Quantum and Molecular Dynamics Study for Binding of Macrocyclic Inhibitors to Human α-Thrombin

Emilia L. Wu,* Ye Mei,† KeLi Han,* and John Z. H. Zhang†‡

*Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China; †Institute of Theoretical and Computational Chemistry, Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, China; and ‡Department of Chemistry, New York University, New York, New York

ABSTRACT Molecular dynamics simulations followed by quantum mechanical calculation and Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) analysis have been carried out to study binding of proline- and pyrazinone-based macrocyclic inhibitors (L86 and T76) to human α-thrombin. Detailed binding interaction energies between these inhibitors and individual protein fragments are calculated using DFT method based on a new quantum mechanical approach for computing protein-ligand interaction energy. The analysis of detailed interaction energies provides insight on the protein-ligand binding mechanism. Study shows that T76 and L86 bind to thrombin in a very similar "inhibition mode" except that T76 has relatively weaker binding interaction with Glu^{217}. The analysis from quantum calculation of binding interaction is consistent with the MM-PBSA calculation of binding free energy, and the calculated free energies for L86/T76-thrombin binding agree well with the experimental data.

INTRODUCTION

Human α-thrombin is a multifunctional member of the trypsin family of serine proteases, and it is the final enzyme of the blood coagulation cascade. The activity of thrombin in this cascade is responsible for the cleavage of fibrinogen to form fibrin and the activation of platelets via the thrombin receptor (1–3). Fibrin then polymerizes to form a network of fibers, entrapping the platelets and leading to clot formation and cessation of blood flow. Thrombin-induced clot formation is a necessary part of the wound-healing process, but it is also associated with many disease states including myocardial infarction, pulmonary embolism, and stroke (4). While anticoagulant therapeutics have traditionally been available in the form of heparin, which is administered by injection and the orally active coumadin, both of these treatments have significant limitations in safety and efficacy (5). An injectable form of the thrombin inhibitor argatroban has recently been approved by the FDA, but only for the relatively rare condition of heparin-induced thrombocytopenia (6). The direct inhibition of thrombin by an orally available small-molecule inhibitor is therefore still a high priority in medicinal chemistry research. The inhibition of any of the closely related serine proteases in the coagulation cascade might be therapeutically effective, and eventually it may be possible to selectively inhibit these proteases for optimal treatment of specific conditions. Thrombin has also been implicated in some events at the cellular level such as the activation of blood platelets and in a number of other processes (7). However, at present, the ease with which thrombin can be purified and studied in soluble form, along with the wealth of structural information available and its remarkable variety of functions in hemostasis (8), has made it the most intensely studied target for anticoagulation therapy (9–11).

Crystallographic investigations (12) have revealed that thrombin is composed of disulfide-linked A and B chains. It also has an anion binding exosite positioned ~20 Å from the active site along a groove. The active site of thrombin is mainly defined by the specificity (S1) pocket, the hydrophobic proximal (S2) pocket, and the hydrophobic distal (S3) pocket. A schematic representation of the active site is shown in Fig. 1 (13). The specificity pocket consists of a hydrophobic channel with the carboxylic acid of Asp^{189} (14) and two backbone carbonyls in the bottom of the pocket, and the former forms strong ionic interactions with amine-, guanidine-, or amidine-type structures located at the terminus of a hydrophobic spacer. The proximal pocket is defined on three sides by the $Tyr^{60}A$ and $Trp^{60}D$ side chains of the 60-insertion loop, the imidazole ring of His^{57}, and the isobutyl group of Leu^{99} in the enzyme. The larger distal pocket is mainly made up of the side chains of Trp^{215}, Ile^{174}, and Leu^{99}. Other important interactions with potential inhibitors include hydrogen bonding to the β-sheet segment from Ser^{214}-Trp^{215}-Gly^{216} (15). The catalytic triad of Asp^{102}, His^{57}, and Ser^{195} is responsible for the proteolytic activity.

To this end, numerous theoretical studies have been performed to estimate the activities of a series of thrombin inhibitors. The methods used in these studies include correlation of binding affinity with interaction energy (16), active site mapping (17), comparative molecular field analysis (18), and linear responses calculations (12). All of these analyses have their merits as well as weaknesses. A complete understanding of the binding mechanism of the thrombin inhibitor complexes is essential for discovery of new and

Submitted October 10, 2006, and accepted for publication February 12, 2007.

Address reprint requests to J. Z. H. Zhang, E-mail: john.zhang@nyu.edu; or K. Han, E-mail: klhan@dicp.ac.cn.

0006-3495/07/06/4244/10 $2.00

doi: 10.1529/biophysj.106.099150

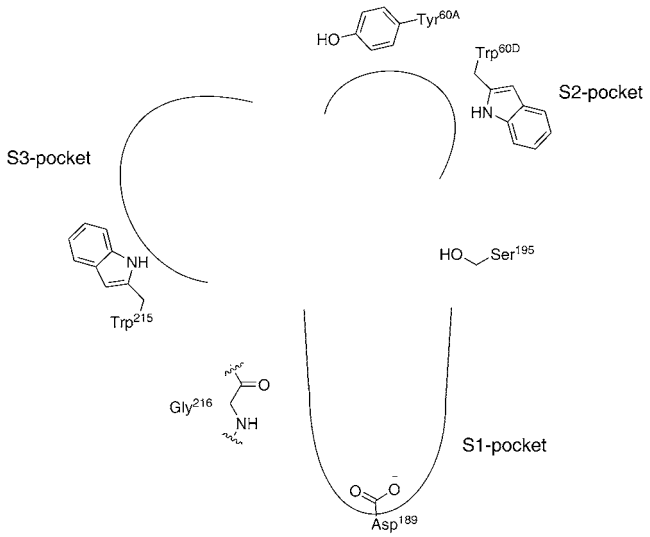

FIGURE 1 The schematic representation of the active site of thrombin.

improved inhibitors. Due to large size of the enzyme, computational studies of thrombin-inhibitor binding are primarily based on molecular mechanics force-field approaches. Nonetheless, a quantitative analysis of binding interaction of the thrombin-inhibitor complex at quantum mechanical level is highly desirable. To aid in the structure-based (or rational) design of new thrombin inhibitors, we carry out systematic quantum mechanical studies together with Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) free energy calculations to gain molecular insight into the binding mechanism of two thrombin inhibitors: the proline- and pyrazinone-base macrocyclic inhibitors (19–21). The full quantum mechanical computation of thrombin-inhibitor binding interaction energy presented in this study is made possible by applying a recently developed efficient molecular fractionation with conjugate caps (MFCC) approach (22,23). A full quantum study of binding of nevirapine and efavirenz to HIV-1 RT using the MFCC approach was recently reported (24,25). In the MFCC approach, the entire protein is decomposed into amino-acid fragments that are properly capped and the protein-ligand interaction energy can be calculated at various levels of quantum mechanical theory (23). In particular, the method scales linearly, is computationally efficient, and is particularly suitable for computation on multiprocessor computer systems.

