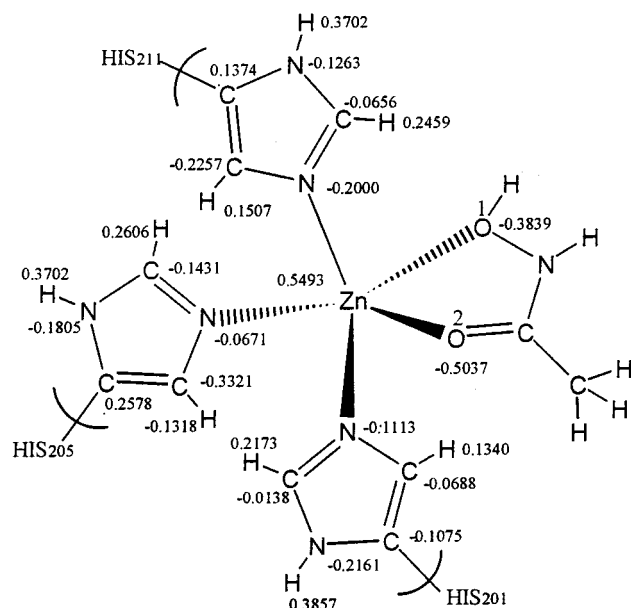


Figure 3. The partial charges for the active site model from the RESP fitting calculations.

TABLE 2: Force Field Parameters for the Catalytic Zinc Ions and Some Groups Associated with Hydroxamates

| 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} (degree) | |
| NB–Zn–NB | 20.0 | 105.0 | |
| NB–Zn–O | 20.0 | 115.0 | |
| CR–Zn–NB | 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 | |
| C–O–Zn | 24.0 | 131.2 | |
| CV–NB–Zn | 20.0 | 126.0 | |
| HO–OH–Zn | 28.9 | 110.3 | |
| Dihedral Parameters | | | |
| dihedral | ν (kcal/mol) | γ (degree) | n |
| HO–OH–N–C | 6.0 | 180 | 1 |

the other atoms were determined from the ESP fitting. For the unbound inhibitors, no Lagrange constraints were applied. ESP charges for the studied inhibitors as the bound and unbound states were provided as Supporting Information in the form of the AMBER database files.

Molecular Dynamics. All MD simulations were performed using AMBER 6.0. The three bonds to imidazole rings and two bonds to hydroxamate oxygens were considered conserved. The charges resulted from the ESP calculations were incorporated into the AMBER force field. From the calculations, the partial charges on three histidines coordinated with catalytic zinc ion are quite different, meanwhile, they are significantly different from the standard AMBER values, so we need to make some modifications to the AMBER database file and PDB file to properly recognize these three different residues. In AMBER, we defined three kinds of new residue: HIA, HIB, and HIC, and they represented His 201, His 205, and His 211 in 1QIB, respectively. The partial charges on these three newly defined

residues were revised from the default values in AMBER to the ESP charges in this paper. Moreover, we defined a new zinc unit to represent the catalytic zinc center. Each ligand studied was defined as a new type of residue and added to the AMBER database files. All force field parameters in Table 1 were added to the AMBER force field file.

All MD simulations were carried out at 300 K. The explicit solvent model TIP3P water was used. For each system a pair of simulations was performed, one with the ligand in 20 Å sphere of water, the other with the ligand bound to the protein with a cap of waters around the complex filled up to 20 Å from the center of mass of the ligand. The SHAKE procedure was employed to constrain all bonds involving at least one hydrogen atom. The time step of the simulations was 1.5 fs with a cutoff 12 Å for the nonbonded interactions. The nonbonded pairs were updated every 30 steps. All protein atoms within 16 Å of the center of mass of ligand were allowed to move freely. Prior to the MD simulations, the system was minimized with harmonic position constraints for all protein heavy atoms. The constraints are 5000, 1000, 100, and 10 kcal/mol/Å². Subsequently, a cycle of minimization was done to relax all the atoms without constraints. The maximum minimization steps were 10 000 and the convergence criterion for energy gradient was 0.5 kcal/mol/Å². In MD simulations of bound ligand in protein, all residues within 16 Å were allowed to move, while the other atoms were restrained by a 50 kcal/mol/Å harmonic force. In MD simulation of an unbound ligand in aqueous solution, a position constraint for the ligand heavy atom that is closest to the center of mass of the ligand was applied. MD simulation procedures for the ligand in water and in protein involved (1) 50 ps of MD simulations for equilibrium, (2) 100 ps of MD simulations for data collection. In the data collection stage, every 200 fs, the snapshot was recorded in the trajectory file. During the data collection run, the interaction energies of the ligand with all solvent atoms (protein and solvent) were calculated and printed out for use in the subsequent analysis.

Construction of the LIE Models. In this work, a genetic algorithm (GA) was used to fit the present energetic components to the experimental ΔG_b values. During the GA optimizations, the multiple linear regression coefficient (r) was defined as the fitness score to evaluate the LIE models. The reliabilities of the models were tested by the leave-one-out cross-validation technique.

3. Results and Discussion

Partial Charges. The resulting partial charges for the catalytic zinc ion are summarized in Figure 3. Figure 3 clearly demonstrates that the modeling zinc with a formal $+2|e|$ point charge is unrealistic. The calculated partial charge on the catalytic zinc ion is significantly smaller than $+2|e|$. When zinc ion is coordinated with three histidine nitrogens and two hydroxamate oxygens, the electron will be transferred from the dentates to the zinc center. So, the partial charges for the coordinated ligands from the two-stage ESP fittings with two Lagrange constraints for O1 and O2 atoms should be more accurate than those from the direct ESP fittings. In several previous paper of parametrization about zinc ions, the point charge on zinc seem to bear some differences. In the paper of Hoops et al.,²¹ the author selected the metalloprotein human carbonic anhydrase II (HCAII) as the test metalloprotein system for ESP fitting calculations. When zinc ion adopted tetrahedral geometry with coordination to three histidine nitrogens and hydroxide oxygen, the partial charges on zinc ion from MNDO calculations is $+0.688|e|$. In Toba's work,³⁵ when catalytic zinc center adopt a trigonal–bipyramidal

