Calculation and Prediction of Binding Free Energies for the Matrix Metalloproteinases

Oreola A. T. Donini and Peter A. Kollman*

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0446

Received January 31, 2000

The zinc-dependent matrix metalloproteinases are drug targets of interest for diseases ranging from arthritis to cancer. Unfortunately, the use of computational rational drug design has been limited by the challenges introduced by the zinc center. We present an extension of the MM/PB/SA methodology which allows us to calculate the relative binding energies of six known nanomolar carboxylate ligands of MMP-1. We are able to rank the neutral and charged ligands correctly. We further illustrate the utility of our approach by modifying the best-binding ligand of our set and predicting a better binding ligand.

1. Introduction

The matrix metalloproteinases (MMPs) are zinc-dependent enzymes that have been implicated in a variety of disease states ranging from cancer[1,2] to arthritis and multiple sclerosis.[1,3] These tightly regulated enzymes are primarily involved in the degradation of the extracellular matrix (ECM) that forms the connective material between cells and around tissues. Disease processes associated with the MMPs are generally related to an imbalance between the inhibition and activation of the MMPs resulting in excessive degradation of the ECM.[4] Thus, inhibitors of the MMPs make attractive pharmacological drug targets. The MMPs can be further partitioned into subfamilies according to which components of the extracellular matrix they regulate. More recent research has begun to identify MMP subfamilies with disease processes; although, this process has been hampered by a lack of specific inhibitors for the MMPs.[4]

There are three major components to most MMP inhibitors—the zinc binding group (ZBG), the peptidic (or peptidomimetic) backbone, and the pocket-occupying side chains. Most MMP inhibitors are classified according to their ZBG. An empirical ranking of the potency of the ZBGs is approximately hydroxamate ≫ sulfydryl > phosphinate > aminocarboxylate > carboxylate.[4] Many of the currently designed inhibitors for the MMPs are based on hydroxamate zinc-binding moieties coupled to a peptidic framework. Unfortunately, these inhibitors are associated with significant direct toxicity and biological lability,[5] making their usage problematic, particularly in non-cancer-related therapies.

While the need for specific inhibitors of the MMPs is debatable in cancer treatment, it is clear that in other disease processes and for the purpose of elucidating disease mechanisms, specific inhibitors of the matrix metalloproteinase subfamilies are required. The need for novel, selective MMP inhibitors, as well as the plethora of high quality structural information, makes them an attractive target for the use of molecular modeling.

Unfortunately, metallo centers have long been a challenge for molecular modeling. Zinc in particular is problematic because of its promiscuous ability to assume a variety of coordination states.[6] While there have been both nonbonded[7−10] and bonded[11,12] models suggested for the parametrization of zinc, neither is entirely satisfactory. The nonbonded models tend to generate the highest coordination number possible (i.e., octahedral) while the bonded model requires "freezing" a specific zinc coordination from the outset and potentially reduces the flexibility of ligand conformational sampling. In neither case can we predict the correct binding geometry a priori using molecular mechanics models. Alternatively, it has recently been suggested that the zinc ion be modeled as a tetrahedral charge distribution;[13] however, this model requires the interpretation that all zinc coordination states are tetrahedral and that the higher coordination states ostensibly seen in the crystal structures are the result of the length of time over which the "snapshot" is taken.[14]

With the eventual aim of producing novel, selective inhibitors of the matrix metalloproteinases, we present a modeling methodology that allows us to rank six known carboxylate MMP-3 inhibitors with nanomolar inhibition constants (see Figure 1, Table 3)[15] despite the challenges presented by the catalytic zinc center. The free energy of binding includes contributions from the protein and ligand flexibility in the complex, the cost of desolvating the ligand and receptor, and the entropic cost of forming the complex. The flexibility of the complex is modeled using molecular dynamics (MD) in the presence of explicit solvation, while the desolvation energy is calculated using a continuum model by post-processing the MD trajectory stripped of the explicit waters (using the MM-PBSA approach)[16−18] and the entropy is estimated with a simple single-point normal mode calculation. This methodology has been previously used to successfully model relative DNA and RNA stabilities,[16−18] protein−RNA interactions[19] (Reyes and Kollman, submitted), protein−protein interactions,[20] and protein−small molecule interactions.[21] This is the first application of this methodology to ranking the relative binding constants of potential drug molecules and the first system in which an inherent metal ion is considered.

10.1021/jm000040d CCC: $19.00
Published on Web 10/14/2000

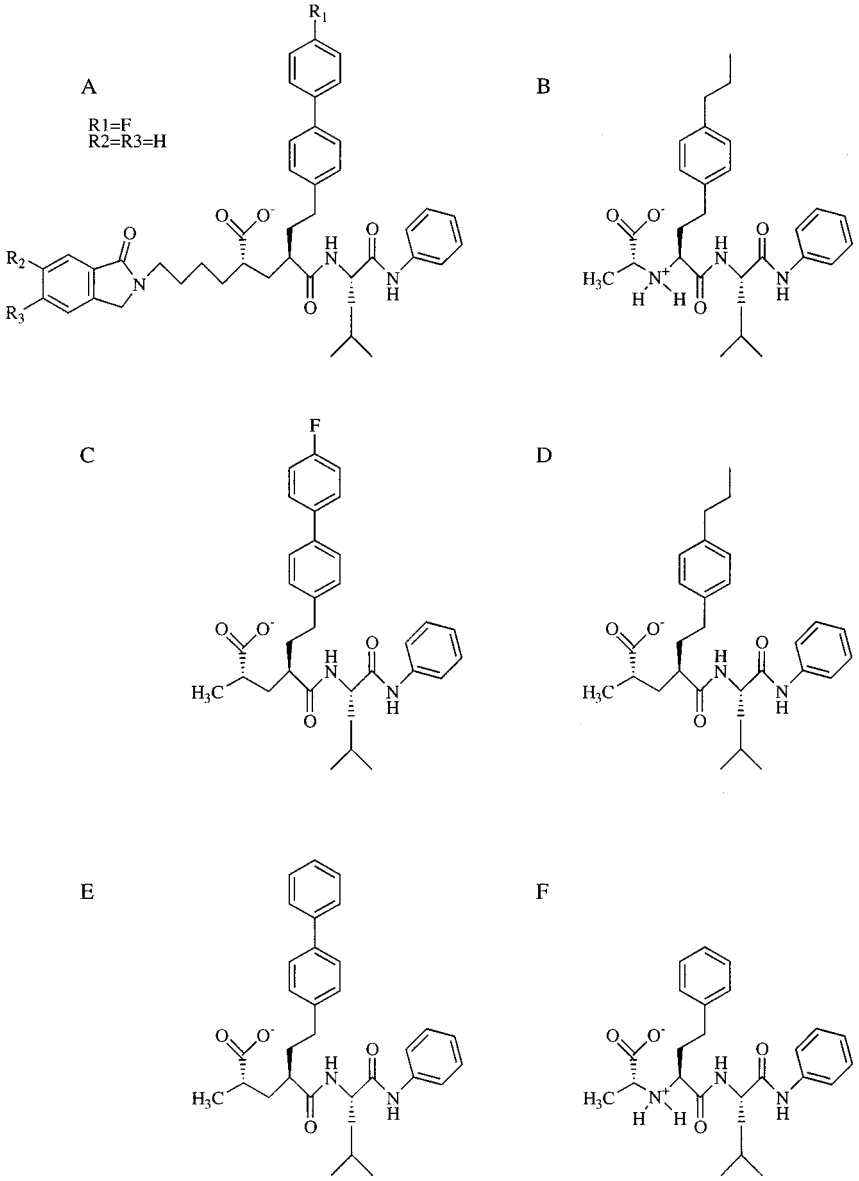

Figure 1. Structure of ligands.[15] The binding constants and relative binding energies are given in Tables 3, 5, and 6. Note that in all cases the zinc binding group is the carboxylate, generally believed to be tetrahedrally coordinated using only one oxygen of the carboxylate. Ligands are shown in the charge states that were used in the computations.