THEORY AND METHODS

The MFCC method

The MFCC approach is developed to compute quantum mechanical properties of biological molecules such as protein-ligand interaction (23–27). Below we give a brief description of the MFCC method as applied to the computation of protein-ligand interaction energy. The main idea of the MFCC approach is to divide a protein molecule into amino-acid fragments that are properly capped (23,26). Using the fragmentation scheme, the interaction energy between the protein and another molecule (ligand) can be computed by separate calculations of individual fragments interacting with the ligand. A crucial feature of the MFCC approach is that a pair of conjugate caps (or concaps) is inserted at the location of cut. These caps are introduced to serve two purposes:

1. They cap the cutoff fragments to satisfy the valency requirement.
2. They mimic the local chemical environment of the original protein to the capped fragments.

Fig. 2 illustrates the MFCC scheme in which a peptide bond is cut and the fragments are capped.

By cutting the peptide bond of the protein into amino-acid fragments and inserting a pair of concaps, CH_3CO-$NHCH_3$, at the location of cut to cap the fragments, the interaction energy for protein-ligand binding system ($E_{\text{P-L}}$) is given by the expression (23–27)

$$E_{\text{P-L}} = \sum_{i=1}^{N} E_{\text{F}_i,\text{L}} - \sum_{i=1}^{N-1} E_{\text{CC}_i-\text{L}} - \sum_{i=1}^{N} E_{\text{F}_i} + \sum_{i=1}^{N-1} E_{\text{CC}_i} - \sum_{i=1}^{Nd} E_{\text{DC}_i-\text{L}} + \sum_{i=1}^{Nd} E_{\text{DC}_i} + (Nd-1)E_{\text{L}}, \quad (1)$$

where $E_{\text{F}_i,\text{L}}$ denotes the ith capped fragment-ligand energy, $E_{\text{CC}_i-\text{L}}$ the ith concap-ligand energy, E_{F_i} and E_{CC_i} are, respectively, the self-energy of the capped ith fragment and ith concap, and E_{L} is the ligand self-energy. For a protein with N amino acids in a single chain, there are N-1 concaps needed. $E_{\text{DC-L}}$ denotes the disulfide concap-ligand energy as introduced in Chen et al. (27), and N_{d} is the number of disulfide bonds (27).

In this study, we employ the MFCC method to compute binding interaction of two thrombin inhibitors (19–21). The structures of the two inhibitors are shown in Fig. 3, and we also show the orientations of the two inhibitors in the active site in the same figure. Using the MFCC approach, we first decompose the 276 amino acid thrombin into 276 fragments by cutting all the backbone peptide (C-N) bonds, and there are also 275 concaps formed by the fusion of pairs

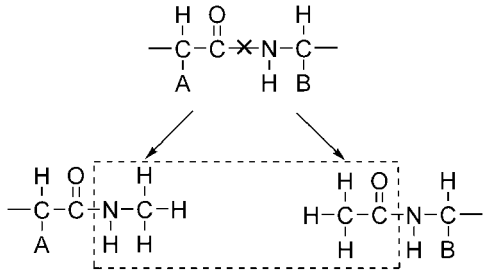

FIGURE 2 The MFCC scheme in which a peptide bond is cut and the fragments are capped.

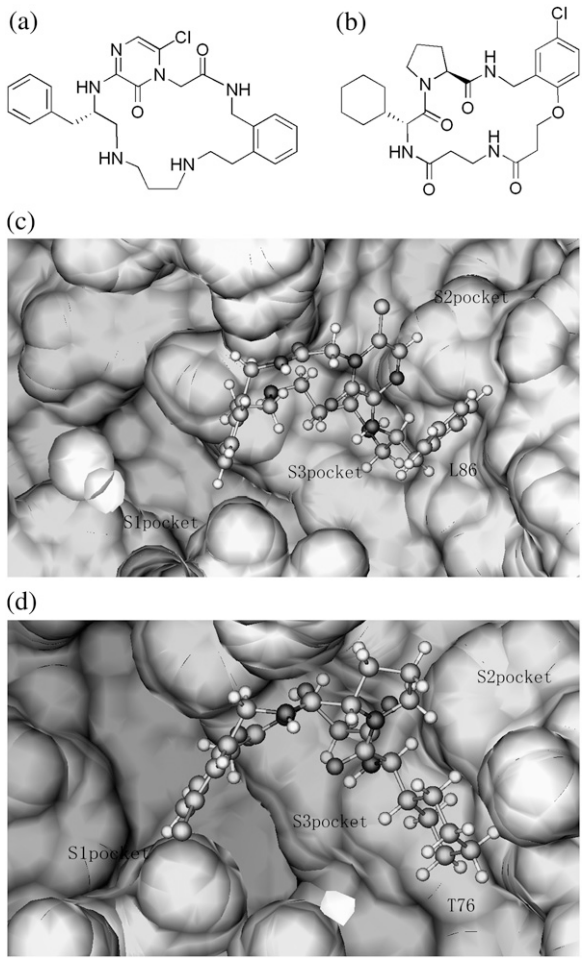

FIGURE 3 Molecular structures of the inhibitors: (*a*) L86, (*b*) T76, and the orientations in the active site of the two inhibitors (*c*) L86 and (*d*) T76. L86 and T76 are displayed in a ball-and-stick representation, and thrombin is displayed in a surface representation.

of conjugate caps. In addition, there are four disulfide bonds that are also cut and capped with the SCH_3 group (27).

It is worthwhile to mention that the individual fragment-inhibitor interaction energy from the MFCC calculation is then decapped to remove the interaction due to the cap component of the fragment before the interaction spectrum. Most cap components have negligible interactions with the inhibitor; only those cap components associated with strong binding fragments can have appreciable interaction with the inhibitors. These extra interactions are largely of hydrogen bonding in nature and are associated with the CO- and NH-groups in concaps. Although, the MFCC scheme essentially cancels these ''extra'' interaction energies to give the correct total binding energy, it is desirable to remove these extra energies from the individual fragment energies to show the interaction spectrum more clearly. The procedure to remove these cap effects was described in the Appendix of He et al. (24).

MOLECULAR DYNAMICS SIMULATION

Before the calculations of protein-ligand interaction energy using MFCC and free energy calculation, a 5-ns molecular dynamics simulation is performed for these two thrombin-inhibitor systems to relax the structure and investigate the dynamics of the binding complex using AMBER8 (28) para99 (29). Two PDB files are used as initial thrombin-inhibitor complex structures for molecular dynamics (MD) simulation: 1NM6.pdb (19) for THROMBIN-L86 and 1NT1.pdb (19) for THROMBIN-T76. Hydrogen atoms are added and their positions are optimized. The crystallographic water molecules in the PDB files are discarded and counterions are added to maintain the electroneutrality of the system. This starting structure was then placed in a truncated octahedral periodic box of TIP3P water molecules. The distance between the edges of the water box and the closest atom of the solutes is at least 8 Å. The system was minimized by steepest descent followed by conjugate gradient minimization. The particle mesh Ewald method (30) is used to treat long-range electrostatic interactions in a periodic boundary condition, and bond lengths involving bonds to hydrogen atoms are constrained using SHAKE (31). The time-step for all MD simulations is 2 fs, with a direct-space, nonbonded cutoff of 10 Å. Applying position restraints with a force constant to all solute atoms and using the Langevin dynamics to control the temperature with a collision frequency of 1.0 ps^{-1}, 20 ps MD, is carried out at constant volume, during which the system is heated from 0 to 300 K. Subsequent isothermal isobaric MD simulation is used for 5 ns to adjust the solvent density without any restraints on all the solute atoms. Finally, conformations are collected every 1 ps for the last 100-ps simulation, and 100 snapshots are collected for the gas-phase binding energy calculations (32).

