



Published in final edited form as:

J Comput Aided Mol Des. 2017 January ; 31(1): 71–85. doi:10.1007/s10822-016-9968-2.

Absolute binding free energy calculations of CBClip host–guest systems in the SAMPL5 blind challenge

Juyong Lee¹, Florentina Tofoleanu¹, Frank C. Pickard IV¹, Gerhard König¹, Jing Huang^{1,2}, Ana Damjanovi¹, Minkyung Baek³, Chaok Seok³, and Bernard R. Brooks¹

¹Laboratory of Computational Biology, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20852, USA

²Department of Pharmaceutical Science, School of Pharmacy, University of Maryland, Baltimore, MD 20852, USA

³Department of Chemistry, Seoul National University, Seoul, Republic of Korea

Abstract

Herein, we report the absolute binding free energy calculations of CBClip complexes in the SAMPL5 blind challenge. Initial conformations of CBClip complexes were obtained using docking and molecular dynamics simulations. Free energy calculations were performed using thermodynamic integration (TI) with soft-core potentials and Bennett’s acceptance ratio (BAR) method based on a serial insertion scheme. We compared the results obtained with TI simulations with soft-core potentials and Hamiltonian replica exchange simulations with the serial insertion method combined with the BAR method. The results show that the difference between the two methods can be mainly attributed to the van der Waals free energies, suggesting that either the simulations used for TI or the simulations used for BAR, or both are not fully converged and the two sets of simulations may have sampled difference phase space regions. The penalty scores of force field parameters of the 10 guest molecules provided by CHARMM Generalized Force Field can be an indicator of the accuracy of binding free energy calculations. Among our submissions, the combination of docking and TI performed best, which yielded the root mean square deviation of 2.94 kcal/mol and an average unsigned error of 3.41 kcal/mol for the ten guest molecules. These values were best overall among all participants. However, our submissions had little correlation with experiments.

Keywords

Absolute binding free energy calculation; Hamiltonian replica exchange; Thermodynamic integration; Bennett’s acceptance ratio; Double decoupling scheme; Host-guest complexes; Constant-pH simulation

Introduction

Accurate estimations of binding free energies of small molecules to macromolecules are essential in computational structure-based drug design and discovery [1–4]. Over the past few decades, various free energy calculation methods have been suggested [5, 6] and, at the same time, force field parameters have been improved [4]. Recently, from a large-scale benchmark test, it was shown that the latest free energy simulation (FES) approach yields accurate predictions of relative binding free energies over a broad range of protein–ligands complexes (with an error less than 2 kcal/mol for most cases) [7]. This indicates that, for protein–ligand complexes, current computational methods are practical and reliable tools for drug design and lead optimization when the crystal structure of a reference complex with a known binding free energy is available.

There is another important direction of research in drug development using molecular containers (host molecules) [8] to enhance the solubility of poorly soluble drug-like molecules (guests). Although the use of host–guest complexes is being investigated extensively, the accuracy of free energy calculations of these systems have been tested less frequently than protein–ligand systems [9]. From a theoretical point of view, host–guest systems are appealing because of their relatively small size and the possibility to sample all relevant degrees of freedom.

Recently, the SAMPL blind challenge has been introduced as a community-wide platform for the comparison of various free energy calculation methods [9–36]. In SAMPL3, the binding affinities between three hosts [37, 38] (an acyclic cucurbit[6]uril (CB[6]), CB[7] and CB[8]) and nine guests were predicted by 10 different groups [10–13]. The root mean square deviation (RMSD) values obtained with molecular dynamics (MD) simulations ranged from 2.6 to 45.2 kcal/mol. Interestingly, the solvated interaction energy (SIE) method [39], which used an empirical free energy estimation function without extensive conformational sampling made predictions with the lowest RMSD value (1.4 kcal/mol), while the best molecular-dynamics-based approaches yielded RMSD of about 2.5 kcal/mol [12]. In SAMPL4, the binding affinities of CB[7] and Octa-acid complexes were predicted and the overall accuracies were similar to SAMPL3. In the previous two challenges, it was identified that no single method stood out as the top performer, indicating the need for improvement in host–guest binding affinity calculations [9].

CBClip is an acyclic CB[n]-type receptor. It has three characteristic structural features; (1) a central glycoluril oligomer induces curvature, (2) terminal aromatic walls enhance binding to drugs with aromatic rings, and (3) sulfonate arms enhance solubility [40, 41]. The Isaacs group measured the binding affinities CBClip and ten guest molecules. The structures and the protonation states of CBClip and the guest molecules are displayed in Fig. 1. All binding affinity measurements were performed using NMR in 20 mM sodium phosphate buffer at pH 7.4 at a temperature of 298 K.

In this paper, we present our absolute binding free energy calculation protocol for CBClip [40, 41] and its 10 guest molecules in the SAMPL5 challenge. The CB[n] and its analogs have been widely studied for their ability to enhance solubility and specific binding with

their target [42, 43]. In addition, CB[n] complexes are considered one of the most promising drug delivery systems because of their low toxicity [37, 41]. Thus, accurate binding free energy predictions of CB[n]-type host and guests will facilitate the design and discovery of new drug candidates delivered by this host. In addition, understanding the underlying physical mechanism of host–guest binding will provide us with insights into the highly specific mechanisms of CB[n].

Methods and materials

Outline of the binding free energy calculations of CBClip and guest molecules

The overall workflow of our SAMPL5 blind binding free energy calculations is shown in Fig. 2. First, we generated the putative binding poses of CBClip and guest molecules in two ways: (1) by slowly creating a guest molecule inside CBClip in the gas phase with a harmonic restraint to keep the guest near the center of mass of CBClip (referred to as “-MD” sets) and (2) by performing docking simulations using the GalaxyDock program (referred to as “-dock” sets), which finds the putative binding poses of protein–ligand, protein–protein and host–guest complexes via highly efficient global optimization [44–47] of the AutoDock4 scoring function [45, 48, 49]. We obtained the initial conformations for the MD simulations from scratch and did not use the conformations provided by the organizers. To obtain the initial conformations for the MD simulations, CBClip and guests were centered at the origin and rotated so that the first principle axis of inertia matches with X-axis using the “COOR ORIENT” command in CHARMM. All initial complex structures were solvated and neutralized with explicit water molecules and counter ions followed by brief molecular dynamics simulations to equilibrate. During the MD simulations, harmonic restraints were imposed to the complexes to keep the initial conformation. The last snapshots of the equilibrium runs were used as the input structures for absolute binding free energy calculations based on the double decoupling scheme [50, 51], which requires alchemical free energy calculations of the guest molecules in the bound and free states.

Free energy calculations were carried out in two ways: (1) a combination of Hamiltonian replica change (HREM) [52–54] and Bennett acceptance ratio (BAR) [11, 18, 55, 56] and (2) thermodynamic integration (TI) [57, 58]. After the binding free energy calculations, if a guest molecule has one or more titratable groups whose pKa values are in the range of experimental condition, we performed constant-pH calculations of the bound and free states to consider the effect of possible protonation states [59]. Based on the combinations of the two docking and two free energy calculation methods, we submitted four combinations of prediction schemes: ‘TI-dock’, ‘BAR-dock’, ‘TI-MD’ and ‘BAR-MD’. In addition, we submitted the set that consists of the minimum free energy predictions from the four sets: ‘All’.

Obtaining the initial bound conformations via docking simulation

With the refined CBClip force field parameters, we performed 10 ns MD simulations of the host solvated with explicit water. Using the ART-2 clustering algorithm [60] implemented in CHARMM, we clustered the sampled CBClip conformations from the MD simulation and selected the four largest clusters for following docking simulations.

For each of the four host conformations obtained from MD simulations, up to three bound conformations were selected after host–guest docking with each guest molecule. The docking method, called GalaxyDock-HG, performs global optimization of the AutoDock4 energy [61] with the conformational space annealing (CSA) algorithm [44–49, 62], as illustrated in Fig. 3. The AutoDock4 energy function consists of van der Waals energy, directional hydrogen bond energy, Coulomb electrostatic energy, and desolvation free energy as follows:

$$\begin{aligned}
 E_{AutoDock4} = & w_{vdW} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \\
 & + w_{hbond} \sum_{i,j} h(t_{ij}) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^6} \right) \\
 & + w_{qq} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \\
 & + w_{desolv} \sum_{i,j} (S_i V_j + S_j V_i) \exp \left(-\frac{r_{ij}}{2\sigma^2} \right) \quad (1)
 \end{aligned}$$

$\{A_{ij}, B_{ij}\}$ and $\{C_{ij}, D_{ij}\}$ are parameters for the van der Waals energy and hydrogen bond energy, respectively. $h(t)$ is the weight factor for hydrogen bond to describe its directionality. q_i are partial charges, and $\epsilon(r)$ is a distance-dependent dielectric constant. $\{S_i, V_i, \sigma\}$ are desolvation energy parameters. AutoDock4 energy parameters were used in the all-hydrogen topology except that the atomic partial charges of CGenFF were used.

GalaxyDock-HG was developed based on the GalaxyDock protein–ligand docking program [48] with the following modifications: (1) the potential energy was evaluated in the continuous space instead of interpolating energy values at the grid points because both host and guest molecules are treated flexibly and the host–guest systems were small enough to run global optimization efficiently in the continuous space, and (2) the initial set of conformations for CSA, called initial bank, was generated by randomly perturbing initial structures instead of using the geometry-based pre-docking method of GalaxyDock because more intensive pre-docking did not seem necessary for the small systems. More specifically, the initial structures were generated by randomly perturbing the flexible torsion angles of the host molecule first and then by randomly placing a guest molecule onto each host conformation by perturbing the values of translational, rotational, and torsional degrees of freedom for the guest molecule. A total of 100 conformations were selected as the initial bank after local energy minimization, and the bank was evolved by the CSA algorithm as implemented in GalaxyDock. The same docking run parameters were used as GalaxyDock except that the number of seed conformations used to generate trial conformations was increased from 25 to 50 because the bank size was increased from 50 to 100.

