CHEMICAL REVIEWS

Review

Cite This: *Chem. Rev.* XXXX, XXX, XXX−XXX

pubs.acs.org/CR

End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design

Ercheng Wang,[†,∇] Huiyong Sun,[†,∇] Junmei Wang,[‡] Zhe Wang,[†] Hui Liu,[†] John Z. H. Zhang,[*,§,∥,⊥,#] and Tingjun Hou[*,†]

[†]Hangzhou Institute of Innovative Medicine, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China

[‡]Department of Pharmaceutical Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States

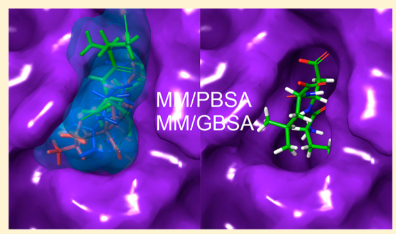
[§]Shanghai Engineering Research Center of Molecular Therapeutics & New Drug Development, Shanghai Key Laboratory of Green Chemistry & Chemical Process, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200062, China

[∥]NYU−ECNU Center for Computational Chemistry, NYU Shanghai, Shanghai 200122, China

[⊥]Department of Chemistry, New York University, New York, New York 10003, United States

[#]Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China

ABSTRACT: Molecular mechanics Poisson−Boltzmann surface area (MM/PBSA) and molecular mechanics generalized Born surface area (MM/GBSA) are arguably very popular methods for binding free energy prediction since they are more accurate than most scoring functions of molecular docking and less computationally demanding than alchemical free energy methods. MM/PBSA and MM/GBSA have been widely used in biomolecular studies such as protein folding, protein−ligand binding, protein−protein interaction, etc. In this review, methods to adjust the polar solvation energy and to improve the performance of MM/PBSA and MM/GBSA calculations are reviewed and discussed. The latest applications of MM/GBSA and MM/PBSA in drug design are also presented. This review intends to provide readers with guidance for practically applying MM/PBSA and MM/GBSA in drug design and related research fields.



CONTENTS

1. INTRODUCTION

In thermodynamics, free energy refers to the amount of internal energy of the system that can be used to do work, and it determines the direction of the thermodynamic process as well as the probability that the system will remain in a given

Received: January 21, 2019

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

state. Since the free energy drives all molecular processes, such as protein folding, molecular association, chemical reaction, etc., accurate determination of the free energy is one of the most important tasks in biomolecular studies. At present, molecular dynamics (MD) simulation is the most important tool used to obtain free energies of molecular systems.[1] MD simulation not only helps one to understand the physical processes of the system at the atomic level but also enables one to uncover the hidden states of the system that cannot be detected experimentally.[2−6] Since experimental measurements of the thermodynamic properties of biomolecular systems are often expensive and time-consuming, accurate theoretical calculations of their free energies by numerical simulation are becoming more and more important in many research fields, such as rational drug design, protein folding, protein−protein interactions (PPIs), etc.[1]

In a real molecular system, such as the binding of a drug to its receptor in a cell, the thermodynamic process is carried out under isothermal−isobaric (NPT) conditions, and the free energy of the system is given by[7]

$$F = -k_B T \ln Z \tag{1}$$

where k_B is Boltzmann's constant, T is the temperature of the system, and Z is the partition function of the system. Here the system is assumed to be in thermodynamic equilibrium, so the partition function Z can be expressed as[7−9]

$$Z = \frac{1}{V_0 N! h^{3N}} \int \exp\left(-\frac{PV}{k_B T}\right) \mathrm{d}V \iint \exp\left(-\frac{H(\mathbf{p}, \mathbf{r})}{k_B T}\right) \mathrm{d}\mathbf{p}\, \mathrm{d}\mathbf{r} \tag{2}$$

where V_0 is a constant that has units of volume, N is the number of particles in the system, h is Planck's constant, and P and V are the pressure and volume of the system, respectively. The factor $N!$ in eq 2 appears only for indistinguishable particles. The integration is performed over all phase space ($3N$ positions $\mathbf{r}$ and conjugate momenta $\mathbf{p}$). The Hamiltonian $H(\mathbf{p}, \mathbf{r})$ is the system's total energy for a given configuration with known momenta and coordinates.

The absolute free energy in eq 1 can be directly computed for a limited number of cases only,[1,10,11] for which the systems are usually small, governed by simple Hamiltonians, and analytical expressions for the partition functions exist. For larger systems with complex interactions between particles, it is often impossible to obtain analytical formulas for the partition functions, and thus, their absolute free energies cannot be directly computed using eqs 1 and 2. In most cases, it is often more practical to compute the difference between the free energies of the targeted state and a reference one. Of course, in special cases, if the free energy of the reference state is known, the absolute free energy of the system could still be obtained. For example, an analytical formula for the partition function can be acquired by neglecting the interactions between particles (such as ideal gases) or as a result of the symmetry between the particles (such as ideal crystals). However, most biological events occur in liquid solution, and it is more difficult to define an appropriate reference state for a liquid-phase system than those ideal systems. Therefore, in areas such as drug design,[12−15] protein−protein/ligand interactions,[16,17] solubility of small molecules,[18,19] protein−ligand binding affinities,[20−24] protein folding,[25,26] and conformational changes of biological macromolecules,[27] it is more convenient and realistic to calculate the free energy difference for an event or the relative free energy of two states. In general, combining the statistical mechanical expressions for F to estimate ΔF_{BA}, the free energy difference between two (or possibly a series of) states A and B, gives[28]

$$\Delta F_{BA} = F_B - F_A = -k_B T \ln \frac{Z_B}{Z_A} = -k_B T \ln \frac{p_B}{p_A} \tag{3}$$

where p_A and p_B are the probabilities that the system is state A or B, respectively.

In drug design, the binding free energy is often used to characterize the binding strength between a receptor and a drug molecule. The fundamental goal of structure-based drug design is to find new lead compounds that bind as tightly as possible to macromolecular receptors. Compared with the experimental methods, computational methods can significantly reduce the time and cost of designing new drug molecules. Various theoretical methods have been successfully utilized in drug design/discovery.[28,29] The proper choice of these methods largely depends on the stage at which the design will be needed, and there is usually a trade-off between the accuracy and efficiency.[24]

The most widely used computational method in structure-based drug design is molecular docking.[30−32] The main application of this technique in drug design is to predict the binding poses of candidate compounds in a defined binding pocket and discriminate binders from nonbinders. Although molecular docking is computationally efficient and low-cost, its predictions of binding poses and, particularly, binding free energies as measured by docking scores are usually not of high accuracy,[33−36] and it has difficulty in reliably distinguishing compounds with comparable binding affinities.

Alchemical free energy (AFE) methods,[15,37−39] which are also called pathway methods, require the interconversion of the system from the initial state to the final state via infinitesimal alchemical changes of the energy function, and they are theoretically rigorous and accurate. Free energy perturbation (FEP)[40−43] and thermodynamic integration (TI)[44−47] are two techniques that are widely utilized in alchemical free energy computations. However, one critical issue of the alchemical methods is the slow convergence of the free energy differences and high computational cost.[37] The convergence is particularly difficult in systems involving slow structural transitions or large environmental reorganizations. Hence, these methods are based on Monte Carlo (MC) or MD simulations and require sufficient sampling of complexes, ligands, and intermediate states in solution, resulting in huge computational cost in practical applications. Moreover, the setup of systems for TI and FEP calculations is complicated and requires experience. Recently, software for performing TI and FEP calculations using graphics processing units (GPUs) has begun to emerge.[48,49] Even though the computational efficiency is greatly improved in comparison to the CPU versions, TI/FEP with GPU is still not suitable for large-scale virtual screenings and is mostly used in the stage of lead optimization in a drug design campaign.

The above two sets of methods do not perform well in terms of the balance between accuracy and efficiency. On the other hand, end-point free energy methods[50−55] have been extensively utilized in structure-based drug design. As the name indicates, end-point methods are based on samplings of the final states of a system, and therefore, they are much less expensive than the pathway methods and more accurate than

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

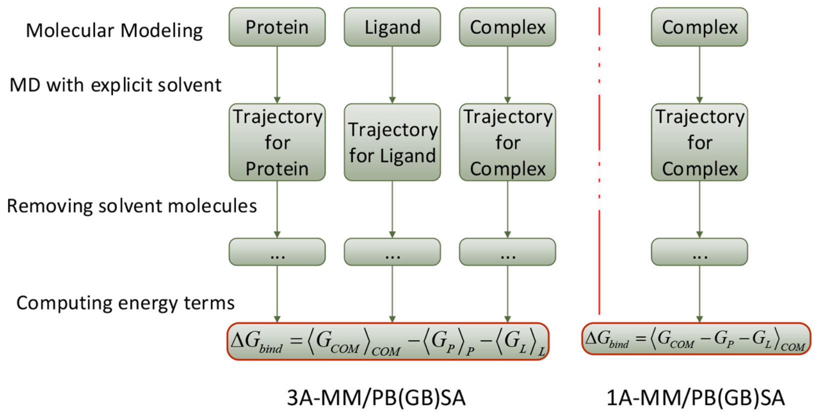


Figure 1. Typical flowchart for calculating binding free energy using the three-average and one-average MM/PB(GB)SA protocols (called 3A-MM/PB(GB)SA and 1A-MM/PB(GB)SA, respectively, by Genheden and Ryde[51]).

most docking scoring functions. The most well known end-point free energy methods are molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) and molecular mechanics generalized Born surface area (MM/GBSA), developed by Kollman et al.,[56−58] which achieve a good balance between computational efficiency and accuracy and thus are the focus of this review. Since the PB solution[56,58,59] is computationally time-consuming, a set of more efficient approximation methods based on the GB model[59−62] have been developed and have attracted more and more attention.[63−67] Another popular method with intermediate performance is linear interaction energy (LIE)[55,68−70], whose computational efficiency is second only to that of the scoring function, but we will not discuss it in this review. MM/PBSA and MM/GBSA have been widely used to evaluate docking poses, determine structural stability, and predict binding affinities and hotspots. In addition, MM/PBSA and MM/GBSA allow analysis of the contributions from individual residues or energy terms by free energy decomposition analysis,[63,71,72] which gives detailed residue-specific energetic contributions to the system binding, identifies dominant interactions in the binding process, and thereby facilitates individualized drug design.

Earlier reviews[24,73−76] mainly focused on how to apply MM/PBSA and MM/GBSA to calculate binding free energies and how to improve the methods from both the efficiency and accuracy perspectives. Unfortunately, there is not much consensus on these techniques so far, mainly because the computational performance of the methods depends on the system being studied. In this review, we first briefly describe the methodologies of the MM/PBSA and MM/GBSA approaches and then discuss how to improve their performance, with a special emphasis on calculation of the polar solvation energy. Finally, we discuss the latest applications of the MM/GBSA and MM/PBSA methods in the fields of small-molecule drug design and macromolecule interactions. We hope that the current review will provide readers with helpful guidance in comprehensive understanding of the methodologies as well as practical applications of the MM/GBSA and MM/PBSA approaches.

2. METHODOLOGY

In the MM/PBSA or MM/GBSA approach, the free energy for binding of the ligand (L) to the protein receptor (R) to form the complex (RL),

$$\Delta G_{\mathrm{bind}} = G_{\mathrm{RL}} - G_{\mathrm{R}} - G_{\mathrm{L}} \tag{4}$$

can be decomposed into contributions of different interactions and expressed as[58]

$$\Delta G_{\mathrm{bind}} = \Delta H - T\Delta S = \Delta E_{\mathrm{MM}} + \Delta G_{\mathrm{sol}} - T\Delta S \tag{5}$$

in which

$$\Delta E_{\mathrm{MM}} = \Delta E_{\mathrm{int}} + \Delta E_{\mathrm{ele}} + \Delta E_{\mathrm{vdW}} \tag{6}$$

$$\Delta G_{\mathrm{sol}} = \Delta G_{\mathrm{PB/GB}} + \Delta G_{\mathrm{SA}} \tag{7}$$

$$\Delta G_{\mathrm{SA}} = \gamma \cdot \mathrm{SASA} + b \tag{8}$$

where ΔE_{MM}, ΔG_{sol}, and $-T\Delta S$ are the changes in the gas-phase molecular mechanics (MM) energy, solvation free energy, and conformational entropy upon ligand binding, respectively. ΔE_{MM} includes the changes in the internal energies ΔE_{int} (bond, angle, and dihedral energies), electrostatic energies ΔE_{ele}, and the van der Waals energies ΔE_{vdW}. ΔG_{sol} is the sum of the electrostatic solvation energy $\Delta G_{\mathrm{PB/GB}}$ (polar contribution) and the nonpolar contribution ΔG_{SA} between the solute and the continuum solvent. The polar contribution is calculated using either the PB or GB model, while the nonpolar energy is usually estimated using the solvent-accessible surface area (SASA).[77,78] However, the GB method gives an analytical expression for the polar solvation energy and is thus much faster than the PB method. The change in conformational entropy $-T\Delta S$ is usually calculated by normal-mode analysis[58] on a set of conformational snapshots taken from MD simulations. However, because of the heavy computational cost, changes in the conformational entropy are usually neglected when only the relative binding free energies of similar ligands are needed.

The typical steps of using MM/PB(GB)SA combined with MD simulations to calculate the binding free energy include the following: (1) An MD simulation of the protein–ligand complex is performed using an explicit solvent model, as the implicit solvent simulations have been shown to produce less accurate results.[79] (2) All of the solvent molecules and charged ions are deleted from each MD snapshot, and the implicit PBSA or GBSA solvent model is used to evaluate the solvation energy. (3) Optionally, the solute conformational entropy change can be computed from a chosen set of snapshots. The final binding free energy is then acquired by a simple summation of these individual energy components. Figure 1 shows the flowchart of the three steps with two computational

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

protocols (the separate trajectory protocol on the left and the single trajectory protocol on the right). It is noted that the reason for applying an implicit solvation model to calculate the free energy of a system is to avoid the large fluctuation of potential energies when explicit water molecules are used in the calculation.[80] However, the above computational protocol of applying both the explicit and implicit solvation models in the first two steps is arguably inconsistent, as the MD simulation and energy calculation share the inconsistent energy functions, thus requiring reweighting of some energy terms. Although one could in principle avoid this inconsistency by also performing the MD simulation with an implicit solvent model, the result of such an approach is generally less reliable, and the simulation may even lead to dissociation of a ligand from its receptor or a protein subunit from the others.[79] Ryde et al. claimed that explicit solvent model for MD simulation is essential and that simulations with implicit water molecules often yield poor results.[79] Since water plays an important role in protein–ligand interactions (e.g., water molecules can form hydrogen bonds between the protein and the ligand), MD simulation in an explicit water solvent system is essential for reliable prediction of the binding free energies of protein–ligand systems.

As shown in Figure 1, there are mainly two protocols to generate the conformations in the first step above: (1) performing independent MD simulations for the isolated ligand, apo protein, and bound protein–ligand complex and (2) performing an MD simulation for the bound protein–ligand complex and using a single MD trajectory of the bound complex to obtain the structures of all three components (i.e., the ligand, apo protein and protein–ligand complex).[81] In practice, the second protocol is more popular and preferable, and it gives more accurate results with lower standard errors.[24] However, the second approach is based on the assumption that the structures sampled for the protein–ligand complex in solution are sufficiently similar to those sampled for the apo protein and the isolated ligand. This assumption is valid in most cases of protein–ligand binding but could be invalid in some cases, e.g., when the protein–ligand binding is associated with large conformational changes.[50] A main attractive advantage of the second approach is that the internal energy errors of the systems can be canceled because the energy difference between the protein–ligand complex and its individual components (apo protein and ligand) are computed using exactly identical configurations. However, for the first protocol, the energy difference taken between the averages produced from separate bound and unbound trajectories may result in additional errors or noise due to large internal energy errors at different conformations or structures, and these errors are difficult to eliminate simply by running longer MD simulations. In general, it is a challenge to correctly determine whether the simulation has reached convergence.[82] In the second protocol, however, the fluctuations of the energy terms are much smaller because of the cancellation of internal energy errors, but this protocol may suffer from inadequate sampling of the apo receptor and/or the ligand.

The energy terms in eq 5 are averaged over multiple configurations or several MD snapshots (typically a few dozen or hundred structures for ΔE_{MM} and ΔG_{sol}) to improve the predictive accuracy of the binding free energies. Depending on the extent of configurational fluctuations of the system, convergence to the stable states possibly requires relatively longer (multinanosecond) MD simulations. Genheden and Ryde investigated the convergence issue[83] and found that it is better to average the results predicted from several independent MD trajectories, and that opinion is also supported by other works.[84−86] For avidin, with a 200 ps production time for each MD run, 5−50 independent MD simulations are needed to reach a statistical accuracy of ~1 kJ/mol for the seven biotin analogues.[83] Several published works reported that the results of the MM/PB(GB)SA approaches are dependent on the length of the MD simulations.[87,88] Johnson et al. found that with different PB radii, the results of their calculations using the MM/PBSA method based on 0.25−2.00 ns MD trajectories were satisfactory (Pearson correlation coefficient $r^2 > 0.70$),[87] but their calculations based on MD trajectories longer than 2 ns gave less accurate results. In a previous work,[88] we explored the effect of the length of MD simulations (ranging from 400 to 4800 ps) on binding free energy predictions. The results showed that the length of the MD simulations significantly impacts the accuracy of the predicted binding free energies and that in order to obtain better predictions, longer MD simulations are not necessarily beneficial, with simulation lengths less than 5 ns considered to be reasonable. Virtanen et al., however, concluded that the length of the MD simulations is not of critical importance to the accuracy of the calculations.[89] It seems that the impact of the MD simulation length on free energy calculations is system-dependent.

A problem with MM/PB(GB)SA is the occurrence of several substates that are seldom sampled during the simulations. In such a case, a binding free energy with a larger standard error may be obtained beyond expectation,[90] indicating that even longer (maybe >10 ns) or several independent simulations should be performed to yield improved results from a better-equilibrated simulation.[83] Recently, Tsuda et al. proposed a useful machine learning approach (called *Best Arm Identify*) to optimally regulate the minimum number of MD trajectories for protein–ligand systems.[91] Interestingly, an important improvement in MM/PBSA predictions was also obtained by filtering MD snapshots through prescoring of the protein–ligand complexes with a machine-learning-based approach (SVMSP).[92]

Although many studies emphasized the importance of MD sampling,[93] the minimized conformations could frequently yield predictions as good as or even better than those from MD simulations in practice.[94−97] That is to say, MD sampling does not seem to be essential in binding free energy prediction in some cases. Binding free energy can be calculated by MM/PB(GB)SA on the basis of a single minimized structure rather than abundant MD snapshots. Undoubtedly, that approach costs much less computational time, whereas at the same time it ignores the dynamical effects, resulting in predictions that are extremely dependent on the initial structures and losing all of the information about the statistical accuracy[24] of the approach. It seems that the standard deviation in the statistics cannot be utilized to estimate the precision of the predicted binding free energy and various energy terms from multiple MD snapshots. Ryde et al. tested this approach by minimizing selected MD snapshots and showed that the results based on single minimized configurations are similar to those based on MD trajectories but that sometimes unrealistic structures need to be eliminated in order to avoid unpredictable and incorrect binding affinities.[79]

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

Table 1. Assessments of the Performance of the MMPBSA and MMGBSA Methods since 2010[a]

system	method	content	conclusions	year[ref]
six protein−ligand systems	MM/PBSA and MM/GBSA	solute dielectric constant; length of MD simulations; conformational entropy	(1) The length of the MD simulation obviously affects the predictions of binding free energies, and longer simulations are not always necessary to achieve better predictions. (2) The predictions based on different solute dielectric constants are quite different. (3) Large fluctuations of conformational entropy often appear in MD trajectories, and a large number of snapshots are necessary to obtain stable predictions. (4) MM/PBSA exhibits better performance than MM/GBSA in predicting absolute binding free energies, but it may be not a better choice to calculate relative binding free energies.	2011[88]
98 protein−ligand complexes	MM/PBSA and MM/GBSA	MD sampling; performance of MM/PBSA and MM/GBSA	(1) MM/GBSA outperforms MM/PBSA and most popular scoring functions in identifying the correct binding conformation. (2) MD simulation is essential to correctly identify the binding structures for some systems. (3) MM/GBSA shows good performance in predicting both binding poses and binding free energies.	2011[93]
four protein−ligand systems	MM/GBSA	different initialization	(1) The predictions of MM/GBSA based even on structures generated by different MD simulation packages are reasonably reproducible. (2) Solvating the systems with different equilibrated water boxes can also sample extensive conformations, in addition to the common use of different initial velocities.	2011[103]
seven biotin analogues binding to avidin and nine inhibitors binding to factor Xa (fXa)	MM/PBSA and MM/GBSA	solute dielectric constant; electrostatic contributions	(1) The predicted accuracy of the two approaches (3A-MM/GBSA and 1A-MM/GBSA) depends on the tested case and the solvation model. (2) Combining MM/PB(GB)β with the stricter LRA approach yields the best predictions for avidin, whereas the pure MM/PB(GB)SA approaches yield the best predictions for fXa. (3) The precision can be improved with a relatively higher dielectric constant, especially for the 3A-MM/PBβ method. (4) The optimum solute dielectric constant for PB is slightly higher than that for GB.	2012[51]
eight inhibitors binding to fXa, nine inhibitors binding to ferritin, and two ligands binding to HIV-1 protease	MM/GBSA	system truncation; buffer region; solute dielectric constant	(1) Truncation of the protein is not innocent. However, it almost always improves the predicted precision for the truncated systems. (2) The predictions with a distance-based dielectric constant give a rather constant MAD. (3) A shell of fixed water molecules and protein residues is crucial in stabilizing the energies for the truncated systems. (4) Omitting the entropy effect is not always appropriate. Calculating entropies using a protein truncated beyond a distance of 8 Å from the ligand with a 4 Å buffer of fixed protein residues and water molecules together with a dielectric constant of 1 gives a good balance between accuracy and computational efficiency, only for predicting the relative binding affinities.	2012[104]
46 small molecules binding to five proteins	MM/PBSA and MM/GBSA	force fields; ligand charge models	(1) For short MD simulations (<1 ns), ff03 yields the best predictions using both MM/GBSA and MM/PBSA. For medium MD simulations (2−4 ns), MM/GBSA with ff99 yields the best predictions, while MM/PBSA with ff99SB performs the best. Longer simulations (>5 ns) may not be quite necessary. (2) MM/PBSA with Tan's parameters overcomes MM/GBSA (GB^{OBC1}) in terms of ranking ability. (3) RESP charges yield the optimal performance for MM/PBSA and MM/GBSA, and the results based on AM1-BCC and ESP charges are also fairly satisfactory.	2013[97]
five ligands binding to gal3	MM/GBSA	MD with explicit and implicit solvent	(1) The predicted accuracy and precision are extremely influenced by the choice of solvent model both in the MD simulations and in the energy calculations. (2) RMSD analysis could hardly distinguish the generated ensembles. (3) MM/GBSA based on minimizations may not yield results similar to those from standard MM/GBSA based on MD simulations. The predictions based on minimizations show worse convergence.	2013[105]
1872 protein−ligand complexes	MM/PBSA and MM/GBSA	simulation protocols; solute dielectric constant	(1) Both MM/GBSA and MM/PBSA calculations based on a dielectric constant of 4 show the best unbiased accuracies (the one-protein-family/one-binding-ligand case) and the best overall accuracies (based on the 1864 protein−ligand complexes) for MM/GBSA predictions. (2) The accuracies of both the MM/GBSA and MM/PBSA predictions decrease with increasing ligand total formal charge. (3) MM/PBSA with higher systematic sensitivity is potentially more suitable for individual-target-level prediction than MM/GBSA.	2014[98]
three tyrosine kinases (ABL, ALK, and BRAF)	MM/PBSA, MM/GBSA, and docking	solute dielectric constant	(1) A higher solute dielectric constant (ε = 2 or 4) is preferred to improve the rescoring accuracy for both MM/GBSA and MM/PBSA. (2) MD simulations may be unnecessary to improve enrichment, but optimizations are indeed necessary for MM/GBSA or MM/PBSA rescoring.	2014[94]

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

Table 1. continued

system	method	content	conclusions	year[ref]
16 inhibitors binding to FabI	MM/PBSA, MM/GBSA, and QM−MM/GBSA	radii sets; GB models; entropy; QM methods; MD sampling	(1) MM/PBSA results based on 0.25 to 2.00 ns MD trajectories were satisfactory in all bondi, mbondi, mbondi2, and PARSE radii sets. However, using ≥2 ns MD trajectories for MM/PBSA seemed to lower the accuracy of prediction with all four radii sets. For MM/GBSA, the critical MD length is 1 ns. (2) It was clearly seen that bondi and mbondi2 gave almost identical predictive power. MM/GBSA is more sensitive to various radii sets than MM/PBSA, and the mbondi radii set was a nonrecommended radii setting for GB^{OBC}. (3) GB^{HCT} offered the best agreement with the experimental binding results. (4) MM/GBSA and QM−MM/GBSA provide better accuracy and higher efficiency than MM/PBSA using either the multiple independent sampling method or the traditional single long simulation method.	2015[87]
46 protein−protein complexes	MM/PBSA and MM/GBSA	force fields; solute dielectric constant; GB models	(1) The MM/GBSA calculations yield the best r_P of −0.647, which is better than that for MM/PBSA. (2) The rescoring accuracy for MM/GBSA is improved with a lower solute dielectric constant. (3) The PBA of the binding interface of a protein−protein complex can possibly be applied as a criterion to choose the optimal solute dielectric constant in MM/GBSA calculations to recognize the correct binding poses.	2016[93]
20 inhibitors binding to PLK1	docking, FEP, MM/GBSA, and QM−MM/GBSA	length of MD simulations; radii sets; GB models; multiple short MD simulations	(1) For MM/PBSA, even 10 ns is not enough to get the optimal r_S, while for MM/GBSA, 8 ns seems to be a plausible choice to get the ideal result. (2) The ranking capability of FEP is the best. (3) The r_S obtained from the 8 ns MD sampling-based MM/GBSA score can reach 0.767, which is lower than that of FEP by 0.087, but its computational cost is much less than that of FEP. (4) The ranking performance of MM/GBSA can be obviously improved by treating the ligands with a QM method. (5) mbondi is more sensitive to the GB method, and the recommended value of *igb* is 5. (6) Among the three kinds of traditional docking scoring functions, the ranking performance of a force-field-based scoring function is much better for congeneric compounds.	2017[106]
RSL lectin bound to MeFuc	MM/PBSA, MM/GBSA, and QM−MM/GBSA	GB models; QM methods; entropic contributions	(1) The binding free energies using implicit solvent methods are sensitive to the simulation length, radii set, GB model, and QM Hamiltonian. (2) The MM/PBSA and MM/GBSA calculations from a 7 ns MD simulation provided statistically good agreement with the experimental binding affinities. (3) MM/PBSA using the mbondi radii sets offers the best correlation. (4) Out of 12 QM Hamiltonians tested, PM6, DFTB, and their variants proved to be more efficient than other semiempirical methods. (5) The inclusion of entropic terms may be reasonable only when aiming to obtain a quantitatively better agreement with experimental values.	2018[107]
1508 protein−ligand complexes	MM/PBSA and MM/GBSA	force fields; solute dielectric constant; entropy effects	(1) The ff03 force field always performs the best across all six tested force fields for each calculation strategy at any dielectric constant. (2) One should be cautious to estimate ligand-binding free energies with NMA when using minimized structures. (3) The inclusion of truncated NMA entropies can improve the performance of both MM/GBSA and MM/PBSA results at a relatively high dielectric constant ($\varepsilon = 4$).	2018[99]
148 protein−RNA complexes	MM/PBSA, MM/GBSA, and docking	solute dielectric constant; GB models	(1) The best prediction comes from MM/GBSA calculated using the GB^{GBn1} model with a dielectric constant of 8 based on 1 ns MD simulations in TIP3P water solvent. (2) The computational protocol of simple minimization of binding poses in explicit solvent using the ff14SB force field and calculation of the binding free energies using the GB^{GBn1} model with a dielectric constant of 2 yields the satisfied correlation. (3) MM/GBSA is a powerful scoring function for protein−RNA docking studies.	2018[100]

[a]Abbreviations: LRA, linear-response approximation; MAD, mean absolute deviation; QM, quantum mechanics; r_P, Pearson correlation coefficient; PBA, polar buried area; r_S, Spearman correlation coefficient; PM6, parametrized model number 6; DFTB, density functional theory-based tight binding; NMA, normal-mode analysis; RMSD, root-mean-square deviation.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

3. ASSESSING THE PERFORMANCE OF MM/PBSA AND MM/GBSA

In a series of published works, we have systematically evaluated the prediction capability of the MM/PBSA and MM/GBSA methods.[88,93,94,97−101] The results suggest that the prediction of binding free energy strongly depends on the force field,[79,97] charge model,[97] continuum solvation method,[102] interior dielectric constant,[94] sampling method,[79] and conformational entropy.[99] In Table 1, we summarize some benchmark reports assessing the performance of the MMPBSA and MMGBSA methods since 2011. The general conclusions include the following: (1) 1 ns MD simulation in explicit solvent with the AMBER ff03 force field is critical for the MM/GBSA calculation; (2) a relatively higher solute dielectric constant (ε_{in} = 2 or 4) is preferable to improve the rescoring accuracy of MM/PB(GB)SA; (3) MM/PBSA is more sensitive to the investigated systems than MM/GBSA; (4) the Hawkins−Cramer−Truhlar GB model (GB^{HCT}) gives the best agreement with the experimental values of binding affinities, and the mbondi radii set is not recommended for the Onufriev−Bashford−Case GB model (GB^{OBC}).

