Validation and Use of the MM-PBSA Approach for Drug Discovery

Bernd Kuhn,* Paul Gerber, Tanja Schulz-Gasch, and Martin Stahl

Molecular Design, Pharmaceutical Division, F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland

Received November 15, 2004

The MM-PBSA approach has become a popular method for calculating binding affinities of biomolecular complexes. Published application examples focus on small test sets and few proteins and, hence, are of limited relevance in assessing the general validity of this method. To further characterize MM-PBSA, we report on a more extensive study involving a large number of ligands and eight different proteins. Our results show that applying the MM-PBSA energy function to a single, relaxed complex structure is an adequate and sometimes more accurate approach than the standard free energy averaging over molecular dynamics snapshots. The use of MM-PBSA on a single structure is shown to be valuable (a) as a postdocking filter in further enriching virtual screening results, (b) as a helpful tool to prioritize de novo design solutions, and (c) for distinguishing between good and weak binders ($\Delta pIC_{50} \geq 2-3$), but rarely to reproduce smaller free energy differences.

Introduction

Common tasks in structure-based computational drug design involve (a) virtual screening of large databases to select smaller subsets for biological testing, (b) de novo design of novel scaffolds to generate new, unpatented starting points for medicinal chemistry, and (c) library design or, on a smaller scale, substituent prioritization for a predefined scaffold. Due to the large number of molecules that often need to be processed in relatively short time, fast docking and scoring methods are routinely used for these tasks.[1,2] A common conclusion from numerous studies involving different scoring functions is that they are accurate enough to yield substantial statistical enrichment over random or diverse compound or substituent selection[3−5] but are not suited to rank-order structurally related compounds.[6] Numerous reasons can be made responsible for failures of docking studies. Among these are unconsidered solvation effects, unpredictable induced protein fit, deficiencies of the energy function, and ambiguity in protonation states.[2]

More rigorous theoretical approaches that partly address these issues are available, but they are hampered by their often large computational cost. These methods include free energy perturbation,[7] ligand interaction energy approach,[8,9] λ-dynamics,[10] chemical Monte Carlo/molecular dynamics,[11] ligand interaction scanning,[12] or MM-PBSA.[13] As almost any larger pharmaceutical company has recently invested in considerable computing power, in form of Linux clusters or distributed GRID computing,[14] these more CPU-intensive methods might become more interesting in the future. In this study, we focus on the MM-PBSA free energy approach and try to assess its usefulness in a pharmaceutical environment.

MM-PBSA is an acronym for a method developed by Kollman and Case that combines molecular mechanical (MM) energies, a continuum solvent Poisson−Boltzmann (PB) model for polar solvation, and a solvent-accessible surface area (SA) dependent nonpolar solvation term.[13] The total free energy of a biomolecular system is expressed as a sum of these energy contributions plus an additional solute entropy term, $-TS_{solute}$:

$$G = E_{MM} + G_{PB} + G_{SA} - TS_{solute} \quad (1)$$

The solute entropy is usually approximated by a combination of classical statistical expressions and normal-mode analysis. The sum of molecular mechanical energies, E_{MM}, can be further divided into contributions from electrostatic (E_{es}), van der Waals (E_{vdW}), and internal energies (E_{int}):

$$E_{MM} = E_{es} + E_{vdW} + E_{int} \quad (2)$$

Using eqs 1 and 2, the binding free energy of a noncovalent association, ΔG_{bind}, can be computed as:

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) \quad (3)$$

The coordinates for the complex are typically obtained in an explicit solvent environment by simple minimization of the structure or by generation of an ensemble of molecular conformations using molecular dynamics (MD). In the following, we will distinguish these two variants as MM-PBSA (single, relaxed structure) and MD-PBSA (molecular dynamics ensemble). In both cases, the energies of the uncomplexed reactants are obtained by turning off the interactions between the atoms of the protein and ligand. Several successful examples of the MM(D)-PBSA approach for the calculation of relative free energies of macromolecules and complex binding free energies have been reported.[15−20] However, an evaluation of this method on a larger scale involving a number of proteins of pharmaceutical interest is still missing and is the topic of this study.

Within this evaluation, we examine the use of the MM(D)-PBSA approach for several common tasks in structure-based computational drug design, namely for the ranking of related ligands binding to the same

* To whom correspondence should be addressed. Phone: +41 61 6889773. Fax: +41 61 6886459. e-mail: bernd.kuhn@roche.com.

10.1021/jm049081q CCC: $30.25
Published on Web 05/10/2005

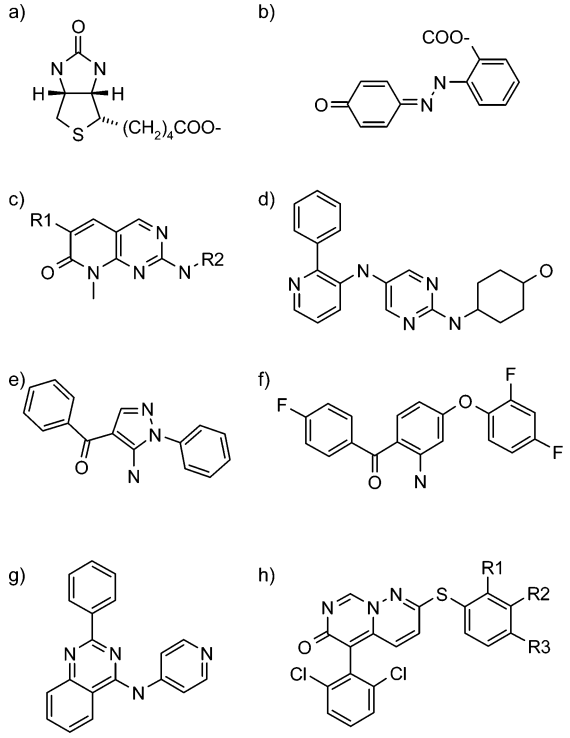

Figure 1. Chemical structures of representative ligands and scaffolds used in this study: (a) biotin, (b) 2-(4′-hydroxyazobenzene)benzoic acid (HABA), (c) 2-amino-8-methyl-8*H*-pyrido[2,3-*d*]pyrimidin-7-one scaffold, (d) representative of 2-aminopyrimidine scaffold, (e) representative of 3-amino-4-ketopyrazole scaffold, (f) representative of 2-ketoaniline scaffold, (g) representative of 4-aminoquinazoline scaffold, (h) 1,7,8a-triazanaphthalen-6-one scaffold. The structures of the seven biotin derivatives[16] and sixteen 1,7,8a-triazanaphthalen-6-one ligands[6] have been published elsewhere.

protein site, as a postdocking filter for the virtual screening of compound collections, and as a tool to prioritize de novo design suggestions. We compare this method, both applying it to a single minimized structure and to molecular dynamics ensembles, with standard docking/scoring and with a molecular mechanics energy function including distance-dependent dielectric screening (MM-RDIEL). The last approach is a very simple representation of solvent effects and was recently shown to perform as well as more accurate solvation models in recognizing near-native binding configurations among a set of decoys and in ranking binding affinities of protein−ligand complexes.[21] Details of our implementation of MM(D)-PBSA and of the computational procedures are given in the following section. The performances of the different theoretical approaches for a panel of eight proteins are discussed in the Results and Discussion section.

