

# Interpretation of the Binding Affinities of PTP1B Inhibitors with the MM-GB/SA Method and the X-Score Scoring Function

Xinglong Zhang, Xun Li, and Renxiao Wang\*

State Key Laboratory of Bioorganic Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, P. R. China

Received December 6, 2008

We have studied the binding affinities of a set of 45 small-molecule inhibitors to protein tyrosine phosphatase 1B (PTP1B) through computational approaches. All of these compounds share a common oxalylamino benzoic acid (OBA) moiety. The complex structure of each compound was modeled by using the GOLD program plus the ASP scoring function. Each complex structure was then subjected to a molecular dynamics (MD) simulation of 2 ns long by using the AMBER program. Based on the configurational ensembles retrieved from MD trajectories, both MM-GB/SA and MM-PB/SA were employed to compute the binding free energies of all 45 PTP1B inhibitors. The correlation coefficient between the MM-GB/SA results and experimental binding data was 0.87 and the standard deviation was 0.60 kcal/mol. The performance of MM-PB/SA was slightly inferior to that of MM-GB/SA. Several aspects of the MM-GB(PB)/SA method were explored in our study to obtain optimized results. The X-Score scoring function was found to produce equally good results as MM-GB/SA on both the complex structures prepared by molecular docking and the configurational ensembles obtained through lengthy MD simulations. The structure–activity relationship of this set of compounds is also discussed based on the computed results. The computational approaches validated in our study are hopefully applicable to the study of other classes of PTP1B inhibitors.

## 1. INTRODUCTION

Reversible tyrosine phosphorylation is catalyzed by the coordinated actions of protein tyrosine kinases (PTKs) and phosphatases (PTPs), which are of paramount importance to many biological processes, such as cell growth and differentiation, cell-cell communication, cell migration, gene transcription, ion channel activity, immune response, apoptosis/survival decision, and so on.<sup>1–3</sup> Protein tyrosine phosphatase 1B (PTP1B) is a member of the PTP family. It is regarded as a major negative regulator of both insulin- and leptin-stimulated signal transduction pathways. Specific PTP1B inhibitors may enhance insulin and leptin sensitivity and thus may be used as effective therapeutics for the treatment of type II diabetes, insulin resistance, and obesity.<sup>4</sup>

Development of PTP1B inhibitors has received extensive attention due to their potential pharmaceutical applications. The full-length PTP1B has 325 amino acid residues. The conserved residues forming the active site of PTP1B include Cys215, Ser216, and Arg221. Because all PTPs share a high degree of structural conservation in the active site, designing inhibitors with both high affinity and selectivity for PTP1B poses a challenge. In 1997, Zhang et al. discovered a second binding site adjacent to the main active site.<sup>5</sup> The key residues at the adjacent binding site include Arg24 and Arg254, which are less conserved. This observation has suggested a novel paradigm for the development of potent and specific PTP1B inhibitors, namely bidentate ligands that bind to both the active site and the unique peripheral site. Based on this paradigm, a number of classes of potent small-

molecule PTP1B inhibitors have been developed. Among them, the PTP1B inhibitors containing an oxalylamino benzoic acid (OBA) moiety are particularly notable, which were first discovered by Møller et al. from high throughput screening.<sup>6</sup> The OBA moiety closely mimics the native structure of a phosphorylated tyrosine. In addition, the relatively small size of the OBA moiety leaves plenty of room for further structural modifications. Møller et al. optimized the chemical structures of their lead compounds and obtained a few inhibitors with high activity as well as selectivity.<sup>7</sup> Liu et al. synthesized and tested a series of OBA-containing PTP1B inhibitors.<sup>8–10</sup> Jirousek et al. adopted a fragment-based strategy to obtain a PTP1B inhibitor with an inhibition constant of 22 nM, which also exhibited good selectivity within the PTP family.<sup>11</sup>

The primary aims of our study are to model the binding of PTP1B inhibitors to PTP1B and to establish a reliable way for the computation of their binding affinities. In this work, we choose to study the structure–activity relationship of the well-cited OBA-containing PTP1B inhibitors reported by Liu et al.<sup>8–10</sup> These compounds, 45 in total, exhibit a remarkable structural diversity except for the common oxalylamino benzoic acid moiety, and their binding affinities span nearly 4 orders of magnitude. As proved by crystal structures,<sup>9,10</sup> this class of PTP1B inhibitors occupies the main active site as well as the adjacent peripheral site and thus serves as a good case for validating the fragment-based strategy in structure-based drug design. We have combined molecular docking and molecular dynamics simulation to deduce the complex structures of these PTP1B inhibitors, and we have employed two different methods, viz., MM-GB(PB)/SA and X-Score, in binding affinity computation.

\* Corresponding author phone: 86-21-54925128; e-mail: wangrx@mail.sioc.ac.cn.

The optimal parameters and settings of MM-GB(PB)/SA and X-Score on these PTP1B inhibitors are explored along the process. To the best of our knowledge, our study is the first computational study of this class of PTP1B inhibitors. Understanding the interactions between this class of compounds and PTP1B at the atomic level may also inspire the development of new classes of PTP1B inhibitors with better potency and selectivity.

## 2. METHODS

In this work, we studied the binding of a set of 45 OBA-containing compounds to PTP1B through computational approaches. We first validated the docking/scoring tools available to us with 26 diverse PTP1B-inhibitor complexes with known three-dimensional structures and reliable binding data. Then, the complex structures of 45 PTP1B inhibitors were modeled by using the most suitable docking/scoring scheme identified at the previous step. Each complex model was subsequently refined by 2 ns long molecular dynamics (MD) simulation in explicit solvent. The MM-GB(PB)/SA method as well as the X-Score scoring function were applied to compute the binding affinities of these complexes. The structure–activity relationship of this set of PTP1B inhibitors was also analyzed based on the computation results.

**2.1. Validation of Docking/Scoring Methods.** Molecular docking is the very essential approach for modeling the binding of a ligand molecule to its receptor. Currently, a number of computer programs, such as GOLD,<sup>12–14</sup> FlexX,<sup>15</sup> Dock,<sup>16</sup> Autodock,<sup>17–19</sup> Glide,<sup>20,21</sup> and LigandFit,<sup>22</sup> are capable for this task with a reasonable accuracy. In our study, we chose the GOLD program (version 3.2) to conduct all molecular docking jobs since some comparative evaluations of molecular docking programs<sup>23,24</sup> have demonstrated that GOLD has a robust performance on various systems. There are three scoring functions implemented in the GOLD program, i.e. GoldScore,<sup>13</sup> ChemScore,<sup>14</sup> and ASP.<sup>25</sup> Besides these internal scoring functions, we chose the X-Score program (version 1.2)<sup>26</sup> as an additional option for binding affinity computation since it was demonstrated to be among the relatively successful scoring functions by some previous evaluations.<sup>27,28</sup>

In order to determine the optimal docking/scoring method for reproducing the binding modes and binding affinities of PTP1B inhibitors, we tested all four possible combinations on a set of PTP1B complexes with known three-dimensional structures and reliable binding data. For this purpose, we searched among the “refined set” of the PDBbind database (version 2007).<sup>29,30</sup> The results were a total of 26 complexes formed between PTP1B and small-molecule ligands, including 1BZC, 1BZJ, 1C83, 1C84, 1C86, 1C87, 1C88, 1ECV, 1G7F, 1G7G, 1KAV, 1NL9, 1NNY, 1NO6, 1NZ7, 1ONY, 1ONZ, 1PYN, 1QXK, 1XBO, 2AZR, 2B07, 2F6T, 2H4G, 2H4K, and 2HB1. All of these complex structures are determined through X-ray crystal diffraction with resolution better than 2.5 Å, and the binding affinities of these complexes ( $K_i$  or  $K_d$ ) are experimentally measured. The processed structural file of each complex was also retrieved from the PDBbind database. In brief, each complex was separated into a protein molecule saved in a PDB-format file and a ligand molecule saved in a Mol2-format file so that they could be readily used as inputs to the GOLD

**Table 1.** RMSD Distribution of the Best-Scored Docking Solutions of 26 PTP1B Complexes in PDB Produced by the GOLD Program

rmsd	scoring function applied in docking		
	ASP	GoldScore	ChemScore
≤1.0 Å	11	10	12
≤2 Å	20	20	19
≤4 Å	26	23	22
>4 Å	0	3	4

**Table 2.** Correlations between the Experimental Binding Constants and the Binding Scores of 26 PTP1B Complexes in PDB Computed by Four Scoring Functions

scoring function	$R^a$	SD <sup>b</sup>
X-Score::HScore	0.72	0.72
X-Score::HPScore	0.69	0.75
X-Score::HMScore	0.63	0.80
GOLD::ChemScore	0.47	0.91
GOLD::ASP	0.26	1.00
GOLD::GoldScore	0.19	1.01

<sup>a</sup> Pearson correlation coefficient between the experimental binding constants and the computed binding scores. <sup>b</sup> Standard deviation (in log  $K_i$  units) in fitting the experimental binding constants and the computed binding scores.

program. We further examined the chemical structure of the ligand molecule in each complex to ensure that all atom types and bond types were correctly assigned.

Our validation of docking/scoring methods consisted of two parts. The first test was to examine the GOLD program in combination with GoldScore, ChemScore, and ASP, respectively, to determine which scheme was able to best reproduce the observed ligand binding poses in all 26 PTP1B complexes. For this purpose, the GOLD program was employed to perform molecular docking of the ligand in each complex to the PTP1B structure contained in the same complex. The binding pose of the ligand observed in the crystal structure was used as the reference for defining the binding site, and the radius of the binding site was set to 10 Å. The “GA runs” parameter was set to 30, and the “GA operations” parameter was set to 500000. Other parameters were set to the default values of the program. Distributions of the rmsd values between the best-scored docking solutions and the observed ligand binding poses of all 26 PTP1B complexes produced by each docking/scoring scheme are summarized in Table 1. As indicated in this table, the success rates of all three scoring functions in producing acceptable docking solutions (e.g., rmsd < 2.0 Å) were almost identical on these PTP1B complexes. Nevertheless, application of ASP did not produce any significant outliers (e.g., rmsd > 4.0 Å). Therefore, the ASP scoring function was applied in all of our subsequent molecular docking jobs.

In the second test, we tested all four scoring functions, i.e. GoldScore, ChemScore, ASP, and X-Score, directly on the crystal structures of all 26 PTP1B complexes to determine which scoring function could produce the highest correlation between its binding scores and the experimentally measured binding affinities of these complexes. The results are summarized in Table 2. It needs to be mentioned that X-Score actually provides three options, namely HPScore, HMScore, and HScore, for the computation of protein–ligand binding affinities. These three options differ in the algorithm for computing the hydrophobic effect.<sup>26</sup> As indicated in Table

2, all of them performed generally better in this test than the three scoring functions implemented in GOLD. The best option was HScore, which produced a standard deviation of 0.72 log units in fitting the known binding constants of these 26 PTP1B complexes. This level of accuracy corresponds to 0.98 kcal/mol in terms of standard binding free energy at room temperature.

**2.2. Construction of the Structural Models of PTP1B-Inhibitor Complexes.** The PTP1B inhibitors under our study together with their inhibition constants ( $K_i$ ) are summarized in Table 3. It needs to be mentioned that among all these compounds, three have crystal structures in complex with PTP1B: compounds **4** (PDB entry: 1ONZ), **24** (PDB entry: 1ONY), and **34** (PDB entry: 1NZ7). We chose PDB entry 1ONY as the template for modeling the complex structures of the other compounds since compound **24** was of intermediate size among this set of PTP1B inhibitors. Accordingly, each molecule (except for **4**, **24**, and **34**) listed in Table 3 was sketched by making necessary modifications to the three-dimensional structure of **24** with the Sybyl software.<sup>31</sup> Each sketched molecule was then minimized to convergence with the Powell method using the Tripos force field and the Gasteiger-Hückel partial charges.