We choose the lowest energy structures obtained from 5-ns MD simulation as the binding structure for MFCC calculation of thrombin-inhibitor binding energy. The residues in 1NM6 and 1NT1 are labeled from 1 to 276, according to the respective PDB files. The interaction energies between inhibitors and missing residues are set to zero. Because the missing residues in the crystallographic structure are all far away from the binding site, we just ignored their interactions with the inhibitors. Considering the computational cost, the MFCC calculation is performed for all the fragment-inhibitor interaction pairs at B3LYP/6-31G* level for the thrombin-inhibitor complex using the Gaussian03 package (33). The MFCC calculation at B3LYP/6-31G* level generates a L86-THROMBIN interaction spectrum showing binding interactions with individual residues as shown in Fig. 4. The overwhelming majority of enzyme fragments have negligible interaction with L86, as shown in Fig. 4, and the dominant binding residues are Ser^{214}, Trp^{215}, Gly^{216}, Asp^{102}... in order of reducing strength.

Table 1 lists the energies of the dominant enzyme fragment-inhibitor interaction pairs for the two complex structures,

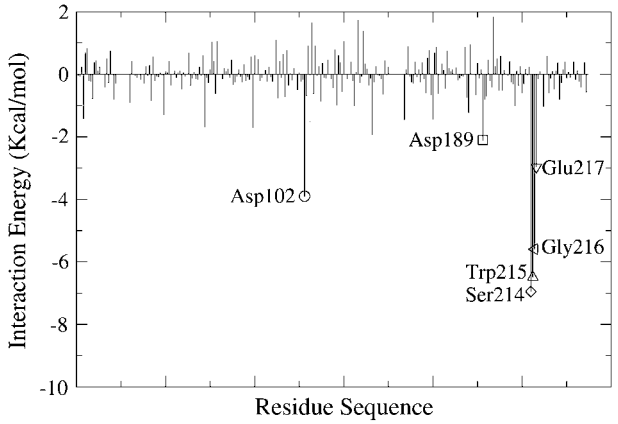

FIGURE 4 MFCC computed interaction energy spectrum (showing interaction between the inhibitor and individual protein residues) for L86/thrombin complex at B3LYP/6-31G* level.

which shows the binding mode of the inhibitors to thrombin and will be discussed in detail later in the article.

THE MM-PBSA METHOD

Although free energy perturbation and thermodynamic integration calculations should produce more accurate binding free energies, they are extremely time-consuming and require sufficient statistical sampling. The heavy computational cost prevents free energy perturbation and thermodynamic integration from being routinely used for free energy calculation in structure-based drug design (34–36). In this work, the binding free energy is calculated using MM-PBSA (37) and normal mode analysis. We chose 100 snapshots evenly from the last 100 ps MD simulation and calculate the binding free energies between the two inhibitors and thrombin with MM-PBSA module in AMBER8. In MM-PBSA, the free energy of $A + B \rightarrow AB$ is calculated through the thermodynamic cycle[36] as shown in Fig. 5. The absolute binding free energy in condensed phase can be calculated according to the equation

$$\begin{aligned}\Delta G_{\text{binding}} &= \Delta G_{\text{gas}} - \Delta G^{\text{A}}_{\text{solv}} - \Delta G^{\text{B}}_{\text{solv}} + \Delta G^{\text{AB}}_{\text{solv}} \\ &= \Delta H_{\text{gas}} - T\Delta S - \Delta G^{\text{A}}_{\text{PBSA}} - \Delta G^{\text{B}}_{\text{PBSA}} + \Delta G^{\text{AB}}_{\text{solv}} \\ &= \Delta H_{\text{gas}} - T\Delta S + \Delta\Delta G_{\text{PB}} + \Delta\Delta G_{\text{SA}}, \end{aligned} \quad (2)$$

where the definition of various energy terms are clear from Fig. 5. ΔH_{gas} in Eq. 2 is the total molecular mechanics energy in the gas phase, and it is

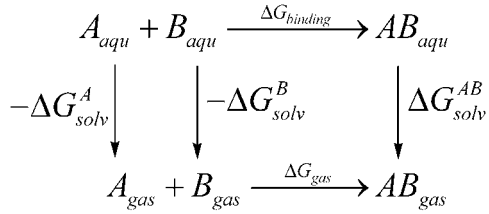

FIGURE 5 Thermodynamic cycle for absolute binding free energy calculation. $\Delta G^{\text{A}}_{\text{solv}}$, $\Delta G^{\text{B}}_{\text{solv}}$, and $\Delta G^{\text{AB}}_{\text{solv}}$ are solvation free energy of *A*, *B*, and *AB*, respectively. ΔG_{gas} and $\Delta G_{\text{binding}}$ are binding free energy in gas phase and condensed phase, respectively.

calculated for the unsolvated molecule using the standard AMBER force field with the Sander module of the AMBER program. The electrostatic contribution to the free energy ($\Delta\Delta G_{\text{PB}}$) is calculated using a finite difference solution to the Poisson-Boltzmann (PB) equation, which is solved by DELPHI program (38). The hydrophobic contribution to the solvation free energy ($\Delta\Delta G_{\text{SA}}$) is determined from the equation

$$\Delta G_{\text{nonpolar}} = \gamma SA + b, \quad (3)$$

where *SA* is the solvent-accessible surface area that is determined with Paul Beroza's Molsurf program, which is based on the analytical ideas primarily developed by Mike Connolly (39). The values γ and b are the empirical constants used in this work. We also used 0.00542 kcal mol^{-1} Å^{-2} and 0.92 kcal mol^{-1}, respectively, as is standard in the MM-PBSA work that has been published (40). AMBER99 charge parameters are used again for thrombin atoms and radius parameters are those taken from the PARSE parameter set (41). Atomic partial charge parameters for inhibitors are scaled to reproduce the solvation free energy calculated by PCM method.

Finally, entropy contributions arising form changes in the degrees of freedom (translational, rotational, and vibrational) of the solute molecules are included applying classical statistical thermodynamics (32,42). Since the normal mode calculation of entropy is extremely time-consuming for large systems, only 20 snapshots (every fifth snapshot of the 100 snapshots) for each inhibitor are used to estimate the contribution of the entropies to lower the computational time. Contributions to the vibrational entropy are obtained by normal-mode analysis. After minimization of each snapshot in the gas-phase using the conjugated gradient method with a distance-dependent dielectric of $4r$ (with r being the distance between two atoms) until the root mean-square of the elements of the gradient vector is $<10^{-4}$ kcal mol^{-1} Å^{-1}, frequencies of the vibrational modes are computed at 300 K for these minimized structures using an harmonic approximation of the energies. The nmode module of the AMBER8 package is used to perform this part of the calculation (43).