TABLE 3: Averaged Interaction Energies from the Bound and Unbound MD Simulations (kcal/mol)

| | ligand | $\langle U_{\text{coord}} \rangle$ | $\langle U_{\text{vdw}} \rangle$ | $\langle U_{\text{elec}} \rangle$ |
|-----|--------------|------------------------------------|----------------------------------|-----------------------------------|
| a1 | bound | 24.936 | -63.907 | -58.337 |
| | unbound | | -42.150 | -71.758 |
| | ΔE^a | 24.936 | | |
| a2 | bound | 24.189 | -62.167 | -55.298 |
| | unbound | | -41.863 | -75.427 |
| | ΔE^a | 24.189 | | |
| a3 | bound | 26.332 | -65.752 | -63.437 |
| | unbound | | -45.757 | -76.873 |
| | ΔE^a | 26.332 | | |
| a4 | bound | 19.310 | -72.609 | -56.634 |
| | unbound | | -47.086 | -65.412 |
| | ΔE^a | 19.310 | | |
| a5 | bound | 28.935 | -81.010 | -52.475 |
| | unbound | | -47.527 | -74.627 |
| | ΔE^a | 28.935 | | |
| a6 | bound | 25.983 | -75.992 | -51.914 |
| | unbound | | -49.138 | -78.443 |
| | ΔE^a | 25.983 | | |
| a7 | bound | 25.100 | -77.526 | -54.524 |
| | unbound | | -45.268 | -78.198 |
| | ΔE^a | 25.100 | | |
| a8 | bound | 24.926 | -80.470 | -58.108 |
| | unbound | | -49.182 | -77.362 |
| | ΔE^a | 24.926 | | |
| a9 | bound | 22.664 | -74.989 | -55.716 |
| | unbound | | -45.477 | -79.196 |
| | ΔE^a | 22.664 | | |
| a10 | bound | 20.200 | -78.364 | -55.248 |
| | unbound | | -49.773 | -79.196 |
| | ΔE^a | 20.200 | | |
| a11 | bound | 20.322 | -73.560 | -67.751 |
| | unbound | | -41.287 | -84.212 |
| | ΔE^a | 20.322 | | |
| a12 | bound | 22.329 | -75.714 | -56.392 |
| | unbound | | -44.504 | -88.603 |
| | ΔE^a | 22.329 | | |
| a13 | bound | 22.316 | -84.236 | -57.647 |
| | unbound | | -53.836 | -79.442 |
| | ΔE^a | 22.316 | | |
| a14 | bound | 21.765 | -59.132 | -49.217 |
| | unbound | | -39.006 | -76.876 |
| | ΔE^a | 21.765 | | |
| a15 | bound | 22.523 | -55.449 | -47.980 |
| | unbound | | -38.863 | -64.349 |
| | ΔE^a | 22.523 | | |

$$^a \Delta E = \langle U_{\text{bound}} \rangle - \langle U_{\text{unbound}} \rangle.$$

coordination sphere, the point charge on zinc ion is about $+0.80|e|$. In Ryde's work,⁴² the zinc charge varied by only $\pm 0.1|e|$ for four- and five-coordinate systems with the same basis sets, and the partial charge determined for zinc ion is $+0.488|e|$. In the current work, the partial charge for the five-coordinate zinc ion is $+0.549|e|$. We think that the differences are not surprising, because some elements will affect the calculated results, including the active site models extracted from the different enzyme/ligand systems and the usage of different Hamiltonians or different basis sets in quantum chemical calculations.

Optimization of the LIR Models. The average electrostatic, van der Waals, and coordinate bond interaction energies between the inhibitors (bound and free) and their respective environments were determined during the MD simulations (see Table 3).

The initial model proposed by Åqvist (eq 1) needs only two parameters, one for the electrostatic part and the other for the van der Waals part. Moreover, in some LIE calculations published in the literature, $\alpha = 0.5$ has been used for the electrostatic interaction energy, while β is calibrated according to different systems. In Table 4, the LIE expression (eq 1) and

TABLE 4: Comparison of Experimental and Calculated ΔG_b (kcal/mol) Using Different Fitted Models

| no. | ΔG_b | | | | |
|-------|--|--|---|--|---|
| | model1 $\alpha = 0.5,$ $\beta = 0.827$ | model2 $\alpha = 0.057,$ $\beta = 0.510$ | model3 $\alpha = 0.09,$ $\beta = 0.348,$ $\gamma = -0.024$ | model4 $\alpha = 0.191,$ $\beta = 0.827,$ $\delta = -7.405$ | model5 $\alpha = 0.178,$ $\beta = 0.339,$ $\kappa = -0.132,$ $\delta = -10.236$ |
| | | | | | |
| a1 | -11.28 | -10.33 | -9.48 | -12.20 | -11.93 |
| a2 | -5.10 | -8.20 | -8.43 | -9.76 | -9.67 |
| a3 | -9.82 | -9.43 | -10.24 | -11.61 | -11.14 |
| a4 | -16.72 | -12.51 | -13.03 | -14.37 | -14.77 |
| a5 | -16.61 | -15.81 | -15.32 | -14.50 | -13.82 |
| a6 | -8.94 | -12.18 | -12.13 | -11.42 | -11.18 |
| a7 | -14.84 | -15.10 | -14.49 | -13.80 | -13.64 |
| a8 | -16.25 | -14.85 | -15.35 | -14.32 | -14.12 |
| a9 | -13.48 | -13.80 | -13.54 | -13.22 | -13.35 |
| a10 | -11.67 | -13.21 | -14.03 | -12.50 | -12.99 |
| a11 | -18.46 | -15.52 | -15.41 | -15.18 | -15.56 |
| a12 | -16.71 | -14.87 | -14.53 | -14.49 | -14.62 |
| a13 | -14.24 | -14.26 | -14.24 | -13.53 | -13.71 |
| a14 | -2.81 | -8.68 | -8.34 | -8.93 | -9.97 |
| a15 | -5.53 | -7.52 | -8.25 | -9.89 | -9.26 |
| r^2 | 0.633 | 0.713 | 0.710 | 0.851 | 0.877 |
| q^2 | 0.532 | 0.667 | 0.715 | 0.854 | 0.751 |
| rms | 2.94 | 1.53 | 1.44 | 0.79 | 0.72 |

α (0.5) and β (0.827) were ineffective at producing the experimental ΔG_b values for inhibitors **a1–15** with the energetic results from the present simulations. The r^2 of the model is 0.633, and the q^2 is relatively poor (0.532). The rms derivation between the computed and observed values is 2.94 kcal/mol. The statistical validation shows that the two-parameter model with $\alpha = 0.5$ does not bear acceptable statistical significance and predictive ability. Then, another two-term equation was reparametrized, in which the values of α and β were allowed to be optimized. In this model, β is equal to 0.510, and α is only 0.06, which is greatly different from the usually used value 0.5 (see model2 in Table 4). The new fit resulted in substantial improvement ($r^2 = 0.713$, $q^2 = 0.667$), but the obtained rms derivation (1.53 kcal/mol) was still not satisfactory. In the current work, we used the AMBER force field, while Åqvist used the GROMOS force field, but we do not think that the influence of the force fields is the main factor, because Wang et al. compared the GROMOS force field and AMBER force field, and they found that these two force fields could generally produce comparative binding free energies. So, it seems that the influence of protein–ligand systems is rather more remarkable than that of force fields. We believe that the binding free energies for the studied protein–ligand systems in this paper cannot be well computed by a simple two-parameter model.

