Affinity Calculations of Cyclodextrin Host-Guest Complexes: Assessment of Strengths and Weaknesses of End-Point Free Energy Methods

Dimas Suárez,(*) Natalia Díaz.

Departamento de Química Física y Analítica. Universidad de Oviedo.

Avda. Julián Clavería 8. Oviedo (Asturias). 33006. SPAIN

dimas@uniovi.es

KEYWORDS: CYCLODEXTRINS · BINDING ENERGY · END-POINT METHODS · MOLECULAR DYNAMICS.

ABSTRACT

The end-point methods like MM/PBSA or MM/GBSA estimate the free energy of a biomolecule by combining its molecular mechanics energy with solvation free energy and entropy terms. Their performance largely depends on the particular system of interest and despite numerous attempts to improve their reliability that have resulted in many variants, there is still no clear alternative to improve their accuracy. On the other hand, the relatively small cyclodextrin host-guest complexes, for which high quality binding calorimetric data are usually available, are becoming reference models for testing the accuracy of free energy methods. In this work, we further assess the performance of various MM/PBSA-like approaches as applied to cyclodextrin complexes. To this end, we select a set of complexes between β-CD and 57 small organic molecules that has been previously studied with the binding energy distribution analysis method in combination with an implicit solvent model (Wickstrom, L.; He, P.; Gallicchio, E.; Levy, R. M. *J. Chem. Theory Comput.* **2013**, 9, 3136-3150). For each complex, a conventional 1.0 μs MD simulation in explicit solvent is carried out. Then we employ semiempirical quantum chemical calculations, several variants of the MM-PB(GB)SA methods, entropy estimations, etc., in order to assess the reliability of the end-point affinity calculations. The best end-point protocol in this study, which combines DFTB3 energies with entropy corrections, yields estimations of the binding free energies that still have substantial errors (*RMSE*=2.2 kcal/mol), but exhibits a good prediction capacity in terms of ligand ranking (R^2=0.66) that is close or even better than that of rigorous free energy methodologies. Our results can be helpful to discriminate between the intrinsic limitations of the end-point methods and other sources of error, such as the underlying energy and continuum solvation methods.

INTRODUCTION

End-point free-energy methods[1];[2] directly estimate free energies from unrestrained simulations sampling the phase space of individual states. This family of computational methods is quite diverse. Thus, there are theoretical approaches like the M2 method,[3] which are based on the relation between free energy and configurational integral Q (*i.e.*, $G=-RT\ln Q$), adopt solvent continuum models and rely on the exhaustive sampling of the phase space so that they are mainly restricted to study relatively small molecules. Another end-point technique results from the combination of the Protein-Dipoles Langevin-Dipoles (PDLD) semi-microscopic solvent model with the linear response approximation (LRA) to calculate free energy differences in Perturbation Theory, giving rise to the PDLD/S-LRA method to calculate binding free energies.[4] Conceptually similar to LRA, the Linear Interaction Energy method (LIE) evaluates electrostatic and nonpolar binding energy components by linearly scaling the average values of molecular mechanics energy terms with parameters fitted to reproduce more rigorous calculations.[5] In contrast, the most popular end-point techniques, which are derived from the MM/PBSA protocol (Molecular Mechanics-Poisson-Boltzmann Surface Area) originally proposed by Kollman and Massova,[6] lack a strong theoretical foundation as they *directly* approximate the free energy of a solute molecule by combining its molecular mechanics (MM) energy with estimations of its solvation free energy and entropy. Some theoretical support can be derived from a first-order approximation to the calculation of configurational integrals combined with the assumption that the phase space volume of the host and ligand molecules remain unaltered upon complexation, what leads to binding energy expressions resembling those of the MM/PBSA method.[1] However, there are many variants of the MM/PBSA approach that have been employed in a broad range of applications [7];[8] so that these methods may be better introduced as physically-based *scoring* functions.

Several reviews[7, 9-10] have been published during the last years that examine the various approximations underlying the MM/PBSA-like approaches and focus on recent developments and applications. Nonetheless, the diversity of MM/PBSA approaches is briefly introduced here to better outline the goals of the present work. Thus, MM/PBSA-like energy (*G*) can be derived either from single-point calculations (*e.g.*, on X-ray structures, docking poses) or from averages of calculations on snapshots extracted from molecular dynamics (MD) simulations (after having removed all solvent molecules and counterions in explicit solvent simulations). These energies are computed according to the following kind of equations:

$$G \approx 3RT + E + \Delta G_{solv}$$

where E is the gas-phase energy, the $3RT$ contribution is due to six translational and rotational degrees of freedom and ΔG_{solv} is the solvation energy, which is typically decomposed into polar and nonpolar contributions. Many specific methods and approximations have been used to calculate E and ΔG_{solv}. Most commonly, the gas-phase energy is obtained using conventional MM force fields, preferably the same force field used in the previous simulations or geometry relaxations, but quantum mechanical (QM) and hybrid QM/MM methods have been also employed.[11-13] The electrostatic solvation energy (ΔG_{solv}^{el}) is normally estimated by means of implicit solvent methods like the Poisson-Boltzmann (PB)[14] and Generalized-Born (GB) [15] methods. Both PB and GB represent the solute molecule in terms of a set of atomic charges and parameterized radii, but while PB determines the solute electrostatic potential through the numerical solution of the Poisson equation, GB is based on a modified and further parameterized Coulombic expression that is computationally faster.

The electrostatic PB(GB) solvation energies must be complemented with the corresponding nonpolar part (ΔG_{solv}^{np}) due to cavity formation and van der Waals interactions between the solute and the solvent molecules. In principle, a surface area term, $\gamma A + b$, can approximate

ΔG_{solv}^{np}, but this simple recipe is better suited to estimate cavity costs rather than solute-solvent dispersion. Thus, effective nonpolar solvent models that treat separately the cavity (repulsive) and dispersion (attractive) energies have been derived.[16] The post-processing of explicit solvent MD simulations also allows the evaluation of the dispersion part by computing the average van der Waals interaction energy between solute and the water molecules within a shell of 12-15 Å thickness.[17] Alternatively, the PB(GB) and nonpolar solvation energies can be replaced by those calculated with the three-dimensional reference interaction site model (3D-RISM).[18] This method,[19] which is based on a first-principles statistical mechanics theory, yields the 3D distribution functions of the solvent atoms around each solute site, allowing thus specific descriptions of polar and non-polar solvation effects for each solute-solvent interaction site, albeit with larger computational cost than that of the PB(GB)SA methods. Mixed explicit/implicit solvation models have also been used to include specific solvent effects in the MM/PBSA-like protocols by augmenting the solute coordinates with those of selected water molecules located nearby important binding sites.[20-21] This approach, which has improved agreement with experiment for various protein-ligand complexes, introduces some arbitrariness in the selection of the location and number of explicit waters.

Addition of absolute entropy terms to the *G* expressions can be a key ingredient for a successful application of the MM/PBSA-like methods. However, the accurate determination of single-molecule entropies is still challenging due to the incomplete sampling performed by the molecular simulation methods and/or theoretical difficulties in the determination of entropy correlation effects.[22] In the context of the end-point free energy methods, entropy has been normally accounted for using the statistical thermodynamic formulas derived within the rigid-rotor harmonic-oscillator (RRHO) approximations, but again different choices and approximations have been considered. Thus, the absolute entropy of a flexible molecule can be approximated by:

$$S \approx \langle S_{RRHO} \rangle + S_{conform}$$

where $\langle S_{RRHO} \rangle = \langle S_{trans} + S_{rot} + S_{vib} \rangle$ is the average RRHO entropy over the various conformers that can be decomposed into the translational, rotational and vibrational parts, the latter one being computationally demanding because strict energy relaxations followed by Hessian calculations are required. For this reason, truncated models and simplified Hessian methods have been devised.[23-24] On the other hand, $S_{conform}$, which is the entropy that arises from the population of the various conformers, is usually neglected in MM/PBSA-like studies even though this entropy contribution can play an important role in binding and folding processes. The sum of $S_{conform}$ and $\langle S_{vib} \rangle$ can be also termed as *configurational* entropy[25] and this can be estimated using other methods like the quasi-harmonic (QH) technique in which the Hessian matrix elements are approximated by the covariance matrix of the Cartesian coordinates around the average configuration.[26] Although the QH method is attractive and computationally efficient, it largely overestimates the entropies of flexible molecules.[27] Therefore, a satisfactory approach to the inclusion of absolute entropies in MM/PBSA calculations is still lacking.

In principle the binding free energy estimate obtained with end-point methods consists of the difference of average G energies derived from three independent simulations, that is, $\Delta G_{bind} \approx \langle G(cmplx) \rangle - \langle G(host) \rangle - \langle G(lig) \rangle$. Another source of variability in the application of the MM/PBSA-like techniques can be found in the adoption, or not, of the one-trajectory approximation (also known as one-average approximation). For a host-ligand complex, the interaction energy can be computed by averaging relative G values: $\Delta G_{int} \approx \langle G(cmplx) - G(host^*) - G(lig^*) \rangle$. In this expression, the asterisk superscript means that the G energies of the host and ligand molecules are evaluated using their distorted geometries in the complex. In the case of small ligands bound to proteins, the one-trajectory approximation equals the full binding energy ΔG_{bind} by ΔG_{int}, what minimizes statistical

uncertainty and facilitates the estimation of relative affinities for a series of ligand molecules. This same approximation has been adopted to define the so-called interaction entropy derived from the fluctuations of the host-ligand interaction energy (*i.e.*, ΔE_{int} as calculated with the MM force fields).[28] However, for relatively large complexes in which a substantial structural rearrangement takes place upon binding, the one-trajectory approximation ignores the distortion effects.

The review of the recent literature[7, 9-10] shows that the MM/PBSA-like methods are quite popular in molecular modelling not only for estimating the binding energy of protein-ligand, protein-peptide or protein-protein complexes studied by MD simulations, but also to refine the energy scorings of docking poses and/or validate fast energy scoring functions.[29] However, the broad diversity of MM/PBSA-like energy functions hampers their systematic assessment. In the recent reviews, one can find mixed impressions concerning the overall reliability and usefulness of these methods. On one hand, Ryde and co-workers have concluded that they "may be useful to improve the results of docking and virtual screening or to understand observed affinities and trends".[9] However, these authors point out that MM/PBSA and the related approaches rely on questionable approximations, tend to overestimate energy differences and lack the required accuracy for predictive drug design due to their large statistical uncertainty and systematic errors that usually amount to few kcal/mol. On the other hand, Luo and co-workers[10] have noticed that there is "a large volume of valuable predictive results in a wide variety of studies". They admit that end-point free energies are less accurate than those provided by computationally-expensive thermodynamic integration or free energy perturbation methods, but also emphasize that "the qualitative agreement is often good enough" in the sense that ligand rakings based on MM-PBSA-like scorings may have enough predictive capacity despite of their statistical and systematic errors. This view is well illustrated by a recent work on the comparative performance of MM/PBSA and absolute binding free energy

calculations on 47 protein-ligand complexes, finding that MM/PBSA exhibits a reasonable correlation with experimental affinities at lower computational cost.[30] In any case, it is clear that the performance of the MM/PBSA-like approaches depends both on the nature of the studied systems and on the details of the calculation and there is still no clear alternative to improve their accuracy while keeping their simplicity and moderate computational cost.