The relative inexpense of this method lies in the fact that while other methodologies rank the free energy of ligands in a pairwise manner (e.g., thermodynamic integration[22]) or require multiple trajectories and system-dependent parameters (e.g., Aqvist linear interaction energy approach[23]), this methodology has the potential to rank tens of ligands simultaneously. In this case, we rank six ligands based on a single trajectory. The single trajectory is generated with the largest and best binding ligand. Selected conformations from the trajectory are then used to assess the binding of other ligands. Particularly in the context of the challenges presented by the zinc ion, this proves to be an extremely efficient approach.

Finally, we attempt to propose a better binding ligand by modifying the best ligand of our test set. This process involves an increasing scale of computations from PROFEC analysis[24] to the approach highlighted here (substitution into the previously generated trajectory) to analysis of individual trajectories. The overall approach is a hierarchical one in which only the best scoring ligands of the previous step are carried through to the final calculations.

2. Methods

2.1. Force Field Description. All calculations were performed using the Amber5.0 molecular modeling package[25] with the parm94 parameter set.[26] Ligand charges were determined by RESP[27] analysis of the electrostatic potential generated by ab initio (Gaussian94,[28] HF 6-31G*) calculations on common fragments among the ligands (see Supporting Information). Some ligand geometric parameters were also fitted to ab initio calculations (see Supporting Information).

The nonbonded model for the zinc ion was adopted because of considerations involved in the MM-PBSA approach. The MM interaction energy and PB solvation energy are calculated for both the complex and the isolated receptor. In the bonded model, the zinc charge is parametrized in a manner dependent on the identity of the complexing groups. Since the complex would have four bonds to zinc (three histidines, one ligand) while the receptor would have three (three histidines), the bonded model would give rise to a discrepancy in the assigned atomic charges between the complex and the isolated ligand and receptor.

The zinc and calcium ions were assigned the full formal charge of 2+. The calcium ion VDW radius was taken from the Aqvist parameter set.[29] The zinc ion VDW radius was taken from the work of Stote and Karplus.[8]

2.2. Molecular Dynamics. The crystal structure of MMP-3 (Stromelysin-1) with the "A" ligand was used as the starting point for all simulations (1HFS,[15] 1.7 Å resolution). The structures of ligands B, C, D, E, and F were built from the coordinates of ligand A as found in the crystal structure using the *xleap* module of Amber.

Molecular dynamics trajectories were run (*sander* module of Amber) in the presence of a solvent cap of water extending 24 Å around the catalytic zinc ion. Three sodium ions were added as counterions to neutralize the system in the presence of a negatively charged ligand. The water positions were energy minimized in the presence of the frozen solute. Thereafter, only atoms within 16 Å of the catalytic zinc ion were allowed to move. All simulations included a SHAKE constraint on covalent bonds to hydrogen, a time step of 1 fs, a nonbonded residue-based cutoff of 14 Å, and separate temperature coupling (Berendsen algorithm[30]) of solvent and protein. Simulations were run at a constant temperature of 300 K and included 2000 steps of minimization, and heating and equilibration for 80 ps. Simulations were run in the presence of a "query ligand", and snapshots collected over the course of the trajectory were used to rank the relative binding energies of the other ligands.

For comparison purposes, we ran the trajectories with the A (500 ps), E (200 ps), and F (100 ps) ligands (Figure 1). Two trajectories were run for each of the E and F ligands. In the initial trajectories for E and F, the waters and solute were minimized concurrently, resulting in zinc coordination with the sixth ligand as glu202 (OE2 for E, OE1, for F). These trajectories are identified as E_{glu} and F_{glu}. In the second set of trajectories (E_{hoh} and F_{hoh}), the water positions were energy minimized in the presence of the frozen solute (as was done for the A trajectory), and the sixth coordinating ligand for zinc was water. The E and F trajectories were generated from the starting structure of the A trajectory (complete with water and counterions) by replacing the A ligand with E or F. Since the F ligand is neutral (Figure 1), the net charge of the simulated system was +1. These trajectories allow us to determine the effect of varying both the query ligand and the zinc coordination.

2.3. Determination of the Free Energy of Binding. The total free energy of binding is described by three contributions (eq 1). These include the molecular mechanical description of the interaction energy for each snapshot, the continuum energy assessment of the cost of desolvation, and the entropic contribution of complex formation.

$$\Delta G_{bind} = \langle E_{MM} \rangle + \Delta G_{solv} - T\Delta S \quad (1)$$

The gas-phase molecular mechanical (MM) energy is averaged over all the snapshots and includes contributions from the electrostatic and van der Waals interactions for the complex, receptor, and ligand. All MM calculations (*anal* module of Amber) are performed with a nonbonded cutoff of 99 Å and a dielectric constant of 1 in the absence of any solvent or counterions. The receptor and ligand geometries are taken from that of the complex, and thus there is no internal energy (i.e., bonds, angles, and dihedrals) contribution to the net MM average.

The entropic contribution is included by minimizing a structure of the complex, receptor, and ligand for each ligand in the presence of a distance-dependent dielectric ($\epsilon = 4r_{ij}$) and a 99 Å nonbonded cutoff to a tolerance in the rms gradient of 8×10^{-5} kcal/(mol Å) using conjugate gradient minimization (*sander* module of Amber). Subsequent normal-mode analysis (*nmode* module of Amber) results in an estimate of the entropy of the complex, receptor, and ligand. To determine the error involved in the entropic calculation, we followed this procedure for both the first and last snapshot of the mmpD complex (difference ~0.5 kcal/mol) and therefore calculated an error in the net entropic contribution of ~1.5 kcal/mol (= 3 × 0.5 kcal/mol to account for errors in the receptor, complex, and ligand entropy determinations, Table 1). However, since normal-mode analysis is based on the harmonic approximation, there is also likely to be a significant systematic error that is not included. This calculation of the entropy is to be considered a crude estimate only.