TABLE 1 Interaction energies (kcal/mol) between inhibitors and selected amino-acid fragments of thrombin from ab initio calculations at B3LYP/6-31G* level

Inhibitor/Residue	Asp^{102}	Glu^{146}	Asp^{189}	Ser^{214}	cap^{214}	Trp^{215}	cap^{215}	Gly^{216}	cap^{216}	Glu^{217}
L86	−3.9	—‡	−2.1	−6.9	−9.7	−6.5 (−1.8¶)	−5.8	−5.6	−6.5	−3.0
T76	−2.7	−2.5	−5.7	−5.4	−7.1	−3.6 (3.6¶)	—	−4.7	−6.9	—
L86_MP2*	−4.8	—	−2.7	−11.2	NA§	−10.2	NA	−14.8	NA	−5.6
T76_MP2*	−5.7	0.83	−8.5	−10.8	NA	−6.2	NA	−10.6	NA	—
L86_protonated†	−32.7	—	−26.9	−10.6	NA	−14.5	NA	−12.5	NA	−39.5

These values are calculated at MP2/6-31+G level.

†The interaction energies between protonated L86 and thrombin at B3LYP/6-31G* level.

‡The dash means there is no significant interaction between Glu^{146} and L86, and other dashes in the table also mean there are no strong interactions between the corresponding inhibitor and residue.

§We do not calculate the interaction between inhibitors and the caps with the level MP2/6-31+G*, because of the huge computation cost.

¶The value in the parentheses is calculated with small concap CH_3-CH_3.

RESULTS AND ANALYSIS

L86 binding to thrombin

L86 and T76 are proline- and pyrazinone-based small molecules, which inhibit thrombin with a high degree of potency and selectivity. They are designed by linking the P1 and P3 groups and are synthesized by Merck Research Laboratories (West Point, PA) (19). There are two amines in L86, and as with many diamines, one of the two amines is usually protonated in aqueous solution. In this study, however, the neutral form of the inhibitor is used for the calculation. One reason is that the active site of thrombin is mostly hydrophobic, and L86 is inside the cavity free of water. Another reason is that reliable solvation energy of a charged ligand is difficult to calculate at present.

Fig. 4 plots the interaction spectrum generated from the MFCC calculation at B3LYP/6-31G* level for L86/THROMBIN complex, and it shows that the main binding attractions come from approximately six residues with individual gas-phase binding energies >2 kcal/mol. Fig. 6 plots the relative position of L86 in the binding complex with the residues to which it has strong interactions. As can be seen from the two figures, the dominant binding interactions between the L86 and thrombin are the bindings to Ser^{214}, Trp^{215}, Gly^{216}, Glu^{217}, Asp^{102}, and Asp^{189} residue (the numbering of the thrombin residues used here is based on the topological equivalence with chymotrypsin as described by (14)). The amide NH forms an H-bond with the backbone carbonyl oxygen of Ser^{214}, and the amide NH from the pyrazinone and pyrazinone oxygen form two H-bonds with Gly^{216}.

According to Fig. 6, the distance between the corresponding oxygen atom and nitrogen atom is 3.01 Å, 3.12 Å, and 3.15 Å, and thus these groups have quite favorable geometries for hydrogen bonding interactions with L86. Fig. 4 shows strong interaction between the inhibitor and Trp^{215}. Trp^{215} is in a very special situation, in that the backbone atoms of Ser^{214} and Gly^{216}, which are on the sides of Trp^{215}, both form strong H-bonds with L86; so we checked this energy with the smaller concap CH3-CH3 instead of the larger cap CH3CO-NHCH3. In the test calculation, this interaction energy between L86 and Trp^{215} is substantially reduced, and the data are shown in the parentheses in Table 1. Thus the strong interaction energy in the B3LYP/6-31G* level is due to the use of the larger concap. Because the concap coordinates of Trp^{215} adopt the coordinates of the backbone amide nitrogen of Gly^{216} and the backbone carbonyl oxygen of Ser^{214}, the calculated binding interaction of L86 to Trp^{215} contains the effect of two extra hydrogen bonds, which are from the concaps and should be removed.

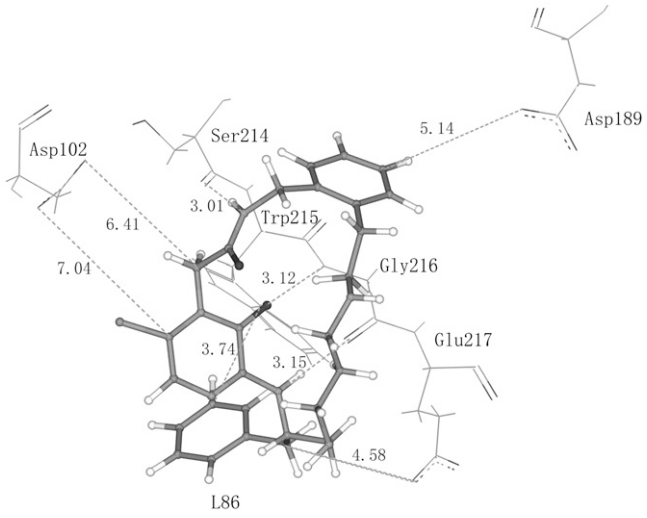

FIGURE 6 Relative geometries and distances of L86 and relevant residues of thrombin in L86/thrombin binding complex.

L86 also has significant ion-dipole attractive interactions with Glu^{217}, Asp^{189}, and Asp^{102}, which is one of the catalytic triad, as shown in Fig. 6. (L86 is a polar molecule, and its permanent dipole is ~5.6 Debye.) Because the carbon atoms that interact with Asp^{102} are directly bonded to chlorine and nitrogen atoms, there are more positive charges on those carbon atoms than on the carbon atoms that interact with Asp^{189} and Glu^{217}. Thus, the binding to Asp^{102} is more favorable than that to Asp^{189} and Glu^{217} as shown in Table 1, which lists the binding energies calculated at the level of DFT B3LYP/6-31G*. We believe B3LYP/6-31G* is capable of giving fairly good results on the relative binding energies of different residues in the gas phase.

The MFCC method is also applied with the protonated form of L86 for comparison, and we randomly choose an amine group to add a hydrogen atom. The results are also shown in Table 1. We only calculate the binding energies of amino-acid fragments, and the energies between the concaps and thrombin are not calculated to reduce the computational cost. As we expected, all the interaction energies are stronger than the neutral form, especially those amino acids with negative charges. This is because MFCC calculation only gives the gas-phase interaction, and the electrostatic interactions between ions will be overestimated without considering the solvent effects.