Cyclodextrins (CDs) are cyclic oligosaccharides usually formed by 5-7 glucopyranoside units.[31] The CD structure resembles a truncated cone with an inner partially hydrophobic cavity lined by hydroxyl groups at the narrow and wide sides. In this way, CDs can host different kinds of guest molecules by forming inclusion complexes, which have many practical applications in industry and pharmaceutical chemistry.[32] More interestingly, the CD inclusion complexes are very attractive from the point of view of basic research because their binding thermodynamics can be determined accurately with high precision calorimetry while the relatively small size of CDs facilitates the intensive application of molecular modeling methods including free energy methodologies. Thus, CD inclusion complexes are becoming reference models for testing the accuracy of simulation methods in the context of noncovalent binding.

An interesting computational study that relies on experimental binding data of CDs to validate free energy calculations has been reported by Levy and co-workers[33], who calculate the absolute binding energy of 57 β-CD complexes with small organic molecules. They use the so-called binding energy distribution analysis method (BEDAM) in combination with the OPLS-AA force field and the analytical generalized Born plus nonpolar (AGBNP2) solvent model (this GB method includes also hydrogen-bonding corrections and it is augmented with two hydration sites specific for β-CD). The BEDAM technique estimates the free energy cost of connecting the coupled and decoupled states of the host-guest molecules confined within a predefined binding site volume immersed in the solvent continuum. The authors find a reasonable agreement with experimental data (global R^2=0.43 and RMSE=1.44 kcal/mol) and

conclude that, for the most flexible ligands, the enthalpic and entropic reorganization effects are essential to obtain a better affinity ranking than those provided by methods based only on host-guest interaction energies. Similarly, Gilson and co-workers, who advocate for the inclusion of host-guest systems in blinded tests,[34] have recently simulated 43 complexes of α-CD/β-CD with known ΔG_{bind}/ΔH_{bind} values using the attach-pull-release (APR) method and explicit solvent.[35] In the APR technique, the ligand is pulled by imposing adequate harmonic restraints at varying host-guest distances and the authors focus on the performance of different force fields for the water solvent and the solute molecules. Interestingly, it turns out that none of the tested force field combinations exhibits a clear advantage with respect to the others, the corresponding R^2 coefficients for ΔG_{bind} varying between 0.43 and 0.60 with RMSE values between 0.85 and 1.80 kcal/mol. Another set of 75 neutral ligands in complex with β-CD has been selected by van der Spoel and coworkers[36] to examine the goodness of the MM/GBSA method either to energetically rank docking poses or to calculate binding energies using a potential of mean force (PMF) protocol. Although various generic sets of GB parameters are considered, both the docking and PMF calculations give affinity energies that are weakly correlated with experimental data (R^2~0.3), what has been ascribed to limitations of the implicit solvent models. The separate evaluation of enthalpy and entropy changes as well as the rates of association/dissociation between β-CD and a small set of 7 ligands, have been attempted by Tang and Chang[37] using multi-microsecond MD simulations in explicit solvent followed by elaborate post-processing of the trajectories. These authors conclude that β-CD/ligand binding is mainly stabilized by the nonpolar interactions and the entropy gain of desolvated water molecules.

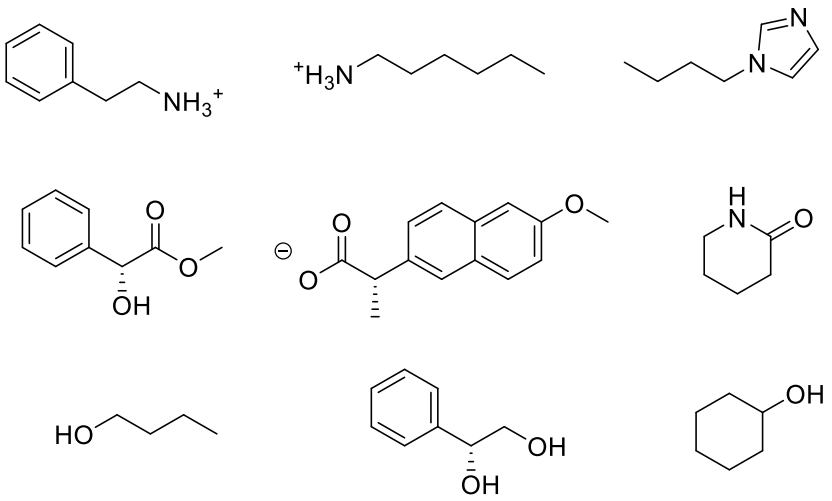
In line with the recent theoretical studies, we aim in this work to further assess the weaknesses and strengths of the MM/PBSA-like approaches as applied to host-guest complexes. In particular, we select the same reference set of β-CD complexes that has been

studied by Levy.[33] For each of the 57 CD-ligand pairs, we carried out a conventional 1.0 μs MD simulation in explicit solvent that samples exhaustively the CD-ligand interactions and their internal motions, minimizing thus potential uncertainties in the determination of molecular geometries. Analogously, the phase space of the separate fragments was explored by means of long MD simulations. Subsequently, MM/PBSA-like calculations were done using various settings and method choices for the gas-phase energy and the solvation free energy as well as by including normal-mode and conformational entropy estimations. Semiempirical QM (SQM) calculations were also considered using the self-consistent charges density functional tight-binding method including third order terms with dispersion and hydrogen-bond corrections. The comparative performance of various energy and solvation methods and the role played by entropy corrections is assessed in terms of their ligand ranking ability rather than in the systematic errors of the ΔG_{bind} estimations. The comparison with former calculations can help discriminate between the intrinsic limitations of the end-point protocol for the free energy estimation and other sources of error, such as the underlying energy and continuum solvation methods.

Methods

Ligand parameterization

As mentioned in the Introduction, we used the set of CD ligands that has been previously used by Levy and co-workers in their validation study of the BEDAM method. This set comprises 57 guests of varying size, charge and chemical groups (structural formulas of representative ligands are shown in Scheme 1; the full list of ligand chemical names and formulas is provided in the Supporting Information). In this work, the ligands are grouped into various classes depending on the identity of their main functional groups (e.g., alcohol/phenol (20), ammonium cations (21), acetate anion (3), imine heterocycle (3), etc.).



Scheme 1. Examples of β-CD ligands

Starting with the 2D chemical structures of each ligand, we obtained its SMILE code and converted it into a 3D structure with the help of the Openbabel toolbox [38]. Using then the *sqm* and *antechamber*[39] programs available in the AMBER16 package,[40] the semiempirical AM1 method[41] is used to relax the internal geometry and to obtain atomic charges (AM1-BCC). These charges together with AMBER-1999 Lennard-Jones parameters were fed into the *MS-DOCK* program[42] that carried out a fast conformational search. After energetic and structural filtering to remove high energy and redundant conformers, the resulting conformers (~1-5 structures depending on the ligand identity) were optimized at the HF/6-31G* level of theory using the *gaussian03* package.[43] The *antechamber* program was used again to assign GAFF atom types[44] and atomic charges, which were built with the RESP methodology and taking into account the HF/6-31G* electrostatic potential of each conformer.[45] The dihedral parameters for internal rotations in which at least one of the central or terminal atoms is electronegative were refined using the *paramfit* program included in AMBER16. To this end, the most stable HF/6-31G* conformer of the ligand was perturbed by rigid rotation around the corresponding bond(s) in 30º steps. After removing the geometries with steric clashes, the QM energies of the remaining structures were evaluated by means of single-point RI-MP2/aug-cc-pVTZ calculations done with the ORCA program.[46] The resulting QM data and geometries were

introduced into *paramfit*, which optimized additional dihedral parameters applying a least-squares fitting procedure. All the ligand parameters are available in the Supporting Information.

MD simulations

The initial geometry of the β-CD molecule was taken from the Cambridge Structural Database (ref. 1107195).[47] β-CD was described with the GLYCAM-06j parameter set,[48] which is usually employed to simulate carbohydrates in combination with the AMBER/GAFF force fields. Each ligand molecule was placed in the β-CD cavity by superposing the center-of-mass of the two molecules and aligning their principal axes of inertia. Using the *tLEaP* program in AMBER16, the CD/ligand complexes were centered in an octahedral box of TIP3P [49] water molecules that extended at least 14 Å from the solute atoms. Na^+/Cl^- counterions were added to the solvent box by *tLEaP* in order to neutralize the systems when negatively/positively charged ligands are included. For a typical β-CD complex, these settings resulted in a simulation box containing a total of ~8800±100 atoms (~5600±45 atoms for the box containing the isolated ligand).

Periodic boundary conditions were applied to simulate continuous systems and long-range interactions were described by the Particle-Mesh-Ewald method with a grid spacing of ~1 Å and a non-bonded cutoff of 9 Å. A flat bottom potential was imposed to restrict the center of mass of the ligand to sample a binding region within 6 Å from the center-of-mass of the β-CD. Water molecules were relaxed by means of energy minimizations and 100 ps of MD at 300 K with the *sander* program. The full systems were minimized and heated gradually to 300 K during 120 ps of MD (constant NVT) with a time step of 1.0 fs. Subsequently, the systems were pressurized (1 bar) by running a 2.0 *ns* NPT simulation. The production phase of the simulations comprised 1.0 μs, which were run at NVT conditions with a time step of 2 fs and using the accelerated version of the *pmemd* code for Graphical Processing Units.[50-51]

Coordinates of the solute atoms were saved for analysis every 2.5 ps and those of the water molecules every 50 ps. During the MD simulations, Langevin dynamics[52] was employed to control the temperature with a collision frequency of 2 ps^{-1} and the length of all R-H bonds was constrained with the SHAKE algorithm.[53] In the NPT run, the pressure of the system was controlled by a Monte Carlo barostat as implemented in AMBER16.

The MD simulations of the unbound states of the ligand molecules were executed with analogous settings to those employed in the case of the β-CD/ligand simulations, the production phase extending up to 0.5 μs. For the unbound state of β-CD, we collected MD snapshots from our previous simulation work[54] on the conformational and entropic properties of α-, β- and γ-CDs, which were run for 5.0 μs and using comparable MD settings.

Structural analyses were performed using the *cpptraj* module[55] of AMBER16. The coordinates of the β-CD atoms along the MD trajectories were clustered using *cpptraj* with the average-linkage clustering algorithm and a sieve of 50 frames. The distance metric between frames was calculated via best-fit coordinate root mean square deviation (RMSD) using the two ether oxygens (O4 and O5) and the five carbon atoms (C1-C5) in the sugar rings, and the clustering was finished when the minimum distance between clusters was greater than 1.5 Å.