Table 1. Entropic Contributions to the Binding Energy[a]

complex	S complex (cal/mol K)	S receptor (cal/mol K)	S ligand (cal/mol K)	$-T\Delta S$ (kcal/mol)
MmpA	6232.1 (0.056)	5994.6 (0.032)	308.8 (0.115)	21.4
MmpB	6158.1 (0.053)		239.1 (0.000)	22.5
MmpC	6148.2 (0.045)		244.0 (0.041)	26.9
MmpD	6140.8 (0.059)		238.2 (0.025)	27.4
MmpE	6144.8 (0.037)		239.1 (0.012)	26.5
MmpF	6136.0 (0.052)		215.8 (0.026)	22.2
MmpD last	6142.8 (0.057)		as above	27.0

[a] The entropy of the complex, receptor (calculated once), and ligand is given with the largest frequency of the first six modes indicated in brackets. "Last" indicates that normal-mode analysis was repeated using the last snapshot from the A trajectory. All other initial conformations were taken from the first snapshot of the A trajectory.

The continuum approximation to the electrostatic desolvation energy includes the iterative, finite-difference solution of the linear Poisson−Boltzmann (PB) equation for the isolated complex, receptor, and ligand (DelPhi2.0[31]). The charges used in the PB calculation were taken from the Amber parameter set. DelPhi radii were taken from the PARSE parameter set.[32] In the case where no PARSE radii existed, we used the following parameters: F (1.45 Å), Ca^{2+} (1.97 Å), and Zn^{2+} (1.10 or 1.40 or 1.60 Å). The resulting absolute binding energy estimates are very dependent on the value of the zinc radius (Table 3). The Zn^{2+} radius of 1.40 Å was used in all further calculations (see section 3.3 for discussion). The iterative DelPhi calculations were run for 1000 steps, although the relative differences between receptor, complex, and ligand are expected to converge well before this point.[18] The probe molecular radius was 1.4 Å, and the PB contribution was calculated assuming a dielectric constant of 1 within the solute and 80 outside of the solute. No salt contribution was included in the PB calculation. The ligand and receptor geometries are taken from that of the complex. The hydrophobic contribution to the desolvation is calculated by using the solvent accessible surface area (SA)[33] of the isolated complex, receptor, and ligand (ΔG_{np} = SA × 0.00542 kcal/mol Å^2 + 0.92 kcal/mol).

The MM-PBSA calculations described above are carried out for the complex structures taken directly from the trajectory (relevant for the query ligand only), for the complex structures with the respective ligand energy optimized with 2000 steps of conjugate gradient minimization, and for the complex structures with the ligands and surrounding 5 Å belly (identified with the A ligand) of the enzyme energy optimized for 2000 steps of conjugate gradient minimization. All minimizations were carried out in the presence of the water and counterions.

2.4. Ligand Design. As a demonstration of the proposed utility of our approach and to provide experimentally verifiable predictions, we suggest improvements to the A ligand, based on our current methodology. We begin by proposing modifications using PROFEC analysis. These modifications are then tested using the minimization procedure based on the A trajectory. Promising candidates are further analyzed by running individual trajectories.

2.4.1. PROFEC. PROFEC (or pictorial representation of free energy changes) is an approximate method to calculate the free energy cost of adding a particle to a ligand. A series of grid points are constructed around the relevant part of the ligand, and the van der Waals potential energy is calculated at each point assuming a particle of $R = 2.0$ Å, and $E = 0.15$ kcal/mol was inserted there. The grids are constructed for snapshots of the ligand within the enzyme as well as the ligand solvated by a cap of water, and the free energy is calculated as an average over all the snapshots in each case. Comparison

Table 2. MM-PBSA Contributions to the Binding Energy Using the A Trajectory

complex[a]	$\Delta E\langle$elec$\rangle$	$\Delta E\langle$vdw$\rangle$	$\Delta G\langle$PB$\rangle$	$\Delta G\langle$SA$\rangle$	$\Delta G\langle$total$\rangle$
MMPA 0	−193.3 ± 9.2	−64.7 ± 4.6	264.9 ± 9.7	−8.5 ± 0.3	−1.6 ± 5.5
MMPA 1	−194.0 ± 4.8	−69.2 ±1.7	263.8 ±8.9	−8.0 ± 0.4	−7.3 ± 6.2
MMPB 1	−177.5 ± 10.5	−52.3 ± 4.5	236.5 ± 5.7	−6.2 ± 0.4	1.9 ± 6.1
MMPC 1	−185.1 ± 11.5	−53.8 ± 8.8	234.4 ± 6.9	−6.5 ± 0.4	−9.7 ± 5.4
MMPD 1	−182.6 ± 4.7	−49.7 ± 1.6	230.8 ± 6.0	−6.8 ± 0.3	−8.2 ± 3.6
MMPE 1	−185.3 ± 9.5	−53.1 ± 4.5	235.8 ± 6.6	−6.9 ± 0.3	−8.2 ± 5.1
MMPF 1	−175.5 ± 5.8	−43.1 ± 1.4	229.7 ± 5.2	−6.0 ± 0.3	5.1 ± 4.8
MMPA 2	−194.0 ± 2.4	−71.2 ± 1.0	261.6 ± 4.8	−8.5 ± 0.0	−12.0 ± 4.8
MMPB 2	−188.3 ± 4.0	−51.7 ± 0.8	229.9 ± 3.4	−6.6 ± 0.1	−16.7 ± 2.9
MMPC 2	−181.3 ± 2.7	−55.7 ± 1.0	229.9 ± 3.0	−7.0 ± 0.0	−14.1 ± 2.9
MMPD 2	−180.2 ± 3.2	−51.3 ± 1.0	225.1 ± 3.4	−6.7 ± 0.1	−13.0 ± 2.8
MMPE 2	−181.5 ± 2.7	−55.1 ± 1.1	231.0 ± 3.2	−6.9 ± 0.0	−12.5 ± 3.0
MMPF 2	−188.4 ± 3.8	−41.8 ± 0.8	222.3 ± 3.2	−5.9 ± 0.1	−13.8 ± 2.6

[a] "0" indicates no minimization occurred after the trajectory snapshots were obtained, "1" represents minimization of the ligand only in the presence of the protein and solvent cap, while "2" indicates minimization of a 5 Å belly in the presence of the solvent cap. The average value over the 49 (48 for MMPC 1) snapshots is reported ± the standard deviation. For all cases, $R_{Zn}^{PB} = 1.40$ Å.

of the two grids then illustrates whether there is a free energy benefit to the introduction of a particle at each grid point. The electrostatic contribution can also be determined by calculating the electrostatic potential at each grid point. The procedure is described in more detail in ref 24.

The PROFEC grid construction was carried out for the A trajectory run previously, as well as for the A ligand simulated in a cap of water. The simulation of the A ligand in water included the A ligand surrounded by a 24 Å cap of water. The water and ligand positions were energy optimized with 3000 steps of conjugate gradient minimization, and then all the atoms were heated and equilibrated for 35 ps. The final collection phase was run for 220 ps and yielded 23 snapshots.

Grid calculations were performed around the three rigid ring structures in the A ligand: the biphenyl group, the phenyl group, and the morpholino group (Figure 1). In each case the reference atoms were three ring atoms, and the grids were at least 21 Å × 21 Å × 15 Å in size. The grid calculations were performed at a temperature of 300.0 K and with a nonbonded cutoff of 14 Å.

2.4.2. Minimization. The proposed ligands (A_1−A_5, see Table 8, Figure 1) are substituted into the A trajectory. The ligand and the enzyme residues within 5 Å of the respective ligand (determined individually for each of A_1 through A_5) are energy minimized for 2500 steps of conjugate gradient optimization. MM-PBSA analysis and entropic contributions are calculated as described above.

