Comprehensive Evaluation of End-Point Free Energy Techniques in Carboxylated-Pillar[6]arene Host-guest Binding: II. Regression and Dielectric Constant

Xiao Liu[1*], Lei Zheng[2], Yalong Cong[3], Zhihao Gong[4,5], Zhixiang Yin[1], John Z.H. Zhang[2,3,6,7*], Zhirong Liu[8], Zhaoxi Sun[8*]

[1]*School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620, China*

[2]*NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China*

[3]*School of Chemistry and Molecular Engineering, East China Normal University, Shanghai, 200062, China*

[4]*School of Micro-Nano Electronics, Zhejiang University, Hangzhou 310027, China*

[5]*Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou 310027, China*

[6]*Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China*

[7]*Department of Chemistry, New York University, NY, NY 10003, USA*

[8]*College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China*

*To whom correspondence should be addressed:

Xiao Liu, liuxiaode2013@163.com

John Z. H. Zhang, john.zhang@nyu.edu

Zhaoxi Sun, proszx@163.com

Abstract

End-point free energy calculations as a powerful tool have been widely applied in protein-ligand and protein-protein interactions. It is often recognized that these end-point techniques serve as an option of intermediate accuracy and computational cost compared with more rigorous statistical mechanic models (e.g., alchemical transformation) and coarser molecular docking. However, it is observed that this intermediate level of accuracy does not hold in relatively simple and prototypical host-guest systems. Specifically, in our previous work investigating a set of carboxylated-pillar[6]arene host-guest complexes, end-point methods provide free energy estimates deviating significantly from the experimental reference, and the rank of binding affinities is also incorrectly computed. These observations suggest the unsuitability and inapplicability of standard end-point free energy techniques in host-guest systems, and alteration and

development are required to make them practically usable. In this work, we consider two ways to improve the performance of end-point techniques. The first one is the PBSA_E regression that varies the weights of different free energy terms in the end-point calculation procedure, while the second one is considering the interior dielectric constant as an additional variable in the end-point equation. By detailed investigation of the calculation procedure and the simulation outcome, we prove that these two treatments (i.e., regression and dielectric constant) are manipulating the end-point equation in a somehow similar way, i.e., weakening the electrostatic contribution and strengthening the non-polar terms, although there are still many detailed differences between these two methods. With the trained end-point scheme, the RMSE of the computed affinities is improved from the standard ~12 kcal/mol to ~2.4 kcal/mol, which is comparable to another altered end-point method (ELIE) trained with system-specific data. This phenomenon along with the extremely efficient optimized-structure computation procedure suggests the regression (i.e., PBSA_E as well as its GBSA_E extension) as a practically applicable solution that brings end-point methods back into the library of usable tools for host-guest binding. However, the dielectric-constant-variable scheme cannot effectively minimize the experiment-calculation discrepancy for absolute binding affinities, but is able to improve the calculation of affinity ranks. This phenomenon is somehow different from the protein-ligand case and suggests the difference between host-guest and biomacromolecular (protein-ligand and protein-protein) cases. Therefore, the spectrum of tools usable in protein-ligand cases could be unsuitable for host-guest binding, and numerical validations are necessary to screen out really workable solutions in these ‘prototypical’ situations.

Keywords: Pillar[*n*]arenes, Host-guest Binding, End-point Free Energy Methods, Regression, Dielectric Constant

1. Introduction.

Computational techniques are widely applied in drug discovery.[1-8] Considering the balance between computational cost and prediction accuracy, often a hierarchical scheme is employed.[9-12] A large set of potentially useful candidates are first screened with crude methods, after which more detailed techniques are employed to pick promising molecules from a smaller screened set. The costliest technique on this computational ladder would be the alchemical method, which constructs a thermodynamic cycle with several artificial transformation legs connecting states of interest (e.g., the bound and unbound states).[13-17] However, as such rigorous treatment requires extensive sampling in the configurational space, the computational cost of the alchemical technique is too huge to be practically usable when a number of molecules are under investigation. Methods that are a bit less computationally demanding are end-point free energy methods of MM/PBSA and MM/GBSA.[18-20] These end-point techniques are computationally feasible due to their approximated calculation procedure. The most widely employed end-point scheme uses the single-trajectory approximation and only samples the well-defined bound conformation. The enthalpic and entropic contributions are computed with snapshots from this single trajectory and ultimately combined to derive the free energy of binding.[21-25] Compared with end-point techniques that rely on all-atom force fields and molecular simulation, an even coarser and cheaper technique is molecular docking. The method incorporates simpler energy/scoring functions and sampling methods, which enables fast screening and highly efficient conformational sampling.[26-28]

Compared with protein-ligand complexes, host-guest systems are smaller in size and simpler in structural features. Specifically, the host molecules are often cylindrical macrocycles with symmetrical rims that have a limited number of rotatable bonds, which leads to their limited conformational flexibility and thus smaller conformational space.[29-34] Further, the number of functional groups is also small in these macromolecular containers, which leads to their limited capability of forming distinct interaction patterns with the external agents. Host-guest systems are often considered prototypical cases of protein-ligand complexes. Thus, macrocyclic hosts are widely employed in many laboratory and industrial applications, e.g., drug delivery, catalyzed synthesis, and molecular machinery.[35-39] Pillar[*n*]arenes as a host family popular in modern chemical research have satisfactory ligand-binding ability.[40-44] However, their low aqueous solubility limits the practical application in biomedical treatments. Chemical modifications of the host cavity are often considered an effective solution.[45-48] For instance, carboxylated-pillar[6]arenes (WP6) with enhanced water-solubility, strong guest-binding ability, tunability and intermediate cavity volume and entrance size are promising instruments in drug delivery and reservoir.[49-52]

Despite the massive application of end-point free energy techniques in protein-ligand and protein-protein interactions, understandings of their practical performance in prototypical host-guest binding are rather limited. This calls for systematic benchmark calculations with detailed investigations of the underlying physics. Based on the obtained insights, constructive solutions to improve their performance can finally be proposed, making them practically usable tools in the computational investigation of host-guest binding. In a recent work, we report a thorough evaluation of end-point methods with the standard (popular) procedure in a set of WP6 host-guest systems.[53] Specifically, we consider four modelling details including the scoring function used in docking (two scoring functions), the charge scheme for solutes (three charge sets), the water model for solvation (two water models), and the end-point method used in free energy estimation (two end-point schemes). The resulting calculation regimes include 24 combinations, covering almost all popular selections in modern researches. However, all of these combined end-point calculation schemes could neither reproduce the experimental binding affinities nor correctly compute the rank of affinities, which proves that end-point free energy calculations are actually inapplicable to host-guest binding. This is in stark contrast to the recognized performance of end-point techniques in protein-ligand complex. Possible reasons causing the underperformance of end-point calculations in host-guest complex along with further directions for developments have been discussed extensively in our previous work.[53] In the current work, building on previous observations of the standard end-point procedure, we expect to extend the evaluation by considering two altered regimes that could lead to improved performance. Specifically, we consider the regression and dielectric-constant-variable alterations that change the weights of free energy terms in the end-point calculation. In the PBSA_E regression,[54] the weights of free energy terms are pre-trained with end-point and experimental results of a huge number of protein-ligand complexes. The fitted formula achieves better reproductions of both the absolute binding affinity and the rank of affinities (e.g., with RMSE ~2 kcal/mol).[54] However, due to the existence of many differences between the designed usage of PBSA_E and the current host-guest calculation (e.g., structure generation and force field), whether the weighting factors in this transferable scheme are suitable for host-guest systems remains unknown. In the dielectric-constant-variable scheme,[12, 55-57] the polar interactions inside the complex are scaled by an effective internal dielectric constant of the heterogeneous dielectric environment. It is often observed that values slightly larger than the standard unity could lead to pronounced improvements in the prediction quality of end-point calculations, but whether this treatment applies to host-guest systems remains unclarified. Extensive numerical experiments performed in this work suggest that both techniques could effectively improve the quality of prediction. Further, through a detailed analysis of the calculation

procedure of the two regimes, we provide insights into the relationships between the two altered schemes and explanations of their dependence on various modelling parameters.

2. Modelling Details.

2.1. Model Construction.

The system construction follows exactly the same protocol used in our previous work.[53] Briefly, the 3D chemical structures of all molecules shown in Fig. 1 are grabbed from the GitHub site of the SAMPL9 challenge.[58] To make our benchmark test of end-point calculations thorough, we consider a set of parameter/force-field combinations. The changeable modelling details include the charge scheme for solutes, the water model, and the scoring function used in the generation of the initial bound structure.

For atomic charges, two popular fixed-charge models including AM1-BCC[59] and restrained electrostatic potential (RESP)[60] are considered. Two RESP charge sets are generated due to their noticeable difference in the electrostatic potential (ESP) around each molecule observed in our previous work.[53] The fitting target of the RESP-1 charge set is the molecular ESP scanned at the traditional HF[61-63]/6-31G*, while the target level of the second RESP charge set (RESP-2) is B3LYP[64-66]/def2-TZVPP with the IEFPCM implicit solvent. All the other missing force-field parameters are obtained from the second generation of the general AMBER force field (GAFF2)[67] and a recent publication adding the parameters for Si-related species.[68]

For water molecules, we employed two popular 3-point models including TIP3P[69, 70] and SPC/E.[71] As all parameters of these two water models (e.g., charges and force constant) differ, the solvation environments produced by these water models differ, which ultimately leads to the water-model-dependent behavior of the conformational/energetic preference of each molecule involved in the inter-molecular coordination and alters the binding thermodynamics.

The bound structure of each host-guest complex remains unknown, which is often solved via molecular docking. Following our previous work,[53] we use the Autodock Vina[72] program to generate the initial guess, i.e., picking the top-1 (the most stable) structure provided by docking. As the starting configuration could have a significant impact on the simulation outcome, especially when the simulation is performed in an unbiased way. Thus, we consider two scoring functions including Autodock4 (AD4) and Vina[72, 73] in molecular docking. The two scoring functions differ in both the definition of energy terms and computational complexity, and show somehow different behaviors in many benchmark calculations.[72, 74, 75] Also, according to the analysis of binding modes produced by the two scoring functions reported in the first paper of this WP6 end-point series, for many host-guest pairs there are obvious structural differences.[53]

The consideration of three charge schemes for solutes, two water models, and two scoring functions leads to a total of 12 combinations of model-construction protocols. Using each protocol, for each host-guest pair, the most stable bound configuration (i.e., the binding pose with the highest docking score) is selected. Then, the complex is parametrized (e.g., the above-mentioned charge schemes) and solvated (with TIP3P or SPC/E) to create a simulation box. Non-polarizable monovalent spherical counter ions[76, 77] of Na^+ or Cl^- are added for neutralization. Periodic boundary conditions are employed to replicate the unit cell in the whole space.

2.2. Molecular Simulations.

For each simulation box constructed with the above-mentioned protocols, we perform some level of molecular simulation to accumulate configurational snapshots. Note that this procedure is only employed in the dielectric-constant-variable case, as the PBSA_E model uses a single optimized structure in free energy estimation. In molecular simulations, starting from the docking-produced binding pose, we perform 5000 steps minimization, 300 ps constant-volume heating with weak harmonic restraints on solutes and 1 ns NPT equilibration to reach the physiological condition. For production, we perform 100 ns NPT simulations with a sampling interval of 10 ps. We employ the SHAKE constraints on bonds involving hydrogen atoms to remove high-frequency motions.[78, 79] We use Langevin dynamics[80] with the collision frequency of 2 ps^{-1} for temperature regulation and a time step of 2 fs. The cutoff for non-covalent interactions is set to 10 Å, and the PME method is used to treat long-range electrostatics.[81] The GPU version of the pmemd engine with the hybrid precision (SPFP) in the AMBER[82] suite is used for dynamics propagation.