End-point free energy estimations

MM/PB(GB)SA calculations

The various MM/PB(GB)SA calculations reported in this work were carried out on 10000 snapshots extracted from each MD simulation every 100 ps after having removed all solvent molecules and counterions. The gas-phase MM energy was evaluated using the GLYCAM/GAFF force field with no-cutoff and the *sander/pbsa* programs. To calculate the electrostatic PB solvation energy, the non-linear Poisson-Boltzmann equation[56] was solved on a cubic lattice by using an iterative finite-difference method and taking the atomic charges and

modified Bondi atomic radii automatically assigned by the *tLEaP* program. These calculations were done with the *pbsa* program, choosing a grid spacing of 0.33 Å, null ionic strength and building the dielectric boundary as the contact surface between the radii of the solute and the radius (1.4 Å) of a water probe molecule. The internal and external dielectric constant values ε_{int}=1 and ε_{out}=80, respectively, were used in the PB calculations (calculations choosing ε_{int}=4 were also performed). For computing the GB solvation energy, we used the *sander* program and the Hawkins-Cramer-Truhlar pairwise Generalized-Born (HCT-GB) model was selected. The total PBSA and GBSA energies include also the implicit non-polar terms: the dispersion term and the cavity term according to the model of Tan *et. al*.[16] In combination with the PB energies, we also tested the approximation introduced by Gohlke and Case[17] to estimate the non-polar solvation energy:

$$\Delta G_{solv}^{np} = \Delta H_{solute-solvent}^{vdW} + \gamma MSA$$

where $\Delta H_{solute-solvent}^{vdW}$ is the van der Waals interaction energy between solute and the water molecules within a shell of 12 Å thickness, while the cavitation free energy contribution to the non-polar solvation energy is determined by a molecular surface area (MSA) dependent term. This approach includes the favorable van der Waals dispersion between all the solute atoms and the nearby solvent molecules that may be especially relevant when assessing the stability of host-guest complexes in which the number of solvent-exposed and buried atoms can differ considerably.

MM/RISM calculations

The MM/RISM approach replaces the PBSA or GBSA polar and nonpolar solvation terms by the solvation free energy calculated with the RISM theory. More specifically, we solved the 3D-RISM integral equations derived with the so-called Kovalenko-Hirata (KH) closure relation[57] on the MD snapshots sets. The calculations were performed with the *sander* program available in AMBER16 coupled with the TIP3P water model for solute-solvent and solvent-

solvent interactions. The 3D-RISM-KH solvent distributions were solved iteratively until a residual tolerance of 10^{-5} on a cubic grid with 0.33 Å spacing and a solute-solvent interaction cutoff of 15 Å. The Pressure Correction Plus (PC+/3D-RISM)[58] was added to mitigate the overestimation of the non-polar component of the solvation energy. MM/RISM is much more computationally demanding than MM/PB(GB)SA and for this reason the calculations were carried out on 1000 MD snapshots, equally-spaced along the MD trajectories (500 snapshots for the isolated ligands).

Semiempirical QM/MM and QM calculations

We carried out semiempirical QM (SQM) calculations in order to analyze their impact on the accuracy of the β-CD/ligand affinity rankings. Thus, we employed the self-consistent charges density functional tight-binding (SCC-DFTB) method[59-60], which is a fast and general SQM method based on density functional theory (DFT) parameterized against high-level calculations. In particular we used the DFTB3 version,[61] which improves the original SCC-DFTB method by better describing electron-electron interactions and including third-order terms of the Taylor series expansion of the DFT exchange-correlation energy. The Slater-Koster parameters embedded into the DFTB3 Hamiltonian were extracted from the 3OB set[62-63] which has been optimized for organic and biological molecules. In addition, dispersion and hydrogen-bonding corrections (dubbed as D3H4)[64] were taken into account, the resulting energies being thus denoted as DFTB3-D3H4.

From each simulation of the β-CD bound and unbound ligands, we selected 1000 solvated snapshots which were first truncated by removing water molecules >15.0 Å away from the solute atoms. The internal geometry of the β-CD/ligand fragments was relaxed by performing DFTB3/MM geometry optimizations (50 cycles) with the *sander* program: the solute atoms were in the SQM region while the fixed water molecules were described with the TIP3P force field.

Upon removal of the coordinates of water molecules from the SQM/MM relaxed structures, three DFTB3-based protocols for computing the ΔG_{bind} and ΔG_{int} energies were assessed. In the first one, the DFTB3-D3H4 energy of the relaxed structures was calculated using the Cuby4 framework[65] coupled with the DFTB3 program.[66] The gas-phase DFTB3-D3H4 energy was augmented with the solvation energy computed with the PM6[67] SQM level coupled with the COSMO implicit solvent model[68] as implemented in the MOPAC program[69]. The sum of the DFTB-D3H4 energy and the PM6-COSMO solvation term gives the SQM/COSMO scoring function,[64] which we refer to as DFTB3/COSMO in this work. In the second scoring protocol, we computed the DFTB3/GBSA energy of the structures using the SQM GB implementation in *sander* which, as described by Pellegrini and Field,[70] takes the Mulliken SQM charges. In the third protocol, we employed the *pbsa* program to calculate the PBSA solvation energy using the Mulliken atomic charges derived from the DFTB3/GBSA density, the modified Bondi atomic radii and identical settings as those used in the MM/PBSA calculations. The PBSA energy was combined with the DFTB3 energy extracted from the DFTB3/GBSA calculation with *sander.* The dispersion and hydrogen-bond corrections (D3H4) were also added to the DFTB3/GBSA and DFTB3/PBSA energies.

We quickly examined the performance of other SQM scorings that combine PM6[67]gas-phase energies with PBSA solvation terms as computed with the PM6 Mulliken charges. The PM6 Hamiltonian was enhanced with hydrogen-bonding corrections and standard dispersion energy,[71] resulting in the so-termed PM6-DH+ method. Just single-point PM6-DH+/PBSA calculations were carried out on the geometries of the β-CD/ligand solute(s) as extracted from the MD snapshots and using the *sander* and *pbsa* programs.

For comparative purposes, the gas-phase interaction energy of selected β–CD complexes was calculated using the GLYCAM/GAFF force field, the DFTB3-D3H4 method and the B3LYP-D3/6-31+G** level of theory. The MM and DFTB3 energies were computed following

the same prescriptions as followed in the end-point energy calculations. The B3LYP/6-31+G** energies,[72-73] which were computed using the *TeraChem* software,[74-75] correspond to single-point calculations on structures that were previously relaxed by performing hybrid HF/6-31G*/MM geometry optimizations (50 cycles) with the *sander* program coupled with the *gaussian03* program. Again in these QM/MM optimizations, only the QM solute atoms were allowed to move while the TIP3P water molecules were frozen. The B3LYP energies were augmented with the Grimme's D3 potential which yields pairwise dispersion energy corrections.[76] The basis set superposition error in the B3LYP interaction energies was corrected with the counterpoise method.[77]

Entropy calculations

To complement the diverse combinations of energy and solvent methods, we estimated the absolute entropy ($S \approx \langle S_{RRHO} \rangle + S_{conform}$) of the β-CD/ligand complexes and the separate fragments by taking advantage of the MD sampling. The details of the corresponding protocols leading to the $\langle S_{RRHO} \rangle$ and $S_{conform}$ terms are summarized below.

RRHO entropy calculations

The RRHO entropy calculations were performed on 10000 MD snapshots extracted from the MD trajectories and using the MM force field employed in the explicit solvent simulations. To carry out the normal mode calculations yielding the harmonic frequencies, the structures must be subject to geometry optimization setting a tight tolerance in the energy gradient, but avoiding that the minima deviate significantly from the sampled structures in solution. In this respect, two different protocols were considered. In the first one, the solvated MD snapshots were truncated by keeping the coordinates of the solute atoms and those of a buffer layer of waters with a ~6 Å thickness around the solute atoms. The *sander* program was used to relax the internal geometry of the solute until the root mean squared deviation (RMSD) of the elements in the gradient vector is less than 10^{-6} kcal mol^{-1} Å^{-1}. Subsequently, the calculation

of the Hessian matrix of second derivatives and its diagonalization to obtain the normal modes were restricted to the active region comprising the solute atoms using a locally modified version of the *nmode* program, which is also included in AMBER16. A similar protocol has been shown to reduce the statistical uncertainty of the entropy terms and give improved MM/PBSA scorings.[78] In the second protocol for the RRHO calculations, we used the HCT-GB implicit solvent model for removing the explicit consideration of solvent degrees of freedom. The selected MD snapshots were then post-processed through the removal of all waters and counterions and the geometries of the systems were minimized in the implicit solvent until the RMSD in the gradient elements was below 10^{-6} kcal mol^{-1} Å^{-1}. These GBSA minimizations were carried out via the *sander* program while the second derivatives of the MM and GBSA energies included in the Hessian matrix were calculated analytically with the *NAB* program.[79] In the two protocols, the translational entropy is 6.4 cal mol^{-1} K^{-1} lower than the entropy value obtained for the standard state of an ideal gas, owing to the change in concentration from 0.045 M (ideal gas) to 1.00 M (solution).

Conformational entropy calculations

Conformational entropies (S_{conform}) were calculated using the CENCALC program.[80] This program selects a set of rotatable dihedral angles using both trajectory coordinates and topology information. It discretizes the time evolution of the selected dihedral angles by evaluating their continuous probability density functions (PDFs) represented by a von Mises kernel density estimator, which depends on a concentration parameter κ (a κ=0.35 value was chosen here). By finding the maxima and minima of the PDF, the time series containing the values of the corresponding dihedral angle during the MD simulation is transformed into an array of integer numbers labelling the accessible conformational states which, in turn, is processed to estimate the rate of conformational change[54] along the MD trajectory.

CENCALC calculates first unidimensional probability mass functions and readily computes the marginal (first-order) conformational entropy of each dihedral. Correlation effects into $S_{conform}$ were estimated using an expansion technique, termed multibody local approximation (MLA),[81] which can be formally expressed as a sum of conditional entropies that resembles the so-called Maximum Information Spanning Tree (MIST) approach.[82] Prior to the MLA calculations, the internal rotations were categorized as having "fast", "medium" or "slow" conformational rates following the prescriptions that have been previously described for cyclodextrin systems.[54] In this way, the MLA entropy calculations included correlation effects between dihedral angles belonging to the same group, leading thus to a significant reduction in the number of conformational degrees of freedom that have to be considered. In addition, the bias of the MLA entropy due to finite sampling was minimized by shuffling the elements of the arrays of integer numbers labelling the conformational states.

Interaction Entropy Calculations

Besides the absolute entropy calculations, we also estimated the interaction entropy (ΔS_{IE}) introduced by Duan et al., which has been proposed as an efficient and reliable measure of the binding entropy for non-covalent complexes.[28] The ΔS_{IE} formula is based on the fluctuations of the gas–phase interaction energy evaluated along the MD trajectory of the complex (i.e., the one trajectory approach):

$$-T\Delta S_{IE} = RT \ln \left\langle exp\left(\frac{\Delta E_{int} - \left\langle \Delta E_{int} \right\rangle}{RT} \right) \right\rangle$$

It must be noted that this classical entropy term is conceptually different from the single-molecule configurational entropy underlying the S_{RRHO} and $S_{conform}$ calculations. We computed the interaction entropy on the ~10000 frames extracted from the MD simulations of the β-CD/ligand complexes to carry out the MM/PB(GB)SA calculations. The interaction energies of each MD snapshot, $\Delta E_{int} = E\left(cmplx\right) - E\left(host*\right) - E\left(lig*\right)$, were computed using the

GLYCAM/GAFF force field with no cutoff and the *sander* program. For the sake of consistency, the $-T\Delta S_{IE}$ term was combined only with the interaction ΔG_{int} terms evaluated with the MM/PB(GB)SA method.

