Published in final edited form as:

J Mol Recognit. 2010 ; 23(2): 173–182. doi:10.1002/jmr.989.

Including Receptor Flexibility and Induced Fit Effects into the Design of MMP-2 Inhibitors

Jacob D. Durrant^{1,2}, César Augusto F. de Oliveira^{2,3,*}, and J. Andrew McCammon^{2,3}

¹ Biomedical Sciences Program, University of California San Diego, La Jolla, California 92093-0365

² Department of Chemistry & Biochemistry, Department of Pharmacology, and NSF Center for Theoretical Biological Physics, University of California San Diego, La Jolla, California 92093-0365

³ Howard Hughes Medical Institute, University of California San Diego, La Jolla, CA 92093-0365

Abstract

Matrix metalloproteinases (MMPs) comprise a class of flexible proteins required for normal tissue remodeling. Overexpression of MMPs is associated with a wide range of pathophysiological processes, including vascular disease, multiple sclerosis, Alzheimer's disease, and cancer. Nearly all MMP inhibitors have failed in clinical trials, in part due to lack of specificity. Due to the highly dynamic molecular motions of the MMP-2 binding pockets, the rational drug design of MMP inhibitors has been very challenging. To address these challenges, in the current study we combine computer docking with molecular dynamics (MD) simulations in order to incorporate receptor-flexibility and induced-fit effects into the drug-design process. Our strategy identifies molecular fragments predicted to target multiple MMP-2 binding pockets.

Introduction

Matrix metalloproteinases (MMPs) comprise a class of flexible proteins (Moy, Pisano et al. 1997; Moy, Chanda et al. 1998; Lovejoy, Welch et al. 1999; Moy, Chanda et al. 2000; Zhang, Gonnella et al. 2000) for which computational drug design has been very challenging (Matter and Schudok 2004). There are over 20 MMP enzymes (Chang and Werb 2001) capable of degrading nearly all the components of the extracellular matrix, as well as many other extracellular proteins (Ray and Stetler-Stevenson 1995; Chambers and Matrisian 1997; Kahari and Saarialho-Kere 1999; Kleiner and Stetler-Stevenson 1999; Coussens, Fingleton et al. 2002). MMPs are required for normal tissue remodeling and are involved in many physiological processes, including embryonic growth, signal regulation, inflammation, angiogenesis (Chang and Werb 2001; Hu, Van den Steen et al. 2007), uterine involution, bone resorption, wound healing (Ray and Stetler-Stevenson 1995; Chambers and Matrisian 1997; Kahari and Saarialho-Kere 1999; Kleiner and Stetler-Stevenson 1999), CNS regeneration (Yong 2005), and ovulation (Whittaker, Floyd et al. 1999).

Overexpression of MMPs has been associated with a wide range of pathophysiological processes as well, including vascular disease (Hu, Van den Steen et al. 2007), asthma (Mautino, Oliver et al. 1997; Cataldo, Munaut et al. 2001; Prikk, Maisi et al. 2002), arthritis, multiple sclerosis, Alzheimer's disease (Coussens and Werb 1996), and cancer (Denis and Verweij 1997; Wojtowicz-Praga, Dickson et al. 1997; Brown 1999), which requires

*Correspondence to: Department of Chemistry & Biochemistry, University of California San Diego, 9500 Gilman Drive, Mail Code 0365, La Jolla, CA 92093-0365, cesar@mccammon.ucsd.edu, 858-534-3038 (Office), 858-534-4974 (Fax).

degradation of the extracellular matrix for tumor progression and metastasis. The development of MMP inhibitors for the prevention of cancer metastasis has been a particularly active field of study (Hidalgo and Eckhardt 2001). MMP-2 (gelatinase A) and other gelatinases have received special attention because they are most consistently detected in malignant tissues and are associated with tumor aggressiveness, metastatic potential, and poor prognosis (Hidalgo and Eckhardt 2001; Mook, Frederiks et al. 2004). Nearly all MMP inhibitors have failed in clinical trials, in part due to lack of specificity (Pirard 2007).

Any rational design of MMP-2 inhibitors must take into account the general shape and characteristics of the MMP-2 active site (Pirard 2007). MMP inhibitors are typically comprised of a zinc-binding group (ZBG) with attached moieties that interact with surrounding binding pockets. A number of zinc-chelating groups have been proposed in the literature, including hydroxamates (Whittaker, Floyd et al. 1999), thiols (Skiles, Gonnella et al. 2004), pyrone derivatives (Puerta and Cohen 2003; Puerta, Lewis et al. 2004; Puerta, Griffin et al. 2006), carbamoylphosphonates (Farkas, Katz et al. 2004; Reich, Katz et al. 2005), hydroxyureas (Michaelides, Dellaria et al. 2001; Campestre, Agamennone et al. 2006), hydrazines (Auge, Hornebeck et al. 2003), β -lactams and squaric acids (Cainelli, Galletti et al. 2003; Onaran, Comeau et al. 2005), and nitrogenous ligands (dipicolylamine) (Lim, Schuster et al. 2005; Jacobsen, Lewis et al. 2006). Among these chelators, hydroxamic acid is by far the most frequently used (Whittaker, Floyd et al. 1999; Skiles, Gonnella et al. 2004).

Cohen et al. recently used a bioinorganic approach to discover new classes of zinc-chelating compounds that show improved potency and novel selectivity relative to similar hydroxamate-based ZBGs (Puerta, Lewis et al. 2004; Puerta, Mongan et al. 2005; Jacobsen, Lewis et al. 2006). These ZBGs were crystallized bound to a hydrotris(3,5-phenylmethylpyrazolyl)borate bioinorganic model complex whose zinc-coordinating nitrogen atoms are geometrically positioned in a way that is analogous to the three zinc-coordinating histidine nitrogen atoms of the MMPs (Puerta and Cohen 2003), allowing the ZBG to be modeled within any MMP-2 active site *in silico* by aligning analogous model and protein atoms.