Methods

(a) Preparation of Target Systems. Ranking of Related Ligands. We applied MM(D)-PBSA to three different test sets involving the proteins avidin (8 ligands) and p38 MAP kinase (12 and 16 ligands, respectively) and additionally compared the results to those obtained from docking runs using FlexX and the ScreenScore energy function[3] as well as to MM-RDIEL. Experimental and calculated binding data for all three test systems can be found in the Supporting Information. First, the binding of avidin to biotin (Figure 1a) and to six biotin derivatives (for structures see previous publication[16]) as well as to 2-(4′-hydroxyazobenzene)benzoic acid (HABA, Figure 1b) were compared. Starting structures were generated from an avidin−biotin complex (PDB: 1avd)[22] by manual modeling of the six biotin-like ligands and from an avidin−HABA complex.[23]

In a second test set, we looked at twelve different Roche p38 MAP kinase inhibitors of a 2-amino-8-methyl-8*H*-pyrido-[2,3-*d*]pyrimidin-7-one series (Figure 1c) and from other classes for which representative structures are shown in Figure 1d−g. Starting structures for these complexes were obtained by manual modeling using the X-ray structure of the p38/Sb218655 complex (PDB: 1bmk).[24]

Third, we looked at sixteen different 1,7,8a-triaza-naphthalen-6-one ligands (Figure 1h) that had been used in a study by Pearlman and Charifson.[6] Structural modifications within this scaffold are small with R1−R3 covering: H, CH_3, Br, Cl, F, OH, NH_2. Using the p38 MAP kinase apo structure of Wilson et al. (PDB: 1wfc),[25] the common scaffold was docked into the active site of this crystal conformation and the substituents R1−R3 attached. As the ortho- and meta-substituted phenyl rings (R1 or R2 ≠ H) can be oriented in two ways in the binding pocket, a second binding conformation was taken into account in which this ring was flipped by 180° against the remainder of the molecule. For these molecules, the lower of the two calculated MM(D)-PBSA free energies was used in the analysis. The exact substitution patterns of all molecules have been published previously.[6]

As we would like to assess the prediction accuracy for ranking different inhibitors, we use the predictive index (PI) measure, which is a sum over all pairwise ligand comparisons, introduced by Pearlman:[6]

$$\mathrm{PI} = \sum_{j>i}\sum_{i} w_{ij}C_{ij} / \sum_{j>i}\sum_{i} w_{ij} \tag{4}$$

with

$$w_{ij} = |E(j) - E(i)| \tag{5}$$

and

$$C_{ij} = \begin{cases} 1 \text{ if } [E(j) - E(i)]/[P(j) - P(i)] < 0 \\ -1 \textit{ if } [E(j) - E(i)]/[P(j) - P(i)] > 0 \\ 0 \text{ if } [P(j) - P(i)] = 0 \end{cases} \tag{6}$$

E(i) is the experimental binding free energy and *P(i)* is the calculated energy value of a ligand *i*, hence, w_{ij} is a weighting term that is proportional to the difference between the experimental values of the two ligands that are compared. A predictive index of +1 indicates perfect prediction for a test set, −1 arises from predictions that are always wrong, and 0 reflects a completely random prediction.

Virtual Screening. To assess the performance of MM(D)-PBSA as a filter after docking of compound libraries, we use an experiment in which a random subset of the WDI database is seeded with known actives for a panel of seven proteins. The panel comprises a wide range of binding sites from closed, lipophilic (cyclooxygenase-2, COX-2) to solvent-exposed, polar (neuraminidase) characteristics. Details of the ligand collections and protein structures for each target have been reported previously.[3] Explicit water molecules bridging protein−ligand interactions were included for gyrase (Hoh231 next to Asp73) and thrombin (Hoh247 adjacent to Tyr228 in the S1 pocket) and were treated as part of the receptor in all subsequent calculations. For each of the seven targets the seeded library was docked into the corresponding binding site using the docking engine FRED[26] and the scoring function ChemScore,[27,28] which has demonstrated a good overall performance for this protein panel.[4] From each of the docking runs, the top 200 solutions were retained, resulting in a reduced number of actives: COX-2 (42), estrogen receptor (27), p38 MAP kinase (23), gyrase (7), thrombin (36), gelatinase A (9), neuraminidase (12). After running MM-PBSA and MD-PBSA calculations on

Table 1. Validation of the MAB* and GAFF Force Fields for Test Sets of Conformational Energies and Intermolecular Interaction Energies[a]

	MMFF94s	MM3*	CVFF	MAB*	GAFF
Conformational Energies I (19 comparisons)[38]					
rmsd	0.74	0.78	2.86	**1.70**	**0.50**
max. dev.	1.45	2.27	5.78	**3.58**	**0.87**
Conformational Energies II (37 comparisons)[37]					
rmsd	0.38	0.72	2.36	**1.93**	**1.17**
max. dev.	0.99	2.57	6.11	**4.27**	**2.88**
Intermolecular Interaction Energies (66 comparisons) [37]					
rmsd	0.76	3.81	4.36	**2.09**	–
max. dev.	2.52	13.31	16.39	**7.17**	–

[a] The data for the MMFF94s, MM3*, and CVFF force fields for these test sets are taken from the literature.[37] rmsd stands for the root-mean-square deviation of the calculated energies relative to the experimental or ab initio reference values and max. dev. denotes the maximal deviation. The force fields used in this study are shown in bold. No nonbonded cutoff and a dielectric constant of 1 were used.

these systems, as described below, enrichment curves were calculated for each method. For the multiconformer MM-PBSA calculations the docking poses were clustered using a root-mean-square distance (rmsd) of 2.0 Å and the three highest scoring poses were retained after clustering.

De Novo Design. Using the de novo design program Skelgen,[29,30] 400 different molecules were generated for the target COX-2. The Skelgen run parameters, protein structure, and pharmacophore constraints within the binding site have been described in a previous publication.[31] Starting from the binding modes suggested by Skelgen, MM-PBSA energies were calculated for all 400 molecules and their ranking determined. To find out whether MM-PBSA is able to prioritize the Skelgen suggestions, we modeled known COX-2 ligands into the binding site and compared them to the generated solutions.