The force field plays a central role in molecular simulations since it determines all of the interactions of a system.[108] Ryde et al. studied the effects of the force field on the predicted ligand binding affinities calculated by MM/PBSA using three different versions of the nonpolarizable AMBER force field (AMBER ff94, ff99, and ff03) and obtained similar results.[79] Better force fields such as polarizable force fields are desirable for use in MM/PBSA, as the current implicit solvent models were developed for nonpolarizable force fields. In our previous work, 46 small molecules targeting five different protein targets were used to test the effects of five AMBER force fields and four different charge models.[97] We reached the following conclusions: (1) for short MD simulations (<1 ns), the AMBER ff03 force field yields the optimal prediction by both MM/GBSA and MM/PBSA; (2) for medium-length MD simulations (2−4 ns), MM/GBSA with ff99 and MM/PBSA with ff99SB give the optimal predicted results. It should be noted that these conclusions are restricted to the set of force fields used in that study (AMBER ff99, ff99SB, ff99SB-ILDN, ff03, and ff12SB). These results indicate that the force field used in MD simulations and MM/PB(GB)SA calculations should be consistent and that mixing of different force fields may cause inaccurate predictions.

Gohlke and Case studied in detail how the results of binding free energy prediction depend on the polar solvation energy.[66] The results suggested that the PB predictions were largely impacted by the radii used, as the multiple variants of GB radii gave very different results. Among the four radii sets studied (bondi, mbondi, mbondi2, and PARSE), the results obtained using PARSE showed the lowest correlation with the experimental data (r^2 = 0.62−0.77), and unexpectedly, the difference between the predictions based on mbondi and mbondi2 was very small.[87] Recently, Ryde et al. proposed a method called MM/3D-RISM and compared the predictions of the new method with those of four GB and two PB methods.[102] Their results indicated that the accuracy of the MM/3D-RISM approach (Pearson correlation coefficient r_P = 0.90) was comparable to that of the best MM/PBSA method (r_P = 0.93), but the mean absolute deviation (MAD) was significantly worse (37 kJ/mol compared with 16 kJ/mol). Comparison of the MM/3D-RISM and MM/PBSA approaches was also made by Pandey et al., and in general, MM/PBSA performed better than MM/3D-RISM in the predictions of the relative binding free energies for three protein systems.[109] We evaluated the accuracy of the MM/GBSA predictions of three GB models,[88] and the results showed that the GB model developed by Onufriev, Bashford, and Case[110] is the best one for ranking the binding affinities of the inhibitors. In most cases, the MM/PBSA calculation with Tan's PB parameters shows better ranking performance than MM/GBSA (GB^{OBC1}). Restrained electrostatic potential (RESP) charges yield the optimal capability for MM/PBSA and MM/GBSA, while the predictions based on the AM1-BCC and ESP charges are also fairly satisfactory.

Johnson et al. found that the use of multiple short MD trajectories did not decrease the MM/PB(GB)SA performance compared with a single long MD simulation trajectory,[87] indicating that multiple independent dynamic samplings can offset the errors due to the force field. Therefore, the use of multiple independent samplings is recommended, which tends to average out the difference between different GB methods.

In addition, we also assessed the performance of MM/GBSA and MM/PBSA in reproducing the absolute binding free energies for a large data set.[88,99] The results showed that MM/GBSA gives worse predictions than MM/PBSA in calculating the absolute binding free energies. However, MM/GBSA produces better performance in ranking the binding affinities for systems without metals. Thus, MM/GBSA, with its much better computational efficiency, can be a powerful tool in drug design, where researchers frequently pay attention to the rational ranking of inhibitors.

Although MM/GBSA and MM/PBSA are reliable and efficient methods to estimate binding free energies, their weaknesses should also be noticed. In particular, one major source of errors is the conformational entropy. This is often computed by normal-mode analysis (NMA),[111−113] which dominates the computational cost of the MM/PB(GB)SA methods. Thus, for structurally similar molecules whose contributions to the conformational entropy are similar, this entropic term is usually ignored, and only the relative binding free energies of those molecules are calculated. Moreover, to decrease the computational cost, Genheden et al.[104] and Duan et al.[114] developed high-performance entropy calculation algorithms named truncated entropy and interaction entropy, respectively, to estimate the entropy changes for receptor−ligand interactions in MM/PB(GB)SA calculations, and these methods showed improved accuracy in an extensive computational study.[99] We developed a very efficient method called weighted solvent-accessible surface area (WSAS) to reproduce the conformational entropies computed by NMA.[196] The WSAS method requires no minimization of MD snapshots prior to conformational entropy calculations. It has been successfully used to predict the absolute binding affinities for a variety of systems, including orexin receptor 2[115] and cannabinoid receptor 1.[116]

Besides the contribution from entropy, the quality of the MM/PB(GB)SA calculations is dependent on the quality of the MD snapshots in representing the entire conformational space sampled as well as several parameters used for the description of the molecular system, such as the force field, the dielectric constants, and the set of atomic radii. Moreover, MM/PBSA and MM/GBSA show some limitations[74] in the estimation of binding free energies of highly polar or charged molecules, since the uncertainty in the calculation of the

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

solvation energy is proportional to the polarity of the considered molecules. Furthermore, contributions of structural water molecules, which bridge the key receptor−ligand interactions, are not taken into account well for predicting the binding free energies by implicit solvation models.[117]

Pentikäinen et al.[89] investigated the performance of the MM/GBSA, MM/PBSA, and solvation interaction energy (SIE)[118,119] approaches in terms of their virtual screening efficiency and ability to predict the binding affinities of five different protein targets. Protein−ligand complexes were prepared by two important methods in structure-based drug design: molecular docking and ligand-based similarity search methods. The results show a significant difference between different binding energy calculation methods. They suggested that these techniques should be used with caution in virtual screening or binding affinity estimation. Moreover, another work assessed the performance of four docking scoring functions and the FEP, MM/GBSA, and QM−MM/GBSA methods on a series of PLK1 inhibitors.[106] The ranking performance of FEP is optimal (Spearman correlation coefficient $r_S = 0.854$) and MM/GBSA, which requires much less simulation time (about one-eighth that of FEP), gives a comparable prediction ($r_S = 0.767$). In addition, the ranking performance can be significantly improved by treating the ligands with a quantum mechanics (QM) method. The combination of QM/MM molecular docking[120−123] and MM/GBSA calculations has been successfully utilized to reproduce the X-ray geometries of protein−ligand complexes with halogen bonding.[124]

Besides, binding free energy prediction is strongly influenced by the solute dielectric constant. Therefore, the energy should be carefully determined on the basis of the characteristics of the protein−protein/ligand binding interfaces. The following section will discuss this issue in detail.

4. THE POLAR SOLVATION ENERGY AND ENTROPY TERMS IN MM/PB(GB)SA CALCULATIONS

Many chemical reactions and biological processes are carried out in solvents, especially in water. Solvation effects are therefore critical to investigate the structures and functions of biological systems such as proteins, DNAs, and RNAs and interactions between them. In the MM/PB(GB)SA methods, the solvation energy is divided into polar and nonpolar contributions. We focus on the polar solvation energy in this section.

4.1. The Polar Solvation Energy Term in MM/PBSA

The polar solvation term in eq 7 was originally calculated to solve the Poisson−Boltzmann equation (PBE) numerically using a finite difference (FD) solution.[125−128] In a biomolecular system with no mobile ions, the Poisson equation is described as[129]

$$\nabla\cdot\varepsilon(\mathbf{r})\nabla\varphi(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \tag{9}$$

where $\varepsilon(\mathbf{r})$ is the dielectric distribution function for the solvated molecular system, $\varphi(\mathbf{r})$ is the electrostatic potential distribution function, and the $\rho(\mathbf{r})$ is the fixed atomic charged density based on the solute atom positions $\mathbf{r}$. However, in most cases of biomolecular systems, because of the presence of salt in the solution, the electrostatic potential $\varphi(\mathbf{r})$ is obtained by solving the following equation:[129]

$$\nabla\cdot\varepsilon(\mathbf{r})\nabla\varphi(\mathbf{r}) + 4\pi\lambda(\mathbf{r})\sum_{i=1}^{N} z_i e c_i \exp\left[-\frac{z_i e\varphi(\mathbf{r})}{k_B T}\right] = -4\pi\rho(\mathbf{r}) \tag{10}$$

where $\lambda(\mathbf{r})$ is the predefined ion-exclusion function, which has a value of 0 within the Stern layer and the molecular interior and a value of 1 outside the Stern layer, z_i is the charge and c_i is the bulk number density of ion type i far from the solute at a given temperature T, and e is the electron charge. The summation in eq 10 is over all of the different ion types, and when both the ionic strength and electric field are weak, the nonlinear PBE can be linearized for easier numerical solutions:[129]

$$\nabla\cdot\varepsilon(\mathbf{r})\nabla\varphi(\mathbf{r}) - \varepsilon_{sol}\kappa^2\varphi(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \tag{11}$$

where $\kappa^2 = \frac{8\pi e^2 I}{\varepsilon_{sol} k_B T}$ is the modified Debye−Hückel parameter, ε_{sol} is the solvent dielectric constant, and I is the ionic strength of the solution. The salt term in the PBE can be linearized when the exponent of the Boltzmann factor is close to zero. However, the approximation apparently does not hold in highly charged biomolecular systems.[130,131] Thus, it is recommended that a full nonlinear PBE solver should be used for such systems.

Obtaining analytical solutions of the linearized and nonlinear PBEs is extremely complicated, even in the few simple cases for which they exist. In the past decades, however, several computational methods have been developed[126] to solve the PBE. The classical FD method,[127,128] based on the superimposition of a regular rectangular Cartesian mesh over the system where the PBE can be solved, involves the following steps: (1) mapping atomic charges to the FD grid points; (2) assigning nonperiodic/periodic boundary conditions, i.e., electrostatic potentials on the boundary surfaces of the FD grid; (3) applying a dielectric model to define the high-dielectric (e.g., water) and low-dielectric (the solute interior) regions and mapping them to the FD grid edges. As a result of the developments in computational algorithms and hardware in recent years, several investigations of the efficiency and accuracy of numerical methods for the linearized equation have appeared,[128,132,133] and over the past few years a few new algorithms have been developed for the numerical solution of PBE.[134−136] In *pbsa*[137] (a module in the AMBER package,[138,139] one of the most popular computer tools to solve PBE), four common linear FD PBE solvers are implemented:[140] modified incomplete Cholesky conjugate gradient (ICCG), geometric multigrid, conjugate gradient (CG), and successive over-relaxation (SOR). In addition, six nonlinear FD solvers are implemented:[141] inexact Newton (NT)/modified ICCG, NT/geometric multigrid, CG, SOR, and its improved versions, adaptive SOR and damped SOR. The progress made in developing more accurate and efficient solutions to model the electrostatics in biomolecular systems, such as the finite element[142−145] and boundary element[146−148] methods, was recently reviewed by Alexov et al.[126]

The PBE is mathematically a three-dimensional second-order nonlinear elliptic partial differential equation. Using FD schemes for solving the PBE, various solvers have been developed, including PBSA,[137] MIBPB,[149] DelPhi,[150] UHBD,[151] ZAP,[152] and many others. The ZAP algorithm, developed by Nicholls et al., was incorporated into the CHARMM[153] package, providing a fast, stable, smooth permittivity model for implicit solvation energy calculations.[154]

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

Unfortunately, many popular PB methods do not converge. Specifically, their solution changes dramatically when the grid mesh is refined. This happens because of the discontinuous dielectric constants used across the solvent−solute interface.[155] To our knowledge, the only existing second-order convergent PB method for realistic protein surfaces is the MIBPB approach (https://weilab.math.msu.edu/MIBPB/). That is, the accuracy increases 4 times when the grid mesh size is halved.[149] In MIBPB, the grid mesh size should be set in the range of 0.2 to 1.2 Å, and a default value of 0.8 Å is used if the size is not specified. The convergence of MIBPB for protein−ligand binding analysis has been carefully tested recently, showing a relative error of 0.4%.[156]

Although numerical algorithms have been implemented to solve the PBE, currently all serial PBE solvers on CPUs are capable of calculating the electrostatics of only relatively small biomolecules and systems because of the intensive demand of computational resources (both time and memory) required to calculate the electrostatics of large systems (such as biomolecules and the complexes between them).[157−159] To obtain accurate results for large systems, even the fastest solvers, such as the DelPhi program,[150] typically take more than half a day to perform the calculations at the minimum requirement of grid resolution. Such computational efficiency is insufficient to meet current researchers' requirements in practical applications. Therefore, significant acceleration is required to make these serial algorithms suitable for studying large systems. In addition to developing new algorithms, there are two ways to speed up the current PBE solvers: (1) soft acceleration, that is, running parallel solvers on multiple CPUs, and (2) GPU acceleration. Several popular PBE solvers have been parallelized via different techniques to allow users to perform intensive calculations on parallel computers/clusters, such as *pbsa*,[137,160] APBS,[161] and DelPhi.[159] Solving the PBE on CUDA-based GPUs is much more efficient.[157,158,162] A FD scheme with the successive over-relaxation approach in the DelPhi package was implemented on NVIDIA GPUs, achieving a speedup of ∼10 times in both the linearized and nonlinear cases.[157] In the AMBER 2018 release, two new solvers were added to use NVIDIA GPUs to accelerate the FD PB calculations.[158] The GPU version of *pbsa* is called *pbsa.cuda*. It should be pointed out that only the numerical solvers are ported to the GPU platforms, while the other *pbsa* system building routines remain unchanged. A fully GPU-enabled *pbsa* and associated MM/PBSA functions are still under development. More details on the GPU version of *pbsa* are provided in the AMBER 2018 reference manual.[139]

The advantage of the PBE is that the water in the solution is reduced to a dielectric medium with a uniform dielectric constant. This treatment of the solvent greatly simplifies simulations of biomacromolecules. However, the disadvantage of implicit solvent is the use of the mean-field approximation. When a certain concentration of high-valent ions in the aqueous solution leads to ionic interactions and correlation enhancement, the PBE cannot accurately describe those kinds of systems.[130,131]

4.2. The Polar Energy Solvation Term in MM/GBSA

In MD applications, the associated computational costs are often very high, as the PBE needs to be solved every time the conformation of a molecule changes. To solve the problem, the GB model, a faster and more efficient approximation of PBE, has been developed. In a GB model, atoms are described as charged spheres whose internal dielectric constant is lower than that of the environment.[163,164] The screening that each atom experiences is determined by the local environment. The more an atom is surrounded by other atoms, the less its electrostatics will be screened since it is surrounded by lower-dielectric material. This property is called descreening of one atom by another. Different GB models calculate atomic descreening differently. Descreening is used to calculate the Born radius of each atom, and thus, the Born radius of an atom describes the degree of descreening. A large Born radius represents small screening (strong electric field), as if the atom is in vacuum. A small Born radius represents large screening (weak electric field), as if the atom is in bulk water. The canonical GB equation[165] with the absence of salt can be written as[166]

$$\Delta G_{\mathrm{GB}} = -\left(\frac{1}{\varepsilon_{\mathrm{in}}} - \frac{1}{\varepsilon_{\mathrm{sol}}}\right)\sum_{i,j}\frac{q_i q_j}{f_{\mathrm{GB}}} \tag{12}$$

in which

$$f_{\mathrm{GB}} = \sqrt{r_{ij}^{\;2} + \alpha_{ij}^{\;2}\exp\left(\frac{r_{ij}^{\;2}}{4\alpha_{ij}^{\;2}}\right)} \tag{13}$$

where r_{ij} is the distance between atoms i and j, q_i and q_j are the partial charges of those atoms, and α_{ij} is the geometric average of the efficient Born radii α_j and α_j. It is assumed that the atom is uniformly filled with a material with a low dielectric constant ($\varepsilon_{\mathrm{in}} = 1$) and the molecule is surrounded by a solvent with a high dielectric constant ($\varepsilon_{\mathrm{sol}} = 80$ for water at 300 K).

Case et al. also derived an extension of the basic GB model that allows for the treatment of mobile ions[167] by modification of eq 12 to

$$\Delta G_{\mathrm{GB}} = -\left[\frac{1}{\varepsilon_{\mathrm{in}}} - \frac{\exp(-\kappa f_{\mathrm{GB}})}{\varepsilon_{\mathrm{sol}}}\right]\sum_{i,j}\frac{q_i q_j}{f_{\mathrm{GB}}} \tag{14}$$

From the GB equation, it can be seen that the GB calculation is strongly dependent on the efficient Born radii. The first GB model implemented in the AMBER software package, which is called the GB$^{\mathrm{HCT}}$ model (*igb* = 1), was developed by Hawkins, Cramer, and Truhlar[168,169] with the parameters described by Tsui and Case.[170] Another widely used GB model, GB$^{\mathrm{OBC}}$, was developed by Onufriev, Bashford, and Case[62,110] (*igb* = 2 or 5 in AMBER). In this model, the effective Born radii are readjusted to account for the interstitial spaces between atom spheres missed by the GB$^{\mathrm{HCT}}$ approximation. As such, GB$^{\mathrm{OBC}}$ has a closer approximation to true molecular volume than GB$^{\mathrm{HCT}}$, albeit in an average sense. The GBn models (*igb* = 7 or 8 in AMBER) yield results in considerably better agreement with PB and explicit solvent than the GB$^{\mathrm{OBC}}$ models on molecular surfaces of MD snapshots under numerous circumstances.[171] The GBn model, parametrized for peptides and proteins, is not recommended for nucleic acids. The GB models have also been implemented in CHARMM,[153] referring to the works reported by Brooks et al.[163,172]

More recently, Onufriev et al. presented a grid-based surface implementation of the "R6" integration[173,174] of the GB model, named GBNSR6,[175] in which the effective Born radii were calculated numerically. The model can be described as

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

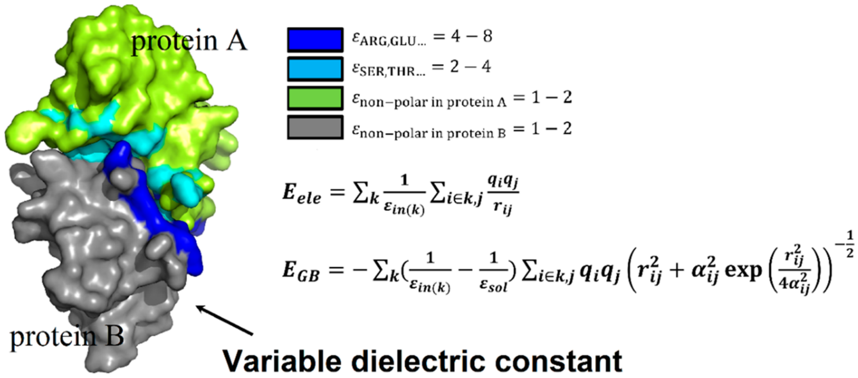


Figure 2. Graphical representation of the variable dielectric constant MM/GBSA method.

$$R_i^{-3} = -\frac{1}{4\pi} \oint_{\partial V} \frac{\mathbf{r} - \mathbf{r}_i}{|\mathbf{r} - \mathbf{r}_i|^6} \cdot d\mathbf{S} \tag{15}$$

where ∂V represents the molecular surface, d$\mathbf{S}$ is the infinitesimal surface element vector, and $\mathbf{r}_i$ and $\mathbf{r}$ are the positions of atom i and the infinitesimal surface element, respectively. To reflect the physiological conditions, the ionic strength is set to 0.145 M. The results demonstrate that the accuracy of GBNSR6 with a relatively coarse grid resolution of h = 0.5 Å in computing binding free energies for a set of small protein−ligand complexes remains in the range of $k_BT \sim 0.6$ kcal/mol, relative to the grid limit ($h \to 0$). Therefore, the default grid resolution of h = 0.5 Å is recommended because the calculations are reasonably fast on a personal computer. Compared with the second-order convergent PB solver (MIBPB[149]), GBNSR6 gives highly correlated results with r^2 = 0.97 and a root-mean-square error of 1.43 kcal/mol. Recent benchmarks show that the electrostatic binding energies computed by GBNSR6 are in good agreement with the numerical PB reference.[175,176]

4.3. Theory, Implementation, and Limitations of the Variable Dielectric Model in MM/GBSA

The continuum model is used to calculate the polar solvation energy of a system by solving the PB or GB equation. In practice, the most common user-tunable parameters for MM/PB(GB)SA include the solvent and solute dielectric constants. Among them, the solvent dielectric constant ε_{sol} represents the nature of the solvent used in the MD simulations (ε_{sol} = 80 for water). However, the solute interior dielectric constant ε_{in} is especially important in calculating the polar solvation energy, which indirectly affects the accuracy of the binding free energy prediction.[177] The solute dielectric constant ε_{in} is generally fixed with a value of 1 by default.[178] Since ligand−receptor complexes are not continuous uniform dielectric media, the choice of using a single solute dielectric constant is controversial[51,88,94,179,180] and could even lead to large errors. In particular, when sorting the ligand−receptor binding free energies, it was observed that the use of ε_{in} = 1 resulted in an overestimation of the ligand−receptor electrostatic interaction for some systems.[180−183]

Since the atomic charges used to calculate the polar solvation energy have fixed values, they cannot be adapted to respond to the dielectric changes when a solute is solvated in the solvent. Therefore, a charge model that takes the solvent effect into account is critical for the accurate calculation of solvation free energies. Applying a single dielectric constant to describe the heterogeneous dielectric environment of a solute may be problematic. However, for the sake of simplicity, a single dielectric constant is usually used for the whole solute in both the PB and GB models. Instead of using the default dielectric constant, two improved methods have been developed to find the best dielectric constant for a given molecular system that is expected to achieve the best prediction of the polar solvation energy. The first method is based on systematic scanning, in which the solute dielectric constant is scanned from 1 to 25, and the best dielectric constant strongly depends on the characteristics of the whole system.[51,88,101,184] The second approach applies variable dielectric constants for different types of residues.[179,180] For the first method, several papers have been published that explore the dependence of binding free energy predictions on the solute dielectric constants and suggest that the results are potentially improved by using a larger dielectric constant.[74,78,185−187] Genheden and Ryde estimated the optimum dielectric constant and obtained diverse results (ε_{in} = 1−25) depending on the solvation model and the tested proteins.[51] Such results have also been observed in other studies, and although the optimum value of ε_{in} is dependent on the characteristics of the binding site (a higher ε_{in} for a highly charged binding site and a lower ε_{in} for a hydrophobic site),[78,88] frequently the calculations are best with ε_{in} = 2−4,[94,101,188] especially in larger data sets of diverse proteins.[98,184]

The solvation free energy prediction method based on variable dielectric constant (Figure 2) was first tested by Ravindranathan et al.[180] on six pharmaceutically relevant targets, namely, CDK2, fXa, p38_u, PDE10A, human carbonic anhydrase, and p38_pp, in complex with several ligands. They assigned five different ε_{in} values (1, 2, 4, 8, and 20) for each type of polar or ionizable residue (Ser, Thr, Asn, Gln, His, Lys, Arg, Asp, or Glu) and assigned the same dielectric constant for the other types of residues. Then, for each system, the best set of dielectric constants evaluated in terms of r^2 and predictive index (PI) was selected and discussed. However, this approach results in only a small improvement in the r^2 and PI values compared with the standard electrostatic treatment. Especially for the systems whose binding sites composed of nonpolar residues and the ligand−receptor electrostatic interactions are negligible (PDE10A and p38_pp), the predictions are not significantly improved. Mulakala and Viswanadhan[189] predicted the binding free energies for two distinct data sets using SGB-NP,[190] VSGB-1.0,[191] and VSGB-2.0[192] (with a variable dielectric model and a novel energy function) and found that the VSGB-2.0 model may approach the accuracy needed for

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

determining the absolute free energy via linear regression without any conformational sampling.