The trajectories are used as input to the end-point equation to extract binding affinities. The standard end-point free energy estimation follows the following formula,

$$\Delta G_{\text{binding}} = \Delta E_{\text{elec}} + \Delta E_{\text{vdW}} + \Delta G_{\text{solv,polar}} + \Delta G_{\text{solv,non-polar}} - T\Delta S_{\text{gas}} \quad (1).$$

The first two components arise from the gas-phase enthalpic contribution, i.e., inter-molecular electrostatic and vdW interactions. The third and fourth terms come from the (de)solvation effects. The last term is the entropy change upon the formation of the inter-molecular coordination. The gas-phase enthalpic contribution (i.e., the first two terms) can be directly computed with all-atom force fields. The solvation terms are often computed with implicit solvent models. In our calculations, the polar part $\Delta G_{\text{solv,polar}}$ is computed with PB[83] or the popular GB^{OBC} model (with the second set of modified Bondi radii),[84, 85] and the non-polar contribution is treated with the solvent-accessible surface area method.[86] The last term can be calculated with normal mode analysis (NMA),[87] quasi-harmonic approximation[88] and so on.[89-92] In our

previous WP6 work, the end-point calculation using the standard procedure selects the NMA method due to its popularity, and 50 snapshots equally spaced in the 100 ns unbiased sampling are included in calculation.[53]

3. Results and Discussions.

3.1. Altering the weights of free energy terms.

The first alteration of the end-point calculation procedure is regression, which adjusts the weights of free energy terms in the end-point calculation to improve the results. Such parameter adjustment does not expect to improve the accuracy of every term involved in the calculation, but seeks for error cancellations that minimize the discrepancy between calculation and experiment. In the current work, we do not fit the weights of terms in Eq. (1) with numerical data in the current WP6 host-guest dataset or similar host-guest complexes, as this system-specific regression procedure would limit the transferability of the calculation procedure. By contrast, we use the pre-fitted formula of PBSA_E,[54] which is obtained by applying the regression analysis on a huge set of protein-ligand complexes. The PBSA_E method divides the free energy terms in Eq. (1) into four groups, the detailed definitions and weights of which are presented below.

$$\Delta G_{\text{binding}} = 0.03037(\Delta E_{\text{elec}} + \Delta G_{\text{solv,polar}}) + 0.07791\Delta E_{\text{vdW}} + 1.2193\Delta G_{\text{solv,non-polar}} + 0.1854N_{\text{rot}} \qquad (2).$$

The polar contributions (both gas-phase and solvation) are included in the first term, which can thus be considered the net electrostatic contribution to the final binding free energy. The second term is the inter-molecular vdW interaction, the third term is the non-polar solvation. These two terms are often grouped into non-polar interactions. The last entropic contribution, which is often computed with the costly NMA in standard end-point calculations, is substituted/approximated with a computationally cheaper scheme, the number of rotatable bonds of the ligand/guest. Further, previous applications of the PBSA_E method suggest that the free energy estimate obtained with a single optimized configuration could be closer to the experimental value than that with extensive sampling. Thus, the method is often used with a single energy-minimized bound structure. In many protein-ligand complexes, the PBSA_E method outperforms popular scoring functions and also standard end-point methods.[54] Overall, the above features of the PBSA_E scheme make it a low-cost end-point scheme. Aside from the original PBSA_E formula, we also consider the GB extension, i.e., directly substituting the polar solvation term of PB with GB solvation. The GB extension of the PBSA_E scheme is named GBSA_E, the validity of which would be discussed later.

The regression procedure of the PBSA_E model differs from the current WP6 host-guest cases in many aspects. First, in the training of the PBSA_E model, the solutes are described with the RESP-1 charge

scheme and an older version of the general AMBER force field (i.e., GAFF), while in the current calculations we consider three charge sets (AM1-BCC, RESP-1 and RESP-2) and the latest version of the transferable force field (i.e., GAFF2). Second, the PBSA_E model is trained with a large set of protein-ligand complexes. These inter-molecular coordinations between biomacromolecules and small external agents obviously differ from the current WP6 host-guest cases. Third, the training and application of the PBSA_E scheme always use the crystal structure of the complex, while in the current WP6 host-guest case the experimental coordination feature is unavailable and we are actually dealing with docking-produced bound structures. Due to the existence of these obvious differences between the training procedure and the designed usage of the PBSA_E weights and the current investigation of WP6 host-guest binding, it is difficult to say whether the educated end-point regime could perform satisfactorily in the current calculations. Thus, some numerical experiments are needed.

Some insights from weighting factors.

Before discussing numerical results, we first analyze the weighting factors in detail to grab some insights into what the numerical regression really does to the standard end-point calculation. The weight of the polar interactions is the smallest one in Eq. (2) and is much smaller than 1, which suggests the suppression of the electrostatic contributions to the net free energy change. As for the non-polar contribution, the contribution from the gas-phase vdW interaction is also weakened (smaller than 1) but is still much larger than the electrostatics, and the non-polar part of solvation receives the largest weighting factor in the formula. This phenomenon suggests that the PBSA_E regression actually emphasizes the non-polar contribution (especially the non-polar solvation) but weakens the impact of electrostatics in end-point free energy estimation.

Based on these observations, the impacts of variations in the modelling details could be predicted. It is expected that the charge scheme would not have a significant impact on the PBSA_E free energy outcome, as the electrostatic interactions are scaled down significantly. Second, for the water models, as only geometry optimization is performed starting from the docking-produced bound configuration, the perturbation to the initial configuration is expected to be minimal and the water model does not have a significant impact on the bound structure used in the PBSA_E calculation. Third, the docking procedure is expected to have some impacts on the PBSA_E estimate, as the initial conditions show some differences due to the differences in scoring functions. However, the magnitude of this docking-induced variation is difficult to quantify without some numerical experiments. Finally, for the GBSA_E extension of the original PBSA_E formula, as the polar contribution of solvation is also scaled down significantly and the PB and GB

solvation terms do not exhibit huge differences, the free energy estimates obtained with these two solvation models (i.e., the original PBSA_E and the extension GBSA_E) are also expected to be similar. Overall, none of the modelling details (charge scheme for solutes, water model, docking procedure and implicit solvent) is expected to have a significant impact on the free energy outcome in the trained end-point scheme.

Numerical performance of educated end-point regimes.

The PBSA_E and GBSA_E estimates with all combinations of modelling parameters (i.e., charge schemes, water models, and scoring functions) are computed. The numerical results along with the experimental values[93] are summarized in Table S1-S4. We use the root-mean-squared error (RMSE), the mean signed error (MSE), the Kendall τ rank coefficient[94] and the Pearlman's predictive index (PI)[95] to evaluate the quality of computation. The RMSE and Kendall τ for all combinations of modelling parameters are compared in Fig. 2. The results of another regressed PBSA scheme named extended linear interaction energy (ELIE)[96] and the best estimate of standard end-point calculations obtained in our previous work[53] (AD4+RESP-1+SPC/E+MM/GBSA) are also provided for comparison.

For the error sizes presented in Fig. 2a, we can clearly see that the trained weights in PBSA_E improve the reproduction of the absolute binding affinities significantly. The ~12 kcal/mol RMSE under the standard end-point calculation is improved to ~2.4 kcal/mol. The error sizes of different modelling protocols (charge scheme, water model and scoring function) are similar, which agrees with the above analysis of weighting factors in Eq. (2). Interestingly, we also observe that the PBSA_E RMSE is similar to the ELIE result. As PBSA_E uses pre-fitted weights rather than system-specific parameters in ELIE, the transferable PBSA_E scheme can be considered a more robust option in end-point free energy calculations for host-guest binding and also protein-ligand interaction.

As for the ranking coefficient shown in Fig. 2b, it is still observed that the trained end-point schemes outperform the standard procedure significantly. For the pre-fitted PBSA_E and GBSA_E schemes, the charge scheme for solutes, the water model, and the scoring function used in docking all have some influences on the free energy outcome. Although it is difficult to conclude which parameter combination performs best, considering the consistent error sizes in Fig. 2a, we believe that each of the parameter combinations is usable. The value of τ obtained with ELIE is larger than any pre-fitted scheme, which is expected due to the system-specific training procedure of the ELIE regime. However, it should be noted that the improvement is only observed for ranking coefficients and the magnitude of improvement is limited. It is reasonable to believe that if we further add WP6-related experimental data to the PBSA_E training set, the calculation accuracy would also be improved to the ELIE level.

A detailed view of trained end-point estimates.

The above comparison of error metrics and ranking coefficients only provides insights about the overall performance of the whole dataset, but the behaviors of detailed free energy estimates remain unknown. Thus, the PBSA_E estimates obtained with different modelling parameters are compared in Fig. 3a to provide a detailed view of the modelling-parameter dependence. We only present the results obtained with the AD4 scoring function for clarity. It is clearly shown that in most cases, the charge scheme for solutes and the water model do not have significant impacts on the PBSA_E estimates. However, we can still identify several cases with ~1 kcal/mol variation when the modelling parameters are varied. The Vina-produced results are similar and thus would not be discussed again. Similar observations could also be obtained for the GBSA_E case in Fig. 3b. Overall, the impact of variations of modelling parameters is limited (negligible in most cases). Considering the above modelling-parameter-independent behavior of the PBSA_E outcome, it can be concluded that as long as a reasonable combination of modelling parameters is employed, the PBSA_E/GBSA_E method would produce free energy estimates of good quality. As GB calculations are faster than PB, the GBSA_E scheme serves as a faster yet accurate alternative to the original PBSA_E.

The quality metrics of the pre-fitted PBSA_E and GBSA_E schemes are similar to those of the ELIE scheme trained with system-specific data. An interesting question to ask is whether these methods are actually producing similar free energy predictions. To answer this question, we compare the free energy estimates with the three trained end-point schemes in Fig. 3c. Due to the modelling-parameter-independent behavior of pre-fitted PBSA_E and GBSA_E schemes, for these two pre-fitted schemes we only present the results obtained with the AD4+AM1-BCC+TIP3P combination of modelling parameters. It can be seen that the PBSA_E and GBSA_E estimates are very similar, which agrees with the implicit-solvent-independent behavior from analysis of weighting factors in Eq. (2). By contrast, the ELIE results obviously differ from those with the two pre-fitted regimes, which suggests the dissimilarities of the pre-fitted PBSA_E and GBSA_E regimes and the system-specific ELIE and provides an explanation of the better ranking reproduction of the ELIE scheme in Fig. 2b.

Overall, tuning the weighting factors in end-point calculation of host-guest binding could lead to improved performance. The transferable PBSA_E scheme does not have significant dependence of the modelling parameters, and the extended GBSA_E regime performs similarly to PBSA_E. As GB calculations are relatively faster than PB solvation, this GBSA_E scheme serves as a more efficient option than PBSA_E. Training the end-point calculation with system-specific experimental data could be helpful, but this introduces an extra fitting step and is inapplicable to newly encountered systems without existing

results. As the transferable PBSA_E and GBSA_E could produce free energy estimates with small errors, we recommend them as practically applicable tools in end-point screening of host-guest complexes.

3.2. Varying the interior dielectric constant.

The interior dielectric constant ε_{in} as an additional tunable parameter in the end-point equation is often altered to improve the accuracy of end-point free energy calculations. It is often recommended to enlarge this effective internal dielectric constant from the standard unity to something like 2 or 4 in a uniform or residue-type-specific manner, especially for polar and charged regions.[97-99] In this section, we investigate whether this scheme would improve the calculation accuracy in WP6 host-guest binding and provides some insights into its connections with the trained PBSA_E regime.

The dielectric-constant-variable form of the end-point equation can be expressed as

$$\Delta G_{\text{binding}} = \Delta E_{\text{elec}}\left(\varepsilon_{\text{in}}\right) + \Delta E_{\text{vdW}} + \Delta G_{\text{solv,polar}}\left(\varepsilon_{\text{in}}\right) + \Delta G_{\text{solv,non-polar}} - T\Delta S_{\text{gas}} \quad (3).$$

It can be clearly seen that the dielectric constant of the interior involves in the calculation of the gas-phase electrostatic contribution and the polar part of the (de)solvation effect, which is similar to the alteration of weighting factors in PBSA_E, i.e., Eq. (2). The gas-phase electrostatics exhibit a reciprocal dependence on the dielectric constant (c.f., Coulomb equation), while the polar solvation term shows a more complex non-linear behavior (see references[55-57, 100] for details). Thus, the dielectric-constant-variable scheme is somehow different from the PBSA_E regression. However, it can be safely to conclude the trend that each component of the polar contributions (i.e., either gas-phase or solvation) would become weakened upon increased dielectric constant, although the net change could be non-monotonic. As the standard weight of the electrostatic contribution is 1 and the refitted weights in PBSA_E is ~0.03 (approximately 1/33), in our numerical experiment of dielectric constants we scan values between 1 and 40 to cover this region. We specifically test dielectric constants including 2, 4, 6, 8, 10, 20, 30, and 40. It should be noted that in the PBSA_E equation, the weights of the other terms (e.g., vdW) are also varied, which also has impacts on the end-point estimates. Therefore, the dielectric-constant-variable scheme is expected to behave somehow differently compared with the educated end-point regime.

We first discuss about the results obtained with GB solvation in the dielectric-constant-variable calculations. As the best parameter combination observed in our previous work is AD4+RESP-1+SPC/E(+MM/GBSA),[53] in the current dielectric-constant-variable test, we first include this parameter combination in the benchmark calculation. Numerical experiments are also performed for another water

model (producing the AD4+RESP-1+TIP3P set), another charge set (AD4+RESP-2+SPC/E and AD+RESP-2+TIP3P) and another scoring function (Vina+RESP-1+SPC/E). The detailed free energy estimates along with quality metrics obtained with the above-mentioned interior dielectric constants are summarized in Table S5-S9. We first check the dielectric-constant dependence of end-point estimates under the AD4+RESP-1+SPC/E and AD4+RESP-2+TIP3P parameter combinations in Fig. 4a-b, in order to understand what the variation of interior dielectric constant does to each system. With the increase of the dielectric constant, the individual free energy estimate first experiences some monotonic (increasing or decreasing) behavior. Most of the changes happen when ε_{in} is smaller than 10 (or 6 for some cases), and when ε_{in} reaches a large value the free energy estimate seems 'converged' on this degree of freedom, which is in accordance with the expectation from analysis of Eq. (3). This dependence is generally valid for all parameter combinations, as shown in the numerical data in Table S5-S9. From this dependence, we can expect the quality metrics to vary significantly for small ε_{in} and become relatively unchanged/stable for large ε_{in}. The dielectric-constant-variable results under PB solvation for the selected five modelling parameter combinations are summarized in Table S10-S14. The dielectric-constant dependence under AD4+RESP-1+SPC/E and AD4+RESP-2+TIP3P parameter combinations are presented in Fig. 4c-d, where a behavior similar to the GB case could be observed and thus would not be discussed further.