(b) Molecular Simulations and MM(D)-PBSA Calculations. Ligand parameters and charges were determined with the *antechamber* module of Amber 7[32] based on the general atom force field (GAFF)[33] and the AM1-BCC charge scheme.[34,35] Since for approximately 10% of all ligands no complete GAFF parameter set could be generated, we devised an alternative, second parametrization in which parameters for bonds, valence, dihedral, and improper angles were taken from the MAB force field[36] and combined with the van der Waals parameters from GAFF and AM1-BCC charges for electrostatic interactions. The MAB and GAFF force field expressions for the internal degrees of freedom are similar enough to warrant such a combination. This new force field, MAB*, is designed to deliver parameters for all organic ligands. To validate the MAB* parametrization, we compared calculated conformational energy differences and hydrogen-bond interaction energies from the standard test sets of Halgren[37] and Gundertofte et al.[38] with other common force fields. As can be seen from Table 1, the MAB* force field shows medium accuracy for both conformational energies and intermolecular interactions. The less complete GAFF reveals a very good performance for conformational energies comparable to the best force field in this study, MMFF94s. Additional GAFF validation data for reproducing crystal structures as well as intermolecular and conformational energies have been published previously.[33] We combined these two small molecule force fields with the parm94 protein force field.[39] For gelatinase A, the nonbonded Zn parameters ($R^* = 1.1$ Å; $\epsilon = 0.0125$ kcal/mol) were taken from the parm99 force field.

All protein−ligand complexes were immersed in a 24 Å sphere of TIP3P water[40] and neutralized by adding counterions. Using the *sander_classic* module of AMBER 6,[41] the solvated systems were minimized for 1000 steps with a flexible belly region of 12 Å and a nonbonded cutoff of 12 Å. The relaxed structures were used as input for the MM-PBSA and MM-RDIEL calculations. To obtain MD-PBSA free energies, we equilibrated each system with molecular dynamics for 150 ps and saved snapshots every 5 ps for additional 50 ps, yielding an ensemble of 10 structures each. All MD simulations were performed at $T = 300$ K with a time step of 1.5 fs and the SHAKE algorithm.[42] For the ranking of related ligands, we further investigated the influence of minimizing the MD snapshots (1000 steps) prior to free energy averaging and distinguish this variant as MD-PBSA* from the conventional MD-PBSA approach.

Complex binding free energies for a single structure (MM-PBSA) or averaged over several structures (MD-PBSA, MD-PBSA*) were calculated from the difference in MM-PBSA free energies for the complex and uncomplexed reactants according to eqs 1−3. The molecular mechanical energy, E_{MM}, was determined with the *anal* module of AMBER with no cutoff for nonbonded interactions. The PB calculation was done with *solvate* of the MEAD program package[43] with interior and exterior dielectric constants of 1 and 80, respectively. We used electrostatic focusing with a final grid spacing of 0.25 Å extending 50% beyond the dimensions of the ligands. G_{SA} was calculated from

$$G_{SA} = 0.00542 \text{ kcal}/(\text{mol}\cdot\text{A}^2) \times \text{SA} + 0.92 \text{ kcal/mol} \quad (7)$$

using the surface area estimation of the in-house program XSAE. To estimate the solute entropy change upon association, $-TS_{solute}$, we used the *mfebd* module of MOLOC. This involved the use of the standard classical statistical expressions for translational and rotational entropies and normal-mode analysis for the vibrational contribution.[15] For the latter calculation, receptor atoms were fixed outside of 6 Å of the bound ligand. We performed solute entropy averaging for six snapshots while the other MD-PBSA free energy terms of eq 1 were averaged over 10 snapshots.

For the calculation of MM-RDIEL binding free energies, we used eqs 1−3 with the solvation term G_{PB} replaced by an additional distance-dependent dielectric function, $\epsilon(r) = 4r$, as part of the electrostatic component E_{es}. This method was applied to single, minimized complex structures using the MAB* force field.

Results and Discussion

Ranking of Related Ligands. Accurately predicting the rank-ordering of homologous series of molecules binding to the same protein site has been a long standing challenge for computational chemistry.[44] Due to the relatively small number of ligands that are typically involved, many of the previous MM(D)-PBSA application examples have focused on this task.[18] Using our previously published example of ranking binding affinities of biotin derivatives and HABA to avidin as a test case,[16] we first investigate whether a single-structure MM-PBSA calculation and our modified MD-PBSA protocol are useful alternatives for the considerably longer earlier MD-PBSA procedures. We then extend our comparison to two p38 MAP kinase test systems in which the ligands span a smaller range of binding affinities than in the avidin case.

The avidin example is the least challenging problem as there are only eight ligands and their complexation energies cover a wide range, from −4.5 to −20.4 kcal/mol, i.e. $\Delta\Delta G_{bind} = 15.9$ kcal/mol. The results in Table 2 reveal that very high PI values between 0.8−0.9 can be obtained for MM(D)-PBSA, which are considerably higher than the docking study using FlexX/ScreenScore (PI = 0.49) but slightly lower than the force field approach with a simple dielectric screening function (PI = 0.99). Only minor differences are detectable between the two force fields MAB* and GAFF, and a slight additional improvement is achieved when minimizing

Table 2. Performance Comparison in Ranking Similar Ligands for Three Different Test Sets Using MM-PBSA, MD-PBSA, and MM-RDIEL with the MAB* and GAFF Force Fields[a]

	MM-PBSA MAB*	MD-PBSA MAB*	MD-PBSA* MAB*	MM-PBSA GAFF	MD-PBSA GAFF	MM-RDIEL MAB*	FlexX ScreenScore
				Avidin (8 ligands)			
PI	0.81	0.80	0.91	0.90	0.86	0.99	0.49
R^2	0.63	0.68	0.87	0.87	0.73	0.59	0.24
				p38 MAP Kinase (test set II, 12 ligands)			
PI	0.53	0.31	0.27	0.51	0.37	−0.06	0.31
R^2	0.21	0.11	0.07	0.20	0.17	0.00	0.12
				p38 MAP Kinase (test set III, 16 ligands)			
PI	0.27	0.31	−0.22	−0.01	0.15	−0.14	−0.10
R^2	0.04	0.12	0.01	0.01	0.03	0.03	0.00

[a] MD-PBSA* indicates free energy averaging over minimized MD snapshots. PI stands for the predictive index defined in eqs 4−6, and R^2 is the correlation coefficient between experiment and computation. A previous MD-PBSA study on the avidin system using a considerably more demanding computational setup yielded PI = 0.99 and R^2 = 0.92.[16]