Docking calculations

Ligands were properly placed within the β-CD ring using the LMOD (Low_MODe) method[83] as coupled with the *sander* program, which allows for minimization, conformational searching and flexible docking based on eigenvector following of low frequency vibrational modes. We performed the LMOD calculations in the gas phase using the GLYCAM/GAFF force field and restraining the position of the seven etheric oxygens of the sugar rings to maintain an open cyclodextrin, but allowing the reorientation of the β-CD alcohol groups to interact with the ligands. A total of 10 LMOD iterations were computed by exploring 3 low-frequency vibrational modes. Eigenvectors were recalculated every 3 LMOD iterations. Translation/rotation of the ligands were performed by random variations between 0.0-0.1 Å and 0-180^{o}, respectively. The LMOD calculations generated a total of 10 low energy structures for each cyclodextrin/ligand complex. Inspection of these structures confirmed that, in most cases, the ligands are placed within the hydrophobic cavity. For each complex, the structure with the lowest LMOD energy was then selected to perform single point calculations of the interaction energy.

Statistical uncertainty

The instantaneous values of the MM/PBSA-like energy components and relative differences oscillate rapidly on the sub-nanosecond timescales and exhibit lower-amplitude oscillations on longer timescales. Therefore, besides the standard error (*se*) estimation, we also assessed the statistical uncertainty of the average values by computing the block averaged standard errors of the mean (*be*).[84] Thus, the MD trajectories were divided into segments (‘‘blocks’’) with a block size *M* that ranges from 1 up to a quarter of the total number of frames. We report the

limiting value of *be* versus block size, which may be taken as an upper limit to the statistical uncertainty.

Results

The 1.0-μs length of the MD simulations sampled intensively the conformational space of the β-CD and ligand molecules bound in the inclusion complexes, as well as of the direct or water-mediated binding interactions that stabilize the complexes. The structural and energetic stability of the β-CD/ligand simulations is illustrated in Figures 1 and S1, which show the evolution along the MD trajectory of the separation between the center of mass (CM) of the fragments, the molecular surface, the MM/PBSA interaction energy between β-CD and ligand and the Euler angles characterizing the relative orientation of the principal axis of inertia of the ligand with respect to β-CD. Views of MD snapshots for representative ligands are also represented in Figure 1.

< Figure 1 here >

Depending on the parallel or antiparallel axis orientation, two possible binding modes in the β-CD cavity are usually distinguished that are called the primary and secondary orientation referring to the location of the ligand polar groups at the wide and narrow cavity entrances, respectively. All the simulations were started at an initial structure in which the CM of the two fragments were coincident and their principal axis of inertia were aligned, the initial primary/secondary binding modes being thus assigned randomly. During the simulations, the only geometrical constraint was the harmonic penalty that acted when the distance between the CMs was longer than 6.0 Å. In general, we found that the β-CD inclusion complexes are structurally stable along the conventional MDs and that the impact of the CM restriction was small, acting in less than 1-2 % of the MD frames excepting in a few amine systems (see below).

The typical dynamic behavior of small ligands like butanol is represented in Figure 1. Thus, inspection of the MD snapshots reveals that the ligand molecule undergoes reorientation events so that the polar group of butanol (**oh1** in Figure 1) is placed either at the wide or narrow entrances of the β-CD cavity lined by the secondary and primary β-CD hydroxyl groups, respectively. As a matter of fact, the butanol molecule occupies the hydrophobic cavity (see Figure 1a) and simultaneously changes its relative orientation with respect to the β-CD: essentially all possible orientations are explored as characterized by the uniform distribution of the Euler angles (ϕ, θ, ψ) within their accessible ranges (0-2π, 0-π) all along the trajectory (see the three dial plots in Figure 1a). Bulkier ligands like pinanediol (**oh4**) exhibit a similar dynamic behavior, undergoing frequent "up/down flips" on the ns timescale, although they also exhibit less fluctuations in the CM distance and tend to adopt a parallel alignment between their principal axis and that of the β-CD (see for example the θ,ψ dial plots in Figure 1b). Therefore, it turns out that the 1-μs MD sampling is effectively independent of the initial β-CD/ligand orientation for most of the examined ligands.

A few ligands bearing positively charged ammonium groups gave complexes with β-CD that are structurally less stable as suggested by the larger abundance of MD frames affected by the CM constraint (~5-10%) and the appearance of non-inclusive complexes in the course of the simulations. This is the case of 1*R*,2*R*-pseudoephedrine (**am01**, see Figure 1c) that occupies the β-CD cavity in a secondary binding mode (~90% abundance), but also forms a non-inclusive complex (~10% abundance) characterized by a CM separation above 5 Å and a lower interaction MM/PBSA energy. The non-inclusive structures transform back into inclusive ones after 20-80 ns, probable due in part to the presence of the CM restraint. Other three amine ligands (3,4-dimethoxyphenethylammonium, **am7**; phenethylammonium, **am20**; and 3-methoxyphenethylammonium, **am21**) display also phases of non-inclusive binding during their MD trajectories. For the subsequent energy analysis, the non-inclusive structures have a very

small or negligible contribution to the mean values of the MM/PBSA-like energy components although their presence increments the statistical deviations. Thus, we removed them from the set of MD frames used for energy averaging.

Several complexes exhibited a steady primary binding mode all along the production phase of the MD trajectories, that is, they did not flip between the primary and secondary modes. They correspond to relatively large ligands like nabumetone (**ke3**) and naproxen (**ac3**). The structural stability of their complexes is well observed in the plot of the CMs separation (< 2 Å) and the narrower ranges of variation of the Euler angles (see Figure 1d). To make sure that the observed binding mode of these and other ligands (**am1**, **am6**, **am10**, **am11**, **am13**, and **am18** in Table S1) is not an artefact of the initial β-CD/ligand alignment, we run a second 1.0 μs MD simulation starting from an antiparallel alignment of the respective principal axis of inertia. We found that, for all the ligands considered in the repetition of the simulations, the same primary binding mode was readily adopted along the pressurization phase and kept stable during the production phase. The structural and energetic data rendered by the second simulations matched the data from the first simulations (for the sake of brevity, the results of the second simulations are not reported). Overall, we conclude that the MD sampling approach used for the 57 complexes reveals a varied dynamic behavior of the accommodated ligands (*i.e.*, ranging from ligands that adopt a very stable positioning within the host cavity to other ligands that flip their orientation and/or undergo several exit/entrance events) and provides well equilibrated structures of host-guest inclusion complexes suitable for the end-point energy analyses.

End-point energy calculations using MM methods

As described in the Methods section, we used several protocols for computing the gas-phase energy and the solvation energies contributing to the various MM/PBSA-like scorings. Prior to the energy & solvation calculations, 10000 structures (1000 in the case of the most expensive DFTB3 and RISM calculations) were extracted from the MD trajectories. To remove the statistical noise due to the formation of transient non-inclusive (lateral) complexes, only the structures having an inter-fragment CM separation below 5.0 Å were preserved. This filtering process keeps the large majority (~99%) of the extracted MD snapshots, excepting for a few amine ligands as above mentioned that retain around 85-90%. The average values of the binding (or interaction) energies and their statistical uncertainties are graphically compared with the experimental binding energies in the correlation plots shown in Figure 2. The uncertainties for the experimental binding affinities range from 0.10 to 0.30 kcal/mol as noticed in ref. 25. Table 1 collects various statistical descriptors (determination coefficient, Spearman coefficient, etc.) for all the end-point protocols examined in this work. To better assess the overall and specific performances of the energy and solvation methods, the statistical indexes are segregated according to the charge state of the ligands (neutrals and charged). A large subset comprising 21 closely related alcohol ligands is also distinguished.

< Table 1 and Figure 2 here >

The first correlation plot in Figure 2 corresponds to the MM/PBSA estimates of the binding energy (ΔG_{bind}) computed with the average MM/PBSA energies (*i.e.*, $3RT + E^{MM} + \Delta G_{solv}^{PBSA}$) from the three separate MD simulations (complex, host, guest). In consonance with former MM/PBSA assessments, we find a moderate correlation (R^2~0.47) between the computed affinities and the experimental ones. The linear regression slope of ΔG_{calc} vs ΔG_{exp} is well above 1, which is indicative of overestimation in the computed absolute and relative affinities. However, both the correlation plot showing colored data points and the statistical

measurements for the three categories of ligands (neutrals, alcohols and charged) reveal that the MM/PBSA scoring performs clearly better in the case of the neutral ligands with an acceptable R^2 value of 0.65. In the set of neutral ligands, the most abundant functional groups are hydroxyl and ketone functionalities, although some diversity is also present by including amide, benzene, imidazole and ester functionalities. The degree of correlation improves for the homologous series of 21 alcohol ligands, yielding an R^2 coefficient of 0.81.

For comparative purposes, we adopt the MM/PBSA scoring as a suitable reference in order to assess the effect of other computational choices. The first comparison can be established between the conventional three-average (3A) MM/PBSA binding energy and the one-average (1A) MM/PBSA interaction energy (ΔG_{int}). In Table 1 we see that the correlation between 1A MM/PBSA ΔG_{int} and the experimental data is above that of the 3A MM/PBSA ΔG_{bind} values for the neutral ligands and the alcohol series (*e.g.*, R^2= 0.71 vs 0.65 for neutrals), but its performance for the whole set turns out to be worse (R^2= 0.39 vs 0.47). This suggests that host/ligand reorganization effects may play a minor role in determining the relative affinity of the neutral β-CD ligands, but they may be more important for the charged amine and carboxylate ligands. The average statistical noise expressed in the mean values of the ΔG_{int} is 0.24 kcal/mol, which is below that of ΔG_{bind} data (0.37 kcal/mol) although the difference is not large because of the μs sampling achieved in the simulations. Therefore, we consider that the 3A MM/PBSA protocol would be more robust provided that extensive sampling is available.