We then check the numerical results of the quality metrics in Fig. 5. Under the GB solvation, in Fig. 5a, with the increase of the interior dielectric constant, the error size increases monotonically for all parameter combinations, although for AD4+RESP-2+TIP3P some fluctuation is observed at ε_{in}=2. If the aim is to reproduce the absolute values of binding affinities, then varying the interior dielectric constant would not be a good choice. The monotonic increase of the error size somehow differs from existing experiences accumulated in many protein-ligand and protein-protein complexes, where monotonic decreasing behaviors or some complicated dependences are observed.[55, 56, 98-103] Such differences could be attribute to the differences between the interior microenvironments in protein-ligand and host-guest systems. The former is more heterogeneous and hydrophobic, while the latter is simpler, more solvent-exposed and thus more hydrophilic. This phenomenon suggests that the accumulated experiences in protein-ligand situations including applicable alterations could be inapplicable in host-guest complexes, and numerical experiments are necessary to validate the usage of these solutions. As for the ranking information in Fig. 5b, it is clearly shown that the increase of ε_{in} significantly improves the ranking calculation. Most of betterments happen when ε_{in} is smaller than 6, and the larger-than-6 ranking coefficients are somehow similar and only fluctuating behaviors could be observed. Considering the monotonically increasing behavior of RMSE,

adjusting ε_{in} to some values close to 6 could be a balanced choice that provides good ranking information without further devastating the absolute values of binding affinities. The improved ranking coefficients under the five parameter combinations are all about 0.32, which is similar to the PBSA_E/GBSA_E results. The dielectric-constant dependence of RMSE and the ranking coefficient τ under PB solvation is presented in Fig. 5c-d. The behaviors are still similar to the GB case. Thus, the dielectric constant as another adjustable parameter in the end-point equation is indeed helpful to improve the performance of end-point methods. However, it should be noted that the trained end-point schemes also ensure better reproduction of the absolute values of binding affinities, which makes them a more robust choice than the dielectric-constant-variable scheme.

4. Concluding Remarks.

Host-guest systems, due to their small size and relatively simple structural feature, are considered prototypical cases for protein-ligand and protein-protein complexes. The hosts are often cylindrical macrocycles with symmetrical rims and the guests are drug-like molecules. In drug screening within biomacromolecular systems, end-point free energy techniques are widely used as a practical tool with balanced accuracy and efficiency. However, their performance in host-guest binding is often not well understood. In our previous work, a comprehensive evaluation of standard end-point calculations in WP6 host-guest binding is presented. Specifically, the protocol for model construction and configuration generation is obtained by combining three modelling details, including the docking procedure (two scoring functions), the solute charge set (three charge sets), and the solvent model (two water models). The resulting 12 combinations cover popular selections in modern end-point free energy calculations. The configurations sampled with the 12 modelling protocols are analyzed with two popular end-point free energy techniques, i.e., MM/PBSA and MM/GBSA. Astonishingly, the free energy estimates from all modelling and free energy estimation protocols show significant deviations from the experimental reference, with RMSE ~12 kcal/mol and τ ~0.08. Error sizes of this level suggest the inapplicability of standard end-point calculations in host-guest systems.

The failure of the standard end-point calculation procedure calls for alterations and developments. In this work, we consider two altered schemes including the trained weights of free energy terms in end-point free energy estimation (i.e., PBSA_E and its GBSA_E extension) and the dielectric-constant-variable scheme. By investigating the detailed weighting factors in the PBSA_E and GBSA_E calculations, we notice that the trained end-point schemes are severely weakening the polar contribution and marginally strengthening

different components in the non-polar part. This treatment shares some similarities with the dielectric-constant-variable regime, where a larger-than-unity interior dielectric constant is selected to weaken the polar contributions to the end-point free energy estimate. However, the differences of the remaining parts in the educated end-point and the dielectric-constant-variable schemes lead to significant differences in their accuracies.

Extensive numerical experiments are performed to obtain free energy estimates with the two altered schemes. Significant improvements are observed for the trained end-point schemes (PBSA_E and GBSA_E). The RMSE reaches ~2.4 kcal/mol and the ranking coefficient τ is ~0.35. Such prediction quality is comparable to the ELIE regime trained with system-specific data (i.e., published WP6 host-guest binding affinities) and is similar to that in the protein-ligand situations. As the weighting factors of free energy terms in our trained end-point regimes are transferable, the calculation is only performed on optimized bound structure and the costly normal mode calculation is replaced by the number of rotatable bonds of the guest, the PBSA_E and GBSA_E models serve as an extremely efficient and practically usable tool in virtual screening of host-guest complexes. As for the dielectric-constant-variable scheme, the reproduction of experimental values of absolute binding affinities is degraded upon the increase of the interior dielectric constant ε_{in}, but the calculation of the rank of binding affinities is improved when ε_{in} is increased from the standard unity to ~6. For values larger than 6, the ranking coefficients are not improved and stay in the neighborhood of 0.32, which is comparable to the trained schemes (i.e., PBSA_E and GBSA_E). The monotonic increase of error size (RMSE) is somehow different from biomacromolecular cases (e.g., protein-protein and protein-ligand), which could be attribute to the differences between the interior microenvironments in protein-ligand and host-guest complexes. The differences between the 'prototypical' host-guest binding and the complex protein-ligand situations indicate that end-point investigations of host-guest binding cannot fully follow the accumulated experiences in protein-ligand binding. Usable tools extensively tested in protein-ligand cases still require comprehensive numerical validations to consolidate their usage in host-guest binding.

Overall, both the PBSA_E regression and the dielectric-constant-variable scheme are helpful alterations of the standard end-point procedure. For accuracy, PBSA_E and its GBSA_E extension simultaneously ensure the reproduction of absolute binding affinities and a better ranking calculation, while the variation of the interior dielectric constant improves the ranking calculation but degrades the reproduction of absolute binding affinities. For efficiency, the trained end-point schemes using optimized bound structure and avoiding the costly normal mode calculation are much more efficient than the dielectric-constant-variable

regime. Therefore, we consider the trained end-point schemes as a more powerful tool. Aside from the two alterations of end-point calculations considered in this work, it should be noted that there are still other potentially workable techniques that could improve the prediction quality, e.g., enhanced sampling, entropy estimation and using other docking procedures instead of the current top-1 structure from Autodock. In the following works of our WP6 host-guest series, comprehensive benchmarks on these remaining solutions would be presented.

Acknowledgement

This work was supported by National Natural Science Foundation of China (Grant No. 21633001), Beijing Natural Science Foundation (Grant No. 7224357), National Natural Science Foundation of China (Grant No. 22107063) and National Natural Science Foundation of China (Grant No. 62072296). Part of the simulation was performed on the high-performance computing platform of the Center for Life Science (Peking University). We thank Dr. Zhe Huai (XtalPi) and the anonymous reviewers for valuable comments and critical reading.

Conflict of Interest Statement

There are no conflicts of interest to declare.

Supporting Information Description

The detailed values of PBSA_E and GBSA_E estimates of WP6 host-guest binding free energies obtained with three charge schemes, two water models, initial bound structures from AD4 and Vina docking, and the dielectric-constant-dependent MM/GBSA and MM/PBSA estimates for five combinations of modelling parameters are given in the supporting information.

References

1. Zeng, J.; Huang, Z., From Levinthal's Paradox to the Effects of Cell Environmental Perturbation on Protein Folding. *Curr. Med. Chem.* **2019**, 26, 7537-7554.
2. Panday, S. K.; Alexov, E., Protein–Protein Binding Free Energy Predictions with the MM/PBSA Approach Complemented with the Gaussian-Based Method for Entropy Estimation. *ACS Omega* **2022**.
3. Chen, J.; Zeng, Q.; Wang, W.; Hu, Q.; Bao, H., Q61 mutant-mediated dynamics changes of the GTP-KRAS complex probed by Gaussian accelerated molecular dynamics and free energy landscapes. *RSC Advances* **2022**, 12, 1742-1757.
4. Sotriffer, C. A.; Sanschagrin, P.; Matter, H.; Klebe, G., SFCscore: scoring functions for affinity prediction of protein–ligand complexes. *Proteins: Structure, Function, and Bioinformatics* **2008**, 73, 395-419.
5. Pecina, A.; Eyrilmez, S. M.; Köprülüoğlu, C.; Miriyala, V. M.; Lepšík, M.; Fanfrlík, J.; Řezáč, J.; Hobza, P., SQM/COSMO Scoring Function: Reliable Quantum-Mechanical Tool for Sampling and Ranking in Structure-Based Drug Design. *ChemPlusChem* **2020**, 85, 2362-2371.
6. Wang, X., Conformational Fluctuations in GTP-Bound K-Ras: A Metadynamics Perspective with Harmonic Linear Discriminant Analysis. *J. Chem. Inf. Model.* **2021**.
7. Goel, H.; Yu, W.; MacKerell, A. D., hERG Blockade Prediction by Combining Site Identification by Ligand Competitive Saturation and Physicochemical Properties. *Chemistry* **2022**, 4, 630-646.
8. Nicolaï, A.; Petiot, N.; Grassein, P.; Delarue, P.; Neiers, F.; Senet, P., Free-Energy Landscape Analysis of Protein-Ligand Binding: The Case of Human Glutathione Transferase A1. *Applied Sciences* **2022**, 12, 8196.
9. Giordano, D.; Biancaniello, C.; Argenio, M. A.; Facchiano, A., Drug Design by Pharmacophore and Virtual Screening Approach. *Pharmaceuticals* **2022**, 15, 646.
10. Homeyer, N.; Gohlke, H., Free Energy Calculations by the Molecular Mechanics Poisson−Boltzmann Surface Area Method. *Molecular Informatics* **2012**, 31, 114-122.
11. 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.
12. 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: Structure, Function, and Bioinformatics* **2011**, 79, 1940-1951.
13. Huai, Z.; Yang, H.; Sun, Z., Binding thermodynamics and interaction patterns of human purine nucleoside phosphorylase-inhibitor complexes from extensive free energy calculations. *J. Comput.-Aided Mol. Des.* **2021**.
14. Khalak, Y.; Tresadern, G.; de Groot, B. L.; Gapsys, V., Non-equilibrium approach for binding free energies in cyclodextrins in SAMPL7: force fields and software. *J. Comput.-Aided Mol. Des.* **2021**, 35, 49-61.
15. Sun, Z.; Wang, X.; Zhang, J. Z. H., BAR-based Optimum Adaptive Sampling Regime for Variance Minimization in Alchemical Transformation. *Phys. Chem. Chem. Phys.* **2017**, 19, 15005-15020.
16. Sun, Z.; Wang, M.; He, Q.; Liu, Z., Molecular Modelling of Ionic Liquids: Force-Field Validation and Thermodynamic Perspective from Large-Scale Fast-Growth Solvation Free Energy Calculations. *Adv. Theory Simul.* **2022**, 2200274.
17. Sun, Z.; Gong, Z.; Zheng, L.; Payam, K.; Huai, Z.; Liu, Z., Molecular Modelling of Ionic Liquids: General Guidelines on Fixed-Charge Force Fields for Balanced Descriptions. *Journal of Ionic Liquids* **2022**, 2, 100043.
18. Ferrari, A. M.; Degliesposti, G.; Sgobba, M.; Rastelli, G., Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. *Biorg. Med. Chem.* **2007**, 15, 7865-7877.
19. Rapp, C.; Kalyanaraman, C.; Schiffmiller, A.; Schoenbrun, E. L.; Jacobson, M. P., A Molecular Mechanics Approach to Modeling Protein–Ligand Interactions: Relative Binding Affinities in Congeneric Series. *J. Chem. Inf. Model.* **2011**, 51, 2082-9.
20. Miller, B. R.; Mcgee, T. D.; 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 8:3314-3321. *J. Chem. Theory Comput.* **2012**, 8, 3314-3321.