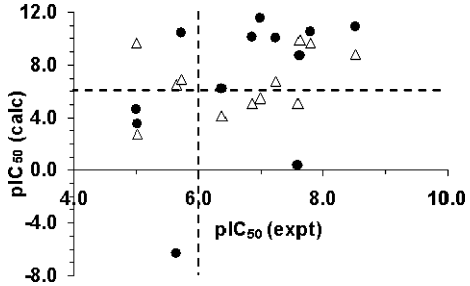

Figure 2. Graphical representation of calculated vs experimental pIC_{50} for twelve Roche p38 MAP kinase inhibitors (test set II). We compare MM-PBSA using the MAB* force field (filled circles) with FlexX/ScreenScore (empty triangles). The calculated binding free energies are shifted by their average differences to the experimental values, i.e. −5.1 kcal/mol and +28.6 score units for MM-PBSA (MAB*) and FlexX, respectively, and converted to pIC_{50} values. Dashed lines at pIC_{50} = 6.0 indicate a threshold of IC_{50} = 1 μM.

molecular dynamics snapshots prior to free energy sampling. The previous study[16] on this system involving a much longer MD simulation time, wider free energy sampling, and manual force field parametrization yielded PI = 0.99, slightly higher than our best MM-PBSA results. This small difference suggests that our optimized automated procedure for running MM(D)-PBSA is a good substitute for the more time-consuming earlier setups.

In the second test set, we are comparing different p38 MAP kinase inhibitors (Figure 1c−g) that bind to the ATP binding site and cover a range of ΔpIC_{50} = 3.5 (pIC_{50} = 5.0−8.5) corresponding to a smaller span in free energy of binding of $\Delta\Delta G_{bind}$ = 4.8 kcal/mol compared to the avidin example. Table 2 contains the numerical results and Figure 2 illustrates the performance of MM-PBSA (MAB*) vs FlexX/ScreenScore graphically. We find a similar behavior in the PI values of the MAB* and GAFF force fields and even a slight degradation when using molecular dynamics compared to the single structure results. Docking with FlexX/ScreenScore performs comparably well with a PI of 0.31 but MM-RDIEL shows a worse than random prediction capability (PI = −0.06). The poorer performance of MD-PBSA vs MM-PBSA in this case is interesting, as in the previous applications free energy averaging over an ensemble of molecular dynamics snapshots was usually employed. We will discuss the MD-PBSA performance in more detail in the next section in which we see a similar behavior.

In the lead optimization phase of a typical drug discovery project usually a screening cascade exists in which several thresholds have to be overcome in order for a compound to be considered of further interest. Figure 2 shows such a threshold at a value of pIC_{50} = 6.0, meaning that compounds with binding affinities of IC_{50} < 1 μM are further followed up. Using this threshold, the graph of Figure 2 can be divided into four regions, in which the upper left and lower right areas represent false positives and false negatives, respectively. With this qualitative measure the docking calculation would yield a total of seven wrong predictions (three false positives, four false negatives) while MM-PBSA (MAB*) would result in only two wrong predictions (one false positive and negative each). Our results indicate that single-structure MM-PBSA is able to distinguish the four weak from the eight strong binders relatively well. In this example, further differentiation of smaller binding energy differences in the range of 1−2 kcal/mol seems not possible with this method.

The last test set involves 16 ligands of a congeneric series (Figure 1h) of p38 MAP kinase binders in which the structural modifications are small, differing by at most one heteroatom. The pIC_{50} values range from 5.7−7.4, corresponding to a free energy difference of only $\Delta\Delta G_{bind}$ = 2.3 kcal/mol, which is not surprising given the small molecular differences. As expected, PI numbers are small for this problem and peak at only 0.3 with slightly higher values when free energy averaging over unrelaxed MD snapshots is used. The MAB* force field shows a better performance than GAFF and is superior to both FlexX/ScreenScore docking and the MM-RDIEL approach, but the prediction capability for selecting substituents in advance of chemical synthesis seems of modest use with all methods investigated. In a previous study, it was reported that thermodynamic integration (TI) calculations for this test set reached a predictive index of 0.84. Apparently, for small structural changes conventional TI or related free energy perturbation approaches are still unrivaled.

In summary, we see comparable performances of the two force fields over the three test sets of related ligands, indicating that the MAB* parametrization is a useful substitute for the less complete GAFF. We are not able to consistently improve the prediction accuracy by free energy averaging over molecular dynamics snapshots relative to MM-PBSA, and the minimization of MD snapshots prior to ensemble averaging also does

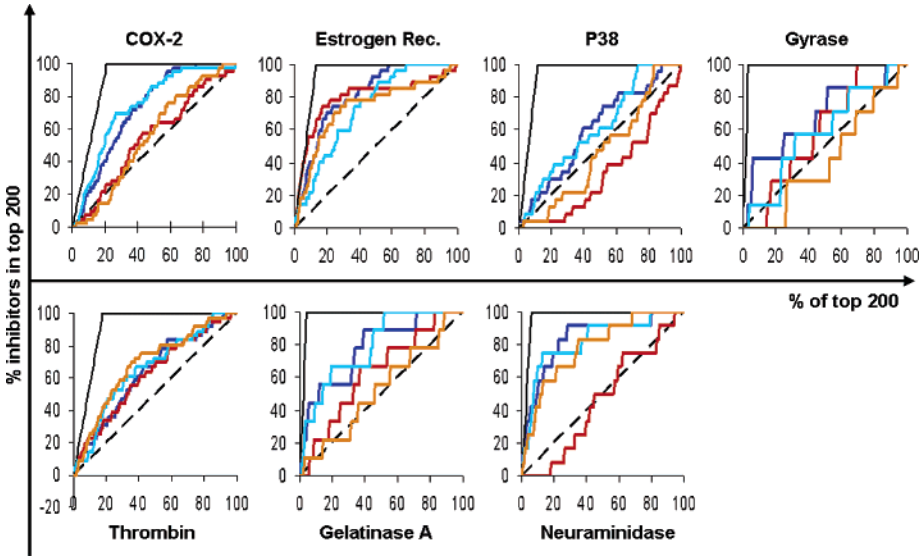

Figure 3. Comparison of virtual screening performance for seven different proteins. The enrichment curves show the percentage of inhibitors retrieved as a function of the percentage of total molecules in the database ($N = 200$). We compare random selection (black, dashed), ideal performance (black, solid), FRED/ChemScore ranking (red), MM-RDIEL ranking with MAB* force field (orange), MM-PBSA ranking with MAB* force field (blue), and MM-PBSA ranking with GAFF force field (cyan).

not lead to better overall predictions. Finally, both the MM-RDIEL and the docking/scoring approaches show poorer performance than MM-PBSA, except for the avidin example in which the force field method with distance-dependent dielectric performs particularly well.