The GOLD program was then used to perform molecular docking of all 42 modeled compounds into the PTP1B binding site. The PTP1B structure retrieved from PDB entry 1ONY was used in docking. The ASP scoring function, as suggested by the previous validation step, was applied in these docking jobs. Other parameters were the same as those used in the previous validation step. A total of 30 binding poses for each compound were produced by the GOLD program. The final binding pose of each compound for subsequent analyses was then selected with human interpretation. For this purpose, we assumed that this class of compounds had similar binding modes with PTP1B as compounds **4**, **24**, and **34**. The conserved H-bond interactions between this class of compounds and PTP1B observed in the crystal structures of 1ONZ, 1ONY, and 1NZ7 are sketched in Figure 1. These H-bond interactions include the ones formed between the two carboxylic groups on the ligand and the amide hydrogen atoms in the active site of PTP1B (site 1), the ones between the carboxylic group on the ligand and the Arg24 residue at the peripheral binding site (site 2), and the ones between the amide hydrogen atom on the ligand and the Asp48 residue at the linker region. Accordingly, the binding poses of each compound that did not contain such key H-bonds were ignored. Among the remaining candidates, the one with the highest ASP binding score was selected as the final binding pose for subsequent analyses.

**2.3. Molecular Dynamics Simulations.** We then applied MD simulations to refine the models of the PTP1B-inhibitor complexes prepared at the previous step. All MD simulations were conducted by using the AMBER (version 9.0) program.<sup>32</sup> To set up the MD job for each complex, the ligand was processed by using the Antechamber<sup>33</sup> module in AMBER and assigned the AM1-BCC charges.<sup>34,35</sup> Atoms on PTP1B were assigned the PARM99 charges, and all ionizable residues were set at their default protonation states at neutral pH. The complex was soaked in a box of TIP3P<sup>36</sup> water molecules with a margin of 10 Å along each dimension. An appropriate number of counterions were added to neutralize the whole system. The entire system was then

minimized at three rounds, each of which consisted of 1000 steps, with harmonic constraints on all non-hydrogen atoms. The force constant was set to 100, 10, and 0 kcal/(mol\*Å<sup>2</sup>) at each round, respectively. Such a stepwise procedure was adopted to gradually resolve the steric clashes between the complex and solvent molecules.

The MD simulation was started by heating the entire system from 0 to 300 K in 100 ps and equilibrating at 300 K for another 100 ps. A subsequent 2 ns production run was performed under a constant temperature of 300 K and a constant pressure of 1 atm. The time interval was set to 2 fs. The General AMBER Force Field (GAFF)<sup>37</sup> was applied in simulation with the Particle Mesh Ewald (PME) method<sup>38</sup> for handling long-range electrostatic interactions and the SHAKE algorithm<sup>39</sup> for constraining all covalent bonds connecting hydrogen atoms. No other constraint was applied to either the protein or the ligand during the entire MD simulation. Note that the "ELEC" parameter was set to 1.2 instead of the default value of 1.0 in the MD simulations of all PTP1B complexes, including 42 modeled structures and 3 crystal structures. The reason will be explained in the *Results and Discussion* section later in this manuscript. All these MD simulations were performed on an Intel Xeon 5345-based Linux cluster.

**2.4. MM-GB/SA and MM-PB/SA Calculations.** The MM-GB(PB)/SA method is a popular method for the computation of protein–ligand binding free energies, which was first developed by Kollman et al. in 2000.<sup>40</sup> A good number of successful applications of the MM-GB(PB)/SA methods have already reported in the literature.<sup>41–49</sup> The MM-GB(PB)/SA method implemented in the AMBER9 program was applied to compute the binding free energies ( $\Delta G_{\text{bind}}$ ) of all 45 PTP1B complexes. The MM-GB(PB)/SA method can be conceptually summarized as

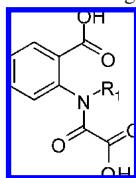
$$\Delta G_{\text{bind}} = \langle G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}} \rangle \quad (1)$$

$$G \approx E_{\text{gas}} - TS_{\text{conf}} + G_{\text{sol}} \quad (2)$$

$$E_{\text{gas}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{vdw}} + E_{\text{elec}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{elec}} + G_{\text{non-polar}} \quad (4)$$

Here, the binding free energy change in protein–ligand binding ( $\Delta G_{\text{bind}}$ ) is computed as the difference between the free energies of the complex ( $G_{\text{complex}}$ ), the protein ( $G_{\text{protein}}$ ), and the ligand ( $G_{\text{ligand}}$ ). The free energy of each molecule is calculated through eq 2 by summing up its internal energy in the gas phase ( $E_{\text{gas}}$ ), the solvation free energy ( $G_{\text{sol}}$ ), and a vibrational entropy term ( $TS$ ) computed by the normal-mode analysis.  $E_{\text{gas}}$  is a standard force field energy, including strain energies from covalent bonds and torsion angles as well as noncovalent van der Waals and electrostatic energies (eq 3). The solvation free energy ( $G_{\text{sol}}$ ) is computed as the sum of an electrostatic component and a nonpolar component (eq 4). The electrostatic component can be computed either by the Poisson-Boltzmann method (PB)<sup>50,51</sup> or the Generalized Born (GB) method,<sup>52</sup> while the nonpolar component is assumed to be proportional to the solvent-accessible surface area (SASA) of the molecule under consideration. As indicated by the angled brackets in eq 1, a distinctive feature of MM-GB(PB)/SA is that it computes the averages of all of the above properties over a configurational ensemble,

**Table 3.** Chemical Structures and Inhibition Constants of the OBA-Containing PTP1B Inhibitors Considered in This Study<sup>a</sup>

No.	<i>R</i> <sub>1</sub>	<i>K</i> <sub>i</sub> (μM) <sup>a</sup>	No.	<i>R</i> <sub>1</sub>	<i>K</i> <sub>i</sub> (μM)
1		93	7		35
2		24	8		24 ± 1
3		39	9		17 ± 4
4		17 ± 7	10		11 ± 2
5		74	11		14 ± 1
6		60 ± 16			
No.	<i>R</i> <sub>2</sub>	<i>R</i> <sub>3</sub>	<i>R</i> <sub>4</sub>	<i>K</i> <sub>i</sub> (μM) <sup>a</sup>	
12	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		1.2 ± 0.3	
13	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.54 ± 0.14	
14	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		1.2	
15	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		4.7 ± 1.6	
16	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		2.5 ± 0.2	
17	H	t-BuOCO-		17.3 ± 1.9	
18	H	CH <sub>3</sub> CO-		9.8 ± 2.1	
19	H	CH <sub>3</sub> SO <sub>2</sub> -		9.1	
20		CH <sub>3</sub> CO-		1.1 ± 0.5	
21	-CH(CH <sub>3</sub> ) <sub>2</sub>	CH <sub>3</sub> CO-		1.2 ± 0.1	
22	-CH <sub>2</sub> CH <sub>2</sub> OH	CH <sub>3</sub> CO-		1.5 ± 0.7	

Table 3. Continued

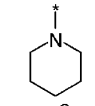
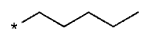
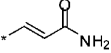
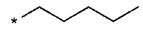
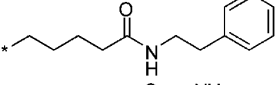
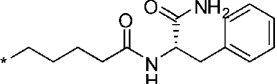
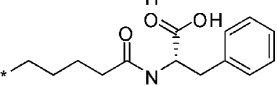
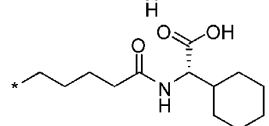
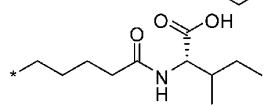
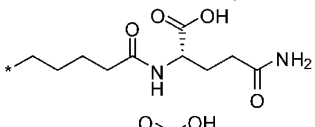
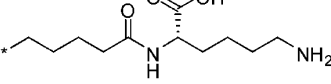
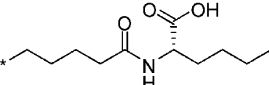
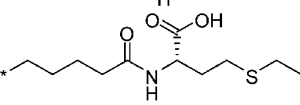
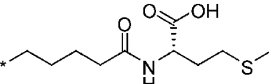
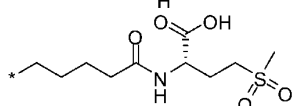
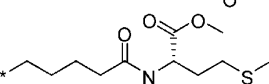
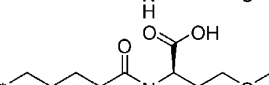
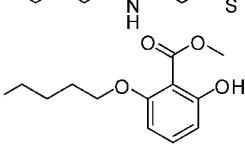
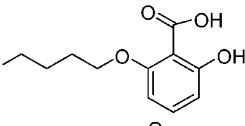
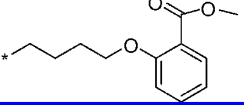
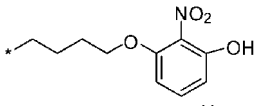
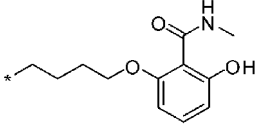
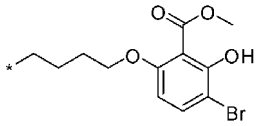
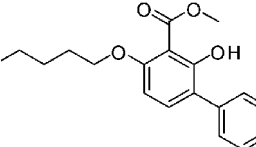
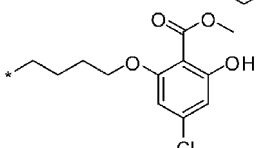
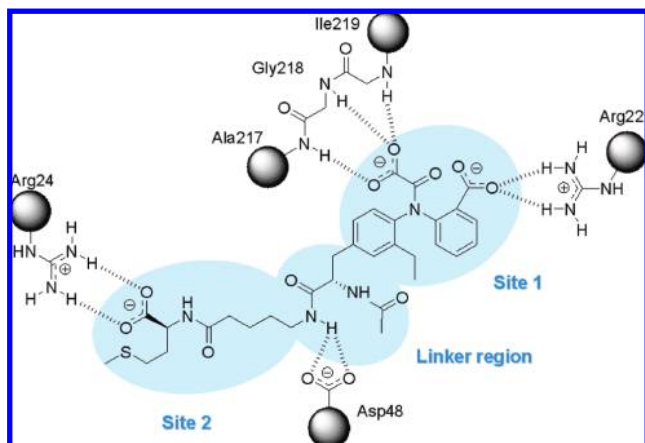
No.	$R_2$	$R_3$	$R_4$	$K_i(\mu\text{M})^a$
23		CH <sub>3</sub> CO-		2.0 ± 1.5
24		CH <sub>3</sub> SO <sub>2</sub> -		0.17 ± 0.07
25	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		3.4 ± 0.90
26	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		1.3 ± 0.40
27	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.14 ± 0.02
28	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.25 ± 0.10
29	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.43 ± 0.2
30	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.33 ± 0.16
31	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.71 ± 0.28
32	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.12 ± 0.07
33	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.13 ± 0.01
34	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.076 ± 0.015
35	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.47 ± 0.19
36	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.54 ± 0.19
37	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		1.7 ± 0.7
38	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.018 ± 0.004
39	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		0.042 ± 0.019
40	-CH <sub>2</sub> CH <sub>3</sub>	CH <sub>3</sub> CO-		1.60 ± 0.86



Table 3. Continued

No.	$R_2$	$R_3$	$R_4$	$K_i (\mu\text{M})^a$
41	$-\text{CH}_2\text{CH}_3$	$\text{CH}_3\text{CO}-$		$0.13 \pm 0.05$
42	$-\text{CH}_2\text{CH}_3$	$\text{CH}_3\text{CO}-$		$0.37 \pm 0.12$
43	$-\text{CH}_2\text{CH}_3$	$\text{CH}_3\text{CO}-$		$0.39 \pm 0.15$
44	$-\text{CH}_2\text{CH}_3$	$\text{CH}_3\text{CO}-$		$0.56 \pm 0.23$
45	$-\text{CH}_2\text{CH}_3$	$\text{CH}_3\text{CO}-$		$0.040 \pm 0.012$

<sup>a</sup> Experimentally measured  $K_i$  values of these compounds are cited from refs 8, 9, and 10. These data are for a racemic mixture with respect to the chiral center indicated by the asterisk. In our modeling, this chiral center in all these compounds was sketched in the *R* configuration as prompted by the known crystal complexed structures of three compounds among them.