In the works of Jones-Hertzog et al.,²² the authors proposed another LIE model with three parameters. They added another term concerned with the solvent-accessible surface area named SASA, as an index for cavity formation (see eq 2).

$$\Delta G_{\text{bind}} = \alpha(\Delta U_{\text{elec}}) + \beta(\Delta U_{\text{vdw}}) + \gamma(\Delta \text{SASA}) \quad (2)$$

Addition of the SASA term and reparametrization of α , β , and γ could slightly improve the correlation. The obtained values of α , β , and γ are 0.131, 0.131, and 0.014, respectively. The introduced SASA was considered to be directly connected with the free energies of hydration (ΔG_{hyd}). In the current work, we also introduced another term named separation surface area (SSA) (see eq 3).

$$\Delta G_{\text{bind}} = \alpha(\Delta U_{\text{elec}}) + \beta(\Delta U_{\text{vdw}}) + \gamma(\Delta \text{SSA}) \quad (3)$$

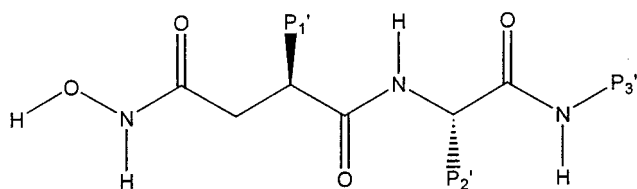


Figure 4. The common structure of the hydroxamate inhibitors.

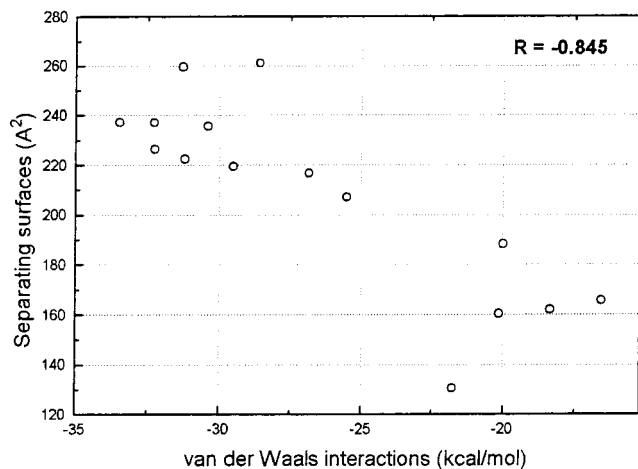


Figure 5. The relationships between van der Waals interactions with the separating surfaces.

This principle of separating surface was proposed by Keil et al. in 1998.⁴⁸ The separating surface is defined by a set of points located halfway on the shortest distance vectors between surface points of the two molecular partners. The surface is generated using a grid-based algorithm. The distance to the nearest atom is stored on the grid points and an isosurface is generated forming the separating surface. Size and shape of the surface characterizes the complex interface. We believe that the entropic contribution in ΔG_{hyd} is more closely connected with SSA rather than SASA. In this paper, we considered only the SSA between the P_1' groups of hydroxamates and the $S1'$ subsite of proMMP-2, because the surface contacts between the hydrophobic P_1' groups of hydroxamates and the hydrophobic $S1'$ subsite of proMMP-2 were crucial for the hydrophobic interactions. Because the P_2' and P_3' groups of the inhibitors are solvent exposed, while the other parts besides the P_1' groups are polar dominated, which can produce strong electrostatic interactions with proMMP-2. The separating surfaces were modeled by using the MOLCAD module in Tripos 6.5 molecular simulations package.⁴⁹ Addition of the SSA term and reparametrization of α , β , and γ for eq 3 further promoted the value of q^2 to 0.715 and lowered the rms derivation to 1.44 kcal/mol (model3 in Table 4). As being well-known, for the analysis of multiple linear regression (MLR), the data material must be reduced to fewer and less correlated variables. The cross-correlated descriptors would make the linear model not disclose the actual relationship between the binding free energies and these descriptors. After correlation studies, we found that the SSA and ΔU_{vdw} in eq 3 were highly correlated. Figure 5 shows the relationship between SSA and ΔU_{vdw} . So from the statistical point of view, model3 in Table 3 is unreasonable.

A new fit was made to eq 3 with the SSA term replaced by a constant, δ (see eq 4).

$$\Delta G_{\text{bind}} = \alpha(\Delta U_{\text{elec}}) + \beta(\Delta U_{\text{vdw}}) + \delta \quad (4)$$

The new fit resulted in significant improvement, indicated by the high q^2 (0.854) and low rms derivation (0.79). Three

parameters for model4 were $\alpha = 0.191$, $\beta = 0.339$, and $\delta = -7.405$, respectively. The average unsigned error for this fit with 15 inhibitors and three parameters are 0.63 kcal/mol, which can be compared with the work reported by other groups. In the work of Åqvist et al., the average unsigned errors for small two sets of four endothiopepsin and three HIV proteinase inhibitors are 0.4 and 1.8 kcal/mol, respectively. In Jones-Hertzog's work of seven sulfonamide inhibitors for human thrombin, an average error of 0.77 kcal/mol was obtained using a three-parameter equation. In an expanded study for a total of 18 protein-inhibitor complexes, Hansson et al. have obtained an average error of 0.72 kcal/mol using eq 1 with both α and β variable and an average error of 0.54 kcal/mol with a three-parameter equation.

In the MD simulations, a bond model was used to represent the potentials of the catalytic zinc center and the concerned coordinate bonds. So the systems studied in this paper are quite different from the reported references concerned with free energy calculations based on LIE approach. Besides the electrostatic and van der Waals interactions, the coordinate bond energy should also be included in the interactions between hydroxamates and proMMP-2, so we performed a correlation after adding the coordinate bond energy term (eq 5).

$$\Delta G_{\text{bind}} = \alpha(\Delta U_{\text{elec}}) + \beta(\Delta U_{\text{vdw}}) + \kappa(\Delta U_{\text{coord}}) + \delta \quad (5)$$

The fitting produced the highest r^2 . The high r^2 does not mean this model is better than model4. Because this model includes four terms, while model4 only includes three terms, and the more probable random correlation in model5 will produce higher r^2 and lower rms derivation. According to the values of q^2 , model5 was better than model4, indicating that model5 bears better predictive ability. So, adding the coordinate bond term cannot improve the predictive ability of the LIE model. Moreover, κ adopts a negative value, which is physically unreasonable. Although the coordinate bonds between the hydroxamate oxygens and the catalytic zinc center are crucial for ligand binding, but the coordinate bond energies for the studied inhibitors do not show obvious differences, and they seem not to obviously influence the variations of the binding affinities. It is naturally deduced that the coordinate bonds introduce equal effects to all ligand-receptor systems, which can be implicitly embedded in the constant term in eq 4.