We also assessed the quality gain in the affinity ranks yield by the MD sampling by computing single-point MM/PBSA scorings on individual structures of the β-CD/ligand complexes. Three alternative ways of selecting the structures were considered: picking up the representative structure of the most populated cluster according to the β-CD conformation, the most favorable pose predicted by docking calculations using the LMOD method or a random MD snapshot having a interfragment CM distance < 1.0 Å. The statistical measurements

collected in Table 1 indicate that the single-point MM/PBSA calculations capture little correlation for the neutral ligands (R^2= 0.34-0.56) and that the corresponding scorings for the charged complexes are essentially uncorrelated, deteriorating thus the overall performance. The best single-point scorings are obtained on the MD cluster representatives, which seems to confirm the convenience of carrying out MD simulations as a prerequisite to obtain more reliable average MM/PBSA ΔG_{int} data. To assess the influence of the amount of MD sampling, we repeated the MM/PBSA ΔG_{int} and ΔG_{bind} calculations using only the last 10% of the available sampling (~1000 frames in the case of β–CD/ligand complexes). The corresponding statistical measurements indicate that the performance of the ΔG_{int} rankings is hardly affected (R^2 ~0.39) by the reduced sampling while that of the ΔG_{bind} values is slightly deteriorated (*e.g.*, the global R^2 decreases from 0.47 to 0.41). Hence, the extended sampling is more beneficial for the ΔG_{bind} scorings than for the ΔG_{int} ones. Such dissimilar behavior is probably due to the fact that ΔG_{bind} accounts also for the deformation of the β-CD ring upon ligand binding that, in turn, depends on the slow conformational motions of the macrocycle.[54]

The comparison among the MM/PBSA, MM/GBSA and MM/RISM variants can be of particular interest (see Table 1). With respect to the PBSA solvation energies computed with ε_{int}=1, the other alternatives rendered ΔG_{bind} energies that have a weaker correlation with experimental data (similar results were obtained for the one-trajectory ΔG_{int} scorings; data not shown for brevity). For example, choosing an internal dielectric constant ε_{int}=4 gives PB solvation energies that result in a poor global correlation (R^2~0.11) and in worse rakings for neutral and alcohol ligands. The replacement of the non-polar (SA) solvation term of Tan et al. by the explicit solute-solvent vdW energy term plus a cavity term as proposed by Gohlke and Case, results in the MM/PB(vdW) protocol (see Table 1), which leads to comparable results in terms of the various statistical descriptors. Thus, the total R^2 values are similar 0.41 (MM/PB(vdW)) and 0.47 (MM/PBSA), and those for the neutrals are 0.68 and 0.65,

respectively. The two non-polar solvation methods, which treat separately the attractive dispersion and repulsive cavitation contributions, render quite similar results in this test. Since the practical implementation of Tan's approach is simpler as it depends only on the coordinates of the solute atoms, it seems reasonable to adopt it preferably.

The MM/GBSA method constitutes nowadays one of the most favorite end-point energy approaches because of its low computational cost and similar or better performance to that of MM/PBSA for protein-ligand systems.[85] In particular we employed the HCT GB model (*igb*=1 option in the *sander* program). The average ΔG_{bind} MM/GBSA energies have R^2 values of 0.24, 0.44, 0.68 and 0.08 for the full set of ligands, neutral, alcohol and charged ligands, respectively, which are all well below the equivalent MM/PBSA results. On the other hand, the statistical RISM theory of solvation is a more sophisticated (and costly) approach that is presumed to offer a more balanced description of polar and non-polar solvent effects. As described in Methods, we used the RISM implementation available in the *sander* program and obtained averages over 1000 MD frames instead of the 10000 frames used in the MM/PB(GB)SA calculations. This lower number of snapshots has a moderate effect in the statistical uncertainty of the ΔG_{bind} MM/RISM energies, which have a mean statistical error of 0.38 kcal/mol (block error average) similar to that of MM/PBSA. However, the overall performance of MM/RISM in terms of R^2 coefficients, 0.36 (global), 0.60 (neutrals), 0.77 (alcohols), 0.07 (charged), turns out to be below that of MM/PBSA. Thus, in comparison with MM/GBSA and MM/RISM, we see that the MM/PBSA affinity rankings of the β-CD complexes are more reliable at a reasonable computational cost.

End-point energy calculations using SQM methods

Computer hardware improvements, linear scaling techniques and other algorithmic advances have paved the way for the rapid evaluation of energy and molecular properties using SQM methods, typically with a 10x cost (or lower) over that of MM methods. Thus, as mentioned in

the Introduction, the SQM methods are gaining popularity as energy scoring functions that outperform empirical functions applied to rank docking poses.[64, 86] In this work, we calculated the ΔG_{bind} for the 57 β-CD/ligand complexes using the DFTB3 method augmented with dispersion and H-bond corrections and in combination with various solvation methods. These calculations were done over equally-spaced 1000 MD (500 for isolated ligands) snapshots, which were relaxed in the explicit solvent by means of hybrid SQM/MM calculations followed by single-point DFTB3-D3H4 and PBSA/GBSA/COSMO calculations on the solute geometries. Again the averaging of the results was done for the structures having an inter-fragment CM separation below 5.0 Å.

Figure 2b shows the correlation plot between the DFTB3/PBSA ΔG_{bind} and the experimental data. As described in Methods, the DFTB3/PBSA energy is the combination of the gas-phase DFTB3-D3H4 energy with the corresponding PB solvation energy derived from DFTB3 Mulliken charges. The calculated ΔG_{bind} lie in the (-10, -25 kcal/mol) interval and the linear regression slope is close to 2.0, showing thus the DFTB3/PBSA ΔG_{bind} significantly overestimate the experimental binding free energies. Nevertheless, the overall R^2 value, 0.54, indicates a moderate improvement with respect to the MM/PBSA value (0.47). Further gains in the reliability of the DFTB3/PBSA scorings are observed in the reinforced correlation for the subsets: 0.80 (neutrals), 0.91(alcohols) and 0.14 (charged). The predictive ability of the theoretical DFTB3/PBSA ΔG_{bind} values may be better judged in terms of the Spearman rank coefficient which, for example, has moderate/high values of 0.64, 0.85, 0.94 for the total, neutrals and alcohol sets, respectively.

According to data in Table 1, the PM6 COSMO solvation model gives results that are closely similar to those provided by the PBSA method. For example, the DFTB3/COSMO R^2 values are 0.52 (full set), 0.77 (neutrals) and 0.89 (alcohol ligands), slightly lower than the DFTB3/PBSA counterparts. It must be noted, however, that the PBSA solvation calculations

done with the *pbsa* program are faster than the PM6-COSMO calculations done with MOPAC. Taking also into account that in the DFTB3/PBSA scheme only one SQM method is used (DFTB3), it seems that this approach could be a more balanced scoring. Concerning the SQM GBSA solvation energy, the performance of the associated DFTB3/GBSA scoring is also improved with respect to that of MM/GBSA (*e.g.*, the overall R^2 increases from 0.24 to 0.33), but it is clearly less favorable than that of the DFTB3/PBSA scheme. In this way, either MM or DFTB3 methods lead to better results for cyclodextrin systems when they are combined with PBSA solvation.

It may be worth commenting on two technical issues regarding the DFTB3 end-point calculations. First, the preliminary DFTB3/MM relaxation of the solute structures may be avoided as we obtained similar rankings by means of single-point calculations on the unrelaxed geometries of the MD snapshots. For example, the single-point DFTB3/PBSA calculations give ΔG_{bind} data that have R^2 values of 0.51 (global), 0.82 (neutrals) and 0.93 (alcohol ligands), only the overall R^2 is slightly lower than that from the relaxed structures (0.54). This comparison is very much alike for the other DFTB3 scorings (data not shown for brevity). In addition, we also found that the interaction energies, ΔG_{int}, computed over the snapshots of the complex trajectory have also less correlation with the experimental data (DFTB3/PBSA R^2=0.10, 0.58 and 0.81 and 0.01 for the full set, neutrals, alcohol series and charged ligands, respectively). This result, which contrasts with the more similar rankings in terms of the MM/PBSA ΔG_{int} and ΔG_{bind} values, is related to the DFTB3 distortion energy *(i.e.*, $\Delta G_{dis}=G(frag^*)-G(frag)$), especially that of the β-CD molecule, which can vary several kcal/mol.

Although a systematic analysis of the performance of different SQMs is beyond the scope of this work, we decided to examine the performance of the PM6-DH+/PBSA ΔG_{bind} energies (computed on the unrelaxed MD geometries), which have R^2 values of 0.42 (global), 0.66

(neutrals), 0.86 (alcohols) and 0.15 (charged ligands). These values are below the equivalent DFTB3/PBSA R^2 coefficients and similar to the MM/PBSA ones. More particularly, the global MM/PBSA ΔG_{bind} R^2 (0.47) is superior to the PM6-DH+/PBSA one (0.42). Hence, a particular SQM method may or not offer an improved affinity ranking with respect to a MM force field, some previous assessment being thus recommendable.

Entropy corrections

Solute entropy contributions were first estimated by means of energy minimizations and normal mode calculations leading to the RRHO absolute entropies (S_{RRHO}) as well as by conformational entropy calculations using the CENCALC program ($S_{conform}$), We tested two different protocols for obtaining S_{RRHO} as described in Methods. On one hand, solute molecules surrounded by a shell of explicit water molecules were minimized over the set of MD snapshots used for energy calculations. In this approach, the Hessian matrix is then calculated only for the solute atoms using the GLYCAM/GAFF force fields to derive the entropy terms labelled as S_{RRHO}^{MM}. In the second approach, the GBSA model accounts for implicit solvent effects. After removing the coordinates of water molecules, the solutes were relaxed using the MM/GBSA method followed by MM/GBSA normal mode calculations (this gives the $S_{RRHO}^{MM/GBSA}$ terms). Concerning the $S_{conform}$ entropies, they were computed with the MLA method and included correlation effects among sets of dihedral angles clustered by their rate of conformational change as described elsewhere.[54] The $S_{conform}$ estimates are derived from all the available sampling (1.0 μs/0.5 μs for complexes and isolated ligands, respectively) after having filtered the coordinates of MD frames in which the CM distance between β-CD and ligand is above 5.0 Å. The quality of the $S_{conform}$ calculations depends on the degree of convergence of the first-order conformational entropy, which turns out it is not uniform for all the β-CD/ligand simulations. Most of the trajectories render first-order entropy curves that slightly fluctuate

while approaching towards the 1.0 μs limit, although there are other systems that have worse convergence properties (see Figure S3).

< Figure 3 and Table 2 here >

Figure 3 displays correlation plots of the MM/PB(GB)SA and DFTB3/PB(GB)SA scorings augmented with the $S_{RRHO}^{MM/GBSA}$ terms while Table 2 collects the various statistical measurements of the different energy and entropy combinations (e.g., MM/PBSA + $S_{RRHO}^{MM/GBSA}$, MM/PBSA + S_{RRHO}^{MM}+$S_{conform}$, etc.). By comparing the correlation plots in Figures 2-3, we see mixed effects on the affinity rakings upon inclusion of the entropy corrections. Thus, MM/PBSA scorings of the β-CD/ligand affinities slightly improve when the *T*-weighted difference of average RRHO MM/GBSA entropies ($-T\Delta S_{RRHO}^{MM/GBSA}$) is added. The R^2 coefficients of the resulting MM/PBSA+ $S_{RRHO}^{MM/GBSA}$ scoring are 0.48, 0.67, 0.77 and 0.23 for the full set, neutrals, alcohols and charged ligands, very close to those of the simple MM/PBSA scoring (see Tables 1 and 2). These determination coefficients are hardly modified by the replacement of the $S_{RRHO}^{MM/GBSA}$ entropies by the S_{RRHO}^{MM} ones. Curiously, addition of the $S_{conform}$ estimates has a negative impact upon the correlation between theoretical and experimental data, the overall R^2 of the MM/PBSA + $S_{RRHO}^{MM/GBSA}$ + $S_{conform}$ scoring being only 0.36.