Virtual Screening. Having seen that our MD-PBSA setup and especially the single-structure MM-PBSA calculations provide a fast and comparably accurate alternative for the earlier, more time-consuming setups, we can evaluate these methods on more molecule-extensive drug discovery applications, such as virtual screening. While the use of MM(D)-PBSA for large corporate databases of several hundred thousand compounds is out of reach, its application as a secondary filter after docking might be rewarding. To test the usefulness of such a postdocking filter for virtual screening applications, we have applied MM(D)-PBSA to 200 top-ranking compounds from docking runs against seven protein targets. Figure 3 shows the enrichment curves for docking, MM-RDIEL, and MM-PBSA relative to random and ideal selection. While a previous study for these seven targets has found that docking with FRED/ChemScore is able to obtain significant enrichment factors for a large database of inactive molecules seeded with a small number of actives,[4] little discrimination over random selection can be seen for the top 200 molecules. A better than random performance is achieved for three receptors (estrogen receptor, thrombin, gelatinase A) while the results for p38 kinase are worse. Overall slightly better results than with docking are obtained with the simple solvation model MM-RDIEL especially in the case of neuraminidase. The best overall performance is seen with MM-PBSA leading in all cases to better than random prediction accuracy and improving the rank-ordering for five and four out of seven targets relative to docking and MM-RDIEL, respectively. On average, few differences are seen between MM-PBSA using the MAB* and the GAFF force fields.

How can we explain the better performance of MM-PBSA relative to the other methods? In the case of neuraminidase, the correct binding mode for the inhibitors is generally found by docking. Nevertheless, the ChemScore energy function is not able to further differentiate between true and false inhibitors. A number of the top ranking inactives display direct interactions between nonpolar ligand parts and the hydrophilic binding pocket. These molecules are presumably ranked too high because of unconsidered energy penalties for the desolvation of the mismatched polar protein patches in the scoring function. The big improvement seen with MM-RDIEL supports this analysis and indicates that already a very simple solvation model can alleviate this problem. For p38 kinase, a substantial improvement for MM-PBSA relative to docking is observed. Figure 4 shows that for 11 out of the 23 inhibitors in the docked top 200 molecules the correct binding mode can be found. Again, the more robust MM-PBSA energy function leads to a better ranking. The importance of including a proper treatment of solvation into the energy function is illustrated in Figure 5, which shows on the left side the binding mode of a known p38 kinase inhibitor suggested by FRED/ChemScore docking. While the hydrogen bonding interactions with the kinase hinge region are well reproduced, the placement of the positively charged piperidine substituent into the lipophilic back pocket is obviously energetically unfavorable. With MM-PBSA, this structure drops from rank 7 to rank 122 and obtains a positive binding energy ($\Delta G_{bind} = +7.8$ kcal/mol). Rotating the bond between the pyrimidine and imidazole rings into the correct binding mode (Figure 5b) improves the MM-PBSA result (Rank 57, $\Delta G_{bind} = -2.6$ kcal/mol), as expected for a more robust energy function.

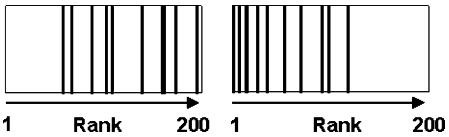

Figure 4. Ranking of p38 MAP kinase inhibitors with correct binding mode ($N = 11$). The diagrams show the ranking obtained with FRED/ChemScore (left) and MM-PBSA (right).

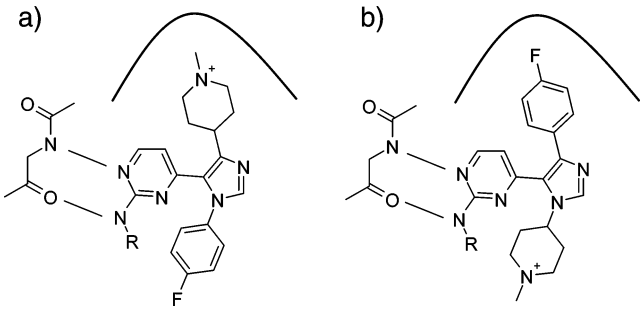

Figure 5. Illustration of handling of solvation effects. (a) Binding mode suggested by FRED/ChemScore (Rank 7); corresponding MM-PBSA result (Rank 122, $\Delta G_{bind} = +7.8$ kcal/mol). (b) X-ray binding mode obtained by rotation around the pyrimidine−imidazole bond (MM-PBSA: Rank 57, $\Delta G_{bind} = -2.6$ kcal/mol). R = CH_2-phenyl.

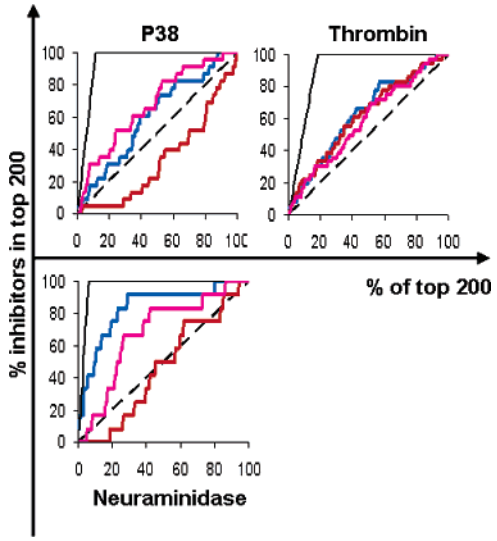

Figure 6. Assessment of multipose MM-PBSA. Shown are random selection (black, dashed), ideal performance (black, solid), FRED/ChemScore ranking (red), MM-PBSA (MAB*) ranking using the top scored docking pose (blue), and MM-PBSA (MAB*) ranking using the three highest scored docking poses (magenta). For axis nomenclature, see Figure 3.

The observation that MM-PBSA is able to better rank the different binding modes for the p38 kinase inhibitor example suggests that an additional improvement in performance might be achieved by applying MM-PBSA calculations not only to rank different ligands but also multiple docking poses. Including the three highest scored poses, after structural clustering (rmsd = 2.0 Å), into the MM-PBSA ranking leads to the enrichment curves of Figure 6. We find higher enrichment relative to the single-pose MM-PBSA results for p38 kinase. However, for both thrombin and neuraminidase, results are unchanged or even worse. The improvement for p38 kinase can be attributed to the fact that for several molecules, as in the example shown above, the correct binding modes are sampled, but not scored highest. In neuraminidase, the correct binding mode is correctly identified by docking for the large majority of the inhibitors and adding additional poses does in most cases not lead to improved binding energies. This is a good sign as it indicates that this method could also be used for the determination of native binding modes.[19] As the MM-PBSA energy function is not perfect, including additional poses for the inactives also increases their

