ELSEVIER

Docking studies of matrix metalloproteinase inhibitors: zinc parameter optimization to improve the binding free energy prediction

Xin Hu, William H. Shelver*

Department of Pharmaceutical Science, North Dakota State University, Fargo, ND 58105, USA

Abstract

Docking of metalloproteinase inhibitors remains a challenge due to the zinc multiple coordination geometries and the lack of appropriate force field parameters to model the metal/ligand interactions. In this study, we explore the docking accuracy and scoring reliability for the docking of matrix metalloproteinase (MMP) inhibitors using AutoDock 3.0. Potential problems associated with zinc ion were investigated by docking 16 matrix metalloproteinase ligands to their crystal structures. A good coordination between the zinc binding group (ZBG) and the zinc was shown to be a prerequisite for the ligand to fit the binding site. A simplex optimization of zinc parameters, including zinc radius, well depth, and zinc charges, was performed utilizing the 14 MMP complexes with good docking. The use of optimized zinc parameters (zinc radius: 0.87 Å; well depth: 0.35 kcal/mol; and zinc charges: +0.95 e) shows improvement in both docking accuracy at the zinc binding site and the prediction of binding free energies. Although further improvement in the docking procedure, particularly the scoring function is needed, optimization of zinc parameters provides an efficient way to improve the performance of AutoDock as a drug discovery tool.
© 2003 Elsevier Inc. All rights reserved.

Keywords: Docking; AutoDock; Matrix metalloproteinase inhibitors; Zinc parameters; Simplex algorithm; Structure-based drug design

1. Introduction

Matrix metalloproteinases (MMPs) belong to a family of zinc-dependent endopeptidases responsible for the metabolism of extracellular matrix proteins [1,2]. These structurally related enzymes play a crucial role in matrix remodeling and wound healing [3]. Disease processes associated with the MMPs are involved in an imbalance between the activation and inhibition of the MMPs, and alteration in level of MMPs is implicated in a wide range of pathological states such as osteoarthritis, rheumatoid arthritis, cancer, joint destruction, Alzheimer, and multiple sclerosis [4,5]. Therefore, MMP represents an attractive pharmacological drug target and inhibition of MMPs has become a promising therapeutic strategy for the treatment of such diseases.

On the basis of substrate specificity and primary sequence similarities, MMPs can be grouped into five subfamilies: collagenases (MMP-1, -8, and -13), stromelysins (MMP-3, -10, and -11), gelatinases (MMP-2, and -9), membrane-type MMPs (MMP-14–17), and others (MMP-7) [6]. Presently, about 22 human MMPs are known. Structures of many of

these MMPs, such as MMP-1–3, -7–9, and -11–14 have been determined. More than 70 X-ray crystal structures or NMR solution structures complexed with various small molecular inhibitors have been deposited in Protein Data Bank (PDB) [7]. MMP inhibitors are generally characterized by a ZBG coupled to a framework with varying numbers of pocket-occupying side chains. Common zinc binding groups (ZBG) of MMP inhibitors include hydroxamates, carboxylates, phosphinates, and thiolates [8–10]. Of these, inhibitors with the hydroxamic acid have been found to be most potent. However, concerns exist about the unfavorable pharmacokinetics, poor solubility, and potential toxicities presumably from metabolism of the hydroxamate group, making their clinical use problematic [11].

As important therapeutic drug targets, MMPs have recently attracted great interest in the search of potent and selective inhibitors using computer-aided molecular modeling and docking techniques [12–16]. Docking as an efficient in silico tool is playing an ever increasing role in structure-based drug design [17–19]. However, docking to MMPs remains a challenge due to the multiple coordination geometries of zinc and the lack of appropriate force field parameters to model the metal/ligand interactions. Most of those widely used docking programs such as DOCK and

* Corresponding author. Tel.: +1-701-232-3071.
E-mail address: william.shelver@ndsu.nodak.edu (W.H. Shelver).

AutoDock lack reliable zinc parameters. In practice, zinc parameters are usually adopted from other studies [20,21], such as Stote and Karplus' work [20]. These zinc parameters are derived for different purposes and their use in a docking program without validation may lead to erroneous results. Furthermore, metal ions are associated with a variety of coordination states. Zinc is commonly found to be four-coordinated with a tetrahedral geometry. Five and six coordinated geometries such as octahedral, trigonal bipyramidal, and square-base pyramidal are also observed in zinc metalloproteinases [22]. These geometries play an important role in the metal/ligand binding. Modeling of zinc/ligand interactions with such multifarious coordination is quite problematic. The metal-associated problems in docking have been studied recently [23–26].

In conjunction with our efforts in the design of novel potent and selective MMP inhibitors, we undertook a docking study on metalloproteinases using AutoDock [27–29]. Preliminary studies showed that AutoDock is a robust approach with good docking accuracy and reliability in the docking of metalloproteinase inhibitors. However, problems associated with zinc were found, causing a substantial error in the prediction of binding affinities. In this study, we investigated the effects of zinc ion on the docking results. The docking accuracy and scoring reliability of AutoDock were evaluated by reproducing a set of MMP X-ray structure complexes. To improve the prediction of binding free energy, we performed a simplex optimization of zinc parameters including zinc radius, well depth, and zinc charges. The use of optimized zinc parameters for AutoDock showed significant improvement in both docking accuracy at the zinc binding site and the prediction of binding free energies, making the program more efficient and reliable to be utilized in structure-based drug design.

2. Material and methods

2.1. Dataset of MMP/ligand complexes

Three-dimensional coordinates of MMPs structures were obtained from the PDB [7]. The complexes were selected based on the following criteria: (1) crystal structures with resolution better than 2.6 Å; (2) temperature factor of zinc lower than 20; (3) the dissociation equilibrium constant (K_i) available. In addition, the structural and functional diversities of both metalloproteinases and ligands were considered [38]. Table 1 lists the selected 16 metalloproteinases/ligand crystal structures. Structures of ligands are shown in Fig. 1 and includes examples of ligands containing all of the common ZBGs. The flexibility of these ligands was highly variable containing anywhere from 6 to 19 rotatable bonds.

The Sybyl 6.8 [39] package was used to prepare the docking files. The ligand was extracted from protein complex in the PDB file and hydrogen atoms were added. A short minimization (100 steepest descent steps using Tripos force fields with a gradient convergence value of 0.05 kcal/mol Å) was performed to release any internal strain. The partial atomic charges were calculated and the Gasterger-Marstil charges [40] were assigned. All the ligands were used in the charged form, i.e., the hydroxamate, carboxylate, phosphinate, and thiol were deprotonated. For hydroxamates that exist in a number of different tautomeric forms, the one in agreement with FlexX [41] was selected. For the protein, the hetero atoms including cofactors, water molecules, and the ligand were removed. The zinc ion at the active site was retained. Polar hydrogens were added and Kollman united atom charges were assigned. The desolvation parameter was assigned to protein using the AutoDock module ADDSOL. All the residues and charges were visually inspected to