**Figure 1.** Schematic illustration of the H-bonds formed between PTP1B and compound **34** (PDB entry 1NZ7). This compound occupies both the active site (site 1) and the peripheral binding site (site 2).

which is typically generated through MD simulations of the given protein–ligand complex.

In our study, the internal energy ( $E_{\text{gas}}$ ) of each given molecule was computed by the General AMBER Force Field (GAFF). Note that the “ELEC” parameter was restored to the default value of 1.0 to produce the correct force field energies at this step. The electrostatic term in the solvation free energy ( $G_{\text{elec}}$ ) was calculated by both the PB and GB models implemented in the PBSA module in the AMBER9 program. For PB computations, the intrinsic atom radii set provided by AMBER9 was adopted. The dielectric constant was set to 1 inside the solute and 80 in solvent. For GB

computations, Onufriev’s GB model was chosen since it provides a robust molecular volume correction and is supposedly better than earlier pairwise GB models.<sup>53</sup> The same set of atomic charges on each complex used in MD simulation was used in both PB and GB computations. The nonpolar term in solvation free energy was computed as  $G_{\text{nonpolar}} = \beta \times \text{SASA} + \gamma$ . Here, SASA is the solvent-accessible surface area (SASA) of each given molecule, which was computed using the MOLSURF<sup>54</sup> module in AMBER9 with a solvent probe radius of 1.4 Å. The parameter  $\beta$  was set to 0.0072 kcal/(mol·Å<sup>2</sup>), and the parameter  $\gamma$  was set to zero according to the default setting by AMBER9.

For each complex, a total of 126 snapshots were retrieved from the last 1000 ps segment on the MD trajectory with an interval of 8 ps. According to the MM-GB(PB)/SA method, the arithmetic averages of the internal energies and solvation free energies of each species in eq 1 were computed based on this configurational ensemble. The configurational entropy term ( $TS_{\text{conf}}$ ) of each molecule was computed by using the NMODE<sup>55</sup> module in AMBER9. Since this term was time-consuming to compute, it was computed in our study based on a total of 21 snapshots for instead, which were retrieved from the last 1000 ps segment of the MD trajectory with an even interval of 50 ps. Structural minimization as well as normal-mode analysis were carried out with a distance-dependent dielectric function ( $\epsilon = 4R$ ) to mimic the impact of solvent.

**2.5. Calculation of Binding Constants by X-Score.** The X-Score scoring function was also applied to the computation of the binding affinities of all 45 PTP1B complexes. X-Score

is a typical empirical scoring function,<sup>26</sup> which combines terms accounting for van der Waals interaction, hydrogen bonding, hydrophobic effect, and deformation effect. It can be expressed conceptually as

$$S_{\text{binding}} = C_1 \times S_{\text{vdw}} + C_2 \times S_{\text{H-bond}} + C_3 \times S_{\text{hydrophobic}} + C_4 \times S_{\text{deformation}} + C_0 \quad (5)$$

The X-Score program (version 1.2) actually provides three scoring functions, namely HPscore, HMscore, and HSscore. They share the same “master equation” above but differ in the algorithm for computing the hydrophobic effect term. These algorithms consider hydrophobic pairwise contacts between the complex (HPscore), hydrophobic matching between the complex (HMscore), and hydrophobic molecular surface buried upon binding (HSscore), respectively. The binding scores computed by X-Score are given in the  $-\log K_d$  units, which can be converted to standard binding free energies using the following equation

$$\Delta G_{\text{bind}}^0 = 2.303 \times RT \log K_d \quad (6)$$

where  $R$  is the ideal gas constant and  $T$  is the temperature. In our computation, all parameters in X-Score were set as the default values. Technically, X-Score only needs a single configuration of the given complex structure as input. Nevertheless, in order to make a head-to-head comparison, X-Score was also applied to the same configurational ensembles produced by MD simulations for applying MM-GB/SA and MM-PB/SA. Both sets of results will be reported in this manuscript.

### 3. RESULTS AND DISCUSSION

**3.1. Acceleration of MD Simulations for the Application of MM-GB(PB)/SA.** As mentioned earlier in the *Methods* section, the MM-GB(PB)/SA method computes the average free energy change in protein–ligand binding over a configurational ensemble. Ideally, this configurational ensemble should be sampled around a quasi-equilibrium state. In practice, a lengthy MD simulation on the given protein–ligand complex structure is typically employed for this purpose. Snapshots used in subsequent MM-GB(PB)/SA computations are retrieved from the resulting MD trajectory only after the complex structure becomes stable. The adequate length of the MD simulation for establishing this quasi-equilibrium state, however, cannot be predicted in prior. It seems that the only practical approach for determining this parameter is trial-and-error. Since MD simulation is definitely the most time-consuming step for applying MM-GB(PB)/SA, it will be very helpful if this process can be accelerated somehow.

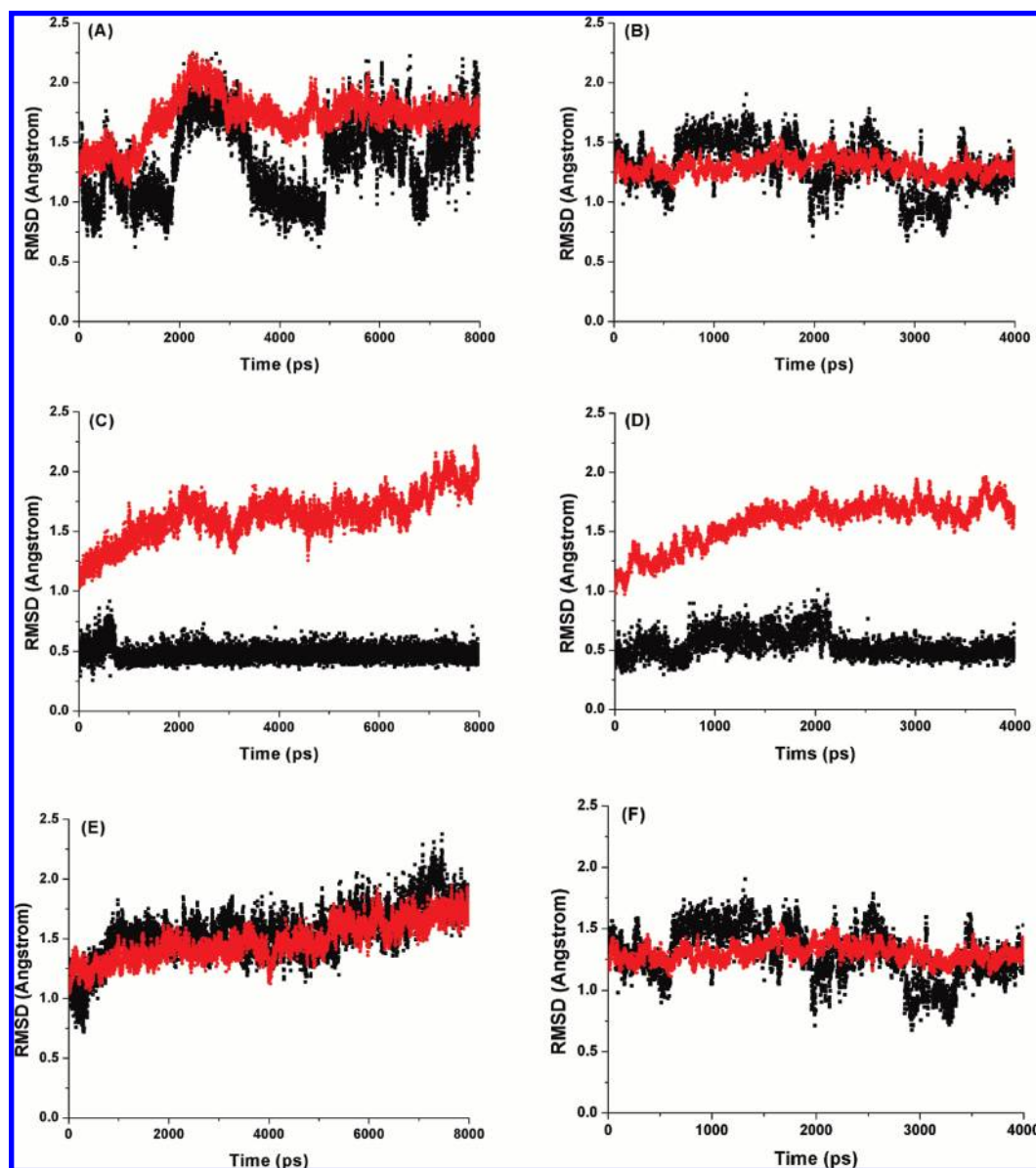
In order to determine the appropriate parameters used in the MD simulations for the PTP1B complexes under our study, we first carried out MD simulations on three PTP1B complexes with known crystal structures, viz., PDB entries 1ONY, 1ONZ, and 1NZ7. Considering that this class of compounds form several characteristic H-bonds with PTP1B (Figure 1), it is reasonable to expect that the electrostatic factor in their direct interactions with PTP1B ( $E_{\text{elec}}$  in eq 3) plays a critical role. Accordingly, a bias on this energy term will hopefully help these PTP1B complex structures reach stability in MD simulations more quickly. Technically, this

can be done by adjusting the “ELEC” parameter used by AMBER9, which defines the weight factor of the electrostatic term in the GAFF force field. Therefore, increasing the value of this parameter will reinforce the contribution of the electrostatic energy term to the total internal energy computed by the GAFF force field.

We actually conducted two sets of MD simulations for each of these three complexes: the first set of simulations was 8-ns long each, in which the “ELEC” parameter was set to 1.0 (the default value); while the second set was 4-ns long each, in which the “ELEC” parameter was set to 1.2. The rmsd values of each complex with respect to its crystal structure in these two sets of MD simulations are illustrated in Figure 2. One can see that when ELEC = 1.0, PDB entries 1NZ7 and 1ONY reached structural stability after 1 ns and 5 ns, respectively, whereas PDB entry 1ONZ did not even after 8 ns. It was thus unclear if MD simulations of 8 ns long would be adequate for refining the structural models of the other 42 PTP1B complexes since they were obviously less reliable than crystal structures. In contrast, when ELEC = 1.2, both 1NZ7 and 1ONY maintained an acceptable structural stability along almost the entire MD trajectory. Even the troublesome 1ONZ reached a structural stability after 2 ns. The snapshot at 2 ns on the MD trajectory of each complex (when ELEC = 1.2) was compared to its crystal structure in Figure 3. One can see that both the overall structure of PTP1B and the binding mode of each given small-molecule inhibitor were well-preserved after MD simulations under such setting. In fact, the overall rmsd values of both the protein and the ligand were below 2.0 Å in all three cases. Visual inspection revealed that the key H-bonds between PTP1B and the inhibitors in these three complexes were also well-preserved.

In order to examine if our approach produced structures valid for MM-GB(PB)/SA computations, we carried out MM-GB/SA computations based on the MD trajectories of these three complexes obtained with ELEC = 1.2. The detailed methods were the same as those given in the *Methods* section. Note that the ELEC parameter was restored to its default value of 1.0 in our MM-GB(PB)/SA computations so that the electrostatic components are at the normal level to produce the correct force field energies. The binding free energies given by our MM-GB/SA computations for 1ONY, 1ONZ, and 1NZ7 were  $-79.89$ ,  $-61.96$ , and  $-92.67$  kcal/mol, respectively. Given that the experimentally measured binding free energies of these three complexes are  $-9.29$ ,  $-6.55$ , and  $-9.77$  kcal/mol, respectively, at least the order of these computed binding free energies was in accordance with experimental data.