2.4.3. Trajectories. Individual trajectories of the most promising candidates among A_1−A_5 are then calculated. In this case, the equilibrated structure from the A trajectory is modified to include the appropriate ligand, and the dynamics is continued for 400−500 ps. Snapshots are collected after 200 ps and are analyzed as previously described.

3. Results

3.1. Zinc Geometry. While the crystal structure of MMP-3 indicates a tetrahedral binding geometry for the catalytic zinc involving one carboxylate oxygen from the ligand and the three surrounding histidines (Figure 2), minimization and molecular dynamics invariably resulted in an octahedral geometry involving the three histidines, both carboxylate oxygens of the ligand, and either a water oxygen or a carboxylate oxygen from glu202. Once coordinated, the identity of the sixth ligand was constant for a given trajectory.

The incorrect binding geometry for zinc is potentially problematic. However, all the ligands in question (Figure 1) coordinate with Zn with the same functional group. Therefore any errors in the energetic or conformational assessment of the interaction between zinc and the ligand should be constant. The caveats to this assumption involve the movement of the side chains of the ligand in response to the change in binding coordination and the role of glu202 in interacting with the ligand versus the zinc ion. The first assumption was tested by running a 200 ps trajectory. When the water in the solvent cap is initially minimized, the side chains of the A ligand do not deviate from their pockets nor appear to respond to changes in the binding coordination. The second assumption is addressed with the trajectories run on ligands E and F discussed below.

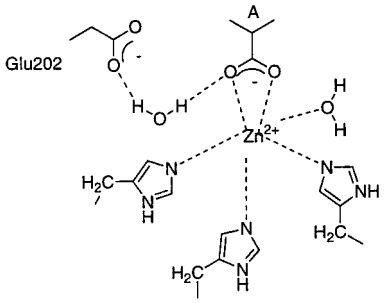

Figure 2. Orientation of zinc in active site.[15] The coordination of the catalytic zinc in a representative molecular dynamics structure with the A ligand is shown, along with the position of glu202. Note that only the ZBG of the ligand A is shown. In the A trajectory, a water molecule forms a bridging hydrogen bond between glu202 carboxylate and an oxygen of the carboxylate of the ligand, the carboxylate of the ligand is positioned symmetrically above the zinc, and a water molecule also comes within 2.2 Å of the zinc. In the E_{glu} and F_{glu} trajectories, the glu202 carboxylate approaches within 2.2 Å of the catalytic zinc while the carboxylate of the ligand rotates. The E_{hoh} and F_{hoh} trajectories have geometries similar to the A trajectory.

3.2. A Trajectory. Throughout the A trajectory, the sixth ligand coordinating the zinc was a tightly bound water molecule. The MM-PBSA contributions to the binding energy for each ligand with each minimization protocol are given in Table 2. The resulting net binding energy and ranking are shown in Table 3. Note that as an additional error check, the total binding energy of the ligand only minimization protocol with the A ligand was performed with the A ligand starting structure taken both directly from the trajectory (different conformation for each snapshot) and from the initial crystal structure (same conformation for each snapshot). The protocols yielded negligible differences in the MM-PBSA contributions (∼0.4 kcal/mol, Table 3) and indicate that minimization is sufficient to sample the conformations of the A ligand. This further indicates that the sampling

Table 3. Binding Free Energy Prediction: A Trajectory[a]

ligand	K_i[15] ($\Delta\Delta G$[b]) nM (kcal/mol)	no min ΔG $R_{Zn}^{PB} = 1.4$ Å	ligand only ΔG (rank) $R_{Zn}^{PB} = 1.4$ Å	belly ΔG (rank) $R_{Zn}^{PB} = 1.4$ Å	belly ΔG (rank) $R_{Zn}^{PB} = 1.1$ Å	belly ΔG (rank) $R_{Zn}^{PB} = 1.6$ Å
A	2 (0)	19.8 ± 5.7	14.1 ± 6.4 (1) 13.7 ± 6.4[c]	9.4 ± 5.0 (3)	71.3 ± 12.3 (6)	−17.7 ± 3.4 (1)
B	18 (1.3)		24.4 ± 6.3 (5)	5.8 ± 3.3 (1)	52.2 ± 5.6 (1)	−15.6 ± 3.0 (2)
C	38 (1.8)		17.2 ± 5.6 (2)	12.8 ± 3.3 (4)	62.4 ± 4.5 (4)	−9.7 ± 2.8 (4)
D	68 (2.1)		19.2 ± 3.9 (4)	14.4 ± 3.2 (6)	60.7 ± 5.8 (3)	−8.4 ± 2.7 (5)
E	82 (2.2)		18.3 ± 5.3 (3)	14.0 ± 3.4 (5)	63.6 ± 4.4 (5)	−8.4 ± 2.7 (5)
F	110 (3.0)		27.3 ± 5.0 (6)	8.4 ± 3.0 (2)	55.3 ± 5.4 (2)	−12.9 ± 2.7 (3)

[a] All ΔG values are in kcal/mol, all values are reported ± the standard deviation over 49 snapshots, where the standard deviation is calculated assuming a 1.5 kcal/mol error in the entropic contribution. [b] Calculated from experimental K_i values relative to ligand A. [c] Calculated binding energy when a single conformation of A is substituted into each snapshot and energy minimized.

Table 4. Stability of MM-PBSA Calculations Using the A Trajectory, Belly Minimized[a]

snapshots	average	std dev	range
1−49	−12.1	4.8	22.8
1−10	−10.8	6.7	22.8
11−20	−14.8	4.3	15.3
21−30	−13.4	3.0	9.9
31−40	−10.2	3.5	9.7
41−49	−11.4	4.9	14.3

[a] All values in kcal/mol.

of the B, C, D, E, and F ligands will not be adversely affected by the substitution of a single conformation in each trajectory snapshot.

To assess the number of snapshots required to determine the binding energy, the A trajectory was broken down into series of 10 consecutive snapshots. The average, standard deviation, and range of the MM/PB/SA contributions to the binding energy was assessed and is reported for the A ligand in Table 4 (results are similar for the other ligands). This indicates that trajectories run for 100 ps (after 80 ps equilibration) yield a stable average for the binding energy.

3.3. Zinc Radius. The PB analysis of the cost of desolvation during complex formation is particularly sensitive to the radius used for the divalent zinc ion. We attempted these calculations using three radii: 1.10 Å from the Stote and Karplus parametrization,[8]1.40 Å as suggested for the surface area calculation in the msms program,[33] and 1.60 Å in an attempt to achieve more reasonable absolute binding energies (i.e., values < 0). While the absolute energies change dramatically with the particular value of the zinc radius, they cannot be used for calibration purposes. The total binding energy is formed from both the MM and PB estimates, and the MM estimate is known to be incorrect, since the binding geometry is incorrect. Therefore, we can only calibrate to the relative energy differences. In particular the binding energy difference between the A and the C, D, and E ligands is a definitive test. Because the A ligand is significantly larger than the C, D, and E ligands and because the morpholino group approaches within 6 Å of the structural zinc ion, the effect of the zinc desolvation parameters is particularly apparent for this ligand. The best results are obtained with a zinc radius of 1.40 Å, which has a range of ∼5.0 kcal/mol in binding energy between the A and C, D, and E ligands as compared to the experimental range of 2.1 kcal/mol. The Zn radii of 1.60 Å and 1.10 Å yield ranges of ∼9.3 kcal/mol ∼−10 kcal/mol, respectively (Table 3). $R_{Zn}^{PB} = 1.40$ Å is used for all other calculations (Table 3). Interestingly, using a Born model of solvation for Zn^{2+} and assuming a solvation free energy of −480 to −485 kcal/mol,[8] would result in a calculated radius of 1.35 to 1.38 Å, in accord with our final radius of 1.40 Å.