The MM/PBSA method of variable dielectric constant has also been used to rank the inhibitory activities of a set of viral inhibitor peptide (VIRIP) mutants that have known IC_{50} values against HIV-1 gp41 fusion peptide.[179] The authors originally assigned the dielectric constant of the wild-type VIRIP−gp41 complex to 2 and set varying dielectric constants for the mutated residues. In contrast to the previously reported scheme, the dielectric constant of each mutated residue was assigned using the following rules: a value of 2 was assigned for the dielectric constant of nonpolar residues, a value of 3 for polar residues, and a value of 4 for charged residues. The authors obtained an improved correlation between the experimental activities and MM/PBSA binding energies compared with that provided by the standard method in which a single dielectric constant of 2 was used for all complexes.[179] More recently, Zhang et al. proposed a new strategy by combining the interaction entropy approach recently developed for efficient computing of the entropy change with the use of residue-type-specific dielectric constants in the framework of MM/GBSA, and they obtained optimal results for the predictions of protein−protein binding affinities with optimal ε_{in} values of 2.7 for charged residues and 1.1 for noncharged residues (r_P = 0.78 and mean absolute error = 2.8 kcal/mol).[193]

In fact, the choice of the solute dielectric constant is strictly system-dependent and requires precise study of the binding sites to obtain the most suitable ε_{in}.[88] The Spearman correlation coefficient r_S is often used to assess the correlation between experimental and predicted binding free energies. We systematically studied the effect of the solute dielectric constant on binding free energy calculations for a set of six different protein systems (α-thrombin, avidin, cytochrome *c* peroxidase, neuraminidase, P450cam, and penicillin).[88] Three different dielectric constants, ε_{in} = 1, 2, and 4, were evaluated in the PB and GB calculations. We found that the best dielectric constant is system-dependent. For the neuraminidase and α-thrombin systems, which are characterized by highly charged binding sites and the ability to form ion−ion interactions with negatively charged ligands, using ε_{in} = 4 is necessary to achieve good correlation for the MM/PBSA calculations (r_S = 0.68 and 0.81, respectively). A slightly better result was obtained using the GB^{HCT} model with ε_{in} = 4 (r_S = 0.78 and 0.90, respectively). MM/GBSA achieved good results for α-thrombin with ε_{in} = 2 (r_S = 0.88 and 0.91 for GB^{HCT} and GB^{OBC}, respectively). Yang et al. obtained consistent results for α-thrombin by applying MM/PB(GB)SA with ε_{in} of 1 and 4 to calculate the binding free energies for 28 ligands.[184] In this case, ε_{in} = 4 gave the best correlation (r^2 = 0.74 for PB and 0.72 for GB). Additionally, we have carried out related research work on the prediction of binding free energies based on variable dielectric constants and made some progress (data not published). We believe that the application of variable dielectric constants can help to improve the accuracy of binding free energy predictions.

4.4. Comparison between PB and GB

As mentioned above, the PB calculations are significantly time-consuming, especially when a finer grid mesh and a larger number of energy calculations are used to achieve better convergence.[83,149] Alternatively, the GB method, which is considered as a simple approximation to the PB method, requires much less computer resources than the PB method. For example, we performed a test and found that to perform a complete analysis for an ensemble of 100 snapshots from a constant-temperature MD simulation at 300 K, it takes several minutes to obtain the binding free energy of the popular Ras−Raf system[72] using MMPBSA.py[194] with a GB model, whereas the computational duration is ∼50 times longer with the PB model. The accuracy of the calculated energy using the GB approach is compromised at the expense of computational speed. However, the correlation and the computational demands make the GB approach attractive, especially for qualitative analysis, though the GB method in principle is not as accurate as PB.[178]

Many studies comparing the performance of MM/PBSA and MM/GBSA indicated that the predicted result strongly depends on the system being studied. Ryde et al. predicted the binding free energies of seven biotin analogues and avidin with the MM/PBSA and MM/GBSA methods.[79] They found that the GB calculation is much faster than the PB calculation but gives a less accurate result, namely, a MAD of 35 kJ/mol for MM/GBSA compared with 16 kJ/mol for MM/PBSA. Moreover, the estimated ΔG_{bind} values with MM/GBSA are coincidentally lower than the experimental data (by 8−71 kJ/mol), whereas the results with MM/PBSA are fairly well dispersed around the experimental values (average error = −0.5 kJ/mol). In our previous work, we comprehensively studied how the ranking performance of the binding free energies is influenced by the force field and partial charge model in MM/PBSA and MM/GBSA calculations.[97] In most cases, the ranking capability of MM/PBSA with Tan's parameters is better than that of MM/GBSA (GB^{OBC1}). However, several studies reported that the results predicted by MM/PBSA are worse than those by MM/GBSA.[101,195,196] The optimal prediction of MM/GBSA with a solute dielectric constant of 2.0 (r_P = 0.66) is better than using MM/PBSA (r_P = 0.49) for 98 protein−ligand complexes.[101] In addition, the MM/PBSA results are of similar quality,[96,98,197] compared with MM/GBSA. Gohlke and Case studied how the predictions depend on the polar solvation energy and suggested that the radii selected in calculations strongly impact the MM/PBSA results. Moreover, different variants of GB models result in different predictions.[66]

4.5. Efficient Entropy Calculation Methods To Estimate the Entropy Change upon Ligand Binding

It is well-known that entropy plays important role in characterizing the absolute binding free energy upon ligand−protein interaction,[198−206] but entropy estimation is usually very time-consuming. For example, NMA, which is one of the most widely used entropy estimation methods, needs to expand the covariance matrix of internal coordinates for all of the degrees of freedom and therefore is not suitable for large systems (i.e., systems with protein length >350 residues).[198] Meanwhile, many other methods also need very long simulation times to provide convergent predictions of entropy.[202,203,206] Consequently, numerous studies simply ignore the entropy term in end-point free energy calculations, which, however, leads to serious overestimation of the predicted binding free energies. Fortunately, several simplified calculation methods have been proposed in recent years, such as the truncated NMA entropy method[96] and the interaction entropy method.[114] In the truncated NMA entropy method, the protein−ligand complex is truncated into a smaller

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

structure with the center at the center of mass (CoM) of the ligand, and the residues around the ligand are retained. Usually, the truncation radius is set to 8−16 Å. The truncated structures are subsequently subjected to a traditional NMA calculation. Because of the reduced structures, the NMA calculations require much less time than those for the entire complex structures. We also assessed the performance of the truncated method in reproducing the absolute NMA entropy of the full-length structure, in which the use of a radius cutoff of 9 Å for the truncated structures is sufficient to reproduce the absolute entropies of the full-length structures for most cases.[99] Moreover, our additional evaluation on a large data set (PDBbind data set with >1700 structures) also showed that the end-point binding free energy calculations incorporating the truncated NMA entropy can, to some extent, improve the prediction accuracy both in terms of the absolute binding free energies and the correlation with experimental data.[99]

The interaction entropy method is a more recently developed entropy estimation method that considers only the fluctuations of the ligand−receptor interactions during the MD simulations and thus does not need additional computational cost.[114] This method has been successfully used in many aspects of molecular interactions such as calculating hot spots for protein−protein interactions,[193,207−210] predicting absolute binding free energies for ligand−protein complexes,[211−215] etc. In one important work, Aldeghi et al. systemically compared the accuracy of MM/PBSA and the alchemical method (TI)[211] and found that incorporation of the interaction entropy into MM/PBSA can significantly improve its performance across all of the tested protocols, indicating that the interaction entropy method is a highly efficient method to improve the performance of the end-point binding free energy calculations. Moreover, we also systemically assessed the performance of the interaction entropy method on the PDBbind data set with >1700 structures for six commonly used AMBER force fields (with 1 ns MD simulation for each system). We found that this approach can significantly improve the overall performance for both MM/GBSA and MM/PBSA in any investigated cases (including different interior dielectric constants, different force fields, etc.) for MD simulations, suggesting that this approach is a very useful tool for entropy estimation.

5. THE NONPOLAR SOLVATION ENERGY TERM IN MM/PB(GB)SA CALCULATIONS

The nonpolar contribution of the solvation energy results from solute cavity formation within the solvent and van der Waals interactions between the solute and the solvent around the cavity. The nonpolar solvation free energy is typically given by an empirical formula that is proportional to the solvent accessible surface area of the solute:

$$\Delta G_{\mathrm{np}}^{\mathrm{SA}} = \gamma \cdot \mathrm{SASA} + b \tag{16}$$

where γ is the surface tension constant and b is a correction constant (γ = 0.00542 kcal·mol^{-1}·Å^{-2} and b = 0.92 kcal/mol in the AMBER package). Regardless of the poor accuracy of the SASA model, it has been widely used in the simulations of molecular mechanics and binding affinity predictions.

The limitations of this simplified SASA model have been demonstrated previously. The total nonpolar solvation energy, with a small difference between two large components, is independent of the solute surface area or volume but nevertheless is correlated with the repulsive and attractive components of the nonpolar contribution computed from the TIP3P simulations.[216] To solve that problem, an improved method has been proposed[217,218] in which atom-specific surface tension parameters are adopted:

$$\Delta G_{\mathrm{np}}^{\mathrm{SA}} = \sum_{i=1}^{N} \gamma_i \cdot \mathrm{SASA}_i \tag{17}$$

A more modern method in which the nonpolar solvation energy is divided into cavity and dispersion (CD) terms was reported by Luo et al.[219] A cavity capable of accommodating the solute in the solvent is created, and then the nonpolar solute is introduced into the cavity. The energy for cavity formation is often estimated using a linear relation to the molecular surface (MS), similar to the SASA model. Hence, the nonpolar solvation energy should be described as

$$\Delta G_{\mathrm{np}}^{\mathrm{CD}} = \gamma \cdot \mathrm{MS} + b + \Delta G_{\mathrm{disp}} \tag{18}$$

A solvent-accessible volume (or surface) integration can be utilized to calculate the dispersion term (ΔG_{disp}). The scaling factors are typically set to γ = 0.0378 kcal·mol^{-1}·Å^{-2} and b = −0.569 kcal/mol in the AMBER package.

The polarizable continuum model (PCM),[220] with separate terms for cavitation, dispersion, and repulsion energies, usually gives more accurate results than SASA. Genheden and Ryde proposed that the nonpolar solvation energy computed by SASA is 3−8 times smaller and of the opposite sign compared with the same energy computed by PCM.[24] They studied the binding of benzene to an engineered nonpolar cavity in T4 lysozyme and found that the SASA and CD models yield similar results and that the PCM model is slightly better.[221]

Although several attempts have been made, none of the above-mentioned methods (namely, the SASA, CD and PCM methods) can yield accurate predictions for systems with more water-exposed binding sites[222] because the continuum models ignore all information about water molecules (including the number and entropy changes) before and after ligand binding.[24] One approach to solve this problem is to treat the water molecules as a part of the receptor, and improved results have been obtained for some cases;[223−225] however, the performance is strongly impacted by the number of explicit water molecules,[226] and sometimes this approach yields worse predictions.[227] Another way is to replace the desolvation in MM/GBSA by the free energy combined with displacement of binding-site water molecules upon ligand binding estimated by the WaterMap approach, which yields varying results.[183,228−230]

Unfortunately, the nonpolar distribution of the solvation free energy has received less attention, to a certain extent because the smaller value of the energy compared with the polar solvation energy. As far as we know, no relevant works on assessment of nonpolar solvation energy predictions have been reported in recent years, and we hope that some will be published in the future because the nonpolar contribution is crucial for obtaining accurate estimates of absolute hydration free energies using implicit solvent models.

6. NEW TOOLKITS AND WEB SERVERS FOR MM/PB(GB)SA CALCULATIONS

Over the past few years, numerous computational toolkits for MM/PB(GB)SA calculations have been developed and released. *MMPBSA.py* is a user-friendly Python script implemented in AMBER that automates energy analysis of

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

the snapshots extracted from an MD trajectory using ideas generated from the continuum solvent models.[194] An older Perl script called *mm_pbsa.pl* has functionality similar to that of *MMPBSA.py*, but first implicit snapshots (without the water) must be extracted from the production runs for use in the MM/PBSA and MM/GBSA calculations. Free Energy Workflow (FEW) is another set of Perl scripts (including the main script *FEW.pl* and other Perl modules) developed in AMBER to automate free energy calculations based on TI, MM/PBSA, or LIE.[231,232] To integrate high-throughput MD simulations with binding energy calculations, *g_mmpbsa* was developed as one part of the Open Source Drug Discovery (OSDD) consortium, and it implements the MM/PB(GB)SA approaches using subroutines written in-house or sourced from the GROMACS and APBS packages.[233] GMXPBSA is another user-friendly suite of Bash/Perl scripts for streamlining the MM/PBSA calculations on structural ensembles derived from GROMACS trajectories to automatically compute binding free energies for protein−protein or protein−ligand complexes.[234] The iAPBS interface written in C/C++/Fortran allows access to the APBS functionality from NAMD.[235] This module can be used to perform implicit solvent MD simulations, to write out electrostatic maps for the purpose of visualization, and to perform MM/PBSA calculations directly with NAMD. In addition, the iAPBS interface can serve as the linker between NAMD and other popular software packages like AMBER and GROMACS. FESetup is a tool to automatically set up alchemical free energy simulations for protein−ligand complexes like TI and FEP.[236] Postprocessing methods, such as MM/PBSA and LIE, are also supported.[236] In a previous work, to facilitate the prediction of binding affinities for protein−protein/ligand systems, we released an easy-to-use pipeline tool named Calculation of Free Energy (CaFE) to perform MM/PBSA and LIE calculations.[237] CaFE, powered by the VMD and NAMD programs, is capable of handling numerous static coordinates and MD trajectory file formats created by diverse molecular simulation packages and supports various force field parameters.

In addition to the MM/PB(GB)SA toolkits listed above, several Web servers based on MM/PB(GB)SA have been developed and are open to all users. ACFIS, a Web server for fragment-based drug discovery, was developed in order to improve the effectiveness of drug discovery.[238] CAND_GEN, one of the three computational modules in ACFIS, is a tool to generate hit candidates. Users can choose a binding free energy calculation method (MM/PBSA or MM/GBSA) to rescore the hit candidates. SAMPDI was designed to predict changes in protein−DNA binding free energies upon missense mutations using a modified MM/PBSA approach.[239] More recently, we developed a Web server for fast AMBER rescoring for PPI inhibitors (farPPI) that offers a freely available service for rescoring the docking poses for PPI inhibitors using MM/PB(GB)SA methods.[240]

Although MM/GBSA is more computationally efficient than most end-point free-energy calculation methods, it still takes much more time than the scoring functions commonly used in protein−protein docking. Hence, some scoring functions based on energetic terms extracted from the MM/PB(GB)SA methods have been developed and applied for molecular docking.[187] PBSA_E, a new free energy estimator based on the MM/PBSA descriptors, was developed by Zhang et al.[241] Chowdhury et al. refined docking protocols using shape complementarity, electrostatics affinity functions, and knowledge-based interface propensity and utilized the GBSA solvation energy to rerank the structures.[242] The time cost of MM/GBSA is mainly used to calculate the polar desolation energy term based on the GB model. In response to that point, our group developed HawkRank, a force-field-based scoring function with energy terms similar to those in MM/GBSA.[243] Our results show that HawkRank yields better predictions than three other scoring functions, namely, ZRANK,[244] FireDock,[245] and dDFIRE,[246] according to the total number of hits and modified success rate (MSR). Moreover, MM/GBSA rescoring is competent to distinguish correct protein−protein binding structures from decoys, and the use of HawkRank followed by MM/GBSA rescoring is an efficient protocol to improve the predictions of protein−protein docking.

7. APPLICATIONS IN SMALL-MOLECULE DRUG DESIGN

With the advantage of requiring much lower computational cost while giving prediction accuracy comparable to that of the much more time-consuming pathway methods (i.e., FEP and TI), MM/PBSA and MM/GBSA have been widely used in the field of small-molecule drug design, such as in postprocessing of structure-based virtual screening. Moreover, they are also very useful tools for analyzing the binding details of drug−target interactions since they can be conveniently decomposed into different energy terms (eqs 5−8) to capture vital region/residue receptor−ligand interactions, which are very important for rational drug design. In this section, we will review the associated progress of using MM/PB(GB)SA in small-molecule drug design.

7.1. Applications in Virtual Screening

In the early stages of structure-based virtual screening, large compound databases are usually screened using scoring functions from molecular docking to identify promising drug candidates. In molecular docking, with the conformation of a protein target determined from X-ray, NMR, or theoretical modeling, a ligand is brought close to the specific binding site of the target, and then the possible poses and conformations of the ligand are sampled. For the sake of fast computation, simple scoring functions are usually employed to estimate the binding affinity for a given docking pose. Some important energy terms, such as the solvation free energy, are simplified or totally ignored in most docking scoring functions. Therefore, a single docking scoring function may have difficulty in correctly predicting the binding poses and binding affinities in virtual screening.[247] Hence, more advanced computational methods are needed for molecular docking in rational virtual screening.[248−254] In many cases, the combination of molecular docking and MM/PB(GB)SA rescoring has proven to be a promising strategy in both the identification of the correct binding poses and the correct ranking of the binding affinities of a series of ligands.[94,96,101,255−261] For example, Sgobba et al. assessed the "screening power" of the MM/PB(GB)SA approaches in rescoring the docking conformations of ligands targeting six drug targets,[96] and they found that in most cases MM/GBSA can give a higher area under the curve (AUC) value and enrichment factor compared with the traditional docking approaches. A similar conclusion was derived by Zhang et al., who enlarged the assessment to 38 drug targets in the DUD database with up to ~0.7 million actives and decoys.[258] Besides assessing the screening power of the MM/PB(GB)SA rescoring, our group

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

also assessed the "docking power" of MM/PB(GB)SA to 98 targets, and we found that MM/GBSA rescoring can markedly improve the ratio of finding the correct binding poses of ligands in most cases (successful rate = 69.4%).[101] The definition of each assessment based "power" in molecular docking is fully presented in Liu's work.[262]

In the past decades, the use of MM/PB(GB)SA in virtual screening has been limited to as many as a few hundred of the top docking hits.[95] However, with the dramatic increase ing computer power in recent years, MM/PBSA and MM/GBSA have been applied for rescoring of thousands of compounds prescreened by molecular docking.[94,258,259] The potential power of the MM/PB(GB)SA approaches in discriminating true binders from a much larger number of decoys (the so-called screening power) has been demonstrated in high-throughput virtual screening studies.[96,263−268] For instance, using the MM/GBSA approach, Amato et al. identified a set of chemical fragments targeting PHD zinc fingers (a target once considered with low ligand ability), and they also successfully crystallized the first complex with a chemical fragment binding in the anchoring pocket of the histone binding site of PHD zinc fingers.[265] Moreover, Li's group carried out docking-based virtual screening with MM/GBSA rescoring for human dihydroorotate dehydrogenase (hDHODH).[264] They successfully identified a series of hDHODH inhibitors, and the best inhibitor from their initial virtual screening has an IC_{50} of 110 nM. A similar example was also reported by Ferreira de Freitas et al., who used MM/GBSA rescoring for the initial virtual screening and found a series of low-micromolar inhibitors targeting the HDAC6 zinc-finger ubiquitin binding domain.[266] Besides, our group employed MM/GBSA to identify actives targeting different drug targets, such as macrophage migration inhibitory factor (MIF)[269] and anaplastic lymphoma kinase (ALK).[270] All of the studies found nanomole-level actives, implying that the use of end-point methods such as MM/GBSA is indeed a very promising approach in virtual screening.

Besides the applications in virtual screening, the end-point approaches have been also used in the lead optimization stage of drug design campaigns for fast and accurate prediction of the binding affinities of the newly modified compounds.[68,73,271−274] Recently the capability of MM/PB(GB)SA rescoring in lead optimization has been investigated, and more and more advanced molecular simulations and free energy calculations with MM/PB(GB)SA have been successfully applied to the optimization of lead compounds.[73,275−277] For instance, by employing molecular docking and MM/GBSA rescoring, Xu et al. successfully found and optimized several novel-scaffold selective inhibitors targeting PfDHODH,[268] where the best optimized inhibitor has an IC_{50} of 6 nM with 40% oral bioavailability and >14000-fold species selectivity over hDHODH. Taddei et al. also used MM/GBSA to design and optimize a series of inhibitors with the 1,4,5-trisubstituted 1,2,3-triazole scaffold targeting Hsp90,[277] in which one compound, SST0287CL1, was shown to have in vitro and in vivo activities comparable to those of the clinically tested Hsp90 inhibitor NVP-AUY922. Moreover, our group also employed MM/GBSA to analyze, design, and optimize antiresistant ALK inhibitors, and in that work we found what is to our knowledge the best antiresistant inhibitor (IC_{50} = 0.27 nM) with very high binding selectivity in 35 kinases.[270]

The above discussions have shown numerous successful cases of using the end-point methods for rational drug design or lead optimization. However, as a theoretical method whose parameters come largely from finite experimental sources, these methods may also be biased to some well-tested systems and fail to correctly identify the true binders from the decoys in many cases. To alleviate this problem, our group proposed an energy-decomposition-based virtual screening method named MIEC-SVM,[277,278] which combines molecular interaction energetic components (MIECs) derived from the MM/GBSA decomposition and machine learning methods (support vector machine, SVM) to construct personalized prediction models to distinguish actives from decoys. This method has been successfully used in many cases such as designing novel inhibitors for ALK[279] and distinguishing binders from nonbinders for luciferase,[280] HIV-1 protease,[278,281,282] etc. In the case of ALK, we successfully identified seven strong novel inhibitors (<10 μM), four of which show nanomole-level activities.[279] All of the cases shown above imply that the end-point methods are indeed very promising tools for rational drug design.

Although MM/PBSA and MM/GBSA have been successfully applied in virtual screening, optimization of lead compounds, and detailed binding analyses, the prediction results are sensitive to many calculation conditions, such as the atomic charges, interior dielectric constants, MD simulation length, entropy calculations, etc.[88,97,99] Different settings may result in very different binding affinities even for the same system in study. Thus, one should keep in mind that system dependence always exists and that the selection of appropriate computational techniques depends on the characteristics of the studied system and the information available. Currently there is no universally accurate and reliable solution within reach, and innovative approaches are definitely needed, e.g., to tackle the problems arising from target flexibility and solvent molecules residing inside or around the binding pocket.

7.2. Analysis of Critical Interactions for Rational Drug Design

As mentioned above, MM/PBSA and MM/GBSA are powerful tools in optimizing lead compounds because they can quantitatively characterize the binding details (such as analyzing critical interactions) of ligand−receptor systems.[270,283−288] In drug−target interactions, it has been suggested that the electrostatic interaction dominates the non-covalent binding in molecular recognition.[289] However, this is not generally true, as it is well-known that shape complementarity is also very important.[290] Molecular recognition can therefore be attributed to contributions from both electrostatic and van der Waals interactions, solvation/desolvation and entropy effects. With a computing framework in hand, numerous studies have employed the MM/PB(GB)SA approaches to analyze the critical interactions in ligand−receptor pairs.[277,283−285,291] For example, by using MD simulations and MM/GBSA free energy decomposition, Jiang et al. analyzed the binding mechanism of known inhibitors targeting Keap1[285] and then designed a series of high-binding-affinity inhibitors to disrupt the protein−protein interaction between Keap1 and Nrf2 (a target that modulates many kinds of cancers and other chronic diseases[292,293]). In the designed inhibitors, one compound (compound **2**) for the first time reached single-digit-nanomolar activity (K_d = 3.59 nM) and showed better pharmacological properties as well. Barril et al. used molecular docking, MD simulations, and MM/PBSA calculations to investigate the binding modes for agonists targeting REV-ERBα/NCoR,[294] and among the four tested

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

compounds, three were validated with activities in their experiments. Moreover, in Kocakaya's work, MD simulations were performed for PTP1B to reveal possible mechanisms of ligand recognition and inhibition,[295] and MM/GBSA free energy decomposition was performed to give the detailed binding mechanism. The energy decomposition analysis suggested that potent and selective PTP1B inhibitors could possibly be designed by targeting the surface residues. The results were extremely consistent with the experimental work,[296] in which residues such as Arg47, Asp48, Val49, Lys120, Ala217, Ile219, Gly220, Met258, and Thr263 had been confirmed to play an important role in modulating the activities. Besides, Mena-Ulecia et al. performed a comprehensive analysis of the binding specificity of 177 thrombin inhibitors using a three-dimensional quantitative structure−activity relationship (3D-QSAR) and end-point free energy calculations.[297] Through these analyses, they inferred the effects of van der Waals contacts, electrostatic interactions, and solvation on the effectiveness of thrombin inhibitors. In addition, our group elucidated the binding mechanisms of type-$I_{1/2}$ ALK inhibitors using umbrella sampling (US) simulations and MM/GBSA binding free energy decomposition analysis.[298] We found that several residues in the hinge region (Leu1122, Leu1198, Gly1202, and Glu1210) and the allosteric pocket (Glu1197, Ile1171, Phe1174, Ile1179, His1247, Ile1268, Asp1270, and Phe1271) of ALK play vital roles in determining the relative binding strength of the studied inhibitors. The above examples suggest that the MM/PB(GB)SA methods are powerful tools in analyzing the vital regions/residues for drug−target interactions, which have been proven to provide important information for rational drug design.

Moreover, with the capability of fast characterization of the vital protein−ligand interactions, another highly useful application of MM/PB(GB)SA is to analyze the drug resistance mechanism, as they can provide more details on the energetic difference between the wild-type and mutated systems.[288,291,299−305] By virtual mutagenesis technology, one can explain how a specific mutation influences the binding of a drug to its target since the drug-resistant mutants can usually reduce the binding affinity of a drug to its target and/or change the pocket conformation of the target.[301,304,306−314] For instance, numerous drug-resistant mutations have been detected in ALK tyrosine kinase for nearly all of the launched drugs.[315−320] Using the MM/GBSA approach, a recent study revealed that the L1198F mutation of ALK results in a conformational change of the binding pocket and alters the binding affinity of ALK to the launched drugs crizotinib and lorlatinib,[321] where the critical amino acids identified by MM/GBSA free energy decomposition are in agreement with the experimental results.[319,322,323] Besides, Zhang's group investigated the resistance mechanisms of ALK mutations (I1171N, V1180L, and L1198F) to alectinib by means of MD simulations and end-point binding free energy calculations.[324] They presented a "key and lock" mechanism between the ethyl group at position 9 of alectinib and a recognition cavity in the hinge region of ALK to illustrate the major molecular origin of drug resistance. Our group has also done several works to analyze the crizotinib resistance mechanisms for ALK using the end-point methods, including the effects of the C1156Y, G1202R, R1152L, and S1206Y mutations, and the simulation results show that both the entropy effect and long-range indirect interactions can attenuate the binding of crizotinib.[307,325]

Thanks to the theoretical analysis tools, many pioneering works have successfully designed a number of antiresistant inhibitors to overcome drug-resistant mutations, such as L1196M, G1202R, C1156Y, etc. in ALK[270,326−328] and A421V, A156T, R155K, etc. in HCV,[287,329] which is encouraging for researches who want to use the end-point binding free energy calculation approaches for rational antiresistant drug design.