Based on all of the above observations, we thus concluded that it was appropriate to set the ELEC parameter to 1.2 in the MD simulations of all 45 PTP1B complexes, and a MD simulation of 2 ns long was adequate for each given complex. Indeed, all 45 PTP1B complex structures established acceptable stability in our MD simulations under such setting. Reducing MD simulations from 8 ns to 2 ns certainly had saved a large amount of CPU time, hard disk space, and human labor. We propose that adjusting the “ELEC” parameter could be an effective approach applicable to other systems as well as for accelerating the MD simulations required by MM-GB(PB)/SA if electrostatic interactions also play a critical role in those protein–ligand binding processes.



**Figure 2.** Root-mean-square deviations (rmsd) observed along the MD trajectories for three PTP1B complexes with known crystal structures. The results of PDB entries 1ONY, 1ONZ, and 1NZ7 when ELEC = 1.0 are given in (A), (C), and (E), while the corresponding results of three structures when ELEC = 1.2 are given in (B), (D), and (F). (In each figure, the red line indicates the rmsd values of PTP1B computed on all C $_{\alpha}$  atoms, while the black line indicates the rmsd values of the inhibitor computed on all heavy atoms. Both sets of rmsd values were computed with respect to the crystal structure.)

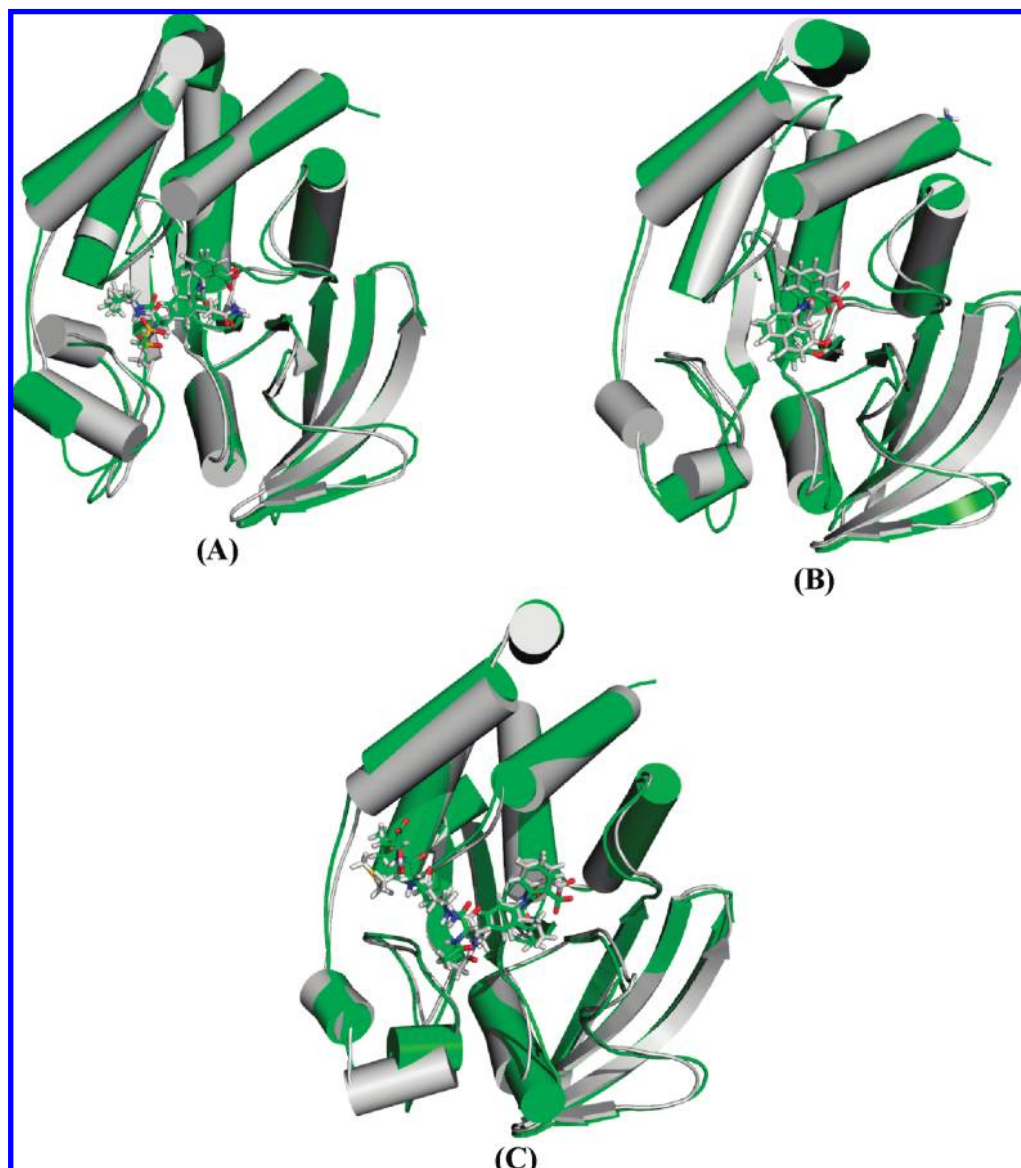
**3.2. Performance of MM-GB/SA and MM-PB/SA.** Both the Poisson–Boltzmann (PB)<sup>50,51</sup> and the General Born (GB)<sup>52</sup> methods are based on the continuum solvent assumption, but they follow different approaches to the computation of the electrostatic component in the solvation free energy. Both methods have been extensively assessed regarding their performance.<sup>56,57</sup> Our study provides a comparison of MM-GB/SA and MM-PB/SA with respect to their capabilities in reproducing protein–ligand binding free energies. Considering that all other energy terms in MM-GB/SA and MM-PB/SA (eqs 2–4) are identical, this comparison actually reflects the relative performance of PB and GB methods in such scenarios.

Tables 4 and 5 itemize our MM-GB/SA and MM-PB/SA results of all 45 PTP1B complexes together with the experimental binding data. If the configurational entropy term is not considered, the binding free energies ( $\Delta G_{\text{MM-GB/SA}}$ ) calculated by MM-GB/SA agree with the experimental values

with a correlation coefficient of 0.87 and a standard deviation of 0.67 kcal/mol; while the binding free energies calculated by MM-PB/SA ( $\Delta G_{\text{MM-PB/SA}}$ ) agree with the experimental values with a correlation coefficient of 0.79 and a standard deviation of 0.83 kcal/mol (Table 6). The scatter plots of the MM-GB/SA and MM-PB/SA results versus experimental binding data are given in Figure 4. One can see that both MM-GB/SA and MM-PB/SA produced acceptable linear responses to the experimental binding data, albeit the former was clearly better than the latter. The intercorrelation between the MM-GB/SA and MM-PB/SA results has a high correlation coefficient of 0.965. The absolute values of the MM-GB/SA and MM-PB/SA results on this class of PTP1B inhibitors are also at exactly the same range.

Our finding indicates that GB is not necessarily less rigorous than PB, although the latter is computationally more complicated. In fact, MM-GB/SA performed even better than MM-PB/SA in our study. Thus, GB at least has some





**Figure 3.** Superimposition of the crystal structure with the snapshot at 2 ns on the MD trajectory of PDB entries 1ONY (A), 1ONZ (B), and 1NZ7 (C) when ELEC = 1.2. In each figure, the crystal structure and the MD snapshot are colored in green and gray, respectively.

technical advantages over PB. The MM-GB/SA method implemented in AMBER9 is an improved version based on the study by Mongan et al.<sup>58</sup> They reported a robust method to correct the computation of molecular volume. The core idea was to introduce a free energy barrier to the separation of nonbonded atoms. Their study indicated that the addition of the correction term yielded protein hydrogen bond length distributions and polypeptide configurational ensembles that were in better agreement with results obtained in explicit solvent than earlier pairwise models. Another observation is that the standard deviation in the MM-GB/SA results for each complex is smaller by 1–2 kcal/mol on average than the counterpart of MM-PB/SA (Tables 4 and 5). This indicates that GB results are less sensitive to the conformational changes of the given molecule than PB, which is another technical advantage of GB.

It should be mentioned that the binding free energies of 45 PTP1B complexes computed by both MM-GB/SA and MM-PB/SA are all much larger in magnitude than the experimental data, although the correlation between two sets of data is acceptable. This is a phenomenon frequently

noticed in our own applications of MM-GB(PB)/SA<sup>43</sup> and similar studies by other researchers.<sup>41,44,46,49</sup> As indicated by eq 2, the free energy of each species, i.e. the protein, the ligand, or the complex, computed by MM-GB(PB)/SA consists of an internal energy term ( $E_{gas}$ ), a configurational entropy term ( $TS_{conf}$ ), and a solvation energy term ( $G_{solv}$ ). The first two terms are computed with a force field, whereas the last term is computed with a GB/SA or PB/SA solvation model. Note that a GB/SA or PB/SA solvation model is normally calibrated independently from any force field. Consequently, it is possible that when combined together, the energies computed by force field and solvation model do not reconcile very well to produce “realistic” energies. Therefore, our opinion is that to reproduce the absolute values of binding free energies may not be the primary aim for applying the MM-GB/SA or MM-PB/SA method to the study of protein–ligand binding.

Interestingly, we also observed that inclusion of the configurational entropy term led to worse results (Table 6): for MM-GB/SA, the correlation coefficient decreases from 0.87 to 0.81; while for MM-PB/SA, the correlation coefficient