Table 1
The 16 MMP/ligand complexes used in the docking studies

PDB	Protein/ligand	Ligand type	pK_i	Resolution (Å)	B (Zn)	Geometry[a]	Distance[b]		RMSD[c]	Reference
966C	MMP-1/RS2	Hydroxamate	7.64	1.90	5.77	SP	2.10	2.21	21.79	[30]
1B8Y	MMP-3/IN7	Carboxylate	7.85	2.00	16.77	TB	2.02	2.85	18.55	[31]
1CAQ	MMP-3/DPS	Carboxylate	7.44	1.80	14.59	TB	1.99	2.83	18.28	[31]
1CIZ	MMP-3/DPS	Carboxylate	7.72	1.64	10.96	TB	2.09	2.62	17.69	[31]
1HFS	MMP-3/L04	Carboxylate	8.70	1.70	13.92	TB	1.84	2.62	16.40	[32]
1MMP	MMP-7/RSS	Carboxylate	6.07	2.30	12.36	TB	1.96	2.66	19.64	[33]
1MMQ	MMP-7/RRS	Hydroxamate	7.52	1.90	17.71	SP	2.16	2.20	18.18	[33]
1MMR	MMP-7/SRS	Sulfodiimine	5.40	2.30	13.27	TH	2.00	–	13.80	[33]
1A85	MMP-8/HMI-Asn-BNN	Hydroxamate	4.59	2.00	8.06	TB	2.35	2.16	15.17	[34]
1A86	MMP-8/HMI-Asp-BNN	Hydroxamate	3.87	2.00	10.11	TB	1.87	2.41	15.60	[34]
1JAO	MMP-8/BTP-Asp-GM1	Thiolate	5.92	2.40	9.37	TH	2.24	–	11.07	[35]
1JAQ	MMP-8/HMP-Asp-GM1	Hydroxamate	4.48	2.25	10.04	TB	2.45	2.10	17.41	[35]
1JAP	MMP-8/ProLeuGly-HOA	Hydroxamate	4.72	1.82	16.64	SP	2.18	1.91	15.20	[36]
1HV5	MMP-11/RCX	Phosphinate	8.05	2.60	19.61	TB	2.38	3.00	23.46	[37]
456C	MMP-13/RS-113456	Hydroxamate	9.77	2.40	15.22	SP	1.89	1.96	17.86	[30]
830C	MMP-13/RS-130830	Hydroxamate	9.28	1.60	11.83	SP	2.04	1.85	17.35	[30]

[a] Geometry: TB-trigonal bipyramidal, SP-square-based pyramidal, TH-tetrahedral.

[b] Distances are measured between zinc and coordinated atoms of ligand.

[c] The geometry fit was measured by root mean square deviation of angles, RMSD $= \left[\sum (\alpha_i - \alpha_{ideal})^2 / n\right]^{1/2}$, where α_i is the measured angle (L–Zn–L); α_{ideal} the angle in ideal coordination geometry; and n is the total number of angles of the corresponding geometry (6-TH, 10-TB and SP).
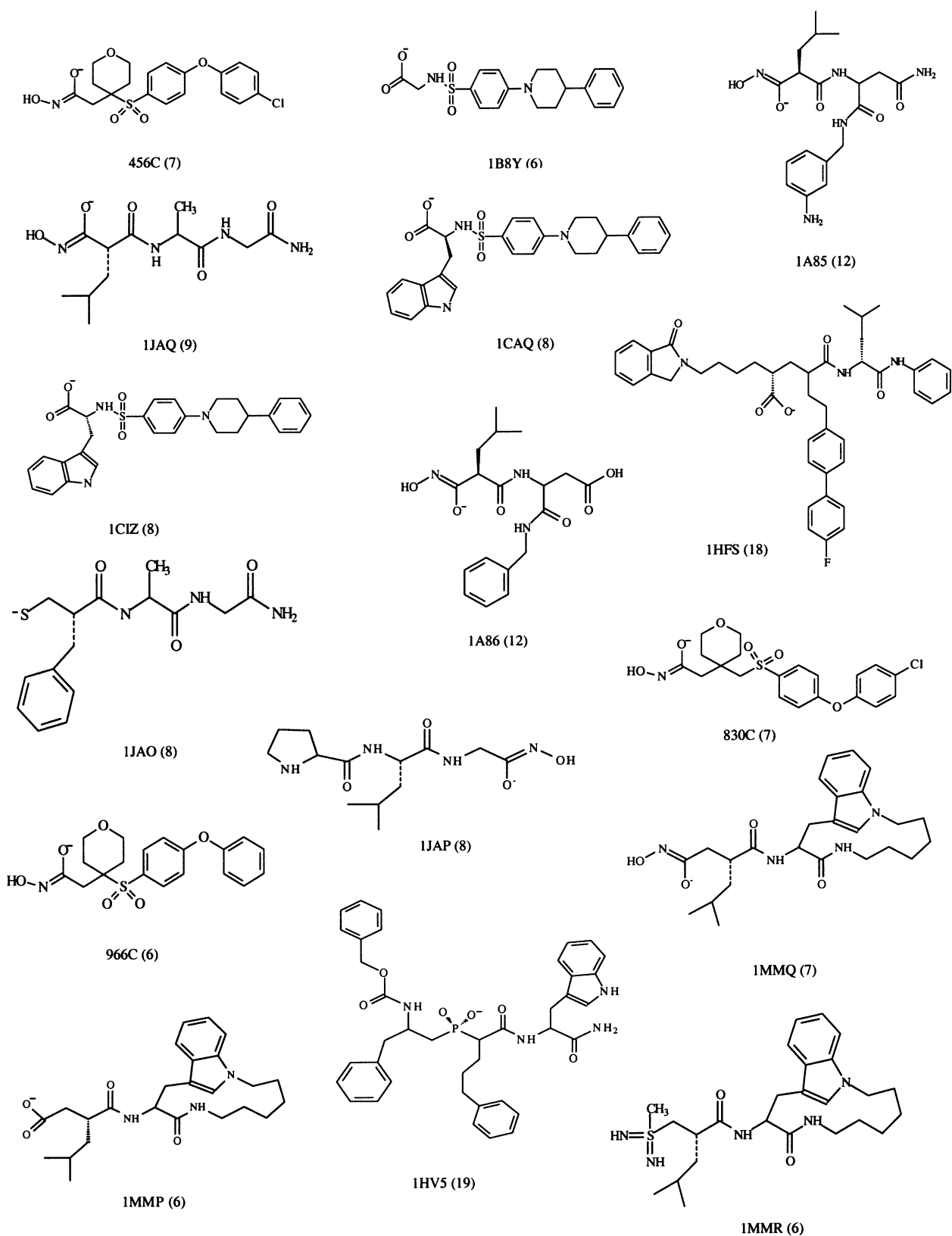
Fig. 1. Structures of the 16 MMP ligands. The number of rotatable bonds is given in parenthesis.

insure consistency and reasonable charges. The three zinc-coordinated histidines were treated as the neutral form with the hydrogen on ND1, while other histidines used the default option with hydrogen on NE2. Glutamates were treated as charged form as default, except the glutamate at the second shell of zinc binding site (Glu 202 in MMP-1, Glu219 in MMP-3, and Glu222 in MMP-8), that were protonated form with the hydrogen on the oxygen nearest the Zn (OE2).

2.2. Docking using AutoDock 3.0

AutoDock 3.0 offers the option of three search algorithms to explore the space of active binding with different efficiency. We used the Lamarckian genetic algorithm (LGA) in this study. Preliminary experiments using simulated annealing showed deficiencies for highly flexible ligands. The active site was defined using AutoGrid. The grid size was set to $90 \times 90 \times 90$ points with a grid spacing of 0.375Å centered on the original ligand in crystal structure complex. The grid box includes the entire binding site of the enzyme and provides enough space for the ligand translational and rotational walk. Step sizes of 1Å for translation and $50°$ for rotation were chosen and the maximum number of energy evaluations was set to 2,000,000. Thirty runs were performed. For each of the 30 independent runs, a maximum number of 27,000 GA operations were generated on a single population of 50 individuals. Operator weights for crossover, mutation, and elitism were default parameters (0.80, 0.02, and 1, respectively).

Metal ions are modeled in AutoDock by Amber force field potentials [42]. However, there are no zinc parameters defined in the program. After examination of zinc parameters (zinc radius and well depth) used in literature, we selected the one from Stote and Karplus' work ($r = 1.1 \text{Å}$, $\varepsilon = 0.25 \text{kcal/mol}$) [20] and a charge of $+2.0$ e as the initial parameters.