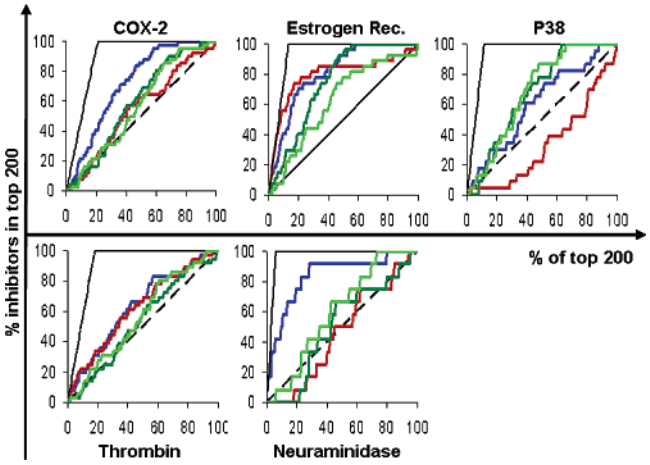

Figure 7. Comparison of MM-PBSA vs MD-PBSA performance for five different proteins. We compare random selection (black, dashed), ideal performance (black, solid), FRED/ChemScore ranking (red), MM-PBSA (MAB*) ranking (blue), MD-PBSA ranking with the MAB force field (green), and MD-PBSA ranking with the GAFF force field (light green). For axis nomenclature see Figure 3.

chance of a better binding energy score, which in other words corresponds to a dilution of the percentage of actives in this data set. Consequently, we see a shift of the best calculated ΔG_{bind} from −1.2 kcal/mol to −14.4 kcal/mol when single-pose and multipose setups are compared, with improved binding free energies occurring predominantly for the inactives. In conclusion, the strategy of multipose MM-PBSA cannot be recommended in general, only if reasons exist that the correct binding mode is contained within the higher scored docking conformations but is not captured with a single pose.

In the previous section about the ranking of related ligands, we have not seen a generally improved performance of MD-PBSA vs MM-PBSA. To extend this comparison to the virtual screening application, we have calculated MD-PBSA results for the five targets for which the statistics is most meaningful, i.e. targets for which the number of known inhibitors in the top 200 list is higher than 10. Figure 7 shows that the MD-PBSA results with the MAB* and GAFF force fields are on average worse than the computationally less demanding MM-PBSA calculations. Only for p38 kinase the results improve slightly while for three other proteins (COX-2, estrogen receptor, neuraminidase) prediction quality drops. Again as in the MM-PBSA comparison, MAB* and GAFF force fields behave rather similar. This is more remarkable in the MD-PBSA application as the two force fields mainly differ in the torsional parameters, which can considerably affect the molecular dynamics trajectory. The reasons for this disappointing behavior of MD-PBSA are manifold. In the case of several estrogen receptor−inhibitor complexes, the use of MD leads to explicit water molecules breaking up polar protein−ligand interactions. As the explicit solvent is stripped from the complex and replaced by the continuum solvent in the subsequent free energy calculation the net interaction energy is reduced. Likewise, in neuraminidase a Na^+-ion is interfering with protein−ligand contacts after the molecular dynamics run. In general, we do not observe that ligands move out of the binding pocket during molecular

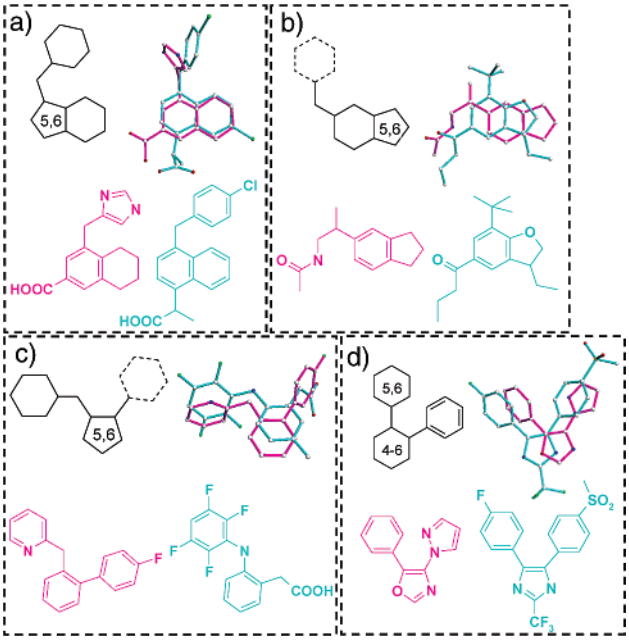

Figure 8. Superimposed pairs of modeled COX-2 ligands (cyan) and molecules designed by Skelgen (magenta). 2D representations of the general inhibitor topologies are displayed in black. Numbers inside the rings indicate that active ligands with alternative ring sizes are known. The MM-PBSA ranking of the Skelgen structures are: (a) 2, (b) 3, (c) 5, (d) 11.

dynamics, which would be an indication of an improper MD setup. It should also be noted that we have rarely observed cases in which MD significantly improves the binding mode of a ligand. In summary, our results indicate that the use of MD introduces additional structural uncertainties, which are often not favorable, and that free energy averaging over a number of different snapshots does not add sufficient accuracy to compensate for these effects. Our molecular dynamics runs are performed in a 24 Å sphere of water with a nonbonded cutoff of 12 Å and for a total of 200 ps. While our setup is of reasonable size we cannot exclude that very long MD simulations on the ns scale with a more rigorous treatment of electrostatics, such as the Particle Mesh Ewald (PME) method,[45] would give drastically improved results. At least the explicit ion artifact in neuraminidase should be remedied as charges are smeared out over the periodic box in PME. However, one should bear in mind that the considerably higher cost in CPU time would prohibit its broader use in pharmaceutical industry, in particular for virtual screening applications.

De Novo Design. De novo design of novel, unpatented chemical scaffolds is an important component of modern drug discovery and has gained attention through algorithmic developments in recent years.[31,46] While existing programs can quickly generate new scaffolds that are consistent with binding site and additional pharmacophore constraints, there remains the challenge of prioritizing the suggestions before synthesis. Using the COX-2 system as an example we look at the capability of MM-PBSA to solve this problem. A large part of the known COX-2 inhibitors can be represented by four distinct topological classes, which are shown in the black 2D depictions of Figure 8. The numbers indicate possible ring sizes, various annulation positions are omitted for clarity. All compounds share an angular form with typically three ring systems connected in various ways. It has been shown previously that all four topological classes can be redesigned by the de novo program Skelgen.[31]

For a test set of 400 de novo designed molecules, we determined a ranking list based on their calculated MM-PBSA ligand-COX 2 interaction energies. As can be seen from Figures 8a−c, we can identify three of the four topological classes within the top five ranked molecules, with binding modes that are well reproduced by Skelgen. The closest analogue of the scaffold with two ring systems directly connected in ortho-position to a central unit (Figure 8d) is ranked at position 11. As the original 400 de novo-designed molecules do not contain the exact structure of an active compound we observe slight differences in atom types and substituents in the comparisons of Figure 8.