From the analyses, model4 in Table 4 was determined as the best statistical model. The observed versus calculated binding free energies for the 15 compounds are shown in Figure 6. The constant δ , -7.405 , yields an attractive term of -7.405 kcal/mol. The constant attractive term is providing a base value for these inhibitors that is then modulated by U_{elec} and U_{vdw} terms. We believe that the constant term indeed mainly includes three parts of contributions from enthalpic and entropic effects. First, the coordinate bonds introduce relatively equal effects to all inhibitors, which can be roughly absorbed in the constant term in eq 4. Second, it is important to point out that the U_{elec} and U_{vdw} terms only include the changes in unbound ligand-solvent vs bound ligand-solvent plus bound ligand-protein interaction energies. Other terms including the changes in protein-solvent and solvent-solvent interactions are adsorbed in the α and β parameters and in the roughly constant term. The changes in protein-solvent and solvent-solvent interactions should be included in the desolvation energy. Third, the entropy contribution from the overall translational and rotational entropy, rearrangement of the water molecules, as well as the internal entropy changes of the inhibitors should be significant.^{50,51} In the current work, the hydrophobic P_1' group can produce large

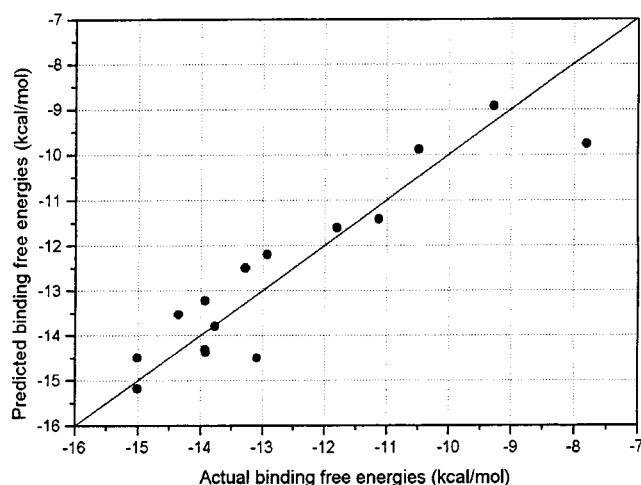


Figure 6. Observed versus calculated binding free energies for the 15 compounds binding to proMMP-2 using the three-parameter model5 in Table 4.

surface contacts with the S1' subsite of proMMP-2. When the inhibitors associate with proMMP-2, the water molecules near the P1' groups and in the S1' cavity will be rearranged. The rearrangement of the water molecules will lead to a significant entropy decrease. Another issue we must consider is the treatment of internal entropy changes deriving from, e.g., torsional degrees of freedom. Generally, the entropy contributions in the molecular docking processes are quite complicated, and they are very difficult to be estimated quantitatively.

In eq 1, the interpretation of the polar part is rather straightforward, which is represented as the free energy arising from electrostatic interactions between the ligand and its surroundings, while the nonpolar term is less obvious interpreted. Clearly, it is not the actual "van der Waals energy" itself that is the physical origin of this free energy contribution. In the work of Hansson et al., the authors thought that $\langle U_{\text{vdw}} \rangle$ measured two very important quantities, namely (1) the size of the solute and (2) the heavy atom number density (ρ) of the surroundings. The quantity $\langle U_{\text{vdw}} \rangle$ measures not only packing effects, but also the fraction of protein around the ligand. This fraction can be expected to correlate with hydrophobicity. Hydrophobicity combined with ligand size, in turn, determines the magnitude of the hydrophobic effect, which probably constitutes the major part of the $\Delta G_{\text{b}}^{\text{nonpolar}}$ term. But until now, it has not been validated that the quantity $\langle U_{\text{vdw}} \rangle$ can be well correlated with the entropic contribution in desolvation energy. We also believed that the $\langle U_{\text{vdw}} \rangle$ cannot fully consider the entropic effects in the docking processes. The exact implication of the quantity $\langle U_{\text{vdw}} \rangle$ needs more profound studies.

In principle, for different receptor–ligand systems the entropy contribution to ligand binding should be different. That is to say, the constant δ should vary according to different systems. The studied inhibitors here are quite similar and their entropic contributions to the binding free energies do not bear large differences, so introducing a constant term will improve the correlation. Jones-Hertzog and co-workers have added a SASA term for correlation.²² They found that the reduction in SASA for these systems falls in a relatively narrow range, 564–698 Å. Multiplication times the γ of 0.014 yields an attractive term between 7.9 and 9.8 kcal/mol. In their work, a fit was also made to eq 2 with the SASA term replaced by a constant, δ ; the result was $\alpha = \beta = 0.124$, $\delta = -8.23$ kcal/mol. The δ value from Jorgensen et al. is quite close to that from us in this paper. The

entropy contribution to ligand binding are sensitive to the combination of the protein's structure and the ligand's structure, and that value for different systems should be different. So, the precondition of introducing the constant term is that the studied inhibitors must possess similar structures and adopt a binding mode similar to that of the receptor. In the works of Hansson et al., the authors found that the consideration of the constant term could not improve the correlation.¹⁹ In their work, the authors employed different kinds of ligand–receptor systems. The entropy contributions for these systems have remarkable differences, and it certainly cannot be simply expressed by a constant term. It seems that if we cannot propose an efficient way to represent the entropic effects, it is very difficult to obtain a universal model to predict the ligand binding; furthermore, the applications and the future developments of the LIE approach will be greatly limited.

It may be noted that the value of α obtained here is only 0.191, which is much smaller than 0.5 from the Born equation for the solvation of atomic ions in a uniform dielectric continuum. In previous work, values of α near 0.3 and 0.2 have been found to be appropriate for a small neutral solute in the MD and MC studies with explicit solvent models for water and chloroform, respectively.^{18,19,51,52} In Jones-Hertzog's work, the mix of zwitterionic inhibitors and inhibitors with a net charge of +1 has led to further damping of the electrostatic contribution to 0.146.²² In the current work, due to the charge transfer from the hydroxamate oxygens to the catalytic center, the studied inhibitors are positive charged. So the combination of the large solute size, the specific protein environment, and the positive charged inhibitors has led to the relatively small value of α .