3. Results and discussion

3.1. Assessment of docking accuracy and scoring reliability

The docking results of the 16 MMP complexes using AutoDock are shown in Table 2. The top-ranked pose (lowest docked energies) was selected from the 30 runs. Eight out of 16 top-ranked poses had root mean square deviation (RMSD) values less than 1 Å, while 14 out of 16 are considered as well-docked with the RMSD less than 2.0 Å. Only two cases (1JAP and 1HV5) failed to dock correctly with RMSD values of 7.43 and 4.65 Å, respectively. Among the eight dockings with RMSD <1.0 Å, 1MMQ, 1MMR, and 1MMP had an RMSD lower than 0.5 Å. This excellent result occurred because of the relatively rigid cyclic structure of ligands limiting the choices of docking poses and making the energy differences between poses relatively large. A typical example is 1MMP Fig. 2A), in which the excellent agreement between the docked structure and crystal structure is demonstrated. The remaining complexes with RMSD <1.0 Å, 1CAQ, 1B8Y, 830C, 456C, and 966C are characterized with a long P1′ side chain and lack of the P2′–P3′ binding. Fig. 2B shows 830C (RMSD = 0.93) in which the P1′ fits rather well, but the zinc binding hydroxamate group is misplaced. An interesting comparison is 1CIZ and 1CAQ, in which differ only in their configuration. The RMSD value of 1CAQ, the S isomer, is less than 0.60 Å while for the R isomer (1CIZ) the RMSD value is 1.78 Å (Fig. 2C and D).

Table 2
Docking results of 16 MMP inhibitors using AutoDock 3.0 with the initial zinc parameters and the optimized parameters

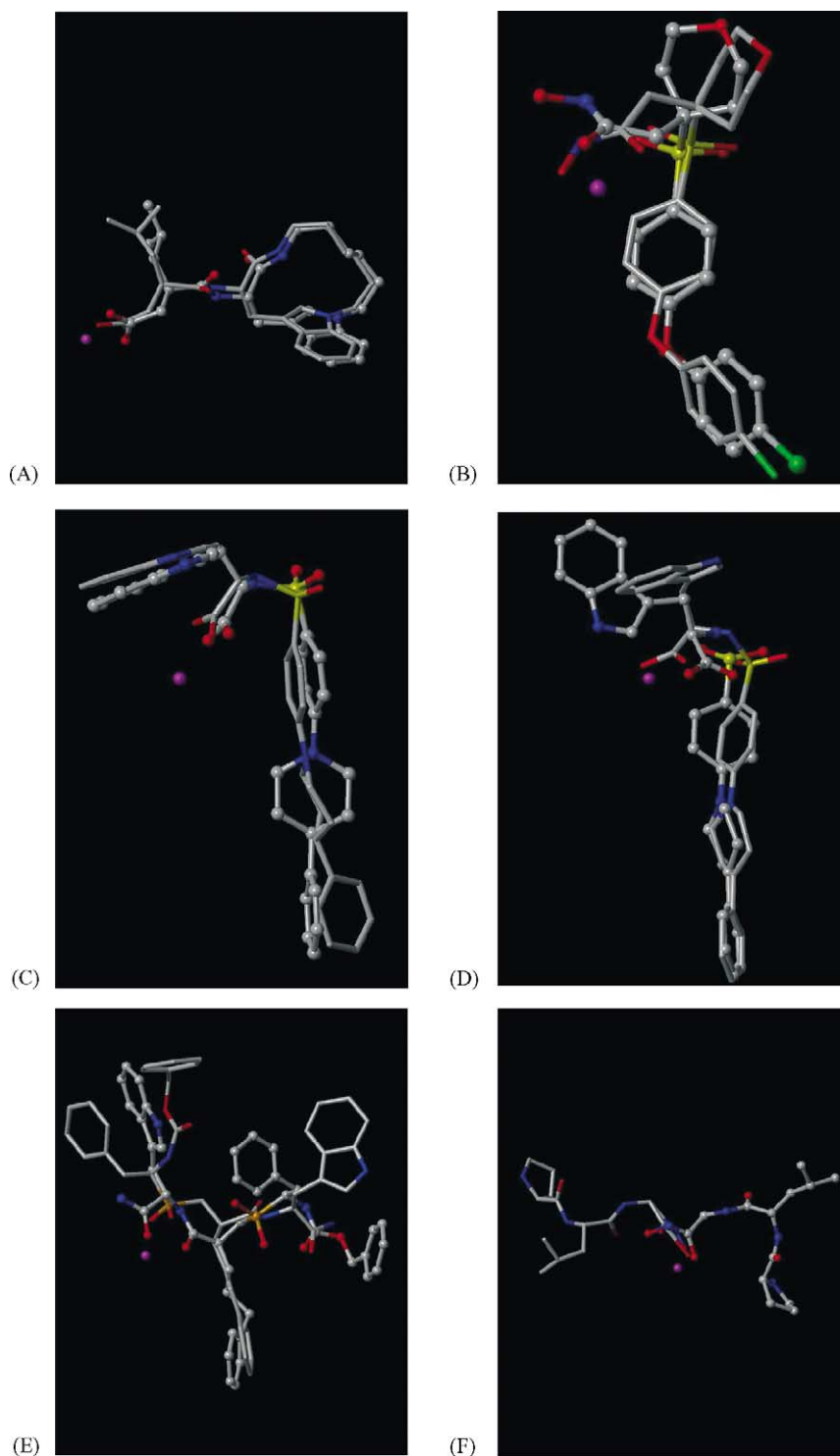| MMP	$\Delta G_{\text{experimental}}$ (kcal/mol)	$r = 1.1 \text{Å}$, $\varepsilon = 0.25 \text{kcal/mol}$, $q = +2.0$ e			$r = 0.87 \text{Å}$, $\varepsilon = 0.35 \text{kcal/mol}$, $q = +0.95$ e		
		RMSD (Å)	$\Delta G_{\text{predicted}}$ (kcal/mol)	$\Delta G_{\text{predicted}} - \Delta G_{\text{experimental}}$ (kcal/mol)	RMSD (Å)	$\Delta G_{\text{predicted}}$ (kcal/mol)	$\Delta G_{\text{predicted}} - \Delta G_{\text{experimental}}$ (kcal/mol)
966C	−10.42	0.88	−12.81	−2.39	0.78	−11.75	−1.33
1B8Y	−10.71	0.75	−13.42	−2.71	0.63	−12.27	−1.56
1CAQ	−10.15	0.56	−16.54	−6.39	0.69	−13.98	−3.83
1CIZ	−10.53	1.87	−17.22	−6.69	1.13	−15.78	−5.25
1HFS	−11.87	1.94	−17.95	−6.08	1.57	−14.52	−2.65
1MMP	−8.28	0.36	−13.57	−5.29	0.38	−11.95	−3.67
1MMQ	−10.26	0.42	−14.23	−3.97	0.56	−11.48	−1.22
1MMR	−7.37	0.50	−12.69	−5.32	0.71	−8.75	−1.38
1A85	−6.26	1.47	−12.87	−6.61	1.21	−8.78	−2.52
1A86	−5.28	1.75	−14.01	−8.73	1.43	−9.32	−4.04
1JAO	−8.08	1.40	−11.50	−3.42	1.15	−11.46	−3.38
1JAQ	−6.11	1.28	−9.33	−3.22	0.94	−9.57	−3.46
456C	−13.33	0.75	−14.85	−1.52	0.85	−12.45	0.88
830C	−12.66	0.93	−13.58	−0.92	0.94	−11.68	0.98
1JAP	−6.44	7.43	−9.77	−3.33	6.48	−8.31	−1.87
1HV5	−10.98	4.65	−17.83	−6.85	5.74	−14.45	−3.47

Fig. 2. Superposition of the docked ligand (balls and sticks) with the experimental conformation (sticks): (A) 1MMP (RMSD = 0.36), (B) 830C (RMSD = 0.93), (C) 1CAQ (RMSD = 0.56), (D) 1CIZ (RMSD = 1.87), (E) 1HV5 (RMSD = 4.65), and (F) 1JAP (RMSD = 7.43).

The R isomer docking was unable to fit the zinc binding site appropriately, but placed the P1′ group well. Visual inspection of the six inhibitors with RMSD between 1.0 and 2.0 Å showed that the S1′ groups are generally placed correctly, but at the remaining parts, often in solvent-exposed areas present relatively poor fits.