One identified ZBG, 3-hydroxy-2-methyl-4-pyrone (maltol), is amenable to the addition of moieties that simultaneously interact with multiple MMP binding pockets. This is especially true when maltol is modified to 3-hydroxy-*N*-methyl-4-oxo-4*H*-pyran-2-carboxamide. Molecular modeling suggests that two of the hydrogen atoms on the terminal methyl group of this modified maltol point towards the S1' and S2' pockets, respectively (Figure 1). S1' is a cavernous, flexible, hydrophobic pocket that is important for substrate selectivity (Pirard 2007). The S2' pocket has the greatest geometric variability of all sub-pockets, as judged by force-field interaction energies used to compare the structures of 24 MMPs (nine X-ray structures and 15 homology models). As the shape of the S2' pocket is not well conserved across the many different MMPs, targeting this pocket may lead to better therapeutic specificity (Lukacova, Zhang et al. 2004).

Due to the high flexibility of the MMP binding pockets (Moy, Pisano et al. 1997; Moy, Chanda et al. 1998; Lovejoy, Welch et al. 1999; Moy, Chanda et al. 2000; Zhang, Gonnella et al. 2000), predicting the binding affinity of MMP inhibitors *in silico* is particularly difficult. Although structural information extracted from the several available X-ray structures can be used, in principle, to design and improve the selectivity of a given inhibitor towards a specific MMP, the high flexibility of the MMP binding pockets drastically limits the accurate modeling and prediction of inhibitor-protein complexes (Yuan, Marshall et al. 1999; Cuniasse, Devel et al. 2005).

In order to assist future drug design, our group recently studied the flexibility of the *apo* MMP-2 binding pockets *via* a molecular dynamics (MD) simulation (de Oliveira, Zissen et al. 2007). The MD simulation revealed that two states, characterized as open and closed, are preferentially sampled. By analyzing the population of those states along the trajectory, we observed that while the closed state is heavily sampled, the open state, in which the S1' pocket adopts a tunnel-like shape, appears only rarely. The S1' loop, which has the highest sequence variability among different MMPs, is among the most mobile segments of MMP tertiary structure. Loops S2' and S3', which encompass solvent-exposed pockets S2' and S3', are also highly flexible (Figure 1).

Previously, our group presented a novel virtual-screening approach called the “relaxed-complex” scheme that accounts for receptor flexibility (Lin, Perryman et al. 2002). In this scheme, protein flexibility is taken into account by docking candidate inhibitors into an ensemble of protein configurations extracted from MD simulations of either the *apo* or *holo* receptor (Lin, Perryman et al. 2003; Amaro, Baron et al. 2008). This approach has been applied to several proteins including HIV-1 integrase (Schames, Henchman et al. 2004; Adcock and McCammon 2006), *Trypanosoma brucei* RNA editing ligase 1 (Amaro, Schnaufer et al. 2008), and influenza neuraminidase (Cheng, Amaro et al. 2008).

Given the dynamics of the MMP-2 binding pockets observed previously, a relaxed-complex technique was employed in the current study. Protein configurations were extracted from the *apo* MMP-2 simulation and the ZBG was positioned into each frame using *in silico* modeling based on the Cohen bioinorganic model (Puerta, Lewis et al. 2004; Puerta, Mongan et al. 2005; Jacobsen, Lewis et al. 2006). LUDI, a fragment-based *de novo* drug-design algorithm (Bohm 1994), was then used to dock fragments into the MMP-2 active site such that the fragments could be easily merged with the ZBG while maintaining interactions with the S1', S2', and S3' pockets across multiple protein configurations extracted from the *apo* MD simulation.

Next, we extended the relaxed-complex scheme by performing secondary MD simulations of the *holo* receptor (receptor in complex with each of the thirteen best LUDI-predicted ligands) in order to account for induced-fit effects. These thirteen compounds were then re-ranked according to their ensemble-average predicted binding energy. Re-ranking by the ensemble-average score, which accounts in part for the open-closed dynamics of the active site, significantly reordered the compounds, suggesting that induced-fit effects can have a substantial impact on the computational identification of potential enzyme inhibitors.

Computational Details

Initial molecular dynamics simulations were performed on a 2.80 Å resolution MMP-2 crystal structure (Dhanaraj, Williams et al. 1999) (PDBID: 1QIB). Crystal waters and inhibitors were removed from the structure. Basic residues, i.e. Arg and Lys, were protonated, and acidic residues, i.e. Asp and Glu, were deprotonated. Except for one histidine residue located in the structural domain (His 179), which was protonated in the epsilon position, all histidine residues were protonated in the delta position. To maintain the correct orientation of all zinc-chelating histidine residues and to prevent the escape of zinc atoms to the solvent, harmonic restraints were applied on the distances and bending angles defined by the zinc and the nitrogen atoms of the appropriate histidine residues. The partial charges of the histidine imidazole rings and zinc atoms were calculated using the RESP program. The molecular electrostatic potential was calculated at the HF/6-31G* level.

Following the initial system preparation, we applied 500 steps of steepest descent followed by 1000 steps of conjugate gradient minimization. The relaxed structure was then solvated in a cubic box of pre-equilibrated TIP3P water molecules (Jorgensen, Chandrasekhar et al.

1983), which extended 10 Å further than the protein atoms themselves. To neutralize the system, sodium counterions (Na^+) were added to balance the charge of the MMP-2 protein. To relax the waters, the system was then again minimized for 500 of steepest descent followed by 2000 steps of conjugate gradient minimization using constant volume periodic boundaries, with the protein and counterion atoms fixed. To bring the system to the correct density, we carried out an additional 50 ps MD simulation in which only water molecules were allowed to move and *NPT ensemble* ($T = 298\text{K}$, $P = 1\text{ bar}$) was applied. To relax the protein in the solvent environment, all coordinates were then optimized by employing 500 steps of steepest descent followed by 1000 steps of conjugate gradient minimization. We then performed a 500 ps MD simulation to heat the system from 0K to 300K using the *NVT ensemble* ($T = 298\text{K}$). All subsequent simulations were carried out using the *NVT ensemble* ($T = 298\text{K}$) and, except for the internal restraints mentioned before and SHAKE constraints on bonds to hydrogen atoms (Ryckaert, Ciccotti et al. 1977), the coordinates of all atoms were allowed to move freely. A final MD simulation of 30 ns was performed. The first 15 ns were used to equilibrate the system, and the last 15 ns were used as the productive run.