The Predicted Structures of the Complexes. Almost all ligands contain the hydrophobic side chains on P1' and P2' substituent sites (see Figure 4), respectively. The P1' groups of the inhibitors is in the protein interior, whereas the P2' group is solvent-exposed. Most of the effective ligands have a hydrophobic group in the P1' position and take advantage of the favorable van der Waals and hydrophobic interactions with the S1' pocket of MMPs. Hydrophilic interactions such as hydrogen bonds between the ligands and receptor also significantly contribute to ligand binding. Figure 7 shows the stereoplot of compound a4 in the active site. The illustrations are taken from the 100ps MD data collection stage in the averaging runs. All water molecules have been removed for clarity. As reflected in Figure 7, the following structural requirements are demonstrated: (1) the inhibitors must possess strong zinc-binding functionalities; (2) the hydrophobic P1' group with suitable size will produce extensive surface contacts in the S1' hydrophobic pocket; (3) several hydrogen bonds are observed to be maintained throughout the simulations. During the entire simulation time, the coordination of the inhibitor's hydroxamate group to the catalytic zinc atom was maintained very well, which can be suggested by the low rms fluctuations of these five coordinate bonds. The hydrophobic P1' group of a4 is buried deeply in the protein and contacts closely with the hydrophobic S1' pocket. The P2' and P3' groups are partly solvent-exposed. After careful observations, it can be found that an optimal steric complementarity between the P1' group and the S1' subsite is produced. In the S1' pocket, four residues, including Leu 197, Val 198, Leu 218, and Tyr 223 constitute a relatively large hydrophobic core, which can generate strong van der Waals and hydrophobic interactions with the inhibitors. Moreover, a4 can form five hydrogen bonds with Gly 162, Leu 164, Glu 202, Pro 221, and Tyr 223 of proMMP-2, which is very important to stabilize the bound a4 in the active site.

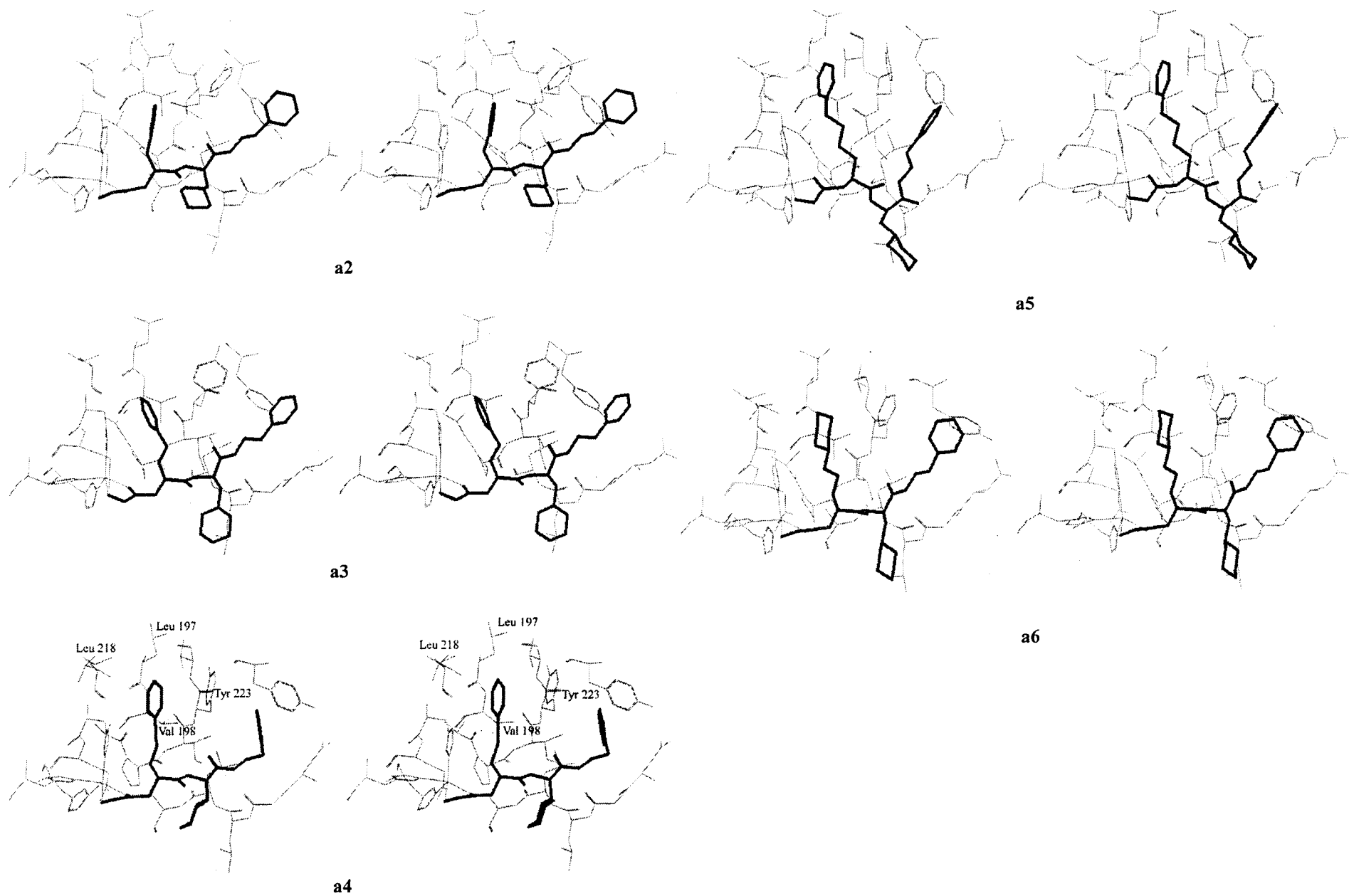


Figure 7. Stereoplots of the binding site for the complexes of a2, a3, a4, a5, and a6 with proMMP-2 from averaging the MD trajectories.