Examination of the two incorrect dockings indicates a potential reason for the failure to find the correct structure. Docking of 1HV5 is quite challenging due to the high flexibility. The predicted conformation with the lowest free energy was −17.83 kcal/mol with RMSD of 4.65 Å (Fig. 2E). Analyses of the ligand binding show that several binding

modes were adopted, the four pocket binding groups (P1, P2, P1′, P2′) were misplaced in most binding modes. To further investigate the failure of docking, the free energy of the correct binding conformation was computed, and was higher by 3.28 kcal/mol than the lowest energy docked structure. The ligand in 1JAP Fig. 2F) was docked in the prime side pockets as the most energetically stable conformation instead of non-prime pockets found in the crystal structure. The lowest energy docked structure places an isopropyl group in the hydrophobic S1′ pocket. The experimental non-prime mode (RMSD = 0.53 Å) was found in docking groups with much higher energy (ranked 22 of the 30 runs). The free energies difference between the experimentally observed left binding mode and the predicted right binding mode is 2.66 kcal/mol. Therefore, for both complexes, the failure of the docking resulted from the flaw of the scoring function to correctly compute the binding energies.

Incorrectly docked geometries cannot be expected to give accurate binding energies, so the two wrong-docked cases were excluded in the evaluation of scoring reliability. The docking results in Table 2 show that for most of docking of MMP complexes, the binding free energies predicted by AutoDock were greatly overestimated with seven cases with errors greater than 5 kcal/mol and a mean error of −4 kcal/mol. The correlation coefficient (R^2) between the predicted free energies by AutoDock and experimental free energies for the 14 MMP complexes that are docked properly was 0.31 with a standard error of 3.17 kcal/mol. The results are considerably worse than reported in the original AutoDock publication, indicating a potential problem with docking ligands with metalloproteinases.

3.2. Effects of zinc ion on the docking results

We focused on the docking accuracy at the zinc binding site due to the dominant importance of ZBG binding to the scoring reliability of MMP ligands. Examinations showed that a number of docking results suffer from problems associated with zinc. For example, in the case of 1CIZ docking (Fig. 2D), only one oxygen atom of carboxylate binds to zinc, the other oxygen points to the opposite position away from the zinc. Similar phenomena were also observed for 1HFS and 1B8Y. In the case of hydroxamate docking, a monodentate binding geometry rather than the experimentally observed bidentate mode was often generated, as shown in Fig. 2B. In the worst case of 1HV5 (Fig. 2E), no binding between the phosphinic group and zinc was found. However, when distances between zinc and ZBG were restricted, the program easily found the correct binding mode, indicating that the ZBG binding is one of the primary determinants for MMP ligand docking. Clearly, the coordination geometries of zinc complexes are not well reproduced in AutoDock docking using the parameters from Stote and Karplus, even though these parameters allowed generally successful docking. Since Stote and Karplus derived their parameters for molecular dynamics where they have been demonstrated to work well, the failure in AutoDock is not surprising as the docking computations are quite different from the molecular dynamic calculations.

To provide an objective assessment of ZBG binding, we combined two measures. We used the distances between zinc and the ligating atoms as a measure of a bond between zinc and docked ligand. If this distance exceeded 3.0 Å, no bonding existed and the quality of the interaction between the ZBG and Zn was rated as poor. If the distance was less than 3.0, we assessed the overall geometry by using the RMSD of angles between docked and ideal geometries with the tolerance extracted from the experimental values from Table 1. Table 3 shows the resulting classification scheme for assessing ZBG binding. The ZBG binding is considered as good if both distance and geometry fit well, and fair if only the distance is within the reasonable range but geometry is poor. Fig. 3 illustrates typical examples of the three categories of ZBG binding quality.

A summary of the results of the ZBG docking accuracy of the 16 MMP complexes predicted by AutoDock are shown in Fig. 4. The percentage of docking with good ZBG binding was only 40%, with 20% of the dockings showing no interaction between the zinc and the ligand. Although the overall docking accuracy was good, such a poor ZBG docking accuracy confirmed a problem exists concerning docking the ligand to the zinc.

ZBG docking accuracy markedly affects the predicted binding affinities. Analysis of the docking experiments shows that the electrostatic interactions between zinc and ZBG contribute a large portion of binding energies. For example, of a total of 17 kcal/mol of predicted binding free energies for 1CAQ, the zinc/carboxylate interactions contributed 4.75 kcal/mol, most coming from electrostatic interactions. Similar results were also observed for other complexes. The strong electrostatic interactions might be one of the major factors in overestimation of the total predicted binding energies, as observed in the MMP docking.

Table 3
Criteria to evaluate the docking accuracy of ZBG binding

	Rating		
	Good	Fair	Poor
Distance (Å)[a]	<3.0	<3.0	>3.0
Geometry[b] RMSD			
Tetrahedral	<15	>15	>15
Trigonal bipyramidal	<25	>25	>25
Square-based pyramidal	<25	>25	>25

[a] A distance of 3.0 Å is selected to indicate the minimum distance needed for coordination with the zinc atom. The cutoff is based on the largest distance found in the experimental structures of the 16 MMP complexes used in the initial docking experiments.

[b] RMSD is the root mean square deviation from the ideal geometry about the zinc atom. The cutoff for each geometry is based on the largest RMSD found in the experimental structures of the 16 MMP complexes used in the initial study.
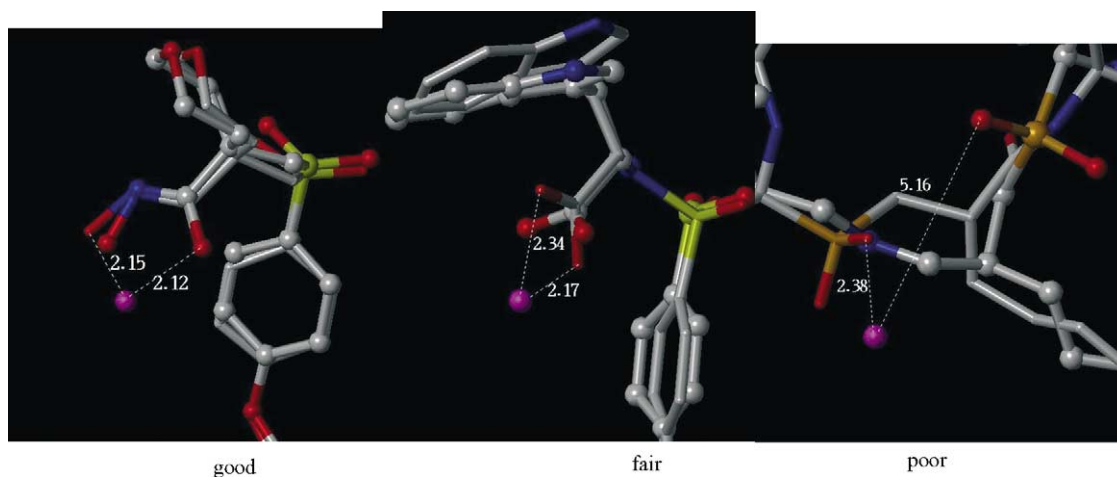
Fig. 3. Examples of docking accuracy for the zinc binding group (ZBG) for each of the categories described in the text.

3.3. Zinc parameter optimization

AutoDock uses the "non-bonded approach" [20,21] to model the zinc/ligand interactions. A suitable representation of charges is critical to the success of docking and scoring. Quantum mechanic calculations reveal that the zinc charges are close to +1.0 e instead of the formal charge of +2 e in the binding complex [43,44]. To investigate the effects of zinc charges on the docking result, we reduced the zinc charges in docking. While the zinc charges were reduced to +0.9 e, the 14 MMP complexes can still be docked properly. However, the proper dockings were less than half if the zinc charges were reduced to below +0.5 e. With no charges on the zinc, more than 70% of the dockings failed.