To better assess the performance of the MM-PBSA ranking, it is necessary to know how many different topologies are contained in our test set of 400 molecules and how they are oriented in the binding site. To this end, we clustered the de novo designed molecules with a proximity index p that takes into account similarity in both topology and binding site coverage:[31]

$$p = \frac{2 \times \mathrm{rmsd} - N_{\mathrm{mcs}}}{0.5(N_1 + N_2)} \qquad (8)$$

N_{mcs} is the number of atoms in the maximum common substructure (mcs), rmsd is the root-mean-square deviation of these atoms, and N_1 and N_2 are the numbers of heavy atoms in the two molecules that are compared. Subsequent clustering of the similarity matrix using Ward's method[47] (threshold = −2.4) resulted in 15 different topologies/binding site coverages. If we assume that medicinal chemists start synthesis for 3 of the 15 different classes of the initial de novo set the probability of embarking on a scaffold that is quite far away from an active topology would be relatively high ($11/15 \times 10/14 \times 9/13 = 36\%$). In comparison, the application of the MM-PBSA filter would yield two starting scaffolds (Figures 8a,b) that are quite close to active ones. In this respect, the combination of Skelgen de novo design with MM-PBSA postprocessing to prioritize the initial suggestions appears to be a promising approach.

With the improved performance of the MM-PBSA filter relative to a random selection of de novo designed molecules, the question remains whether this is sufficient to have an impact in the lead generation process. We think it is, if the scaffold selection process is carefully planned and synthetic chemistry experts are involved at an early stage. Due to the problem of sampling a huge chemical space and the limitations in accurately predicting protein−ligand binding affinities, one can currently not expect that the molecules generated by a combination of de novo design and MM-PBSA postprocessing are already potent inhibitors. Hence, it is hardly effective to make a particular designed ligand if the synthetic route is complicated. Rather, we suggest a strategy in which, after clustering of the de novo solutions, essential and variable structural features contained in these clusters are identified. In close interaction with organic chemists, this informa-

tion can be used to select novel molecules that have relatively straightforward synthetic access and whose topology and pharmacophore profile satisfies the molecular modeler. This approach seems more rewarding than to focus on particular individual molecules. Database searching techniques to find close analogues of designed structures are of additional value in this phase.[48]

Summary and Conclusions

Our extensive validation of MM(D)-PBSA on a panel of eight proteins reveals several new observations that affect the use of this popular approach for drug design applications. We find that the MM-PBSA energy function, eqs 1−3, is more robust and shows considerably less performance fluctuation than typical scoring functions for a wide range of binding site characteristics. Several examples indicate a more realistic treatment of solvation effects penalizing unfavorable nonpolar/polar protein−ligand interactions that frequently haunt docking solutions. The Poisson−Boltzmann model contributes to the improved accuracy as a simpler solvation model involving distance-dependent dielectric screening decreases the overall prediction capability. Our results further show that applying the MM-PBSA energy function to a single, relaxed complex structure is an adequate and sometimes better approach than the standard free energy averaging over molecular dynamics snapshots. This is noteworthy as almost all previous applications used the considerably higher CPU expense of MD-PBSA. We find very little gain in accuracy in extending our computational MD-PBSA procedure (MD: 200 ps, averaging over 10 snapshots) to longer molecular dynamics runs and wider free energy sampling (MD: 500 ps, averaging over 50 snapshots) as performed in several previous applications.[16,19] Moreover, including an additional minimization step before free energy averaging does not lead to generally better rankings of protein−ligand binding affinities. It may be that a more sophisticated MD procedure involving longer simulation times and the use of PME electrostatics further improves the results; however, this time-consuming effort would be of greatly reduced interest for drug design applications. Our results indicate that some artifacts arise when the explicit solvent is replaced by the continuum description. Rather than longer simulation times, it seems that including water molecules from the first solvation shell into the free energy expression could be more helpful.

Concerning the usefulness of MM(D)-PBSA for the three computational drug design applications investigated we can draw the following conclusions. First, without the need to perform time-consuming molecular dynamics and free energy sampling, MM-PBSA becomes attractive as a postdocking filter for the virtual screening of larger databases. As the CPU time for the calculation of a single structure MM-PBSA binding energy (∼15 min) is only by a factor of 10−100 slower than current docking programs, several ten thousand compounds can be processed on a medium-sized compute cluster in a week. This is sufficiently fast to be useful for the selection of focused subsets, which are typically in the range of a few 100−1000 compounds. Faster PB algorithms[49] or efficient generalized Born implementations[50] could be a useful replacement of our solvation model and should further increase the throughput for this task. Second, MM-PBSA appears to be useful as a prioritization filter for the manifold of solutions generated by de novo design. Finally, our results on the ranking of similar ligands to a given protein, in agreement with previous studies, indicate that weak and strong binders that differ by 2−3 orders of magnitude in IC_{50} can often be distinguished by MM-(D)-PBSA. Reproducing smaller free energy differences is in our experience, however, more the exception than the rule. More rigorous free energy approaches, such as TI or newer variants thereof,[51] seem to remain unchallenged, but have their well-known limitations in handling molecular diversity and compute time.

Acknowledgment. We thank Clemens Broger for providing the surface area module of his program XSAE and David M. Goldstein for helpful discussions. Donald Bashford is gratefully acknowledged for making the MEAD program available to us.