**Table 4.** Binding Free Energies Computed by the MM-GB/SA Method for the 45 PTP1B Inhibitors<sup>a</sup>

no.	$\Delta E_{\text{elec}}$	$\Delta E_{\text{vdw}}$	$\Delta G_{\text{GB}}$	$\Delta G_{\text{SA}}^b$	$\Delta G_{\text{elec}}^c$	$\Delta G_{\text{np}}^d$	$\Delta G_{\text{MM-GB/SA}}^e$	$T\Delta S$	$\Delta G_{\text{total}}^f$	$\Delta G_{\text{expt}}^g$
1	-349.65	-25.50	335.98	-4.40	-13.67	-29.90	-43.57 ± 3.89	-24.77 ± 8.78	-18.80	-5.53
2	-375.15	-30.29	361.14	-4.66	-14.01	-34.95	-48.96 ± 4.02	-25.78 ± 3.07	-23.18	-6.34
3	-360.85	-28.32	348.87	-4.53	-11.98	-32.85	-44.83 ± 2.97	-23.08 ± 5.60	-21.75	-6.05
4	-382.53	-29.29	354.63	-4.77	-27.90	-34.06	-61.96 ± 4.30	-20.95 ± 6.05	-41.01	-6.55
5	-354.31	-30.22	342.01	-4.50	-12.30	-34.72	-47.02 ± 3.55	-21.12 ± 5.51	-25.90	-5.67
6	-353.12	-27.63	337.33	-4.56	-15.79	-32.19	-47.98 ± 4.65	-20.33 ± 7.22	-27.65	-5.80
7	-393.21	-26.34	379.71	-4.29	-13.50	-30.63	-44.14 ± 3.41	-24.99 ± 6.11	-19.15	-6.12
8	-350.00	-36.87	340.53	-5.16	-9.47	-42.03	-51.49 ± 3.49	-23.50 ± 4.50	-27.99	-6.34
9	-347.63	-26.62	332.99	-4.41	-14.64	-31.03	-45.67 ± 3.63	-23.63 ± 5.42	-22.04	-6.55
10	-396.42	-29.45	376.19	-4.98	-20.23	-34.43	-54.66 ± 4.91	-22.05 ± 8.63	-32.61	-6.81
11	-383.11	-26.52	357.88	-4.80	-25.23	-31.32	-56.55 ± 3.67	-24.09 ± 5.29	-32.46	-6.66
12	-387.65	-43.00	376.21	-6.89	-11.44	-49.89	-61.33 ± 3.79	-26.26 ± 7.65	-35.07	-8.13
13	-435.98	-43.17	413.99	-7.04	-21.99	-50.21	-72.20 ± 4.37	-26.75 ± 6.67	-45.45	-8.60
14	-425.17	-36.65	398.42	-6.51	-26.75	-43.16	-69.91 ± 5.29	-28.52 ± 9.35	-41.39	-8.13
15	-417.28	-41.95	395.32	-7.24	-21.96	-49.19	-71.15 ± 4.12	-32.47 ± 7.53	-38.68	-7.31
16	-482.60	-40.99	466.00	-6.90	-16.60	-47.89	-64.49 ± 4.04	-33.56 ± 5.98	-30.93	-7.69
17	-413.18	-42.22	397.07	-6.86	-16.11	-49.08	-65.19 ± 4.47	-29.31 ± 4.50	-35.88	-6.54
18	-365.71	-39.39	355.85	-6.40	-9.86	-45.79	-55.66 ± 3.68	-26.77 ± 8.44	-28.89	-6.88
19	-364.58	-39.04	347.43	-6.34	-17.15	-45.38	-62.52 ± 4.01	-29.92 ± 5.88	-32.60	-6.92
20	-413.06	-44.34	386.37	-7.01	-26.69	-51.35	-78.05 ± 4.74	-28.49 ± 5.39	-49.56	-8.18
21	-396.09	-46.29	387.47	-7.31	-8.62	-53.60	-62.22 ± 4.00	-31.34 ± 6.22	-30.88	-8.13
22	-383.17	-41.80	357.40	-6.78	-25.77	-48.58	-74.35 ± 4.10	-26.81 ± 7.73	-47.54	-8.00
23	-369.46	-44.75	353.15	-7.41	-16.31	-52.16	-68.47 ± 4.30	-27.58 ± 6.88	-40.89	-7.82
24	-423.08	-42.58	392.72	-6.95	-30.36	-49.53	-79.89 ± 4.44	-30.38 ± 7.06	-49.51	-9.29
25	-440.33	-45.68	409.69	-7.25	-30.64	-52.93	-83.57 ± 4.49	-32.73 ± 7.20	-50.84	-7.51
26	-455.40	-50.78	417.94	-8.29	-37.46	-59.07	-96.52 ± 4.48	-28.19 ± 9.36	-68.33	-8.08
27	-490.37	-54.37	480.97	-8.98	-9.40	-63.35	-72.74 ± 4.90	-34.57 ± 7.52	-38.17	-9.41
28	-523.99	-46.88	490.69	-8.16	-33.30	-55.04	-88.34 ± 5.35	-39.25 ± 7.79	-49.09	-9.06
29	-509.22	-54.49	480.57	-8.56	-28.65	-63.05	-91.70 ± 5.25	-36.53 ± 6.72	-55.17	-8.74
30	-574.53	-51.54	545.01	-8.53	-29.52	-60.07	-89.59 ± 5.52	-36.26 ± 6.45	-53.33	-8.90
31	-505.43	-50.89	470.92	-8.48	-34.51	-59.37	-93.87 ± 5.10	-35.41 ± 8.86	-58.46	-8.44
32	-544.37	-49.00	517.62	-8.53	-26.75	-57.53	-84.27 ± 4.31	-28.69 ± 8.00	-55.58	-9.50
33	-574.60	-51.22	533.09	-9.07	-41.51	-60.29	-101.80 ± 5.54	-37.29 ± 10.42	-64.51	-9.45
34	-540.58	-57.79	514.96	-9.26	-25.62	-67.05	-92.67 ± 5.54	-36.02 ± 7.53	-56.65	-9.77
35	-531.03	-51.91	503.02	-8.42	-28.01	-60.33	-88.34 ± 4.22	-34.33 ± 6.72	-54.01	-8.69
36	-390.60	-47.76	376.72	-7.91	-13.88	-55.67	-69.56 ± 4.23	-33.56 ± 7.79	-36.00	-8.60
37	-494.87	-54.01	481.28	-8.88	-13.59	-62.89	-76.48 ± 4.28	-37.49 ± 8.11	-38.99	-7.92
38	-461.93	-55.08	425.62	-8.52	-36.31	-63.60	-99.91 ± 4.78	-36.17 ± 8.56	-63.74	-10.63
39	-545.15	-46.70	511.99	-8.23	-33.16	-54.93	-88.08 ± 5.39	-34.61 ± 6.51	-53.47	-10.13
40	-439.73	-49.36	414.51	-8.11	-25.22	-57.47	-82.70 ± 4.18	-31.35 ± 7.15	-51.35	-7.96
41	-462.79	-49.04	427.90	-8.09	-34.89	-57.13	-92.03 ± 4.52	-35.92 ± 6.60	-56.11	-9.45
42	-407.72	-55.24	397.33	-8.63	-10.39	-63.87	-74.26 ± 5.23	-33.69 ± 8.04	-40.57	-8.83
43	-405.12	-59.82	391.82	-8.74	-13.30	-68.56	-81.85 ± 4.39	-33.51 ± 7.22	-48.34	-8.80
44	-405.51	-60.38	404.31	-9.51	-1.20	-69.89	-71.09 ± 4.20	-29.88 ± 5.54	-41.21	-8.58
45	-463.79	-53.42	434.25	-8.71	-29.54	-62.13	-91.68 ± 4.55	-34.49 ± 7.65	-57.19	-10.16

<sup>a</sup> All energies are in kcal/mol units. <sup>b</sup>  $\Delta G_{\text{SA}} = \beta * \text{SASA} + \gamma$ ,  $\beta = 0.0072 \text{ kcal}/(\text{mol} \cdot \text{\AA}^2)$ . <sup>c</sup>  $\Delta G_{\text{elec}} = \Delta E_{\text{elec}} + \Delta G_{\text{GB}}$ . <sup>d</sup>  $\Delta G_{\text{np}} = \Delta E_{\text{vdw}} + \Delta G_{\text{SA}}$ . <sup>e</sup>  $\Delta G_{\text{MM-GB/SA}} = \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}}$ . <sup>f</sup>  $\Delta G_{\text{total}} = \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}} - T\Delta S$ . <sup>g</sup>  $\Delta G_{\text{expt}} = RT \ln K_i$  ( $T = 300 \text{ K}$ ).

decreases from 0.79 to 0.69. This observation is in conflict with many successful cases reported in literature, in which inclusion of the configurational entropy term led to overall improvements. The configurational entropy term computed by normal-mode analysis is another distinctive feature of the MM-GB(PB)/SA method. It is not quite possible for us to provide an in-depth discussion why this term is not helpful in our study. We however would like to point out that computation of configuration entropies, even in gas phase, is a complicated problem which remains largely unsolved. For this matter, we recommend to the readers a nice study published by Gilson et al. recently,<sup>59</sup> in which they discussed the computation of the configurational entropy loss of a small-molecule ligand upon protein binding. Our observation is at least a fair warning to the concept that inclusion of this configurational entropy term is always desirable in the application of MM-GB(PB)/SA.

When we analyzed the energy terms in the MM-GB/SA and MM-PB/SA results, we found that the nonpolar component in the solvation free energy ( $G_{\text{nonpolar}}$  in eq 4) was much smaller than the electrostatic component ( $G_{\text{elec}}$  in eq 4). In fact,  $\Delta G_{\text{SA}}$  is smaller than 5% of either  $\Delta G_{\text{PB}}$  or  $\Delta G_{\text{GB}}$  in magnitude for all of the PTP1B complexes under our study (Tables 4 and 5). Considering this, the difference between the solvation free energies of two molecules will be dominated solely by the electrostatic component, which is somewhat contradictory to commonsense. Therefore, we believe that the default parameters used in computing the nonpolar component in solvation energy have underestimated the contribution of this component ( $G_{\text{nonpolar}} = \beta \times \text{SASA} + \gamma$ ). Those parameters ( $\beta$  and  $\gamma$ ) were typically derived from GB/SA or PB/SA analyses of the solvation free energies of simple small-molecule organic compounds<sup>54</sup> and thus may not be applicable to protein–ligand binding events. A simple

**Table 5.** Binding Free Energies Computed by the MM-PB/SA Method for the 45 PTP1B Inhibitors<sup>a</sup>

no.	$\Delta E_{\text{elec}}$	$\Delta E_{\text{vdw}}$	$\Delta G_{\text{PB}}$	$\Delta G_{\text{SA}}^b$	$\Delta G_{\text{elec}}^c$	$\Delta G_{\text{np}}^d$	$\Delta G_{\text{MM-PB/SA}}^e$	$T\Delta S$	$\Delta G_{\text{total}}^f$	$\Delta G_{\text{expt}}^g$
1	-349.65	-25.50	334.37	-4.40	-15.28	-29.90	-45.18 ± 5.55	-24.77 ± 8.78	-20.41	-5.53
2	-375.15	-30.29	366.45	-4.66	-8.70	-34.95	-43.65 ± 6.10	-25.78 ± 3.07	-17.87	-6.34
3	-360.85	-28.32	344.36	-4.53	-16.49	-32.85	-49.34 ± 4.05	-23.08 ± 5.60	-26.26	-6.05
4	-382.53	-29.29	353.91	-4.77	-28.62	-34.06	-62.68 ± 4.93	-20.95 ± 6.05	-41.73	-6.55
5	-354.31	-30.22	340.72	-4.50	-13.59	-34.72	-48.31 ± 4.73	-21.12 ± 5.51	-27.19	-5.67
6	-353.12	-27.63	332.93	-4.56	-20.19	-32.19	-52.37 ± 7.44	-20.33 ± 7.22	-32.04	-5.80
7	-393.21	-26.34	375.45	-4.29	-17.76	-30.63	-48.40 ± 3.89	-24.99 ± 6.11	-23.41	-6.12
8	-350.00	-36.87	338.12	-5.16	-11.88	-42.03	-53.90 ± 5.08	-23.50 ± 4.50	-30.40	-6.34
9	-347.63	-26.62	329.63	-4.41	-18.00	-31.03	-49.04 ± 6.31	-23.63 ± 5.42	-25.41	-6.55
10	-396.42	-29.45	368.57	-4.98	-27.85	-34.43	-62.28 ± 5.10	-22.05 ± 8.63	-40.23	-6.81
11	-383.11	-26.52	349.71	-4.80	-33.40	-31.32	-64.72 ± 4.59	-24.09 ± 5.29	-40.63	-6.66
12	-387.65	-43.00	373.84	-6.89	-13.81	-49.89	-63.70 ± 6.48	-26.26 ± 7.65	-37.44	-8.13
13	-435.98	-43.17	413.99	-7.04	-21.99	-50.21	-72.20 ± 5.74	-26.75 ± 6.67	-45.45	-8.60
14	-425.17	-36.65	389.45	-6.51	-35.72	-43.16	-78.87 ± 6.66	-28.52 ± 9.35	-50.35	-8.13
15	-417.28	-41.95	392.15	-7.24	-25.13	-49.19	-74.32 ± 5.03	-32.47 ± 7.53	-41.85	-7.31
16	-482.60	-40.99	464.10	-6.90	-18.50	-47.89	-66.39 ± 5.89	-33.56 ± 5.98	-32.83	-7.69
17	-413.18	-42.22	399.82	-6.86	-13.36	-49.08	-62.43 ± 6.84	-29.31 ± 4.50	-33.12	-6.54
18	-365.71	-39.39	356.62	-6.40	-9.09	-45.79	-54.89 ± 6.84	-26.77 ± 8.44	-28.12	-6.88
19	-364.58	-39.04	339.59	-6.34	-24.99	-45.38	-70.37 ± 4.50	-29.92 ± 5.88	-40.45	-6.92
20	-413.06	-44.34	389.79	-7.01	-23.27	-51.35	-74.63 ± 5.02	-28.49 ± 5.39	-46.14	-8.18
21	-396.09	-46.29	387.92	-7.31	-8.17	-53.60	-61.77 ± 7.39	-31.34 ± 6.22	-30.43	-8.13
22	-383.17	-41.80	352.41	-6.78	-30.76	-48.58	-79.34 ± 4.41	-26.81 ± 7.73	-52.53	-8.00
23	-369.46	-44.75	356.32	-7.41	-13.14	-52.16	-65.30 ± 5.17	-27.58 ± 6.88	-37.72	-7.82
24	-423.08	-42.58	384.75	-6.95	-38.33	-49.53	-87.86 ± 5.26	-30.38 ± 7.06	-57.48	-9.29
25	-440.33	-45.68	400.55	-7.25	-39.78	-52.93	-92.71 ± 5.44	-32.73 ± 7.20	-59.98	-7.51
26	-455.40	-50.78	410.40	-8.29	-45.00	-59.07	-104.07 ± 4.95	-28.19 ± 9.36	-75.88	-8.08
27	-490.37	-54.37	485.61	-8.98	-4.76	-63.35	-68.10 ± 7.87	-34.57 ± 7.52	-33.53	-9.41
28	-523.99	-46.88	491.92	-8.16	-32.07	-55.04	-87.10 ± 7.06	-39.25 ± 7.79	-47.85	-9.06
29	-509.22	-54.49	479.11	-8.56	-30.11	-63.05	-93.16 ± 5.87	-36.53 ± 6.72	-56.63	-8.74
30	-574.53	-51.54	553.27	-8.53	-21.26	-60.07	-81.33 ± 8.68	-36.26 ± 6.45	-45.07	-8.90
31	-505.43	-50.89	466.57	-8.48	-38.86	-59.37	-98.22 ± 5.89	-35.41 ± 8.86	-62.81	-8.44
32	-544.37	-49.00	521.10	-8.53	-23.27	-57.53	-80.80 ± 5.49	-28.69 ± 8.00	-52.11	-9.50
33	-574.60	-51.22	531.94	-9.07	-42.66	-60.29	-102.94 ± 6.86	-37.29 ± 10.42	-65.65	-9.45
34	-540.58	-57.79	511.43	-9.26	-29.15	-67.05	-96.19 ± 8.46	-36.02 ± 7.53	-60.17	-9.77
35	-531.03	-51.91	500.71	-8.42	-30.32	-60.33	-90.65 ± 5.36	-34.33 ± 6.72	-56.32	-8.69
36	-390.60	-47.76	380.72	-7.91	-9.88	-55.67	-65.55 ± 5.50	-33.56 ± 7.79	-31.99	-8.60
37	-494.87	-54.01	479.39	-8.88	-15.48	-62.89	-78.37 ± 4.28	-37.49 ± 8.11	-40.88	-7.92
38	-461.93	-55.08	422.63	-8.52	-39.30	-63.60	-102.91 ± 5.35	-36.17 ± 8.56	-66.74	-10.63
39	-545.15	-46.70	509.20	-8.23	-35.95	-54.93	-90.88 ± 7.45	-34.61 ± 6.51	-56.27	-10.13
40	-439.73	-49.36	407.69	-8.11	-32.04	-57.47	-89.52 ± 4.81	-31.35 ± 7.15	-58.17	-7.96
41	-462.79	-49.04	418.78	-8.09	-44.01	-57.13	-101.15 ± 5.04	-35.92 ± 6.60	-65.23	-9.45
42	-407.72	-55.24	405.79	-8.63	-1.93	-63.87	-65.80 ± 6.56	-33.69 ± 8.04	-32.11	-8.83
43	-405.12	-59.82	395.38	-8.74	-9.74	-68.56	-78.30 ± 5.29	-33.51 ± 7.22	-44.79	-8.80
44	-405.51	-60.38	415.69	-9.51	10.18	-69.89	-59.71 ± 7.36	-29.88 ± 5.54	-29.83	-8.58
45	-463.79	-53.42	432.98	-8.71	-30.81	-62.13	-92.94 ± 6.02	-34.49 ± 7.65	-58.45	-10.16