In addition to the charges, AutoDock requires the van der Waal (vdW) radius and well depth of zinc ion. Preliminary studies indicated that these parameters are strongly co-dependent. We therefore undertook a zinc parameter optimization including the vdW radius, well depth, and charge of zinc ion using the simplex method (MultiSimplex [45,46]), which is capable of simultaneously optimizing multiple parameters. The correlation coefficient (R^2) between predicted and experimental binding energies was used as the opti-

mization variable. The experimental binding energies derived from K_d for the inhibitor are subject to considerable error because they are from different laboratories using different conditions, but these errors are probably considerably less then the errors in the predicted values from the docking program. The diversity of the data set will help to insure a median value for the zinc parameters. The choice of using R^2 for the optimization rather than the conventional RMSD allowed a direct attempt to improve binding energy computations from the docking program. In addition, improvement in zinc parameters should lead to improvement in the zinc binding. The starting parameters were obtained by selecting three sets of reasonable values for the charge, radii, and well depth, carrying out docking studies for the 14 MMP complexes, and evaluating R^2 for each set. The MultiSimplex then computed a fourth set of parameters as the next step in the simplex operation, and these were used to dock the 14 complexes and obtain a value for R^2. After 34 trials were performed, no further significant improvement in the R^2 value was noted and the parameters of the trial with the best R^2 were selected for use in docking. The final optimized parameters were a zinc radius of 0.87 Å, a well depth of 0.35 kcal/mol, and a zinc charge of +0.95 e, giving an R^2 of 0.50. Interestingly, the optimized charge of zinc ion agrees quite well with the charge computed from sophisticated quantum computations, even though the Simplex optimization is completely unrelated to quantum calculations.

The docking results with the optimized parameter set were compared with the results using the initial parameters ($r = 1.1$ Å, $\varepsilon = 0.25$ kcal/mol, $q = +2.0$ e) in Table 2. Regression analysis showed that the standard error was reduced from 3.17 kcal/mol to 2.1 kcal/mol, while the R^2 was improved from 0.31 to 0.50. More importantly, the original parameters had 8 examples where the error in estimating the free energy of binding exceeded 5 kcal/mol, while using the optimized parameters only one example exceeded the 5 kcal/mol error. The mean absolute error of 4.52 kcal/mol was reduced to 2.58 kcal/mol using the optimized parameters. The critical
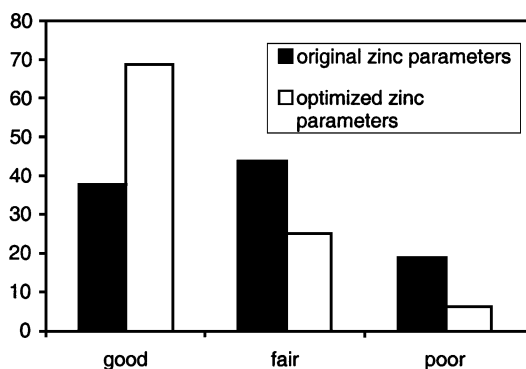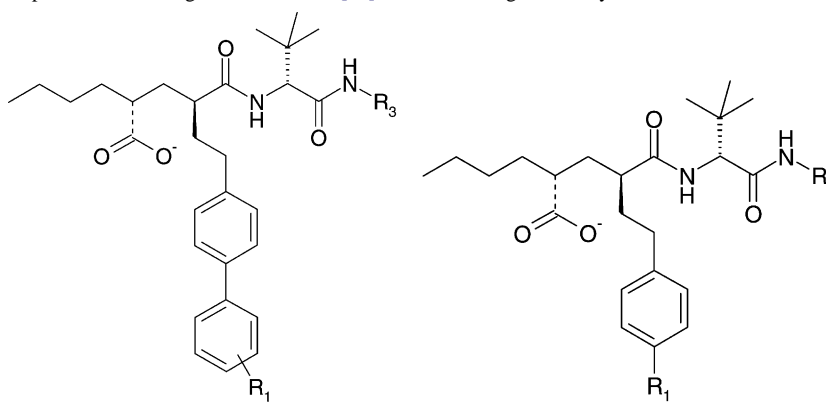

Fig. 4. Docking accuracy of the zinc binding group (ZBG) for the initial zinc parameters and the optimized zinc parameters.

Table 4
Experimental binding constants from [47] for a set of eight carboxylate MMP-3 inhibitors

Compound number	R_1	R_3	K_i (μM)	Compound number	R_1	R_3	K_i (μM)
A1	4-F	Ph	0.004	A5	CH_2Ph	Ph	1.4
A2	4-Ph	CH_3	0.039	A6	Oph	CH_3	0.2
A3	2-F	CH_3	0.011	A7	H	CH_3	3.8
A4	2,6-diCH$_3$	CH_3	1.2	A8	OCH_2Ph	CH_3	0.074

docking of the ZBG (Fig. 4) also markedly improved with nearly 70% classified as good using the optimized parameters as opposed to less than 40% with the original parameters. There was a significant reduction (from 20% to 7%) in poor dockings with the new parameters. Consequently, the optimized parameters achieved a marked improvement in binding energy computation with a simultaneous improvement of the docking geometry of the ZBG.

However, while improvement was obtained within our 14 training complexes, the two cases (1JAP and 1HV5) still failed in docking even using the optimized zinc parameters. As described above, the two failed dockings result from a potential deficiency of AutoDock scoring function, apparently outside of the zinc parameters. The use of optimized zinc parameters can improve the quality of docking, minimize the excessive energy caused by zinc metal, and thus improve the prediction of binding free energies; however, it does not eliminate all problems with the scoring function.

3.4. Parameter testing and applications in structure-based drug design

We elected to compare the original with the optimized zinc parameters with data not in the original set to show the general applicability of the optimized zinc parameters. Three different applications were used to demonstrate the diversity for which the optimized zinc parameters would perform: docking of closely related analogs similar to a lead optimization; selectivity for different enzymes; and application to other zinc enzymes. The comparison of the two parameter sets was carried with three tests: comparison of the quality of the binding geometry of the ZBG with the theoretical geometry; comparison of the absolute error in the calculated versus experimental binding energies; and com-

parison of regression analysis of the two parameter sets. The quality of the binding geometries of the ZBG reflects a proper balance between the effect of the zinc parameters and the other components of the force field. A reduced absolute error in binding energies represents an improvement in docking making the program much more useful for drug discovery. Improvements in regression analysis, increased R^2, slope close to 1.0, and decreased standard error likewise indicates more accurate binding energy calculations.

We selected eight carboxylate inhibitors [47] of MMP-3 to test the optimized zinc parameters used for AutoDock. These test compounds represent mainly a modification of the P1′ group to improve the potency of the lead compound, a typical medicinal chemistry problem (Table 4). Table 5 shows the docking results of predicted binding free energies as well as the docking accuracy at the zinc binding site. The four structures showing only fair ZBG binding geometry using the original parameters improved to demonstrating good binding geometry when the optimized zinc parameters were used. The optimized parameters allowed all complexes to show good binding geometry. The average absolute error of the predicted energies was 3.38 kcal/mol using the old zinc parameters. With the use of the optimized zinc parameters, the average error was reduced to 1.25 kcal/mol. An improved correlation between predicted binding free energy and experimental was also obtained with the optimized zinc parameters ($R^2 = 0.84$ with standard error of 0.79 kcal/mol) compared with the original parameters ($R^2 = 0.61$ with standard error of 2.7 kcal/mol). Obviously, such improvement makes the program more reliable for lead optimization.