The combination of MM/GBSA energies with $S_{RRHO}^{MM/GBSA}$ data substantially reinforces the correlation with experimental data both for the whole set of ligands and for the various subgroups (see Figure 3 and Table 2). From the MM/GBSA ranking to the MM/GBSA + $S_{RRHO}^{MM/GBSA}$ one, the global R^2 coefficient increases from 0.24 to 0.52 and the average RMS error is reduced from 8.0 to 3.2 kcal/mol. This positive effect of the entropy corrections to the MM/GBSA energies is smaller if the S_{RRHO}^{MM} entropies are used instead, the corresponding R^2 value for the MM/GBSA + S_{RRHO}^{MM} scoring being 0.36. Therefore, the details of the RRHO

protocol for the absolute entropy calculations can have a significant influence on the end-point free energy estimations. This is most probably due to the enthalpy-entropy compensation[87] characteristic of binding processes that, perhaps not surprisingly, seems better described by combining the MM/GBSA energy and MM/GBSA RRHO entropy. Concerning the inclusion of the $S_{conform}$ corrections, the global agreement of the MM/GBSA+ $S_{RRHO}^{MM/GBSA}$ +$S_{conform}$ is not improved (R^2=0.30) although the R^2 values for the neutrals and alcohol ligands do not change much with respect to those of the MM/GBSA+ $S_{RRHO}^{MM/GBSA}$ data.

We also tested the interaction entropy corrections ($-T\Delta S_{IE}$) on the MM/PBSA ΔG_{int} energies because this entropy method, which has no additional computational cost, is becoming widely applied.[88] For a given complex, ΔS_{IE} depends on the distribution of energy fluctuations that depends on the amount of sampling. For most of the β-CD/complexes, the $-T\Delta S_{IE}$ convergence plots led to stable entropy values (see Figure S3 in the Supporting Information). For the sake of consistency, the $-T\Delta S_{IE}$ corrections were added to the ΔG_{int} (1A) average energies obtained with the MM/PBSA method (see Table 2 and Figure 2). Inspection of the correlation plot and the statistical measurements indicates that the interaction entropies slightly improve the ranking of the neutral ligands, but have a significant favorable effect on the scorings of the charged ligands (e.g., R^2 increases from 0.07 to 0.47). However, the overall correlation including all ligands is worsened after including the $-T\Delta S_{IE}$ corrections (R^2 varies from 0.39 to 0.26) because the binding entropic penalty for the charged ligands is largely overestimated (see Figure 2 and the linear regression parameters in Table 2). This is not entirely unexpected since the interaction energy between the ionic ligands and β-CDs is larger in absolute value and fluctuates more widely than that of neutral ligands. The improved ranking within the subset of charged amine ligands is probably due to the anticorrelation between the structural and energetic flexibility of the complexed ligands and the stability of the corresponding complex.

Overall, we conclude that the interaction entropy approach yields an unbalanced description of entropy effects for the set of β-CD complexes considered in this work.

According to data in Table 2, the predictive ability of the SQM-based scorings benefits from the inclusion of the RRHO entropies regardless of the solvation method. The R^2 coefficients of the DFTB3/PBSA scoring varies from 0.54 (without $S_{RRHO}^{MM/GBSA}$) to 0.68 (with $S_{RRHO}^{MM/GBSA}$) while those of the DFTB3/GBSA method change from 0.33 to 0.56. The RRHO entropy terms reduce the RMS errors to ~4-6 kcal/mol and confine the absolute ΔG_{bind} values within an interval of [~0, ~ -15] kcal/mol, which still overestimates the ΔG_{exp} values restricted within the [0, -5] kcal/mol range. More particularly, the performance gain in the DFTB3/PBSA+ $S_{RRHO}^{MM/GBSA}$ ranking is quite significant as the global R^2 and Spearman coefficients (~0.7) are clearly higher than the rest of scorings. Such improvements are also evident in the three ligand families: neutrals, alcohol series and charged species, which have DFTB3/PBSA+ $S_{RRHO}^{MM/GBSA}$ R^2 values of 0.85, 0.92 and 0.24. The combination of the DFTB3/PB(GB)SA energies with the S_{RRHO}^{MM} terms is good as compared with the absence of the RRHO term (e.g., R^2 for the DFTB3/PBSA + S_{RRHO}^{MM} amounts to 0.62), but the improvement is less important than that achieved by the MM/GBSA entropy. Hence, it seems confirmed that the $S_{RRHO}^{MM/GBSA}$ calculations capture more efficiently the β-CD and ligand entropy variations. Finally, we found that the addition of the $S_{conform}$ differences has a small negative effect on the global correlation by specifically reducing the already weak correlation observed in the case of the charged ligands. Thus, the DFTB3/PBSA+ $S_{RRHO}^{MM/GBSA}$ + $S_{conform}$ protocol results in R^2 values of 0.66, 0.88, 0.95, 0.16 for the full set, neutrals, alcohol series and charged, respectively, so that the affinity ranking is then improved for the neutrals and alcohol ligands. However, inclusion of the $-TS_{conform}$ estimations is clearly beneficial for getting the computed ΔG_{bind} values closer to the experimental data, the corresponding RMSE values being 2.1-2.6 kcal/mol (~3-4 kcal/mol in the absence of $S_{conform}$).

Similar trends are observed in the case of the DFTB3/GBSA+ $S_{RRHO}^{MM/GBSA} + S_{conform}$ protocol (see Table 2).

The case of amine ligands

The analysis of the statistical measurements in Tables 1 and 2 point out that all the tested end-point methods exhibit a dissimilar performance depending on the subgroup of β-CD ligands being examined. While the computed affinity rankings for the neutral ligands tend to exhibit a moderate or even strong correlation with experiment, the resulting correlation in the scorings of the charged ligands is in general very poor. This set comprises three negatively charged carboxylate ligands and twenty-one positively charged amine ligands, the lack of correlation emerging entirely from the subset of amine ligands.

As mentioned in the Introduction, the inclusion of explicit hydration shells in the MM/PBSA-like calculations on protein-ligand complexes may increase correlation with experiment. To find out if the lack of explicit solvation may negatively affect the scoring of the amine ligands, we performed again the MM/PBSA ΔG_{bind} calculations on the amine ligand subset while including the 3, 10 and 20 closest water molecules to the ammonium groups in each of the frames extracted from the MD simulations. The explicit water molecules, which were selected by the *cpptraj* program, were assigned to the ligand fragment. However, the calculated ΔG_{bind} MM/PBSA values with 3, 10 or 20 waters do not show any improvement (Figure S4) and, therefore, the question arouse whether or not the poor performance on the amine ligands was due to the intrinsic limitations of the MM or SQM methods. To answer it, we calculated the β-CD/ligand interaction energy in the gas-phase for a set of representative snapshots extracted from the MD simulations using MM, DFTB3 and B3YLP-D3/6-31+G** methods. Only one snapshot is taken from each trajectory requiring that its inter-fragment CM distance is minimal

and its MM/PBSA ΔG_{int} matches the average ΔG_{int} of the trajectory. Then the MM or DFTB3 energies are compared with the presumably more accurate B3LYP-D3/6-31+G** ones.

<Figure 4 and Table 3 here>

The correlation plots in Figure 4 clearly show that both the GLYCAM/GAFF force field and the SQM DFTB3 method yield host-guest interaction energies that are quite correlated (R^2~0.90-0.91) with those predicted by the B3LYP-D3/6-31+G** calculations. There are, however, some differences between the MM and DFTB3 energies (see Table 3). In the case of the MM ones, the subset correlations against B3LYP-D3/6-31+G** are quite similar, $R^2 = 0.93$, 0.93 and 0.96 for the amine group, neutral and alcohol series, respectively. The SQM energies exhibit a different behavior depending on the type of ligands as the best R^2 coefficient (0.92) results in the complexes with charged amines, being only 0.68 and 0.84 for complexes with neutrals and alcohol ligands. Nevertheless, these differences do not seem large enough to account for the discrepancy in the MM/PBSA or DFTB3/PBSA scorings between amines and the rest of ligands because data in Figure 4 and Table 3 point out that neither the MM force field nor the DFTB3 method are strongly biased towards one set of ligands or another. Therefore, it may be concluded that other free energy terms like solvation free energy and configurational entropy would be responsible for the varying performance of the MM/PB(GB)SA and DFTB3/PB(GB)SA methods.

Discussion

Clearly, the inclusion complexes formed between cyclodextrins and a wide variety of organic and drug molecules are appropriate reference systems for designing and performing validation studies of end-point free energy methodologies. Thus, our calculations point out that the same reasons that justify the interest of the CD systems in the validation studies of more rigorous free energy methods (*i.e.*, availability of experimental binding free energies, small size, …) apply equally well to the case of MM/PBSA-like methods. In this way, we assess a series of

end-point approaches that, although it is far from being a comprehensive list of the many possible MM/PBSA-like variants, it may constitute a representative sample of MM, SQM, solvation and entropy methods usually combined in the end-point protocols. Furthermore, it has been feasible to carry out extended 1.0 μs MD simulations of the 57 β-CD complexes that help minimize the statistical uncertainty of the various scorings and better discern their inherent performance. Another advantage of the very long conventional MD simulations is that they allow us to estimate the conformational entropy of flexible molecules along the fast (sub *ns*) and slow (supra *ns*) time scales.

Besides the advantages of selecting cyclodextrins as reference systems, our simulations reveal potential disadvantages in such a way that the corresponding host-guest complexes are somehow challenging systems. Thus, the dynamic properties of the β-CD-ligand complexes along the 1.0 μs trajectories underline the frequent and fast reorientation of the small and medium-sized ligands bound within the hydrophobic cavity so that they alternate between the primary and secondary binding sites. Some amine ligands tend also to form short-lived non-inclusion complexes. Furthermore, the β-CD ring in the unbound state experiences conformational changes due to large amplitude motions around the rotatable bonds that expand over time scales from ~10^{-1} ps^{-1} to ~10^{-6} ps^{-1}. The slowest motions can contribute significantly to the configurational entropy, require very long simulations (~5μs) to be properly sampled, and are dampened down to a different extent upon ligand binding.[54] All these factors seem to confirm the convenience of extensive sampling of cyclodextrin complexes for applying either the MM/PBSA-like approaches or the sophisticated free energy methods.

According to data in Tables 1-2, some of the examined choices in the construction of the MM(SQM)/PB(GB)SA scorings may be more effective than others for improving the prediction capacity of the models over heterogeneous host and/or ligand molecules. Thus, we believe that conformational sampling is beneficial even though single-point calculations on

representative cluster structures or docking poses may exhibit significant correlation (up to R^2~0.6) for the neutral and alcohol ligands. The correlation gain of the averaged energies over the single-point ones may seem modest, but it consistently improves the R^2 of the whole set. Similarly, the three-average (3A) ΔG_{bind} values give also a better affinity ranking for the whole set of ligands with respect to that of the one-average (1A) ΔG_{int} interaction energies. These sampling requirements, which certainly are computationally expensive and introduce statistical uncertainty, incorporate both conformational and distortion effects that increase the reliability of the global rankings.