The more reliable results calculated at MP2/6-31+G* levels are shown in Table 1, and we only calculate the interaction energies of 12 amino-acid fragments for two inhibitors, which have strong interactions at B3LYP/6-31G* level, due to the heavy computation cost of MP2 calculation. The basis set superposition error, which results from the use of an incomplete basis set, is corrected (44). The interaction energies from MP2/6-31+G* calculation are slightly higher than the results from B3LYP/6-31G*. It is more important, however, that the relative binding energies of different residues are essentially similar from different calculations. The interaction energies calculated at B3LYP/6-31G* level are in quite good agreement with results computed at the MP2/6-31+G* level expect for Trp^{215}. Here the MP2

calculation also uses the small concap CH_3-CH_3 for Trp^{215}. It is worth noting that the binding interaction at MP2 level is much stronger than at B3LYP level with the same small concap. Because other groups found DFT unreliable for π-stacking, and incapable of including the van der Waals interaction (45–47), we believe there is also strong hydrophobic and aromatic stacking interaction between the phenyl moiety of the inhibitor and the indole ring of Trp^{215}, and this kind of interaction with Trp^{215} also mentioned in Costanzo et al. (9).

We further investigate the dynamics and fluctuation of the binding complex in the 5 ns MD simulation in which the particle mesh Ewald method is carried out in explicit water for the L86/THROMBIN system. For 100 snapshots extracted from the last 100 ps of the stable trajectory, gas-phase binding energies are calculated and averaged using AMBER force field. Fig. 7 shows the comparison between two force-field (from AMBER force-field) interaction energy spectra. The spectrum in Fig. 7 *a* is calculated using the x-ray crystallography structure, while the spectrum in Fig. 7 *b* is the time-averaged result over the MD collected snapshots. Here, we found there is no significant difference between the two spectra, and the residues of Asp^{102}, Ser^{214}, Trp^{215}, Gly^{216}, and Glu^{217} are still the primary bindings to L86. Some minor differences between these two spectra can be observed, because MD can relax the experimental structure and give a more stable conformation comparing with the crystal structure from the point of view of the complex conformation as a whole. Thus, the simulation is reasonable and good enough to obtain lowest energy structure.

Comparing the quantum energy spectrum in Fig. 4 and those from the force field in Fig. 7 for L86-thrombin binding, it is assuring to see that the dominant interaction profiles are quite similar, with difference only in some details. This is not surprising since the dominant binding interactions here are those of hydrogen bonding, which can also be reasonably represented by the force field. Thus, we can also use force-field calculation to provide additional information on the binding mechanism. Fig. 8 plots the time-dependence of five dominant L86-thrombin residue interaction energies calculated using the AMBER force field. Fluctuation in the interaction energy represents the sensitivity of these values to conformational details. As shown in Fig. 8, the binding of L86 to these residues are quite stable.

T76 binding to thrombin

Fig. 9 is similar to Fig. 7 except for T76/thrombin complex, and it shows the comparison between two force-field interaction energy spectra, one from the x-ray crystal structure and the other from the time-averaged structure over the collected MD snapshots. As we can see from the figure, the interaction between T76 and Glu^{146} has changed from repulsion to attraction after a 5-ns MD run due to the relaxation, and we obtain more reasonable complex conformation from the MD simulation. Overall, the main interactions between T76 and residues of thrombin are consistent, so the simulation process is good enough to allow us to extract structures for MFCC and MM-PBSA calculations.

We follow the same numerical procedure as described previously to perform MFCC calculation. The computed interaction spectrum at B3LYP/6-31G* level in Fig. 10 shows that the dominant interactions between T76 and thrombin are the bindings to Asp^{189}, Ser^{214}, Trp^{215}, Gly^{216}, Asp^{102}, and Glu^{146}. Fig. 11 plots the relative positions of T76 in complex with these residues having strong binding interactions.

From the interaction spectrum in Fig. 10 and Table 1, we notice that T76 and L86 bind to thrombin in a very similar inhibition mode due to their similar chemical structures. By

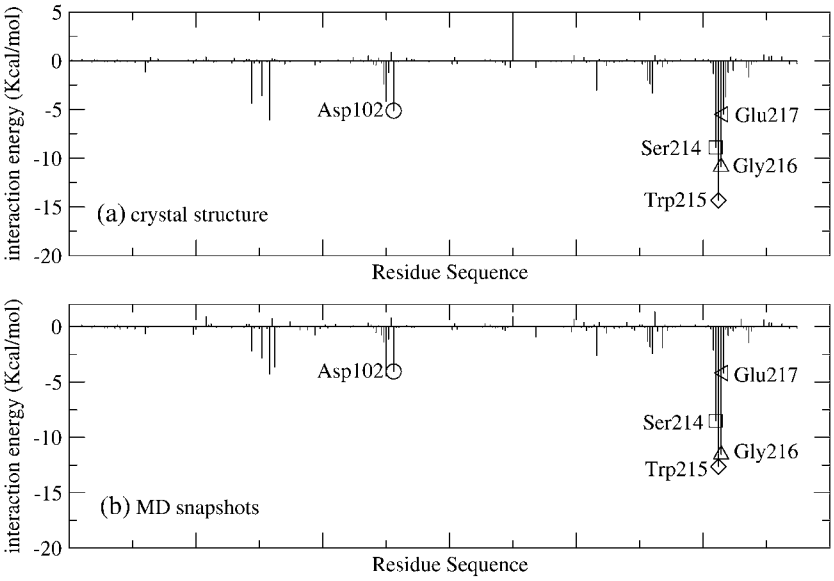

FIGURE 7 The interaction spectrum for the L86/THROMBIN binding complex calculated using the force field from (*a*) the crystal structure and (*b*) the averaged structures over the MD collected snapshots.

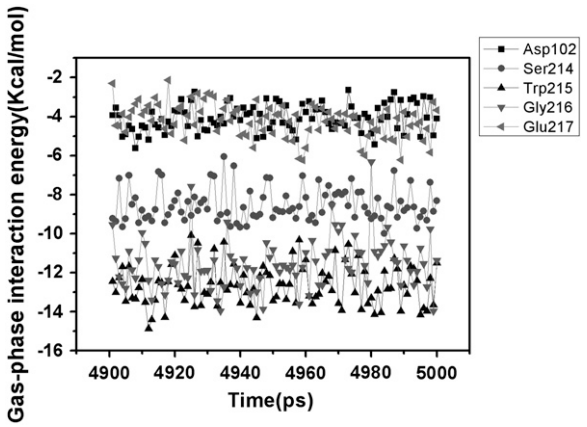

FIGURE 8 The time-dependence of interaction energies calculated using AMBER force field for the five dominant residues binding to L86 for 100 consecutive snapshots of the L86/thrombin complex.