Selectivity is desirable but difficult to achieve in the design of MMP inhibitors. We selected MMP inhibitor RS-113456 and docked to eight MMPs with both zinc parameters. RS-113456 is a sulfohydroxamate inhibitor

Table 5
Comparison of initial and optimized zinc parameters for docking of eight carboxylate inhibitors on MMP-3

Compound number	$G_{experimental}$	Initial zinc parameters ($r = 1.1$ Å, $\varepsilon = 0.25$ kcal/mol, $q = +2.0$ e)			Optimized zinc parameters ($r = 0.87$ Å, $\varepsilon = 0.35$ kcal/mol, $q = +0.95$ e)		
		$G_{predicted}$	ZBG	$G_{predicted} - G_{experimental}$	$G_{predicted}$	ZBG	$G_{predicted} - G_{experimental}$
Al	−11.46	−13.58	Good	−2.12	−12.64	Good	−1.18
A2	−10.86	−16.28	Fair	−5.42	−11.57	Good	−0.71
A3	−10.11	−15.97	Fair	−5.86	−12.32	Good	−2.21
A8	−9.73	−13.21	Good	−3.48	−10.90	Good	−1.17
A6	−9.14	−10.23	Good	−1.09	−11.23	Good	−2.09
A4	−8.08	−12.61	Fair	−4.53	−8.97	Good	−0.89
A5	−7.99	−9.51	Good	−1.52	−9.58	Good	−1.59
A7	−7.39	−10.43	Fair	−3.04	−7.24	Good	0.15

Table 6
Comparison of initial and optimized zinc parameters for docking of RS-113456 in a MMP selectivity study

MMP (PDB)	$G_{experimental}$	Initial zinc parameters ($r = 1.1$ Å, $\varepsilon = 0.25$ kcal/mol, $q = +0.95$ e)			Optimized zinc parameters ($r = 0.87$ Å, $\varepsilon = 0.35$ kcal/mol, $q = +2.0$ e)		
		$G_{predicted}$	ZBG	$G_{experimental} - G_{predicted}$	$G_{predicted}$	ZBG	$G_{predicted} - G_{experimental}$
MMP-1 (1HFC)	−9.76	−11.86	Poor	−2.1	−10.84	Poor	−1.08
MMP-2 (1QIB)	−14.01	−15.22	Fair	−1.21	−13.81	Good	0.20
MMP-3 (1HFS)	−11.30	−14.97	Good	−3.67	−12.54	Good	−1.24
MMP-7 (1MMQ)	−9.03	−13.28	Poor	−4.25	−10.05	Fair	−1.02
MMP-8 (I76)	−13.49	−18.35	Good	−4.86	−12.70	Good	0.79
MMP-9 (1GKC)	−13.90	−17.13	Good	−3.23	−15.12	Good	−1.22
MMP-12 (1JK3)	−13.40	−14.27	Fair	−0.87	−13.38	Good	0.02
MMP-13 (456C)	−13.33	−14.85	Good	−1.52	−12.54	Good	0.79

Table 7
Comparison of initial and optimized zinc parameters for 20 zinc-dependent metalloproteinases

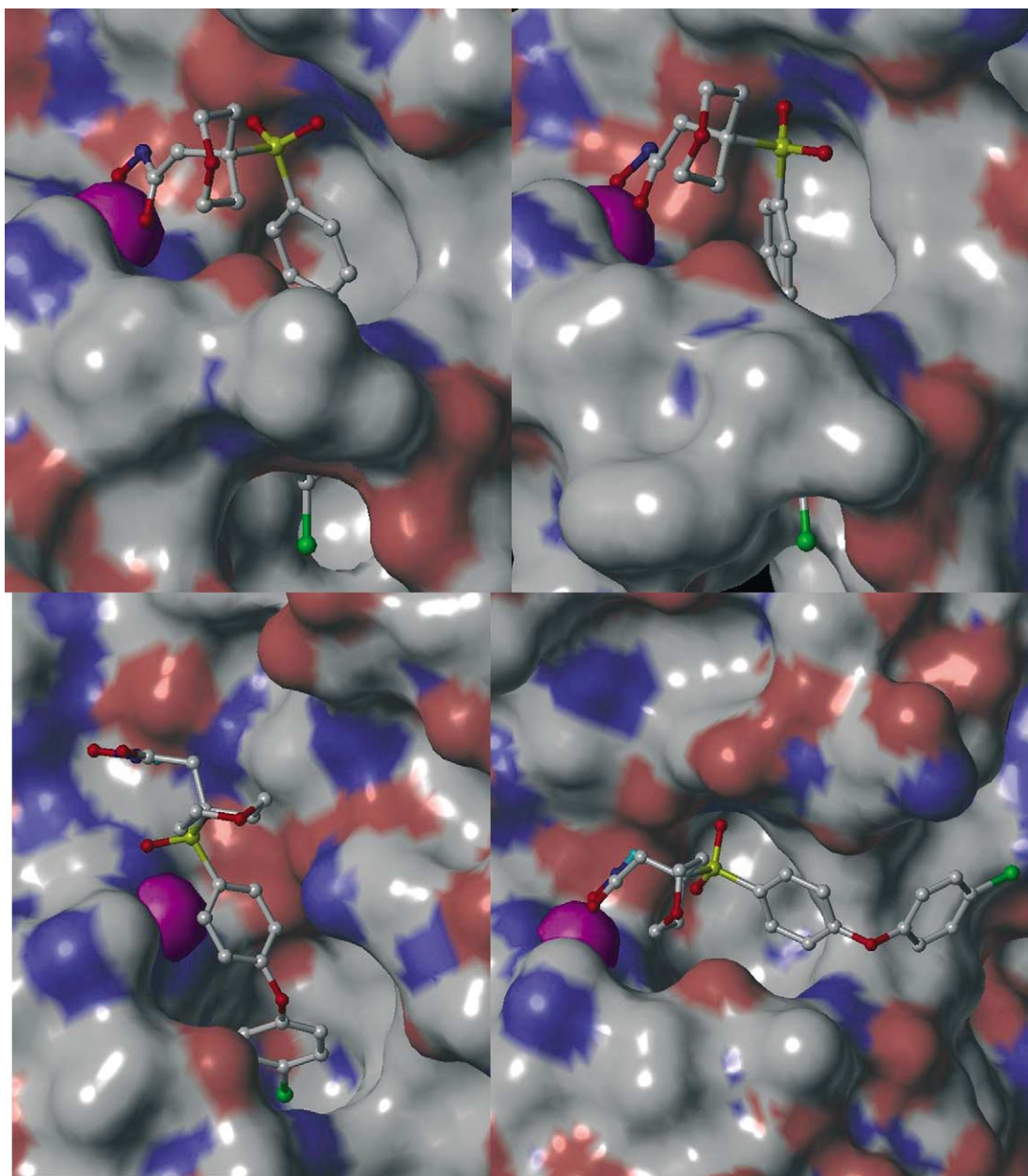| PDB	$G_{experimental}$	Initial zinc parameters ($r = 1.1$ Å, $\varepsilon = 0.25$ kcal/mol, $q = +0.95$ e)			Optimized zinc parameters ($r = 0.87$ Å, $\varepsilon = 0.35$ kcal/mol, $q = +2.0$ e)		
		$G_{predicted}$	ZBG	$G_{predicted} - G_{experimental}$	$G_{predicted}$	ZBG	$G_{predicted} - G_{experimental}$
1TLP	−10.31	−8.54	Poor	1.77	−7.85	Poor	2.46
1TMN	−10.19	−9.19	Poor	1.00	−10.31	Fair	−0.12
2TMN	−8.04	−6.31	Fair	1.73	−7.49	Good	0.55
3TMN	−8.05	−9.63	Good	−1.58	−8.79	Good	−0.74
4TMN	−13.88	−10.74	Poor	3.14	−9.28	Poor	4.60
4TLN	−5.08	−6.62	Good	−1.54	−4.94	Good	0.14
5TMN	−10.97	−6.74	Poor	4.23	−7.83	Fair	3.14
5TLN	−8.69	−7.03	Fair	1.66	−7.69	Good	1.00
6TMN	−6.89	−4.49	Poor	2.40	−6.81	Fair	0.08
7TLN	−3.37	−5.73	Fair	−2.36	−4.26	Fair	−0.89
1QF0	−10.07	−12.46	Good	−2.39	−9.21	Good	0.86
1QF1	−9.99	−12.08	Good	−2.09	−9.36	Good	0.63
1QF2	−8.08	−11.27	Good	−3.19	−8.74	Good	−0.66
1CBX	−8.60	−13.47	Poor	−4.87	−9.31	Good	−0.71
1CPS	−9.09	−12.95	Good	−3.86	−12.13	Good	−3.04
2CTC	−5.31	−7.36	Good	−2.05	−6.27	Good	−0.96
3CPA	−5.30	−8.68	Poor	−3.38	−6.94	Good	−1.64
6CPA	−15.72	−12.93	Poor	2.79	−10.24	Poor	5.48
7CPA	−19.10	−13.46	Poor	5.64	−12.18	Poor	6.92
8CPA	−12.48	−9.97	Poor	2.51	−8.83	Poor	3.65