Three thousand and one frames were next extracted at regularly spaced intervals from the 15 ns productive portion of the MD simulation. For each frame, the complex formed between the ZBG and the receptor was built by aligning each frame with the crystal-structure coordinates of maltol (3-hydroxy-2-methyl-4-pyrone) bound to $[(\text{TpPh,Me})\text{ZnOH}]$ ($\text{TpPh,Me} = \text{hydrotris}(3,5\text{-phenylmethylpyrazolyl})\text{borate}$), a complex that models the zinc(II) ion coordination within the MMP active site (Puerta and Cohen 2002; Puerta, Schames et al. 2003). Maltol was positioned within the protein active site by minimizing the RMSD distance between the appropriate zinc and the coordinating nitrogen atoms of the protein active site and those of the bioinorganic model (Figures 2A and 2B).

LUDI was then used to dock potential interacting fragments into the many receptor configurations extracted from the *apo* MD simulation. The terminal methyl hydrogen atoms of the ZBG were selected as the link sites (Figure 1). Initial docking used the default Discovery-Studio LUDI link fragment library (Accelrys) with the following parameters: maximum alignment angle 20° ; maximum alignment RMSD 0.6 Å; search radius 11 Å; rotate bonds two at a time; preselect 4.0; minimum separation 3.0; lipophilic density 40; polar density 40; minimum surface 0; link weight 1.0; lipophilic weight 1.0; H-bond weight 1.0; aliphatic aromatic off; reject bifurcated off; no unpaired polar off; electrostatic check off; minimum score 0; maximum fits 8000; maximum hits all; maximum unfilled cavity 0; energy estimate 1 scoring function (Bohm 1994); and best fit (Puerta, Mongan et al. 2005).

Following the LUDI-guided identification of interacting fragments, the fragments were combined with the ZBG and additional MD simulations of 5 ns were carried out on the resulting protein-ligand complexes comprised of the best LUDI-generated ligand candidates and the MMP-2 receptor. An MD protocol similar to that described above was used. Ligand force-field parameters were determined using the antechamber program inside the AMBER package. Ligand partial charges were calculated using the AM1-BCC charge scheme (Jakalian, Bush et al. 2000). In order to guarantee that the zinc-chelating group of each candidate ligand remain bound to the zinc cation of the active site, harmonic restraints were applied to the distances between the two ligand chelating oxygen atoms and the catalytic zinc cation.

For all MD simulations, the equations of motion were integrated every 2 fs using the Verlet Leapfrog algorithm (Hockney 1968). During MD, temperature and pressure were controlled *via* a weak coupling to external bath with a coupling constant of 0.5 and 1 ps, respectively (Berendsen, Postma et al. 1984). The center-of-mass motion was removed at regular intervals of 5 ps. The PME summation method was used to treat the long-range electrostatic

interactions in the minimization and simulation steps of the solvated systems (Essmann, Perera et al. 1995). The short-range nonbonded interactions were truncated using an 8 Å cutoff and the nonbonded pair list was updated every 10 steps. All calculations were performed in the AMBER package (Case, Darden et al. 2006) with the ff99 force field (Wang, Cieplak et al. 2000).

In order to estimate the ensemble-average binding energy of each protein-ligand complex, seventy frames were extracted at regular intervals from the last 3.5 ns of each 5 ns *holo* simulation. The AutoDock 4.0 scoring function was used to estimate the ligand-protein binding energy of each frame. AutoDockTools (ADT) version 1.5.2 (Sanner 2005) was used to remove non-polar receptor and ligand hydrogen atoms and to assign Gasteiger charges (Gasteiger and Marsili 1980) to all the atoms of each frame. The MMP-2 zinc cations were manually assigned charges of +2.0 *e*. AutoGrid 4.0 was used to generate affinity grids centered on the active site. Each grid enclosed a volume of 48.00 Å × 48.00 Å × 48.00 Å with 0.375 Å spacing. For each protein-ligand system, all corresponding AutoDock-predicted binding energies were averaged to give the ensemble-average predicted binding energy (Score^{*holo*}).

For all controls, a procedure similar to the above was used, except that the ZBG 3-hydroxy-4-oxo-4*H*-pyran-2-carboxamide was used instead of 3-hydroxy-*N*-methyl-4-oxo-4*H*-pyran-2-carboxamide, and an extra methyl group was placed on the docked fragments to compensate.

Results and Discussions

Accounting for protein flexibility

The methodology presented here identifies potential inhibitors in a way that accounts for protein flexibility. The approach presented in this work, illustrated in Figure 2, aims to identify ligand candidates that not only interact favorably with conformations typical of the *apo* form of the protein, but also with potential novel configurations induced by ligand binding.

Based on the ideas behind the relaxed-complex scheme (Lin, Perryman et al. 2002; Lin, Perryman et al. 2003; Amaro, Baron et al. 2008), we first performed a molecular dynamics simulation of the *apo* protein in order to account for protein flexibility. Three thousand and one regularly spaced frames sampling both open and closed MMP-2 conformations were extracted from the last 15 ns of the 30 ns simulation, and the zinc-binding group (ZBG, 3-hydroxy-*N*-methyl-4-oxo-4*H*-pyran-2-carboxamide) was manually placed into each protein configuration *via* structural alignment of the protein structure with a crystal-structure bioinorganic model of the MMP-2 active site (Puerta, Lewis et al. 2004; Puerta, Mongan et al. 2005; Jacobsen, Lewis et al. 2006) (Figures 2A and 2B).

Next, we selected two strategically positioned ZBG hydrogen atoms as fragment linking points (Figure 1, inset picture). The first hydrogen atom is generally oriented toward the S1' pocket and the second is located in the vicinity of the S2' and S3' pockets. By targeting different binding pockets, we aim to identify fragments that simultaneously interact favorably and specifically with the protein. This approach can be easily extended to enable future combinatorial drug design that achieves high affinity by targeting two or more pockets concurrently.