3.4. E and F Trajectories. To determine the dependence of the result on the particular query ligand and zinc coordination, the analysis was repeated using E and F (Figure 1) as the query ligands. Trajectories were generated with the sixth zinc coordinating ligand identified as either glu202 or water. Results for the final calculated binding free energies are shown in Tables 5 and 6.

3.5. Ligand Design. PROFEC analysis of the regions around the rigid ring structures of the A ligand indicated very few areas for modification. As has been shown in the literature, the biphenyl ring fits quite tightly into the P1′ pocket, leaving little room for substitution along the sides of the rings.[15] There is room for replacement of the fluorine atom on the tip of the biphenyl ring (position R_1 in Figure 1), and substitution with CH_2OH was proposed in an attempt to recruit some hydrogen bonding with the enzyme (A_1). However, the use of the hydroxyl group is often discouraged due to the cost of desolvation. Substitution of the phenyl group

Table 5. Binding Free Energy Prediction Using Trajectory E[a]

L	expt[b] $\Delta\Delta G$	E_{glu} no min ΔG	E_{glu} ligand only ΔG (rank)	E_{glu} belly ΔG (rank)	E_{hoh} no min ΔG	E_{hoh} ligand only ΔG (rank)	E_{hoh} belly ΔG (rank)
A	0		3.5 ± 19.2[c] (4)	−15.1 ± 4.7 (1)		14.0 ± 8.1 (1)	7.2 ± 7.4 (3)
B	1.3		20.9 ± 4.6 (5)	3.1 ± 4.4 (5)		26.2 ± 5.6 (5)	6.1 ± 3.4 (1)
C	1.8		0.8 ± 4.4 (2)	−3.9 ± 2.7 (4)		20.0 ± 4.7 (3)	14.0 ± 4.6 (5)
D	2.1		0.3 ± 4.1 (1)	−5.3 ± 2.3 (2)		19.4 ± 4.5 (2)	13.9 ± 3.6 (4)
E	2.2	3.8 ± 6.5	1.3 ± 4.6 (3)	−4.2 ± 3.4 (3)	25.8 ± 4.1	21.5 ± 4.3 (4)	14.7 ± 4.4 (6)
F	3.0		24.5 ± 3.9 (6)	9.0 ± 6.1 (6)		30.6 ± 5.6 (6)	8.9 ± 3.3 (2)

[a] All ΔG values are in kcal/mol, and all values are reported ± the standard deviation over 20 snapshots, where the standard deviation is calculated assuming a 1.5 kcal/mol error in the entropic contribution. L = ligand, $R_{Zn}^{PB} = 1.40$ Å. [b] Calculated from K_i values in ref 15, see Table 3. [c] Average and standard deviation over 19 snapshots.

Table 6. Binding Free Energy Prediction Using Trajectory F[a]

		F_{glu}			F_{hoh}		
L	expt[b] $\Delta\Delta G$	no min ΔG	ligand only ΔG (rank)	belly ΔG (rank)	no min ΔG	ligand only ΔG (rank)	belly ΔG (rank)
A	0		11.9 ± 9.5 (3)	−8.0 ± 3.8 (1)		45.3 ± 33.9 (3)	15.3 ± 7.6 (3)
B	1.3		7.2 ± 3.6 (1)	−0.1 ± 3.5[c] (2)		60.4 ± 80.3 (5)	6.0 ± 4.6 (1)
C	1.8		18.9 ± 4.4 (5)	7.8 ± 3.1 (5)		44.2 ± 10.9 (2)	17.0 ± 2.7 (4)
D	2.1		16.6 ± 4.7 (4)	5.3 ± 2.9[c] (4)		365 ± 904 (6)	17.4 ± 3.3[d] (5)
E	2.2		18.9 ± 4.5 (5)	9.4 ± 3.9 (6)		46.1 ± 11.3 (4)	17.8 ± 2.7 (6)
F	3.0	11.7 ± 3.4	6.6 ± 3.8 (2)	3.0 ± 2.7 (3)	20.3 ± 3.9	17.5 ± 9.0 (1)	6.5 ± 2.7 (2)

[a] All ΔG values are in kcal/mol. All values are reported ± the standard deviation over 10 snapshots, where the standard deviation is calculated assuming a 1.5 kcal/mol error in the entropic contribution. $R_{Zn}^{PB} = 1.40$ Å. [b] Calculated from K_i values in ref 15, see Table 3. [c] Eight snapshots. [d] Nine snapshots.

is not favored since it is largely solvent exposed and would not introduce much differentiation between binding to the enzyme and solvation. Finally positions around the morpholino group were examined. The carbonyl side is largely solvent exposed, while substitutions on the other side of the morpholino group result in steric clashes with the enzyme. This leaves investigation of the R_2 and R_3 positions as indicated in Figure 1. A hydroxyl substitution of R_2 was proposed (ligand A_2), as the electrostatic map indicated this would be a favorable placement, and there was the potential to recruit hydrogen bonding with the enzyme. A series of fluorine substitutions were proposed for both R_2 and R_3 since the PROFEC contour map indicated there was some room for substitution (ligands A_3 through A_5) and the fluorine atom is a small substitution which is often favorable. The result of the PROFEC analysis was a series of five ligands, and the resulting binding energies are given in Table 8.

4. Discussion

4.1. Ability To Predict Relative Ranking. The success of our methodology for ranking multiple ligands simultaneously can be assessed both in terms of the relative order of ranking as well as the range of free energies predicted among the ligands. The ligand only minimization technique for the A trajectory ranks the negatively charged ligands (A, C, D, and E) correctly within error. The neutral ligands (B and F) are also ranked correctly with respect to each other, although they appear to be disfavored with respect to the negatively charged ligands. This trend occurs because the interaction of glu202 with the NH_2^+ moiety of the ligand has not been properly optimized (the enzyme was frozen during the minimization). The range in free energies among all the ligands, and among the negatively charged ligands, is generally too large. Most of these problems are removed when the enzyme, as well as the ligand, is allowed to relax in the presence of the particular ligand (column 5, Table 3). Although the neutral and negatively charged ligands are again not ranked well with respect to each other, the range in free energies has decreased. In particular, it appears that the A ligand is slightly over favored while the range among the C, D, and E ligands is 1.6 kcal/mol (vs 0.4 kcal/mol experimentally) and well within the standard deviation. Similarly, the free energy difference in binding between B and F is predicted to be 2.6 kcal/mol vs 1.7 kcal/mol experimentally. Given that the error (standard deviation) in our methodology is ∼4 kcal/mol, these results are well within the error bars.