Fig. 5. RS-113456 docked into different MMPs using AutoDock 3.0, illustrating the good fit for MMP-2 (top left) and MMP-13 (top right) with long P1′ pocket and poor fits for MMP-1 (bottom left) and MMP-7 (bottom right) with short P1′ pocket.

characterized with a long S1′ substituent. As shown in Table 6, the use of the optimized Zn parameters improved the geometry quality from 50% good fits to 75% good ZBG geometry. Furthermore, the absolute error in the predicted binding energies was reduced from 2.71 kcal/mol to 0.8 kcal/mol indicating excellent agreement with experimentally determined energies with the optimized zinc parameters. Experimentally, RS-113456 shows high affinity toward gelatinases but a relatively low activity on MMP-7 and -1. Fig. 5 shows the binding of RS-113456 to different MMPs as predicted by AutoDock. The long biphenyl group of RS-113456 docked into the S1′ pocket of MMP-2 and -13 with good zinc binding, the predicted binding mode

matched the X-ray crystal structure. However for MMP-1 and -7, the biphenyl group either inserted halfway into the S1′ or adopted a distorted geometry binding into the cleft. The docking results are consistent well with the known structural differences between the P1′ pockets in MMPs. More promisingly, the selectivity of inhibitor RS-113456 on different MMPs was predicted well.

To evaluate the generality of the optimized zinc parameters for other zinc-dependent metalloproteinases, 13 thermolysins and 7 carboxypeptidase ligand–enzyme structures were selected for docking using the two sets of zinc parameters. The results are summarized in Table 7. With the original parameters, only 35% are good and 50% of the dockings

showed poor ZBG binding. However, using the optimized zinc parameters, the docking accuracy with good ZBG binding increased to 55% and the number with poor ZBG dropped to 25%. The absolute error in the predicted binding energies dropped from 2.56 kcal/mol to l.91 kcal/mol. However, poor zinc binding was observed with five complexes with both parameters and inclusion of these complexes in the analysis is not appropriate. Exclusion of these complexes from the analysis gave a change in absolute errors from 2.36 to 1.01 and an increase in R^2 from 0.25 to 0.62, demonstrating excellent performance of the optimized zinc parameters even with different zinc enzymes. The large number of incorrectly docked structures, even with the optimized parameters, indicates the need for validation of the docking for each different enzyme system, although binding energy calculations are good for all the systems we examined.

4. Conclusions

Docking of ligands to zinc proteinases remains a challenge, due to the various polyhedral coordination geometries assumed by zinc and the lack of a force field capable of correctly predicting the behavior of ligands for these enzymes. In this study, we demonstrated optimization of the zinc parameters greatly improved the performance of the docking program in predicting the appropriate geometry of the docked complex as well as the binding free energy. The optimized zinc charge was close to the charge predicted by *ab intio* quantum calculation indicating that a more realistic zinc charge should be used in docking instead of assigning full formal charges. The charge is particularly critical for metals with multiple positive charges and computations in which electrostatic interactions are computed. Because the computation of the binding energy using AutoDock 3.0 uses a unique method of computing the energy, the use of parameters not optimized for use in the program will likely not produce optimum results. Our optimized parameters do not mean the parameters from Stote and Karplus are not optimum for the purpose for which they were derived, and our parameters should not be used for other types of computation or even other docking programs without adequate testing. The optimized zinc parameters improve the use of AutoDock for lead optimization or prediction of selectivity for MMPs, may show potential for the docking of other zinc metalloproteinases.

Acknowledgements

References

[1] J.F. Woessner Jr., Matrix metalloproteinases and their inhibitors in connective tissue remodeling, FASEB J. 5 (1991) 2145–2154.

[2] A.E. Yu, R.E. Hewitt, E.W. Connor, W.G. Stetler-Stevenson, Matrix metalloproteinases: novel targets for directed cancer therapy, Drugs Aging 11 (1997) 229–244.

[3] P.E. Gottschall, S. Deb, Regulation of matrix metalloproteinase expression in astrocytes, microglia and neurons, Neuroimmuno-modulation 3 (1996) 69–75.

[4] N.J. Clendeninn, K. Appelt, Matrix Metalloproteinase Inhibitors in Cancer Therapy, Humana Press, Totowa, NJ, 2001.

[5] A. Maeda, R.A. Sobel, Matrix metalloproteinases in the normal human nervous system, microglial nodules, and multiple sclerosis lesions, J. Neuropathol. Exp. Neurol. 55 (1996) 300–309.

[6] R.P. Becket, A.H. Davidson, A.H. Drummond, P. Huxley, M. Whittaker, Recent advances in matrix metalloproteinase inhibitor research, Drug Discuss. Today 1 (1996) 16–26.

[7] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.

[8] R.E. Babine, S.L. Bender, Molecular recognition of protein–ligand complexes: applications to drug design, Chem. Rev. 97 (1997) 1359–1472.

[9] M. Whittaker, C.D. Floyd, P. Brown, A.J.H. Gearing, Design and therapeutic application of matrix metalloproteinase inhibitors, Chem. Rev. 99 (1999) 2735–2776.

[10] L.M. Coussens, B. Fingleton, L.M. Matrisian, Matrix metalloproteinase inhibitors and cancer: trials and tribulations, Science 295 (2002) 2387–2390.

[11] E.J. Jacobsen, M.A. Mitchell, S.K. Hendges, Synthesis of a series of stromelysin-selective thiadazole urea matrix metalloproteinase inhibitors, J. Med. Chem. 42 (1999) 1525–1536.

[12] S. Toba, K.V. Damodaran, K.M. Merz Jr., Binding preference of hydroxamate inhibitors of the matrix metalloproteinase human fibroblast collagenase, J. Med. Chem. 42 (1999) 1225–1234.

[13] H. Matter, W. Schwab, D. Barbier, G. Billen, B. Haase, B. Neises, M. Schudok, W. Thorwart, H. Schreuder, V. Brachvogel, P. Lonze, K.U. Weithmann, Quantitative structure–activity relationship of human neutrophil collagenase (MMP-8) inhibitors using comparative molecular field analysis and X-ray structure analysis, J. Med. Chem. 42 (1999) 1908–1920.

[14] R. Kiyama, Y. Tamura, F. Watanabe, H. Tsuzuki, M. Ohtani, M. Yodo, Homology modeling of gelatinase catalytic domains and docking simulations of novel sulfonamide inhibitors, J. Med. Chem. 42 (1999) 1723–1738.

[15] E.A. Amin, W.J. Welsh, Three-dimensional quantitative structure–activity relationship (3D-QSAR) models for a novel class of piperazine-based stromelysin-1 (MMP-3) inhibitors: applying a "divide and conquer" strategy, J. Med. Chem. 44 (2001) 3849–3855.

[16] J.M. Chen, F.C. Nelson, J.I. Levin, D. Mobiho, F.J. Moy, R. Nilakantan, A. Zask, R. Powers, Structure-based design of a novel, potent, and selective inhibitor for MMP-13 utilizing NMR spectroscopy and computer-aided molecular design, J. Am. Chem. Soc. 122 (2000) 9648–9654.