Performance of GalaxyDock-HG was tested on a set of 68 host–guest complex structures selected from those generated by Wang and Pang [63]. Successful docking results with guest-RMSD < 2 Å were obtained in 64.7 %/76.5 % of the test set complexes when top 1/top 3 complex structures were selected by clustering the final bank conformations and selecting

the lowest-energy conformations among the lowest-energy representatives of the clusters. These success rates were higher than those when the conformations were selected from the final bank conformations by energy only (64.7 %/70.6 %) or by the cluster size (52.9 %/76.5 %). Therefore, we selected up to three host–guest docking conformations for each of the initial host conformation based on energy after clustering, which resulted in 12 conformations for each guest. Among the 12 conformations, the top two scored conformations were used as initial docked conformations for the absolute binding free energy calculations.

Absolute binding free energy calculation

Binding affinities of all guest-CBClip complexes were calculated using the double-decoupling scheme [50, 51] (Fig. 4). This scheme allows one to calculate the absolute binding free energy of a given conformation of a complex. The key to obtaining well-converged results with this scheme is to restrain the relative geometry of a ligand to prevent it from being separated from the host when guest–host interactions become very weak. The thermodynamic cycle shows that the absolute binding free energy of a complex is the difference between the decoupling free energy of a ligand in the bound state and the solvation free of a ligand in the monomeric state:

$$\Delta G_{\text{bind}} = -\Delta G_{\text{rest-on}}^C - \Delta G_{\text{elec}}^C - \Delta G_{\text{vdW}}^C - \Delta G_{\text{rest-off}}^C + \Delta G_{\text{elec}}^M + \Delta G_{\text{vdW}}^M, \quad (2)$$

where $\Delta G_{\text{rest-on}}^C$ and $\Delta G_{\text{rest-off}}^C$ are the free energies of turning on and turning off restraints, ΔG_{elec}^C and ΔG_{vdW}^C are the free energies of decoupling the electrostatic and the vdW interactions of a guest in the bound state, and ΔG_{elec}^M and ΔG_{vdW}^M are those of a guest in the monomeric state.

To calculate the free energies of a complex, six geometric restraints, one distance, two angle, and three dihedral restraints, are gradually turned on between a guest and a host. With the presence of restraints, the electrostatics and vdW free energies between the guest and host are calculated by decoupling the interactions. The free energy of removing restraints $\Delta G_{\text{rest-off}}^C$ was calculated by the following formula:

$$\Delta G_{\text{rest-off}}^C = -kT \ln \frac{8\pi V (K_r K_{\vartheta_A} K_{\vartheta_B} K_{\phi_A} K_{\phi_B} K_{\phi_C})^2}{r_0^2 \sin \vartheta_{A,0} \sin \vartheta_{B,0} (2\pi K T)^3}, \quad (3)$$

where V is the volume of the unit box, r_0 is a reference distance of a distance restraint, ϑ_A and ϑ_B are the reference angles of angle restraints, K values are the force constants of the six restraints.

All simulations were performed with CHARMM 39b2 [64]. The force field parameters of all ligands were generated with the CGENFF server [65, 66]. The dihedral parameters and atomic partial charges with high penalty scores were further optimized using the standard

CGENFF procedure [67, 68]. All complex structures were solvated in a cubic box whose dimensions were $40 \times 40 \times 40 \text{ \AA}^3$ with TIP3P water molecules.

To calculate the free energy contributions of the electrostatic and van der Waals (vdW) interactions, we performed a Hamiltonian replica exchange (HREM) [52–54] simulation and post processed the results using Bennett's acceptance ratio (BAR) [11, 18, 55, 56] method. For the HREM simulation, we used 32 replicas. The electrostatic interactions were scaled down linearly using 12 replicas. The vdW interactions were gradually turned off based on the serial insertion scheme [69] (i.e., without using soft-core potentials). Twenty nonlinear scaling factors for vdW interactions were used to make exchange ratios between replicas uniform: 0.995, 0.985, 0.960, 0.930, 0.900, 0.850, 0.800, 0.750, 0.700, 0.650, 0.600, 0.550, 0.500, 0.450, 0.390, 0.330, 0.270, 0.220, 0.180, and 0.0. All HREM simulations were performed for 1 ns, which makes 32 ns in total. The free energy calculations of unbound ligands were performed using the same procedures as the complexed calculations. All simulations were repeated three times and averages and standard deviations were calculated from the three trials.

To compare the performances of different free energy calculation methods, we also performed thermodynamic integration calculations. To calculate the free energy cost of turning on restraints between the host and the ligand, we performed thermodynamic integration calculations using 19 windows with lambda values of 0.0005, 0.002, 0.004, 0.00625, 0.01, 0.01875, 0.03, 0.05, 0.075, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. The electrostatic and van der Waals free energy values were calculated using 15 windows with lambda values of 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.675, 0.725, 0.775, 0.825, 0.875, 0.925, 0.975, 1.0. For each window, equilibrium and production runs were performed for 20 and 100 ps, respectively, which amounts to 6 ns of sampling time in total for each target. The λ values were determined from a series of short TI calculations with G1 so that the variances of $\langle \frac{\partial U(\mathbf{x}, \lambda)}{\partial \lambda} \rangle_{\lambda}$ remain lower than 1 kcal/mol, which are specified by the DIFFLC variable in the log of the PERT command in CHARMM.

After the SAMPL5 submission, we performed additional longer TI calculations using the docked conformations (“-dock” set) to check the convergence of the TI calculations. For the TI-long simulations, we extended the simulation time of each step by a factor of five, which amounts to 29.5 ns of simulation time in total. Production simulations were run in the NVT ensemble using the Nose–Hoover thermostat [70–72]. All simulations were performed at temperature 300 K. Direct space nonbonded interactions were truncated with a 14.0 \AA cutoff, whereas long-range electrostatics were handled with the PME method [73]. SHAKE constraints were applied to bonds involving hydrogen, and the simulation time step was set to 1 fs [74].

Free energy methods

The Hamiltonian replica exchange method (HREM) uses multiple replicas to sample a phase space simultaneously, while replicas are swapped periodically [52]. Each replica corresponds to a different environmental condition. It is well known that proper exchanges between Hamiltonians enhance sampling efficiency while preserving the Boltzmann

distribution [53, 54, 75, 76]. In HREM, the ratio between the probability of exchanging \mathbf{x} in the m th replica and \mathbf{y} in n th replica $W(\mathbf{x}, E_m; \mathbf{y}, E_n)$ and that of the reverse process $W(\mathbf{y}, E_m; \mathbf{x}, E_n)$ must follow the relation: $\frac{W(\mathbf{x}, E_m; \mathbf{y}, E_n)}{W(\mathbf{y}, E_m; \mathbf{x}, E_n)} = \exp(-\beta \Delta)$, where $\Delta = E_m(\mathbf{y}) - E_m(\mathbf{x}) + E_n(\mathbf{x}) - E_n(\mathbf{y})$ and this condition can be satisfied if the Metropolis-type criterion is used for exchanges; if $\Delta > 0$, $W(\mathbf{x}, E_m; \mathbf{y}, E_n) = \exp(-\beta \Delta)$, otherwise $W(\mathbf{x}, E_m; \mathbf{y}, E_n) = 1$.

TI calculates the free energy difference between two states 0 and 1 by using the derivatives of the potential energy with respect to a state variable λ [6, 58], $\frac{\partial U(\mathbf{x}, \lambda)}{\partial \lambda}$. Using integration,

the free energy difference between state 0 and 1 is $\Delta G^{TI} = \int_0^1 \langle \frac{\partial U(\mathbf{x}, \lambda)}{\partial \lambda} \rangle_\lambda d\lambda$. In practice, multiple simulations are performed at different λ values to evaluate $\langle \frac{\partial U(\mathbf{x}, \lambda)}{\partial \lambda} \rangle_\lambda$. The total free energy difference is obtained from the obtained ensemble averages of the derivatives using numerical integration schemes, such as the trapezoidal rule or Simpson's rule [77, 78]. With the trapezoidal rule, the free energy difference is obtained from the following equation:

$$\Delta G = \int_0^1 g(\lambda) d\lambda = \frac{1}{2} \sum_{i=1}^{n-1} (\lambda_{i+1} - \lambda_i) (g_i + g_{i+1})$$

where $g(\lambda) = \langle \frac{\partial U(\mathbf{x}, \lambda)}{\partial \lambda} \rangle_\lambda$ and n is the total number of intermediate states between state 0 and 1. By using Simpson's rule with the generalization of Brun that allows to use non-equidistant λ values [79], the integration can be performed using the following equation:

$$\int_{\lambda_i}^{\lambda_{i+2}} g(\lambda) d\lambda = \frac{\lambda_{i+2} - \lambda_i}{6} (g_i + 4g_{i+1} + g_{i+2}) + \frac{\lambda_i - 2\lambda_{i+2} + \lambda_{i+2}}{3} (g_{i+2} - g_i).$$

In contrast to TI, BAR uses information from a pair of simulations. From the ensembles sampled obtained with two simulations, the free energy difference is calculated according to

$$\Delta G^{BAR} = -\beta^{-1} \ln \left(\frac{\langle f(U_0 - U_1 + C) \rangle_1}{\langle f(U_1 - U_0 - C) \rangle_0} \right) + C,$$

where f is the Fermi function and C is a constant that can be obtained from an iterative procedure [80, 81]. Here, for the van der Waals free energy calculations using BAR, we used the serial insertion scheme [69]. In the serial insertion scheme, van der Waals interactions are turned on or off on an "atom by atom" basis. The advantage of this scheme is that it does not require using soft-core potentials preventing from using fast energy calculation routines.