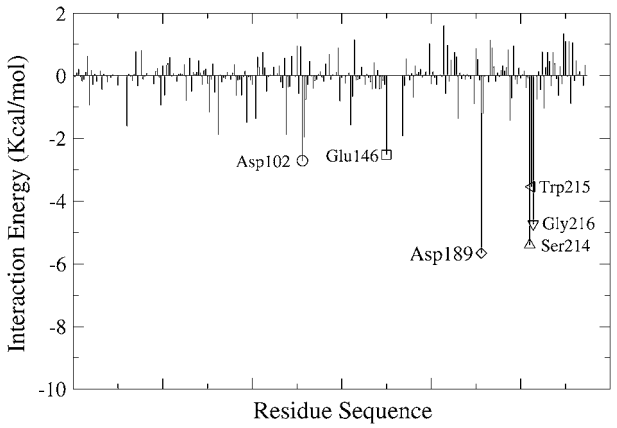

FIGURE 10 MFCC computed interaction energy spectrum for T76/THROMBIN complex at B3LYP/6-31G* level.

comparing results in Fig. 10 and geometries in Fig. 11, we can explain the binding character of T76 to thrombin as follows. Similar to L86, T76 also forms one H-bond to backbone carbonyl oxygen of Ser^{214}, two H-bonds to Gly^{216} in an antiparallel β-stand fashion, and ion-dipole attractive interaction (T76 is also a polar molecule, and its permanent dipole is ~4.0 Debye) with Asp^{189} and Asp^{102}. In addition, because Asp^{189} interact with the carbon atoms that are directly bonded to chlorine and oxygen atoms with the distance of 5.13 Å, 5.71 Å, and 6.56 Å, the binding energy between Asp^{189} and T76 is significantly larger than that between Asp^{189} and L86.

We also calculated individual thrombin fragments-T76 interaction energies for several dominant binding residues at MP2/6-31+G* level as shown in Table 1. As we can see clearly from the table, most of the calculated interaction energies using both MP2/6-31+G* and B3LYP/6-31G* method are in close agreement with each other, with the MP2/6-31+G* energies being slightly higher in general. Similarly, we believe there is also significant interaction between T76 and Trp^{215}, as in the case for L86 according to the MP2 result.

The interaction between Glu^{146} and T76 at different computational levels is a little different. The MP2 result shows there should be weak repulsion between them, while their interaction should be attractive interaction at the B3LYP level. We also check the structure from the MD simulation, and there is no bridge water between them to form extra H-bonds to reduce the repulsion. Although the results at different levels are not exactly consistent, the strong repulsion between them in the crystal structure is highly reduced after MD simulation. Examination of the relative geometry in Fig. 11 shows that the distance is 6.15 Å

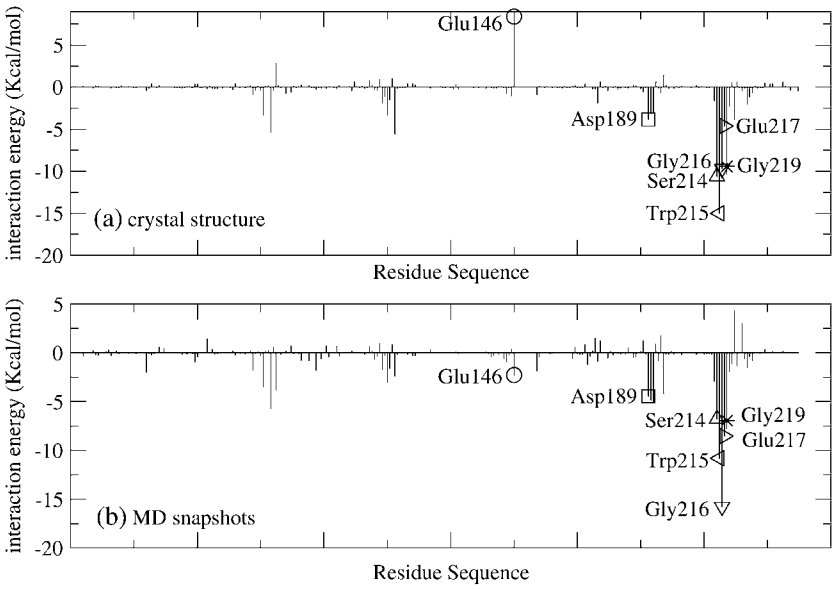

FIGURE 9 Similar to Fig. 7 except for the T76/thrombin complex.

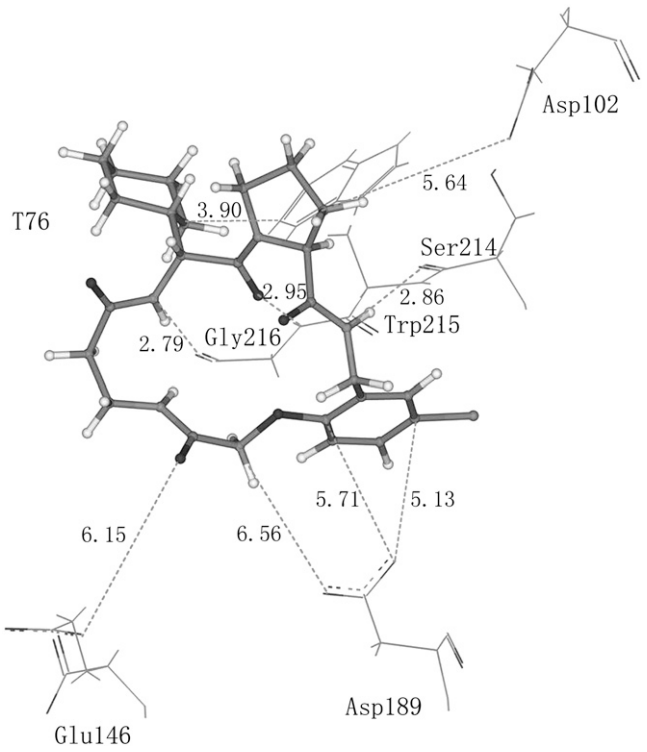

FIGURE 11 Relative geometries and distances of T76 and THROMBIN fragments in T76/THROBMBIN complex.

between the oxygen atom of T76 and the negatively charged carboxyl group of Glu^{146}, and this may be the source of the weak repulsion.

DISCUSSION AND CONCLUSION

The MFCC method has been applied to study binding mechanism of L86 and T76 to human α-thrombin at DFT level. In addition, molecular dynamics simulation has been carried out to investigate the dynamics and stability of the thrombin-inhibitor complex and to calculate binding free energy using the MM-PBSA approach. Our study shows that the strong bindings of L86 to Ser^{214}, Gly^{216}, Trp^{215}, Glu^{217}, Asp^{102}, and Asp^{189} are the primary mechanism of drug binding to thrombin, and T76 binds to thrombin in a very similar inhibition mode to L86, except that T76 has a relatively weaker interaction with Glu^{217} in the quantum calculation.