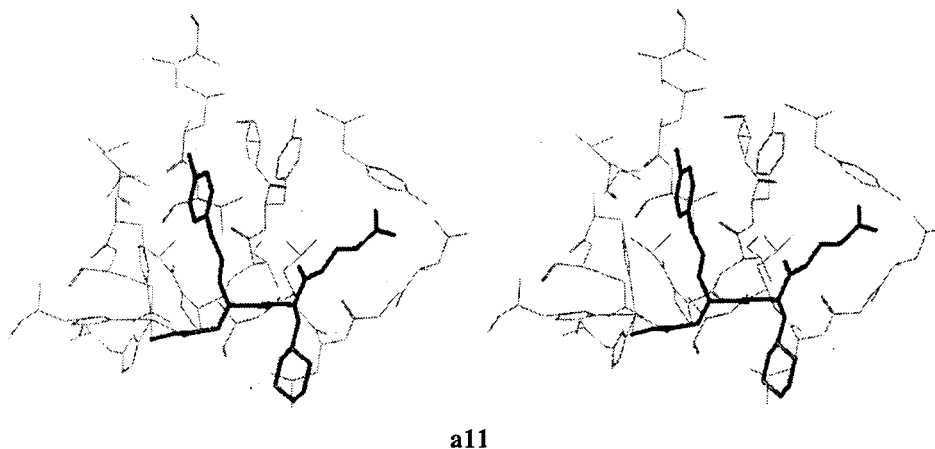


Figure 8. Stereoplots of the binding site for the complexes of all with proMMP-2 from averaging the MD trajectories.

Insights on the variation in the observed binding affinities for a1–15 emerge from the structures of the complexes obtained from the MD simulations. For examples, the change of the binding affinities of a2 to a5 clearly demonstrate the variation of binding affinities with the change of P_1' groups. The hydroxamates a2 to a4 show a clear stepwise increase in MMP-2 binding affinities according to the straight elongation of the P_1' group. While the P_1' of a4 is elongated of P_1' by one CH_2 to $(\text{CH}_2)_4\text{Ph}$, the binding free energy of a5 is decreased. Figure 7 shows the stereoplots of the inhibitors a2, a3, a4, and a5. In viewing the structure for the a2 complex in Figure 7, it is obvious that the nonpolar P_1' group is in only close contact with the hydrophobic side chain of Val 198 and Tyr 223. The relatively short P_1' group of a2 seems not to move into the interior of the S_1' subsite.

In viewing the structures for the a3 complex in Figure 7, two key contributors to the 1000-fold reduction in K_i relative to a2 are evident. First, the longer P_1' group can be located in the deeper interior of the S_1' pocket. The longer P_1' groups can produce effective surface contacts with not only the nonpolar side chains of Val 198 and Tyr 223, but also those of Leu 197 and Leu 218. As a result, a3 can produce more favorable van der Waals interaction (-65.752 kcal/mol) with its environments (including protein and solvent) than a4 (-62.167 kcal/mol) can do. Second, a3 can form four stable hydrogen bonds with proMMP-2, while a2 can only form three stable hydrogen bonds. The electrostatic interaction of a3 with its environments (-63.437 kcal/mol) is obviously stronger than that of a2 with its environments (-55.298 kcal/mol). The structure for the complex of a4 in Figure 7 shows the P_1' group of a4 can form more sufficient surface contacts with the S_1' pocket, which leads to a large reduction of van der Waals interactions (-72.609 kcal/mol). From the MD trajectories for a4, we find that the benzene rings of the inhibitors and Tyr 223 prefer to be parallel. The distance between their mass centers of two benzene rings is about $4\sim 5$ Å, and we believe that the pair parallel benzene rings can produce relatively strong aromatic stacking interactions. The stronger van der Waals interaction of a4 leads to a further 36-fold increase in binding relative to a3.

Compared with a4, the binding affinity of a5 shows a 4-fold reduction. From the complex of a5 in Figure 7, the P_1' group of a5 can also produce good surface contacts with the S_1' subsite. The van der Waals interaction of a5 (-81.01 kcal/mol) with its environment is favorable, which is stronger than that of a4. So, only from the analysis of the van der Waals interactions, the binding capability of a5 should be stronger than that of a4. The reduced binding affinity of a5 may be caused by the unfavorable electrostatic interactions. Table 3 shows that

the electrostatic interaction (-51.48 kcal/mol) and coordinating bonding energy (28.94 kcal/mol) of a5 with its environments are really weaker than those of a4 with its environments. Because the volume of the P_1' group of a5 is larger than that of a4, if the P_1' group of a5 are fully buried in the S_1' pocket, bad van der Waals contact between a5 and proMMP-2 will be formed. To produce the best steric complementarity between a5 and proMMP-2, some part of the P_1' group of a5 will move outside the S_1' subsite; consequently, some hydrogen bonds and the coordinate bonds of a5 will be weakened.

For the compound a6 in Figure 7, displacement of the benzene group of a4 to the cyclohexyl group produces a 100-fold reduction of binding affinity. The van der Waals interactions between a6 with its environments (-75.992 kcal/mol) are slightly stronger than those between a4 with its environments (-72.609 kcal/mol). Considering the van der Waals interaction between the unbound a6 with the water molecules also slightly stronger (-49.138 kcal/mol) than those between the unbound a4 with the water molecules (-47.086), the ΔU_{vdw} for these two molecules are similar. But the electrostatic interactions between a6 and its environments are weaker than those between a4 with its environments. Moreover, from the electrostatic interactions for the unbound inhibitors, a6 demonstrates a stronger capability of hydration, so the ΔU_{elec} of a6 is unfavorable to improve the binding affinity. In going from a2 to a6 in Figure 7, the strong dependence of ligand binding to length of the P_1' group in the S_1' hydrophobic pocket is clearly demonstrated. The buried P_1' group should bear adequate length and adopt specific orientation to produce the optimal steric complementarity. Too short P_1' groups cannot produce good steric complementarity and strong van der Waals interactions with the S_1' subsite, while too long P_1' groups will affect the overall conformation of the inhibitors, and weaken the electrostatic interactions between ligand and receptor. Meanwhile, the hydration ability for the unbound inhibitor is also very important. Too strong hydration ability is unfavorable to the ligand binding.

Compounds a7 to a10 bear similar structures, the only differences for them are the substituents linked to the benzene rings of the P_1' groups. Meanwhile, their binding free energies do not show obvious differences. The predicted complexes for them are quite similar to that of compound a4. We think that it is not very meaningful to perform further structural and energetic analyses for them. Compared with compounds a1 to a10, a11 and a12 adopt the different P_3' groups. The structural analyses of the MD trajectories illustrate that two additional hydrogen bonds can be generated between the carboxyl oxygens of a11 and proMMP-2 (Figure 8). The two carboxyl oxygen atoms of a11 can form two hydrogen bonds with the O atom of Pro 221