LUDI docking calculations were then performed on the 3001 snapshots of the *apo* MD trajectory using both linking hydrogen atoms (Figure 2C). Because the best candidates obtained from LUDI docking calculations may change according to the conformation of the

receptor, two strategies were applied to select the “best” interacting fragments. The best fragments were considered to be those that I) were frequently selected as binders across multiple protein configurations regardless of LUDI score or II) had the highest ensemble-wide maximum LUDI score regardless of frequency (Figure 2D). Candidates were considered “frequent binders” if they were selected and docked into at least 60% of the protein configurations, regardless of score. Candidates were considered “strong binders” if their ensemble-maximum LUDI scores fell within 100 units of the highest score of any receptor/fragment combination. Finally, we discarded from our list any fragments that failed to bind in the targeted pockets or that contained long linear aliphatic chains, as compounds with long linear aliphatic chains contain many rotatable bonds and, consequently, large entropic penalties upon binding.

Including Induced-Fit Effects: Rescoring Candidate Ligands

In order to account for the induced-fit changes in protein flexibility that occur upon ligand binding, we performed additional MD simulations of each of the LUDI-generated ligands selected previously in complex with the MMP-2 protein. To model each ligand, we iterated through both the frequent and strong binding fragments and in each case identified the fragment and receptor conformations associated with the ensemble maximum LUDI score. Each candidate fragment, posed in its optimal configuration, was subsequently positioned within the corresponding optimal receptor conformation from the *apo* simulation and merged with the corresponding ZBG.

A 5 ns MD simulation of each protein-ligand system was then performed to simulate induced-fit effects (Figure 2E). For each system, 100 regularly spaced frames were extracted from the MD trajectory. The AutoDock 4.0 scoring function was used to estimate the ligand energy of binding associated with each frame; visualization of the data revealed that the score converged sufficiently after the first 1.5 ns of each simulation. To predict the binding affinity of each ligand, the AutoDock scores associated with the seventy frames of the last 3.5 ns of each of MD simulations were averaged, yielding an ensemble-average predicted binding energy for each ligand (Figure 2F). Candidate ligands were then re-ranked by this ensemble average. Comparing the ensemble-average AutoDock-predicted binding affinity of each system ($\text{Score}^{\text{holo}}$) with the predicted binding affinity of the ligand within the *apo* active site following 5000 steps of steepest-descent minimization ($\text{Score}^{\text{apo}}$) reveals that accounting for induced-fit conformational changes significantly alters the predicted binding energies of some compounds (Figure 4).

Table 1 shows the potential ligands obtained by applying the procedure depicted in Figure 2. We identified nine fragments predicted to bind in the S1' pocket and three predicted to bind in the S2' or S3' pockets, hereafter referred to jointly as S2'/S3'. As expected, fragments targeting the S1' pocket were generally elongated and hydrophobic. Due to the less hydrophobic and more solvent-exposed characteristics of the S2'/S3' sites, two of the three candidates targeting these pockets were polar and charged. The third, T12 (uncharged), is also a S1' binder; in fact, ensemble-average predicted binding energies suggest T12 prefers the S1' pocket to the S2'/S3' pockets. The predicted binding energies calculated from the *apo* protein ($\text{Score}^{\text{apo}}$) were only loosely correlated with the ensemble-average predicted binding energies ($\text{Score}^{\text{holo}}$) (R^2 less than 0.3), indicating that altered receptor flexibility due to induced-fit effects has a significant impact on the predicted energy of binding.

For ligands such as those associated with fragments UD2b, UE8, and U97, significant changes in the predicted binding energy occur when the subtle changes in protein configuration that occur upon binding are considered. For instance, the ligand associated with fragment UE8, which was selected as a reliable and strong binder due to its high LUDI frequency and its maximum LUDI score, was among the compounds with the highest

binding affinity when only conformations of the *apo* protein were considered (Table 1, see Score^{apo}). However, when induced-fit effects were included in our scoring procedure, the ligand associated with fragment UE8 became the second worst candidate.

Interesting results were also obtained for the ligands associated with the fragments UD2a and UD2b. The high LUDI scores of these tautomers qualified both as strong binders. Nevertheless, after re-scoring the compounds using the AutoDock 4.0 scoring function, the ligand associated with fragment UD2b became the least potent candidate on the list ($\text{Score}^{apo} = -4.40$ kcal/mol), though the ligand corresponding to fragment UD2a still scored well. As expected, after equilibrating each ligand-protein complex, the ensemble-average scores (Score^{holo}) of the tautomers (UD2a and UD2b) converged to nearly the same value.

To measure the subtle changes in the active-site configuration that yield such drastic changes in the predicted binding energy (i.e. the “induced-fit effects”), we further analyzed the top four predicted inhibitors, those associated with fragments UD2, T12, UB9, and UG1. For each of these inhibitors, the frame from the initial *apo* MD simulation that gave the highest LUDI score, the same frame following minimization with ligand present, and the corresponding 3.5 ns productive portion of each associated *holo* trajectory were all aligned by minimizing the RMSD between the active-site non-hydrogen protein atoms of 26 residues lining the active site of the MMP-2 crystal structure (PDB: 1QIB, Table 2). The results suggest that even small changes in the active-site configuration, as judged by the small changes in the RMSD to the crystal-structure active site, can lead to substantial changes in predicted binding affinity, as judged by $\text{Score}^{apo} - \text{Score}^{holo}$. Clearly, even limited protein flexibility can have a major impact on the predicted energy of binding.

An analysis of the *holo* trajectories of the top four compounds also reveals a number of common ligand binding modes. The H201 side chain plays an important role in the binding of all these compounds. All four form critical pi-pi interactions with H201; additionally, UD2, UB9, and UG1 also form hydrogen bonds with the H201 side chain. Protein backbone carbonyl oxygen atoms are also critical for binding, as they accept hydrogen bonds from three of the top four ligands; UD2, UB9, and UG1 all donate hydrogen bonds to the backbone carbonyl oxygen atoms of residues I222 and A220. Additionally, UD2 and UG1 also form hydrogen bonds with the backbone carbonyl oxygen atoms of residues L218 and L197. Finally, the backbone carbonyl oxygen atom of P221 and the backbone amine of I222 form hydrogen bonds with UD2 and UG1, respectively.