8. APPLICATIONS IN MACROMOLECULAR INTERACTIONS

As powerful end-point binding free energy calculation tools, MM/PBSA and MM/GBSA have also been widely used in many other fields besides small-molecule drug design. For example, a very useful application of MM/PB(GB)SA is to predict the interactions between macromolecules, such as protein−protein,[93,209,243,330−335] protein−peptide,[278,283,336−341] and protein−nucleic acid interactions.[100,239,342−352] At present, calculations of the absolute binding free energies for these problems remain very challenging for alchemical methods.

8.1. Applications in Protein−Protein Interactions

PPIs play crucial roles in most biological processes in living cells,[353−355] and numerous PPIs have been considered as potential drug targets.[356−358] The 3D structure of a protein−protein complex can provide a global scope of how and where one protein interacts with another. Nowadays, a variety of experimental techniques have been developed to explore whether there are interactions between two proteins.[359] However, it is hard to determine how two proteins interact through most of the biophysical and/or biochemical techniques without detailed structural information.[360−362] Although there are limitations, X-ray crystallography, cryogenic electron microscopy (cryo-EM), and NMR techniques are able to determine the native structure of a protein monomer or a protein−protein complex at the atomic level,[363] but solving the high-resolution structures for all of the PPIs is a considerably more difficult or even impossible task.[364,365] Therefore, computational methods have become alternatively popular to explore the interactions between two proteins in a complex.[366,367]

Protein−protein docking, which can predict the binding affinities and binding modes between individual protein structures, is a powerful approach to predict PPIs.[368] In principle, the stable conformation of a protein−protein complex can be computationally determined to be the structure with the minimum free energy on the potential energy surface.[359] To date, many efforts have been made to develop protein−protein docking methods to determine and analyze PPIs, though using just molecular docking techniques to determine the energy minima from the huge conformational space is arduous.[359] In practice, two phases are included in most protein−protein docking approaches: the docking phase and the ranking phase. In the first phase, extensive conformations are generated by molecular docking, and potential binding poses (also called decoys) are sampled from them. In the ranking stage, the decoys sampled during the first stage are scored and ranked using diverse scoring functions.[368] To allow fast screening of large molecular databases, a hierarchical scoring scheme is usually applied by

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

using a coarse scoring function as a rapid filter of the entire database in the initial docking phase and then using a more rigorous but time-consuming scoring function to rescore the top hits to produce the final ranked queries.[257]

For rescoring of macromolecule interactions, the MM/PB(GB)SA methods are thought to be greatly effective technologies since they often achieve a good balance between computational speed and accuracy compared with the alchemical methods (low in speed) and traditional molecular docking scoring functions (low in accuracy).[93,243,331] For instance, Maffucci and Contini determined the relative binding free energies for a data set of 20 protein–protein pairs using a modified MM/PBSA approach (called Nwat-MM/PBSA)[331] that explicitly incorporates the effects of water on the binding of the protein–protein pairs, and significant improvement was shown for the correlation between the predicted results and the experimental data (r^2 = 0.77, compared with 0.45 for the traditional method). We also assessed the ability of MM/PBSA and MM/GBSA to predict the binding affinities for 46 protein–protein complexes.[93] Our results show that MM/GBSA rescoring is better than ZDOCK scoring for differentiating the correct binding structures from the decoys. Therefore, considering the low computational cost and relatively high prediction accuracy, MM/GBSA is potentially an alternatively powerful tool to predict binding affinities and identify correct binding structures for protein–protein systems.

Moreover, to understand the details of protein–protein interactions, one can utilize sequence, structure, and energy-based features to reveal the binding mechanisms, such as vital residues, mutation effects, and hot spots.[207,369,370] Petukh et al. proposed a new method called SAAMBE that combines structure minimization and a modified MM/PBSA approach to estimate the effects of single and multiple mutations on protein–protein binding and suggested a crucial role of the water model and polar solvation energy in predicting the binding affinity.[371] The core of this modified MM/PBSA method is to use a residue-specific dielectric constant protocol to characterize the mutation effects, and it achieves a good correlation (r_P = 0.75) between the predictions and the experimental data for 1300 mutations in 43 proteins. Besides, alanine scanning, proposed by Massova and Kollman in 1999,[372] is another useful technology to analyze hot spots for PPIs. Combined with the MM/PB(GB)SA methods, this technology has proven to be a very useful approach for revealing specific features of protein–protein interactions, and it has been successfully applied to insulin dimer,[373] IGF-II/IGF2R,[374] Ras–Raf and Ras–RalGDS,[72] etc. For example, the hot spots in the IGF-II/IGF2R complex identified by MM/PB(GB)SA-based alanine scanning[374] are consistent with the reported experimental mutagenesis data.[375−378]

Besides their applications in PPI predictions, the MM/PB(GB)SA methods are usually used in studies of protein–peptide interactions (PpIs), which is also a very important field for molecular design since PpIs can mimic the interaction patterns of protein–protein systems to regulate biological processes.[278,283,336−340] For example, using MD simulations and MM/GBSA free energy analyses, Xu et al. studied the binding pattern of an eight-residue peptide targeting TNKS and subsequently designed several constrained peptide inhibitors (called macrocyclized extended peptides) that can effectively inhibit the biological activity of the target and overcome the off-target phenomenon.[283] Moreover, the above-mentioned MIEC-SVM method proposed by our group has been successfully used in many protein–peptide systems, such as chromodomain-methyllysine-binding peptides,[339] ABL1-SH3-binding peptides,[278,340,341] PKA-RIIα-binding peptides,[338] etc., where many predicted tight-binding peptides have been validated by experimental assays.[339−341]

8.2. Applications in Protein–Nucleic Acid Interactions

Besides protein–protein/peptide interactions, protein–nucleic acid (RNA/DNA) interactions also play crucial roles in many biological processes, such as regulation of gene expression, RNA splicing, protein synthesis, etc. Most RNAs function only when in complex with specific proteins. Therefore, revealing the specific protein–RNA recognitions and binding patterns is crucial for both understanding the important processes of life and designing new drugs.[379] Partly because the existing scoring functions for protein–RNA interactions (PRIs) are unreliable, accurately predicting the 3D structures and binding affinities for PRIs is still quite difficult. To solve this problem, the MM/PB(GB)SA approaches have also been employed in studying the PRIs.[347,348,352] For instance, Orr et al. presented a computational tool to accurately characterize the interactions between proteins and RNA with post-transcriptional modifications,[348] in which they used MM/GBSA to predict whether an RNA modification is favorable for binding with the target protein, and the predictions showed a very high correlation with the experimental data (r^2 > 0.9). Moreover, we systemically investigated the performance of MM/PBSA and MM/GBSA to predict the binding affinities and identify the near-native binding structures for 148 protein–RNA systems with different solvent models and solute dielectric constants.[100] The results showed that MM/GBSA rescoring efficiently improves the prediction capability of the scoring functions for protein–RNA systems (r_P = 0.58), especially for the binding poses generated from ZDOCK_M, a modified ZDOCK program.[380]

With regard to protein–DNA interactions (PDIs), several recent theoretical works involving mechanistic analyses[350,351,381−383] and methodology development[239,349] have been reported. For example, Peng et al. developed a modified MM/PBSA method to predict the binding free energy difference arising from missense mutations to protein–DNA complexes,[239] and a high correlation coefficient (r_P = 0.72) was reached for the test set containing 105 mutations covering 13 protein–DNA systems. Moreover, combining fast side-chain optimization algorithms and the MM/PBSA approach, Li's group developed a new algorithm called PremPDI to predict the effects of missense mutations on the binding of protein–DNA complexes.[349] In a data set of 49 protein–DNA complexes containing 219 mutations, they also achieved a high Pearson correlation coefficient between the predicted results and the experimental data (r_P = 0.71). All of these examples imply that the use of end-point binding free energy calculation methods to predict macromolecule interactions is feasible.

The ribosome is the place where protein synthesis occurs, and in bacteria it consists of small (30S) and large (50S) subunits with tens of proteins and a sequence of rRNAs.[384,385] The synthesis of a protein starts with binding of an mRNA to the ribosomal 30S subunit, and then in the elongation phase the nascent peptide is extended with local folding from the peptidyl transferase center (PTC) through an internal tunnel into the large ribosomal 50S subunit.[386,387] Finally, the nascent peptide escapes from the tunnel and folds into its native

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

structure with the help of other factors.[388] This process involves multiple interactions, including the binding of peptide/antibiotics to rRNAs or proteins.[389−392] The MM/PBSA method was first used to investigate the binding of aminoglycoside derivatives to the A site of the ribosome, and the predicted binding free energies were found to be in good agreement with the experimental values.[393] Moreover, MD simulations suggested that additional stability to the bases A1492 and A1493 in their extrahelical forms is provided by well-designed compounds. A conformational transition in the aminoacyl tRNA site of the bacterial ribosome both in the absence and presence of an aminoglycoside antibiotic was reported by Mobashery et al.[394]

9. QM IN MM/PB(GB)SA CALCULATIONS

Although the MM/PBSA and MM/GBSA methods have been successfully applied to many problems, particularly in estimating free energies for binding of small ligands (drug candidates) to proteins, it is well-known that the MM energy model has some limitations for accurate prediction of free energy. Although sufficient sampling is required for suitable convergence of free energy calculations, the results strongly depend on the quality of the MM potential. For some interesting systems, such as transition states and metal-binding sites, the standard MM potentials may perform poorly. Hence, it is desirable to utilize the more versatile QM approaches for these systems.[395−398] However, it remains true that the highest levels of QM can be used only for reasonably modest system sizes (tens of atoms).[399] A number of general strategies have been employed to further extend the size of systems to which QM calculations can be applied. One of the most popular approaches is to describe a subregion of interest via QM and couple it to its larger environment modeled at the MM level (so-called QM/MM simulation). The fragment molecular orbital (FMO)[400−402] method and linear scaling strategies[398,403] are usually used to increase the reach of QM methods. Another useful strategy is to use a truncated system for the QM calculation. It has been reported that the average truncation error of QM calculations does not reach 1 kJ/mol with the radius of the system truncated to 8.5 Å after the MD simulation of the full-length protein system.[404] Recently, the adaptation of QM algorithms to utilization of GPU architectures has also been reported.[405]

Nowadays, QM has been widely applied in the prediction of protein−ligand docking,[120,123,406−410] protein−ligand binding affinities (scoring),[411−415] and changes in ligand internal energy upon binding (ligand strain).[416−418] For example, Raha and Merz used semiempirical QM to design a scoring function (QMScore) and calculated the solvation free energies and electrostatic interactions for a diverse set of 165 protein−ligand complexes.[411] They obtained encouraging results that the square of the correlation coefficient between the predicted and experimental values reached up to 0.55. Moreover, comparison of QMScore with 11 other scoring functions for a set of 56 protein−ligand complexes from Wang's data set[419] showed that QMScore gave the best performance. Wu et al. studied the binding mechanism of L86 and T76 to human α-thrombin using the molecular fractionation with conjugate caps (MFCC) and MM/PBSA approaches, and the results showed that the L86/T76-thrombin binding interactions given by MFCC and MM/PBSA are consistent and in good agreement with the experimental data.[396]

Ryde et al. proposed an approach to predict free energies of reactions in proteins, called QM−MM/PBSA.[395] In their approach, the internal energy of the reactive site is calculated via QM, while the internal interactions with the surrounding protein are computed at the MM level. They found that QM−MM/PBSA reproduced the results of a strict QM/MM FEP method with a MAD of 8−10 kJ/mol if multiple frames extracted from the MD trajectories were employed and 4−14 kJ/mol if a single QM/MM structure was employed. Sippl et al. applied QM−MM/PBSA rescoring to search for novel Myt1 kinase inhibitors.[410] The QM−MM/GBSA scoring performed better than docking scoring functions or MM/PB(GB)SA in discriminating active from inactive compounds and could be used on a data set with diverse scaffolds. More recently, Mishra and Koca assessed the performance of the MM/PBSA, MM/GBSA, and QM−MM/GBSA approaches on protein−carbohydrate complexes.[107] On the basis of the GB^{HCT} model and the PM6 or DFTB method, QM−MM/GBSA resulted in a marginally improved agreement ($r^2 = 0.96$) with the experimental binding energies compared with MM/PBSA with the mbondi radii set, indicating that the QM Hamiltonian may have a notable impact on the QM−MM/GBSA predictions. They suggested that PM6 may be more suitable for virtual screening involving thousands of compounds because the DFTB/SCC-DFTB calculations are computationally much more demanding. In another work based on 6 ns MD simulation trajectories together with GB^{GBn2}, PM3, and the mbondi2 radii set, QM−MM/GBSA generated the best correlation with the experimental results ($r^2 = 0.88$).[87] However, inclusion of QM methods does not always improve the prediction results of binding affinities and sometimes can even lead to much worse predictions than MM methods.[188,415,420] Ryde et al. used this approach to estimate the binding affinities of ligands to cathepsin B, and the results indicated that the QM−MM/PBSA predictions ($r^2 = 0.59$) were much worse than the predictions based only on gas-phase QM energies ($r^2 = 0.80$),[421] whereas accurate QM−MM/PBSA predictions were obtained for cytochrome P450,[422] highlighting the system dependence of QM−MM/PB(GB)SA. Many more details are provided in the outstanding review of the use of QM in the prediction of ligand binding affinities.[423]

10. COMPARISON WITH OTHER PREDICTIVE METHODS

The LIE method, another popular end-point approach, has usually been used to compare with MM/PB(GB)SA in binding free energy predictions,[51,54,68,181,197] but it is difficult to determine which one is more accurate. However, it seems that the LIE approach is highly system-dependent.[197] For example, the standard LIE method yields no correlation for the CB8 systems but an excellent correlation for the α-CD systems. Ryde et al. compared the computational efficiencies and accuracies of LIE and MM/PBSA and found that LIE is 2−7 times more efficient than MM/PBSA in computational cost.[68]

The much more rigorous alchemical methods (such as FEP and TI) have been compared with MM/PB(GB)SA. because of the inherent limitations of the MM/PB(GB)SA approaches, they are less accurate than the alchemical methods[78,271,273,424,425] in many cases. However, in some cases, the MM/PB(GB)SA methods give comparable[182,426] or even better[427] predictions than the alchemical methods.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

Scoring functions are more computationally efficient and are widely used in the early stage of virtual screening and estimation of binding free energies of protein−ligand interactions. Generally, however, the classical scoring functions are inherently inaccurate. For example, we evaluated the performance of MM/PB(GB)SA to predict the binding free energies and identify the correct binding conformations for 98 protein−ligand complexes, and the results showed that MM/GBSA performed better than almost all of the scoring functions in molecular docking to identify the correct binding structures and rank the binding affinities for the tested protein−ligand systems.[101] Another comparative evaluation study reported by Obiol-Pardo and Rubio-Martinez[428] suggested that compared with XScore,[429] MM/PBSA performed better in identifying small differences upon ligand binding for seven XIAP−peptide complexes. In recent years, machine-learning scoring functions for molecular docking have rapidly developed,[430−434] but reviewing the advances of machine learning in drug design and PPIs is beyond the scope of this review.

11. CONCLUDING REMARKS

Free energy calculations offer an estimation of the energy differences between thermodynamic processes given a series of specific parameters and physical hypotheses. A good free energy estimation approach aims to achieve convergence to a unique value solely determined by the free energy model (such as MM/PBSA or MM/GBSA) itself. This unique value is called the “correct” free energy of a given molecular system for the model. A free energy model can be improved by adjusting the employed parameters to reproduce the experimental data. It is critical to apply the “correct” free energies in the fitting procedure, and care must be taken to avoid overfitting.

MM/PB(GB)SA calculations are popular because they provide a good balance between computational speed and accuracy. The performance of the predictions depends on both the sampling accuracy of the entire conformation space and the quality of the free energy model. To obtain the “correct” free energy for a molecular system with MM/PB(GB)SA calculations, many independent simulations for sampling may be necessary to achieve reasonable convergence. In practice, however, the MM/PB(GB)SA free energies may deviate away from the “correct” values because of inadequate conformational sampling, leading to a poor MAD even though the relative free energies measured by r^2 are still satisfactory. An MM/PB(GB)SA model is a combination of functional forms for different energy terms and the associated parameters, such as the force fields for molecular mechanics energies and the solute dielectric constants and atomic radii for solvation free energies calculated with the PB/GB theories. These parameters fundamentally affect the predictive accuracy of various energy terms.

The solvation free energy is a pivotal term since the solvent effect plays a key role in protein−protein/ligand binding, protein tertiary structure formation, and execution of protein functions. Thus, an accurate treatment of the solvation term is fundamental to compute binding free energies. The polar contribution of the solvation free energy is predicted by solving the PBE. Unfortunately, PBE solvers do not converge with respect to the numerical grid used, except for MIBPB. Theoretically, the PB method for solving the PBE is more accurate than GB. However, the GB model has gained popularity because of its favorable computational speed and comparable or even better accuracy compared with the PB method. The dielectric constant strongly impacts the prediction of the polar solvation free energy, and the variable dielectric model illustrates potential power. The nonpolar contribution of solvation free energy is represented by a linear relation to the SASA. To our knowledge, in recent years nonpolar solvation energy predictions have become less attractive, and we hope that many efforts will be made in the future.

It is noted that the MM/PB(GB)SA free energy model may be unable to perform well when a binding site involves a highly charged environment, for which an accurate treatment of electrostatic interactions is crucial for binding free energy calculations. The calculation uncertainty increases with the polarity of the studied compounds, especially when polar compounds bind to a binding pocket formed by highly charged residues located in the interior of a biomacromolecule. Furthermore, the contribution of structural water molecules residing inside or around the binding pocket cannot be taken into account well in binding free energy calculations with implicit solvation models. Explicit consideration of such water molecules in free energy calculations may significantly improve the accuracy of an MM/PB(GB)SA model. Another main source of error for MM/PB(GB)SA free energy models is the conformational entropies, which are typically estimated by performing NMA. As the anharmonic contribution is neglected in NMA, the conformational entropies obtained using NMA may have systematic errors for some molecular systems. A universally accurate and reliable solution is currently out of reach, and innovative methods are much needed.

Using the PB method to calculate the electrostatics of large systems in high-throughput virtual screening requires large computational resources, including computing time and memory. Therefore, it is necessary to speed up PB calculations via parallelization on multiple CPUs or GPU acceleration. Nevertheless, MM/PBSA and MM/GBSA perform better in determining relative free energies and can be used for postprocessing of docked structures or to rationalize observed differences. However, they cannot serve as a basis for developing more accurate methods and predict true drug candidates in drug design without experimental verification because of their relatively limited accuracy (compared with FEP and TI). Despite this, the end-point free energy techniques are expected to play a more and more important role in detailed energetic investigations of complex formation as the MM/PB(GB)SA free energy models are continually improved in the future.

Lastly, it is worth mentioning that machine learning and mathematical methods, especially the former, have gained much popularity recently. However, it is beyond the scope of the current review to discuss these methods here.

Critical Analysis and Remaining Challenges: With the well-defined algorithms and good balance between computational efficiency and prediction accuracy, MM/PBSA and MM/GBSA have been regarded as competitive methods in the field of binding free energy calculations, and they have been successfully applied in many aspects of drug design, including structure-based virtual screening, lead optimization, and molecular recognition. However, their computational efficiency is achieved by introducing controversial approximations to both the sampling and energy calculation phases, such as using a uniform dielectric constant for the whole solute surrounded by a complicated local microenvironment, ignoring or

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

improperly treating ions and critical water molecules in the binding site, ignoring or applying oversimplified methods to estimate conformational and solvation entropies, etc. Fortunately, the performance of MM/PBSA and MM/GBSA can be further improved by taking some remedial measures and introducing new techniques to them. First, MM/GBSA based on a variable dielectric model, in which variable dielectric constants are assigned to different residues in a protein–protein/ligand complex, could be a good idea. As a matter of fact, we believe that the variable dielectric model is one of the most promising ways to make a breakthrough in the two end-point free energy methods. Second, the implementation of MM/PB(GB)SA on GPUs will greatly accelerate the sampling and energy calculations. Moreover, artificial intelligence, from artificial neural networks to deep learning, may find great applications in MM/PB(GB)SA-based binding free energy calculation and structure prediction in the coming years, especially when big data of high-quality experimental structures and binding data become available.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: tingjunhou@zju.edu.cn (T.H.).
*E-mail: john.zhang@nyu.edu (J.Z.H.Z.).

ORCID

Huiyong Sun: 0000-0002-7107-7481
Junmei Wang: 0000-0002-9607-8229
Hui Liu: 0000-0002-1267-2475
John Z. H. Zhang: 0000-0003-4612-1863
Tingjun Hou: 0000-0001-7227-2580

Author Contributions

$^{\nabla}$E.W. and H.S. contributed equally.

Notes

The authors declare no competing financial interest.

Biographies

Ercheng Wang received his Ph.D. in physics from Huazhong University of Science and Technology (HUST), working on theoretical studies of the cotranslational protein folding mechanism. He is currently a postdoctoral fellow in the College of Pharmaceutical Sciences at Zhejiang University, where his research focuses on computational studies of protein–protein/ligand interactions and method development for binding free energy calculations.

Huiyong Sun received his Ph.D. from the Institute of Functional Nano & Soft Materials (FUNSOM) at Soochow University. Currently he is an associate professor in the School of Pharmacy at China Pharmaceutical University. His research interest is in computer-aided drug design. To date he has coauthored more than 50 publications in peer-reviewed journals.

Junmei Wang received his Ph.D. from Peking University and then studied as a postdoctoral associate with Dr. Peter Kollman at the University of California San Francisco. Currently he is an associate professor of computational chemistry and biophysics and a member of the Computational Chemical Genomics Screening Center in the School of Pharmacy at the University of Pittsburgh. He and his collaborators developed a set of popular AMBER force fields (FFs), such as FF99, GAFF, and polarizable FFs based on atomic dipole interactions, as well as the Antechamber module implemented in AMBER. He has published ~100 papers in peer-reviewed journals with an h-index of 36.

Zhe Wang received his B.S. from Wenzhou Medical University. He is currently a Ph.D. candidate in the College of Pharmaceutical Sciences at Zhejiang University. His research focuses on the development, evaluation, and optimization of combined strategies for structure-based virtual screening.

Hui Liu obtained his B.S. in pharmacology in 2005 and his Ph.D. in biology in 2010 from Wuhan University. Then he undertook postdoctoral research in the laboratory of Tingjun Hou at Zhejiang University, and he is currently an associate professor at Huazhong Agricultural University. His research focuses on molecular simulations and computer-aided drug design.

John Z. H. Zhang received his Ph.D. in chemical physics in 1987. His current research focuses on computational studies of biomolecular systems, including protein–ligand binding, protein–protein interaction, protein design, etc. He is currently the director of the NYU–ECNU Center for Computational Chemistry at NYU Shanghai and a professor of chemistry at NYU Shanghai and ECNU. He has won many awards and fellowships, including the Camille and Henry Dreyfus New Faculty Award (1990), the National Science Foundation Presidential Faculty Fellowship (1994), the Alfred P. Sloan Foundation Research Fellowship (1995), and the Camille Dreyfus Teacher-Scholar Award (1995). He has over 350 publications in peer-reviewed journals with an h-index of 55.

Tingjun Hou received his Ph.D. in computational chemistry from Peking University in 2002. His research focuses on molecular modeling and computer-aided drug design, including the development of structure-based virtual screening methodologies, theoretical predictions of ADMET and drug-likeness, discovery of small-molecule inhibitors toward important drug targets, and multiscale simulations of target–ligand recognition. Currently, he is a full professor in the College of Pharmaceutical Sciences at Zhejiang University and the director of the Drug Informatics and Computational Biology Center in the Hangzhou Institute of Innovative Medicine at Zhejiang University. He has coauthored more than 320 publications in peer-reviewed journals with an h-index of 49.

ACKNOWLEDGMENTS

This work was financially supported by the National Key R&D Program of China (2016YFA0501700 and 2016YFB0201700), the National Natural Science Foundation of China (21575128, 81773632, 21433004, and 91753103), the Innovation Program of Shanghai Municipal Education Commission (201701070005E00020), and the National Institutes of Health (R01-GM079383 and R21-GM097617).

REFERENCES

(1) Christ, C. D.; Mark, A. E.; van Gunsteren, W. F. Basic Ingredients of Free Energy Calculations: A Review. *J. Comput. Chem.* **2009**, *31*, 1569–1582.
(2) Vangunsteren, W. F.; Berendsen, H. J. C. Computer-Simulation of Molecular-Dynamics - Methodology, Applications, and Perspectives in Chemistry. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992–1023.
(3) Hospital, A.; Goni, J. R.; Orozco, M.; Gelpi, J. L. Molecular Dynamics Simulations: Advances and Applications. *Adv. Appl. Bioinform. Chem.* **2015**, *8*, 37–47.
(4) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, *330*, 341–346.
(5) Li, A. J.; Daggett, V. Characterization of the Transition-State of Protein Unfolding by Use of Molecular-Dynamics - Chymotrypsin Inhibitor-2. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 10430–10434.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX–XXX

(6) Plattner, N.; Doerr, S.; De Fabritiis, G.; Noe, F. Complete Protein−Protein Association Kinetics in Atomic Detail Revealed by Molecular Dynamics Simulations and Markov Modelling. *Nat. Chem.* **2017**, *9*, 1005−1011.

(7) Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M. Free-Energy of Hydrophobic Hydration - a Molecular-Dynamics Study of Noble-Gases in Water. *J. Chem. Phys.* **1986**, *85*, 6720−6727.

(8) Han, K. K.; Son, H. S. On the Isothermal−Isobaric Ensemble Partition Function. *J. Chem. Phys.* **2001**, *115*, 7793−7794.

(9) Corti, D. S. Isothermal−Isobaric Ensemble for Small Systems. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2001**, *64*, 016128.

(10) Aldeghi, M.; Bluck, J. P.; Biggin, P. C. Absolute Alchemical Free Energy Calculations for Ligand Binding: A Beginner's Guide. *Methods Mol. Biol.* **2018**, *1762*, 199−232.

(11) Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Predictions of Ligand Selectivity from Absolute Binding Free Energy Calculations. *J. Am. Chem. Soc.* **2017**, *139*, 946−957.

(12) Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303*, 1813−1818.