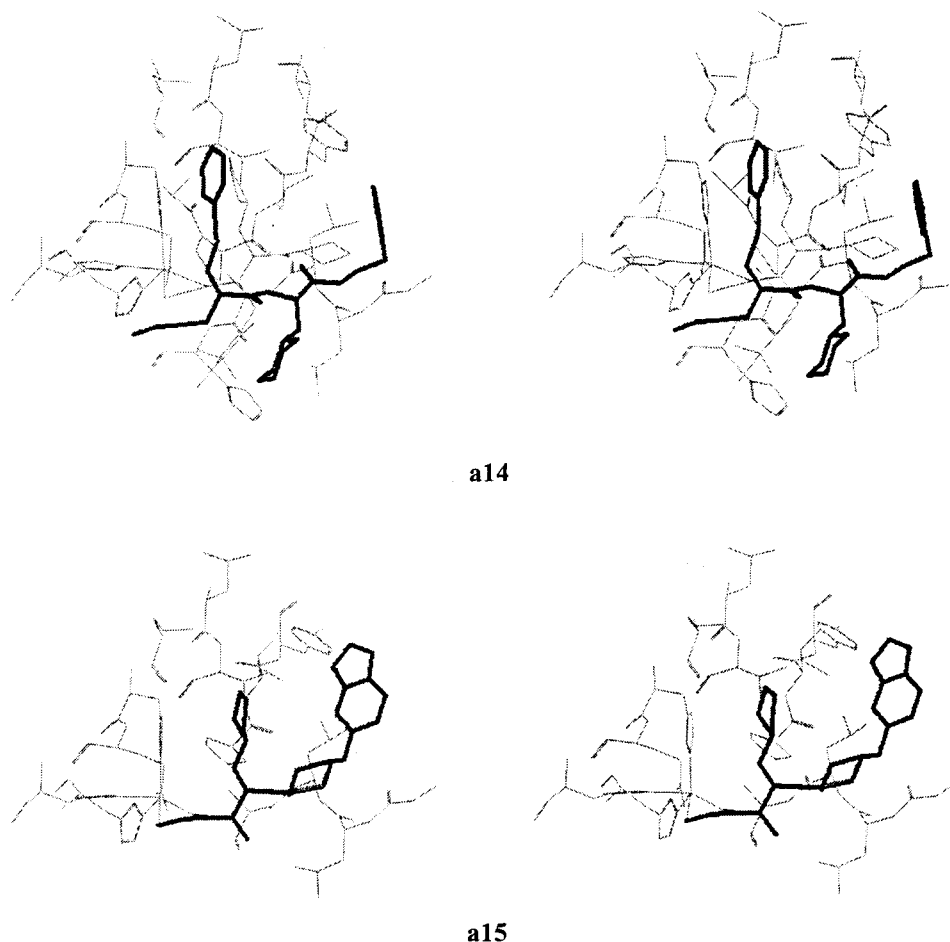
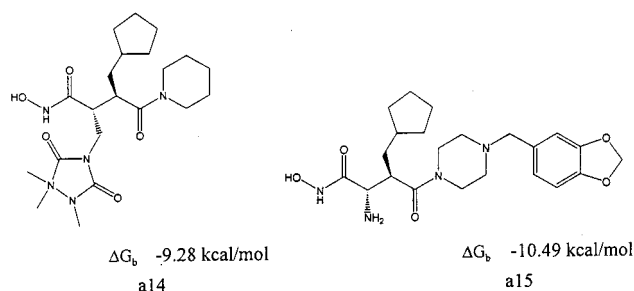


Figure 9. Stereoplots of the binding site for the complexes of a14 and a15 with proMMP-2 from averaging the MD trajectories.

and N atom of Tyr 223. The formation of the hydrogen bonds can significantly increase the electrostatic interactions between a11 and its environments (-67.751 kcal/mol). Although the electrostatic interactions between the unbound a11 and the water molecules are also very stronger, the ΔU_{elec} of a11 is also favorable to improve the ligand binding. For a12 and a13, one hydrogen bond can also be formed between the oxygen or nitrogen atom in the P_3' group with the oxygen atom in Gly 162. So we can conclude that introduction of polar atoms in the P_3' groups will be favorable to produce hydrogen bonds and enhance the binding affinities.

From the energetic analyses, it can be found that the van der Waals interactions between a14 and a15 with their environments are obviously weaker than those between the other inhibitors with their environments. Although the van der Waals interactions between unbound a14 and a15 with the water molecules are slightly weaker than those between the other unbound inhibitors with the water molecules, the higher values of ΔU_{vdw} of a14 and a15 are unfavorable to ligand binding. The structures of the complexes of a14 and a15 show that the P_1' groups cannot form very close surface contacts with the S_1' pocket (Figure 9). The electrostatic interaction between a14 with its environments is similar to that between a15 with its environments. But compared with a15, a14 seems to possess more strong hydration ability, and the electrostatic interaction between unbound a14 with the water molecules (-76.876 kcal/mol) are stronger than that between a15 with the water molecules (-64.980 kcal/mol). That is to say, a14 bears more unfavorable desolvation energy from the enthalpic point of view than a15.

CHART 1: Structures and Experimental Binding Free Energies of a14 and 15



Conclusion

The binding of hydroxamates to proMMP-2 has been studied by means of linear response approaches in conjunction with molecular dynamics simulations. The q^2 from the leave-one-out cross-validation (0.854), rms deviation of 0.79 kcal/mol, and average unsigned error of 0.6 kcal/mol for the observed range of 7.2 kcal/mol in binding affinities shows that the obtained model possesses effective predictive ability. This series of similar inhibitors, a three-parameter LIE model with a constant term is obviously superior to the usually used two-parameter LIE model. The constant term is able to give effective adjustment to the entropy contribution in the binding free energies. From the MD simulations, it can be found that the ligand hydroxamate group coordinating to the catalytic zinc center is very important for ligand binding, but introduction of the coordinate bond energy term cannot improve the correlation.

The coordinate bonds seem to give relatively equal effects to the binding of all studied inhibitors, so the effect of the coordinate bond is implicitly absorbed in the constant term.

The structural and energetic findings from the MD simulations also provided insights into the variations in observed binding affinities. The P₁' group makes extensive van der Waals and hydrophobic contacts with the hydrophobic S1' pocket. The steric complementarity between the P₁' groups and the S1' pocket directly affects the van der Waals interactions between ligand and receptor. The hydrogen bonds between hydroxamates and proMMP-2 are also very important to stabilize the inhibitors in the active site. The generation of the additional hydrogen bonds between P₃' group and proMMP-2 will significantly improve the electrostatic interactions. Though further work is needed, especially the development of the valid estimation of the entropic effects, the illustrated energetic and structural characterization of enzyme inhibition indicates that the linear interaction energy approach based on MD simulations has much promise as a tool for ligand design.

Acknowledgment. We are particularly grateful to Prof. P. A. Kollman of UCSF for his distribution of AMBER 6.0 molecular simulation package and the newest Amber force field parameters. We also thank Dr. J. M. Wang in UCSF for his helpful discussions of the methods about MD simulations and free energy calculations. This project is supported by National Natural Science Foundation of China (NSFC 29992590-2).