Controls

In order to validate our methodology, we included three controls, AM-2, AM-5, and AM-6, which are known to inhibit MMP-2 with IC_{50} values of 9.3 μM , 0.61 μM , and > 50 μM , respectively (Puerta, Mongan et al. 2005). A procedure similar to that used above was repeated with these three controls. Our two positive controls, AM-2 and AM-5, had predicted binding energies of 9.18 ± 0.58 and -9.86 ± 0.56 kcal/mol, respectively. Given that AutoDock has a residual standard error of 2.177 kcal mol⁻¹ (Morris, Goodsell et al. 1998), these positive controls have predicted binding energies comparable to those of the best novel compounds identified (Table 1).

AM-6, included as a negative control, was incorrectly identified as a strong binder (Table 1). In order to understand why the current methodology led to this false positive result, we examined the LUDI dockings of the AM-6 fragment into the protein configurations of the *apo* trajectory. Those dockings with the highest LUDI scores placed the AM-6 fragment deep within the hydrophobic S1' pocket; one of these dockings was subsequently used to construct the initial frame of the *holo* AM-6 MD simulation. However, there were very few of these high-scoring dockings, suggesting that MMP-2 can only rarely accommodate the

large AM-6 fragment. This observation is supported in part by the fact that LUDI was only able to dock AM-6 into the MMP-2 active sites of 24% of the frames extracted from the *apo* MD trajectory (Table 1). Even when docking was successful, docking of the AM-6 fragment into a well defined S1' pocket was rare.

After generating a histogram of all AM-6 LUDI scores, we noted that the most populated cluster corresponded to scores around 450. We therefore repeated the *holo* AM-6 MD simulation, this time using ligand and protein configurations associated with a LUDI score of 450 to create the initial frame. After 1.5 ns of equilibration, we let the simulation run another 1.35 ns and analyzed the AutoDock-predicted binding energy of the productive ensemble, as described above. As expected, the average AutoDock-predicted binding affinity was substantially lower when more representative protein and AM-6 ligand conformations were considered. In cases where the protein configurations associated with the best LUDI scores are rare, we therefore suggest using more representative protein and ligand configurations for subsequent *holo* simulations instead.

Generating Composite Compounds

The current scheme is amenable to the design of composite compounds in which optimum fragments targeting both the S1' and the S2'/S3' pockets are simultaneously added to the same ZBG (described as targeting “both” pockets in Table 1). In this study, a S1'-binding fragment and a S2'/S3'-binding fragment were selected to create a high-affinity composite compound. Fragment UD2a, which had the highest ensemble-average AutoDock binding affinity ($\text{Score}^{\text{holo}}$) of all novel hits, was selected as the S1'-binding fragment. As fragment U97 merged with the ZBG was the only S2'/S3'-binding ligand with a potential linker hydrogen pointed in the direction of the S1' pocket, U97 was selected as the S2'/S3' fragment. The composite compound was created by simultaneously replacing the two hydrogen-atom linkers (Figure 1, inset figure) with the UD2a and U97 fragments, respectively, producing a chiral compound in the S configuration. The initial orientation of each fragment corresponded to the ensemble-maximum LUDI score for each. Two reasonable receptor configurations were envisioned for the ligand-protein complex: the optimal *apo* conformation corresponding to fragments U97 and UD2a, respectively (Figures 3A and 3B). To predict the binding affinity of our composite ligand, two sets of MD and AutoDock-rescoring calculations were performed, as described above, each starting from one of these initial protein structures. Figure 3C displays the final configuration of the MD simulation started from the optimal *apo* conformation for the UD2a fragment. A similar configuration was obtained when the optimal *apo* conformation for the U97 fragment was used.

After equilibrating each of the two MD simulations of the composite compound (UD2a + U97), the ensemble-average scores ($\text{Score}^{\text{holo}}$) of both converged to nearly the same value. Accounting for induced-fit effects clearly has a substantial impact on the predicted binding energy by removing dependency on a single, often unrepresentative receptor conformation.

Conclusion

The rational design of MMP inhibitors is challenging due to the flexibility of the MMP active site. In this work, we combined computer docking with MD simulations in order to account for receptor flexibility, conformational selection mechanisms, and induced-fit effects. Our strategy identified fragments that bind the S1' pocket as well as fragments that bind elsewhere. Subsequently combining both types of fragments with an experimentally validated MMP ZBG group improved predicted potency. Our results indicate the importance of taking into account conformational selection and induced-fit effects when dealing with

highly flexible proteins like the MMPs, as these factors can have a significant impact on the predicted binding energies of candidate compounds.

Acknowledgments

This work was supported in part by grants from the UCSD School of Medicine, NSF, NIH, the Center for Theoretical Biological Physics, the National Biomedical Computation Resource, San Diego Supercomputing Center, and Accelrys, Inc. We thank Dr. Seth Cohen and his group for helpful discussions.