<sup>a</sup> All energies are in kcal/mol units. <sup>b</sup>  $\Delta G_{\text{SA}} = \beta * \text{SASA} + \gamma$ ,  $\beta = 0.0072 \text{ kcal}/(\text{mol} \cdot \text{\AA}^2)$ . <sup>c</sup>  $\Delta G_{\text{elec}} = \Delta E_{\text{elec}} + \Delta G_{\text{PB}}$ . <sup>d</sup>  $\Delta G_{\text{np}} = \Delta E_{\text{vdw}} + \Delta G_{\text{SA}}$ . <sup>e</sup>  $\Delta G_{\text{MM-PBSA}} = \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} + \Delta G_{\text{SA}} + \Delta G_{\text{PB}}$ . <sup>f</sup>  $\Delta G_{\text{total}} = \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} + \Delta G_{\text{SA}} + \Delta G_{\text{PB}} - T\Delta S$ . <sup>g</sup>  $\Delta G_{\text{expt}} = RT \ln K_i$  ( $T = 300 \text{ K}$ ).

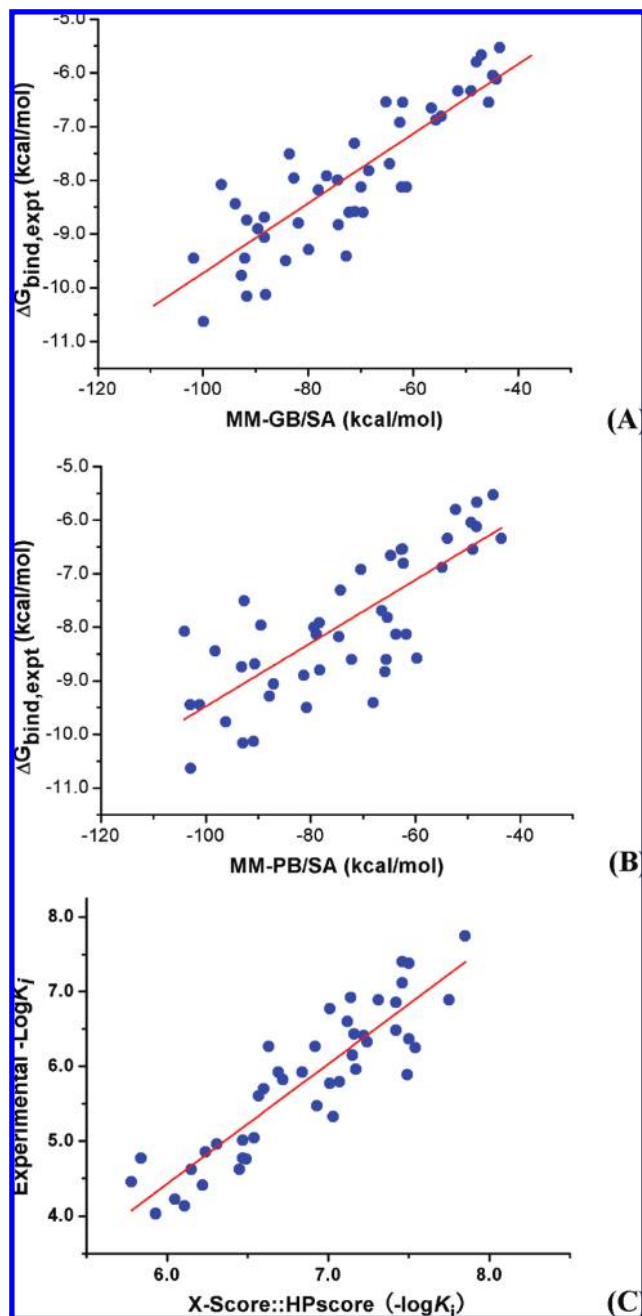
**Table 6.** Correlations between Experimental and Calculated Binding Data Produced by Different Scoring Methods

scoring methods	based on the configurational ensembles from MD simulations		based on the complex structures prepared by molecular docking	
	$R^a$	SD <sup>b</sup>	$R$	SD
MM-GB/SA	0.87	0.67	0.65	1.02
MM-GB/SA+TS	0.81	0.78	0.50	1.16
MM-PB/SA	0.79	0.83	0.62	1.04
MM-PB/SA+TS	0.69	0.97	0.45	1.20
X-Score: HPscore	0.89	0.60	0.85	0.70
X-Score: HScore	0.87	0.66	0.81	0.79
X-Score: HMscore	0.83	0.75	0.71	0.94

<sup>a</sup> Pearson correlation coefficient between the computed binding free energies and the experimental values. <sup>b</sup> Standard deviations (in kcal/mol) in fitting the experimental binding constants and the computed binding scores.

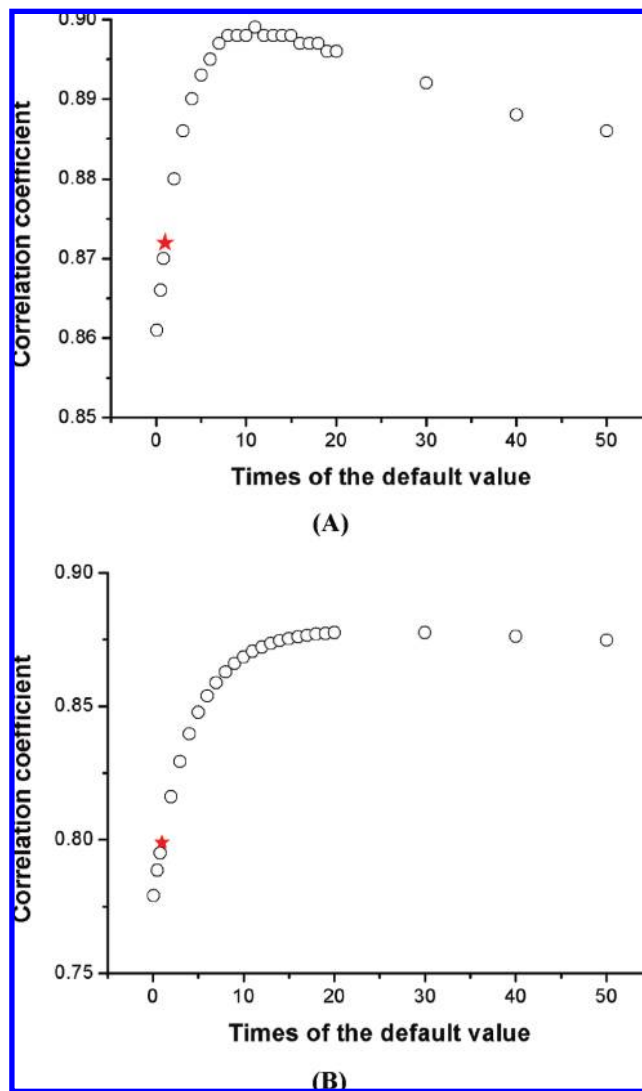
remedy for this problem is to increase the value of parameter  $\beta$  to enhance the contribution of the nonpolar term, which was also proposed in one of our previous applications of the MM-PB/SA method.<sup>43</sup> In this study, we changed the

value of  $\beta$  systematically from 1/10 to 50 times the default value in both our MM-GB/SA and MM-PB/SA computations. The correlation coefficients between computed binding free energies and experimental data as a function of the value



**Figure 4.** Correlations between the experimental binding data of 45 PTP1B complexes and the binding data computed by MM-GB/SA (A), MM-PB/SA (B), and X-Score::HPscore (C) on the same set of configurational ensembles obtained through MD simulations.

of  $\beta$  are given in Figure 5. One can see that increasing the value of  $\beta$  will indeed result in more promising results for both MM-GB/SA and MM-PB/SA. It seems that the optimal value of  $\beta$  for computing the binding affinities of this class of PTP1B complexes is between 10 and 20 times the default value, i.e. around 0.072–0.144 kcal/(mol·Å<sup>2</sup>). Note that  $\Delta G_{SA}$  will be basically comparable to either  $\Delta G_{PB}$  or  $\Delta G_{GB}$  in magnitude after applying this optimized value. Therefore, it may not be a coincidence. In our previous study of FKBP12 inhibitors,<sup>43</sup> the optimal value of this parameter was found to be six times the default value. Although the optimal value of  $\beta$  is not consistent between our previous study and this study, we can still come to the conclusion that the default value of  $\beta$  certainly underestimates the nonpolar component



**Figure 5.** Correlations between the experimental and the computed binding free energies by MM-GB/SA (A) and MM-PB/SA (B) as a function of the  $\beta$  parameter.

in solvation free energy. It is worthwhile to explore the optimal value of this parameter in the application of the MM-GB(PB)/SA method.