Supporting Information Available: The studied inhibitors, the catalytic zinc atom and the histidines coordinated to the zinc atom are defined as new residue types in AMBER. The coordinates and the partial charges from ESP fitting calculations for these new residues are saved as AMBER database files. The missing parameters are added to the AMBER force field file. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- Hou, T. J.; Wang, J. M.; Chen L. R.; Xu, X. J. *Protein Eng.* **1999**, *12*, 639.
- van Gunsteren, W. F. *Protein Eng.* **1988**, *2*, 5.
- Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184.
- Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
- Tomioka, N.; Itai, A.; Iitaka, Y. *J. Comput.-Aided Mol. Design* **1987**, *1*, 197.
- Novotny, J.; Bruccoleri, R. E.; Saul, F. A. *Biochemistry* **1989**, *28*, 4735.
- Williams, D. H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, U.; Kaye, P. T.; Lal, A. R.; Nicholls, I. A.; Salter, C. J.; Mitchell, R. C. *J. Am. Chem. Soc.* **1991**, *113*, 7020.
- Searle, M. S.; Williams, D. H.; Gerhard, U. *J. Am. Chem. Soc.* **1992**, *114*, 10697.
- Horton, N.; Lewis, M. *Protein Sci.* **1992**, *1*, 169.
- Bohacek, R. S.; McMartin, C. J. *Med. Chem.* **1992**, *35*, 1671.
- Krystek, S.; Stouch, T.; Novotny, J. *J. Mol. Biol.* **1993**, *234*, 661.
- Böhm, H. J. *J. Comput.-Aided Mol. Design* **1994**, *8*, 243.
- Head, R. D.; Smythe, M. L.; Oprea, T. I.; Waller, C. L.; Greene, S. M.; Marshall, G. R. *J. Am. Chem. Soc.* **1996**, *118*, 3959.
- Matthew, D. E.; Christopher, W. M.; Timoghy, R. A.; Gaia, V. P.; Roger, P. M. *J. Comput.-Aided Mol. Design* **1997**, *11*, 425.
- Åqvist, J.; Medina, C.; Samuelsson, J. E. *Protein Eng.* **1994**, *7*, 385.
- Åqvist, J.; Mowbray, S. M. *J. Biol. Chem.* **1995**, *270*, 9978.
- Hansson, T.; Åqvist, J. *Protein Eng.* **1995**, *8*, 1137.
- Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512.
- Hansson, T.; Marelius, J.; Åqvist, J. *J. Comput.-Aided Mol. Design* **1998**, *12*, 27.
- Paulsen, M. D.; Ornstein, R. L. *Protein Eng.* **1996**, *9*, 567.
- Wang, J.; Dixon, R.; Kollman, P. A. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 69.
- Jones-Hertzog, D. K.; Jougensen, W. L. *J. Med. Chem.* **1997**, *40*, 1539.
- Woessner, J. F. *FASBE J.* **1991**, *5*, 2145.
- Nagase, H.; Fields, G. B. *Biopolymer* **1996**, *40*, 399.
- Bottomley, K. M. K.; Bradshaw, D.; Nixon, J. S. *Metalloproteinases as Targets for Antiinflammatory Drugs*; Birkhäuser Verlag: Switzerland, 1999.
- Willenbrock, F.; Murphy, G. *Am. J. Respir. Crit. Car. Med.* **1994**, *150*, 5165.
- Green, J.; Wang, M.; Liu, Y. E.; Raymond, L. A.; Rosen, C.; Shi, Y. E. *J. Biol. Chem.* **1996**, *271*, 30375.
- Murphy, G. J. P.; Murphy, G.; Reynolds, J. J. *FEBS Lett.* **1991**, *289*, 4.
- Muller, D.; Quantin, B.; Gesnel, M. C.; Millon-Collard, R.; Abecossis, J.; Breathnach, R. *Biochem. J.* **1988**, *253*, 187–912.
- Liotta, L. A.; Steeg, P. S.; Stetler-Stevenson, W. G. *Cell* **1991**, *64*, 327.
- Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. H. *Chem. Rev.* **1999**, *99*, 2735.
- Stetler-Stevenson, W. G.; Aznavoorian, S.; Liotta, L. A. *Annu. Rev. Cell. Biol.* **1993**, *9*, 541.
- Beckett, R. P.; Davidson, A. H.; Drummond, A. H. *Drug Discov. Today* **1996**, *1*, 16.
- Morgunova, E.; Tuuttila, A.; Bergmann, U.; Isupov, M.; Lindqvist, Y.; Schneider, G.; Tryggvason, K. *Science* **1999**, *284*, 1667.
- Toba, S.; Damodaran, K. V.; Merz, K. M., Jr. *J. Med. Chem.* **1999**, *42*, 1225.
- Hoops, S. C.; Anderson, K. W.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 8262.
- Natchus, M. G.; Cheng, M. Y.; Wahl, C. T. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2077.
- Porter, J. R.; Beeley, N. R. A.; Boyce, B. A.; Mason, B.; Millican, A.; Millar, K.; Leonard, J.; Morphy, J. R.; O'Connell, J. P. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2741.
- Broadhurst, M. J.; Brown, P. A.; Lawton, G.; Ballantyne, N.; Borkakoti, N.; Bottomley, K. M. K.; Cooper, M. I.; Eatherton, A. J.; Kilford, I. R.; Malsher, P. J.; Nixon, J. S.; Lewis, E. J.; Sutton, B. M.; Johnson, W. H. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2299.
- Alpegiani, M.; Bissolino, P.; Absate, F.; Perrone, E.; Corigli, R.; Jabes, D. *Chem. Abstr.* **1999**, *130*, 139360.
- InsightII User Guide, Molecular Simulations Inc., San Diego, 1999.
- Ryde, U. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 40.
- Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
- Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117* (7), 5179.
- MOPAC 7.0 User Guide, Quantum Chemistry Program Exchange (QCPE), Indiana University, 1993.
- Gaussian 98 User Guide, Gaussian, Inc., Pittsburgh, 1998.
- Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *J. Comput. Chem.* **1994**, *16*, 1357.
- Keil, M.; Exner, T.; Brickmann, J. *J. Mol. Model.* **1998**, *4*, 335.
- SYBYL ver. 6.5, Tripos Associates, St. Louis, MO., 1999.
- Janin, J. *Proteins: Struct., Funct., Genet.* **1994**, *24*, R1.
- Carlson, H. A.; Jorgensen, W. L. *J. Phys. Chem.* **1995**, *99*, 10667.
- McDonald, N. A.; Carlson, H. A.; Jorgensen, W. L. *J. Phys. Org. Chem.* **1997**, *10*, 563.