21. 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 (F abl). *J. Comput. Chem.* **2015**, 36, 1859-1873.
22. Sangpheak, W.; Khuntawee, W.; Wolschann, P.; Pongsawasdi, P.; Rungrotmongkol, T., Enhanced stability of a naringenin/2, 6-dimethyl β-cyclodextrin inclusion complex: Molecular dynamics and free energy calculations based on MM-and QM-PBSA/GBSA. *Journal of Molecular Graphics and Modelling* **2014**, 50, 10-15.
23. Tsitsanou, K. E.; Hayes, J. M.; Keramioti, M.; Mamais, M.; Oikonomakos, N. G.; Kato, A.; Leonidas, D. D.; Zographos, S. E., Sourcing the affinity of flavonoids for the glycogen phosphorylase inhibitor site via crystallography, kinetics and QM/MM-PBSA binding studies: comparison of chrysin and flavopiridol. *Food and chemical toxicology* **2013**, 61, 14-27.
24. Huai, Z.; Shen, Z.; Sun, Z., Binding Thermodynamics and Interaction Patterns of Inhibitor-Major Urinary Protein-I Binding from Extensive Free-Energy Calculations: Benchmarking AMBER Force Fields. *J. Chem. Inf. Model.* **2021**, 61, 284-297.
25. Huai, Z.; Yang, H.; Li, X.; Sun, Z., SAMPL7 TrimerTrip host-guest binding affinities from extensive alchemical and end-point free energy calculations. *J. Comput.-Aided Mol. Des.* **2021**, 35, 117-129.
26. Van Zundert, G.; Rodrigues, J.; Trellet, M.; Schmitz, C.; Kastritis, P.; Karaca, E.; Melquiond, A.; van Dijk, M.; De Vries, S.; Bonvin, A., The HADDOCK2. 2 web server: user-friendly integrative modeling of biomolecular complexes. *Journal of molecular biology* **2016**, 428, 720-725.
27. Krammer, A.; Kirchhoff, P. D.; Jiang, X.; Venkatachalam, C.; Waldman, M., LigScore: a novel scoring function for predicting binding affinities. *Journal of Molecular Graphics and Modelling* **2005**, 23, 395-407.
28. Wang, X.; Chong, B.; Sun, Z.; Ruan, H.; Yang, Y.; Song, P.; Liu, Z., More is simpler: Decomposition of ligand-binding affinity for proteins being disordered. *Protein Science* **2022**, 31, e4375.
29. Ahmadian, N.; Mehrnejad, F.; Amininasab, M., Molecular Insight into the Interaction between Camptothecin and Acyclic Cucurbit[4]urils as Efficient Nanocontainers in Comparison with Cucurbit[7]uril: Molecular Docking and Molecular Dynamics Simulation. *J. Chem. Inf. Model.* **2020**, 60, 1791-1803.
30. Mitkina, T.; Naumov, D. Y.; Gerasko, O.; Dolgushin, F.; Vicent, C.; Llusar, R.; Sokolov, M.; Fedin, V., Inclusion of nickel (II) and copper (II) complexes with aliphatic polyamines in cucurbit [8] uril. *Russ. Chem. Bull.* **2004**, 53, 2519-2524.
31. Samsonenko, D.; Virovets, A.; Lipkowski, J.; Geras' ko, O.; Fedin, V., Distortion of the cucurbituril molecule by an included 4-methylpyridinum cation. *J. Struct. Chem.* **2002**, 43, 664-668.
32. Sun, Z.; He, Q.; Li, X.; Zhu, Z., SAMPL6 host–guest binding affinities and binding poses from spherical-coordinates-biased simulations. *J. Comput.-Aided Mol. Des.* **2020**, 34, 589-600.
33. Litim, A.; Belhocine, Y.; Benlecheb, T.; Ghoniem, M. G.; Kabouche, Z.; Ali, F. A. M.; Abdulkhair, B. Y.; Seydou, M.; Rahali, S., DFT-D4 Insight into the Inclusion of Amphetamine and Methamphetamine in Cucurbit[7]uril: Energetic, Structural and Biosensing Properties. *Molecules* **2021**, 26, 7479.
34. Hsiao, Y.-W.; Söderhjelm, P., Prediction of SAMPL4 host–guest binding affinities using funnel metadynamics. *J. Comput.-Aided Mol. Des.* **2014**, 28, 443-454.
35. Jansook, P.; Ogawa, N.; Loftsson, T., Cyclodextrins: structure, physicochemical properties and pharmaceutical applications. *International journal of pharmaceutics* **2018**, 535, 272-284.
36. Gebhardt, J.; Kleist, C.; Jakobtorweihen, S.; Hansen, N., Validation and comparison of force fields for native cyclodextrins in aqueous solution. *J. Phys. Chem. B* **2018**, 122, 1608-1626.
37. Peerannawar, S. R.; Gejji, S. P., Theoretical investigations on vibrational spectra of pillar [5] arene-bis (pyridinium) complexes. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2013**, 104, 368-376.
38. Li, S.-H.; Zhang, H.-Y.; Xu, X.; Liu, Y., Mechanically selflocked chiral gemini-catenanes. *Nature communications* **2015**, 6, 1-7.
39. Qin, S.; Xiong, S.; Han, Y.; Hu, X. Y.; Wang, L., Controllable fabrication of various supramolecular nanostructures based on nonamphiphilic azobenzene derivatives and pillar [6] arene. *Chin. J. Chem .* **2015**, 33, 107-111.
40. Liu, L.; Cao, D.; Jin, Y.; Tao, H.; Kou, Y.; Meier, H., Efficient synthesis of copillar [5] arenes and their host–guest properties with dibromoalkanes. *Organic & Biomolecular Chemistry* **2011**, 9, 7007-7010.

41. Zhang, C.-C.; Li, S.-H.; Zhang, C.-F.; Liu, Y., Size switchable supramolecular nanoparticle based on azobenzene derivative within anionic pillar [5] arene. *Scientific reports* **2016**, 6, 1-9.
42. Xia, B.; He, J.; Abliz, Z.; Yu, Y.; Huang, F., Synthesis of a pillar [5] arene dimer by co-oligomerization and its complexation with n-octyltrimethyl ammonium hexafluorophosphate. *Tetrahedron letters* **2011**, 52, 4433-4436.
43. Yu, G.; Han, C.; Zhang, Z.; Chen, J.; Yan, X.; Zheng, B.; Liu, S.; Huang, F., Pillar [6] arene-based photoresponsive host–guest complexation. *Journal of the American Chemical Society* **2012**, 134, 8711-8717.
44. Ogoshi, T.; Yamafuji, D.; Akutsu, T.; Naito, M.; Yamagishi, T.-a., Achiral guest-induced chiroptical changes of a planar-chiral pillar [5] arene containing one π-conjugated unit. *Chem. Commun.* **2013**, 49, 8782-8784.
45. Strutt, N. L.; Zhang, H.; Schneebeli, S. T.; Stoddart, J. F., Amino-Functionalized Pillar [5] arene. *Chem. Eur. J.* **2014**, 20, 10996-11004.
46. Ma, Y.; Yang, J.; Li, J.; Chi, X.; Xue, M., A cationic water-soluble pillar [6] arene: synthesis, host–guest properties, and self-assembly with amphiphilic guests in water. *RSC advances* **2013**, 3, 23953-23956.
47. Yang, K.; Chang, Y.; Wen, J.; Lu, Y.; Pei, Y.; Cao, S.; Wang, F.; Pei, Z., Supramolecular vesicles based on complex of trp-modified pillar [5] arene and galactose derivative for synergistic and targeted drug delivery. *Chem. Mater.* **2016**, 28, 1990-1993.
48. Strutt, N. L.; Schneebeli, S. T.; Stoddart, J. F., Stereochemical inversion in difunctionalised pillar [5] arenes. *Supramol. Chem.* **2013**, 25, 596-608.
49. Dasgupta, S.; Mukherjee, P. S., Carboxylatopillar [n] arenes: a versatile class of water soluble synthetic receptors. *Organic & Biomolecular Chemistry* **2017**, 15, 762-772.
50. Gu, A.; Wheate, N. J., Macrocycles as drug-enhancing excipients in pharmaceutical formulations. *Journal of Inclusion Phenomena and Macrocyclic Chemistry* **2021**, 100, 55-69.
51. Wheate, N. J.; Dickson, K.-A.; Kim, R. R.; Nematollahi, A.; Macquart, R. B.; Kayser, V.; Yu, G.; Church, W. B.; Marsh, D. J., Host-guest complexes of carboxylated pillar [n] arenes with drugs. *Journal of Pharmaceutical Sciences* **2016**, 105, 3615-3625.
52. Li, Z.; Yang, J.; Yu, G.; He, J.; Abliz, Z.; Huang, F., Water-soluble pillar [7] arene: synthesis, pH-controlled complexation with paraquat, and application in constructing supramolecular vesicles. *Org. Lett.* **2014**, 16, 2066-2069.
53. Liu, X.; Zheng, L.; Qin, C.; Zhang, J. Z. H.; Sun, Z., Comprehensive Evaluation of End-Point Free Energy Techniques in Carboxylated-Pillar[6]arene Host-guest Binding: I. Standard Procedure. *chemrxiv* **2022**.
54. 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.
55. 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.
56. Wang, E.; Weng, G.; Sun, H.; Du, H.; Zhu, F.; Chen, F.; Wang, Z.; Hou, T., Assessing the performance of the MM/PBSA and MM/GBSA methods. 10. Impacts of enhanced sampling and variable dielectric model on protein–protein interactions. *Phys. Chem. Chem. Phys.* **2019**, 21, 18958-18969.
57. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z.; Hou, T., End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chemical reviews* **2019**, 119, 9478-9508.
58. https://github.com/samplchallenges/SAMPL9.
59. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, 23, 1623-41.
60. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1992**, 97, 10269-10280.
61. Mcweeny, R.; Diercksen, G., Self-Consistent Perturbation Theory. II. Extension to Open Shells. *J. Chem. Phys.* **1968**, 49, 4852-4856.
62. Pople, J. A.; Nesbet, R. K., Self-Consistent Orbitals for Radicals. *J. Chem. Phys.* **1954**, 22, 571-572.
63. Roothaan, C. C. J., New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, 23, 69-89.
64. Hertwig, R. H.; Koch, W., On the parameterization of the local correlation functional. What is Becke-3-LYP? *Chem.*

Phys. Lett. **1997**, 268, 345-351.
65. Becke, A. D., Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **1996**, 104, 1040-1046.
66. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, 98, 11623-11627.
67. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, 25, 1157–1173.
68. Dong, X.; Yuan, X.; Song, Z.; Wang, Q., The development of an Amber-compatible organosilane force field for drug-like small molecules. *Phys. Chem. Chem. Phys.* **2021**, 23, 12582-12591.
69. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, 79, 926-935.
70. Price, D. J.; Brooks III, C. L., A Modified TIP3P Water Potential for Simulation with Ewald Summation. *J. Chem. Phys.* **2004**, 121, 10096-10103.
71. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. J., The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, 91, 6269-6271.
72. Eberhardt, J.; Santos-Martins, D.; Tillack, A. F.; Forli, S., AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **2021**, 61, 3891-3898.
73. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, 30, 2785-2791.
74. Nguyen, N. T.; Nguyen, T. H.; Pham, T. N. H.; Huy, N. T.; Bay, M. V.; Pham, M. Q.; Nam, P. C.; Vu, V. V.; Ngo, S. T., Autodock Vina Adopts More Accurate Binding Poses but Autodock4 Forms Better Binding Affinity. *J. Chem. Inf. Model.* **2020**, 60, 204-211.
75. Gaillard, T., Evaluation of AutoDock and AutoDock Vina on the CASF-2013 Benchmark. *J. Chem. Inf. Model.* **2018**, 58, 1697-1706.
76. Joung, I. S.; Cheatham III, T. E., Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, 112, 9020-9041.
77. Joung, I. S.; Cheatham, T. E., Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-Model-Specific Ion Parameters. *J. Phys. Chem. B* **2009**, 113, 13279-13290.
78. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of The Cartesian Equations of Motion of A System with Constraints: Molecular Dynamics of n -alkanes. *J. Comput. Phys.* **1977**, 23, 327-341.
79. Miyamoto, S.; Kollman, P. A., Settle: An Analytical Version of The SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, 13, 952-962.
80. Pastor, R. W.; Brooks, B. R.; Szabo, A., An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Molecular Physics* **1988**, 65, 1409-1419.
81. Tuckerman, M. E.; Berne, B. J.; Martyna, G. J., Molecular dynamics algorithm for multiple time scales: Systems with long range forces. *J. Chem. Phys.* **1991**, 94, 6811-6815.
82. Case, D. A.; Cheatham, T. E.; Tom, D.; Holger, G.; Luo, R.; Merz, K. M.; Alexey, O.; Carlos, S.; Bing, W.; Woods, R. J., The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, 26, 1668-1688.
83. Massova, I.; Kollman, P. A., Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in drug discovery and design* **2000**, 18, 113-135.
84. Onufriev, A.; Bashford, D.; Case, D. A., Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins Structure Function & Bioinformatics* **2004**, 55, 383-94.
85. Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A., Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J. Comput. Chem.* **2004**, 25, 265-84.
86. Weiser, J.; Shenkin, P. S.; Still, W. C., Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J. Comput. Chem.* **1999**, 20, 217-230.
87. Case, D. A., Normal mode analysis of protein dynamics. *Curr. Opin. Struct. Biol.* **2010**, 4, 285–290.