Parameterization of CBClip

The initial force field parameters of CBClip were generated using the ParamChem automation server for the CHARMM general force field (CGenFF) [65, 66]. To sample possible different conformations of apo-CBClip, we performed multiple 10 ns MD simulations with explicit water starting from the initial conformation provided by the organizers (Fig. 5B). In all MD simulations, the two aromatic walls of the CBClip opened

and formed a rather planar conformation mostly within 2 ns (Fig. 5C). In addition, with the opened conformation, the CBClip–guest complexes did not remain stable and become separated in the equilibrium MD simulations, preventing us from performing absolute binding free energy calculations. Thus, we assumed that the opened CBClip conformation is an artifact of the default CGenFF parameters.

We hypothesized that this opened conformation is formed due to the weak force constants of the dihedral angle potentials of 1-3-diazepine rings in CBClip, which were assigned high penalty scores by ParamChem, indicating the possible need to reparameterize. To validate the quality of these dihedral angle parameters, we compared the potential energy surfaces of the 1-3-diazepine rings with the CGenFF parameters and quantum mechanical (QM) calculation at the RIMP2/cc-pvdz level (Fig. 5D). The comparison shows that the initial CGenFF parameters favor the open conformation by about 2 kcal/mol (Fig. 5E). However, the QM calculation favors the closed conformation and has a higher energy barrier than the CGenFF model, suggesting more accurate modeling of the host molecule is needed for the binding affinity calculations.

We optimized the force field parameters of CBClip according to the principle of CGenFF parameterization [67, 68]. The atomic partial charges were optimized by targeting the water interaction computed at the HF/6–31 g* level with model compounds including glycoluril, hexamethyl-glycoluril, and acyclic cucurbit[2]uril. The dihedral parameters for the 1-3-diazepine ring in CBClip were optimized using a substituted glycoluril (SMILES: [C@H]12[C@@H]3NC(=O)N1Cc1c(CN2C(=O)N3)cccc1) as model compound, targeting 2D QM potential energy surfaces computed. With the optimized parameters, the closed conformation is energetically more favorable than the open conformation by 2 kcal/mol, with a barrier height of about 11 kcal/mol (Fig. 5F), suggesting this large-scale conformational change might be of importance in binding free energy calculations.

Considering multiple protonation states using constant-pH simulation

We assumed that the pH of the solvent is the same as the experimental condition pH = 7.4. The protonation states of guests were determined by comparing predicted pKa values and pH. The pKa values were predicted by using the same approach that we used for the hydration free energy calculations in SAMPL4 [18]. If the predicted pKa value of a guest differs from experimental pH by more than 2.0, we assumed that the protonation state of a guest is fixed. Otherwise, we performed constant-pH simulations to consider multiple protonation states.

For guest molecules G1, G2 and G3, we performed the constant-pH simulation of the complex to take into account of the effect pH on the binding affinity using the 2D-HREM-EDS method [59, 75, 76, 82]. The pKa values of G1 and G2 are 8.9 and 9.1, respectively. The pKa values of G3 are 7.95 and 8.82, respectively. The deprotonation free energies of G1, G2, and G3 were calculated using HREM combined with a BAR calculation.

For all constant-pH simulations, replica exchanges between EDS Hamiltonians and different pH conditions were attempted every 100 steps, alternating between exchanging EDS s values and pH values. The same nonbonded parameters were used as in case of HREM

calculations, and simulations were 1 ns long. The simulations were performed at pH 5.4, 7.4 and 9.4 and the EDS temperatures of 0, 7500, 10,000 and 12,000 were used. After the constant-pH simulations of the three complexes were finished, the pKa shifts of the three guests are calculated by fitting the deprotonated fractions to the Hill equation. By using the obtained pKa values, a binding free energy estimated with fixed charges at a given pH is adjusted by using the following equation [59]:

$$\begin{aligned}\Delta G_b(\text{pH}) &= \Delta G_b^0 - k_B T \ln \left(\frac{1 + 10^{\text{pK}_a^C - \text{pH}}}{1 + 10^{\text{pK}_a^M - \text{pH}}} \right) \\ &= \Delta G_b^+ - k_B T \ln \left(\frac{1 + 10^{\text{pH} - \text{pK}_a^C}}{1 + 10^{\text{pH} - \text{pK}_a^M}} \right),\end{aligned}\quad (4)$$

where ΔG_b^0 and ΔG_b^+ are the binding free energies obtained with deprotonated and protonated ligands.

Results and discussion

Overall performances

Among the five submissions, the TI-dock submission, which was obtained from the docking and TI calculations, shows the best agreement with the experimental data in terms of the average unsigned error (AUE), 2.94 kcal/mol, and the root mean squared deviation (RMSD) measures, 3.41 kcal/mol (Table 1). The BAR calculations ('BAR-dock') starting from the same docked conformations with the 'TI-dock' submissions resulted in higher AUE and RMSD values (3.89 kcal/mol and 4.62 kcal/mol). The submissions with initial complex conformations from MD simulations, 'TI-MD' and 'BAR-MD', resulted in worse RMSD than TI-dock. Opposite to the cases of the 'dock' submissions, 'BAR-MD' showed better results than 'TI-MD', which indicates that neither 'TI' or 'BAR' did not outperform the other consistently for the CBClip challenge. We also submitted the "All" set, which consists of the minimum value of all predictions. This set is based on the assumption that if all free energy calculations are accurate and one docked conformation has lower binding free energy than the other; the contribution of the lowest binding free energy conformation will contribute most to the final free energy. However, the RMSD of 'All' is higher than that of 'TI-dock', which indicates that the 'All' set may suffer from the overestimation of binding free energy due to the neglect of the relative free energies of the involved conformations.

To compare different protocols statistically, we estimated the error bounds of the AUE and RMSD values by performing a bootstrapping calculation [83]. We assumed that individual free energy terms, $\Delta G_{\text{rest-on}}^C$, ΔG_{elec}^C , and ΔG_{vdW}^C , obey the Gaussian distribution whose mean and standard deviation are obtained from the three repetitions. Based on the Gaussian distributions, random values were sampled 1000 times. The standard deviations of the AUE and RMSD values were calculated by comparing the random and the experimental values (Table 1). The bootstrapping analysis indicates that the results of all protocols are different from each other in a statistically significant way.

Comparison with other submissions

Our best submission, 'TI-dock', ranked as the top method among all submissions with the RMSD and AUE metrics. With the same metrics, our other submissions, 'All', 'BAR-MD', 'BAR-dock', and 'TI-MD' sets ranked the 3rd, 4th, 6th, and 7th best submissions, which shows that overall performance of our methods are better than other methods. The second best result was submitted by the movable type method devised by the Mertz group [84]. While our submissions were consistently ranked near the top by RMSD and AUE metrics, our results did not correlate very strongly with experiment. Other submissions, especially those by the Michel group, were able to rank order their predictions very well.

Dependency on initial docked conformations

The large difference between the 'dock' and the 'MD' results indicates that absolute free energy calculations using the double decoupling scheme heavily depend on initial docked conformations because the scheme restrains the relative geometry of ligand. If the native conformation is significantly different from the initial conformation, it can be only accessible during the early stages of calculating $\Delta G_{\text{rest-on}}^C$, when the restraints are weak. To avoid this problem, one may use very weak restraints. However, if the restraints are too weak, conformational sampling becomes more extensive. Moreover, a weak restraint may slow down the convergence of calculations and the calculation may become unstable due to severe steric clashes when vdW interactions are almost turned off. Thus, with weak restraints, it would be necessary to perform very long free energy simulations. On the other hand, if restraints are strong, a small conformational basin near an initial conformation will be sampled, which makes the free energy calculation converge faster. However, the contributions of conformations, which differ from the initial structure will be ignored, which biases the free energy calculation [51]. Thus, starting from an accurate docked conformation is one of the most essential steps to obtain an accurate binding free energy using the double decoupling scheme.

Comparison of TI and BAR

To obtain more details on our results, we plotted the computed and the experimental binding free energies of each guest molecule (Fig. 6). Based the docked initial conformations (Fig. 6A), the predictions for G1, G2, G4 and G7 show consistently good agreement with the experimental values regardless of the free energy calculation method. For G8, G9 and G10, the TI calculations result in good or reasonable agreement with experiment, but the BAR calculations significantly underestimate the values. For G3, both predictions made with TI and BAR overestimate the binding free energies by 6 ~8 kcal/mol. Among 10 guest molecules, BAR yielded a better prediction than TI only for G5.