[17] I.D. Kuntz, Structure-based strategies for drug design and discovery, Science 257 (1992) 1078–1082.

[18] J. Drews, Drug discovery: a historical perspective, Science 287 (2000) 1960–1964.

[19] B. Cox, J.C. Denyer, A. Binnie, M.C. Donnelly, B. Evans, D.V. Green, J.A. Lewis, T.H. Mander, A.T. Merritt, M.J. Valler, S.P. Watson, Application of high-throughput screening techniques to drug discovery, Prog. Med. Chem. 37 (2000) 83–133.

[20] R.H. Stote, M. Karplus, Zinc binding in proteins and solution: a simple but accurate nonbonded representation, Proteins 23 (1995) 12–31.

[21] S.C. Hoops, K.W. Anderson, K.M. Merz Jr., Force field design for metalloproteins, J. Am. Chem. Soc. 113 (1991) 8262–8270.

[22] I.L. Alberts, K. Nadassy, S.J. Wodak, Analysis of zinc binding sites in protein crystal structures, Protein Sci. 7 (1998) 1700–1716.

[23] O.V. Buzko, A.C. Bishop, K.M. Shokat, Modified AutoDock for accurate docking of protein kinase inhibitors, J. Comput. Aided Mol. Des. 16 (2002) 113–127.

[24] E. Perola, K. Xu, T.M. Kollmeyer, S.H. Kaufmann, F.G. Prendergast, Y.-P. Pang, Successful virtual screening of a chemical database for farnesyltransferase inhibitor leads, J. Med. Chem. 43 (2000) 401–408.

[25] Y.-P. Pang, E. Perola, K. Xu, F.G. Prendergast, EUDOC: a computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases, J. Comput. Chem. 22 (2001) 1750–1771.

[26] M.L. Verdonk, J.C. Cole, R. Taylor, SuperStar: a knowledge-based approach for identifying interaction sites in proteins, J. Mol. Biol. 289 (1999) 1093–1108.

[27] D.S. Goodsell, A.J. Olson, Automated docking of substrates to proteins by simulated annealing, Proteins 8 (1990) 195–202.

[28] G.M. Morris, D.S. Goodsell, R. Huey, A.J. Olson, Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4, J. Comput. Aided Mol. Des. 10 (1996) 293–304.

[29] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a lamarckian genetic algorithm and empirical binding free energy function, J. Comput. Chem. 19 (1998) 1639–1662.

[30] B. Lovejoy, A.R. Welch, S. Carr, C. Luong, C. Broka, R.T. Hendricks, J.A. Campbell, K.A. Walker, R. Martin, H. Van Wart, M.F. Browner, Crystal structures of mmp-1 and -13 reveal the structural basis for selectivity of collagenase inhibitors, Nat. Struct. Biol. 6 (1999) 217–221.

[31] A.G. Pavlovsky, M.G. Williams, Q.Z. Ye, D.F. Ortwine, C.F. Purchase II, A.D. White, V. Dhanaraj, B.D. Roth, L.L. Johnson, D. Hupe, C. Humblet, T.L. Blundell, X-ray structure of human stromelysin catalytic domain complexed with non-peptide inhibitors: implications for inhibitor selectivity, Protein Sci. 8 (1999) 1455–1462.

[32] C.K. Esser, R.L. Bugianesi, C.G. Caldwell, K.T. Chapman, P.L. Durette, N.N. Girotra, I.E. Kopka, T.J. Lanza, D.A. Levorse, M. MacCoss, K.A. Owens, M.M. Ponpipom, J.P. Simeone, R.K. Harrison, L. Niedzwiecki, J.W. Becker, A.I. Marcy, M.G. Axel, A.J. Christen, J. McDonnell, V.L. Moore, J.M. Olszewski, C. Saphos, D.M. Visco, F. Shen, A. Colletti, P.A. Krieter, W.K. Hagmann, Inhibition of stromelysin-1 (MMP-3) by P1′-biphenylylethyl carboxyalkyl dipeptides, J. Med. Chem. 40 (1997) 1026–1040.

[33] M.F. Browner, W.W. Smith, A.L. Castelhano, Matrilysin-inhibitor complexes: common themes among metalloproteases, Biochemistry 34 (1995) 6602–6610.

[34] H. Brandstetter, R.A. Engh, E.G. Von Roedern, L. Moroder, R. Huber, W. Bode, F. Grams, Structure of malonic acid-based inhibitors bound to human neutrophil collagenase. A new binding mode explains apparently anomalous data, Protein Sci. 7 (1998) 1303–1309.

[35] F. Grams, P. Reinemer, J.C. Powers, T. Kleine, M. Pieper, H. Tschesche, R. Huber, W. Bode, X-ray structures of human neutrophil collagenase complexed with peptide hydroxamate and peptide thiol inhibitors. Implications for substrate binding and rational drug design, Eur. J. Biochem. 228 (1995) 830–841.

[36] W. Bode, P. Reinemer, R. Huber, T. Kieme, S. Schnierer, H. Tschesche, The X-ray crystal structure of the catalytic domain of human neutrophil collagenase inhibited by a substrate analogue reveals the essentials for catalysis and specificity, EMBO J. 13 (1994) 1263–1269.

[37] A.L. Gall, M. Ruff, R. Kannan, P. Cuniasse, A. Yiotakis, V. Dive, M.C. Rio, P. Basset, D. Moras, Crystal structure of the stromelysin-3 (mmp-11) catalytic domain complexed with a phosphinic inhibitor mimicking the transition-state, J. Mol. Biol. 307 (2001) 577–586.

[38] J.C. Spurlino, A.M. Smallwood, D.D. Carlton, T.M. Banks, K.J. Vavra, J.S. Johnson, E.R. Cook, J. Falvo, R.C. Wahl, T.A. Pulvino, 1.56 Å structure of mature truncated human fibroblast collagenase, Proteins 19 (1994) 98–109.

[39] SYBYL Molecular Modeling Software, v. 6.8, Tripos Associates, St. Louis, MO.

[40] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity. A rapid access to atomic charges, Tetrahedron 36 (1980) 3219–3228.

[41] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, A fast flexible docking method using an incremental construction algorithm, J. Mol. Biol. 261 (1996) 470–489.

[42] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ohio, G. Alagona, S. Profeta Jr., P. Weiner, A new force field for molecular mechanical simulation of nucleic acids and proteins, J. Am. Chem. Soc. 106 (1984) 765–784.

[43] U. Ryde, Molecular dynamics simulations of alcohol dehydrogenase with a four- or five-coordinate catalytic zinc ion, Proteins 21 (1995) 40–56.

[44] O. Kleifeld, P.E. Van den Steen, A. Frenkel, F. Cheng, H.L. Jiang, G. Opdenakker, I. Sagi, Structural characterization of the catalytic active site in the latent and active natural gelatinase B from human neutrophils, J. Biol. Chem. 275 (2000) 34335–34343.

[45] J.A. Nelder, R. Mead, A simplex method for function minimization, Comput. J. 7 (1965) 308–313.

[46] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, In Numerical Recipes in Fortran: The Art of Scientific Computing, second ed., Cambridge University Press, Cambridge, 1992, p. 402.

[47] C.K. Esser, R.L. Bugiansi, C.G. Caldwell, K.T. Chapman, P.L. Durette, N.N. Girotra, I.E. Kopka, T.J. Lanza, D.A. Levorse, M. MacCoss, K.A. Owens, M.M. Ponpipom, J.P. Simeone, R.K. Harrison, L. Niedziecki, J.W. Becker, A.I. Marcy, M.G. Axel, A.J. Christen, J. McDonnell, V.L. Moore, J.M. Olszewski, C. Saphos, D.M. Visco, F. Shen, A. Colletti, P.A. Krieter, W.K. Hagmann, Inhibition of stromelysin-1 (MMP-3) by P1′-biphenylethyl carboxyalkyl dipeptides, J. Med. Chem. 40 (1997) 1026–1040.