The overall gas-phase interaction energies of the two inhibitors from MFCC are almost the same to each other (~ -33.9 kcal/mol), and this is likely due to the similar binding mode of the two inhibitors. Calculated free energies of binding from MM-PBSA are listed in Table 2, and they are -14.6 and -13.3 kcal/mol, respectively, for bindings of L86 and T76 to thrombin. Although the current MFCC calculation does not attempt to provide a quantitative comparison between theoretical calculation and the observed

TABLE 2 Binding free energies (kcal/mol) from MM-PBSA calculation and normal mode analysis for two thrombin/inhibitor complexes

Inhibitors	L86	T76
ΔHB_{bB}(s)	−37.7	−34.8
Δ(TS)B_{bB}(s)	−23.1	−21.5
ΔGB_{bB}(s)	−14.6	−13.3
ΔGB_{bB}(experiment)*	−13.7*	−12.2*

*The experimental binding free energies (-13.7 and -12.2) are calculated using K_i, which are provided in Nantermet et al. (19).

experimental binding affinities, both the MM-PBSA and MFCC calculations draw the same conclusion that binding energies of L86/thrombin and T76/thrombin are close to each other because of similar binding modes. Thus, MFCC calculation is consistent with the result from MM-PBSA calculation, though MFCC only gives the gas-phase interaction energy. The free energies for L86/T76-thrombin binding from MM-PBSA calculation are consistent with the experiment values of Nantermet et al. (19).

Current computational study provides an easy means to extract molecular insight of protein-ligand binding by explicitly calculating the interaction energy between individual residues and the ligand at quantum mechanical levels. These individual residue-ligand interaction energies provide detailed quantitative information about specific residue interaction with the ligand that should be extremely useful for our understanding of the molecular nature of protein-ligand binding. This gas-phase binding energy analysis is complementary to MD simulation and free energy calculation, and therefore provides additional insight into the mechanism of protein-ligand binding at quantum mechanical levels.

Some words on the level of quantum chemistry calculation are warranted here. Because current DFT or Hartree-Fock calculations are incapable of giving van der Waals (or dispersion) energies, it is desirable to employ higher level of quantum chemistry methods to perform such calculation. This is possible and practical with the MFCC approach, as was demonstrated in a recent study of water-protein interaction at the MP2 level (48). Future studies will be carried out to investigate the dispersion interaction energies using higher-level quantum chemistry methods or mixed methods with higher-level theory for close contact residues and HF/DFT for more distant residues. Of course, significantly higher computational cost will be needed when higher-level quantum methods with larger basis sizes are employed.

Another issue to be investigated in the future is the solvent effect. How these individual amino acid-ligand interactions energies are affected by the presence of solvent is also very important. A new method (49) is being developed by this group to include the solvent effect in the MFCC study of proteins and such applications are currently being planned.

We thank Changge Ji for helpful discussions. We thank the supercomputer center of Virtual Laboratory of Computational Chemistry, Computer

Network Information Center, Chinese Academy of Sciences for the computational resources.

This work is supported by Chinese Academy of Sciences with the Knowledge Innovation Project and National Natural Science Foundation of China (grants No. 20573110 and 20403019) and National Basic Research Program (grant No. 2004CB719901).

REFERENCES

1. Davie, E. W., K. Fujikawa, and W. Kisiel. 1991. The coagulation cascade—initiation, maintenance, and regulation. *Biochemistry.* 30: 10363–10370.
2. Pierce, A. C., and W. L. Jorgensen. 2001. Estimation of binding affinities for selective thrombin inhibitors via Monte Carlo simulations. *J. Med. Chem.* 44:1043–1050.
3. Vu, T. K. H., D. T. Hung, V. I. Wheaton, and S. R. Coughlin. 1991. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell.* 64:1057–1068.
4. Fenton, J. W., F. A. Ofosu, D. G. Moon, and J. M. Maraganore. 1991. Thrombin structure and function—why thrombin is the primary target for antithrombotics. *Blood Coagul. Fibrinolysis.* 2:69–75.
5. Raskob, G. E., and J. N. George. 1997. Thrombotic complications of antithrombotic therapy: a paradox with implications for clinical practice. *Ann. Intern. Med.* 127:839–841.
6. Thomas, D. J. http://www.fda.gov/cder/foi/appletter/2000/20883ltr.pdf (7/1/2000UT).
7. Bizios, R., L. Lai, J. W. Fenton, and A. B. Malik. 1986. Thrombin-induced chemotaxis and aggregation of neutrophils. *J. Cell. Physiol.* 128:485–490.
8. Steiner, J. L. R., M. Murakami, and A. Tulinsky. 1998. Structure of thrombin inhibited by aeruginosin 298-A from a blue-green alga. *J. Am. Chem. Soc.* 120:597–598.
9. Costanzo, M. J., H. R. Almond, L. R. Hecker, M. R. Schott, S. C. Yabut, H. C. Zhang, P. Andrade-Gordon, T. W. Corcoran, E. C. Giardino, J. A. Kauffman, J. M. Lewis, L. de Garavilla, B. J. Haertlein, and B. E. Maryanoff. 2005. In-depth study of tripeptide-based alpha-ketoheterocycles as inhibitors of thrombin. Effective utilization of the S-1′ subsite and its implications to structure-based drug design. *J. Med. Chem.* 48:1984–2008.
10. Thorstensson, F., I. Kvarnstrom, D. Musil, I. Nilsson, and B. Samuelsson. 2003. Synthesis of novel thrombin inhibitors. Use of ring-closing metathesis reactions for synthesis of P2 cyclopentene- and cyclohexenedicarboxylic acid derivatives. *J. Med. Chem.* 46:1165–1179.
11. Young, M. B., J. C. Barrow, K. L. Glass, G. F. Lundell, C. L. Newton, J. M. Pellicore, K. E. Rittle, H. G. Selnick, K. J. Stauffer, J. P. Vacca, P. D. Williams, D. Bohn, F. C. Clayton, J. J. Cook, J. A. Krueger, L. C. Kuo, S. D. Lewis, B. J. Lucas, D. R. McMasters, C. Miller-Stein, B. L. Pietrak, A. A. Wallace, R. B. White, B. Wong, Y. W. Yan, and P. G. Nantermet. 2004. Discovery and evaluation of potent P-1 aryl heterocycle-based thrombin inhibitors. *J. Med. Chem.* 47:2995–3008.
12. Jones Hertzog, D. K., and W. L. Jorgensen. 1997. Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method. *J. Med. Chem.* 40:1539–1549.
13. Linusson, A., J. Gottfries, T. Olsson, E. Ornskov, S. Folestad, B. Norden, and S. Wold. 2001. Statistical molecular design, parallel synthesis, and biological evaluation of a library of thrombin inhibitors. *J. Med. Chem.* 44:3424–3439.
14. Bode, W., I. Mayr, U. Baumann, R. Huber, S. R. Stone, and J. Hofsteenge. 1989. The refined 1.9 Å crystal structure of human α-thrombin—interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* 8:3467–3475.
15. Nilsson, J. W., I. Kvarnstrom, D. Musil, I. Nilsson, and B. Samuelsson. 2003. Synthesis and SAR of thrombin inhibitors incorporating a novel 4-amino-morpholinone scaffold: analysis of x-ray crystal structure of enzyme inhibitor complex. *J. Med. Chem.* 46:3985–4001.
16. Grootenhuis, P. D. J., and P. J. M. Vangalen. 1995. Correlation of binding affinities with nonbonded interaction energies of thrombin-inhibitor complexes. *Acta Crystallogr. D Biol. Crystallogr.* 51:560–566.
17. Grootenhuis, P. D. J., and M. Karplus. 1996. Functionality map analysis of the active site cleft of human thrombin. *J. Comput. Aided Mol. Des.* 10:1–10.
18. Bohm, M., J. Sturzebecher, and G. Klebe. 1999. Three-dimensional quantitative structure-activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J. Med. Chem.* 42:458–477.
19. Nantermet, P. G., J. C. Barrow, C. L. Newton, J. M. Pellicore, M. B. Young, S. D. Lewis, B. J. Lucas, J. A. Krueger, D. R. McMasters, Y. W. Yan, L. C. Kuo, J. P. Vacca, and H. G. Selnick. 2003. Design and synthesis of potent and selective macrocyclic thrombin inhibitors. *Bioorg. Med. Chem. Lett.* 13:2781–2784.
20. Tucker, T. J., S. F. Brady, W. C. Lumma, S. D. Lewis, S. J. Gardell, A. M. Naylor-Olsen, Y. W. Yan, J. T. Sisko, K. J. Stauffer, B. J. Lucas, J. J. Lynch, J. J. Cook, M. T. Stranieri, M. A. Holahan, E. A. Lyle, E. P. Baskin, I. W. Chen, K. B. Dancheck, J. A. Krueger, C. M. Cooper, and J. P. Vacca. 1998. Design and synthesis of a series of potent and orally bioavailable noncovalent thrombin inhibitors that utilize nonbasic groups in the P1 position. *J. Med. Chem.* 41:3210–3219.
21. Tucker, T. J., W. C. Lumma, A. M. Mulichak, Z. G. Chen, A. M. Naylor-Olsen, S. D. Lewis, R. Lucas, R. M. Freidinger, and L. C. Kuo. 1997. Design of highly potent noncovalent thrombin inhibitors that utilize a novel lipophilic binding pocket in the thrombin active site. *J. Med. Chem.* 40:830–832.
22. Zhang, D. W., Y. Xiang, and J. Z. H. Zhang. 2003. New advance in computational chemistry: full quantum mechanical ab initio computation of streptavidin-biotin interaction energy. *J. Phys. Chem. B.* 107:12039–12041.
23. Zhang, D. W., and J. Z. H. Zhang. 2003. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy. *J. Chem. Phys.* 119:3599–3605.
24. He, X., Y. Mei, Y. Xiang, D. W. Zhang, and J. Z. H. Zhang. 2005. Quantum computational analysis for drug resistance of HIV-1 reverse transcriptase to nevirapine through point mutations. *Proteins Structure Function Bioinformatics.* 61:423–432.
25. Mei, Y., X. He, Y. Xiang, D. W. Zhang, and J. Z. H. Zhang. 2005. Quantum study of mutational effect in binding of efavirenz to HIV-1 RT. *Proteins Structure Function Bioinformatics.* 59:489–495.
26. Zhang, D. W., Y. Xiang, A. M. Gao, and J. Z. H. Zhang. 2004. Quantum mechanical map for protein-ligand binding with application to beta-trypsin/benzamidine complex. *J. Chem. Phys.* 120:1145–1148.
27. Chen, X. H., D. W. Zhang, and J. Z. H. Zhang. 2004. Fractionation of peptide with disulfide bond for quantum mechanical calculation of interaction energy with molecules. *J. Chem. Phys.* 120:839–844.
28. Case, D. A., T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross, and P. A. Kollman. AMBER8. 2004. University of California, San Francisco, CA.
29. Wang, J. M., P. Cieplak, and P. A. Kollman. 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* 21:1049–1074.
30. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald—an $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98: 10089–10092.
31. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical-integration of Cartesian equations of motion of a system with constraints—molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327–341.
32. Gohlke, H., C. Kiel, and D. A. Case. 2003. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RaIGDS complexes. *J. Mol. Biol.* 330: 891–913.
33. Frish, M. J., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam,