When the initial docked poses are obtained with MD, the patterns of estimated binding free energies change substantially (Fig. 6B). Overall, the BAR calculations are in better agreement with experiment than the TI calculations. Both TI and BAR estimated the binding free energies of G1, G4, G5, and G9 reasonably well within an error of 3 kcal/mol. However, TI significantly underestimated the binding free energies of G6 and G10 compared to BAR, which is opposite to the 'TI-dock' and 'BAR-dock' results. The results of four ligands, G2,

G3, G7, and G8, deviate from the experimental values for both TI and BAR. Especially the results of G2 and G7 deteriorate compared to the ‘TI-dock’ and ‘BAR-dock’ results, which indicates that the predicted docked poses of these molecules may have played an important role for the accuracy of the binding free energy calculation. Similar to the ‘TI-dock’ and ‘BAR-dock’ results, the binding of G3 to CBClip was highly overestimated with TI and BAR.

For a more detailed comparison of TI and BAR, the contributions of the free energy terms used in the double decoupling scheme of the ‘TI-dock’ results are listed in Table 2. The major discrepancy between the two methods is observed for the free energy contributions of the vdW interactions in the complex and the monomeric states, ΔG_{vdW}^C and ΔG_{vdW}^M . For all complexes and monomers, BAR exhibits lower vdW free energy values than TI. The average ΔG_{vdW}^C and ΔG_{vdW}^M values estimated with BAR was smaller than those from TI by 2.7 kcal/mol and 3.5 kcal/mol, respectively. The difference between the vdW free energy values, ~ 0.83 kcal/mol, accounts for about 75 % of the average total free energy differences $\langle \Delta G_{\text{total},i}^{\text{TI}} - \Delta G_{\text{total},i}^{\text{BAR}} \rangle$, which is -1.1 kcal/mol. In contrast, the average difference between the electrostatic free energy values is -0.27 kcal/mol.

One possible cause of the difference between TI and BAR is that the numerical quadrature in the TI calculations based on the trapezoidal rule is not converged as we have observed in SAMPL3 [11, 77]. To identify whether this is the case, we re-calculated the vdW free energies from the TI calculations using Simpson’s rule (Table 3). When the TI results are integrated with Simpson’s rule, the vdW free energies changed to some degree. The average absolute differences between ΔG_{vdW}^C values integrated with the trapezoidal rule and Simpson’s rule results for long and short TI calculations are 1.4 and 1.2 kcal/mol, respectively. These results indicate that the error from numerical quadrature is observed in the longer calculations. However, the numerical quadrature error appears not to be the major source of the difference between the TI with soft-core potential calculations and the HREM–BAR with serial-insertion calculations.

Another possibility is that the TI and BAR calculations sampled different conformational spaces and were not fully converged. The major differences in the protocols for the TI and BAR calculations are the use of soft-core potentials (TI) versus serial insertion (BAR) for the van der Waals free energy calculation, and the use of enhanced sampling with HREM for the BAR calculations, which leads to sampling of different conformational spaces. We observed that the discrepancy between the TI and BAR calculations remains even for the longer TI calculations. This suggests that the TI and BAR calculations are not fully converged and the use of soft-core potential may lead to significantly different conformational space sampling, which hinders the convergence of calculations. More systematic investigations are necessary to identify the source of the difference between the convergence of conventional TI calculations with soft-core potentials and free energy calculations using the serial insertion scheme without soft-core potentials [69].

Convergence of TI calculations

The free energy components of the longer (30 ns) TI calculations are listed in Table 4 to compare the convergence of the longer and shorter TI calculations. The largest differences are observed in G3, G7 and G8 and the longer TI calculations resulted in closer agreements with experiments. The absolute errors of (G3, G7, G8) are reduced from (5.6, 1.8, 3.5) kcal/mol to (0.6, 0.3, 1.7) kcal/mol leading to the average AUE of 2.2 kcal/mol for the longer TI calculations. This result suggests that the free energy calculations of these guests were not properly converged with the short TI calculations. The slower convergences of these guests may be attributed to the fact that they have more flexible bonds than the other guests. The differences between the other guests are less or equal than 1.1 kcal/mol.

Effect of the accuracy of force field parameters

One possible source of error in calculating host–guest binding free energies is the accuracy of force field parameters. The automatic CGENFF procedure provides an estimate of the quality of the parameters with empirical penalty scores based on the chemical similarity between a given molecule and reference molecules with pre-calculated parameters. If the penalty score of a guest is correlated with the accuracy of binding free energy, one may obtain the approximate reliabilities of free energy calculations when experimental data are unavailable. To identify how the accuracy of force field parameters affects the accuracy of free energy calculations, we investigated the correlation between the penalty scores of the CGENFF parameters of guests and the average AUE values of four submissions (Fig. 7). When all results are considered, there is little correlation between two values with an R^2 value of 0.09. However, if the results of G3 and G9 are excluded, the correlation becomes much stronger; the R^2 value increases to 0.86. These results suggest that a free energy calculation of a molecule with a lower penalty score is likely to be more accurate than a calculation with a higher penalty score. However, since we investigated the correlation only with 10 guests of SAMPL5, more extensive investigations are necessary. For example, the correlation coefficient between the CGenFF penalty score and AUE for hydration free energies in SAMPL4 was 0.4.

Effect of multiple protonation states

From constant-pH simulations, we obtained the population distributions of different protonation states of G1, G2, and G3. For G1 and G2, the guests in the bound state remained fully protonated for over 99 % of the simulation time at the experimental pH of 7.4. Thus, we did not apply any corrections due to multiple protonation states. From the constant-pH simulation of the CBClip-G3 complex, it was identified that G3 was in the fully protonated state for 98.6 % and in the singly protonated state for 1.4 % of trajectory, which corresponds to a pKa shift of 9.6. Based on the constant-pH result, the total binding free energy was shifted by -0.16 kcal/mol using Eq. 2. Thus, the influence of the protonated states was rather small for the systems under study.

What went right and what went wrong?

From the four different sets of simulations, the relatively small and rigid molecules, G1, G4, and G9 were predicted accurately. The conformations of G1 obtained with docking and MD

simulations are similar to each other (Fig. 8A). The benzene ring of G1 is stacked with the two naphthalene sidewalls of CBClip in a parallel fashion. The two ligand conformations of G4 are the mirror images of each other (Fig. 8B). Thus, they adopt similar interaction patterns with CBClip and resulted in similar binding free energies. However, the two docked conformations of G9 are significantly different from each other although their calculated binding free energies are similar (Fig. 8C). The docking simulation placed G9 between the two sidewalls of CBClip in a parallel fashion. However, G9 is dangling around a sulfonate group via electrostatic interaction. The similar binding free energies of the two conformations suggest that the G9-CBClip complex may adopt diverse docking poses. The sulfonate arms of CBClip may play a role not only in enhancing solubility but also in attracting ligands [40].

The binding free energies of G3, G8, and G10, which are bulky and flexible molecules with many rotatable bonds, were poorly calculated. From both docking and MD simulations, G3 binds to one of the sulfonate arms of CBClip via electrostatic interactions (Fig. 8D). The overestimation of binding free energy of G3 indicates that the electrostatic interactions are significantly exaggerated. On the other hand, G8 is weakly bound to CBClip via interactions with the sulfonate groups and does not form a stable complex (Fig. 8E). The significantly underestimated binding free energy of G8 suggests that the docked conformations are incorrect and more extensive conformational sampling is necessary. The aromatic groups of G10 are stacked in between two sidewalls in the two docked conformations, which is similar to the G9-CBClip complex conformation obtained with the docking simulation (Fig. 8F). When the individual binding free energy components of G9 and G10 are compared, the electrostatic free energies in the complex state (ΔG_{elec}^C) differ the most, which suggests that electrostatic interactions between G10 and CBClip may be underestimated.

Conclusion

The absolute binding free energies of ten CBClip–guest complexes were calculated using the double decoupling method. The initial bound poses were obtained in two ways, using docking and MD simulations. For each initial pose, free energy calculations were carried out using both TI with soft-core potentials and BAR with the serial insertion scheme, which resulted in four different sets of simulations: TI-dock, BAR-dock, TI-MD, BAR-MD, and ALL. TI-dock resulted in the best prediction with an AUE of 2.94 kcal/mol. The AUE of BAR-dock is higher than that of TI-dock by 0.95 kcal/mol although they used the same set of initial docked poses. The error is higher than those of previous SAMPL predictions where the symmetric CB[7] and octa-acid were used as hosts. This suggests that the CBClip host used in SAMPL5 is a more challenging target.

A comparison of long and short TI calculations showed that the short TI calculations of relatively more flexible guest, G3, G7, and G8, were not properly converged and the long TI calculations improved the predictions of those guests. A comparison of the TI calculations with soft-core potentials and HREM–BAR calculations with the serial-insertion scheme showed that all van der Waals free energy calculations were not fully converged even with the small sizes of CBClip–guest complexes. The difference between TI and BAR might be attributed to whether soft-core potentials were used or not for van der Waals free energy

calculations, which might have led to sampling different phase space regions. We also performed the constant-pH simulations of G1, G2, and G3 to consider the effect of pH. However, the free energy contributions of multiple protonation states were small.

Our results suggest that another major source of error in free energy calculations lies in the accuracy of force field parameters. We identified that the penalty score of the force field parameters of the guest molecules and the accuracy of the calculated free energies are correlated. In addition, a careful parameter optimization of CBClip based on the quantum mechanical calculation played an important role in accurate binding free energy calculations. In summary, there is room for improvement in developing better force field parameter fitting schemes and incorporating more sophisticated models, e.g. polarizable force fields for better binding free energy calculations.

Acknowledgments

The authors would like to thank R. Pastor for stimulating discussions. The research was supported by the Intramural Research Program of the NIH, NHLBI. Computational resources and services used in this work were provided by the LoBoS cluster of the National Institutes of Health.