The limiting factors in this assessment include the charge balance and burying hydrogen bonds. Because a continuum approximation to desolvation costs is made, the specific interactions between a given ligand and water molecules (i.e., directional hydrogen bonds) are not completely accounted for. The overstabilization of B and F can be partially attributed to directional hydrogen bonds that would occur between the unique NH_2^+ moiety and water for the free ligand. In particular, because the NH_2^+ group is "overshadowed" by the carboxylate, the derived PB radii may not be adequate to describe the desolvation in the continuum approximation. A possible test of this hypothesis is to include two explicit waters (with their charges scaled 70% since they are no longer considered bulk water molecules) and calculate the resulting desolvation of these two tightly bound water molecules. A quick molecular dynamics simulation of the F ligand in a cap of water reveals that two water molecules are generally complexed to the NH_2^+ moiety. Isolation of these two water molecules, yielding a series of snapshots consisting of $F{\cdot}2H_2O$, and use of these snapshots in the MM-PBSA analysis indicates that these directional hydrogen bonds would contribute ∼8.5 kcal/mol of stability to the free ligand over the continuum approximation and increase the binding energy by a corresponding amount. This would place the B and F ligands within experimental error of the C, D, and E ligands. Despite these limitations, the neutral and charged ligands are tightly grouped, within methodological error, and correctly ranked within each set.

This methodology for ranking the relative binding energy of the ligands is quite robust. As shown in Table 4, the results are not overly sensitive to which subset of conformations are taken to form the estimate of the free energy.

Similar analyses can be made of the E_{glu} and F_{glu} trajectories. In this case, the range in predicted free energies is generally larger, although the order of ranking among the neutral and charged ligands is correct within error. The increased free energy range is in part due to the shielding of the catalytic zinc by glu202, which substantially changes the desolvation cost calculated for the isolated receptor and results in a greater overstabilization of the A ligand in particular. The neutral ligands are not as favored as in the A trajectory because the glu202 carboxylate group is primarily stabilizing the zinc ion as opposed to being completely available for interaction with the NH_2^+ moiety of the ligands. In comparison, the E_{hoh} and F_{hoh} trajectories reveal that when the sixth coordinating

ligand is water, the relative binding energies are similar to those derived with the A trajectory.

4.2. Effect of the Zinc Geometry. The varying coordination of the zinc ion continues to plague attempts to model zinc metalloproteins. However, in the approach undertaken, the energetic and conformational errors involved in incorrect zinc coordination seem to be canceled. In particular, the traditional approach, wherein individual trajectories are run for each ligand, can give rise to incorrect ranking. In this case, the predicted free energy of the A, E, and F ligands using the original A, E_{glu}, and F_{glu} trajectories, respectively, is 19.8, 3.8, and 11.7 kcal/mol. Using the A, E_{hoh}, and F_{hoh} trajectories, the energies are, respectively, 19.8, 25.8, and 20.3 kcal/mol. Since all explicit water molecules and counterions are removed in the MM-PBSA analysis, the effect of a water molecule as the sixth ligand is suppressed. Thus the results in these trajectories are favored over the E_{glu} and F_{glu} trajectories as a result of the particular zinc coordination that occurs. However, when the same "incorrect" coordination of zinc is used to rank all the trajectories, this error canceled out, allowing even the E_{glu} and F_{glu} trajectories to rank the neutral and charged ligands correctly within error.

Nonetheless, the presence of this error does imply that the absolute binding energies will not be correct with this methodology, or any other in which the zinc coordination is incorrect. Thus, the use of this methodology is limited to sets of ligands in which the binding affinity of at least one ligand is known and can be used to "calibrate" the scale of the predicted binding energies.

The cancellation of errors that occurs with this MM/PB/SA approach depends on the invariance of the zinc binding group among all ligands. This restriction to a given ZBG is not particularly severe. Often, considerations such as drug metabolism will limit the identity of the ZBG at the outset of the drug design process. In addition, given the known ranking of the ZBGs, the effect of changing the ZBG is approximately predictable.

Nonetheless, in future MM-PBSA studies on metal binding proteins, it may be worth exploring the use of models that fix the zinc coordination with bonded interactions,[11] direct a tetrahedral geometry[13] or constrain the structure with extra atom−atom constraints.[34] We have shown that our results are not sensitive to zinc coordination geometry, provided that it is common for all ligands, but there is clearly a need for more general and robust methods to describe metal coordination within classical molecular mechanics and dynamics.

4.3. Choice of Query Ligand. One of the most central considerations to the use of this methodology is potentially the choice of the query ligand. Particularly in the case of ligand size, we were concerned whether the use of a small ligand as the query (i.e., E or F) would disfavor the larger ligand (i.e., A).

Comparison of the A, E_{hoh}, and F_{hoh} trajectories reveal that the use of the smaller E and F ligands to generate the trajectory destabilized the larger ligands in the case of the ligand only minimization, particularly with the use of the smallest query ligand, F. However, these destabilizations were removed with the belly minimizations, and the three trajectories gave comparable results. The success shown here in using a variety of query ligands coupled with belly minimizations are still biased since the starting structure was the crystal structure generated with the larger A ligand. In the case where a smaller ligand has been used to determine the crystal structure, we are currently investigating whether straightforward minimization would be able to compensate as was seen here in the simulations performed here with the E and F ligands.

Table 7. Comparative Ranking of the A, E, and F Ligands with Various Protocols

ligand	expt[a] $\Delta\Delta G$	individual trajectories, no minimization $\Delta\Delta G$	A trajectory, belly $\Delta\Delta G$
A	0	0.0 ± 5.7	0.0 ± 5.0
E	2.2	6.0 ± 4.1	4.6 ± 3.4
F	3.0	0.5 ± 3.9	-1.0 ± 3.0

[a] Calculated from K_i values in ref 15.

A final comment can be made on the use of a single trajectory versus multiple trajectories to generate the binding free energies. Table 7 presents the relative binding energies for the A, E, and F ligands as predicted from the individual trajectories (with no minimization) as well as the single trajectory using the A ligand as the query (belly minimization). It is apparent that the single trajectory results are entirely consistent with the multiple trajectory results (when the sixth ligand is water) and that any errors occurring are independent of sampling issues.