88. Karplus, M.; Kushick, J. N., Method for estimating the configurational entropy of macromolecules. *Macromolecules* **1981**, 14, 325-332.
89. Sun, Z.; Yan, Y. N.; Yang, M.; Zhang, J. Z., Interaction Entropy for Protein-Protein Binding. *J. Chem. Phys.* **2017**, 146, 124124.
90. Chen, J.; Pang, L.; Wang, W.; Wang, L.; Zhang, J. Z.; Zhu, T., Decoding molecular mechanism of inhibitor bindings to CDK2 using molecular dynamics simulations and binding free energy calculations. *Journal of Biomolecular Structure and Dynamics* **2019**, 38, 985-996.
91. Zheng, L.; Yang, Y.; Bao, J.; He, L.; Qi, Y.; Zhang, J. Z. H., Discovery of novel inhibitors of CDK2 using docking and physics-based binding free energy calculation. *Chemical Biology & Drug Design* **2022**, 99, 662-673.
92. Procacci, P., Reformulating the entropic contribution in molecular docking scoring functions. *J. Comput. Chem.* **2016**, 37, 1819-1827.
93. Deng, C.-L.; Cheng, M.; Zavalij, P. Y.; Isaacs, L., Thermodynamics of pillararene·guest complexation: blinded dataset for the SAMPL9 challenge. *New J. Chem.* **2022**, 46, 995-1002.
94. Kendall, M. G., A New Measure of Rank Correlation. *Biometrika* **1938**, 30, 81-93.
95. Pearlman, D. A.; Charifson, P. S., Are Free Energy Calculations Useful in Practice? A Comparison with Rapid Scoring Functions for the p38 MAP Kinase Protein System. *Journal of Medicinal Chemistry* **2001**, 44, 3417-3423.
96. He, X.; Man, V. H.; Ji, B.; Xie, X.-Q.; Wang, J., 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.
97. 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.
98. Yan, Y.; Yang, M.; Ji, C. G.; Zhang, J. Z., Interaction entropy for computational alanine scanning. *J. Chem. Inf. Model.* **2017**, 57, 1112-1122.
99. Wang, E.; Liu, H.; Wang, J.; Weng, G.; Sun, H.; Wang, Z.; Kang, Y.; Hou, T., Development and evaluation of MM/GBSA based on a variable dielectric GB model for predicting protein–ligand binding affinities. *J. Chem. Inf. Model.* **2020**, 60, 5353-5365.
100. Simões, I. C.; Costa, I. P.; Coimbra, J. 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.
101. Liu, X.; Peng, L.; Zhang, J. Z., Accurate and Efficient Calculation of Protein–Protein Binding Free Energy-Interaction Entropy with Residue Type-Specific Dielectric Constants. *J. Chem. Inf. Model.* **2018**, 59, 272-281.
102. Genheden, S.; Ryde, U., Comparison of end-point continuum-solvation methods for the calculation of protein–ligand binding free energies. *Proteins: Structure, Function, and Bioinformatics* **2012**, 80, 1326-1342.
103. Wang, C.; Nguyen, P. H.; Pham, K.; Huynh, D.; Le, T.-B. N.; Wang, H.; Ren, P.; Luo, R., Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis. *J. Comput. Chem.* **2016**, 37, 2436-2446.

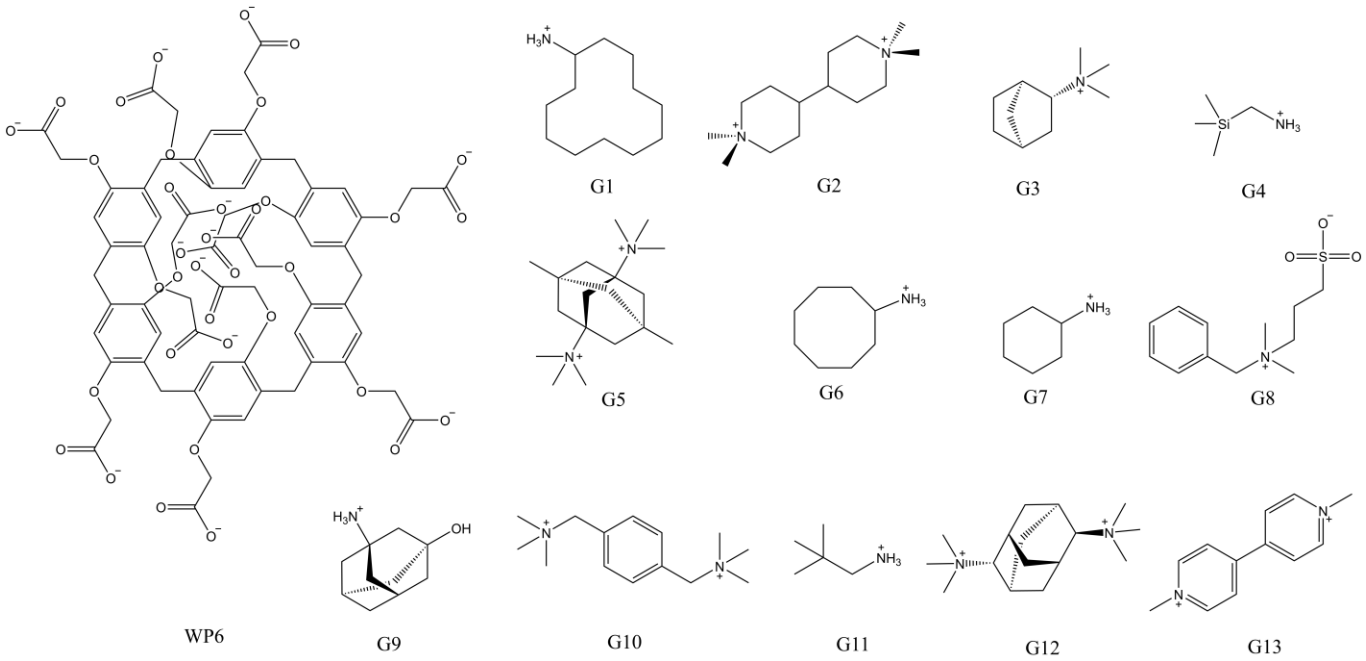

Fig. 1. 2D chemical structures of the macrocyclic host WP6 and 13 guest molecules investigated in the current work.

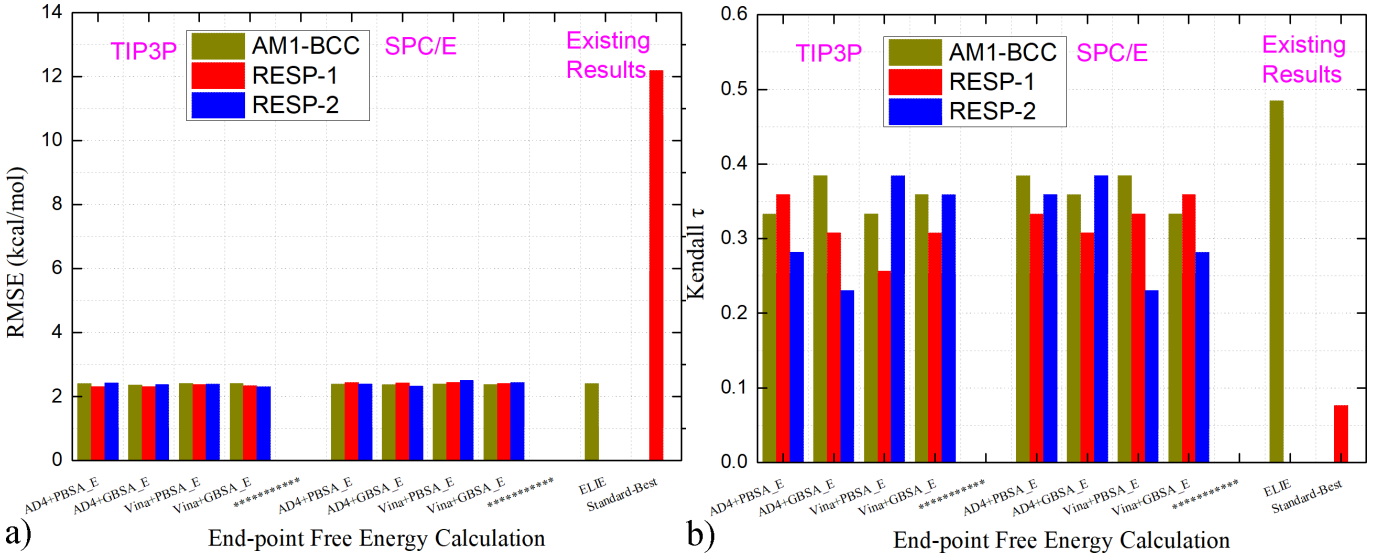

Fig. 2. Quality metrics for PBSA_E and GBSA_E calculations with different modelling parameters: a) RMSE and b) Kendall's ranking coefficient. The results obtained with a similar method named ELIE trained with system-specific data (i.e., published results of WP6 host-guest binding) and the best-performing combination of modelling parameters (AD4+RESP-1+SPC/E+MM/GBSA) are also shown for comparison. In the reproduction of absolute binding affinities, PBSA_E and GBSA_E calculations with all modelling parameters are comparable to the ELIE technique and are much better than standard end-point calculations. As for the calculation of the experimental rank of binding affinities, PBSA_E and GBSA_E are also obviously better than the standard procedure but are a bit worse than ELIE. The reason for the ELIE-better-than-PBSA_E phenomenon is rather simple. ELIE is trained with existing WP6 host-guest data published previously and thus is system-specific, while the weighting factors in our PBSA_E and GBSA_E models are trained with a huge set of protein-ligand complexes and thus are transferable.

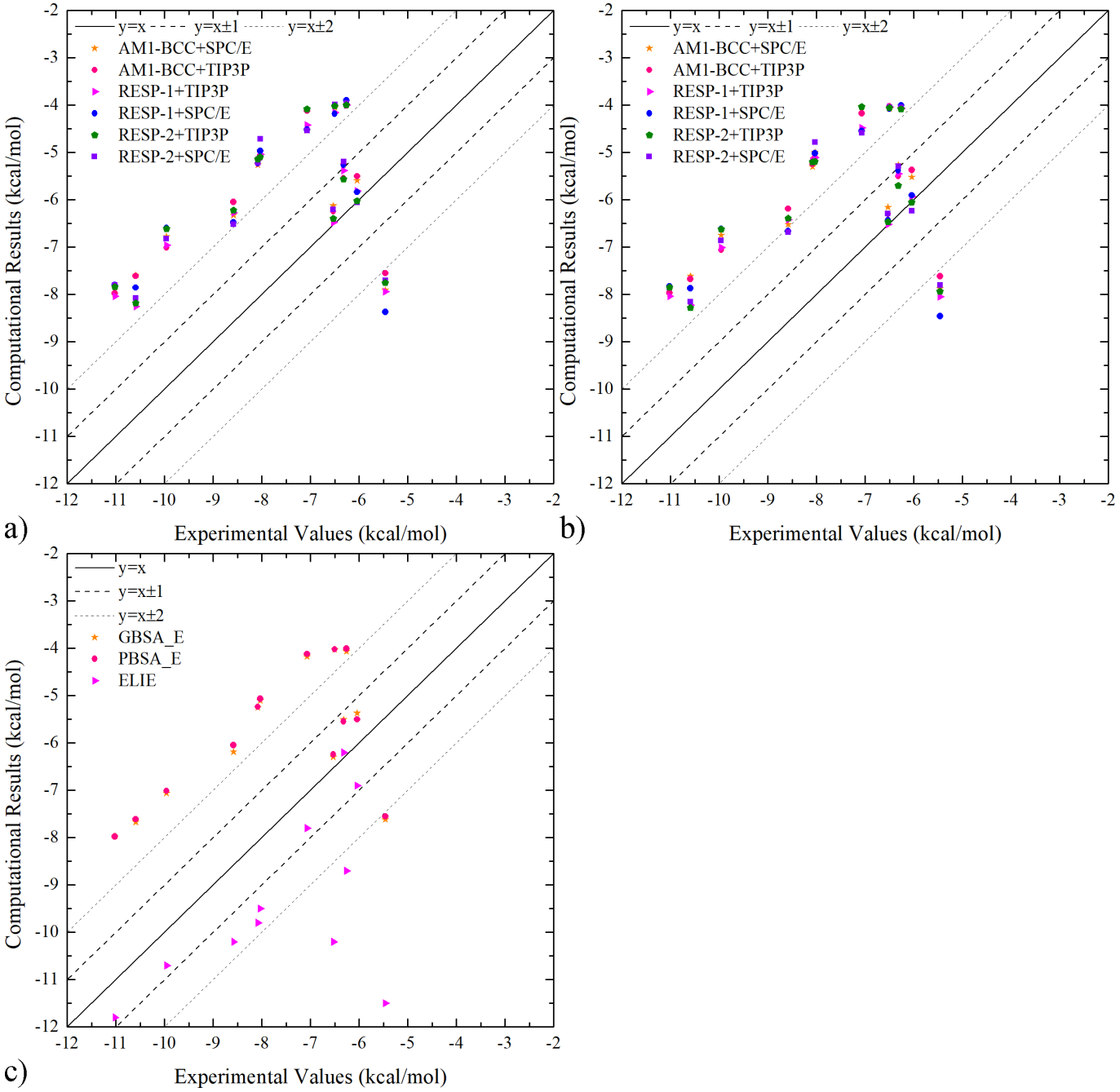

Fig. 3. Correlation between weights-altered end-point estimates and the experimental reference: a) PBSA_E and b) GBSA_E. Here, we only present the results obtained with the AD4 scoring function for clarity. The free energy estimates obtained with different charge schemes and water models are extremely similar, which is in agreement with our detailed analysis of the weighting factors of PBSA_E and GBSA_E equations. c) Comparison between our transferable estimates (PBSA_E and GBSA_E), the system-specific ELIE results and the experimental values. We can see that the PBSA_E estimates agree well with the GBSA_E results, which is expected according to the small weighting factor in the equation. The ELIE estimates differ significantly from those from the other two methods, which suggests that the noticeable differences between the weighting factors in ELIE and those in PBSA_E and GBSA_E. Thus, although the quality metrics of

these trained PBSA/GBSA methods are similar, they do follow different spirits and are expected to produce free energy outcomes with statistically significant differences.