References

1. Liao C, Sitzmann M, Pugliese A, Nicklaus MC. Software and resources for computational medicinal chemistry. *Future Med Chem.* 2011; 3:1057–1085. [PubMed: 21707404]
2. 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. [PubMed: 26588302]
3. Sliwoski G, Kothiwale S, Meiler J, Lowe EW. Computational methods in drug discovery. *Pharmacol Rev.* 2014; 66:334–395. [PubMed: 24381236]
4. Shirts MR, Mobley DL, Brown SP. Free-energy calculations in structure-based drug design. *Drug Des Struct Ligand Based Approaches.* 2010; doi: 10.1017/CBO9780511730412.007
5. Kollman P. Free energy calculations: applications to chemical and biochemical phenomena. *Chem Rev.* 1993; 93:2395–2417.
6. Chipot, C.; Pohorille, A. Free energy calculations. Springer; Berlin: 2007.
7. Wang L, 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; doi: 10.1021/ja512751q
8. Barrow SJ, Kasera S, Rowland MJ, Del Barrio J, Scherman OA. Cucurbituril-based molecular recognition. *Chem Rev.* 2015; 115:12320–12406. [PubMed: 26566008]
9. Muddana HS, Fenley AT, Mobley DL, Gilson MK. The SAMPL4 host–guest blind prediction challenge: an overview. *J Comput Aided Mol Des.* 2014; 28:305–317. [PubMed: 24599514]
10. Gallicchio E, Levy RM. Prediction of SAMPL3 host–guest affinities with the binding energy distribution analysis method (BEDAM). *J Comput Aided Mol Des.* 2012; 26:505–516. [PubMed: 22354755]
11. König G, Brooks BR. Predicting binding affinities of host–guest systems in the SAMPL3 blind challenge: the performance of relative free energy calculations. *J Comput Aided Mol Des.* 2012; 26:543–550. [PubMed: 22198474]
12. Muddana HS, et al. Blind prediction of host–guest binding affinities: a new SAMPL3 challenge. *J Comput Aided Mol Des.* 2012; 26:475–487. [PubMed: 22366955]
13. Muddana HS, Gilson MK. Prediction of SAMPL3 host–guest binding affinities: evaluating the accuracy of generalized force-fields. *J Comput Aided Mol Des.* 2012; 26:517–525. [PubMed: 22274835]

14. Geballe MT, Guthrie JP. The SAMPL3 blind prediction challenge: transfer energy overview. *J Comput Aided Mol Des.* 2012; 26:489–496. [PubMed: 22476552]
15. Reinisch J, Klamt A, Diederhufen M. Prediction of free energies of hydration with COSMO-RS on the SAMPL3 data set. *J Comput Aided Mol Des.* 2012; 26:669–673. [PubMed: 22581451]
16. Kulp JL III, Blumenthal SN, Wang Q, Bryan RL, Guarnieri F. A fragment-based approach to the SAMPL3 Challenge. *J Comput Aided Mol Des.* 2012; 26:583–594. [PubMed: 22290624]
17. Kumar A, Zhang KYJ. Computational fragment-based screening using RosettaLigand: the SAMPL3 challenge. *J Comput Aided Mol Des.* 2012; 26:603–616. [PubMed: 22246345]
18. König G, Pickard FC, Mei Y, Brooks BR. Predicting hydration free energies with a hybrid QM/MM approach: an evaluation of implicit and explicit solvation models in SAMPL4. *J Comput Aided Mol Des.* 2014; 28:245–257. [PubMed: 24504703]
19. Hsiao Y-W, Söderhjelm P. Prediction of SAMPL4 host–guest binding affinities using funnel metadynamics. *J Comput Aided Mol Des.* 2014; 28:443–454. [PubMed: 24535628]
20. Monroe JI, Shirts MR. Converging free energies of binding in cucurbit [7] uril and octa-acid host–guest systems from SAMPL4 using expanded ensemble simulations. *J Comput Aided Mol Des.* 2014; 28:401–415. [PubMed: 24610238]
21. Muddana HS, Yin J, Sapra NV, Fenley AT, Gilson MK. Blind prediction of SAMPL4 cucurbit [7] uril binding affinities with the mining minima method. *J Comput Aided Mol Des.* 2014; 28:463–474. [PubMed: 24510191]
22. Ellingson BA, et al. Efficient calculation of SAMPL4 hydration free energies using OMEGA, SZYBKI, QUACPAC, and Zap TK. *J Comput Aided Mol Des.* 2014; 28:289–298. [PubMed: 24633516]
23. Manzoni F, Söderhjelm P. Prediction of hydration free energies for the SAMPL4 data set with the AMOEBA polarizable force field. *J Comput Aided Mol Des.* 2014; 28:235–244. [PubMed: 24577872]
24. Fu J, Liu Y, Wu J. Fast prediction of hydration free energies for SAMPL4 blind test from a classical density functional theory. *J Comput Aided Mol Des.* 2014; 28:299–304. [PubMed: 24622881]
25. Li L, Dill KA, Fennell CJ. Testing the semi-explicit assembly model of aqueous solvation in the SAMPL4 challenge. *J Comput Aided Mol Des.* 2014; 28:259–264. [PubMed: 24474161]
26. Gallicchio E, et al. BEDAM binding free energy predictions for the SAMPL4 octa-acid host challenge. *J Comput Aided Mol Des.* 2015; 29:315–325. [PubMed: 25726024]
27. Beckstein O, Fourrier A, Iorga BI. Prediction of hydration free energies for the SAMPL4 diverse set of compounds using molecular dynamics simulations with the OPLS-AA force field. *J Comput Aided Mol Des.* 2014; 28:265–276. [PubMed: 24557853]
28. Park H. Extended solvent-contact model approach to SAMPL4 blind prediction challenge for hydration free energies. *J Comput Aided Mol Des.* 2014; 28:175–186. [PubMed: 24554191]
29. Mikulskis P, et al. Free-energy perturbation and quantum mechanical study of SAMPL4 octa-acid host–guest binding energies. *J Comput Aided Mol Des.* 2014; 28:375–400. [PubMed: 24700414]
30. Sure R, Antony J, Grimme S. Blind prediction of binding affinities for charged supramolecular host–guest systems: achievements and shortcomings of DFT-D3. *J Phys Chem B.* 2014; 118:3431–3440. [PubMed: 24588346]
31. Mobley DL, Wymer KL, Lim NM, Guthrie JP. Blind prediction of solvation free energies from the SAMPL4 challenge. *J Comput Aided Mol Des.* 2014; 28:135–150. [PubMed: 24615156]
32. Gallicchio E, et al. Virtual screening of integrase inhibitors by large scale binding free energy calculations: the SAMPL4 challenge. *J Comput Aided Mol Des.* 2014; 28:475–490. [PubMed: 24504704]
33. Mobley DL, et al. Blind prediction of HIV integrase binding from the SAMPL4 challenge. *J Comput Aided Mol Des.* 2014; 28:327–345. [PubMed: 24595873]
34. Perryman AL, Santiago DN, Forli S, Santos-Martins D, Olson AJ. Virtual screening with AutoDock Vina and the common pharmacophore engine of a low diversity library of fragments and hits against the three allosteric sites of HIV integrase: participation in the SAMPL4 protein–ligand binding challenge. *J Comput Aided Mol Des.* 2014; 28:429–441. [PubMed: 24493410]

35. Hogues H, Sulea T, Purisima EO. Exhaustive docking and solvated interaction energy scoring: lessons learned from the SAMPL4 challenge. *J Comput Aided Mol Des.* 2014; 28:417–427. [PubMed: 24474162]
36. Sandberg L. Predicting hydration free energies with chemical accuracy: the SAMPL4 challenge. *J Comput Aided Mol Des.* 2014; 28:211–219. [PubMed: 24550133]
37. Ma D, Zavalij PY, Isaacs L. Acyclic cucurbit[n]uril con-geners are high affinity hosts. *J Org Chem.* 2010; 75:4786–4795. [PubMed: 20540586]
38. Biedermann F, et al. Benzobis(imidazolium)-cucur-bit[8]uril complexes for binding and sensing aromatic compounds in aqueous solution. *Chem A Eur J.* 2010; 16:13716–13722.
39. Naïm M, 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. [PubMed: 17238257]
40. Zhang B, Isaacs L. Acyclic cucurbit[n]uril-type molecular containers: influence of aromatic walls on their function as solubilizing excipients for insoluble drugs. *J Med Chem.* 2014; 57:9554–9563. [PubMed: 25369565]
41. Gilberg L, Zhang B, Zavalij PY, Sindelar V, Isaacs L. Acyclic cucurbit[n]uril-type molecular containers: influence of glycoluril oligomer length on their function as solubilizing agents. *Org Biomol Chem.* 2015; 13:4041–4050. [PubMed: 25731639]
42. Lee JW, Samal S, Selvapalam N, Kim H-J, Kim K. Cucurbituril homologues and derivatives: new opportunities in supramolecular chemistry. *Acc Chem Res.* 2003; 36:621–630. [PubMed: 12924959]
43. Masson E, Ling X, Joseph R, Kyeremeh-Mensah L, Lu X. Cucurbituril chemistry: a tale of supramolecular success. *RSC Adv.* 2012; 2(4):1213–1247.
44. Lee J, Scheraga HA, Rackovsky S. New optimization method for conformational energy calculations on polypeptides: conformational space annealing. *J Comput Chem.* 1997; 18:1222–1232.
45. Shin W-H, et al. LigDockCSA: protein–ligand docking using conformational space annealing. *J Comput Chem.* 2011; 32:3226–3232. [PubMed: 21837636]
46. Lee J, et al. De novo protein structure prediction by dynamic fragment assembly and conformational space annealing. *Proteins Struct Funct Bioinform.* 2011; 79:2403–2417.
47. Lee J, Gross SP, Lee J. Modularity optimization by conformational space annealing. *Phys Rev E.* 2012; 85:056702.
48. Shin WH, Kim JK, Kim DS, Seok C. GalaxyDock2: protein–ligand docking using beta-complex and global optimization. *J Comput Chem.* 2013; 34:2647–2656. [PubMed: 24108416]
49. Shin W-H, Lee GR, Seok C. Evaluation of GalaxyDock based on the community structure—activity resource 2013 and 2014 benchmark studies. *J Chem Inf Model.* 2015; doi: 10.1021/acs.jcim.5b00309
50. Gilson MK, Given JA, Bush BL, McCammon JA. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys J.* 1997; 72:1047–1069. [PubMed: 9138555]
51. Boresch S, et al. Absolute binding free energies: a quantitative approach for their calculation. *J Phys Chem B.* 2003; 107(35):9535–9551.
52. Fukunishi H, Watanabe O, Takada S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction. *J Chem Phys.* 2002; 116:9058.
53. Itoh SG, Okumura H. Hamiltonian replica-permutation method and its applications to an alanine dipeptide and amyloid- β (29–42) peptides. *J Comput Chem.* 2013; 34:2493–2497. [PubMed: 23925979]
54. Itoh SSG, Okumura H, Okamoto Y. Replica-exchange method in van der Waals radius space: overcoming steric restrictions for biomolecules. *J Chem Phys.* 2010; 132:134105. [PubMed: 20387919]
55. Bennett CH. Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys.* 1976; 22:245–268.
56. König G, Hudson PS, Boresch S, Woodcock HL. Multiscale free energy simulations: an efficient method for connecting classical MD simulations to QM or QM/MM free energies using Non-