References

- Adcock SA, McCammon JA. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev.* 2006; 106(5):1589–615. [PubMed: 16683746]
- Amaro RE, Baron R, et al. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *Journal of computer-aided molecular design.* 2008
- Amaro RE, Schnauffer A, et al. Discovery of drug-like inhibitors of an essential RNA-editing ligase in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences.* 2008:17278–17283.
- Auge F, Hornebeck W, et al. Improved gelatinase selectivity by novel zinc binding groups containing galardin derivatives. *Bioorganic & medicinal chemistry letters.* 2003; 13(10):1783. [PubMed: 12729664]
- Berendsen HJC, Postma JPM, et al. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics.* 1984; 81(8):3684.
- Bohm HJ. On the use of LUDI to search the Fine Chemicals Directory for ligands of proteins of known three-dimensional structure. *Journal of computer-aided molecular design.* 1994; 8(5):623. [PubMed: 7876904]
- Brown PD. Clinical studies with matrix metalloproteinase inhibitors. *APMIS: Acta Pathologica, Microbiologica, et Immunologica Scandinavica.* 1999; 107(1):174.
- Cainelli G, Galletti P, et al. 4-Alkylidene-Azetidin-2-Ones: Novel Inhibitors of Leukocyte Elastase and Gelatinase. *Bioorganic & medicinal chemistry.* 2003; 11(24):5391. [PubMed: 14642583]
- Campestre C, Agamennone M, et al. N-Hydroxyurea as zinc binding group in matrix metalloproteinase inhibition: mode of binding in a complex with MMP-8. *Bioorganic & medicinal chemistry letters.* 2006; 16(1):20. [PubMed: 16242329]
- Case DA, Darden TA, et al. *Amber.* 2006; 9
- Cataldo D, Munaut C, et al. Matrix metalloproteinases and TIMP-1 production by peripheral blood granulocytes from COPD patients and asthmatics. *Allergy.* 2001; 56(2):145. [PubMed: 11167375]
- Chambers AF, Matrisian LM. Changing views of the role of matrix metalloproteinases in metastasis. *Journal of the National Cancer Institute.* 1997; 89(17):1260. [PubMed: 9293916]
- Chang C, Werb Z. The many faces of metalloproteases: cell growth, invasion, angiogenesis and metastasis. *Trends in cell biology.* 2001; 11(11):S37. [PubMed: 11684441]
- Cheng LS, Amaro RE, et al. Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. *J Med Chem.* 2008; 51(13):3878–94. [PubMed: 18558668]
- Coussens LM, Fingleton B, et al. Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science (New York, NY).* 2002; 295(5564):2387.
- Coussens LM, Werb Z. Matrix metalloproteinases and the development of cancer. *Chemistry & Biology.* 1996; 3(11):895. [PubMed: 8939708]
- Cuniasse P, Devel L, et al. Future challenges facing the development of specific active-site-directed synthetic inhibitors of MMPs. *Biochimie.* 2005; 87(3–4):393–402. [PubMed: 15781327]
- de Oliveira CA, Zissen M, et al. Molecular dynamics simulations of metalloproteinases types 2 and 3 reveal differences in the dynamic behavior of the S1' binding pocket. *Curr Pharm Des.* 2007; 13(34):3471–5. [PubMed: 18220784]
- Denis LJ, Verweij J. Matrix metalloproteinase inhibitors: present achievements and future prospects. *Investigational new drugs.* 1997; 15(3):175. [PubMed: 9387040]

- Dhanaraj V, Williams MG, et al. X-ray Structure of Gelatinase A Catalytic Domain Complexed with a Hydroxamate Inhibitor. *Croatica chemica acta*. 1999; 72:575.
- Essmann U, Perera L, et al. A smooth particle mesh Ewald method. *The Journal of Chemical Physics*. 1995; 103(19):8577.
- Farkas E, Katz Y, et al. Carbamoylphosphonate-based matrix metalloproteinase inhibitor metal complexes: solution studies and stability constants. Towards a zinc-selective binding group. *Journal of Biological Inorganic Chemistry: JBIC: a publication of the Society of Biological Inorganic Chemistry*. 2004; 9(3):307. [PubMed: 14762707]
- Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity--a rapid access to atomic charges. *Tetrahedron*. 1980; 36(22):3219–3228.
- Hidalgo M, Eckhardt SG. Development of Matrix Metalloproteinase Inhibitors in Cancer Therapy. *JNCI Journal of the National Cancer Institute*. 2001; 93(3):178.
- Hockney RW. Potential calculation. 1968; 13:1747.
- Hu J, Van den Steen PE, et al. Matrix metalloproteinase inhibitors as therapy for inflammatory and vascular diseases. *Nature reviews Drug discovery*. 2007; 6(6):480.
- Jacobsen FE, Lewis JA, et al. A new role for old ligands: discerning chelators for zinc metalloproteinases. *Journal of the American Chemical Society*. 2006; 128(10):3156. [PubMed: 16522091]
- Jakalian A, Bush BL, et al. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *Journal of Computational Chemistry*. 2000; 21(2):132.
- Jorgensen WL, Chandrasekhar J, et al. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*. 1983; 79(2):926.
- Kahari VM, Saarialho-Kere U. Matrix metalloproteinases and their inhibitors in tumour growth and invasion. *Annals of Medicine*. 1999; 31(1):34. [PubMed: 10219712]
- Kleiner DE, Stetler-Stevenson WG. Matrix metalloproteinases and metastasis. *Cancer chemotherapy and pharmacology*. 1999; 43(Suppl):S42. [PubMed: 10357558]
- Lim NC, Schuster JV, et al. Coumarin-based chemosensors for zinc(II): toward the determination of the design algorithm for CHEF-type and ratiometric probes. *Inorganic chemistry*. 2005; 44(6): 2018. [PubMed: 15762729]
- Lin JH, Perryman AL, et al. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *Journal of the American Chemical Society*. 2002; 124(20):5632. [PubMed: 12010024]
- Lin JH, Perryman AL, et al. The relaxed complex method: Accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers*. 2003; 68(1):47. [PubMed: 12579579]
- Lovejoy B, Welch AR, et al. Crystal structures of MMP-1 and -13 reveal the structural basis for selectivity of collagenase inhibitors. *Nat Struct Biol*. 1999; 6(3):217–21. [PubMed: 10074939]
- Lukacova V, Zhang Y, et al. Similarity of binding sites of human matrix metalloproteinases. *The Journal of biological chemistry*. 2004; 279(14):14194. [PubMed: 14732707]
- Matter H, Schudok M. Recent advances in the design of matrix metalloprotease inhibitors. *Curr Opin Drug Discov Devel*. 2004; 7(4):513–35.
- Mautino G, Oliver N, et al. Increased release of matrix metalloproteinase-9 in bronchoalveolar lavage fluid and by alveolar macrophages of asthmatics. *American journal of respiratory cell and molecular biology*. 1997; 17(5):583. [PubMed: 9374109]
- Michaelides MR, Dellaria JF, et al. Biaryl ether retrohydroxamates as potent, long-lived, orally bioavailable MMP inhibitors. *Bioorganic & medicinal chemistry letters*. 2001; 11(12):1553. [PubMed: 11412979]
- Mook ORF, Frederiks WM, et al. The role of gelatinases in colorectal cancer progression and metastasis. *Biochimica et Biophysica Acta (BBA) -Reviews on Cancer*. 2004; 1705(2):69.
- Morris GM, Goodsell DS, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding freeenergy function. *Journal of Computational Chemistry*. 1998; 19(14):1639–1662.