4.4. Ligand Design. Ligands A_1−A_5 were studied on the basis of PROFEC analysis. As expected, the two hydroxyl substitutions were not favorable (Table 8), primarily because the increase in electrostatic interactions was not sufficient to offset the cost of desolvation. The flouro substitutions (A_3−A_5) did not appear to penalize binding, and the belly minimization protocol indicated that all three ligands bound as well as A by itself, within the standard deviation of the method. Further analysis of individual trajectories was therefore undertaken, to allow for any rearrangement due to the electrostatic changes introduced by the fluorine atoms. Analysis of the individual trajectories indicated that relative to A, ligand A_3 would be better by ∼4 kcal/mol, while A_4 is slightly disfavorable and A_5 is slightly favorable. The major difference between all three ligands and the original A ligand is predominately in the desolvation term. Thus for the A ligand $\langle E_{elec} \rangle = -193.3$ kcal/mol, while A_3−A_5 had $\langle E_{elec} \rangle$ ranging from −188.0 to −190.4 kcal/mol. Similarly, the $\langle E_{vdw} \rangle = -64.7$ kcal/mol for A and ranged from −66.1 to −66.7 kcal/mol for A_3−A_5. However, the electrostatic desolvation term was 264.9 kcal/mol for A while A_3 has an electrostatic desolvation cost of 258.4 kcal/mol, and A_4 and A_5 has values of 262.3 and 262.8 kcal/mol, respectively. Thus, A_3 is favored over A_4 and A_5 as a result of decreasing the cost of desolvation. The fluorine substitutions in general result in a decrease of electrostatic contributions and an increase in van der Waals contributions to the binding energy.

In summary, the use of PROFEC analysis allowed us to identify five possible substitutions to improve the binding of the A ligand. The minimization protocol was then used to rule out two of these possibilities. Finally, individual trajectories were run for the final three ligands to allow for the optimal rearrangement of the morpholino pocket in response to the fluorine substitu-

Table 8. Ligand Optimization[a]

ligand	R_1	R_2	R_3	$-T\Delta S$ (kcal/mol)	MM-PBSA BELLY[b] (kcal/mol)	MM-PBSA TRAJ[c] (kcal/mol)	ΔG_{bind} BELLY (kcal/mol)	ΔG_{bind} TRAJ (kcal/mol)
A	F	H	H	21.4	-12.0 ± 4.8	-1.6 ± 5.5[d] (49)	9.4 ± 5.0	19.8 ± 5.7
A_1	CH_2OH	H	H	23.2	-10.9 ± 5.2		12.3	
A_2	F	H	OH	21.6	-9.5 ± 4.9		12.1	
A_3	F	F	H	21.7	-11.9 ± 4.9	$-5.9 + 5.8$[d] (20)	9.8	15.8 ± 6.0
A_4	F	H	F	21.9	-12.4 ± 4.8	-1.7 ± 7.5 (30)	9.5	20.2 ± 7.6
A_5	F	F	F	21.3	-13.0 ± 4.7	-2.3 ± 7.7 (19)	8.3	19.0 ± 7.8

[a] See Figure 1 for definitions of R_1, R_2, R_3. $R_{Zn}^{PB} = 1.40$ Å. [b] Averaged over 49 snapshots. [c] Average number of snapshots used indicated in brackets. [d] The average MM-PBSA binding energies of A and A3 are found to be significantly different ($p < 0.05$) using a *t*-test assuming unequal variance including the Bonferroni correction for multiple comparisons.

tions. As a result of this process, it appears that the substitution of fluorine in the R_2 position (Figure 1) ought to result in a more tightly binding ligand.

4.5. Advantages, Limitations, and Future Directions. The advantages to this methodology are evident in both the CPU savings and in the cancellation of errors. Thus while other approaches (i.e., free energy perturbation[22]) compare ligands in a pairwise manner for *small* changes in ligand structure, this approach has the potential to rank tens of ligands. In addition, the ranking is performed on the basis of a single trajectory, which is a considerable CPU advantage over methods such as Aqvist's linear interaction energy approach.[23] Finally, errors in zinc coordination are canceled within a trajectory as long as the same ZBG is included in all the ligands.

However, these advantages also give rise to limitations. Thus, only one kind of zinc binding group can be considered at a given time. Furthermore, despite the cancellation of errors, it is still necessary to beware of distortions in the enzyme structure due to the zinc ion (i.e., motion of glu202).

One of the problems that has plagued predictions of free energy ranking in the past is ligand diversity. Traditionally, only very small changes in ligand structure could be reliably calculated. In the future we intend to examine this approach to determine what degree of diversity in the ligand (or receptor in the case of selectivity determinations) structure is possible while still maintaining accurate relative binding energy predictions. Issues such as burying differing numbers of hydrogen bonds or partially burying charged functional groups within the enzyme, for example, could be challenging in the context of the continuum approximation to desolvation.

What is the role of this methodology in the structure based design process? In our opinion, the first step when one has a target structure is to use DOCKing on real[34] or virtual databases[35] to broadly span chemical space. Then, when one has reduced the possible structures to a more limited number of templates (10−100), more accurate methods such as molecular dynamics based approaches can be used.

Currently, there are a number of such molecular dynamics based methods available, such as Aqvist's LIE[23] and the MM-PBSA approach, both of which postprocess molecular dynamics trajectories. There are also two promising methods that calculate binding free energy explicitly and do so more efficiently than pairwise free energy calculations—CMC/MD[36] and lambda dynamics[37]—since both of these approaches allow one to consider many molecules at once.

The virtues of the MM-PBSA approach employed here compared to Aqvist's LIE approach is that it requires fewer empirical parameters and, as shown here, can be based on a single MD trajectory on one prototype molecule (query ligand) followed by limited minimization on others. Then PROFEC[24] can be done on the already generated trajectories to search for further improvement. Nonetheless, there still are challenges/difficulties with the MM-PBSA approach to calculating the ΔS contribution to binding. Elsewhere, we have compared the MM-PBSA approach with the LIE approach on biotin analogues interacting with avidin and steptavidin.[38]

5. Conclusion

With the current (Stote et al.[8]) parametrization of the zinc ion, we are unable to reproduce the tetrahedral coordination seen in the crystal structure. Nevertheless we have demonstrated that it is possible to reliably predict the relative ranking of ligands that bind zinc with the same zinc binding group. Six nanomolar inhibitors of MMP-3 were ranked on the basis of a single trajectory run with a representative ("query") ligand. The best results were obtained with the use of the largest, best binding ligand for which the crystal structure was available; however, the results were relatively invariant to the choice of query ligand when a belly minimization of the ligand and enzyme was included in the methodology. While the traditional challenges of the zinc ion are overcome in this approach, the best results are still obtained when the sixth coordinating ligand for zinc is a water molecule (which is removed during the MM-PBSA analysis) as opposed to a neighboring residue within the enzyme (i.e., glu202). Ranking was reliably achieved with ∼4 kcal/mol standard deviations in the free energy results using a belly minimization.

These results indicate that we can now confidently predict, in a feasible time frame, the relative binding energies of ligands with the same ZBG for the matrix metalloproteinases. As a demonstration of the applicability of this methodology, we have used current in-house tools to suggest improvements to the best binding ligand used in the test set. In particular, a fluorine substitution at the R_2 position should give rise to a ligand with comparable or better binding energy than the A ligand. With these tools in place, it is now possible to use molecular modeling to design novel inhibitors of the MMPs.

Acknowledgment. P.A.K. acknowledges Grant GM-29072. O.D. acknowledges an NSERC PDF for salary

support. This research was partially supported through computing resources provided by the National Partnership for Advanced Computational Infrastructure at the San Diego Supercomputer Center (NSF Cooperative agreement ACI-9619020) and by the National Computational Science Alliance utilizing the NCSA SGI/CRAY Origin 2000.

Supporting Information Available: Partial charges and geometrical parameters for all ligands are available free of charge via the Internet at http://pubs.acs.org.