- Boltzmann Bennett reweighting schemes. *J Chem Theory Comput.* 2014; 10:1406–1419. [PubMed: 24803863]
57. Straatsma TP, Berendsen HJ, Postma JPM, Berendsen C, Postma PM. Free energy of hydrophobic hydration: a molecular dynamics study of noble gases in water. *J Chem Phys.* 1986; 85:6720–6727.
58. Straatsma TP, Berendsen HJC. Free energy of ionic hydration: analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics simulations. *J Chem Phys.* 1988; 89:5876.
59. Lee J, Miller BT, Brooks BR. Computational scheme for pH-dependent binding free energy calculation with explicit solvent. *Protein Sci.* 2016; 25:231–243. [PubMed: 26189656]
60. Karpen ME, Tobias DJ, Brooks CL. Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of YPGDV. *Biochemistry.* 1993; 32:412–420. [PubMed: 8422350]
61. Huey R, Morris GM, Olson AJ, Goodsell DS. A semiempirical free energy force field with charge-based desolvation. *J Comput Chem.* 2007; 28:1145–1152. [PubMed: 17274016]
62. Lee J, Lee I-H, Lee J. Unbiased global optimization of Lennard-Jones clusters for $N < \text{or} = 201$ using the conformational space annealing method. *Phys Rev Lett.* 2003; 91:080201. [PubMed: 14525223]
63. Wang Q, Pang YP. Accurate reproduction of 161 small-molecule complex crystal structures using the EUDOC program: expanding the use of EUDOC to supramolecular chemistry. *PLoS One.* 2007; 2(6):e531. [PubMed: 17565384]
64. Brooks BR, et al. CHARMM: the biomolecular simulation program. *J Comput Chem.* 2009; 30:1545–1614. [PubMed: 19444816]
65. Vanommeslaeghe K, et al. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem.* 2010; 31:671–690. [PubMed: 19575467]
66. Yu W, He X, Vanommeslaeghe K, MacKerell AD. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J Comput Chem.* 2012; 33:2451–2468. [PubMed: 22821581]
67. Vanommeslaeghe K, MacKerell AD. Automation of the CHARMM general force field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model.* 2012; 52:3144–3154. [PubMed: 23146088]
68. Vanommeslaeghe K, Raman EP, MacKerell AD. Automation of the CHARMM general force field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model.* 2012; 52:3155–3168. [PubMed: 23145473]
69. Borech S, Bruckner S. Avoiding the van der Waals endpoint problem using serial atomic insertion. *J Comput Chem.* 2011; 32:2449–2458. [PubMed: 21607991]
70. Nose S, Nosé S. A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys.* 1984; 81:511.
71. Hoover W. Canonical dynamics: equilibrium phase-space distributions. *Phys Rev A.* 1985; 31:1695–1697.
72. Martyna GJ, Klein ML. Nose–Hoover chains: the canonical ensemble via continuous dynamics. *J Chem Phys.* 1992; 97(4):2635.
73. Darden T, York D, Pedersen L. Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J Chem Phys.* 1993; 98:10089.
74. Ryckaert JP, Ciccotti G, Berendsen HJC. 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.
75. Lee J, Miller BT, Damjanovi A, Brooks BR. Constant pH molecular dynamics in explicit solvent with enveloping distribution sampling and hamiltonian exchange. *J Chem Theory Comput.* 2014; 10:2738–2750. [PubMed: 25061443]
76. Lee J, Miller BT, Damjanovic A, Brooks BR. Enhancing constant-pH simulation in explicit solvent with a two-dimensional replica exchange method. *J Chem Theory Comput.* 2015; 11:2560–2574. [PubMed: 26575555]

77. De Ruiter A, Boresch S, Oostenbrink C. Comparison of thermodynamic integration and Bennett's acceptance ratio for calculating relative protein–Ligand binding free energies. *J Comput Chem.* 2013; 34:1024–1034. [PubMed: 23335287]
78. Bruckner S, Boresch S. Efficiency of Alchemical free energy simulations. II. improvements for thermodynamic integration. *J Comput Chem.* 2011; 32:1320–1333. [PubMed: 21425289]
79. Brun V. A generalization of the formula of Simpson for non-equidistant ordinates. *Nord Mat Tidskr.* 1953; 1:10–15.
80. König G, Bruckner S, Boresch S. Unorthodox uses of Bennett's acceptance ratio method. *J Comput Chem.* 2009; 30:1712–1718. [PubMed: 19373838]
81. König G, Boresch S. Non-Boltzmann sampling and Bennett's acceptance ratio method: how to profit from bending the rules. *J Comput Chem.* 2011; 32:1082–1090. [PubMed: 21387335]
82. König G, Miller BT, Boresch S, Wu X, Brooks BR. Enhanced sampling in free energy calculations: combining SGLD with the Bennett's acceptance ratio and enveloping distribution sampling methods. *J Chem Theory Comput.* 2012; 8:3650–3662. [PubMed: 26593010]
83. Mooney, CZ.; Duval, RD.; Duval, R. *Bootstrapping: a non-parametric approach to statistical inference.* Sage; NY: 1993.
84. Zheng Z, Ucisik MN, Merz KM. The movable type method applied to protein–ligand binding. *J Chem Theory Comput.* 2013; 9:5526–5538. [PubMed: 24535920]

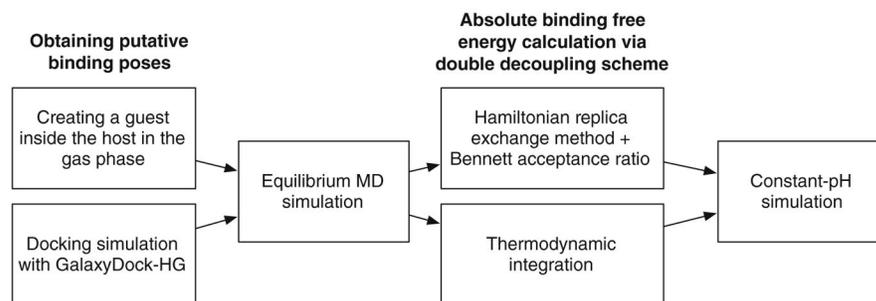


Fig. 1. Overall workflow of the binding free energy calculations of the CBClip–guest complexes in SAMPL5