**3.3. MM-GB/SA versus Scoring Function.** Scoring functions are another class of computational methods for evaluating protein–ligand binding, which have become popular in structure-based drug design since 1990s.<sup>60–63</sup> Unlike MM-GB(PB)/SA, a scoring function only needs a single snapshot of the given protein–ligand complex structure to perform computation. Due to their fast speed, scoring functions are especially suitable for high-throughput jobs, such as virtual screening. In our work, scoring functions were employed in combination with molecular docking program to prepare the structural models of PTP1B-inhibitor complexes. We also examined the performance of the X-Score scoring function in computing the binding affinities of the PTP1B inhibitors under our study, since we found in our validation of docking/scoring tools that X-Score produced promising results in this aspect on a test set of 26 diverse PTP1B complexes. An intriguing comparison of the performance of X-Score and MM-GB(PB)/SA, representing two major classes of methods for binding affinity computation at present, is thus possible.



Our first set of computation was conducted on the PTP1B complex structures prepared by molecular docking, i.e. the starting point for subsequent MD simulation. Each complex structure was minimized to convergence with the SANDER module in the AMBER9 program before being subjected to MM-GB/SA and MM-PB/SA computation. The correlation coefficients ( $R$ ) between the experimentally measured binding free energies of 45 PTP1B complexes and the computed values given by MM-GB/SA and MM-PB/SA in this test were 0.65 and 0.62, respectively (Table 6). This level of accuracy is clearly less promising than the one produced by either MM-GB/SA or MM-PB/SA on the configurational ensembles obtained through MD simulations. This observation indicates that the rough structural models of PTP1B complexes prepared by molecular docking need to be further refined by MD simulations before they can be subjected to MM-GB/SA or MM-PB/SA computations. When X-Score was applied to the same set of PTP1B complex structures, all three options in X-Score produced acceptable correlations between the experimental binding data and the computed binding scores. The best results were given by HPScore, yielding a correlation coefficient ( $R$ ) of 0.85 and a standard deviation ( $SD$ ) of 0.70 kcal/mol in fitting two sets of data (Table 6). This level of accuracy is certainly better than the one achieved by MM-GB/SA and MM-PB/SA on the same set of complex structures and is actually comparable to the ones achieved by MM-GB/SA and MM-PB/SA on configurational ensembles. This is a clear indication that X-Score is not sensitive to small uncertainties in protein–ligand complex structures. This type of robustness is much welcome in practice.

In our second set of computation, X-Score was applied to the same configurational ensemble of each PTP1B complex used in our MM-GB/SA and MM-PB/SA computations, and its final result was computed as the arithmetic average over the results obtained on all snapshots. As compared to the results obtained on the complex structures prepared by molecular docking, all three options in X-Score produced more or less improved results (Table 6). The best results were also given by HPScore, yielding a correlation coefficient ( $R$ ) of 0.89 and a standard deviation ( $SD$ ) of 0.60 kcal/mol in fitting two sets of data (Figure 4C). This level of accuracy is in fact comparable or even slightly better than the one produced by MM-GB/SA or MM-PB/SA on the same set of PTP1B complex structures. It is encouraging to obtain these improved results produced by X-Score although this is not the typical way for applying a scoring function. In addition, applying scoring function to a configurational ensemble has the potential advantage of getting more converged prediction of the protein–ligand binding affinity.

It needs to be mentioned in particular that in a recent study, Martinez et al. compared the performance of MM-PB/SA and X-Score on seven XIAP-peptide complexes through the same approach, i.e. applying scoring function to a configurational ensemble.<sup>42</sup> They concluded based on their results that “both methods are able to predict the experimental binding free energy with acceptable errors, but if one needs to identify slight differences upon binding, MM-PB/SA performs better, although X-Score is not a bad choice taking into account the low computational cost of this method”. In our study, we tested both MM-GB(PB)/SA and X-Score on a much larger and more diverse set of PTP1B inhibitors.

Interestingly, our results indicate that the accuracy of X-Score is at least comparable to MM-GB(PB)/SA besides its clear advantage in terms of computational cost. Therefore, our opinion is that whether to adopt a force field-based method like MM-GB(PB)/SA or a scoring function for binding affinity prediction is still case-dependent, which needs careful validation before hand.

**3.4. Analysis of the Structure-Affinity Relationship of This Class of Compounds.** Now we will analyze the structure–activity relationship of this set of compounds according to the outcomes of our simulations and computations. Our discussion below will be based primarily on our MM-GB/SA results on these compounds because it is possible to summarize the contributions from the nonpolar factors ( $\Delta G_{np}$ ) and the polar factors ( $\Delta G_{elec}$ ) in protein–ligand binding from the MM-GB(PB)/SA results and also because MM-GB/SA performed slightly better than MM-PB/SA in our study. One may wonder why this analysis is still necessary given the fact that the binding affinities of this set of compounds have been measured through experimental means. Note that the primary aim of our study is to establish reliable computational methods for modeling PTP1B inhibitors. If such methods can be validated on this set of compounds, they may be applied to the design of new PTP1B inhibitors in future studies.

Our discussion will continue by dividing the chemical structures of this class of PTP1B inhibitors into three major parts (see Figure 1 for example). (i) The common oxalamino benzoic acid moiety resides at the active site, which consists of two critical carboxylic groups: one forms hydrogen bonds with three amide groups on the backbone of PTP1B, while the other forms a salt bridge with the Arg221 residue. This moiety largely mimics the phosphate group of the native substrate. (ii) The  $R_4$  group occupies the peripheral site. If this group contains an appropriate hydrogen bond acceptor group it may form favorable interactions with the characteristic Arg24 and Arg254 at this site. (iii) A “linker” is also needed to connect the two parts mentioned above. This linker group may also form specific interaction with the Asp48 residue nearby.

As indicated in Table 3, compounds **1–11** lack the linker group as well as the  $R_4$  group. Apparently, they were designed to explore the role of  $R_1$ . The major difference among them is the substituent on the aromatic ring. According to our complex models,  $R_1$  resides mainly at the active site. The aromatic ring forms  $\pi$ – $\pi$  stacking with Tyr46 and Phe182. For example, the binding affinity of compound **2** ( $K_i = 24 \mu\text{M}$ ) is higher than those of **6** ( $K_i = 60 \mu\text{M}$ ) and **1** ( $K_i = 93 \mu\text{M}$ ). Our computation reveals that the  $\Delta G_{np}$  of **2** is lower than those of **6** and **1** by 2.76 and 5.05 kcal/mol, respectively. Apparently, choosing such an aromatic group to maximize the hydrophobic interaction with PTP1B is a good strategy to increase binding affinity.

Compared to the compounds in the first group, compounds **12–24** all have a complete linker moiety. These compounds were designed mainly to explore the role of different  $R_2$ ,  $R_3$ , and  $R_4$  groups. Binding affinities of this group of compounds scatter in a relatively narrow range: most of them exhibit  $K_i$  values between 1 and 10  $\mu\text{M}$ . Our structural models of the complexes formed by these compounds reveal that  $R_2$  and  $R_3$  orient toward the outside rather than the inside of the

binding pocket and therefore do not have a decisive impact on the binding affinities of this class of compounds.

Compounds **25–45** all have the same  $R_2$  and  $R_3$  group, which differ only in the composition of the  $R_4$  group. Compounds **25–37** were designed to explore the appropriate length of the  $R_4$  fragment, all having a characteristic carboxylic group. The carboxylic group is directed toward the Arg24 and Arg254 residues. In particular, compounds **34** and **37** differ only in the configuration of the chiral carbon linking to the carboxylic group (the former *S* and the latter *R*). Our MM-GB/SA computation reveals that  $\Delta G_{elec}$  of **34** is lower than that of **37** by 12.03 kcal/mol. We found by examining the MD trajectories of these two compounds that the relevant carboxylic group on **34** formed H-bond contacts with both Arg24 and Arg254 for most of time, whereas the one on **37** only formed H-bonds contacts with the former. Apart from compounds **25–37**, compounds **38–45** all have a quite different  $R_4$  group. The carboxylic and hydroxyl groups on the phenyl ring are both able to interact with Arg24 and Arg254. Compound **38** has another hydroxyl as compared to **40**, and the  $\Delta G_{elec}$  value is also lower than that of **40** by 11.09 kcal/mol, which confirms the contribution of this hydroxyl to the binding affinity of the compounds in this group. If the ester group is changed to an amide group (**38** vs **42**), that is, a hydrogen acceptor is replaced by a hydrogen donor,  $\Delta G_{elec}$  becomes more positive by 25.92 kcal/mol, which is consistent with the change in the experimental binding data. Based on the above results, one can see that choosing an appropriate hydrogen bond acceptor to maximize the interactions with Arg24 and Arg254 is critical for achieving the binding affinities of compounds **25–45**.

It needs to be emphasized that compounds **25–45** cover both binding sites 1 and 2 nicely, accounting for their obviously higher binding affinities over compounds **1–11**. According to our MM-GB/SA computations, the average  $\Delta G_{MM-GB/SA}$  of compounds **1–11** is  $-49.71$  kcal/mol, whereas the average  $\Delta G_{MM-GB/SA}$  of compounds **25–45** reaches  $-86.24$  kcal/mol. Thus, the idea of fragment-based design is successfully validated by our computational results. Furthermore, the average polar contribution ( $\Delta G_{elec}$ ) of compounds **1–11** is  $-16.25$  kcal/mol, while that of compounds **25–45** is  $-25.56$  kcal/mol. In contrast, the average nonpolar contribution ( $\Delta G_{np}$ ) of compounds **1–11** is  $-33.46$  kcal/mol, while that of compounds **25–45** is  $-60.68$  kcal/mol. This indicates that additional nonpolar interactions contribute more than polar interactions to the enhanced binding affinities of the compounds in the second group.

#### 4. CONCLUSIONS

We have studied the binding affinities of a set of 45 OBA-containing small-molecule inhibitors to PTP1B. In our study, the complex structures of these compounds were prepared through molecular docking by using the GOLD program and subsequent molecular dynamics simulations by using the AMBER program. Their binding affinities were computed with MM-GB/SA, MM-PB/SA, and the X-Score scoring function. Note that the total number of compounds considered in our study is larger than those in most of the MM-GB(PB)/SA studies reported so far. The chemical structures of these compounds are fairly diverse, and their binding affinities span nearly 4 orders of magnitude. This set of compounds thus

represents a challenge for computational analysis. MM-GB/SA yielded a very promising correlation between its results and the experimental binding data of these 45 PTP1B inhibitors with a correlation coefficient of 0.87 and a standard deviation of 0.60 kcal/mol, whereas the performance of MM-PB/SA was slightly inferior. Several aspects of the MM-GB(PB)/SA method were also explored in our study. Our results demonstrate the following: (i) The preparative MD simulations for the application of MM-GB(PB)/SA can be effectively accelerated by adjusting the “ELEC” parameter in AMBER. (ii) The weight factor of the solvent accessible surface area-based nonpolar component in both GB/SA and PB/SA solvation energies would better be increased to 10–20 times of its default value to obtain optimized results. (iii) Inclusion of the configurational entropy term does not necessarily lead to improved results.

Interestingly, X-Score produced equally good results as MM-GB/SA on the complex structures prepared simply through molecular docking. Application of X-Score to the same configuration ensembles used in MM-GB(PB)/SA computation led to even better results. Our results suggest that one can obtain a quick estimation of the binding affinities of PTP1B inhibitors with X-Score. If necessary, one can do MD sampling and then apply either MM-GB/SA or X-Score to obtain more accurate predictions.

As for the structure–activity relationship of this set of compounds, their interactions with PTP1B are characterized by the hydrogen bonds with several key residues, including Arg24, Asp48, Ala217, Gly218, Ile219, Arg221, and Arg254. Our MM-GB(PB)/SA results prompt that as long as these polar interactions are preserved, the difference among the binding affinities of this set of compounds is largely determined by nonpolar interactions. In summary, we have successfully validated a set of computational tools for modeling this set of PTP1B inhibitors, which are hopefully applicable to other classes of PTP1B inhibitors as well.