- Moy FJ, Chanda PK, et al. High-resolution solution structure of the catalytic fragment of human collagenase-3 (MMP-13) complexed with a hydroxamic acid inhibitor. *J Mol Biol.* 2000; 302(3): 671–89. [PubMed: 10986126]
- Moy FJ, Chanda PK, et al. High-resolution solution structure of the inhibitor-free catalytic fragment of human fibroblast collagenase determined by multidimensional NMR. *Biochemistry.* 1998; 37(6): 1495–504. [PubMed: 9484219]
- Moy FJ, Pisano MR, et al. Assignments, secondary structure and dynamics of the inhibitor-free catalytic fragment of human fibroblast collagenase. *J Biomol NMR.* 1997; 10(1):9–19. [PubMed: 9335112]
- Onaran MB, Comeau AB, et al. Squaric acid-based peptidic inhibitors of matrix metalloprotease-1. *The Journal of organic chemistry.* 2005; 70(26):10792. [PubMed: 16356002]
- Pirard B. Insight into the structural determinants for selective inhibition of matrix metalloproteinases. *Drug discovery today.* 2007; 12(15–16):640. [PubMed: 17706545]
- Prikk K, Maisi P, et al. Airway obstruction correlates with collagenase-2 (MMP-8) expression and activation in bronchial asthma. *Laboratory investigation; a journal of technical methods and pathology.* 2002; 82(11):1535.
- Puerta DT, Cohen SM. Elucidating drug-metalloprotein interactions with tris(pyrazolyl)borate model complexes. *Inorganic chemistry.* 2002; 41(20):5075. [PubMed: 12354040]
- Puerta DT, Cohen SM. Examination of novel zinc-binding groups for use in matrix metalloproteinase inhibitors. *Inorganic chemistry.* 2003; 42(11):3423. [PubMed: 12767177]
- Puerta DT, Griffin MO, et al. Heterocyclic zinc-binding groups for use in next-generation matrix metalloproteinase inhibitors: potency, toxicity, and reactivity. *Journal of Biological Inorganic Chemistry: JBIC: a publication of the Society of Biological Inorganic Chemistry.* 2006; 11(2):131. [PubMed: 16391944]
- Puerta DT, Lewis JA, et al. New beginnings for matrix metalloproteinase inhibitors: identification of high-affinity zinc-binding groups. *Journal of the American Chemical Society.* 2004; 126(27):8388. [PubMed: 15237990]
- Puerta DT, Mongan J, et al. Potent, selective pyrone-based inhibitors of stromelysin-1. *Journal of the American Chemical Society.* 2005; 127(41):14148. [PubMed: 16218585]
- Puerta DT, Schames JR, et al. From Model Complexes to Metalloprotein Inhibition: A Synergistic Approach to Structure-Based Drug Discovery. *Angewandte Chemie International Edition.* 2003; 42(32):3772.
- Ray JM, Stetler-Stevenson WG. Gelatinase A activity directly modulates melanoma cell adhesion and spreading. *The EMBO journal.* 1995; 14(5):908. [PubMed: 7534227]
- Reich R, Katz Y, et al. Carbamoylphosphonate matrix metalloproteinase inhibitors 3: in vivo evaluation of cyclopentylcarbamoylphosphonic acid in experimental metastasis and angiogenesis. *Clinical cancer research: an official journal of the American Association for Cancer Research.* 2005; 11(10):3925. [PubMed: 15897594]
- Ryckaert J-P, Ciccotti G, et al. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *Journal of Computational Physics.* 1977; 23:327.
- Sanner MF. A component-based software environment for visualizing large macromolecular assemblies. *Structure.* 2005; 13(3):447–62. [PubMed: 15766546]
- Schames JR, Henchman RH, et al. Discovery of a novel binding trench in HIV integrase. *Journal of medicinal chemistry.* 2004; 47(8):1879. [PubMed: 15055986]
- Skiles JW, Gonnella NC, et al. The design, structure, and clinical update of small molecular weight matrix metalloproteinase inhibitors. *Current medicinal chemistry.* 2004; 11(22):2911. [PubMed: 15544483]
- Wang J, Cieplak P, et al. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry.* 2000; 21(12):1049.
- Whittaker M, Floyd CD, et al. Design and therapeutic application of matrix metalloproteinase inhibitors. *Chemical reviews.* 1999; 99(9):2735. [PubMed: 11749499]

- Wojtowicz-Praga SM, Dickson RB, et al. Matrix metalloproteinase inhibitors. Investigational new drugs. 1997; 15(1):61. [PubMed: 9195290]
- Yong VW. Metalloproteinases: mediators of pathology and regeneration in the CNS. Nature reviews Neuroscience. 2005; 6(12):931.
- Yuan P V, Marshall P, et al. Dynamics of stromelysin/inhibitor interactions studied by 15N NMR relaxation measurements: comparison of ligand binding to the S1-S3 and S'1-S'3 subsites. J Biomol NMR. 1999; 15(1):55–64. [PubMed: 10549133]
- Zhang X, Gonnella NC, et al. Solution structure of the catalytic domain of human collagenase-3 (MMP-13) complexed to a potent non-peptidic sulfonamide inhibitor: binding comparison with stromelysin-1 and collagenase-1. J Mol Biol. 2000; 301(2):513–24. [PubMed: 10926524]

Figure 1.