Supporting Information Available: Experimental and calculated binding data for the comparison of related ligands (test sets I−III). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Muegge, I.; Rarey, M. Small molecule docking and scoring. *Rev. Comput. Chem.* **2001**, *17*, 1−60.
(2) Boehm, H.-J.; Stahl, M. The use of scoring functions in drug discovery applications. *Rev. Comput. Chem.* **2002**, *18*, 41−87.
(3) Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(4) Schulz-Gasch, T.; Stahl, M. Binding site characteristics in structure-based virtual screening: evaluation of current docking tools. *J. Mol. Model.* **2003**, *9*, 47−57.
(5) Wyss, P. C.; Gerber, P.; Hartman, P. G.; Hubschwerlen, C.; Locher, H. et al. Novel Dihydrofolate Reductase Inhibitors. Structure-Based versus Diversity-Based Library Design and High-Throughput Synthesis and Screening. *J. Med. Chem.* **2003**, *46*, 2304−2312.
(6) Pearlman, D. A.; Charifson, P. S. Are Free Energy Calculations Useful in Practice? A Comparison with Rapid Scoring Functions for the p38 MAP Kinase Protein System. *J. Med. Chem.* **2001**, *44*, 3417−3423.
(7) Kollman, P. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(8) Aqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng* **1994**, *7*, 385−391.
(9) Aqvist, J.; Marelius, J. The linear interaction energy method for computation of ligand binding affinities. In *Free Energy Calculations in Rational Drug Design*; Kluwer Academic/Plenum Publishers: New York, 2001; pp 171−194.
(10) Banba, S.; Guo, Z.; Brooks, C. L., III. New free energy based methods for ligand binding from detailed structure−function to multiple-ligand screening. In *Free Energy Calculations in Rational Drug Design*; Kluwer Academic/Plenum Publishers: New York, 2001; pp 195−223.
(11) Eriksson, M. A. L.; Pitera, J.; Kollman, P. A. Prediction of the Binding Free Energies of New TIBO-like HIV-1 Reverse Transcriptase Inhibitors Using a Combination of PROFEC, PB/SA, CMC/MD, and Free Energy Calculations. *J. Med. Chem.* **1999**, *42*, 868−881.
(12) Erion, M. D.; Reddy, M. R. Ligand interaction scanning using free energy calculations. In *Free Energy Calculations in Rational Drug Design*; Kluwer Academic/Plenum Publishers: New York, 2001; pp 225−241.
(13) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(14) Vangrevelinghe, E.; Zimmermann, K.; Schoepfer, J.; Portmann, R.; Fabbro, D. et al. Discovery of a Potent and Selective Protein Kinase CK2 Inhibitor by High-Throughput Docking. *J. Med. Chem.* **2003**, *46*, 2656−2662.
(15) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helixes. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.

(16) Kuhn, B.; Kollman, P. A. Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models. *J. Med. Chem.* **2000**, *43*, 3786−3791.
(17) Donini, O. A. T.; Kollman, P. A. Calculation and Prediction of Binding Free Energies for the Matrix Metalloproteinases. *J. Med. Chem.* **2000**, *43*, 4180−4188.
(18) Kuhn, B.; Donini, O.; Huo, S.; Wang, J.; Kollman, P. A. MM-PBSA applied to computer-assisted ligand design. In *Free Energy Calculations in Rational Drug Design*; Kluwer Academic/Plenum Publishers: New York, 2001; pp 243−251.
(19) Wang, J.; 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.
(20) Huo, S.; Wang, J.; Cieplak, P.; Kollman, P. A.; Kuntz, I. D. Molecular Dynamics and Free Energy Analyses of Cathepsin D-Inhibitor Interactions: Insight into Structure-Based Ligand Design. *J. Med. Chem.* **2002**, *45*, 1412−1419.
(21) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L., III. Assessing Scoring Functions for Protein−Ligand Interactions. *J. Med. Chem.* **2004**, *47*, 3032−3047.
(22) Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M. Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7.ANG. resolution. *J. Mol. Biol.* **1993**, *231*, 698−710.
(23) Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. The structure of the complex between avidin and the dye, 2-(4′-hydroxyazobenzene) benzoic acid (HABA). *FEBS Lett.* **1993**, *328*, 165−168.
(24) Wang, Z.; Canagarajah, B. J.; Boehm, J. C.; Kassisa, S.; Cobb, M. H. et al. Structural basis of inhibitor selectivity in MAP kinases. *Structure* **1998**, *6*, 1117−1128.
(25) Wilson, K. P.; McCaffrey, P. G.; Hsiao, K.; Pazhinisamy, S.; Galullo, V. et al. The structural basis for the specificity of pyridinylimidazole inhibitors of p38 MAP kinase. *Chem. Biol.* **1997**, *4*, 423−431.
(26) McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. Gaussian docking functions. *Biopolymers* **2003**, *68*, 76−90.
(27) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.
(28) Baxter, C. A.; Murray, C. W.; Clark, D. E.; Westhead, D. R.; Eldridge, M. D. Flexible docking using tabu search and an empirical estimate of binding affinity. *Proteins* **1998**, *33*, 367−382.
(29) Todorov, N. P.; Dean, P. M. Evaluation of a method for controlling molecular scaffold diversity in de novo ligand design. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 175−192.
(30) Todorov, N. P.; Dean, P. M. A branch-and-bound method for optimal atom-type assignment in de novo ligand design. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 335−349.
(31) Stahl, M.; Todorov, N. P.; James, T.; Mauser, H.; Boehm, H.-J. et al. A validation study on the practical use of automated de novo design. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 459−478.
(32) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J. et al. *AMBER 7*; 7th ed.; University of California: San Francisco.
(33) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(34) Jakalian, A.; Bush, B. L.; Jack, D. B.; 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) Gerber, P. R.; Mueller, K. MAB, a generally applicable molecular force field for structure modeling in medicinal chemistry. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 251−268.
(37) Halgren, T. A. MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. *J. Comput. Chem.* **1999**, *20*, 730−748.
(38) Gundertofte, K.; Liljefors, T.; Norrby, P.-o.; Pettersson, I. A comparison of conformational energies calculated by several molecular mechanics methods. *J. Comput. Chem.* **1996**, *17*, 429−449.
(39) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr. et al. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(41) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S. et al. *AMBER 6*; 6 ed.; University of California: San Francisco.
(42) 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.
(43) Bashford, D.; Gerwert, K. Electrostatic calculations of the pK_a values of ionizable groups in bacteriorhodopsin. *J. Mol. Biol.* **1992**, *224*, 473−486.
(44) Reddy, M. R.; Erion, M. D., Eds. *Free Energy Calculations in Rational Drug Design*; Kluwer Academic/Plenum Publishers: New York, 2001; 384 pp.
(45) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(46) Lloyd, D. G.; Buenemann, C. L.; Todorov, N. P.; Manallack, D. T.; Dean, P. M. Scaffold hopping in de novo design: ligand generation in absence of receptor information. *J. Med. Chem.* **2004**, *47*, 493−496.
(47) Ward, J. H. Hierarchical grouping to optimize an objective function. *Am. Stat. Assoc. J.* **1963**, *56*, 236−244.
(48) Leach, A. R.; Bryce, R. A.; Robinson, A. J. Synergy between combinatorial chemistry and de novo design. *J. Mol. Graphics Modell.* **2000**, *18*, 358−367.
(49) Grant, J. A.; Pickup, B. T.; Nicholls, A. A smooth permittivity function for poisson-boltzmann solvation methods. *J. Comput. Chem.* **2001**, *22*, 608−640.
(50) Liu, H.-Y.; Kuntz, I. D.; Zou, X. Pairwise GB/SA Scoring Function for Structure-based Drug Design. *J. Phys. Chem. B* **2004**, *108*, 5453−5462.
(51) Oostenbrink, C.; van Gunsteren, W. F. Free energies of binding of polychlorinated biphenyls to the estrogen receptor from a single simulation. *Proteins* **2003**, *54*, 237−246.

JM049081Q