(13) Shirts, M. R. Best Practices in Free Energy Calculations for Drug Design. *Methods Mol. Biol.* **2012**, *819*, 425−467.

(14) Christ, C. D.; Fox, T. Accuracy Assessment and Automation of Free Energy Calculations for Drug Design. *J. Chem. Inf. Model.* **2014**, *54*, 108−120.

(15) Williams-Noonan, B. J.; Yuriev, E.; Chalmers, D. K. Free Energy Methods in Drug Design: Prospects of "Alchemical Perturbation" in Medicinal Chemistry. *J. Med. Chem.* **2018**, *61*, 638−649.

(16) Noskov, S. Y.; Lim, C. Free Energy Decomposition of Protein−Protein Interactions. *Biophys. J.* **2001**, *81*, 737−750.

(17) Rietman, E. A.; Platig, J.; Tuszynski, J. A.; Lakka Klement, G. Thermodynamic Measures of Cancer: Gibbs Free Energy and Entropy of Protein−Protein Interactions. *J. Biol. Phys.* **2016**, *42*, 339−350.

(18) Pitera, J. W.; van Gunsteren, W. F. One-Step Perturbation Methods for Solvation Free Energies of Polar Solutes. *J. Phys. Chem. B* **2001**, *105*, 11264−11274.

(19) Grossfield, A.; Ren, P.; Ponder, J. W. Ion Solvation Thermodynamics from Simulation with a Polarizable Force Field. *J. Am. Chem. Soc.* **2003**, *125*, 15671−15682.

(20) Oostenbrink, B. C.; Pitera, J. W.; van Lipzig, M. M.; Meerman, J. H.; van Gunsteren, W. F. Simulations of the Estrogen Receptor Ligand-Binding Domain: Affinity of Natural Ligands and Xenoestrogens. *J. Med. Chem.* **2000**, *43*, 4594−4605.

(21) Bishop, T. C.; Williams, K. Y. Molecular Dynamics Simulations of the Estrogen Receptor Ligand Binding Domain. *Biophys. J.* **2002**, *82*, 486a.

(22) Williams, K. Y.; Bishop, T. C. Free Energy Simulations of the Estrogen Receptor Ligand Binding Domain Using MM/PBSA. *Biophys. J.* **2003**, *84*, 462a.

(23) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. Direct Calculation of the Binding Free Energies of FKBP Ligands. *J. Chem. Phys.* **2005**, *123*, 084108.

(24) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods To Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discovery* **2015**, *10*, 449−461.

(25) Zhou, R. Free Energy Landscape of Protein Folding in Water: Explicit Vs. Implicit Solvent. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 148−161.

(26) Lei, H.; Wu, C.; Liu, H.; Duan, Y. Folding Free-Energy Landscape of Villin Headpiece Subdomain from Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 4925−4930.

(27) Otzen, D. E.; Kristensen, O.; Proctor, M.; Oliveberg, M. Structural Changes in the Transition State of Protein Folding: Alternative Interpretations of Curved Chevron Plots. *Biochemistry* **1999**, *38*, 6499−6511.

(28) Ytreberg, F. M.; Swendsen, R. H.; Zuckerman, D. M. Comparison of Free Energy Methods for Molecular Systems. *J. Chem. Phys.* **2006**, *125*, 184114.

(29) Brandsdal, B. O.; Osterberg, F.; Almlof, M.; Feierberg, I.; Luzhkov, V. B.; Aqvist, J. Free Energy Calculations and Ligand Binding. *Adv. Protein Chem.* **2003**, *66*, 123−158.

(30) Hu, X.; Shelver, W. H. Docking Studies of Matrix Metalloproteinase Inhibitors: Zinc Parameter Optimization to Improve the Binding Free Energy Prediction. *J. Mol. Graphics Modell.* **2003**, *22*, 115−126.

(31) Jayachandran, G.; Shirts, M. R.; Park, S.; Pande, V. S. Parallelized-over-Parts Computation of Absolute Binding Free Energy with Docking and Molecular Dynamics. *J. Chem. Phys.* **2006**, *125*, 084901.

(32) 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.

(33) Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. Comprehensive Evaluation of Ten Docking Programs on a Diverse Set of Protein−Ligand Complexes: The Prediction Accuracy of Sampling Power and Scoring Power. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964−12975.

(34) Li, Y.; Shen, J.; Sun, X.; Li, W.; Liu, G.; Tang, Y. Accuracy Assessment of Protein-Based Docking Programs against RNA Targets. *J. Chem. Inf. Model.* **2010**, *50*, 1134−1146.

(35) Hou, X. B.; Du, J. T.; Zhang, J.; Du, L. P.; Fang, H.; Li, M. Y. How to Improve Docking Accuracy of Autodock4.2: A Case Study Using Different Electrostatic Potentials. *J. Chem. Inf. Model.* **2013**, *53*, 188−200.

(36) Yuriev, E.; Holien, J.; Ramsland, P. A. Improvements, Trends, and New Ideas in Molecular Docking: 2012−2013 in Review. *J. Mol. Recognit.* **2015**, *28*, 581−604.

(37) Meng, Y.; Sabri Dashti, D.; Roitberg, A. E. Computing Alchemical Free Energy Differences with Hamiltonian Replica Exchange Molecular Dynamics (H-REMD) Simulations. *J. Chem. Theory Comput.* **2011**, *7*, 2721−2727.

(38) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.

(39) Abel, R.; Wang, L.; Mobley, D. L.; Friesner, R. A. A Critical Review of Validation, Blind Testing, and Real-World Use of Alchemical Protein−Ligand Binding Free Energy Calculations. *Curr. Top. Med. Chem.* **2017**, *17*, 2577−2585.

(40) Clark, A. J.; Gindin, T.; Zhang, B.; Wang, L.; Abel, R.; Murret, C. S.; Xu, F.; Bao, A.; Lu, N. J.; et al. Free Energy Perturbation Calculation of Relative Binding Free Energy between Broadly Neutralizing Antibodies and the Gp120 Glycoprotein of HIV-1. *J. Mol. Biol.* **2017**, *429*, 930−947.

(41) Hirono, S.; Kollman, P. A. Calculation of the Relative Binding Free Energy of 2'gmp and 2'amp to Ribonuclease T1 Using Molecular Dynamics/Free Energy Perturbation Approaches. *J. Mol. Biol.* **1990**, *212*, 197−209.

(42) Reddy, M. R.; Erion, M. D. Calculation of Relative Binding Free Energy Differences for Fructose 1,6-Bisphosphatase Inhibitors Using the Thermodynamic Cycle Perturbation Approach. *J. Am. Chem. Soc.* **2001**, *123*, 6246−6252.

(43) Mutyala, R.; Reddy, R. N.; Sumakanth, M.; Reddanna, P.; Reddy, M. R. Calculation of Relative Binding Affinities of Fructose 1,6-Bisphosphatase Mutants with Adenosine Monophosphate Using Free Energy Perturbation Method. *J. Comput. Chem.* **2007**, *28*, 932−937.

(44) Kamath, S.; Coutinho, E.; Desai, P. Calculation of Relative Binding Free Energy Difference of DHFR Inhibitors by a Finite Difference Thermodynamic Integration (FDTI) Approach. *J. Biomol. Struct. Dyn.* **1999**, *16*, 1239−1244.

(45) Wu, K. W.; Chen, P. C.; Wang, J.; Sun, Y. C. Computation of Relative Binding Free Energy for an Inhibitor and Its Analogs Binding

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

with Erk Kinase Using Thermodynamic Integration MD Simulation. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1159−1169.
(46) Lawrenz, M.; Baron, R.; Wang, Y.; McCammon, J. A. Independent-Trajectory Thermodynamic Integration: A Practical Guide to Protein-Drug Binding Free Energy Calculations Using Distributed Computing. *Methods Mol. Biol.* **2012**, *819*, 469−486.
(47) Wang, C. X.; Shi, Y. Y.; Zhou, F.; Wang, L. Thermodynamic Integration Calculations of Binding Free Energy Difference for Gly-169 Mutation in Subtilisin Bpn'. *Proteins: Struct., Funct., Genet.* **1993**, *15*, 5−9.
(48) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; et al. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695−2703.
(49) Lee, T. S.; Hu, Y.; Sherborne, B.; Guo, Z.; York, D. M. Toward Fast and Accurate Binding Affinity Prediction with Pmemdgti: An Efficient Implementation of GPU-Accelerated Thermodynamic Integration. *J. Chem. Theory Comput.* **2017**, *13*, 3077−3084.
(50) Swanson, J. M.; Henchman, R. H.; McCammon, J. A. Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophys. J.* **2004**, *86*, 67−74.
(51) Genheden, S.; Ryde, U. Comparison of End-Point Continuum-Solvation Methods for the Calculation of Protein−Ligand Binding Free Energies. *Proteins: Struct., Funct., Genet.* **2012**, *80*, 1326−1342.
(52) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. Examining Methods for Calculations of Binding Free Energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA Calculations of Ligands Binding to an HIV Protease. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 393−407.
(53) Aqvist, J.; Medina, C.; Samuelsson, J. E. A New Method for Predicting Binding Affinity in Computer-Aided Drug Design. *Protein Eng., Des. Sel.* **1994**, *7*, 385−391.
(54) Adekoya, O. A.; Willassen, N. P.; Sylte, I. Molecular Insight into Pseudolysin Inhibition Using the MM-PBSA and LIE Methods. *J. Struct. Biol.* **2006**, *153*, 129−144.
(55) Hansson, T.; Marelius, J.; Aqvist, J. Ligand Binding Affinity Prediction by Linear Interaction Energy Methods. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27−35.
(56) Srinivasan, J.; Miller, J.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of RNA Hairpin Loops and Helices. *J. Biomol. Struct. Dyn.* **1998**, *16*, 671.
(57) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(58) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate−DNA Helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(59) Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, C. L., 3rd. 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.
(60) Onufriev, A.; Case, D. A.; Bashford, D. Effective Born Radii in the Generalized Born Approximation: The Importance of Being Perfect. *J. Comput. Chem.* **2002**, *23*, 1297−1304.
(61) Bashford, D.; Case, D. A. Generalized Born Models of Macromolecular Solvation Effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129−152.
(62) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712−3720.
(63) Zoete, V.; Irving, M. B.; Michielin, O. MM-GBSA Binding Free Energy Decomposition and T Cell Receptor Engineering. *J. Mol. Recognit.* **2010**, *23*, 142−152.
(64) Bello, M. Binding Free Energy Calculations between Bovine Beta-Lactoglobulin and Four Fatty Acids Using the MMGBSA Method. *Biopolymers* **2014**, *101*, 1010−1018.
(65) Yi, C. H.; Chen, J. Z.; Shi, S. H.; Hu, G. D.; Zhang, Q. G. A Computational Analysis of Pyrazole-Based Inhibitors Binding to Hsp90 Using Molecular Dynamics Simulation and the MM-GBSA Method. *Mol. Simul.* **2010**, *36*, 454−460.
(66) Gohlke, H.; Case, D. A. Converging Free Energy Estimates: MM-PB(GB)SA Studies on the Protein−Protein Complex Ras−Raf. *J. Comput. Chem.* **2004**, *25*, 238−250.
(67) Gohlke, H.; Case, D. A. Insights into Protein−Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition Using a Generalized Born Model. *Abstr. Pap. - Am. Chem. Soc.* **2003**, *225*, U791.
(68) Genheden, S.; Ryde, U. Comparison of the Efficiency of the LIE and MM/GBSA Methods to Calculate Ligand-Binding Energies. *J. Chem. Theory Comput.* **2011**, *7*, 3768−3778.
(69) Nicolotti, O.; Convertino, M.; Leonetti, F.; Catto, M.; Cellamare, S.; Carotti, A. Estimation of the Binding Free Energy by Linear Interaction Energy Models. *Mini-Rev. Med. Chem.* **2012**, *12*, 551−561.
(70) He, X. B.; Man, V. H.; Ji, B. H.; Xie, X. Q.; Wang, J. M. Calculate Protein−Ligand Binding Affinities with the Extended Linear Interaction Energy Method: Application on the Cathepsin S Set in the D3R Grand Challenge 3. *J. Comput.-Aided Mol. Des.* **2019**, *33*, 105−117.
(71) Zoete, V.; Michielin, O. Comparison between Computational Alanine Scanning and Per-Residue Binding Free Energy Decomposition for Protein−Protein Association Using MM-GBSA: Application to the TCR-P-MHC Complex. *Proteins: Struct., Funct., Genet.* **2007**, *67*, 1026−1047.
(72) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein−Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-Ralgds Complexes. *J. Mol. Biol.* **2003**, *330*, 891−913.
(73) Homeyer, N.; Stoll, F.; Hillisch, A.; Gohlke, H. Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. *J. Chem. Theory Comput.* **2014**, *10*, 3331−3344.
(74) Homeyer, N.; Gohlke, H. Free Energy Calculations by the Molecular Mechanics Poisson−Boltzmann Surface Area Method. *Mol. Inf.* **2012**, *31*, 114−122.
(75) Hansen, N.; van Gunsteren, W. F. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10*, 2632−2647.
(76) Ashida, T.; Kikuchi, T. Overview of Binding Free Energy Calculation Techniques for Elucidation of Biological Processes and for Drug Discovery. *Med. Chem.* **2015**, *11*, 248−253.
(77) Gilson, M. K.; Honig, B. Calculation of the Total Electrostatic Energy of a Macromolecular System: Solvation Energies, Binding Energies, and Conformational Analysis. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 7−18.
(78) Wang, J. M.; Hou, T. J.; Xu, X. J. Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. *Curr. Comput.-Aided Drug Des.* **2006**, *2*, 287−306.
(79) Weis, A.; Katebzadeh, K.; Soderhjelm, P.; Nilsson, I.; Ryde, U. Ligand Affinities Predicted with the MM/PBSA Method: Dependence on the Simulation Method and the Force Field. *J. Med. Chem.* **2006**, *49*, 6596−6606.
(80) Zhang, J.; Zhang, H.; Wu, T.; Wang, Q.; van der Spoel, D. Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents. *J. Chem. Theory Comput.* **2017**, *13*, 1034−1043.
(81) Lee, M. S.; Olson, M. A. Calculation of Absolute Protein−Ligand Binding Affinity Using Path and Endpoint Approaches. *Biophys. J.* **2006**, *90*, 864−877.
(82) Daggett, V. Long Timescale Simulations. *Curr. Opin. Struct. Biol.* **2000**, *10*, 160−164.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

(83) Genheden, S.; Ryde, U. How To Obtain Statistically Converged MM/GBSA Results. *J. Comput. Chem.* **2009**, *31*, 837−846.

(84) Sadiq, S. K.; Wright, D. W.; Kenway, O. A.; Coveney, P. V. Accurate Ensemble Molecular Dynamics Binding Free Energy Ranking of Multidrug-Resistant HIV-1 Proteases. *J. Chem. Inf. Model.* **2010**, *50*, 890−905.

(85) Adler, M.; Beroza, P. Improved Ligand Binding Energies Derived from Molecular Dynamics: Replicate Sampling Enhances the Search of Conformational Space. *J. Chem. Inf. Model.* **2013**, *53*, 2065−2072.

(86) Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V. Computing Clinically Relevant Binding Free Energies of HIV-1 Protease Inhibitors. *J. Chem. Theory Comput.* **2014**, *10*, 1228−1241.

(87) Su, P. C.; Tsai, C. C.; Mehboob, S.; Hevener, K. E.; Johnson, M. E. Comparison of Radii Sets, Entropy, QM Methods, and Sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA Ligand Binding Energies of F. Tularensis Enoyl-ACP Reductase (Fabi). *J. Comput. Chem.* **2015**, *36*, 1859−1873.

(88) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods: I. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69−82.

(89) Virtanen, S. I.; Niinivehmas, S. P.; Pentikäinen, O. T. Case-Specific Performance of MM-PBSA, MM-GBSA, and SIE in Virtual Screening. *J. Mol. Graphics Modell.* **2015**, *62*, 303−318.

(90) Stoica, I.; Sadiq, S. K.; Coveney, P. V. Rapid and Accurate Prediction of Binding Free Energies for Saquinavir-Bound HIV-1 Proteases. *J. Am. Chem. Soc.* **2008**, *130*, 2639−2648.

(91) Terayama, K.; Iwata, H.; Araki, M.; Okuno, Y.; Tsuda, K. Machine Learning Accelerates MD-Based Binding Pose Prediction between Ligands and Proteins. *Bioinformatics* **2018**, *34*, 770−778.

(92) Wang, B.; Li, L.; Hurley, T. D.; Meroueh, S. O. Molecular Recognition in a Diverse Set of Protein−Ligand Interactions Studied with Molecular Dynamics Simulations and End-Point Free Energy Calculations. *J. Chem. Inf. Model.* **2013**, *53*, 2659−2670.

(93) Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 6. Capability to Predict Protein−Protein Binding Free Energies and Re-Rank Binding Poses Generated by Protein−Protein Docking. *Phys. Chem. Chem. Phys.* **2016**, *18*, 22129−22139.

(94) Sun, H.; Li, Y.; Shen, M.; Tian, S.; Xu, L.; Pan, P.; Guan, Y.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 5. Improved Docking Performance Using High Solute Dielectric Constant MM/GBSA and MM/PBSA Rescoring. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22035−22045.

(95) Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. Validation and Use of the MM-PBSA Approach for Drug Discovery. *J. Med. Chem.* **2005**, *48*, 4040−4048.

(96) Rastelli, G.; Del Rio, A.; Degliesposti, G.; Sgobba, M. Fast and Accurate Predictions of Binding Free Energies Using MM-PBSA and MM-GBSA. *J. Comput. Chem.* **2009**, *31*, 797−810.

(97) Xu, L.; Sun, H.; Li, Y.; Wang, J.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. The Impact of Force Fields and Ligand Charge Models. *J. Phys. Chem. B* **2013**, *117*, 8408−8421.

(98) Sun, H.; Li, Y.; Tian, S.; Xu, L.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 4. Accuracies of MM/PBSA and MM/GBSA Methodologies Evaluated by Various Simulation Protocols Using PDBbind Data Set. *Phys. Chem. Chem. Phys.* **2014**, *16*, 16719−16729.

(99) Sun, H.; Duan, L.; Chen, F.; Liu, H.; Wang, Z.; Pan, P.; Zhu, F.; Zhang, J. Z. H.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 7. Entropy Effects on the Performance of End-Point Binding Free Energy Calculation Approaches. *Phys. Chem. Chem. Phys.* **2018**, *20*, 14450−14460.

(100) Chen, F.; Sun, H.; Wang, J.; Zhu, F.; Liu, H.; Wang, Z.; Lei, T.; Li, Y.; Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 8. Predicting Binding Free Energies and Poses of Protein−RNA Complexes. *RNA* **2018**, *24*, 1183−1194.

(101) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated from Docking. *J. Comput. Chem.* **2011**, *32*, 866−877.

(102) Genheden, S.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Ryde, U. An MM/3D-RISM Approach for Ligand Binding Affinities. *J. Phys. Chem. B* **2010**, *114*, 8505−8516.

(103) Genheden, S.; Ryde, U. A Comparison of Different Initialization Protocols To Obtain Statistically Independent Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 187−195.

(104) Genheden, S.; Kuhn, O.; Mikulskis, P.; Hoffmann, D.; Ryde, U. The Normal-Mode Entropy in the MM/GBSA Method: Effect of System Truncation, Buffer Region, and Dielectric Constant. *J. Chem. Inf. Model.* **2012**, *52*, 2079−2088.

(105) Godschalk, F.; Genheden, S.; Soderhjelm, P.; Ryde, U. Comparison of MM/GBSA Calculations Based on Explicit and Implicit Solvent Simulations. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7731−7739.

(106) Pu, C.; Yan, G.; Shi, J.; Li, R. Assessing the Performance of Docking Scoring Function, FEP, MM-GBSA, and QM/MM-GBSA Approaches on a Series of Plk1 Inhibitors. *MedChemComm* **2017**, *8*, 1452−1458.

(107) Mishra, S. K.; Koca, J. Assessing the Performance of MM/PBSA, MM/GBSA, and QM−MM/GBSA Approaches on Protein/Carbohydrate Complexes: Effect of Implicit Solvent Models, QM Methods, and Entropic Contributions. *J. Phys. Chem. B* **2018**, *122*, 8113−8121.

(108) Ganesan, A.; Coote, M. L.; Barakat, K. Molecular Dynamics-Driven Drug Discovery: Leaping Forward with Confidence. *Drug Discovery Today* **2017**, *22*, 249−269.

(109) Pandey, P.; Srivastava, R.; Bandyopadhyay, P. Comparison of Molecular Mechanics-Poisson−Boltzmann Surface Area (MM-PBSA) and Molecular Mechanics-Three-Dimensional Reference Interaction Site Model (MM-3D-RISM) Method To Calculate the Binding Free Energy of Protein−Ligand Complexes: Effect of Metal Ion and Advance Statistical Test. *Chem. Phys. Lett.* **2018**, *695*, 69−78.

(110) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 383−394.

(111) Tama, F.; Sanejouand, Y. H. Conformational Change of Proteins Arising from Normal Mode Calculations. *Protein Eng., Des. Sel.* **2001**, *14*, 1−6.

(112) Brooks, B. R.; Janezic, D.; Karplus, M. Harmonic-Analysis of Large Systems. I. Methodology. *J. Comput. Chem.* **1995**, *16*, 1522−1542.

(113) Ma, J. P. Usefulness and Limitations of Normal Mode Analysis in Modeling Dynamics of Biomolecular Complexes. *Structure* **2005**, *13*, 373−380.

(114) Duan, L.; Liu, X.; Zhang, J. Z. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein−Ligand Binding Free Energy. *J. Am. Chem. Soc.* **2016**, *138*, 5722−5728.

(115) Suno, R.; Kimura, K. T.; Nakane, T.; Yamashita, K.; Wang, J.; Fujiwara, T.; Yamanaka, Y.; Im, D.; Horita, S.; et al. Crystal Structures of Human Orexin 2 Receptor Bound to the Subtype-Selective Antagonist EMPA. *Structure* **2018**, *26*, 7−19.

(116) Shao, Z.; Yin, J.; Chapman, K.; Grzemska, M.; Clark, L.; Wang, J.; Rosenbaum, D. M. High-Resolution Crystal Structure of the Human CB1 Cannabinoid Receptor. *Nature* **2016**, *540*, 602−606.

(117) Hayes, J. M.; Skamnaki, V. T.; Archontis, G.; Lamprakis, C.; Sarrou, J.; Bischler, N.; Skaltsounis, A. L.; Zographos, S. E.; Oikonomakos, N. G. Kinetics, in Silico Docking, Molecular Dynamics, and MM-GBSA Binding Studies on Prototype Indirubins, KT5720, and Staurosporine as Phosphorylase Kinase ATP-Binding Site Inhibitors: The Role of Water Molecules Examined. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 703−719.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