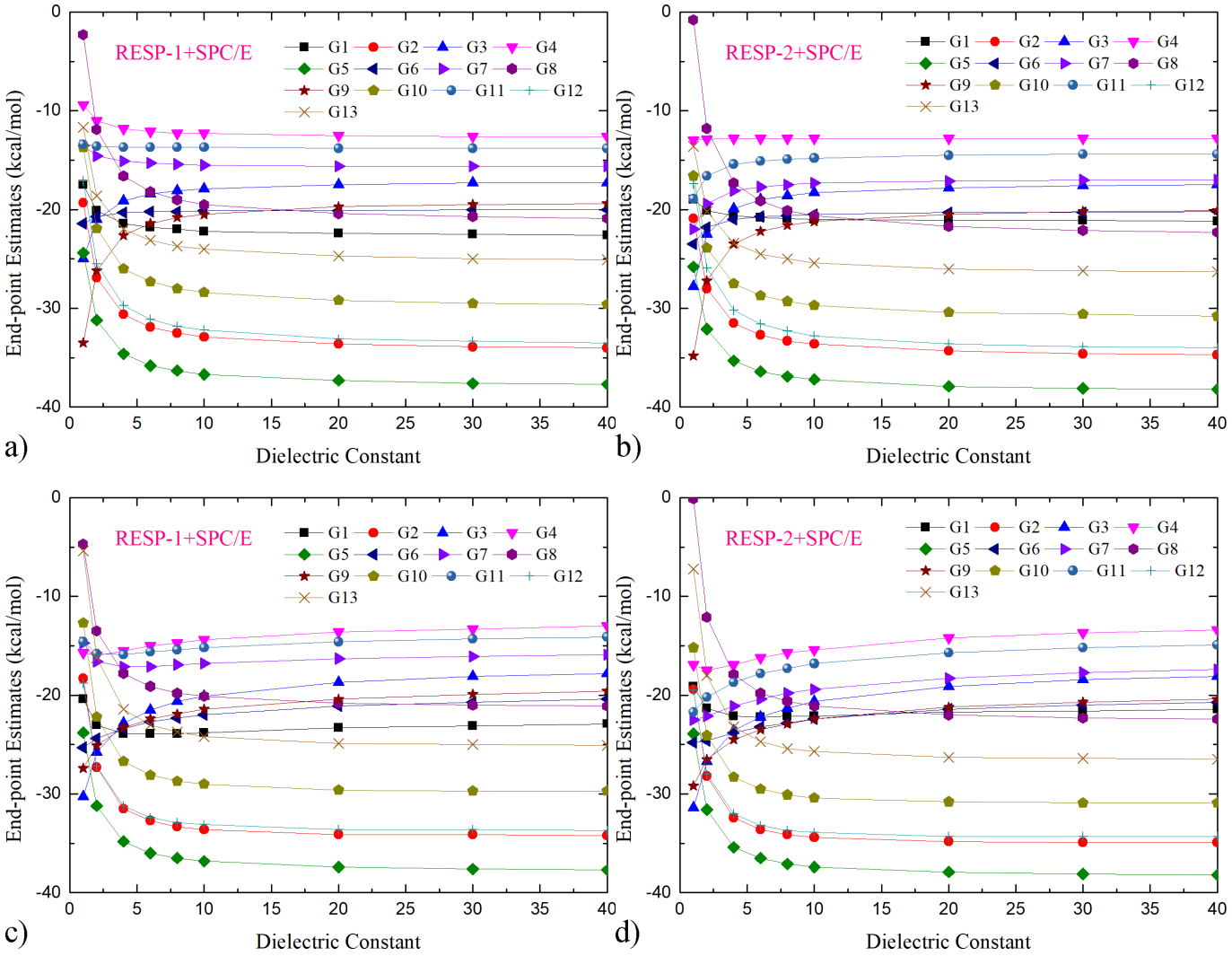

Fig. 4. Dependence of the a-b) MM/GBSA and c-d) MM/PBSA estimates on the dielectric constant for the modelling sets of AD4+RESP-1+SPC/E and AD4+RESP-2+SPC/E. The end-point estimates show significant changes in the small-ε_{in} region, while for ε_{in} larger than 10 the free energy estimates seem converged and do not vary further.

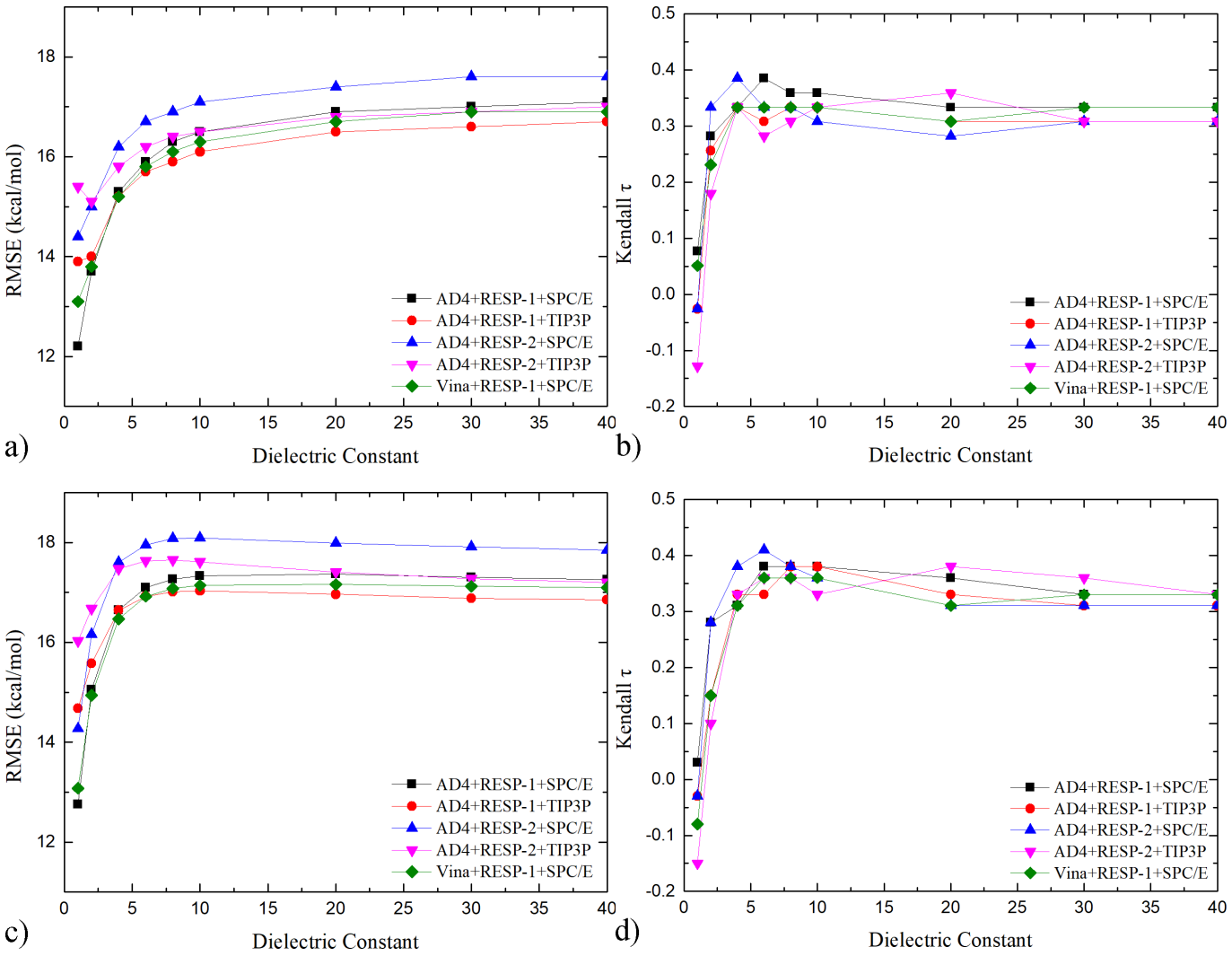

Fig. 5. Quality metrics for dielectric-constant calculations with different modelling parameters: a) RMSE and b) Kendall's ranking coefficient when the polar solvation term is estimated with the GB^{OBC} model, and c) RMSE and d) Kendall's ranking coefficient under PB solvation. With the either solvation model, the results from the standard end-point calculation are those with the interior dielectric constant of 1. The error and ranking metrics exhibit a ε_{in}-dependent behavior similar to the individual free energy estimates shown previously. The increase of the interior dielectric constant worsens the calculation of absolute binding affinities (i.e., increased RMSE), but the ranking coefficients are improved to 0.3~0.4 (comparable to PBSA_E). As long as ε_{in} reaches ~6, the rank calculation achieves a good level of accuracy. Values (ε_{in}) smaller than 6 lead to poorer rank prediction, while values larger than 6 produce predictions of similar quality and thus are not really helpful. The results with PB solvation behave similarly to the GB case.

Supporting Information:

Comprehensive Evaluation of End-Point Free Energy Techniques in Carboxylated-Pillar[6]arene Host-guest Binding: II. Regression and Dielectric Constant

Xiao Liu[1*], Lei Zheng[2], Yalong Cong[3], Zhihao Gong[4,5], Zhixiang Yin[1], John Z.H. Zhang[2,3,6,7*], Zhirong Liu[8], Zhaoxi Sun[8*]

[1]*School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620, China*

[2]*NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China*

[3]*School of Chemistry and Molecular Engineering, East China Normal University, Shanghai, 200062, China*

[4]*School of Micro-Nano Electronics, Zhejiang University, Hangzhou 310027, China*

[5]*Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou 310027, China*

[6]*Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China*

[7]*Department of Chemistry, New York University, NY, NY 10003, USA*

[8]*College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China*

Table S1. PBSA_E estimates obtained with the three charge sets and two water models with the AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements. As PBSA_E uses only a single energy-minimized structure in calculation, there is no statistical uncertainty for free energy estimates.

Host	Guest	Experiment	AM1-BCC		RESP-1		RESP-2	
			TIP3P	SPC/E	TIP3P	SPC/E	TIP3P	SPC/E
WP6	G1	-6.53	-6.2	-6.1	-6.5	-6.4	-6.4	-6.2
	G2	-10.59	-7.6	-7.6	-8.3	-7.9	-8.2	-8.1
	G3	-8.03	-5.1	-5.0	-5.0	-5.0	-5.1	-4.7
	G4	-6.50	-4.0	-4.1	-4.2	-4.2	-4.0	-4.0
	G5	-5.46	-7.6	-7.9	-7.9	-8.4	-7.8	-7.7
	G6	-8.08	-5.2	-5.3	-5.2	-5.2	-5.1	-5.2
	G7	-7.07	-4.1	-4.5	-4.4	-4.5	-4.1	-4.5
	G8	-6.04	-5.5	-5.6	-5.8	-5.8	-6.0	-6.1
	G9	-6.32	-5.5	-5.3	-5.4	-5.3	-5.6	-5.2
	G10	-9.96	-7.0	-6.8	-7.0	-6.6	-6.6	-6.8
	G11	-6.26	-4.0	-4.0	-4.0	-3.9	-4.0	-4.0
	G12	-11.02	-8.0	-8.0	-8.0	-7.8	-7.8	-7.8
	G13	-8.58	-6.0	-6.3	-6.3	-6.5	-6.2	-6.5
RMSE			2.4	2.4	2.3	2.4	2.4	2.4
MSE			-1.9	-1.8	-1.7	-1.8	-1.8	-1.8
τ			0.3	0.4	0.4	0.3	0.3	0.4
PI			0.4	0.4	0.4	0.3	0.4	0.4

Table S2. GBSA_E estimates obtained with the three charge sets and two water models with the AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements. As GBSA_E uses only a single energy-minimized structure in calculation, there is no statistical uncertainty for free energy estimates.

Host	Guest	Experiment	AM1-BCC		RESP-1		RESP-2	
			TIP3P	SPC/E	TIP3P	SPC/E	TIP3P	SPC/E
WP6	G1	-6.53	-6.3	-6.2	-6.5	-6.4	-6.5	-6.3
	G2	-10.59	-7.7	-7.6	-8.2	-7.9	-8.3	-8.2
	G3	-8.03	-5.1	-5.0	-5.1	-5.0	-5.2	-4.8
	G4	-6.50	-4.0	-4.1	-4.0	-4.1	-4.1	-4.1
	G5	-5.46	-7.6	-7.9	-8.0	-8.5	-7.9	-7.8
	G6	-8.08	-5.2	-5.3	-5.2	-5.2	-5.2	-5.2
	G7	-7.07	-4.2	-4.6	-4.5	-4.6	-4.0	-4.6
	G8	-6.04	-5.4	-5.5	-5.9	-5.9	-6.1	-6.2
	G9	-6.32	-5.5	-5.3	-5.5	-5.4	-5.7	-5.3
	G10	-9.96	-7.1	-6.7	-7.0	-6.6	-6.6	-6.9
	G11	-6.26	-4.1	-4.1	-4.1	-4.0	-4.1	-4.0
	G12	-11.02	-8.0	-8.0	-8.0	-7.8	-7.9	-7.9
	G13	-8.58	-6.2	-6.5	-6.4	-6.7	-6.4	-6.7
RMSE			2.4	2.4	2.3	2.4	2.4	2.3
MSE			-1.8	-1.8	-1.7	-1.7	-1.7	-1.7
τ			0.4	0.4	0.3	0.3	0.2	0.4
PI			0.5	0.4	0.4	0.3	0.4	0.4

Table S3. PBSA_E estimates obtained with the three charge sets and two water models with the Vina-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements. As PBSA_E uses only a single energy-minimized structure in calculation, there is no statistical uncertainty for free energy estimates.