#### ACKNOWLEDGMENT

The authors are grateful to the financial supports from the Chinese National Natural Science Foundation (Grants No. 20502031, No. 20772149, and No. 90813006), the Chinese Ministry of Science and Technology (the 863 high-tech project, Grant No. 2006AA02Z337), and the Science and Technology Commission of Shanghai Municipality (Grants No. 06PJ14115 and No. 074319113).

**Supporting Information Available:** Structural models of the 26 PTP1B complexes for validation, the complexed model prepared by molecular docking, and the last MD snapshot for each of the 45 PTP1B inhibitors, together with all of the input parameter files for running the AMBER9 program. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

#### REFERENCES AND NOTES

- (1) Hunter, T. Signaling-2000 and Beyond. *Cell* **2000**, *100*, 113–127.
- (2) Tonks, N. K.; Neel, B. G. From Form to Function: Signaling by Protein Tyrosine Phosphatases. *Cell* **1996**, *87*, 365–368.
- (3) Li, L.; Dixon, J. E. Form, function, and regulation of protein tyrosine phosphatases and their involvement in human diseases. *Semin. Immunol.* **2000**, *12*, 75–84.



- (4) Ukkola, O.; Santaniemi, M. Protein tyrosine phosphatase 1B: a new target for the treatment of obesity and associated co-morbidities. *J. Int. Med.* **2002**, *251*, 467–475.
- (5) Puius, Y. A.; Zhao, Y.; Sullivan, M.; Lawrence, D. S.; Almo, S. C.; Zhang, Z. Y. Identification of a second aryl phosphate-binding site in protein-tyrosine phosphatase 1B: A paradigm for inhibitor design. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13420–13425.
- (6) Andersen, H. S.; Iversen, L. F.; Jeppesen, C. B.; Branner, S.; Norris, K.; Rasmussen, H. B.; Møller, K. B.; Møller, N. P. H. 2-(Oxalylamino)-Benzoic Acid Is a General, Competitive Inhibitor of Protein-tyrosine Phosphatases. *J. Biol. Chem.* **2000**, *275*, 7101–7108.
- (7) Iversen, L. F.; Andersen, H. S.; Møller, K. B.; Olsen, O. H.; Peters, G. H.; Branner, S.; Mortensen, S. B.; Hansen, T. K.; Lau, J.; Ge, Y.; Holsworth, D. D.; Newman, M. J.; Peter, N.; Møller, H. Steric Hindrance as a Basis for Structure-Based Design of Selective Inhibitors of Protein-Tyrosine Phosphatases. *Biochemistry* **2001**, *40*, 14812–14820.
- (8) Liu, G.; Xin, Z.; Liang, H.; Abad-Zapatero, C.; Hajduk, P. J.; Janowick, D. A.; Szczepankiewicz, B. G.; Pei, Z.; Hutchins, C. W.; Ballaron, S. J.; Stashko, M. A.; Lubben, T. H.; Berg, C. E.; Rondinone, C. M.; Trevillyan, J. M.; Jirousek, M. R. Selective Protein Tyrosine Phosphatase 1B Inhibitors: Targeting the Second Phosphotyrosine Binding Site with Non-Carboxylic Acid-Containing Ligands. *J. Med. Chem.* **2003**, *46*, 3437–3440.
- (9) Xin, Z.; Oost, T. K.; Abad-Zapatero, C.; Hajduk, P. J.; Pei, Z.; Szczepankiewicz, B. G.; Hutchins, C. W.; Ballaron, S. J.; Stashko, M. A.; Lubben, T.; Trevillyan, J. M.; Jirousek, M. R.; Liu, G. Potent, selective inhibitors of protein tyrosine phosphatase 1B. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1887–1890.
- (10) Liu, G.; Szczepankiewicz, B. G.; Pei, Z.; Janowick, D. A.; Xin, Z.; Hajduk, P.; Abad-Zapatero, C.; Liang, H.; Hutchins, C. W.; Fesik, S. W.; Ballaron, S. J.; Stashko, M. A.; Lubben, T.; Mika, A. K.; Zinker, B. A.; Trevillyan, J. M.; Jirousek, M. R. Discovery and Structure-Activity Relationship of Oxalylarylamino benzoic Acids as Inhibitors of Protein Tyrosine Phosphatase 1B. *J. Med. Chem.* **2003**, *46*, 2093–2103.
- (11) Szczepankiewicz, B. G.; Liu, G.; Hajduk, P. J.; Abad-Zapatero, C.; Pei, Z.; Xin, Z.; Lubben, T. H.; Trevillyan, J. M.; Stashko, M. A.; Ballaron, S. J.; Liang, H.; Huang, F.; Hutchins, C. W.; Fesik, S. W.; Jirousek, M. R. Discovery of a Potent, Selective Protein Tyrosine Phosphatase 1B Inhibitor Using a Linked-Fragment Strategy. *J. Am. Chem. Soc.* **2003**, *125*, 4087–4096.
- (12) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.
- (13) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (14) Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. Modeling Water Molecules in Protein-Ligand Docking Using GOLD. *J. Med. Chem.* **2005**, *48*, 6504–6515.
- (15) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
- (16) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411–428.
- (17) Goodsell, D. S.; Olson, A. J. Automated docking of substrates to proteins by simulated annealing. *Proteins: Struct., Funct., Genet.* **1990**, *8*, 195–202.
- (18) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 293–304.
- (19) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
- (20) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (21) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
- (22) Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graphics Modell.* **2003**, *21*, 289–307.
- (23) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225–242.
- (24) Zhou, Z.; Felts, A. K.; Friesner, R. A.; Levy, R. M. Comparative Performance of Several Flexible Docking Programs and Scoring Functions: Enrichment Studies for a Diverse Set of Pharmaceutically Relevant Targets. *J. Chem. Inf. Model.* **2007**, *47*, 1599–1608.
- (25) Wijnand, T. M.; Mooij, M. L. V. General and targeted statistical potentials for protein-ligand interactions. *Proteins* **2005**, *61*, 272–287.
- (26) Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11–26.
- (27) Wang, R.; Lu, Y.; Wang, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, *46*, 2287–2303.
- (28) Wang, R.; Lu, Y.; Fang, X.; Wang, S. An Extensive Test of 14 Scoring Functions Using the PDBbind Refined Set of 800 Protein-Ligand Complexes. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2114–2125.
- (29) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind Database: Collection of Binding Affinities for Protein-Ligand Complexes with Known Three-Dimensional Structures. *J. Med. Chem.* **2004**, *47*, 2977–2980.
- (30) Wang, R.; Fang, X.; Lu, Y.; Yang, C.; Wang, S. The PDBbind Database: Methodologies and Updates. *J. Med. Chem.* **2005**, *48*, 4111–4119.
- (31) The SYBYL software (version 7.3), Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri 63144, U.S.A.
- (32) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. AMBER 9; University of California: San Francisco, 2006.
- (33) Wang, J. M.; Kollman, P. A. Automatic parameterization of force field by systematic search and genetic algorithms. *J. Comput. Chem.* **2001**, *22*, 1219–1228.
- (34) Jakalian, A.; Bruce, L. B.; David, B. J.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132–146.
- (35) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.
- (36) William, L. J.; Jayaraman, C.; Jeffry, D. M.; Roger, W. I.; Michael, L. K. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (37) Yang, L.; Tan, C.; Hsieh, M. J.; Wang, J.; Duan, Y.; Cieplak, P.; Caldwell, J.; Kollman, P. A.; Luo, R. New-Generation Amber United-Atom Force Field. *J. Phys. Chem. B* **2006**, *110*, 13166–13176.
- (38) Tom, D.; Darrin, Y.; Lee, P. Particle mesh Ewald: An N(logN) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (39) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (40) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889–897.
- (41) Wang, W.; Lim, W. A.; Jakalian, A.; Wang, J.; Wang, J. M.; Luo, R.; Bayly, C. T.; Kollman, P. A. An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis. *J. Am. Chem. Soc.* **2001**, *123*, 3986–3994.
- (42) Obiol, P. C.; Rubio, M. J. Comparative Evaluation of MMPBSA and XSCORE to Compute Binding Free Energy in XIAP-Peptide Complexes. *J. Chem. Inf. Model.* **2007**, *47*, 134–142.
- (43) Xu, Y.; Wang, R. X. A computational analysis of the binding affinities of FKBP12 inhibitors using the MM-PB/SA method. *Proteins* **2006**, *64*, 1058–1068.
- (44) Wang, W.; Kollman, P. A. Computational study of protein specificity: The molecular basis of HIV-1 protease drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14937–14942.
- (45) Laitinen, T.; Kankare, J. A.; Perakyla, M. Free energy simulations and MM-PBSA analyses on the affinity and specificity of steroid binding to antiestradiol antibody. *Proteins* **2004**, *55*, 34–43.
- (46) Fogolari, F.; Moroni, E.; Wojciechowski, M.; Baginski, M.; Ragona, L.; Molinari, H. MM/PBSA analysis of molecular dynamics simulations of bovine beta-lactoglobulin: Free energy gradients in conformational transitions. *Proteins* **2005**, *59*, 91–103.
- (47) Laitinen, T.; Rouvinen, J.; Perakyla, M. MM-PBSA free energy analysis of endo-1,4-xylanase II (XynII)-substrate complexes: binding

- of the reactive sugar in a skew boat and chair conformation. *Org. Biomol. Chem.* **2003**, *1*, 3535–3540.
- (48) Zou, H. J.; Luo, C.; Zheng, S. X.; Luo, X. M.; Zhu, W. L.; Chen, K. X.; Shen, J. H.; Jiang, H. L. Molecular insight into the interaction between IFABP and PA by using MM-PBSA and alanine scanning methods. *J. Phys. Chem. B* **2007**, *111*, 9104–9113.
- (49) Wang, J. M.; Morin, P.; Wang, W.; Kollman, P. A. 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.* **2001**, *123*, 5221–5230.
- (50) Lamm, G. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Larter, R., Cundari, T. R., Eds.; John Wiley & Sons Inc.: NJ, 2003; Vol. 19, pp 147–365.
- (51) Baker, N. A. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Larter, R., Cundari, T. R., Eds.; John Wiley & Sons Inc.: NJ, 2005; Vol. 21, pp 349–379.
- (52) Bashford, D.; Case, D. A. Generalized born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129–152.
- (53) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.
- (54) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
- (55) Kottalam, J.; Case, D. A. Langevin modes of macromolecules: Applications to crambin and DNA hexamers. *Biopolymers* **1990**, *29*, 1409–1421.
- (56) Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; BrooksIII, C. L. Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J. Comput. Chem.* **2004**, *25*, 265–284.
- (57) Rizzo, R. C.; Aynechi, T.; Case, D. A.; Kuntz, I. D. Estimation of Absolute Free Energies of Hydration Using Continuum Methods: Accuracy of Partial Charge Models and Optimization of Nonpolar Contributions. *J. Chem. Theory Comput.* **2006**, *2*, 128–139.
- (58) Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A. Generalized Born Model with a Simple, Robust Molecular Volume Correction. *J. Chem. Theory Comput.* **2007**, *3*, 156–169.
- (59) Chang, C. A.; Chen, W.; Gilson, M. K. Ligand configurational entropy and protein binding. *Proc. Nat. Acad. Sci.* **2007**, *104*, 1534–1539.
- (60) Muegge, I. and Rarey, M. Small molecule docking and scoring. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH Inc.: NJ, 2001; Vol. 17, pp 1–60.
- (61) Böhm, H. J.; Stahl, M. The use of scoring functions in drug discovery applications. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH Inc.: NJ, 2002; Vol. 18, pp 41–88.
- (62) Gohlke, H.; Klebe, G. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644–2676.
- (63) Schulz-Gasch, T.; Stahl, M. Scoring functions for protein-ligand interactions: a critical perspective. *Drug Discovery Today Technol.* **2004**, *1*, 231–239.

CI8004429