References

(1) Yu, A. E.; Hewitt, R. E.; Connor, E. W.; Stetler-Stevenson, W. G. Matrix Metalloproteinases: Novel Targets for Directed Cancer Therapy. *Drugs Aging* **1997**, *11* (3), 229−244.
(2) Koivunen, E.; Arap, W.; Valtanen, H.; et al. Tumor Targeting with a Selective Gelatinase Inhibitor. *Nature Biotech.* **1999**, *17*, 768−774.
(3) Gottschall, P. E.; Deb, S. Regulation of Matrix Metalloproteinase Expression in Astrocytes, Microglia and Neurons. *NeuroImmunoModulation* **1996**, *3*, 69−75.
(4) Becket, R. P.; Davidson, A. H.; Drummond, A. H.; Huxley, P.; Whittaker, M. Recent Advances in Matrix Metalloproteinase Inhibitor Research. *Drug Discuss. Today* **1996**, *1* (1), 16−26.
(5) Jacobsen, E. J.; Mitchell, M. A.; Hendges, S. K.; et al. Synthesis of a Series of Stromelysin-Selective Thiadazole Urea Matrix Metalloproteinase Inhibitors. *J. Med. Chem.* **1999**, *42*, 1525−1536.
(6) Alberts, I. L.; Nadassy, K.; Wodak, S. J. Analysis of Zinc Binding Sites in Protein Crystal Structures. *Protein Sci.* **1998**, *7* (8), 1700−1716.
(7) Vedani, A.; Huhta, D. W. A New Force Field for Modeling Metalloproteins. *J. Am. Chem. Soc.* **1990**, *112*, 4759−4767.
(8) Stote, R. H.; Karplus, M. Zinc Binding in Proteins and Solution: A Simple but Accurate Nonbonded Representation. *Proteins: Struct., Funct. and Genet.* **1995**, *23*, 12−31.
(9) Clementi, E.; Corongiu, G.; Jonsson, B.; Romano, S. *J. Chem. Phys.* **1980**, *72*, 260−263.
(10) Yongyai,Y. P.; Kokpol, S.; Rode, B. M. *Chem. Phys.* **1991**, *156*, 403−412.
(11) Hoops, S. C.; Anderson, K. W.; Merz, K. M., Jr. Force Field Design for Metalloproteins. *J. Am. Chem. Soc.* **1991**, *113*, 8262−8270.
(12) Toba, S.; Damodaran, K. V.; Merz, K. M., Jr. Binding Preferences of Hydroxamate Inhibitors of the Matrix Metalloproteinase Human Fibroblast Collagenase. *J. Med. Chem.* **1999**, *42*, 1225−1234.
(13) Roe, R. R.; Pang, Y. P. Zinc's Exclusive Tetrahedral Coordination Governed by Its Electronic Structure. *J. Mol. Model.* **1999**, *5*, 134−140.
(14) Pang, Y. P. Novel Zinc Protein Molecular Dynamics Simulations: Steps Toward Antiangiogenesis for Cancer Treatment. *J. Mol. Model.* **1999**, *5*, 196−202.
(15) Esser, C. K.; Bugianesi, R. L.; Caldwell, C. G.; et al. Inhibition of Stromelysin-1 (MMP-3) by P_1'-Biphenylylethyl Carboxyalkyl Dipeptides. *J. Med. Chem.* **1997**, *40*, 1026−1040.
(16) Srinivasan, J.;. Miller, J.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of RNA Hairpin Loops and Helices. *J. Biol. Struct. Dyn.* **1998**, *16* (3), 671−682.
(17) Cheatham 3rd, T. E.; Srinivasan, J.; Case, D. A.; Kollman, P. A. Molecular Dynamics and Continuum Solvent Studies of the Stability of polyG-polyC and polyA-polyT DNA Duplexes in Solution. *J. Biol. Struct. Dyn.* **1998**, *16*, 265−280.
(18) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices. *J. Am. Chem. Soc.* **1998**, *120*, (37), 9401−9409.
(19) Reyes, C. M.; Kollman, P. A. Investigating the Binding Specificity of U1A-RNA by Computational Mutagenesis. *J. Mol. Biol.* **2000**, *295* (1), 1−6.
(20) Massova, I.; Kollman, P. A. Computational Alanine Scanning to Probe Protein−Protein Interactions: A Novel Approach to Evaluate Binding Free Energies. *J. Am. Chem. Soc.* **1999**, *121*, 8133−43.
(21) Chong, L. T.; Duan, Y.; Wang, W.; Massova, I.; Kollman, P. A. Molecular Dynamics and Free-Energy Calculations Applied to Affinity Maturation in Antibody 48G7. *Proc. Natl. Acad. Sci.* **1999**, *96* (25), 14330−14335.
(22) Kollman, P. A. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(23) Hansson, T.; Marelius, J.; Aqvist, J. Ligand Binding Affinity Prediction by Linear Interaction Energy Methods. *J. Comput. Aided Mol. Des.* **1998**, *12*, 27−35.
(24) Radmer, R. J.; Kollman, P. A. The Application of Three Approximate Free Energy Calculations Methods to Structure Based Ligand Design: Trypsin and its Complex with Inhibitors. *J. Comput. Aided Mol. Des.* **1998**, *12*, 215−227.
(25) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.,3rd; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. A. AMBER, a Package of Computer Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Comm.* **1995**, *91*, 1−41.
(26) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr., K.M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(27) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method using Charge Restraints for Atomic Charges: the RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−80.
(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*; Gaussian, Inc.: Pittsburgh, PA, 1995.
(29) Aqvist, J. Ion Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.* **1990**, *94*, 8021−8024.
(30) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(31) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the Electrostatic Potential of Moelcules in Solution − Method and Error Assessment. *J. Comput. Chem.* **1988**, *9*, 327−335.
(32) Sitkoff, D.; Sharp, K.; Honig, B. Accurate Calculation of Hydration Free Energies using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(33) Sanner, M. F.; Olson, A. F.; Spehner, J. C. Reduced Surface − an Efficient Way to Compute Molecular Surfaces. *Biopolymers* **1996**, *38*, 305−320.
(34) Makino, S.; Kuntz, I. D. Automated Flexible Docking and Its Application for Database Search. *J. Comput. Chem.* **1997**, *18*, 1812−1825.
(35) Makino, S.; Ewing, T. J. A.; Kuntz, I. D. DREAMtt: Flexible Docking Program for Virtual Combinational Libraries. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 513.
(36) Eriksson, M.; Pitera, J. Prediction of the Bhinding Free Energy of New TIBO-like HIV-1 Reverse transcriptase Inhibitors Using a Combination of PROFEC, PBISA, CMD/MD and Free Energy Calculations. *J. Med. Chem.* **1999**, *42*, 868−881.
(37) Guo, Z. Y.; Brooks, C. L. Rapid Screening of Binding Affinities: Application of the Lambda-Dynamics Method to a Trypsin Inhibitor System. *J. Am. Chem. Soc.* **1998**, *120*, 1920−1921.
(38) Kuhn, B.; Kollman, P. A. Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models (MM/PBSA). *J. Med. Chem.* **2000**, *43*, 3786−3791.

JM000040D