The choice of the solvation method, typically PBSA vs GBSA, is sometimes not clear for a particular application. Of course GBSA energies are cheaper and the literature normally suggests that GB gives comparable or better results than the PBSA-based energies.[85, 89-90] Nonetheless, the present calculations indicate that the PBSA scorings with ε_{int}=1 perform clearly better than the GBSA ones in all the examined situations. Furthermore, the quality of the PBSA scorings in Table 1 is above that obtained from the more elaborated RISM theory and, therefore, we would support the preferential use of PBSA energies in the scoring of other similar host-guest complexes.

As shown in Figure 2, the R^2 value associated to the MM/PBSA calculations, 0.47, is quite moderate, what is mainly a consequence of the poor correlation for the complexes between β-CD and the charged amine ligands. However, the prediction capacity is enhanced when the MM force field is replaced by the SQM DFTB3 method (augmented with dispersion and hydrogen bond corrections). Thus, the DFTB3/PBSA protocol gives an affinity ranking for the neutral ligands with a good correlation with experiment (R^2=0.80) that results in a global R^2=0.54.

As mentioned in the Introduction, the role played by entropy corrections in the end-point scorings is assessed in terms of the absolute RRHO and conformational entropies of the solute

molecules. The actual impact of the RRHO entropies depends on the details of the normal mode calculations (MM vs MM/GBSA) and on which energy and solvation methods are used. Hence, we find that the goodness of the MM/PBSA energies hardly changes upon addition of $S_{RRHO}^{MM/GBSA}$ or S_{RRHO}^{MM} whereas the MM/GBSA scoring reaches a more acceptable R^2=0.52 value upon addition of the $-TS_{RRHO}^{MM/GBSA}$ entropies. Moreover, the same entropy terms significantly augment the prediction ability of the DFTB3-based scorings, which reaches a moderately good R^2=0.68 value at the DFTB3/PBSA+ $S_{RRHO}^{MM/GBSA}$ level. This disparity reflects again that unbalanced descriptions of enthalpy and entropy terms may arise when combining energy/solvation averages over classical MD ensembles with absolute RRHO entropies. On the other hand, the effect of the conformational entropy corrections, not accounted for by the vibrational contributions to the average RRHO entropies, is diverse. The $S_{conform}$ terms significantly reduce the RMSE values and enhance the correlation between calculated and experimental data for the neutrals and alcohol series of ligands, but deteriorate it in the case of amine and carboxylate ligands. In this respect, we note that the $S_{conform}$ entropy calculations display worse convergence in the case of the amine ligands (see Figure S3). As a matter of fact, either configurational or conformational entropy calculations are extremely difficult to converge and may require multi-microsecond MD sampling for peptide and cyclodextrin systems.[54, 91] Therefore, special attention should be paid to convergence issues if $S_{conform}$ estimations are to be considered.

A well-known weakness of the MM/PBSA-like calculations is that their performance can depend critically on the chemical systems being studied. Certainly, the end-point methodologies likewise any other theoretical method can benefit from error cancellation when studying homologous series of host-guest or protein-ligand complexes. Apparently, they are more prone to show striking performance drops as that observed here in the case of the positively charged amine ligands, whose best scoring amounts only to R^2=0.23. Such limitation cannot be attributed to inaccuracies of the underlying MM or DFTB3 methods as shown before

(see Figure 4) and does not seem to be due to the end-point averaging protocol either. In fact the same problem is found by Levy and coworkers in their BEDAM calculations[33] using their implicit solvent model (AGBNP2) as they note that alkyl-ammonium ligands account for a large part of the deviations between computed and experimental finding. In contrast, the free energy calculations in explicit solvent reported by Gilson and coworkers[35] include 10 congeneric amine ligands in their reference set for which excellent R^2 values of 0.94-0.95 are computed using various force field combinations. Therefore, we consider that the bad performance of the MM/PBSA-like and BEDAM calculations for the amine group is most likely due to the limitations of the implicit solvent models (PBSA, GBSA, and RISM) in order to properly describe the desolvation effects of the ammonium ligands (a similar observation has been noticed in the BEDAM paper).

Let us briefly compare the overall performance of the best end-point protocols in this study, DFTB3/PBSA+ $S_{RRHO}^{MM/GBSA}$ (R^2=0.68, r_S=0.75, RMSE=3.8 kcal/mol) and DFTB3/PBSA+ $S_{RRHO}^{MM/GBSA} + S_{conform}$ (R^2=0.66, r_S=0.75, RMSE=2.2 kcal/mol), with previous binding energy calculations on CD systems. For the same set of 57 ligands proposed by Levy and coworkers, [33] the computational model based on the BEDAM method turns out to be less reliable (R^2=0.43, r_S=0.67, RMSE=1.44 kcal/mol), showing thus that, at least for inclusion complexes, the SQM end-point rankings can be competitive with absolute binding free energies derived from more rigorous free energy methodologies. This seems further confirmed when comparing the DFTB3/PBSA+entropy correlation measurements with those of the thermodynamic free energy calculations using the attach-pull-release technique carried out on a set of 43 α-CD/β-CD complexes: the overall R^2 values range between 0.47 and 0.60 depending on the force field details (RMSE values from 0.8 to 1.8 kcal/mol)[3] The scoring in terms of R^2 determined by the DFTB3/PBSA+ entropy protocol compares favorably with the binding free energies computed by Tang and Chang for a set of 7 β-CD complexes with aliphatic/aromatic alcohols, aspirin

and a small ester.[37] In particular, these authors employed extended simulations in aqueous solvent (1-11 μs depending on the ligand structure) to sample association and dissociation events from which they determined the association and dissociation rate constants. The resulting ΔG_{bind} energies are strongly correlated with experimental data (R^2=0.94) and are highly accurate (RMSE=~0.3 kcal/mol). Nonetheless, the DFTB3/PBSA+ $S_{RRHO}^{MM/GBSA}$ R^2 coefficients for the congeneric series of 21 alcohols is similar (0.92) and amounts to 0.85 for the larger set of neutral ligands. The present results are also in agreement with the recent evaluation of several MM/GBSA models for affinity calculations on β-CD molecules, which report poor correlation between experimental and calculated energies (R^2~0.2-0.3, RMSE~14-20 kcal/mol) either on docking poses or by means of potential of mean force calculations in implicit solvent.[36] These results are similar to our MM/GBSA ΔG_{bind} calculations without entropy corrections that are also weakly correlated (R^2=0.24, RMSE=8.0 kcal/mol) although they can be improved upon the addition of the MM/GBSA RRHO entropies (R^2=0.52, RMSE=3.2 kcal/mol).

Finally, we note that the fact that end-point formulations are physically-based (and computationally-expensive) scoring functions more useful for ligand ranking rather than for predicting absolute binding energies, is confirmed by our ΔG_{bind} estimations that overestimate the experimental binding resulting in RMSE errors of various kcal/mol. This weakness, which has been repeatedly noticed in the literature,[9, 36] is most likely the consequence of some inconsistencies and missing contributions in the MM-PBSA-like enthalpy/entropy terms (*e.g.*, explicit solvent used in structural sampling vs implicit solvent used in solvation energy calculations, lack of explicit solvent effects, classical MD simulations vs absolute entropies, etc.). We found, however, that the SQM and PB methods complemented with entropy corrections produce ΔG_{bind} energies with RMSE values of ~2-3 kcal/mol, which are not far from that achieved in the BEDAM free energy calculations on the same data set (1.4 kcal/mol).

Nevertheless, the end-point energy scorings would be valuable as long as the resulting affinity ranking is reliable. In this respect, the assessments contained in this work indicate that the R^2 indexes produced by some end-point methods are close to or even better than those of free energy methodologies. The validity of such conclusion depends on the availability of extended sampling on the complex and the separate fragments, the adoption of SQM and PBSA methods, the inclusion of entropy corrections, etc., and is also restricted to the case of host-guest complexes that are relatively small systems and lack metal ions.

CONCLUSIONS

On the basis of the computational data for inclusion complexes involving β-CD and 57 structurally-different ligands, we can draw the following conclusions about the performance of a wide variety of MM/PBSA-like protocols aimed to rank non-covalent complexes in terms of their stability. First, the β-CD/ligand complexes are suitable reference systems provided that long MD simulations in explicit solvent are performed, not only to minimize the statistical uncertainty of the energy scorings, but also to sample the slow conformational motions that can affect deformation and entropic terms. Concerning the various settings and choices made in the energy and entropy calculations, the ΔG_{bind} energies derived from three independent MD trajectories (complex, host and ligand) show consistently a better correlation with the experimental data than the ΔG_{int} energies derived from a single MD trajectory or from single-point calculations on representative structures. The PBSA method also improves the correlation with experiment with respect to that of other implicit solvent methods (GBSA and RISM). Similarly, the replacement of the MM force field by DFTB3-D4H4, which combines the SQM DFTB3 Hamiltonian with empirical dispersion and H-bond corrections and introduces polarization effects, results in DFTB3-based rankings that tend to be more reliable. Addition of interaction entropy terms overestimate the entropic penalty of binding for the charged

ligands. The RRHO absolute entropies can have a favorable impact upon the global quality of the results, but their actual effect depends on the particular choice of energy/solvation/entropy settings. In this respect, the RRHO entropies calculated with the MM/GBSA method perform better in combination with MM/GBSA or DFTB3/PBSA energies. In addition, conformational entropy estimations may slightly improve correlation although they may be affected by poor convergence. Altogether, the statistical measurements of the 30 MM(SQM)/PB(GB)SA protocols also confirm two major shortcomings of these approaches: the degree of correlation can be much worse for some ligands than others (e.g., the amine ligands considered in this work) and the computed absolute/relative binding free energies have large errors (e.g., the lowest RMSE observed amounts to 2.2 kcal/mol). It is suggested that these deficiencies arise mainly from limitations in the implicit solvent models and enthalpy/entropy imbalances. Nonetheless, in terms of the affinity rankings, the best end-point protocol, which combines DFTB3 energies, PBSA solvation and RRHO entropy corrections, has a determination coefficient R^2=0.66 that is close or even better than that of rigorous free energy methodologies.

Acknowledgements

This research was supported by the GRUPIN14-049 (FICyT, Spain) grant.

Supporting Information

Table collecting the structural formulas of the 57 ligand molecules. Figure showing the time evolution of structural and energetic properties for the β-CD/ligand complexes. Convergence plots of the *T*-weighed conformational entropy and interaction entropy calculations. Figure showing the correlation plots of MM/PBSA energies with explicit waters. Ligand parameter files in the mol2 and frcmod formats.

References

1. Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A., Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. *Biophys. J.* **2004,** *86*, 67-74.

2. Hansen, N.; van Gunsteren, W. F., Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014,** *10*, 2632-2647.

3. Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A., The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophys. J.* **1997,** *72*, 1047-1069.