S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. Gaussian 03, Rev. B.05. 2003. Gaussian, Inc. Pittsburgh, PA.

34. Beveridge, D. L., and F. M. Dicapua. 1989. Free-energy via molecular simulation—applications to chemical and biomolecular systems. *Annu. Rev. Biophys. Biophys. Chem.* 18:431–492.
35. Kollman, P. 1993. Free-energy calculations—applications to chemical and biochemical phenomena. *Chem. Rev.* 93:2395–2417.
36. Wang, J. M., P. Morin, W. Wang, and P. A. Kollman. 2001. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J. Am. Chem. Soc.* 123:5221–5230.
37. Srinivasan, J., J. Miller, P. A. Kollman, and D. A. Case. 1998. Continuum solvate studies of stability of RNA hairpin loops and helices. *J. Biomol. Struct. Dyn.* 16:671–682.
38. Rocchia, W., E. Alexov, and B. Honig. 2001. Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. *J. Phys. Chem. B.* 105:6507–6514.
39. Connolly, M. L. 1983. Analytical molecular-surface calculation. *J. Appl. Crystallogr.* 16:548–558.
40. Gohlke, H., and D. A. Case. 2003. Converging free energy estimates: MM-Pb(GB)SA studies on the protein-protein complex Ras-Raf. *J. Comput. Chem.* 25:238–250.
41. Sitkoff, D., K. A. Sharp, and B. Honig. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* 98:1978–1988.
42. McQuarrie, D. A. 1976. Statistical Mechanics. Harper & Row, New York.
43. Pearlman, D. A. 2005. Evaluating the molecular mechanics Poisson-Boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP Kinase. *J. Med. Chem.* 48:7796–7807.
44. Boys, S. F., and F. Bernardi. 2002. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. (Reprinted from *Molecular Physics*, 1970, 19:553–566). *Mol. Phys.* 100:65–73.
45. Tauer, T. P., and C. D. Sherrill. 2005. Beyond the benzene dimer: an investigation of the additivity of π–π interactions. *J. Phys. Chem. A.* 109:10475–10478.
46. Ye, X. Y., Z. H. Li, W. Wang, K. Fan, W. Xu, and Z. Hua. 2004. The parallel π–π stacking: a model study with MP2 and DFT methods. *Chem. Phys. Lett.* 397:56–61.
47. Johnson, E. R., R. A. Wolkow, and G. A. DiLabio. 2004. Application of 25 density functionals to dispersion-bound homomolecular dimers. *Chem. Phys. Lett.* 394:334–338.
48. Zhang, D. W., and J. Z. H. Zhang. 2004. Full quantum mechanical *ab initio* computation of protein-molecule interaction energies. *J. Theor. Comput. Chem.* 3:43–49.
49. Mei, Y., C. G. Ji, and J. Z. H. Zhang. 2006. A new quantum method for electrostatic solvation energy of protein. *J. Chem. Phys.* 125: 94906.