Host	Guest	Experiment	AM1-BCC		RESP-1		RESP-2	
			TIP3P	SPC/E	TIP3P	SPC/E	TIP3P	SPC/E
WP6	G1	-6.53	-6.1	-6.6	-6.4	-6.4	-6.4	-5.9
	G2	-10.59	-7.8	-7.7	-7.8	-7.6	-7.7	-7.8
	G3	-8.03	-4.9	-4.9	-5.0	-5.0	-4.8	-4.9
	G4	-6.50	-4.5	-4.3	-4.5	-4.4	-4.3	-4.3
	G5	-5.46	-8.2	-8.0	-8.1	-8.0	-7.9	-7.9
	G6	-8.08	-5.1	-5.4	-5.4	-5.3	-5.3	-5.4
	G7	-7.07	-4.4	-4.5	-4.3	-4.3	-4.6	-4.5
	G8	-6.04	-5.6	-5.2	-5.7	-5.4	-5.1	-5.4
	G9	-6.32	-5.5	-5.5	-5.6	-5.6	-5.5	-5.4
	G10	-9.96	-6.7	-6.7	-6.8	-6.6	-6.7	-6.5
	G11	-6.26	-4.0	-4.2	-4.3	-4.1	-4.2	-4.1
	G12	-11.02	-8.0	-7.9	-7.8	-7.9	-7.9	-7.8
	G13	-8.58	-6.3	-6.3	-6.2	-5.9	-6.4	-5.3
RMSE			2.4	2.4	2.4	2.5	2.4	2.5
MSE			-1.8	-1.8	-1.7	-1.8	-1.8	-1.9
τ			0.3	0.4	0.3	0.3	0.4	0.2
PI			0.4	0.4	0.3	0.4	0.4	0.4

Table S4. GBSA_E estimates obtained with the three charge sets and two water models with the Vina-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements. As GBSA_E uses only a single energy-minimized structure in calculation, there is no statistical uncertainty for free energy estimates.

Host	Guest	Experiment	AM1-BCC		RESP-1		RESP-2	
			TIP3P	SPC/E	TIP3P	SPC/E	TIP3P	SPC/E
WP6	G1	-6.53	-6.2	-6.7	-6.5	-6.4	-6.5	-6.0
	G2	-10.59	-7.7	-7.7	-7.8	-7.6	-7.8	-7.9
	G3	-8.03	-4.9	-4.9	-5.1	-5.0	-4.9	-4.9
	G4	-6.50	-4.5	-4.3	-4.4	-4.4	-4.3	-4.3
	G5	-5.46	-8.1	-8.0	-8.2	-8.1	-8.0	-8.1
	G6	-8.08	-5.2	-5.4	-5.4	-5.4	-5.3	-5.5
	G7	-7.07	-4.5	-4.6	-4.4	-4.4	-4.7	-4.6
	G8	-6.04	-5.3	-4.9	-5.5	-5.3	-5.2	-5.4
	G9	-6.32	-5.5	-5.5	-5.7	-5.8	-5.7	-5.6
	G10	-9.96	-6.7	-6.7	-7.0	-6.6	-6.9	-6.7
	G11	-6.26	-4.1	-4.3	-4.3	-4.2	-4.3	-4.2
	G12	-11.02	-7.9	-7.9	-7.8	-7.9	-8.0	-7.8
	G13	-8.58	-6.3	-6.4	-6.3	-6.1	-6.5	-5.4
RMSE			2.4	2.4	2.3	2.4	2.3	2.5
MSE			-1.8	-1.8	-1.7	-1.8	-1.7	-1.8
τ			0.4	0.3	0.3	0.4	0.4	0.3
PI			0.4	0.4	0.4	0.3	0.4	0.4

Table S5. Dielectric-constant-variable MM/GBSA estimates obtained with the RESP-1 charge set, the SPC/E solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-17.5	-20.1	-21.4	-21.8	-22.0	-22.2	-22.4	-22.5	-22.6
	G2	-10.59	-19.3	-26.9	-30.6	-31.9	-32.5	-32.9	-33.6	-33.9	-34.0
	G3	-8.03	-25.0	-21.0	-19.1	-18.4	-18.1	-17.9	-17.5	-17.3	-17.3
	G4	-6.50	-9.4	-11.0	-11.8	-12.1	-12.3	-12.3	-12.5	-12.6	-12.6
	G5	-5.46	-24.4	-31.2	-34.6	-35.8	-36.3	-36.7	-37.3	-37.6	-37.7
	G6	-8.08	-21.4	-20.7	-20.3	-20.2	-20.2	-20.1	-20.1	-20.0	-20.0
	G7	-7.07	-13.5	-14.6	-15.1	-15.3	-15.4	-15.5	-15.6	-15.6	-15.6
	G8	-6.04	-2.3	-11.9	-16.6	-18.2	-19.0	-19.5	-20.4	-20.7	-20.9
	G9	-6.32	-33.5	-26.2	-22.6	-21.4	-20.8	-20.5	-19.7	-19.5	-19.4
	G10	-9.96	-13.7	-21.9	-26.0	-27.3	-28.0	-28.4	-29.2	-29.5	-29.6
	G11	-6.26	-13.4	-13.6	-13.7	-13.7	-13.7	-13.7	-13.8	-13.8	-13.8
	G12	-11.02	-17.1	-25.5	-29.7	-31.1	-31.8	-32.2	-33.1	-33.3	-33.5
	G13	-8.58	-11.7	-18.6	-22.0	-23.1	-23.7	-24.0	-24.7	-25.0	-25.1
RMSE			12.2	13.7	15.3	15.9	16.3	16.5	16.9	17.0	17.1
MSE			9.4	12.5	14.1	14.6	14.9	15.0	15.3	15.5	15.5
τ			0.1	0.3	0.3	0.4	0.4	0.4	0.3	0.3	0.3
PI			0.0	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3

Table S6. Dielectric-constant-variable MM/GBSA estimates obtained with the RESP-1 charge set, the TIP3P solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-18.8	-19.9	-20.5	-20.7	-20.8	-20.9	-21.0	-21.0	-21.0
	G2	-10.59	-18.9	-26.5	-30.3	-31.6	-32.2	-32.6	-33.3	-33.6	-33.7
	G3	-8.03	-24.3	-20.4	-18.5	-17.9	-17.5	-17.3	-16.9	-16.8	-16.7
	G4	-6.50	-11.7	-12.3	-12.6	-12.7	-12.7	-12.8	-12.8	-12.8	-12.8
	G5	-5.46	-24.1	-30.9	-34.4	-35.5	-36.1	-36.5	-37.1	-37.4	-37.5
	G6	-8.08	-21.3	-20.4	-19.9	-19.7	-19.6	-19.6	-19.5	-19.5	-19.5
	G7	-7.07	-20.9	-18.4	-17.1	-16.7	-16.5	-16.4	-16.1	-16.0	-16.0
	G8	-6.04	-0.9	-11.4	-16.7	-18.4	-19.3	-19.8	-20.9	-21.2	-21.4
	G9	-6.32	-40.4	-29.4	-24.0	-22.2	-21.3	-20.7	-19.6	-19.3	-19.1
	G10	-9.96	-12.5	-20.6	-24.6	-26.0	-26.6	-27.0	-27.9	-28.1	-28.3
	G11	-6.26	-14.4	-13.8	-13.5	-13.4	-13.3	-13.3	-13.2	-13.2	-13.2
	G12	-11.02	-16.9	-25.5	-29.7	-31.2	-31.9	-32.3	-33.2	-33.5	-33.6
	G13	-8.58	-10.3	-17.0	-20.4	-21.5	-22.1	-22.4	-23.1	-23.3	-23.4
RMSE			13.9	14.0	15.2	15.7	15.9	16.1	16.5	16.6	16.7
MSE			10.4	12.8	14.0	14.4	14.6	14.7	14.9	15.0	15.1
τ			0.0	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
PI			-0.1	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3

Table S7. Dielectric-constant-variable MM/GBSA estimates obtained with the RESP-2 charge set, the SPC/E solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-19.0	-20.1	-20.6	-20.8	-20.9	-21.0	-21.1	-21.1	-21.2
	G2	-10.59	-20.9	-28.0	-31.5	-32.7	-33.3	-33.6	-34.3	-34.6	-34.7
	G3	-8.03	-27.8	-22.5	-19.9	-19.0	-18.6	-18.3	-17.8	-17.6	-17.5
	G4	-6.50	-13.0	-12.9	-12.8	-12.8	-12.8	-12.8	-12.8	-12.8	-12.8
	G5	-5.46	-25.8	-32.1	-35.3	-36.4	-36.9	-37.2	-37.9	-38.1	-38.2
	G6	-8.08	-23.5	-21.8	-21.0	-20.7	-20.6	-20.5	-20.3	-20.3	-20.2
	G7	-7.07	-22.0	-19.4	-18.1	-17.7	-17.5	-17.3	-17.1	-17.0	-17.0
	G8	-6.04	-0.8	-11.8	-17.3	-19.1	-20.1	-20.6	-21.7	-22.1	-22.3
	G9	-6.32	-34.8	-27.2	-23.5	-22.2	-21.6	-21.2	-20.5	-20.2	-20.1
	G10	-9.96	-16.6	-23.9	-27.5	-28.7	-29.3	-29.7	-30.4	-30.6	-30.8
	G11	-6.26	-18.9	-16.6	-15.4	-15.1	-14.9	-14.8	-14.5	-14.4	-14.4
	G12	-11.02	-17.4	-25.9	-30.2	-31.6	-32.3	-32.8	-33.6	-33.9	-34.0
	G13	-8.58	-13.6	-20.2	-23.4	-24.5	-25.0	-25.4	-26.0	-26.2	-26.3
RMSE			14.4	15.0	16.2	16.7	16.9	17.1	17.4	17.6	17.6
MSE			11.8	14.0	15.1	15.5	15.6	15.8	16.0	16.0	16.1
τ			0.0	0.3	0.4	0.3	0.3	0.3	0.3	0.3	0.3
PI			0.0	0.3	0.3	0.4	0.4	0.4	0.4	0.4	0.4

Table S8. Dielectric-constant-variable MM/GBSA estimates obtained with the RESP-2 charge set, the TIP3P solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-21.1	-20.9	-20.8	-20.8	-20.7	-20.7	-20.7	-20.7	-20.7
	G2	-10.59	-20.7	-27.9	-31.5	-32.7	-33.3	-33.6	-34.3	-34.6	-34.7
	G3	-8.03	-28.9	-22.9	-19.9	-18.9	-18.4	-18.1	-17.5	-17.3	-17.2
	G4	-6.50	-13.6	-13.2	-13.0	-12.9	-12.8	-12.8	-12.8	-12.8	-12.8
	G5	-5.46	-25.6	-32.0	-35.2	-36.3	-36.8	-37.1	-37.8	-38.0	-38.1
	G6	-8.08	-26.9	-23.2	-21.4	-20.8	-20.5	-20.3	-19.9	-19.8	-19.7
	G7	-7.07	-21.5	-19.1	-17.9	-17.5	-17.3	-17.2	-16.9	-16.8	-16.8
	G8	-6.04	-2.8	-11.8	-16.3	-17.8	-18.5	-18.9	-19.8	-20.1	-20.3
	G9	-6.32	-36.6	-27.9	-23.6	-22.1	-21.4	-21.0	-20.1	-19.8	-19.7
	G10	-9.96	-14.9	-22.1	-25.7	-26.9	-27.5	-27.8	-28.6	-28.8	-28.9
	G11	-6.26	-23.7	-19.1	-16.8	-16.0	-15.6	-15.4	-14.9	-14.8	-14.7
	G12	-11.02	-17.7	-26.1	-30.4	-31.8	-32.5	-32.9	-33.8	-34.0	-34.2
	G13	-8.58	-10.1	-16.4	-19.6	-20.6	-21.2	-21.5	-22.1	-22.3	-22.4
RMSE			15.4	15.1	15.8	16.2	16.4	16.5	16.8	16.9	17.0
MSE			12.6	14.0	14.7	15.0	15.1	15.1	15.3	15.3	15.4
τ			-0.1	0.2	0.3	0.3	0.3	0.3	0.4	0.3	0.3
PI			-0.1	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3