Superposition of equally spaced snapshots extracted from the MD trajectory of the *apo* protein. Fitting was performed by aligning the catalytic zinc cation and histidine residues of each structure. The zinc cation (grey) and coordinating nitrogen atoms of histidine sidechains (blue) are shown in van der Waals representation. The ZBG, maltol (3-hydroxy-2-methyl-4-pyrone), and an amide link are shown in balls and sticks representation. The inset figure shows the ZBG (licorice representation) and the catalytic zinc cation (van der Waals representation) in the MMP-2 active site (surface representation colored according to the protein electrostatic potential). The yellow circles highlight the hydrogen atoms used as LUDI linking points through which interacting fragments were added.

Figure 2.

A and B) A molecular dynamics simulation of the *apo* MMP-2 protein was analyzed in order to account for protein flexibility. Three thousand and one regularly spaced frames were extracted from the simulation, and the ZBG was placed into each protein configuration in harmony with the Cohen bioinorganic model. C) LUDI docking calculations were then performed on the 3001 snapshots of the *apo* MD trajectory to add interacting moieties to the ZBG *via* linking hydrogen atoms. D) The “best” interacting fragments were considered to be those that were either frequently selected as binders across multiple protein configurations regardless of LUDI score or had the highest ensemble-wide maximum LUDI score regardless of frequency. E) A 5 ns MD simulation was then performed on each protein-ligand system selected previously. F) To predict the binding affinity of each ligand, frames were extracted at regular intervals from the last 3.5 ns of each of these secondary MD simulations. For each system, the AutoDock 4.0 scoring function was used to estimate the ligand binding energy for all frames; these predicted binding energies were then averaged, yielding an ensemble-average predicted binding energy for each ligand. A small part of this graphic was obtained from Puerta, Schames et al. 2003.

Figure 3. LUDI-docked fragments added to a ZBG. A) The ligand associated with fragment UD2a. B) The ligand associated with fragment U97. C) The equilibrated structure of the complex formed between the composite compound (UD2a + U97) and the last MMP-2 configuration of a *holo* MD simulation.

Figure 4. Black) Induced-fit energy contribution. The induced-fit energy contribution is given by $\text{Score}^{apo} - \text{Score}^{holo}$. Grey) Standard deviation of the ensemble-average score. Contributions from induced-fit effects are considered significant if $(\text{Score}^{apo} - \text{Score}^{holo}) > \text{standard deviation}$.

Table 1

MMP-2 inhibitor candidates obtained by applying the procedure illustrated in Figure 2.

ID	Molecular Structure	LUDI Frequency	LUDI Max Score	Score ^{apo} (kcal/mol ⁻¹)	Score ^{holo} (kcal/mol ⁻¹)	Pocket
AM-6		0.24	716	-12.36	-11.49 ± 0.65	S1'
AM-5		0.86	635	-10.27	-9.86 ± 0.56	S1'
AM-2		0.77	611	-10.22	9.18 ± 0.58	S1'
UD2a		0.52	712	-8.79	-8.20 ± 0.70	S1'

ID	Molecular Structure	LUDI Frequency	LUDI Max Score	Score ^{dp0} (kcal/mol ⁻¹)	Score ^{dp0} (kcal/mol ⁻¹)	Pocket
UD2b		0.37	722	-4.40	-7.89 ± 0.70	S1'
T12		0.73	481	-7.11	-7.52 ± 0.63	S1'
UB9		0.63	594	-6.67	-7.05 ± 0.48	S1'

ID	Molecular Structure	LUFI Frequency	LUFI Max Score	Score ^{dp} (kcal/mol ⁻¹)	Score ^{bo} (kcal/mol ⁻¹)	Pocket
UG1		0.43	749	-7.53	-6.45 ± 0.52	S1'
UF7		0.35	775	-6.86	-6.44 ± 0.82	S1'
T12		0.65	435	-6.22	-5.99 ± 0.48	S2/S3'
AM-6 ^a		x		-9.05	-5.93 ± 0.60	S1'

ID	Molecular Structure	LUDI Frequency	LUDI Max Score	Score ^{epo} (kcal/mol ⁻¹)	Score ^{bio} (kcal/mol ⁻¹)	Pocket
T02		0.62	398	-4.99	-5.34 ± 0.58	S1'
U97		0.35	576	-6.56	-5.30 ± 0.76	S2/S3'
UA5		0.39	654	-5.01	-5.04 ± 0.55	S2/S3'

ID	Molecular Structure	LUDI Frequency	LUDI Max Score	Score ^{epo} (kcal/mol ⁻¹)	Score ^{bole} (kcal/mol ⁻¹)	Pocket
T03		0.68	416	-5.37	-4.74 ± 0.61	S1'
UE8		0.64	602	-7.53	-4.64 ± 0.64	S1'
T18		0.66	450	-5.10	-4.55 ± 0.76	S1'
U97 + UD2a ^b		-	-	-7.31	-9.28 ± 0.87	Both
U97 + UD2a ^c		-	-	-9.41	-9.20 ± 0.64	Both

^aBased on a more representative LUDI score (450), not the best LUDI score. Associated productive *halo* simulation of only 1.35 ns, not 3.5 ns.

^bUsing the optimal U97 protein configuration.

Using the optimal UD2a protein configuration.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Analysis of active-site structural changes that occur upon ligand binding. The frame from the initial MD simulation that gave the highest LUDI score (“Frame”), the same frame following minimization (“Min”), and the corresponding 3.5 ns productive portion of the associated *holo* trajectory (“Trajectory Average”) were all aligned by minimizing the RMSD between the non-hydrogen atoms of 26 residues lining the active site of the MMP-2 crystal structure (PDB: 1QIB)

ID	Frame (RMSD)	Min (RMSD)	Trajectory Average (RMSD)	Score ^{apo} – Score ^{holo} (kcal/mol)
UD2a	3.57	3.54	3.60 ± 0.08	0.59
UD2b	3.62	3.59	3.62 ± 0.05	–3.49
T12	3.61	3.59	3.37 ± 0.05	–0.41
UB9	3.52	3.51	3.52 ± 0.09	–0.38
UG1	3.50	3.49	3.94 ± 1.15	1.08