(118) Naim, M.; Bhat, S.; Rankin, K. N.; Dennis, S.; Chowdhury, S. F.; Siddiqi, I.; Drabik, P.; Sulea, T.; Bayly, C. I.; et al. Solvated Interaction Energy (SIE) for Scoring Protein−Ligand Binding Affinities. 1. Exploring the Parameter Space. *J. Chem. Inf. Model.* **2007**, *47*, 122−133.
(119) Sulea, T.; Cui, Q.; Purisima, E. O. Solvated Interaction Energy (SIE) for Scoring Protein−Ligand Binding Affinities. 2. Benchmark in the CSAR-2010 Scoring Exercise. *J. Chem. Inf. Model.* **2011**, *51*, 2066−2081.
(120) Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. Importance of Accurate Charges in Molecular Docking: Quantum Mechanical/Molecular Mechanical (QM/MM) Approach. *J. Comput. Chem.* **2005**, *26*, 915−931.
(121) Du, J.; Sun, H.; Xi, L.; Li, J.; Yang, Y.; Liu, H.; Yao, X. Molecular Modeling Study of Checkpoint Kinase 1 Inhibitors by Multiple Docking Strategies and Prime/MM-GBSA Calculation. *J. Comput. Chem.* **2011**, *32*, 2800−2809.
(122) Chaskar, P.; Zoete, V.; Rohrig, U. F. Toward on-the-Fly Quantum Mechanical/Molecular Mechanical (QM/MM) Docking: Development and Benchmark of a Scoring Function. *J. Chem. Inf. Model.* **2014**, *54*, 3137−3152.
(123) Fong, P.; McNamara, J. P.; Hillier, I. H.; Bryce, R. A. Assessment of QM/MM Scoring Functions for Molecular Docking to HIV-1 Protease. *J. Chem. Inf. Model.* **2009**, *49*, 913−924.
(124) Kurczab, R. The Evaluation of QM/MM-Driven Molecular Docking Combined with MM/GBSA Calculations as a Halogen-Bond Scoring Strategy. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2017**, *73*, 188−194.
(125) Davis, M. E.; Mccammon, J. A. Electrostatics in Biomolecular Structure and Dynamics. *Chem. Rev.* **1990**, *90*, 509−521.
(126) Li, C.; Li, L.; Petukh, M.; Alexov, E. Progress in Developing Poisson−Boltzmann Equation Solvers. *Mol. Based Math. Biol.* **2013**, *1*, 42−62.
(127) Gilson, M. K.; Sharp, K.; Honig, B.; Fine, R.; Hagstrom, R. Calculation of Electrostatic Energies in Proteins by a Finite-Difference Method. *Biophys. J.* **1987**, *51*, A234.
(128) Nicholls, A.; Honig, B. A Rapid Finite-Difference Algorithm, Utilizing Successive over-Relaxation to Solve the Poisson−Boltzmann Equation. *J. Comput. Chem.* **1991**, *12*, 435−445.
(129) Sharp, K. A.; Honig, B. Calculating Total Electrostatic Energies with the Nonlinear Poisson−Boltzmann Equation. *J. Phys. Chem.* **1990**, *94*, 7684−7692.
(130) Boschitsch, A. H.; Fenley, M. O. A Fast and Robust Poisson−Boltzmann Solver Based on Adaptive Cartesian Grids. *J. Chem. Theory Comput.* **2011**, *7*, 1524−1540.
(131) Wang, N.; Zhou, S.; Kekenes-Huskey, P. M.; Li, B.; McCammon, J. A. Poisson−Boltzmann Versus Size-Modified Poisson−Boltzmann Electrostatics Applied to Lipid Bilayers. *J. Phys. Chem. B* **2014**, *118*, 14827−14832.
(132) Davis, M. E.; Mccammon, J. A. Solving the Finite-Difference Linearized Poisson−Boltzmann Equation - a Comparison of Relaxation and Conjugate-Gradient Methods. *J. Comput. Chem.* **1989**, *10*, 386−391.
(133) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the Electrostatic Potential of Molecules in Solution - Method and Error Assessment. *J. Comput. Chem.* **1988**, *9*, 327−335.
(134) Xie, D. X. New Solution Decomposition and Minimization Schemes for Poisson−Boltzmann Equation in Calculation of Biomolecular Electrostatics. *J. Comput. Phys.* **2014**, *275*, 294−309.
(135) Fisicaro, G.; Genovese, L.; Andreussi, O.; Marzari, N.; Goedecker, S. A Generalized Poisson and Poisson−Boltzmann Solver for Electrostatic Environments. *J. Chem. Phys.* **2016**, *144*, 014103.
(136) Xie, D. X.; Jiang, Y. A Nonlocal Modified Poisson−Boltzmann Equation and Finite Element Solver for Computing Electrostatics of Biomolecules. *J. Comput. Phys.* **2016**, *322*, 1−20.
(137) Wang, J.; Tan, C.; Tan, Y.-H.; Lu, Q.; Luo, R. Poisson−Boltzmann Solvents in Molecular Dynamics Simulations. *Commun. Comput. Phys.* **2008**, *3*, 1010−1031.
(138) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(139) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 198−210.
(140) Wang, J.; Luo, R. Assessment of Linear Finite-Difference Poisson−Boltzmann Solvers. *J. Comput. Chem.* **2010**, *31*, 1689−1698.
(141) Cai, Q.; Hsieh, M. J.; Wang, J.; Luo, R. Performance of Nonlinear Finite-Difference Poisson−Boltzmann Solvers. *J. Chem. Theory Comput.* **2010**, *6*, 203−211.
(142) Baker, N.; Holst, M.; Wang, F. Adaptive Multilevel Finite Element Solution of the Poisson−Boltzmann Equation II. Refinement at Solvent-Accessible Surfaces in Biomolecular Systems. *J. Comput. Chem.* **2000**, *21*, 1343−1352.
(143) Bowen, R. W.; Sharif, A. O. Adaptive Finite-Element Solution of the Nonlinear Poisson−Boltzmann Equation: A Charged Spherical Particle at Various Distances from a Charged Cylindrical Pore in a Charged Planar Surface. *J. Colloid Interface Sci.* **1997**, *187*, 363−374.
(144) Holst, M.; Baker, N.; Wang, F. Adaptive Multilevel Finite Element Solution of the Poisson−Boltzmann Equation I. Algorithms and Examples. *J. Comput. Chem.* **2000**, *21*, 1319−1342.
(145) Holst, M.; Baker, N.; Wang, F. Adaptive Multilevel Finite Element Solution of the Poisson−Boltzmann Equation I. Algorithms and Examples (Vol 21, Pg 1319, 2000). *J. Comput. Chem.* **2001**, *22*, 475−475.
(146) Bordner, A. J.; Huber, G. A. Boundary Element Solution of the Linear Poisson−Boltzmann Equation and a Multipole Method for the Rapid Calculation of Forces on Macromolecules in Solution. *J. Comput. Chem.* **2003**, *24*, 353−367.
(147) Liang, J.; Subramaniam, S. Computation of Molecular Electrostatics with Boundary Element Methods. *Biophys. J.* **1997**, *73*, 1830−1841.
(148) Lu, B. Z.; McCammon, J. A. Improved Boundary Element Methods for Poisson−Boltzmann Electrostatic Potential and Force Calculations. *J. Chem. Theory Comput.* **2007**, *3*, 1134−1142.
(149) Chen, D.; Chen, Z.; Chen, C.; Geng, W.; Wei, G. W. MIBPB: A Software Package for Electrostatic Analysis. *J. Comput. Chem.* **2011**, *32*, 756−770.
(150) Li, L.; Li, C.; Sarkar, S.; Zhang, J.; Witham, S.; Zhang, Z.; Wang, L.; Smith, N.; Petukh, M.; Alexov, E. DelPhi: A Comprehensive Suite for DelPhi Software and Associated Resources. *BMC Biophys.* **2012**, *5*, 9.
(151) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; et al. Electrostatics and Diffusion of Molecules in Solution - Simulations with the University-of-Houston Brownian Dynamics Program. *Comput. Phys. Commun.* **1995**, *91*, 57−95.
(152) Prabhu, N. V.; Panda, M.; Yang, Q. Y.; Sharp, K. A. Explicit Ion, Implicit Water Solvation for Molecular Dynamics of Nucleic Acids and Highly Charged Molecules. *J. Comput. Chem.* **2008**, *29*, 1113−1130.
(153) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(154) Prabhu, N. V.; Zhu, P. J.; Sharp, K. A. Implementation and Testing of Stable, Fast Implicit Solvation in Molecular Dynamics Using the Smooth-Permittivity Finite Difference Poisson−Boltzmann Method. *J. Comput. Chem.* **2004**, *25*, 2049−2064.
(155) Zhou, Y. C.; Feig, M.; Wei, G. W. Highly Accurate Biomolecular Electrostatics in Continuum Dielectric Environments. *J. Comput. Chem.* **2008**, *29*, 87−97.
(156) Nguyen, D. D.; Wang, B.; Wei, G. W. Accurate, Robust, and Reliable Calculations of Poisson−Boltzmann Binding Energies. *J. Comput. Chem.* **2017**, *38*, 941−948.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

(157) Colmenares, J.; Ortiz, J.; Rocchia, W. GPU Linear and Non-Linear Poisson−Boltzmann Solver Module for DelPhi. *Bioinformatics* **2014**, *30*, 569−570.
(158) Qi, R.; Botello-Smith, W. M.; Luo, R. Acceleration of Linear Finite-Difference Poisson−Boltzmann Methods on Graphics Processing Units. *J. Chem. Theory Comput.* **2017**, *13*, 3378−3387.
(159) Li, C.; Li, L.; Zhang, J.; Alexov, E. Highly Efficient and Exact Method for Parallelization of Grid-Based Algorithms and Its Implementation in DelPhi. *J. Comput. Chem.* **2012**, *33*, 1960−1966.
(160) Luo, R.; David, L.; Gilson, M. K. Accelerated Poisson−Boltzmann Calculations for Static and Dynamic Systems. *J. Comput. Chem.* **2002**, *23*, 1244−1253.
(161) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037−10041.
(162) Colmenares, J.; Ortiz, J.; Decherchi, S.; Fijany, A.; Rocchia, W. Solving the Linearized Poisson−Boltzmann Equation on GPUs Using CUDA. In *21st Euromicro International Conference on Parallel, Distributed, and Network-Based Processing*; IEEE, 2013; pp 420−426.
(163) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. Novel Generalized Born Methods. *J. Chem. Phys.* **2002**, *116*, 10606−10614.
(164) Im, W.; Lee, M. S.; Brooks, C. L., III. Generalized Born Model with a Simple Smoothing Function. *J. Comput. Chem.* **2003**, *24*, 1691−1702.
(165) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(166) Sigalov, G.; Scheffel, P.; Onufriev, A. Incorporating Variable Dielectric Environments into the Generalized Born Model. *J. Chem. Phys.* **2005**, *122*, 094511.
(167) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. Application of a Pairwise Generalized Born Model to Proteins and Nucleic Acids: Inclusion of Salt Effects. *Theor. Chem. Acc.* **1999**, *101*, 426−434.
(168) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996**, *100*, 19824−19839.
(169) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Pairwise Solute Descreening of Solute Charges from a Dielectric Medium. *Chem. Phys. Lett.* **1995**, *246*, 122−129.
(170) Tsui, V.; Case, D. A. Theory and Applications of the Generalized Born Solvation Model in Macromolecular Simulations. *Biopolymers* **2000**, *56*, 275−291.
(171) Roe, D. R.; Okur, A.; Wickstrom, L.; Hornak, V.; Simmerling, C. Secondary Structure Bias in Generalized Born Solvent Models: Comparison of Conformational Ensembles and Free Energy of Solvent Polarization from Explicit and Implicit Solvation. *J. Phys. Chem. B* **2007**, *111*, 1846−1857.
(172) Lee, M. S.; Feig, M.; Salsbury, F. R., Jr.; Brooks, C. L., III. New Analytic Approximation to the Standard Molecular Volume Definition and Its Application to Generalized Born Calculations. *J. Comput. Chem.* **2003**, *24*, 1348−1356.
(173) Aguilar, B.; Onufriev, A. V. Efficient Computation of the Total Solvation Energy of Small Molecules via the R6 Generalized Born Model. *J. Chem. Theory Comput.* **2012**, *8*, 2404−2411.
(174) Aguilar, B.; Shadrach, R.; Onufriev, A. V. Reducing the Secondary Structure Bias in the Generalized Born Model Via R6 Effective Radii. *J. Chem. Theory Comput.* **2010**, *6*, 3613−3630.
(175) Forouzesh, N.; Izadi, S.; Onufriev, A. V. Grid-Based Surface Generalized Born Model for Calculation of Electrostatic Binding Free Energies. *J. Chem. Inf. Model.* **2017**, *57*, 2505−2513.
(176) Izadi, S.; Harris, R. C.; Fenley, M. O.; Onufriev, A. V. Accuracy Comparison of Generalized Born Models in the Calculation of Electrostatic Binding Free Energies. *J. Chem. Theory Comput.* **2018**, *14*, 1656−1670.
(177) Schutz, C. N.; Warshel, A. What Are the Dielectric "Constants" of Proteins and How To Validate Electrostatic Models? *Abstr. Pap. - Am. Chem. Soc.* **2002**, *223*, C75.
(178) Massova, I.; Kollman, P. A. Combined Molecular Mechanical and Continuum Solvent Approach (MM-PBSA/GBSA) to Predict Ligand Binding. *Perspect. Drug Discovery Des.* **2000**, *18*, 113−135.
(179) Venken, T.; Krnavek, D.; Munch, J.; Kirchhoff, F.; Henklein, P.; De Maeyer, M.; Voet, A. An Optimized MM/PBSA Virtual Screening Approach Applied to an HIV-1 Gp41 Fusion Peptide Inhibitor. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 3221−3235.
(180) Ravindranathan, K.; Tirado-Rives, J.; Jorgensen, W. L.; Guimarães, C. R. W. Improving MM-GB/SA Scoring through the Application of the Variable Dielectric Model. *J. Chem. Theory Comput.* **2011**, *7*, 3859−3865.
(181) Mikulskis, P.; Genheden, S.; Rydberg, P.; Sandberg, L.; Olsen, L.; Ryde, U. Binding Affinities in the Sampl3 Trypsin and Host-Guest Blind Tests Estimated with the MM/PBSA and LIE Methods. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 527−541.
(182) Guimarães, C. R. W. A Direct Comparison of the MM-GB/SA Scoring Procedure and Free-Energy Perturbation Calculations Using Carbonic Anhydrase as a Test Case: Strengths and Pitfalls of Each Approach. *J. Chem. Theory Comput.* **2011**, *7*, 2296−2306.
(183) Guimarães, C. R. W.; Mathiowetz, A. M. Addressing Limitations with the MM-GB/SA Scoring Procedure Using the Watermap Method and Free Energy Perturbation Calculations. *J. Chem. Inf. Model.* **2010**, *50*, 547−559.
(184) Yang, T.; Wu, J. C.; Yan, C.; Wang, Y.; Luo, R.; Gonzales, M. B.; Dalby, K. N.; Ren, P. Virtual Screening Using Molecular Simulations. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 1940−1951.
(185) Foloppe, N.; Hubbard, R. Towards Predictive Ligand Design with Free-Energy Based Computational Methods? *Curr. Med. Chem.* **2006**, *13*, 3583−3608.
(186) Singh, N.; Warshel, A. Absolute Binding Free Energy Calculations: On the Accuracy of Computational Scoring of Protein−Ligand Interactions. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1705−1723.
(187) Spiliotopoulos, D.; Kastritis, P. L.; Melquiond, A. S.; Bonvin, A. M.; Musco, G.; Rocchia, W.; Spitaleri, A. dMM-PBSA: A New HADDOCK Scoring Function for Protein−Peptide Docking. *Front. Mol. Biosci.* **2016**, *3*, 46.
(188) Soderhjelm, P.; Kongsted, J.; Ryde, U. Ligand Affinities Estimated by Quantum Chemical Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 1726−1737.
(189) Mulakala, C.; Viswanadhan, V. N. Could MM-GBSA Be Accurate Enough for Calculation of Absolute Protein/Ligand Binding Free Energies? *J. Mol. Graphics Modell.* **2013**, *46*, 41−51.
(190) Zhu, K.; Pincus, D. L.; Zhao, S. W.; Friesner, R. A. Long Loop Prediction Using the Protein Local Optimization Program. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 438−452.
(191) Zhu, K.; Shirts, M. R.; Friesner, R. A. Improved Methods for Side Chain and Loop Predictions Via the Protein Local Optimization Program: Variable Dielectric Model for Implicitly Improving the Treatment of Polarization Effects. *J. Chem. Theory Comput.* **2007**, *3*, 2108−2119.
(192) Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A. The VSGB 2.0 Model: A Next Generation Energy Model for High Resolution Protein Structure Modeling. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 2794−2812.
(193) Liu, X.; Peng, L.; Zhang, J. Z. H. Accurate and Efficient Calculation of Protein−Protein Binding Free Energy-Interaction Entropy with Residue Type-Specific Dielectric Constants. *J. Chem. Inf. Model.* **2019**, *59*, 272−281.
(194) Miller, B. R., 3rd; McGee, T. D., Jr.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314−3321.
(195) Oehme, D. P.; Brownlee, R. T.; Wilson, D. J. Effect of Atomic Charge, Solvation, Entropy, and Ligand Protonation State on MM-PB(GB)SA Binding Energies of HIV Protease. *J. Comput. Chem.* **2012**, *33*, 2566−2580.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

(196) Wang, J.; Hou, T. Develop and Test a Solvent Accessible Surface Area-Based Model in Conformational Entropy Calculations. *J. Chem. Inf. Model.* **2012**, *52*, 1199−1212.
(197) Genheden, S. MM/GBSA and LIE Estimates of Host−Guest Affinities: Dependence on Charges and Solvation Model. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 1085−1093.
(198) Karplus, M.; Kushick, J. N. Method for Estimating the Configurational Entropy of Macromolecules. *Macromolecules* **1981**, *14*, 325−332.
(199) Brooks, B.; Karplus, M. Harmonic Dynamics of Proteins: Normal Modes and Fluctuations in Bovine Pancreatic Trypsin Inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80*, 6571−6575.
(200) Schlitter, J. Estimation of Absolute and Relative Entropies of Macromolecules Using the Covariance Matrix. *Chem. Phys. Lett.* **1993**, *215*, 617−621.
(201) Ben-Shalom, I. Y.; Pfeiffer-Marek, S.; Baringhaus, K.-H.; Gohlke, H. Efficient Approximation of Ligand Rotational and Translational Entropy Changes Upon Binding for Use in MM-PBSA Calculations. *J. Chem. Inf. Model.* **2017**, *57*, 170−189.
(202) Genheden, S.; Akke, M.; Ryde, U. Conformational Entropies and Order Parameters: Convergence, Reproducibility, and Transferability. *J. Chem. Theory Comput.* **2014**, *10*, 432−438.
(203) Hikiri, S.; Yoshidome, T.; Ikeguchi, M. Computational Methods for Configurational Entropy Using Internal and Cartesian Coordinates. *J. Chem. Theory Comput.* **2016**, *12*, 5990−6000.
(204) Sharp, K. Calculation of Molecular Entropies Using Temperature Integration. *J. Chem. Theory Comput.* **2013**, *9*, 1164−1172.
(205) Choi, H.; Kang, H.; Park, H. Computational Prediction of Molecular Hydration Entropy with Hybrid Scaled Particle Theory and Free-Energy Perturbation Method. *J. Chem. Theory Comput.* **2015**, *11*, 4933−4942.
(206) Gyimesi, G.; Závodszky, P.; Szilágyi, A. Calculation of Configurational Entropy Differences from Conformational Ensembles Using Gaussian Mixtures. *J. Chem. Theory Comput.* **2017**, *13*, 29−41.
(207) Huang, D. D.; Qi, Y. F.; Song, J. N.; Zhang, J. Z. H. Calculation of Hot Spots for Protein−Protein Interaction in p53/PMI-MDM2/MDMX Complexes. *J. Comput. Chem.* **2019**, *40*, 1045−1056.
(208) Qiu, L.; Yan, Y.; Sun, Z.; Song, J.; Zhang, J. Z. Interaction Entropy for Computational Alanine Scanning in Protein−Protein Binding. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1342.
(209) Yan, Y.; Yang, M.; Ji, C. G.; Zhang, J. Z. Interaction Entropy for Computational Alanine Scanning. *J. Chem. Inf. Model.* **2017**, *57*, 1112−1122.
(210) Sun, Z.; Yan, Y. N.; Yang, M.; Zhang, J. Z. Interaction Entropy for Protein−Protein Binding. *J. Chem. Phys.* **2017**, *146*, 124124.
(211) Aldeghi, M.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Statistical Analysis on the Performance of Molecular Mechanics Poisson−Boltzmann Surface Area Versus Absolute Binding Free Energy Calculations: Bromodomains as a Case Study. *J. Chem. Inf. Model.* **2017**, *57*, 2203−2221.
(212) Li, Y.; Cong, Y.; Feng, G.; Zhong, S.; Zhang, J. Z.; Sun, H.; Duan, L. The Impact of Interior Dielectric Constant and Entropic Change on HIV-1 Complex Binding Free Energy Prediction. *Struct. Dyn.* **2018**, *5*, 064101.
(213) Duan, L.; Feng, G.; Wang, X.; Wang, L.; Zhang, Q. Effect of Electrostatic Polarization and Bridging Water on CDK2−Ligand Binding Affinities Calculated Using a Highly Efficient Interaction Entropy Method. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10140−10152.
(214) Cong, Y. L.; Li, Y. C.; Jin, K.; Zhong, S. S.; Zhang, J. Z. H.; Li, H.; Duan, L. L. Exploring the Reasons for Decrease in Binding Affinity of HIV-2 against HIV-1 Protease Complex Using Interaction Entropy under Polarized Force Field. *Front. Chem.* **2018**, *6*, 380.
(215) Liu, X.; Peng, L.; Zhou, Y.; Zhang, Y.; Zhang, J. Z. Computational Alanine Scanning with Interaction Entropy for Protein−Ligand Binding Free Energies. *J. Chem. Theory Comput.* **2018**, *14*, 1772−1780.
(216) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. Small Molecule Hydration Free Energies in Explicit Solvent: An Extensive Test of Fixed-Charge Atomistic Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 350−358.
(217) Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. Accessible Surface-Areas as a Measure of the Thermodynamic Parameters of Hydration of Peptides. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 3086−3090.
(218) Eisenberg, D.; McLachlan, A. D. Solvation Energy in Protein Folding and Binding. *Nature* **1986**, *319*, 199−203.
(219) Tan, C.; Tan, Y. H.; Luo, R. Implicit Nonpolar Solvent Models. *J. Phys. Chem. B* **2007**, *111*, 12263−12274.
(220) Barone, V.; Cossi, M.; Tomasi, J. A New Definition of Cavities for the Computation of Solvation Free Energies by the Polarizable Continuum Model. *J. Chem. Phys.* **1997**, *107*, 3210−3221.
(221) Genheden, S.; Kongsted, J.; Soderhjelm, P.; Ryde, U. Nonpolar Solvation Free Energies of Protein−Ligand Complexes. *J. Chem. Theory Comput.* **2010**, *6*, 3558−3568.
(222) Genheden, S.; Mikulskis, P.; Hu, L. H.; Kongsted, J.; Soderhjelm, P.; Ryde, U. Accurate Predictions of Nonpolar Solvation Free Energies Require Explicit Consideration of Binding-Site Hydration. *J. Am. Chem. Soc.* **2011**, *133*, 13081−13092.
(223) Mikulskis, P.; Genheden, S.; Ryde, U. Effect of Explicit Water Molecules on Ligand-Binding Affinities Calculated with the MM/GBSA Approach. *J. Mol. Model.* **2014**, *20*, 2273.
(224) Wallnoefer, H. G.; Liedl, K. R.; Fox, T. A Challenging System: Free Energy Prediction for Factor Xa. *J. Comput. Chem.* **2011**, *32*, 1743−1752.
(225) Lepsik, M.; Kriz, Z.; Havlas, Z. Efficiency of a Second-Generation HIV-1 Protease Inhibitor Studied by Molecular Dynamics and Absolute Binding Free Energy Calculations. *Proteins: Struct., Funct., Genet.* **2004**, *57*, 279−293.
(226) Wong, S.; Amaro, R. E.; McCammon, J. A. MM-PBSA Captures Key Role of Intercalating Water Molecules at a Protein−Protein Interface. *J. Chem. Theory Comput.* **2009**, *5*, 422−429.
(227) Greenidge, P. A.; Kramer, C.; Mozziconacci, J. C.; Wolf, R. M. MM/GBSA Binding Energy Prediction on the PDBbind Data Set: Successes, Failures, and Directions for Further Improvement. *J. Chem. Inf. Model.* **2013**, *53*, 201−209.
(228) Kohlmann, A.; Zhu, X. T.; Dalgarno, D. Application of MM-GB/SA and Watermap to SRC Kinase Inhibitor Potency Prediction. *ACS Med. Chem. Lett.* **2012**, *3*, 94−99.
(229) Yang, Y.; Lightstone, F. C.; Wong, S. E. Approaches to Efficiently Estimate Solvation and Explicit Water Energetics in Ligand Binding: The Use of Watermap. *Expert Opin. Drug Discovery* **2013**, *8*, 277−287.
(230) Cappel, D.; Sherman, W.; Beuming, T. Calculating Water Thermodynamics in the Binding Site of Proteins—Applications of Watermap to Drug Discovery. *Curr. Top. Med. Chem.* **2017**, *17*, 2586−2598.
(231) Homeyer, N.; Gohlke, H. Extension of the Free Energy Workflow FEW Towards Implicit Solvent/Implicit Membrane MM-PBSA Calculations. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 972−982.
(232) Homeyer, N.; Gohlke, H. FEW: A Workflow Tool for Free Energy Calculations of Ligand Binding. *J. Comput. Chem.* **2013**, *34*, 965−973.
(233) Kumari, R.; Kumar, R.; Lynn, A. *g_mmpbsa*—A GROMACS Tool for High-Throughput MM-PBSA Calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951−1962.
(234) Paissoni, C.; Spiliotopoulos, D.; Musco, G.; Spitaleri, A. GMXPBSA 2.0: A GROMACS Tool to Perform MM/PBSA and Computational Alanine Scanning. *Comput. Phys. Commun.* **2014**, *185*, 2920−2929.
(235) Konecny, R.; Baker, N. A.; McCammon, J. A. iAPBS: A Programming Interface to Adaptive Poisson−Boltzmann Solver. *Comput. Sci. Discovery* **2012**, *5*, 015005.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