Table S9. Dielectric-constant-variable MM/GBSA estimates obtained with the RESP-1 charge set, the SPC/E solvation and 100 ns unbiased sampling initiated from Vina-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-14.4	-18.4	-20.4	-21.1	-21.4	-21.6	-22.0	-22.1	-22.2
	G2	-10.59	-19.0	-26.5	-30.3	-31.6	-32.2	-32.6	-33.3	-33.6	-33.7
	G3	-8.03	-25.5	-21.0	-18.7	-18.0	-17.6	-17.4	-16.9	-16.8	-16.7
	G4	-6.50	-9.0	-10.8	-11.6	-11.9	-12.0	-12.1	-12.3	-12.4	-12.4
	G5	-5.46	-24.5	-31.3	-34.7	-35.8	-36.4	-36.7	-37.4	-37.6	-37.7
	G6	-8.08	-18.5	-19.0	-19.3	-19.4	-19.4	-19.5	-19.5	-19.5	-19.5
	G7	-7.07	-12.6	-14.6	-15.7	-16.0	-16.2	-16.3	-16.5	-16.6	-16.6
	G8	-6.04	-3.0	-12.6	-17.4	-19.0	-19.8	-20.3	-21.2	-21.5	-21.7
	G9	-6.32	-40.0	-29.3	-23.9	-22.1	-21.2	-20.7	-19.6	-19.3	-19.1
	G10	-9.96	-12.9	-21.0	-25.0	-26.3	-27.0	-27.4	-28.2	-28.5	-28.6
	G11	-6.26	-15.0	-14.3	-14.0	-13.9	-13.8	-13.8	-13.8	-13.7	-13.7
	G12	-11.02	-16.8	-25.4	-29.7	-31.2	-31.9	-32.3	-33.2	-33.5	-33.6
	G13	-8.58	-11.1	-17.9	-21.3	-22.4	-23.0	-23.3	-24.0	-24.2	-24.3
RMSE			13.1	13.8	15.2	15.8	16.1	16.3	16.7	16.9	16.9
MSE			9.4	12.4	14.0	14.5	14.7	14.9	15.2	15.3	15.3
τ			0.1	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3
PI			-0.1	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3

Table S10. Dielectric-constant-variable MM/PBSA estimates obtained with the RESP-1 charge set, the SPC/E solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-20.4	-23.0	-23.9	-23.9	-23.9	-23.8	-23.3	-23.1	-22.9
	G2	-10.59	-18.3	-27.3	-31.5	-32.7	-33.3	-33.6	-34.1	-34.1	-34.2
	G3	-8.03	-30.3	-25.8	-22.8	-21.5	-20.6	-20.1	-18.7	-18.1	-17.8
	G4	-6.50	-15.7	-16.0	-15.5	-15.0	-14.7	-14.4	-13.6	-13.3	-13.0
	G5	-5.46	-23.8	-31.2	-34.8	-36.0	-36.5	-36.8	-37.4	-37.6	-37.7
	G6	-8.08	-25.3	-24.4	-23.4	-22.7	-22.3	-22.0	-21.1	-20.7	-20.4
	G7	-7.07	-14.7	-16.6	-17.1	-17.1	-16.9	-16.8	-16.3	-16.1	-15.9
	G8	-6.04	-4.7	-13.5	-17.8	-19.1	-19.8	-20.1	-20.8	-21.0	-21.1
	G9	-6.32	-27.4	-25.1	-23.3	-22.4	-21.9	-21.4	-20.4	-19.9	-19.6
	G10	-9.96	-12.7	-22.2	-26.7	-28.1	-28.7	-29.0	-29.6	-29.7	-29.7
	G11	-6.26	-14.6	-15.8	-15.9	-15.6	-15.4	-15.2	-14.6	-14.3	-14.1
	G12	-11.02	-18.8	-27.2	-31.2	-32.4	-32.9	-33.1	-33.6	-33.6	-33.7
	G13	-8.58	-5.4	-16.2	-21.4	-23.0	-23.7	-24.2	-24.9	-25.0	-25.1
RMSE			12.8	15.1	16.7	17.1	17.3	17.3	17.4	17.3	17.3
MSE			10.1	14.1	15.8	16.1	16.2	16.2	16.0	15.9	15.8
τ			0.0	0.3	0.3	0.4	0.4	0.4	0.4	0.3	0.3
PI			0.0	0.2	0.4	0.4	0.4	0.4	0.4	0.4	0.4

Table S11. Dielectric-constant-variable MM/PBSA estimates obtained with the RESP-1 charge set, the TIP3P solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-22.6	-22.8	-22.6	-22.4	-22.2	-22.0	-21.6	-21.4	-21.3
	G2	-10.59	-18.1	-27.0	-31.2	-32.5	-33.0	-33.3	-33.7	-33.8	-33.9
	G3	-8.03	-29.1	-24.9	-22.1	-20.8	-20.0	-19.5	-18.1	-17.5	-17.2
	G4	-6.50	-19.6	-18.2	-16.9	-16.1	-15.6	-15.2	-14.1	-13.6	-13.4
	G5	-5.46	-23.8	-31.1	-34.7	-35.8	-36.3	-36.6	-37.2	-37.4	-37.5
	G6	-8.08	-24.6	-23.8	-22.8	-22.1	-21.7	-21.4	-20.5	-20.1	-19.8
	G7	-7.07	-24.4	-22.3	-20.5	-19.5	-18.9	-18.4	-17.3	-16.7	-16.5
	G8	-6.04	-2.1	-12.5	-17.6	-19.2	-20.0	-20.4	-21.2	-21.4	-21.5
	G9	-6.32	-33.7	-28.3	-24.9	-23.4	-22.5	-21.9	-20.4	-19.7	-19.4
	G10	-9.96	-11.1	-20.8	-25.4	-26.8	-27.4	-27.7	-28.2	-28.3	-28.4
	G11	-6.26	-17.0	-16.8	-16.1	-15.6	-15.2	-15.0	-14.2	-13.8	-13.6
	G12	-11.02	-18.9	-27.4	-31.4	-32.5	-33.1	-33.3	-33.7	-33.8	-33.8
	G13	-8.58	-4.5	-15.0	-20.0	-21.5	-22.2	-22.6	-23.2	-23.4	-23.5
RMSE			14.7	15.6	16.6	16.9	17.0	17.0	17.0	16.9	16.9
MSE			11.5	14.7	15.8	16.0	16.0	15.9	15.6	15.4	15.3
τ			0.0	0.2	0.3	0.3	0.4	0.4	0.3	0.3	0.3
PI			-0.1	0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3

Table S12. Dielectric-constant-variable MM/PBSA estimates obtained with the RESP-2 charge set, the SPC/E solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-19.1	-21.3	-22.1	-22.2	-22.1	-22.1	-21.7	-21.6	-21.4
	G2	-10.59	-19.4	-28.2	-32.4	-33.6	-34.1	-34.4	-34.8	-34.9	-34.9
	G3	-8.03	-31.4	-26.7	-23.6	-22.2	-21.3	-20.6	-19.1	-18.4	-18.1
	G4	-6.50	-16.9	-17.5	-16.9	-16.2	-15.7	-15.4	-14.2	-13.7	-13.4
	G5	-5.46	-23.9	-31.6	-35.4	-36.5	-37.1	-37.4	-37.9	-38.1	-38.2
	G6	-8.08	-24.8	-24.6	-23.8	-23.2	-22.8	-22.4	-21.4	-21.0	-20.7
	G7	-7.07	-22.5	-22.1	-21.1	-20.4	-19.8	-19.4	-18.3	-17.7	-17.4
	G8	-6.04	-0.1	-12.1	-17.9	-19.8	-20.6	-21.1	-22.0	-22.3	-22.4
	G9	-6.32	-29.2	-26.5	-24.5	-23.5	-22.9	-22.4	-21.2	-20.7	-20.4
	G10	-9.96	-15.2	-24.1	-28.3	-29.5	-30.1	-30.4	-30.8	-30.9	-30.9
	G11	-6.26	-21.7	-20.2	-18.7	-17.8	-17.3	-16.8	-15.7	-15.2	-14.9
	G12	-11.02	-19.5	-28.0	-32.0	-33.2	-33.7	-33.9	-34.3	-34.3	-34.3
	G13	-8.58	-7.2	-18.0	-23.2	-24.7	-25.4	-25.7	-26.3	-26.4	-26.5
RMSE			14.3	16.2	17.6	18.0	18.1	18.1	18.0	17.9	17.8
MSE			11.6	15.4	16.9	17.1	17.1	17.0	16.7	16.5	16.4
τ			0.0	0.3	0.4	0.4	0.4	0.4	0.3	0.3	0.3
PI			0.0	0.3	0.4	0.4	0.4	0.4	0.4	0.4	0.4

Table S13. Dielectric-constant-variable MM/PBSA estimates obtained with the RESP-2 charge set, the TIP3P solvation and 100 ns unbiased sampling initiated from AD4-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-20.5	-22.2	-22.6	-22.5	-22.3	-22.1	-21.6	-21.3	-21.1
	G2	-10.59	-19.1	-28.1	-32.4	-33.6	-34.2	-34.4	-34.8	-34.9	-34.9
	G3	-8.03	-32.1	-27.2	-23.8	-22.2	-21.2	-20.6	-18.9	-18.2	-17.8
	G4	-6.50	-20.6	-18.9	-17.3	-16.4	-15.8	-15.4	-14.2	-13.6	-13.3
	G5	-5.46	-23.9	-31.6	-35.3	-36.4	-37.0	-37.3	-37.8	-38.0	-38.1
	G6	-8.08	-30.5	-27.2	-24.9	-23.7	-23.0	-22.5	-21.1	-20.6	-20.2
	G7	-7.07	-22.6	-21.9	-20.8	-20.1	-19.5	-19.1	-18.1	-17.6	-17.3
	G8	-6.04	-3.0	-12.5	-17.1	-18.5	-19.2	-19.6	-20.2	-20.4	-20.5
	G9	-6.32	-31.0	-27.2	-24.7	-23.5	-22.8	-22.2	-20.9	-20.3	-20.0
	G10	-9.96	-12.9	-22.1	-26.4	-27.7	-28.2	-28.5	-29.0	-29.0	-29.1
	G11	-6.26	-28.5	-24.1	-21.0	-19.5	-18.6	-18.0	-16.4	-15.7	-15.3
	G12	-11.02	-19.6	-28.1	-32.0	-33.2	-33.7	-34.0	-34.3	-34.4	-34.4
	G13	-8.58	-4.6	-14.6	-19.4	-20.9	-21.5	-21.8	-22.4	-22.5	-22.5
RMSE			16.0	16.7	17.5	17.6	17.6	17.6	17.4	17.3	17.2
MSE			13.0	15.8	16.7	16.8	16.7	16.5	16.1	15.9	15.7
τ			-0.2	0.1	0.3	0.4	0.4	0.3	0.4	0.4	0.3
PI			-0.1	0.1	0.3	0.3	0.4	0.4	0.4	0.4	0.3

Table S14. Dielectric-constant-variable MM/PBSA estimates obtained with the RESP-1 charge set, the SPC/E solvation and 100 ns unbiased sampling initiated from Vina-produced docking poses. MSE, RMSE, τ, and PI serve as quality measurements.

Host	Guest	Experiment	Interior Dielectric Constant								
			1	2	4	6	8	10	20	30	40
WP6	G1	-6.53	-16.5	-20.1	-21.7	-22.2	-22.3	-22.4	-22.4	-22.4	-22.4
	G2	-10.59	-18.1	-27.0	-31.2	-32.4	-33.0	-33.3	-33.7	-33.8	-33.9
	G3	-8.03	-30.7	-25.9	-22.7	-21.2	-20.3	-19.7	-18.2	-17.6	-17.2
	G4	-6.50	-15.3	-15.6	-15.2	-14.8	-14.5	-14.2	-13.4	-13.0	-12.8
	G5	-5.46	-24.0	-31.3	-34.9	-36.0	-36.5	-36.8	-37.4	-37.6	-37.7
	G6	-8.08	-21.2	-21.9	-21.8	-21.5	-21.3	-21.0	-20.4	-20.1	-19.9
	G7	-7.07	-11.9	-15.8	-17.3	-17.5	-17.5	-17.5	-17.2	-17.0	-16.9
	G8	-6.04	-6.1	-14.6	-18.7	-20.1	-20.7	-21.0	-21.7	-21.8	-21.9
	G9	-6.32	-33.4	-28.1	-24.8	-23.4	-22.5	-21.9	-20.4	-19.8	-19.4
	G10	-9.96	-11.8	-21.2	-25.7	-27.1	-27.7	-28.0	-28.6	-28.7	-28.8
	G11	-6.26	-17.6	-17.4	-16.8	-16.2	-15.8	-15.5	-14.7	-14.3	-14.1
	G12	-11.02	-18.7	-27.3	-31.4	-32.6	-33.1	-33.4	-33.7	-33.8	-33.8
	G13	-8.58	-4.8	-15.5	-20.7	-22.3	-23.0	-23.4	-24.1	-24.3	-24.4
RMSE			13.1	14.9	16.5	16.9	17.1	17.1	17.2	17.1	17.1
MSE			10.0	13.9	15.6	15.9	16.0	16.0	15.8	15.7	15.6
τ			-0.1	0.2	0.3	0.4	0.4	0.4	0.3	0.3	0.3
PI			-0.1	0.1	0.3	0.3	0.3	0.3	0.3	0.3	0.3