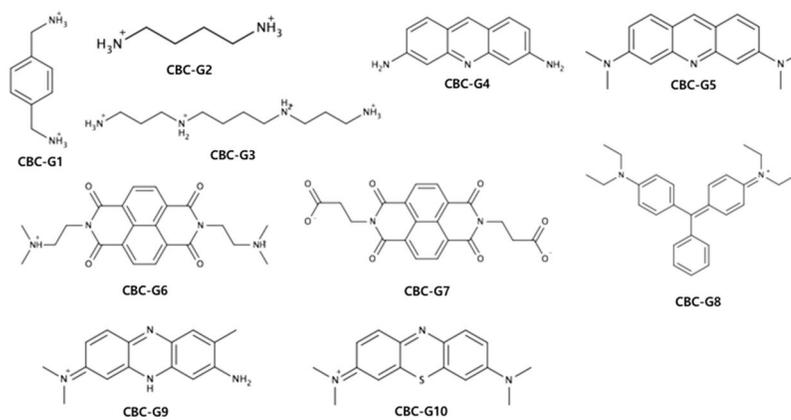


Fig. 2.
The structures and protonation states of guest molecules in this study

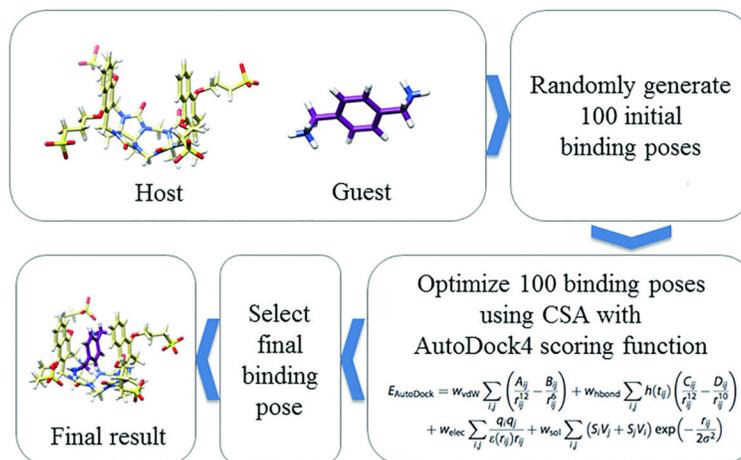


Fig. 3.
A schematic description of the GalaxyDock-HG host-guest docking method

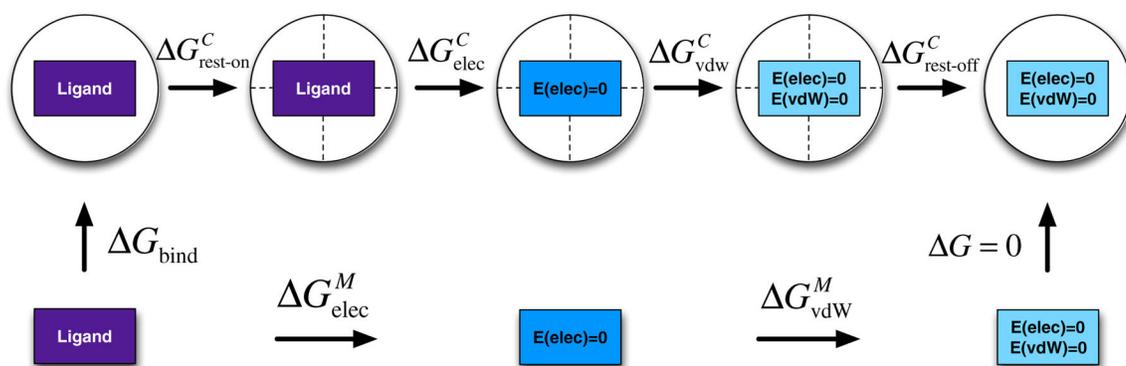


Fig. 4.
The thermodynamic cycle of the double-decoupling scheme for absolute binding free energy calculation

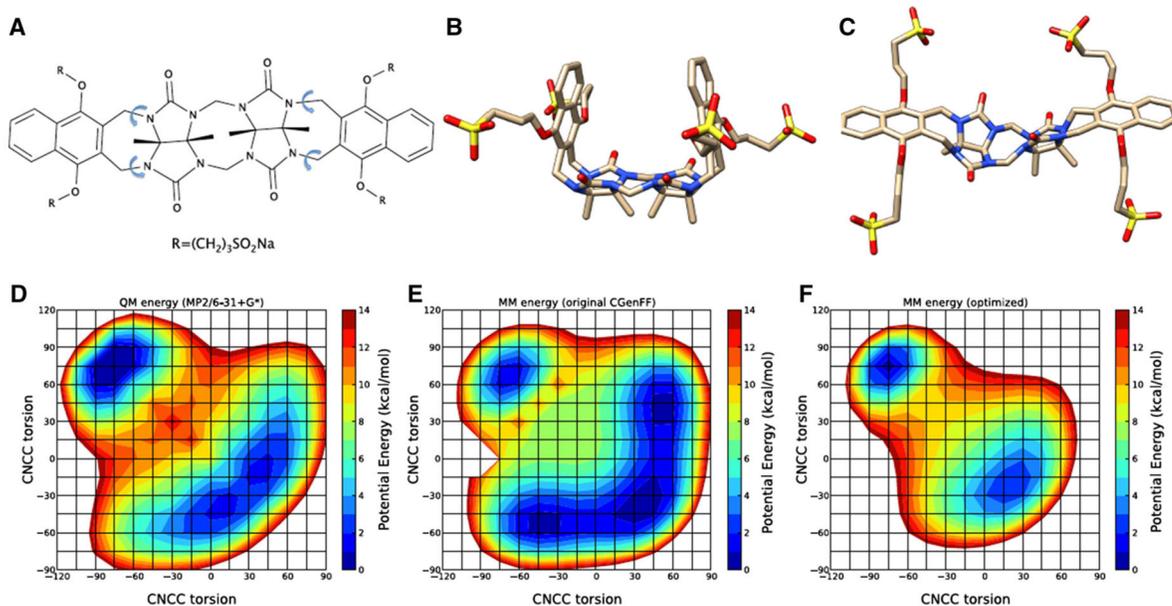


Fig. 5. The 2D and 3D structures of CBClip and potential energy surfaces of CBClip. **A** The chemical structure of CBClip is shown and four parameter optimized dihedral angles, which were chosen for parameter optimization are highlighted with *arrows*. The initial (**B**) and the open (**C**) conformations of CBClip after the equilibrium MD are shown. The potential energy surfaces of the 1-3-diazepin ring calculated with QM (**D**), the original CGENFF (**E**), and the optimized CGENFF (**F**) parameters are shown

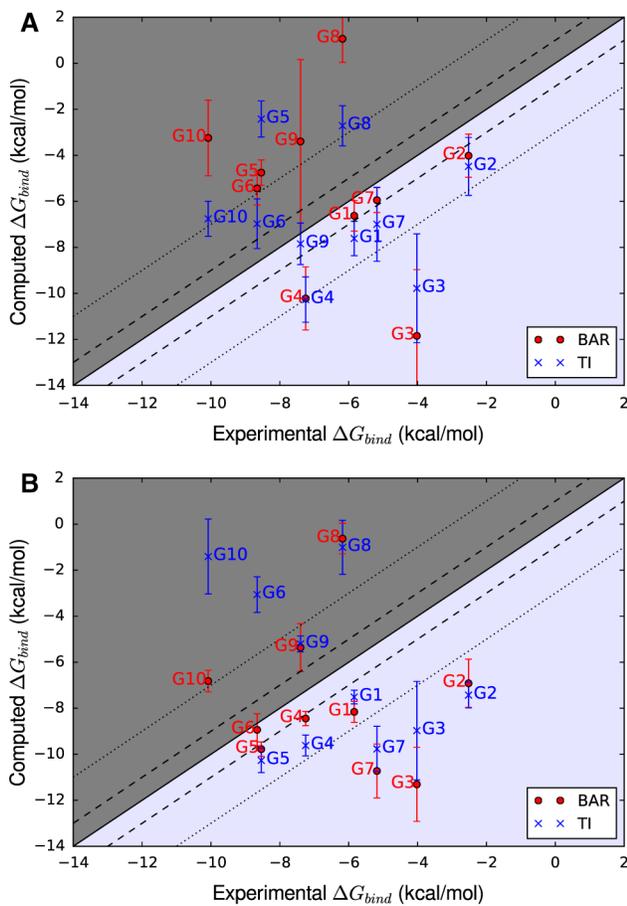


Fig. 6. Comparison of the binding affinity predictions obtained with TI (*blue crosses*) and HREM/BAR (*red circles*) calculations whose initial docked poses were obtained with **A** the docking calculations and **B** the MD simulations. The *solid diagonal line* corresponds to perfect agreement with experiments. The *dashed and dotted lines* correspond to absolute errors of 1 and 3 kcal/mol, respectively

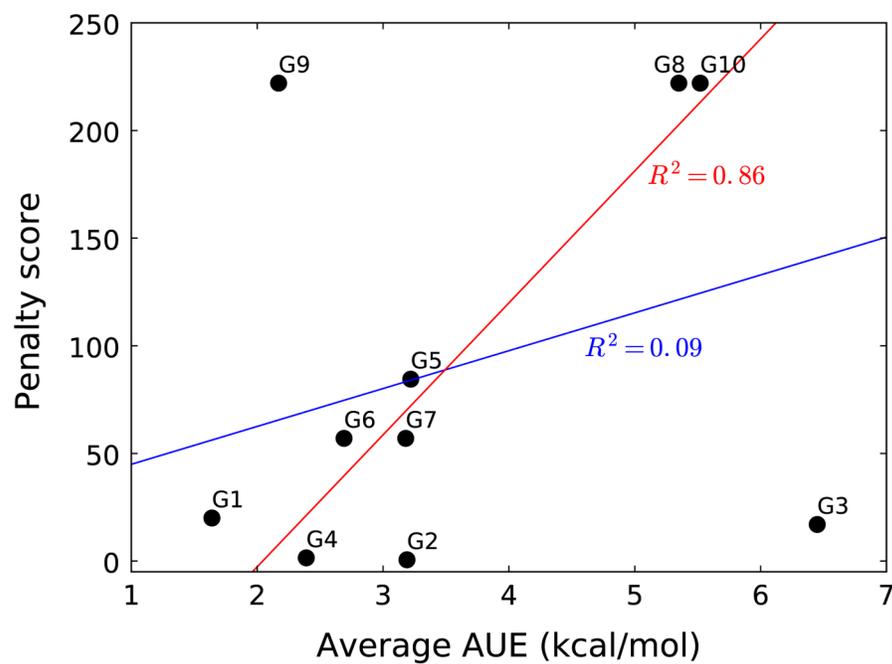


Fig. 7. Correlation between the penalty scores of CGENFF parameters of guests and the average AUE values. The *blue line* is the linear regression of all data with a R^2 value of 0.09. The *red line* is the linear regression obtained by excluding G3 and G9 with a R^2 value of 0.86