(236) Loeffler, H. H.; Michel, J.; Woods, C. Fesetup: Automating Setup for Alchemical Free Energy Simulations. *J. Chem. Inf. Model.* **2015**, *55*, 2485−2490.
(237) Liu, H.; Hou, T. CaFE: A Tool for Binding Affinity Prediction Using End-Point Free Energy Methods. *Bioinformatics* **2016**, *32*, 2216−2218.
(238) Hao, G. F.; Jiang, W.; Ye, Y. N.; Wu, F. X.; Zhu, X. L.; Guo, F. B.; Yang, G. F. ACFIS: A Web Server for Fragment-Based Drug Discovery. *Nucleic Acids Res.* **2016**, *44*, W550−556.
(239) Peng, Y.; Sun, L.; Jia, Z.; Li, L.; Alexov, E. Predicting Protein−DNA Binding Free Energy Change Upon Missense Mutations Using Modified MM/PBSA Approach: SAMPDI Webserver. *Bioinformatics* **2018**, *34*, 779−786.
(240) Wang, Z.; Wang, X.; Li, Y.; Lei, T.; Wang, E.; Li, D.; Kang, Y.; Zhu, F.; Hou, T. farPPI: A Webserver for Accurate Prediction of Protein−Ligand Binding Structures for Small-Molecule PPI Inhibitors by MM/PB(GB)SA Methods. *Bioinformatics* **2019**, *35*, 1777−1779.
(241) Liu, X.; Liu, J.; Zhu, T.; Zhang, L.; He, X.; Zhang, J. Z. PBSA_E: A PBSA-Based Free Energy Estimator for Protein−Ligand Binding Affinity. *J. Chem. Inf. Model.* **2016**, *56*, 854−861.
(242) Chowdhury, R.; Rasheed, M.; Keidel, D.; Moussalem, M.; Olson, A.; Sanner, M.; Bajaj, C. Protein−Protein Docking with F^2Dock 2.0 and GB-Rerank. *PLoS One* **2013**, *8*, e51307.
(243) Feng, T.; Chen, F.; Kang, Y.; Sun, H.; Liu, H.; Li, D.; Zhu, F.; Hou, T. HawkRank: A New Scoring Function for Protein−Protein Docking Based on Weighted Energy Terms. *J. Cheminf.* **2017**, *9*, 66.
(244) Pierce, B.; Weng, Z. ZRANK: Reranking Protein Docking Predictions with an Optimized Energy Function. *Proteins: Struct., Funct., Genet.* **2007**, *67*, 1078−1086.
(245) Andrusier, N.; Nussinov, R.; Wolfson, H. J. FireDock: Fast Interaction Refinement in Molecular Docking. *Proteins: Struct., Funct., Genet.* **2007**, *69*, 139−159.
(246) Yang, Y.; Zhou, Y. Specific Interactions for Ab Initio Folding of Protein Terminal Regions with Secondary Structures. *Proteins: Struct., Funct., Genet.* **2008**, *72*, 793−803.
(247) Plewczynski, D.; Łaźniewski, M.; Augustyniak, R.; Ginalski, K. Can We Trust Docking Results? Evaluation of Seven Commonly Used Programs on PDBbind Database. *J. Comput. Chem.* **2011**, *32*, 742−755.
(248) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. Biomolecular Simulations: Recent Developments in Force Fields, Simulations of Enzyme Catalysis, Protein−Ligand, Protein−Protein, and Protein-Nucleic Acid Noncovalent Interactions. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211−243.
(249) Ahmed, S. M.; Kruger, H. G.; Govender, T.; Maguire, G. E. M.; Sayed, Y.; Ibrahim, M. A. A.; Naicker, P.; Soliman, M. E. S. Comparison of the Molecular Dynamics and Calculated Binding Free Energies for Nine FDA-Approved HIV-1 PR Drugs against Subtype B and C-SA HIV PR. *Chem. Biol. Drug Des.* **2013**, *81*, 208−218.
(250) Åqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. Ligand Binding Affinities from MD Simulations. *Acc. Chem. Res.* **2002**, *35*, 358−365.
(251) Tian, S.; Sun, H.; Li, Y.; Pan, P.; Li, D.; Hou, T. Development and Evaluation of an Integrated Virtual Screening Strategy by Combining Molecular Docking and Pharmacophore Searching Based on Multiple Protein Structures. *J. Chem. Inf. Model.* **2013**, *53*, 2743−2756.
(252) Tian, S.; Sun, H.; Pan, P.; Li, D.; Zhen, X.; Li, Y.; Hou, T. Assessing an Ensemble Docking-Based Virtual Screening Strategy for Kinase Targets by Considering Protein Flexibility. *J. Chem. Inf. Model.* **2014**, *54*, 2664−2679.
(253) Holliday, J. D.; Kanoulas, E.; Malim, N.; Willett, P. Multiple Search Methods for Similarity-Based Virtual Screening: Analysis of Search Overlap and Precision. *J. Cheminf.* **2011**, *3*, 29.
(254) Honarparvar, B.; Govender, T.; Maguire, G. E. M.; Soliman, M. E. S.; Kruger, H. G. Integrated Approach to Structure-Based Enzymatic Drug Design: Molecular Modeling, Spectroscopy, and Experimental Bioactivity. *Chem. Rev.* **2014**, *114*, 493−537.
(255) Lindstrom, A.; Edvinsson, L.; Johansson, A.; Andersson, C. D.; Andersson, I. E.; Raubacher, F.; Linusson, A. Postprocessing of Docked Protein−Ligand Complexes Using Implicit Solvation Models. *J. Chem. Inf. Model.* **2011**, *51*, 267−282.
(256) Steinbrecher, T.; Case, D. A.; Labahn, A. A Multistep Approach to Structure-Based Drug Design: Studying Ligand Binding at the Human Neutrophil Elastase. *J. Med. Chem.* **2006**, *49*, 1837−1844.
(257) Thompson, D. C.; Humblet, C.; Joseph-McCarthy, D. Investigation of MM-PBSA Rescoring of Docking Poses. *J. Chem. Inf. Model.* **2008**, *48*, 1081−1091.
(258) Zhang, X.; Wong, S. E.; Lightstone, F. C. Towards Fully Automated High Performance Computing Drug Discovery: A Massively Parallel Virtual Screening Pipeline for Docking and MM/GBSA Rescoring to Improve Enrichment. *J. Chem. Inf. Model.* **2014**, *54*, 324−337.
(259) Greenidge, P. A.; Kramer, C.; Mozziconacci, J.-C.; Sherman, W. Improving Docking Results Via Reranking of Ensembles of Ligand Poses in Multiple X-Ray Protein Conformations with MM-GBSA. *J. Chem. Inf. Model.* **2014**, *54*, 2697−2717.
(260) Zhu, T.; Lee, H.; Lei, H.; Jones, C.; Patel, K.; Johnson, M. E.; Hevener, K. E. Fragment-Based Drug Discovery Using a Multi-domain, Parallel MD-MM/PBSA Screening Protocol. *J. Chem. Inf. Model.* **2013**, *53*, 560−572.
(261) Karami, M.; Jalali, C.; Mirzaie, S. Combined Virtual Screening, MMPBSA, Molecular Docking and Dynamics Studies against Deadly Anthrax: An in Silico Effort to Inhibit Bacillus Anthracis Nucleoside Hydrolase. *J. Theor. Biol.* **2017**, *420*, 180−189.
(262) Liu, Z.; Su, M.; Han, L.; Liu, J.; Yang, Q.; Li, Y.; Wang, R. Forging the Basis for Developing Protein−Ligand Interaction Scoring Functions. *Acc. Chem. Res.* **2017**, *50*, 302−309.
(263) Okimoto, N.; Futatsugi, N.; Fuji, H.; Suenaga, A.; Morimoto, G.; Yanai, R.; Ohno, Y.; Narumi, T.; Taiji, M. High-Performance Drug Discovery: Computational Screening by Combining Docking and Molecular Dynamics Simulations. *PLoS Comput. Biol.* **2009**, *5*, e1000528.
(264) Diao, Y.; Lu, W.; Jin, H.; Zhu, J.; Han, L.; Xu, M.; Gao, R.; Shen, X.; Zhao, Z.; Liu, X.; et al. Discovery of Diverse Human Dihydroorotate Dehydrogenase Inhibitors as Immunosuppressive Agents by Structure-Based Virtual Screening. *J. Med. Chem.* **2012**, *55*, 8341−8349.
(265) Amato, A.; Lucas, X.; Bortoluzzi, A.; Wright, D.; Ciulli, A. Targeting Ligandable Pockets on Plant Homeodomain (PHD) Zinc Finger Domains by a Fragment-Based Approach. *ACS Chem. Biol.* **2018**, *13*, 915−921.
(266) Ferreira de Freitas, R.; Harding, R. J.; Franzoni, I.; Ravichandran, M.; Mann, M. K.; Ouyang, H.; Lautens, M.; Santhakumar, V.; Arrowsmith, C. H.; Schapira, M. Identification and Structure−Activity Relationship of HDAC6 Zinc-Finger Ubiquitin Binding Domain Inhibitors. *J. Med. Chem.* **2018**, *61*, 4517−4527.
(267) Li, N.; Ainsworth, R. I.; Ding, B.; Hou, T.; Wang, W. Using Hierarchical Virtual Screening to Combat Drug Resistance of the HIV-1 Protease. *J. Chem. Inf. Model.* **2015**, *55*, 1400−1412.
(268) Xu, M.; Zhu, J.; Diao, Y.; Zhou, H.; Ren, X.; Sun, D.; Huang, J.; Han, D.; Zhao, Z.; Zhu, L.; et al. Novel Selective and Potent Inhibitors of Malaria Parasite Dihydroorotate Dehydrogenase: Discovery and Optimization of Dihydrothiophenone Derivatives. *J. Med. Chem.* **2013**, *56*, 7911−7924.
(269) Xu, L.; Zhang, Y.; Zheng, L. T.; Qiao, C. H.; Li, Y. Y.; Li, D.; Zhen, X. C.; Hou, T. J. Discovery of Novel Inhibitors Targeting the Macrophage Migration Inhibitory Factor Via Structure-Based Virtual Screening and Bioassays. *J. Med. Chem.* **2014**, *57*, 3737−3745.
(270) Pan, P.; Yu, H.; Liu, Q.; Kong, X.; Chen, H.; Chen, J.; Liu, Q.; Li, D.; Kang, Y.; Sun, H.; et al. Combating Drug-Resistant Mutants of Anaplastic Lymphoma Kinase with Potent and Selective Type-I1/2 Inhibitors by Stabilizing Unique DFG-Shifted Loop Conformation. *ACS Cent. Sci.* **2017**, *3*, 1208−1220.
(271) Pearlman, D. A. Evaluating the Molecular Mechanics Poisson−Boltzmann Surface Area Free Energy Method Using a Congeneric Series of Ligands to P38 Map Kinase. *J. Med. Chem.* **2005**, *48*, 7796−7807.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

(272) Schmitt, H. P.; Oberwittler, C. Computer-Aided Classification of Malignancy in Astrocytomas. II. The Value of Categorically Evaluated Histologic and Non-Histologic Features for a Numerical Classifier. *Anal. Cell. Pathol.* **1992**, *4*, 409−419.
(273) Genheden, S.; Nilsson, I.; Ryde, U. Binding Affinities of Factor Xa Inhibitors Estimated by Thermodynamic Integration and MM/GBSA. *J. Chem. Inf. Model.* **2011**, *51*, 947−958.
(274) Amaro, R. E.; Cheng, X.; Ivanov, I.; Xu, D.; McCammon, J. A. Characterizing Loop Dynamics and Ligand Recognition in Human- and Avian-Type Influenza Neuraminidases Via Generalized Born Molecular Dynamics and End-Point Free Energy Calculations. *J. Am. Chem. Soc.* **2009**, *131*, 4702−4709.
(275) Guimarães, C. R. W.; Cardozo, M. MM-GB/SA Rescoring of Docking Poses in Structure-Based Lead Optimization. *J. Chem. Inf. Model.* **2008**, *48*, 958−970.
(276) Knight, J. L.; Krilov, G.; Borrelli, K. W.; Williams, J.; Gunn, J. R.; Clowes, A.; Cheng, L.; Friesner, R. A.; Abel, R. Leveraging Data Fusion Strategies in Multireceptor Lead Optimization MM/GBSA End-Point Methods. *J. Chem. Theory Comput.* **2014**, *10*, 3207−3220.
(277) Taddei, M.; Ferrini, S.; Giannotti, L.; Corsi, M.; Manetti, F.; Giannini, G.; Vesci, L.; Milazzo, F. M.; Alloatti, D.; Guglielmi, M. B.; et al. Synthesis and Evaluation of New Hsp90 Inhibitors Based on a 1,4,5-Trisubstituted 1,2,3-Triazole Scaffold. *J. Med. Chem.* **2014**, *57*, 2258−2274.
(278) Hou, T. J.; Li, N.; Li, Y. Y.; Wang, W. Characterization of Domain−Peptide Interaction Interface: Prediction of SH3 Domain-Mediated Protein−Protein Interaction Network in Yeast by Generic Structure-Based Models. *J. Proteome Res.* **2012**, *11*, 2982−2995.
(279) Sun, H.; Pan, P.; Tian, S.; Xu, L.; Kong, X.; Li, Y.; Li, D.; Hou, T. Constructing and Validating High-Performance MIEC-SVM Models in Virtual Screening for Kinases: A Better Way for Actives Discovery. *Sci. Rep.* **2016**, *6*, 24817.
(280) Chen, F.; Sun, H.; Liu, H.; Li, D.; Li, Y.; Hou, T. Prediction of Luciferase Inhibitors by the High-Performance MIEC-Gbdt Approach Based on Interaction Energetic Patterns. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10163−10176.
(281) Ding, B.; Li, N.; Wang, W. Characterizing Binding of Small Molecules. II. Evaluating the Potency of Small Molecules to Combat Resistance Based on Docking Structures. *J. Chem. Inf. Model.* **2013**, *53*, 1213−1222.
(282) Hou, T.; Zhang, W.; Wang, J.; Wang, W. Predicting Drug Resistance of the HIV-1 Protease Using Molecular Interaction Energy Components. *Proteins: Struct., Funct., Genet.* **2009**, *74*, 837−846.
(283) Xu, W.; Lau, Y. H.; Fischer, G.; Tan, Y. S.; Chattopadhyay, A.; de la Roche, M.; Hyvönen, M.; Verma, C.; Spring, D. R.; Itzhaki, L. S. Macrocyclized Extended Peptides: Inhibiting the Substrate-Recognition Domain of Tankyrase. *J. Am. Chem. Soc.* **2017**, *139*, 2245−2256.
(284) Roca, C.; Martinez-Gonzalez, L.; Daniel-Mozo, M.; Sastre, J.; Infantes, L.; Mansilla, A.; Chaves-Sanjuan, A.; González-Rubio, J. M.; Gil, C.; Cañada, F. J.; et al. Deciphering the Inhibition of the Neuronal Calcium Sensor 1 and the Guanine Exchange Factor Ric8a with a Small Phenothiazine Molecule for the Rational Generation of Therapeutic Synapse Function Regulators. *J. Med. Chem.* **2018**, *61*, 5910−5921.
(285) Jiang, Z. Y.; Lu, M. C.; Xu, L. L.; Yang, T. T.; Xi, M. Y.; Xu, X. L.; Guo, X. K.; Zhang, X. J.; You, Q. D.; Sun, H. P. Discovery of Potent Keap1−Nrf2 Protein−Protein Interaction Inhibitor Based on Molecular Binding Determinants Analysis. *J. Med. Chem.* **2014**, *57*, 2736−2745.
(286) Guan, Y.; Sun, H.; Li, Y.; Pan, P.; Li, D.; Hou, T. The Competitive Binding between Inhibitors and Substrates of HCV NS3/4A Protease: A General Mechanism of Drug Resistance. *Antiviral Res.* **2014**, *103*, 60−70.
(287) Guan, Y.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T. Exploring Resistance Mechanisms of HCV NS3/4A Protease Mutations to MK5172: Insight from Molecular Dynamics Simulations and Free Energy Calculations. *Mol. BioSyst.* **2015**, *11*, 2568−2578.
(288) Pan, P.; Li, Y.; Yu, H.; Sun, H.; Hou, T. Molecular Principle of Topotecan Resistance by Topoisomerase I Mutations through Molecular Modeling Approaches. *J. Chem. Inf. Model.* **2013**, *53*, 997−1006.
(289) 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.
(290) Homans, S. W. Water, Water Everywhere–except Where It Matters? *Drug Discovery Today* **2007**, *12*, 534−539.
(291) Zhang, J.; Murugan, N. A.; Tian, Y.; Bertagnin, C.; Fang, Z.; Kang, D.; Kong, X.; Jia, H.; Sun, Z.; Jia, R.; et al. Structure-Based Optimization of N-Substituted Oseltamivir Derivatives as Potent Anti-Influenza a Virus Agents with Significantly Improved Potency against Oseltamivir-Resistant N1-H274Y Variant. *J. Med. Chem.* **2018**, *61*, 9976−9999.
(292) Sporn, M. B.; Liby, K. T. Nrf2 and Cancer: The Good, the Bad and the Importance of Context. *Nat. Rev. Cancer* **2012**, *12*, 564.
(293) Suzuki, T.; Motohashi, H.; Yamamoto, M. Toward Clinical Application of the Keap1−Nrf2 Pathway. *Trends Pharmacol. Sci.* **2013**, *34*, 340−346.
(294) Westermaier, Y.; Ruiz-Carmona, S.; Theret, I.; Perron-Sierra, F.; Poissonnet, G.; Dacquet, C.; Boutin, J. A.; Ducrot, P.; Barril, X. Binding Mode Prediction and MD/MMPBSA-Based Free Energy Ranking for Agonists of REV-Erbalpha/NCoR. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 755−775.
(295) Kocakaya, S. O. The Molecular Modeling of Novel Inhibitors of Protein Tyrosine Phosphatase 1b Based on Catechol by MD and MM-GB (PB)/SA Calculations. *Bull. Korean Chem. Soc.* **2014**, *35*, 1769−1776.
(296) Guo, X. L.; Shen, K.; Wang, F.; Lawrence, D. S.; Zhang, Z. Y. Probing the Molecular Basis for Potent and Selective Protein-Tyrosine Phosphatase 1b Inhibition. *J. Biol. Chem.* **2002**, *277*, 41014−41022.
(297) Mena-Ulecia, K.; Tiznado, W.; Caballero, J. Study of the Differential Activity of Thrombin Inhibitors Using Docking, QSAR, Molecular Dynamics, and MM-GBSA. *PLoS One* **2015**, *10*, e0142774.
(298) Kong, X.; Sun, H.; Pan, P.; Zhu, F.; Chang, S.; Xu, L.; Li, Y.; Hou, T. Importance of Protein Flexibility in Molecular Recognition: A Case Study on Type-I1/2 Inhibitors of ALK. *Phys. Chem. Chem. Phys.* **2018**, *20*, 4851−4863.
(299) Xue, W.; Jin, X.; Ning, L.; Wang, M.; Liu, H.; Yao, X. Exploring the Molecular Mechanism of Cross-Resistance to HIV-1 Integrase Strand Transfer Inhibitors by Molecular Dynamics Simulation and Residue Interaction Network Analysis. *J. Chem. Inf. Model.* **2013**, *53*, 210−222.
(300) Sun, H.-Y.; Hou, T.-J.; Zhang, H.-Y. Finding Chemical Drugs for Genetic Diseases. *Drug Discovery Today* **2014**, *19*, 1836−1840.
(301) Fulle, S.; Saini, J. S.; Homeyer, N.; Gohlke, H. Complex Long-Distance Effects of Mutations That Confer Linezolid Resistance in the Large Ribosomal Subunit. *Nucleic Acids Res.* **2015**, *43*, 7731−7743.
(302) Xue, W. W.; Pan, D. B.; Yang, Y.; Liu, H. X.; Yao, X. J. Molecular Modeling Study on the Resistance Mechanism of HCV NS3/4A Serine Protease Mutants R155K, A156V and D168a to TMC435. *Antiviral Res.* **2012**, *93*, 126−137.
(303) Sun, H.-Y.; Ji, F.-Q.; Fu, L.-Y.; Wang, Z.-Y.; Zhang, H.-Y. Structural and Energetic Analyses of Snps in Drug Targets and Implications for Drug Therapy. *J. Chem. Inf. Model.* **2013**, *53*, 3343−3351.
(304) Hou, T.; Yu, R. Molecular Dynamics and Free Energy Studies on the Wild-Type and Double Mutant HIV-1 Protease Complexed with Amprenavir and Two Amprenavir-Related Inhibitors: Mechanism for Binding and Drug Resistance. *J. Med. Chem.* **2007**, *50*, 1177−1188.
(305) Li, L.; Li, Y.; Zhang, L.; Hou, T. Theoretical Studies on the Susceptibility of Oseltamivir against Variants of 2009 a/H1N1 Influenza Neuraminidase. *J. Chem. Inf. Model.* **2012**, *52*, 2715−2729.
(306) Pan, D. B.; Niu, Y. Z.; Xue, W. W.; Bai, Q. F.; Liu, H. X.; Yao, X. J. Computational Study on the Drug Resistance Mechanism of

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

Hepatitis C Virus NS5B RNA-Dependent RNA Polymerase Mutants to BMS-791325 by Molecular Dynamics Simulation and Binding Free Energy Calculations. *Chemom. Intell. Lab. Syst.* **2016**, *154*, 185−193.

(307) Sun, H.; Li, Y.; Li, D.; Hou, T. Insight into Crizotinib Resistance Mechanisms Caused by Three Mutations in ALK Tyrosine Kinase Using Free Energy Calculation Approaches. *J. Chem. Inf. Model.* **2013**, *53*, 2376−2389.

(308) Sun, H.; Li, Y.; Tian, S.; Wang, J.; Hou, T. P-Loop Conformation Governed Crizotinib Resistance in G2032r-Mutated Ros1 Tyrosine Kinase: Clues from Free Energy Landscape. *PLoS Comput. Biol.* **2014**, *10*, e1003729.

(309) Matthew, A. N.; Leidner, F.; Newton, A.; Petropoulos, C. J.; Huang, W.; Ali, A.; KurtYilmaz, N.; Schiffer, C. A. Molecular Mechanism of Resistance in a Clinically Significant Double-Mutant Variant of HCV NS3/4A Protease. *Structure* **2018**, *26*, 1360−1372.

(310) Kannan, S.; Venkatachalam, G.; Lim, H. H.; Surana, U.; Verma, C. Conformational Landscape of the Epidermal Growth Factor Receptor Kinase Reveals a Mutant Specific Allosteric Pocket. *Chem. Sci.* **2018**, *9*, 5212−5222.

(311) Ragland, D. A.; Whitfield, T. W.; Lee, S.-K.; Swanstrom, R.; Zeldovich, K. B.; Kurt-Yilmaz, N.; Schiffer, C. A. Elucidating the Interdependence of Drug Resistance from Combinations of Mutations. *J. Chem. Theory Comput.* **2017**, *13*, 5671−5682.

(312) Latallo, M.; Cortina, G.; Faham, S.; Nakamoto, R.; Kasson, P. Predicting Allosteric Mutants That Increase Activity of a Major Antibiotic Resistance Enzyme. *Chem. Sci.* **2017**, *8*, 6484−6492.

(313) Cui, Y.-L.; Zheng, Q.-C.; Zhang, J.-L.; Xue, Q.; Wang, Y.; Zhang, H.-X. Molecular Dynamic Investigations of the Mutational Effects on Structural Characteristics and Tunnel Geometry in CYP17A1. *J. Chem. Inf. Model.* **2013**, *53*, 3308−3317.

(314) Özen, A. e. l.; Sherman, W.; Schiffer, C. A. Improving the Resistance Profile of Hepatitis C NS3/4A Inhibitors: Dynamic Substrate Envelope Guided Design. *J. Chem. Theory Comput.* **2013**, *9*, 5693−5705.

(315) Awad, M. M.; Katayama, R.; McTigue, M.; Liu, W.; Deng, Y.-L.; Brooun, A.; Friboulet, L.; Huang, D.; Falk, M. D.; Timofeevski, S.; et al. Acquired Resistance to Crizotinib from a Mutation in Cd74−Ros1. *N. Engl. J. Med.* **2013**, *368*, 2395−2401.

(316) Berry, T.; Luther, W.; Bhatnagar, N.; Jamin, Y.; Poon, E.; Sanda, T.; Pei, D.; Sharma, B.; Vetharoy, W. R.; Hallsworth, A.; et al. The ALK F1174l Mutation Potentiates the Oncogenic Activity of MYCN in Neuroblastoma. *Cancer Cell* **2012**, *22*, 117−130.

(317) Bresler, S. C.; Weiser, D. A.; Huwe, P. J.; Park, J. H.; Krytska, K.; Ryles, H.; Laudenslager, M.; Rappaport, E. F.; Wood, A. C.; et al. ALK Mutations Confer Differential Oncogenic Activation and Sensitivity to ALK Inhibition Therapy in Neuroblastoma. *Cancer Cell* **2014**, *26*, 682−694.

(318) Choi, Y. L.; Soda, M.; Yamashita, Y.; Ueno, T.; Takashima, J.; Nakajima, T.; Yatabe, Y.; Takeuchi, K.; Hamada, T.; Haruta, H.; et al. Eml4-ALK Mutations in Lung Cancer That Confer Resistance to ALK Inhibitors. *N. Engl. J. Med.* **2010**, *363*, 1734−1739.

(319) Heuckmann, J. M.; Holzel, M.; Sos, M. L.; Heynck, S.; Balke-Want, H.; Koker, M.; Peifer, M.; Weiss, J.; Lovly, C. M.; et al. ALK Mutations Conferring Differential Resistance to Structurally Diverse ALK Inhibitors. *Clin. Cancer Res.* **2011**, *17*, 7394−7401.

(320) Shaw, A. T.; Friboulet, L.; Leshchiner, I.; Gainor, J. F.; Bergqvist, S.; Brooun, A.; Burke, B. J.; Deng, Y.-L.; Liu, W.; Dardaei, L.; et al. Resensitization to Crizotinib by the Lorlatinib ALK Resistance Mutation L1198F. *N. Engl. J. Med.* **2016**, *374*, 54−61.

(321) Li, J.; Sun, R.; Wu, Y.; Song, M.; Li, J.; Yang, Q.; Chen, X.; Bao, J.; Zhao, Q. L1198F Mutation Resensitizes Crizotinib to ALK by Altering the Conformation of Inhibitor and ATP Binding Sites. *Int. J. Mol. Sci.* **2017**, *18*, 482.

(322) Tartari, C. J.; Gunby, R. H.; Coluccia, A. M.; Sottocornola, R.; Cimbro, B.; Scapozza, L.; Donella-Deana, A.; Pinna, L. A.; Gambacorti-Passerini, C. Characterization of Some Molecular Mechanisms Governing Autoactivation of the Catalytic Domain of the Anaplastic Lymphoma Kinase. *J. Biol. Chem.* **2008**, *283*, 3743−3750.

(323) Janoueix-Lerosey, I.; Lequin, D.; Brugieres, L.; Ribeiro, A.; de Pontual, L.; Combaret, V.; Raynal, V.; Puisieux, A.; Schleiermacher, G.; et al. Somatic and Germline Activating Mutations of the ALK Kinase Receptor in Neuroblastoma. *Nature* **2008**, *455*, 967−970.

(324) He, M.; Li, W.; Zheng, Q.; Zhang, H. A Molecular Dynamics Investigation into the Mechanisms of Alectinib Resistance of Three ALK Mutants. *J. Cell. Biochem.* **2018**, *119*, 5332−5342.

(325) Sun, H. Y.; Ji, F. Q. A Molecular Dynamics Investigation on the Crizotinib Resistance Mechanism of C1156Y Mutation in ALK. *Biochem. Biophys. Res. Commun.* **2012**, *423*, 319−324.

(326) Hatcher, J. M.; Bahcall, M.; Choi, H. G.; Gao, Y.; Sim, T.; George, R.; Jänne, P. A.; Gray, N. S. Discovery of Inhibitors That Overcome the G1202R Anaplastic Lymphoma Kinase Resistance Mutation. *J. Med. Chem.* **2015**, *58*, 9296−9308.

(327) Shin, S.; Mah, S.; Hong, S.; Park, H. Discovery of Low Micromolar Dual Inhibitors for Wild Type and L1196M Mutant of Anaplastic Lymphoma Kinase through Structure-Based Virtual Screening. *J. Chem. Inf. Model.* **2016**, *56*, 802−810.

(328) Geng, K.; Xia, Z.; Ji, Y.; Zhang, R.; Sun, D.; Ai, J.; Song, Z.; Geng, M.; Zhang, A. Discovery of 2,4-Diarylaminopyrimidines Bearing a Resorcinol Motif as Novel ALK Inhibitors to Overcome the G1202R Resistant Mutation. *Eur. J. Med. Chem.* **2018**, *144*, 386−397.

(329) Xue, W.; Ban, Y.; Liu, H.; Yao, X. Computational Study on the Drug Resistance Mechanism against HCV NS3/4A Protease Inhibitors Vaniprevir and Mk-5172 by the Combination Use of Molecular Dynamics Simulation, Residue Interaction Network, and Substrate Envelope Analysis. *J. Chem. Inf. Model.* **2014**, *54*, 621−633.

(330) Shimba, N.; Kamiya, N.; Nakamura, H. Model Building of Antibody−Antigen Complex Structures Using GBSA Scores. *J. Chem. Inf. Model.* **2016**, *56*, 2005−2012.

(331) Maffucci, I.; Contini, A. Improved Computation of Protein−Protein Relative Binding Energies with the Nwat-MMGBSA Method. *J. Chem. Inf. Model.* **2016**, *56*, 1692−1704.

(332) Simões, I. s. C.; Costa, I. s. P.; Coimbra, J. o. T.; Ramos, M. J.; Fernandes, P. A. New Parameters for Higher Accuracy in the Computation of Binding Free Energy Differences Upon Alanine Scanning Mutagenesis on Protein−Protein Interfaces. *J. Chem. Inf. Model.* **2017**, *57*, 60−72.

(333) Li, M.; Petukh, M.; Alexov, E.; Panchenko, A. R. Predicting the Impact of Missense Mutations on Protein−Protein Binding Affinity. *J. Chem. Theory Comput.* **2014**, *10*, 1770−1780.

(334) Xu, L.; Li, Y.; Li, D.; Xu, P.; Tian, S.; Sun, H.; Liu, H.; Hou, T. Exploring the Binding Mechanisms of Mif to Cxcr2 Using Theoretical Approaches. *Phys. Chem. Chem. Phys.* **2015**, *17*, 3370−3382.

(335) Xu, L.; Li, Y.; Sun, H.; Li, D.; Hou, T. Structural Basis of the Interactions between Cxcr4 and Cxcl12/SDF-1 Revealed by Theoretical Approaches. *Mol. BioSyst.* **2013**, *9*, 2107−2117.

(336) Salmaso, V.; Sturlese, M.; Cuzzolin, A.; Moro, S. Exploring Protein-Peptide Recognition Pathways Using a Supervised Molecular Dynamics Approach. *Structure* **2017**, *25*, 655−662.

(337) Ivanov, S. M.; Huber, R. G.; Warwicker, J.; Bond, P. J. Energetics and Dynamics across the Bcl-2-Regulated Apoptotic Pathway Reveal Distinct Evolutionary Determinants of Specificity and Affinity. *Structure* **2016**, *24*, 2024−2033.

(338) Hou, T.; Li, Y.; Wang, W. Prediction of Peptides Binding to the PKA Riiα Subunit Using a Hierarchical Strategy. *Bioinformatics* **2011**, *27*, 1814−1821.

(339) Hard, R.; Li, N.; He, W.; Ross, B.; Mo, G. C.; Peng, Q.; Stein, R. S.; Komives, E.; Wang, Y.; Zhang, J.; Wang, W. Deciphering and Engineering Chromodomain-Methyllysine Peptide Recognition. *Sci. Adv.* **2018**, *4*, eaau1447.

(340) Xu, Z.; Hou, T.; Li, N.; Xu, Y.; Wang, W. Proteome-Wide Detection of Abl1 SH3-Binding Peptides by Integrating Computational Prediction and Peptide Microarray. *Mol. Cell. Proteomics* **2012**, *11*, O111.010389.

(341) Hou, T.; Chen, K.; McLaughlin, W. A.; Lu, B.; Wang, W. Computational Analysis and Prediction of the Binding Motif and

Protein Interacting Partners of the Abl SH3 Domain. *PLoS Comput. Biol.* **2006**, *2*, e1.
(342) Ottaviani, A.; Iacovelli, F.; Idili, A.; Falconi, M.; Ricci, F.; Desideri, A. Engineering a Responsive DNA Triple Helix into an Octahedral DNA Nanostructure for a Reversible Opening/Closing Switching Mechanism: A Computational and Experimental Integrated Study. *Nucleic Acids Res.* **2018**, *46*, 9951−9959.
(343) Soni, A.; Khurana, P.; Singh, T.; Jayaram, B. A DNA Intercalation Methodology for an Efficient Prediction of Ligand Binding Pose and Energetics. *Bioinformatics* **2017**, *33*, 1488−1496.
(344) Fogolari, F.; Corazza, A.; Esposito, G. Accuracy Assessment of the Linear Poisson−Boltzmann Equation and Reparametrization of the OBC Generalized B Orn Model for Nucleic Acids and Nucleic Acid−Protein Complexes. *J. Comput. Chem.* **2015**, *36*, 585−596.
(345) Zhou, Q.; Sun, X.; Xia, X.; Fan, Z.; Luo, Z.; Zhao, S.; Shakhnovich, E.; Liang, H. Exploring the Mutational Robustness of Nucleic Acids by Searching Genotype Neighborhoods in Sequence Space. *J. Phys. Chem. Lett.* **2017**, *8*, 407−414.
(346) Morrison, E. A.; Bowerman, S.; Sylvers, K. L.; Wereszczynski, J.; Musselman, C. A. The Conformation of the Histone H3 Tail Inhibits Association of the BPTF PHD Finger with the Nucleosome. *eLife* **2018**, *7*, e31481.
(347) Rajagopalan, M.; Balasubramanian, S.; Ramaswamy, A. Insights into the RNA Binding Mechanism of Human L1-Orf1p: A Molecular Dynamics Study. *Mol. BioSyst.* **2017**, *13*, 1728−1743.
(348) Orr, A. A.; Gonzalez-Rivera, J. C.; Wilson, M.; Bhikha, P. R.; Wang, D.; Contreras, L. M.; Tamamis, P. A High-Throughput and Rapid Computational Method for Screening of RNA Post-Transcriptional Modifications That Can Be Recognized by Target Proteins. *Methods* **2018**, *143*, 34−47.
(349) Zhang, N.; Chen, Y.; Zhao, F.; Yang, Q.; Simonetti, F. L.; Li, M. PremPDI Estimates and Interprets the Effects of Missense Mutations on Protein−DNA Interactions. *PLoS Comput. Biol.* **2018**, *14*, e1006615.
(350) Bowerman, S.; Wereszczynski, J. Effects of macroH2A and H2A. Z on Nucleosome Dynamics as Elucidated by Molecular Dynamics Simulations. *Biophys. J.* **2016**, *110*, 327−337.
(351) Beuerle, M. G.; Dufton, N. P.; Randi, A. M.; Gould, I. R. Molecular Dynamics Studies on the DNA-Binding Process of ERG. *Mol. BioSyst.* **2016**, *12*, 3600−3610.
(352) Xu, L.; Kong, R.; Zhu, J.; Sun, H.; Chang, S. Unraveling the Conformational Determinants of Larp7 and 7SK Small Nuclear RNA by Theoretical Approaches. *Mol. BioSyst.* **2016**, *12*, 2613−2621.
(353) Vreven, T.; Hwang, H.; Pierce, B. G.; Weng, Z. Prediction of Protein−Protein Binding Free Energies. *Protein Sci.* **2012**, *21*, 396−404.
(354) Andreani, J.; Guerois, R. Evolution of Protein Interactions: From Interactomes to Interfaces. *Arch. Biochem. Biophys.* **2014**, *554*, 65−75.
(355) Ramakrishnan, G.; Chandra, N. R.; Srinivasan, N. From Workstations to Workbenches: Towards Predicting Physicochemically Viable Protein−Protein Interactions across a Host and a Pathogen. *IUBMB Life* **2014**, *66*, 759−774.
(356) Metz, A.; Ciglia, E.; Gohlke, H. Modulating Protein−Protein Interactions: From Structural Determinants of Binding to Druggability Prediction to Application. *Curr. Pharm. Des.* **2012**, *18*, 4630−4647.
(357) Gonzalez-Ruiz, D.; Gohlke, H. Targeting Protein−Protein Interactions with Small Molecules: Challenges and Perspectives for Computational Binding Epitope Detection and Ligand Finding. *Curr. Med. Chem.* **2006**, *13*, 2607−2625.
(358) Nisius, B.; Sha, F.; Gohlke, H. Structure-Based Computational Analysis of Protein Binding Sites for Function and Druggability Prediction. *J. Biotechnol.* **2012**, *159*, 123−134.
(359) Takemura, K.; Guo, H.; Sakuraba, S.; Matubayasi, N.; Kitao, A. Evaluation of Protein−Protein Docking Model Structures Using All-Atom Molecular Dynamics Simulations Combined with the Solution Theory in the Energy Representation. *J. Chem. Phys.* **2012**, *137*, 215105.
(360) Seidel, S. A.; Dijkman, P. M.; Lea, W. A.; van den Bogaart, G.; Jerabek-Willemsen, M.; Lazic, A.; Joseph, J. S.; Srinivasan, P.; Baaske, P.; et al. Microscale Thermophoresis Quantifies Biomolecular Interactions under Previously Challenging Conditions. *Methods* **2013**, *59*, 301−315.
(361) Syafrizayanti; Betzen, C.; Hoheisel, J. D.; Kastelic, D. Methods for Analyzing and Quantifying Protein−Protein Interaction. *Expert Rev. Proteomics* **2014**, *11*, 107−120.
(362) Ngounou Wetie, A. G.; Sokolowska, I.; Woods, A. G.; Roy, U.; Deinhardt, K.; Darie, C. C. Protein−Protein Interactions: Switch from Classical Methods to Proteomics and Bioinformatics-Based Approaches. *Cell. Mol. Life Sci.* **2014**, *71*, 205−228.
(363) Anishchenko, I.; Kundrotas, P. J.; Tuzikov, A. V.; Vakser, I. A. Protein Models: The Grand Challenge of Protein Docking. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 278−287.
(364) Venkatesan, K.; Rual, J. F.; Vazquez, A.; Stelzl, U.; Lemmens, I.; Hirozane-Kishikawa, T.; Hao, T.; Zenkner, M.; Xin, X. F.; et al. An Empirical Framework for Binary Interactome Mapping. *Nat. Methods* **2009**, *6*, 83−90.
(365) Stelzl, U.; Worm, U.; Lalowski, M.; Haenig, C.; Brembeck, F. H.; Goehler, H.; Stroedicke, M.; Zenkner, M.; Schoenherr, A.; et al. A Human Protein−Protein Interaction Network: A Resource for Annotating the Proteome. *Cell* **2005**, *122*, 957−968.
(366) Kastritis, P. L.; Bonvin, A. M. Are Scoring Functions in Protein−Protein Docking Ready to Predict Interactomes? Clues from a Novel Binding Affinity Benchmark. *J. Proteome Res.* **2010**, *9*, 2216−2225.
(367) Vakser, I. A.; Kundrotas, P. Predicting 3D Structures of Protein−Protein Complexes. *Curr. Pharm. Biotechnol.* **2008**, *9*, 57−66.
(368) Huang, S. Y. Search Strategies and Evaluation in Protein−Protein Docking: Principles, Advances and Challenges. *Drug Discovery Today* **2014**, *19*, 1081−1096.
(369) Xu, D.; Si, Y. B.; Meroueh, S. O. A Computational Investigation of Small-Molecule Engagement of Hot Spots at Protein−Protein Interaction Interfaces. *J. Chem. Inf. Model.* **2017**, *57*, 2250−2272.
(370) Ganesan, A.; Moon, T. C.; Barakat, K. H. Revealing the Atomistic Details Behind the Binding of B7−1 to Cd28 and CTLA-4: A Comprehensive Protein−Protein Modelling Study. *Biochim. Biophys. Acta, Gen. Subj.* **2018**, *1862*, 2764−2778.
(371) Petukh, M.; Li, M.; Alexov, E. Predicting Binding Free Energy Change Caused by Point Mutations with Knowledge-Modified MM/PBSA Method. *PLoS Comput. Biol.* **2015**, *11*, e1004276.
(372) Massova, I.; Kollman, P. A. Computational Alanine Scanning To Probe Protein−Protein Interactions: A Novel Approach To Evaluate Binding Free Energies. *J. Am. Chem. Soc.* **1999**, *121*, 8133−8143.
(373) Zoete, V.; Meuwly, M.; Karplus, M. Study of the Insulin Dimerization: Binding Free Energy Calculations and Per-Residue Free Energy Decomposition. *Proteins: Struct., Funct., Genet.* **2005**, *61*, 79−93.
(374) Guo, J.; Wang, X.; Sun, H.; Liu, H.; Yao, X. The Molecular Basis of IGF-II/IGF2R Recognition: A Combined Molecular Dynamics Simulation, Free-Energy Calculation and Computational Alanine Scanning Study. *J. Mol. Model.* **2012**, *18*, 1421−1430.
(375) Delaine, C.; Alvino, C. L.; McNeil, K. A.; Mulhern, T. D.; Gauguin, L.; De Meyts, P.; Jones, E. Y.; Brown, J.; Wallace, J. C.; Forbes, B. E. A Novel Binding Site for the Human Insulin-Like Growth Factor-II (IGF-II)/Mannose 6-Phosphate Receptor on IGF-II. *J. Biol. Chem.* **2007**, *282*, 18886−18894.
(376) Forbes, B. E.; McNeil, K. A.; Scott, C. D.; Surinya, K. H.; Cosgrove, L. J.; Wallace, J. C. Contribution of Residues A54 and L55 of the Human Insulin-like Growth Factor-II (IGF-II) a Domain to Type 2 IGF Receptor Binding Specificity. *Growth Factors* **2001**, *19*, 163−173.
(377) Roth, B. V.; Burgisser, D. M.; Luthi, C.; Humbel, R. E. Mutants of Human Insulin-Like Growth Factor II: Expression and Characterization of Analogs with a Substitution of Tyr27 and/or a

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

Deletion of Residues 62−67. *Biochem. Biophys. Res. Commun.* **1991**, *181*, 907−914.
(378) Sakano, K.-i.; Enjoh, T.; Numata, F.; Fujiwara, H.; Marumoto, Y.; Higashihashi, N.; Sato, Y.; Perdue, J. F.; Fujita-Yamaguchi, Y. The Design, Expression, and Characterization of Human Insulin-Like Growth Factor II (IGF-II) Mutants Specific for Either the IGF-II/Cation-Independent Mannose 6-Phosphate Receptor or IGF-I Receptor. *J. Biol. Chem.* **1991**, *266*, 20626−20635.
(379) Zhang, Z.; Lu, L.; Zhang, Y.; Hua Li, C.; Wang, C. X.; Zhang, X. Y.; Tan, J. J. A Combinatorial Scoring Function for Protein−RNA Docking. *Proteins: Struct., Funct., Genet.* **2017**, *85*, 741−752.
(380) Iwakiri, J.; Hamada, M.; Asai, K.; Kameda, T. Improved Accuracy in RNA-Protein Rigid Body Docking by Incorporating Force Field for Molecular Dynamics Simulation into the Scoring Function. *J. Chem. Theory Comput.* **2016**, *12*, 4688−4697.
(381) Kim, D. H.; Im, H.; Jee, J. G.; Jang, S. B.; Yoon, H. J.; Kwon, A. R.; Kang, S. M.; Lee, B. J. Beta-Arm Flexibility of Hu from Staphylococcus Aureus Dictates the DNA-Binding and Recognition Mechanism. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 3273−3289.
(382) Ishida, H.; Matsumoto, A. Mechanism for Verification of Mismatched and Homoduplex Dnas by Nucleotides-Bound Muts Analyzed by Molecular Dynamics Simulations. *Proteins: Struct., Funct., Genet.* **2016**, *84*, 1287−1303.
(383) Galindo-Murillo, R.; Cheatham, T. E., 3rd DNA Binding Dynamics and Energetics of Cobalt, Nickel, and Copper Metallopeptides. *ChemMedChem* **2014**, *9*, 1252−1259.
(384) Cabrita, L. D.; Dobson, C. M.; Christodoulou, J. Protein Folding on the Ribosome. *Curr. Opin. Struct. Biol.* **2010**, *20*, 33−45.
(385) Voss, N. R.; Gerstein, M.; Steitz, T. A.; Moore, P. B. The Geometry of the Ribosomal Polypeptide Exit Tunnel. *J. Mol. Biol.* **2006**, *360*, 893−906.
(386) Fedyukina, D. V.; Cavagnero, S. Protein Folding at the Exit Tunnel. *Annu. Rev. Biophys.* **2011**, *40*, 337−359.
(387) Chen, C.; Wang, E.; Liu, P.; Xiao, Y. Simulation Study of the Role of the Ribosomal Exit Tunnel on Protein Folding. *Phys. Rev. E* **2013**, *87*, 022701.
(388) Wang, E.; Wang, J.; Chen, C.; Xiao, Y. Computational Evidence That Fast Translation Speed Can Increase the Probability of Cotranslational Protein Folding. *Sci. Rep.* **2015**, *5*, 15316.
(389) Wolf, A.; Schoof, S.; Baumann, S.; Arndt, H. D.; Kirschner, K. N. Structure-Activity Relationships of Thiostrepton Derivatives: Implications for Rational Drug Design. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 1205−1215.
(390) Yam, W. K.; Wahab, H. A. Molecular Insights into 14-Membered Macrolides Using the MM-PBSA Method. *J. Chem. Inf. Model.* **2009**, *49*, 1558−1567.
(391) Ayoub, A. T.; Elrefaiy, M. A.; Arakawa, K. Computational Prediction of the Mode of Binding of Antitumor Lankacidin C to Tubulin. *Acs Omega* **2019**, *4*, 4461−4471.
(392) Mokdad, A.; Krasovska, M. V.; Sponer, J.; Leontis, N. B. Structural and Evolutionary Classification of G/U Wobble Basepairs in the Ribosome. *Nucleic Acids Res.* **2006**, *34*, 1326−1341.
(393) Murray, J. B.; Meroueh, S. O.; Russell, R. J. M.; Lentzen, G.; Haddad, J.; Mobashery, S. Interactions of Designer Antibiotics and the Bacterial Ribosomal Aminoacyl-Trna Site. *Chem. Biol.* **2006**, *13*, 129−138.
(394) Meroueh, S. O.; Mobashery, S. Conformational Transition in the Aminoacyl T-RNA Site of the Bacterial Ribosome Both in the Presence and Absence of an Aminoglycoside Antibiotic. *Chem. Biol. Drug Des.* **2007**, *69*, 291−297.
(395) Kaukonen, M.; Soderhjelm, P.; Heimdal, J.; Ryde, U. QM/MM-PBSA Method To Estimate Free Energies for Reactions in Proteins. *J. Phys. Chem. B* **2008**, *112*, 12537−12548.
(396) Wu, E. L.; Mei, Y.; Han, K.; Zhang, J. Z. Quantum and Molecular Dynamics Study for Binding of Macrocyclic Inhibitors to Human α-Thrombin. *Biophys. J.* **2007**, *92*, 4244−4253.
(397) Shi, J.; Lu, Z.; Zhang, Q.; Wang, M.; Wong, C. F.; Liu, J. H. Supplementing the PBSA Approach with Quantum Mechanics To Study the Binding between CDK2 and N^2-Substituted O^6-Cyclohexylmethoxyguanine Inhibitors. *J. Theor. Comput. Chem.* **2010**, *9*, 543−559.
(398) Cole, D. J.; Skylaris, C. K.; Rajendra, E.; Venkitaraman, A. R.; Payne, M. C. Protein−Protein Interactions from Linear-Scaling First-Principles Quantum-Mechanical Calculations. *EPL* **2010**, *91*, 37004.
(399) Mucs, D.; Bryce, R. A. The Application of Quantum Mechanics in Structure-Based Drug Design. *Expert Opin. Drug Discovery* **2013**, *8*, 263−276.
(400) Fedorov, D. G.; Nagata, T.; Kitaura, K. Exploring Chemistry with the Fragment Molecular Orbital Method. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7562−7577.
(401) Jia, X.; Wang, X.; Liu, J.; Zhang, J. Z.; Mei, Y.; He, X. An Improved Fragment-Based Quantum Mechanical Method for Calculation of Electrostatic Solvation Energy of Proteins. *J. Chem. Phys.* **2013**, *139*, 214104.
(402) Wang, Y.; Liu, J.; Li, J.; He, X. Fragment-Based Quantum Mechanical Calculation of Protein−Protein Binding Affinities. *J. Comput. Chem.* **2018**, *39*, 1617−1628.
(403) Zhou, T.; Huang, D.; Caflisch, A. Quantum Mechanical Methods for Drug Design. *Curr. Top. Med. Chem.* **2010**, *10*, 33−45.
(404) Misini Ignjatović, M. M.; Mikulskis, P.; Söderhjelm, P.; Ryde, U. Can MM/GBSA Calculations Be Sped Up by System Truncation? *J. Comput. Chem.* **2018**, *39*, 361−372.
(405) Stone, J. E.; Hardy, D. J.; Ufimtsev, I. S.; Schulten, K. GPU-Accelerated Molecular Modeling Coming of Age. *J. Mol. Graphics Modell.* **2010**, *29*, 116−125.
(406) Wang, J. C.; Lin, J. H.; Chen, C. M.; Perryman, A. L.; Olson, A. J. Robust Scoring Functions for Protein−Ligand Interactions with Quantum Chemical Charge Models. *J. Chem. Inf. Model.* **2011**, *51*, 2528−2537.
(407) Chung, J. Y.; Hah, J. M.; Cho, A. E. Correlation between Performance of QM/MM Docking and Simple Classification of Binding Sites. *J. Chem. Inf. Model.* **2009**, *49*, 2382−2387.
(408) Cho, A. E.; Rinaldo, D. Extension of QM/MM Docking and Its Applications to Metalloproteins. *J. Comput. Chem.* **2009**, *30*, 2609−2616.
(409) Fischer, B.; Fukuzawa, K.; Wenzel, W. Receptor-Specific Scoring Functions Derived from Quantum Chemical Models Improve Affinity Estimates for in-Silico Drug Discovery. *Proteins: Struct., Funct., Genet.* **2008**, *70*, 1264−1273.
(410) Wichapong, K.; Rohe, A.; Platzer, C.; Slynko, I.; Erdmann, F.; Schmidt, M.; Sippl, W. Application of Docking and QM/MM-GBSA Rescoring to Screen for Novel Myt1 Kinase Inhibitors. *J. Chem. Inf. Model.* **2014**, *54*, 881−893.
(411) Raha, K.; Merz, K. M., Jr. Large-Scale Validation of a Quantum Mechanics Based Scoring Function: Predicting the Binding Affinity and the Binding Mode of a Diverse Set of Protein−Ligand Complexes. *J. Med. Chem.* **2005**, *48*, 4558−4575.
(412) Hayik, S. A.; Dunbrack, R., Jr.; Merz, K. M., Jr. A Mixed QM/MM Scoring Function to Predict Protein−Ligand Binding Affinity. *J. Chem. Theory Comput.* **2010**, *6*, 3079−3091.
(413) Dobes, P.; Fanfrlik, J.; Rezac, J.; Otyepka, M.; Hobza, P. Transferable Scoring Function Based on Semiempirical Quantum Mechanical Pm6-Dh2Method: Cdk2 with 15 Structurally Diverse Inhibitors. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 223−235.
(414) Grater, F.; Schwarzl, S. M.; Dejaegere, A.; Fischer, S.; Smith, J. C. Protein/Ligand Binding Free Energies Calculated with Quantum Mechanics/Molecular Mechanics. *J. Phys. Chem. B* **2005**, *109*, 10474−10483.
(415) Mikulskis, P.; Genheden, S.; Wichmann, K.; Ryde, U. A Semiempirical Approach to Ligand-Binding Affinities: Dependence on the Hamiltonian and Corrections. *J. Comput. Chem.* **2012**, *33*, 1179−1189.
(416) Ryde, U.; Nilsson, K. Quantum Refinement - a Combination of Quantum Chemistry and Protein Crystallography. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 259−275.
(417) Yu, N.; Yennawar, H. P.; Merz, K. M., Jr. Refinement of Protein Crystal Structures Using Energy Restraints Derived from

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX

Linear-Scaling Quantum Mechanics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 322−332.
(418) Fu, Z.; Li, X.; Merz, K. M., Jr. Accurate Assessment of the Strain Energy in a Protein-Bound Drug Using QM/MM X-Ray Refinement and Converged Quantum Chemistry. *J. Comput. Chem.* **2011**, *32*, 2587−2597.
(419) Wang, R.; Lu, Y.; Wang, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, *46*, 2287−2303.
(420) Mikulskis, P.; Cioloboc, D.; Andrejic, M.; Khare, S.; Brorsson, J.; Genheden, S.; Mata, R. A.; Soderhjelm, P.; Ryde, U. Free-Energy Perturbation and Quantum Mechanical Study of Sampl4 Octa-Acid Host-Guest Binding Energies. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 375−400.
(421) Ciancetta, A.; Genheden, S.; Ryde, U. A QM/MM Study of the Binding of RAPTA Ligands to Cathepsin B. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 729−742.
(422) Lu, H.; Huang, X.; AbdulHameed, M. D.; Zhan, C. G. Binding Free Energies for Nicotine Analogs Inhibiting Cytochrome P450 2A6 by a Combined Use of Molecular Dynamics Simulations and QM/MM-PBSA Calculations. *Bioorg. Med. Chem.* **2014**, *22*, 2149−2156.
(423) Ryde, U.; Soderhjelm, P. Ligand-Binding Affinity Estimates Supported by Quantum-Mechanical Methods. *Chem. Rev.* **2016**, *116*, 5520−5566.
(424) Gouda, H.; Kuntz, I. D.; Case, D. A.; Kollman, P. A. Free Energy Calculations for Theophylline Binding to an RNA Aptamer: Comparison of MM-PBSA and Thermodynamic Integration Methods. *Biopolymers* **2003**, *68*, 16−34.
(425) Wang, L.; Deng, Y.; Knight, J. L.; Wu, Y.; Kim, B.; Sherman, W.; Shelley, J. C.; Lin, T.; Abel, R. Modeling Local Structural Rearrangements Using FEP/REST: Application to Relative Binding Affinity Predictions of Cdk2 Inhibitors. *J. Chem. Theory Comput.* **2013**, *9*, 1282−1293.
(426) 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: Struct., Funct., Genet.* **2004**, *55*, 34−43.
(427) Bea, I.; Cervello, E.; Kollman, P. A.; Jaime, C. Molecular Recognition by *β*-Cyclodextrin Derivatives: FEP vs MM/PBSA Study. *Comb. Chem. High Throughput Screening* **2001**, *4*, 605−611.
(428) Obiol-Pardo, C.; Rubio-Martinez, J. Comparative Evaluation of MMPBSA and XSCORE To Compute Binding Free Energy in XIAP−Peptide Complexes. *J. Chem. Inf. Model.* **2007**, *47*, 134−142.
(429) 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.
(430) Ericksen, S. S.; Wu, H.; Zhang, H.; Michael, L. A.; Newton, M. A.; Hoffmann, F. M.; Wildman, S. A. Machine Learning Consensus Scoring Improves Performance across Targets in Structure-Based Virtual Screening. *J. Chem. Inf. Model.* **2017**, *57*, 1579−1590.
(431) Ain, Q. U.; Aleksandrova, A.; Roessler, F. D.; Ballester, P. J. Machine-Learning Scoring Functions to Improve Structure-Based Binding Affinity Prediction and Virtual Screening. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2015**, *5*, 405−424.
(432) Heck, G. S.; Pintro, V. O.; Pereira, R. R.; de Ávila, M. B.; Levin, N. M. B.; de Azevedo, W. F. Supervised Machine Learning Methods Applied to Predict Ligand-Binding Affinity. *Curr. Med. Chem.* **2017**, *24*, 2459−2470.
(433) de Avila, M. B.; Xavier, M. M.; Pintro, V. O.; de Azevedo, W. F., Jr. Supervised Machine Learning Techniques to Predict Binding Affinity. A Study for Cyclin-Dependent Kinase 2. *Biochem. Biophys. Res. Commun.* **2017**, *494*, 305−310.
(434) Wojcikowski, M.; Ballester, P. J.; Siedlecki, P. Performance of Machine-Learning Scoring Functions in Structure-Based Virtual Screening. *Sci. Rep.* **2017**, *7*, 46710.

DOI: 10.1021/acs.chemrev.9b00055
Chem. Rev. XXXX, XXX, XXX−XXX