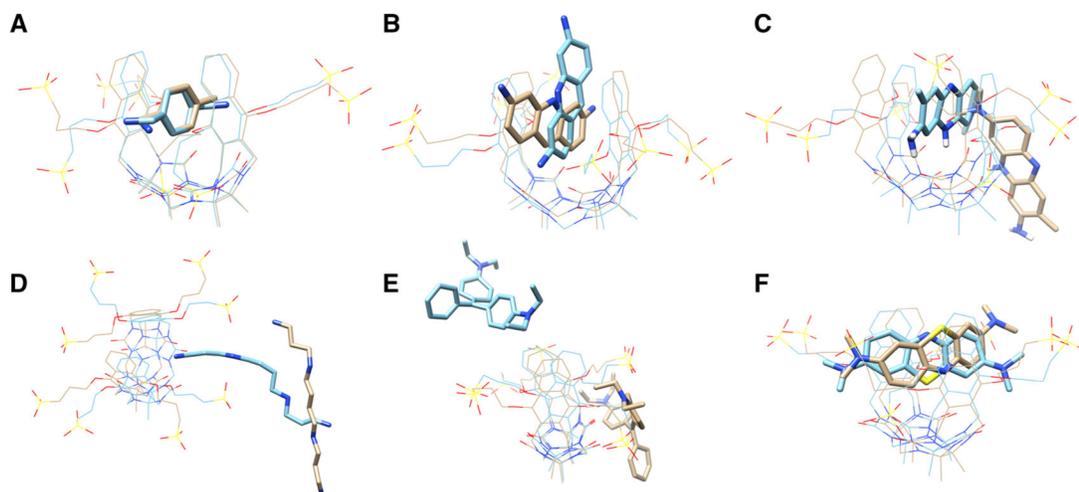


Fig. 8. Docked conformations of six guest molecules. The binding free energies of **A** G1, **B** G4, and **C** G9 were accurately calculated and those of **D** G3, **E** G8, and **F** G10 were incorrectly calculated. The carbon atoms of the conformations obtained with docking simulations are *colored in yellow* and those obtained with MD simulations are colored in *pale blue*. Nitrogen, oxygen, and sulfur atoms are *colored in blue, red, and yellow*, respectively

Table 1

AUE and RMSD of five submissions for the CBClip challenge

	TI-dock	BAR-dock	TL-MD	BAR-MD	All
AUE	2.94 ± 0.37	3.89 ± 0.49	4.18 ± 0.33	3.31 ± 0.28	3.28 ± 0.35
RMSD	3.41 ± 0.45	4.62 ± 0.64	4.68 ± 0.43	3.96 ± 0.37	3.96 ± 0.59

Table 2
Free energy values of individual terms for the double decoupling scheme calculations of TI-dock

Guest	$\Delta G_{\text{rest-on}}^C$	$\Delta G_{\text{rest-off}}^C$	ΔG_{elec}^C			ΔG_{vdW}^C			ΔG_{elec}^M			ΔG_{vdW}^M			G_{bind}		
			TI	BAR	TI	BAR	TI	BAR	TI	BAR	TI	BAR	TI	BAR	TI	BAR	TI
G1	7.9 ± 0.4	-8.8	44.6 ± 0.5	44.7 ± 0.1	-13.1 ± 0.5	-11.9 ± 0.0	41.4 ± 0.1	42.4 ± 0.1	-18.4 ± 0.3	-17.2 ± 0.1	-7.6 ± 0.7	-6.6 ± 0.7					
G2	7.9 ± 0.7	-8.4	76.2 ± 1.2	76.4 ± 0.3	-2.9 ± 0.3	-1.9 ± 0.3	71.4 ± 0.3	72.3 ± 0.0	-3.2 ± 0.2	-2.4 ± 0.0	-4.5 ± 1.3	-4.0 ± 0.9					
G3	13.2 ± 1.2	-7.3	118.5 ± 1.8	120.9 ± 2.5	-5.1 ± 1.5	-2.7 ± 0.8	113.9 ± 0.2	115.6 ± 0.1	-4.2 ± 0.1	-3.1 ± 0.2	-9.6 ± 2.4	-11.7 ± 2.8					
G4	10.0 ± 1.8	-8.9	65.9 ± 0.2	66.4 ± 0.1	-24.1 ± 0.9	-21.3 ± 0.1	65.0 ± 0.2	65.6 ± 0.0	-32.4 ± 0.1	-29.7 ± 0.1	-10.3 ± 1.0	-10.2 ± 1.4					
G5	6.2 ± 0.1	-8.4	25.0 ± 0.2	25.3 ± 0.1	-42.2 ± 0.7	-36.5 ± 0.4	24.2 ± 0.1	24.9 ± 0.0	-46.1 ± 0.3	-43.1 ± 0.1	-2.4 ± 0.8	-4.8 ± 0.6					
G6	7.5 ± 0.5	-8.7	9.2 ± 0.3	9.2 ± 0.1	-42.6 ± 1.0	-38.0 ± 0.2	8.7 ± 0.2	10.2 ± 0.1	-50.2 ± 0.2	-45.7 ± 0.1	-7.0 ± 1.1	-5.5 ± 0.7					
G7	7.4 ± 0.3	-8.6	397.5 ± 0.8	396.0 ± 0.2	-39.3 ± 1.0	-34.9 ± 0.3	396.9 ± 0.9	392.9 ± 0.1	-46.9 ± 0.3	-39.1 ± 0.1	-7.0 ± 1.6	-6.0 ± 0.5					
G8	12.4 ± 0.8	-7.7	19.2 ± 0.1	19.1 ± 0.1	-78.6 ± 0.6	-77.0 ± 0.4	19.6 ± 0.0	20.2 ± 0.0	-77.0 ± 0.6	-72.3 ± 0.2	-2.7 ± 0.9	1.1 ± 1.0					
G9	8.3 ± 0.2	-8.5	-14.1 ± 0.2	-14.0 ± 0.0	-40.6 ± 0.9	-39.3 ± 3.5	-16.7 ± 0.1	-15.8 ± 0.0	-46.0 ± 0.1	-41.1 ± 0.1	-7.8 ± 0.9	-3.4 ± 3.6					
G10	8.9 ± 0.2	-8.5	-7.1 ± 0.2	-7.3 ± 0.2	-41.3 ± 0.6	-39.5 ± 1.6	-11.0 ± 0.0	-10.4 ± 0.0	-43.7 ± 0.3	-39.4 ± 0.1	-6.8 ± 0.8	-3.4 ± 1.6					

Table 3
The vdW free energies obtained with BAR and TI calculations integrated with the trapezoidal and Simpson's rules

Guest	ΔG_{vdW}^C (TI-long)		ΔG_{vdW}^C (TI-short)		ΔG_{vdW}^M	
	TI-Trapezoidal	TI-Simpson	TI-Trapezoidal	TI-Simpson	TI-Trapezoidal	TI-Simpson
G1	13.9	14.4	13.1	13.8	18.4	18.8
G2	2.4	2.8	2.9	3.2	3.2	3.5
G3	7.3	6.6	5.1	5.9	4.2	5.1
G4	24.1	25.1	24.1	25.1	32.4	32.5
G5	42.3	44.7	42.2	44.5	46.1	46.2
G6	43.3	46.4	42.6	44.9	50.2	51.0
G7	40.4	42.0	39.3	40.3	46.9	47.2
G8	78.7	79.2	78.6	79.2	77.0	78.0
G9	40.8	42.3	40.6	41.9	46.0	46.6
G10	41.4	43.2	41.3	43.3	43.7	44.8

Free energy values of individual terms for the double decoupling scheme calculations of the longer TI calculations with the “dock” set

Table 4

Guest	$\Delta G_{\text{rest-on}}^C$	ΔG_{elec}^C	ΔG_{vdW}^C	$\Delta G_{\text{total}}^C$	$\Delta G_{\text{total}}^{\text{TI-long}}$	$\Delta G_{\text{total}}^{\text{TI-short}}$
G1	8.4 ± 0.7	44.4 ± 0.3	-13.9 ± 0.1	-7.2 ± 0.1		0.4
G2	8.1 ± 0.9	76.2 ± 0.2	-2.4 ± 0.3	-5.2 ± 0.3		-0.8
G3	12.3 ± 0.7	116.6 ± 0.7	-7.3 ± 3.4	-4.6 ± 3.4		5.0
G4	10.0 ± 1.3	65.9 ± 0.2	-24.1 ± 0.1	-10.2 ± 0.1		0.0
G5	6.1 ± 0.1	24.0 ± 1.3	-42.3 ± 0.4	-1.3 ± 0.4		1.1
G6	8.1 ± 0.4	9.8 ± 1.1	-43.3 ± 0.1	-7.4 ± 0.1		-0.5
G7	7.4 ± 0.1	396.5 ± 0.5	-40.4 ± 0.5	-4.9 ± 0.5		2.1
G8	14.2 ± 0.8	19.2 ± 0.2	-78.7 ± 0.6	-4.4 ± 0.6		-1.7
G9	8.2 ± 0.3	-14.1 ± 0.2	-40.3 ± 0.3	-7.5 ± 0.3		0.4
G10	9.0 ± 0.2	-7.2 ± 0.2	-41.4 ± 0.1	-6.6 ± 0.1		0.2