4. Singh, N.; Warshel, A., Absolute Binding Free Energy Calculations: On the Accuracy of Computational Scoring of Protein-Ligand Interactions. *Proteins* **2010,** *78*, 1705-1723.

5. Gutiérrez-de-Terán, H.; Aqvist, J., Linear Interaction Energy: Method and Applications in Drug Design. *Methods Mol. Biol.* **2012,** *819*, 305-323.

6. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000,** *33*, 889-897.

7. Homeyer, N.; Gohlke, H., Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. *Molecular Informatics* **2012,** *31*, 114-122.

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

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

10. Wang, C.; Greene, D. A.; Xiao, L.; Qi, R.; Luo, R., Recent Developments and Applications of the MM/PBSA Method. *Front. Mol. Biosci.* **2018,** *4*. 10.3389/fmolb.2017.00087.

11. Díaz, N.; Sordo, T. L.; Suárez, D.; Méndez, R.; Villacorta, J. M.; Simón, L.; Rico, M.; Jiménez, M. Á., Assessing the Protonation State of Drug Molecules: The Case of Aztreonam. *J. Med. Chem.* **2006,** *49*, 3235-3243.

12. Díaz, N.; Suárez, D., Molecular Dynamics Simulations of Matrix Metalloproteinase 2: Role of the Structural Metal Ions. *Biochemistry* **2007,** *46*, 8943-8952.

13. Kaukonen, M.; Soderhjelm, P.; Heimdal, J.; Ryde, U., QM/MM-PBSA Method to Estimate Free Energies for Reactions in Proteins. *J. Phys. Chem. B* **2008,** *112*, 12537-12548.

14. Sharp, K.; Honig, B., Electrostatic Interactions in Macromolecules: Theory and Applications. *Ann. Rev. Biophys. Biophys. Chem.* **1991,** *19*, 301-332.

15. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G., Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *J. Phys. Chem.* **1996,** *100*, 19824-19839.

16. Tan, C.; Tan, Y.-H.; Luo, R., Implicit Nonpolar Solvent Models. *J. Phys.Chem. B* **2007,** *111*, 12263-12274.

17. Gohlke, H.; Case, D. A., Converging Free Energy Estimates: MM-PB(GB)SA Studies on the Protein–Protein Complex Ras –Raf. *J. Comput. Chem.* **2003,** *25*, 238 –250.

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

19. Kovalenko, A., Multiscale Modeling of Solvation in Chemical and Biological Nanosystems and in Nanoporous Materials. *Pure Appl. Chem.* **2013,** *85*, 159-199.

20. Maffucci, I.; Contini, A., Explicit Ligand Hydration Shells Improve the Correlation between MM-PB/GBSA Binding Energies and Experimental Activities. *J. Chem. Theory Comput.* **2013,** *9*, 2706-2717.

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

22. Suárez, D.; Díaz, N., Direct Methods for Computing Single-Molecule Entropies from Molecular Simulations. *WIREs Comput Mol Sci* **2015,** *5*, 1-26.

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

24. Ben-Shalom, I. Y.; Pfeiffer-Marek, S.; Baringhaus, K.-H.; Gohlke, H., Efficient Approximation of Ligand Rotational and Translational Entropy Changes Upon Binding for Use in MM-PBSA Calculations. *J. Chem. Inf. Model.* **2017,** *57*, 170-189.

25. Karplus, M.; Ichiye, T.; Pettit, B. M., Configurational Entropy of Native Proteins. *Biophys. J.* **1987,** *52*, 1083-1085.

26. Andricioaei, I.; Karplus, M., On the Calculation of Entropy from Covariance Matrices of the Atomic Fluctuations. *J. Chem. Phys.* **2001,** *115*, 6289.

27. Chang, C.-E.; Chen, W.; Gilson, M. K., Evaluating the Accuracy of the Quasiharmonic Approximation. *J. Chem. Theory Comput.* **2005,** *1*, 1017-1028.

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

29. Klett, J.; Núñez-Salgado, A.; Dos Santos, H. G.; Cortés-Cabrera, Á.; Perona, A.; Gil-Redondo, R.; Abia, D.; Gago, F.; Morreale, A., Mm-Ismsa: An Ultrafast and Accurate Scoring Function for Protein–Protein Docking. *J. Chem. Theory Comput.* **2012,** *8*, 3395-3408.

30. Aldeghi, M.; Bodkin, M. J.; Knapp, S.; Biggin, P. C., Statistical Analysis on the Performance of Molecular Mechanics Poisson-Boltzmann Surface Area Versus Absolute Binding Free Energy Calculations: Bromodomains as a Case Study. *J. Chem. Inf. Model.* **2017,** *57*, 2203-2221.

31. Crini, G., Review: A History of Cyclodextrins. *Chem. Rev.* **2014,** *114*, 10940-10975.

32. Dodziuk, H., Molecules with Holes – Cyclodextrins. In *Cyclodextrins and Their Complexes.*, Dodziuk, H., Ed. WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, 2006; pp 1-30.

33. Wickstrom, L.; He, P.; Gallicchio, E.; Levy, R. M., Large Scale Affinity Calculations of Cyclodextrin Host-Guest Complexes: Understanding the Role of Reorganization in the Molecular Recognition Process. *J. Chem. Theory Comput.* **2013,** *9*, 3136-3150.

34. Yin, J.; Henriksen, N. M.; Slochower, D. R.; Shirts, M. R.; Chiu, M. W.; Mobley, D. L.; Gilson, M. K., Overview of the Sampl5 Host-Guest Challenge: Are We Doing Better? *J. Comput. Aided Mol. Des.* **2017,** *31*, 1-19.

35. Henriksen, N. M.; Gilson, M. K., Evaluating Force Field Performance in Thermodynamic Calculations of Cyclodextrin Host-Guest Binding: Water Models, Partial Charges, and Host Force Field Parameters. *J. Chem. Theory Comput.* **2017,** *13*, 4253-4269.

36. Zhang, H.; Yin, C.; Yan, H.; van der Spoel, D., Evaluation of Generalized Born Models for Large Scale Affinity Prediction of Cyclodextrin Host–Guest Complexes. *J. Chem. Inf. Model.* **2016,** *56*, 2080-2092.

37. Tang, Z.; Chang, C. A., Binding Thermodynamics and Kinetics Calculations Using Chemical Host and Guest: A Comprehensive Picture of Molecular Recognition. *J Chem Theory Comput* **2018,** *14*, 303-318.

38. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., Open Babel: An Open Chemical Toolbox. *J. Cheminf.* **2011,** *3*, 33.

39. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J Mol Graph Model* **2006,** *25*, 247-260.

40. Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, I., T.E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; L., X.; Kollman, P. A. *Amber 2916*, San Francisco, 2016.

41. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P., Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985,** *107*, 3902-3909.

42. Sauton, N.; Lagorce, D.; Villoutreix, B. O.; Miteva, M. A., Ms-Dock: Accurate Multiple Conformation Generator and Rigid Docking Protocol for Multi-Step Virtual Ligand Screening. *BMC Bioinf.* **2008,** *9*, 184.

43. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Gaussian, Inc.: Wallingford, CT, USA, 2009.

44. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004,** *25*, 1157-1174.

45. 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.* **1993,** *97*, 10269-10280.

46. Neese, F., The ORCA Program System. *WIREs Comput Mol Sci* **2012,** *2*, 73-78.

47. Lindner, K.; Saenger, W., Topography of Cyclodextrin Inclusion Complexes. Xvi. Cyclic System of Hydrogen Bonds: Structure of [Alpha]-Cyclodextrin Hexahydrate, Form (Ii): Comparison with Form (I). *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem* **1982,** *38*, 203-210.

48. Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J., Glycam06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008,** *29*, 622-655.

49. 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.

50. Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013,** *9*, 3878-3888.

51. Le Grand, S.; Götz, A. W.; Walker, R. C., Spfp: Speed without Compromise—a Mixed Precision Model for Gpu Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013,** *184*, 374-380.

52. Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D., Langevin Stabilization of Molecular Dynamics. *J. Chem. Phys.* **2001,** *114*, 2090-2098.

53. 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.

54. Suarez, D.; Diaz, N., Conformational and Entropy Analyses of Extended Molecular Dynamics Simulations of Alpha-, Beta- and Gamma-Cyclodextrins and of the Beta-Cyclodextrin/Nabumetone Complex. *Phys. Chem. Chem. Phys.* **2017,** *19*, 1431-1440.

55. Roe, D. R.; Cheatham, T. E., Ptraj and Cpptraj: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013,** *9*, 3084-3095.

56. Cai, Q.; Hsieh, M.-J.; Wang, J.; Luo, R., Performance of Nonlinear Finite-Difference Poisson−Boltzmann Solvers. *J. Chem. Theory Comput.* **2010,** *6*, 203-211.

57. Kovalenko, A., Molecular Theory of Solvation: Methodology Summary and Illustrations. *Cond. Matt. Phys.* **2015,** *18*, 32601.

58. Sergiievskyi, V.; Jeanmairet, G.; Levesque, M.; Borgis, D., Solvation Free-Energy Pressure Corrections in the Three Dimensional Reference Interaction Site Model. *J. Chem. Phys.* **2015,** *143*, 184116.

59. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G., Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998,** *58*, 7260-7268.

60. Krüger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T., Validation of the Density-Functional Based Tight-Binding Approximation Method for the Calculation of Reaction Energies and Other Data. *J. Chem. Phys.* **2005,** *122*, 114110.

61. Gaus, M.; Cui, Q.; Elstner, M., Dftb3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (Scc-Dftb). *J. Chem. Theory Comput.* **2011,** *7*, 931-948.

62. Gaus, M.; Goez, A.; Elstner, M., Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013,** *9*, 338-354.

63. Kubillus, M.; Kubař, T.; Gaus, M.; Řezáč, J.; Elstner, M., Parameterization of the DFTB3 Method for Br, Ca, Cl, F, I, K, and Na in Organic and Biological Systems. *J. Chem. Theory Comput.* **2015,** *11*, 332-342.

64. Pecina, A.; Haldar, S.; Fanfrlík, J.; Meier, R.; Řezáč, J.; Lepšík, M.; Hobza, P., Sqm/Cosmo Scoring Function at the DFTB3-D3H4 Level: Unique Identification of Native Protein–Ligand Poses. *J. Chem. Inf. Model.* **2017,** *57*, 127-132.

65. Jan, Ř., Cuby: An Integrative Framework for Computational Chemistry. *J. Comput. Chem.* **2016,** *37*, 1230-1237.

66. Aradi, B.; Hourahine, B.; Frauenheim, T., Dftb+, a Sparse Matrix-Based Implementation of the DFTB Method. *J. Phys.Chem. A* **2007,** *111*, 5678-5684.

67. Stewart, J. J. P., Optimization of Parameters for Semiempirical Methods V: Modification of Nddo Approximations and Application to 70 Elements. *J. Mol. Model.* **2007,** *13*, 1173-1213.

68. Klamt, A.; Schüümann, G., Cosmo: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc. Perkin Transactions* **1993,** *2*, 799-805.

69. Stewart, J. J. P. *Mopac2009 Computational Chemistry*, 11.038L; 2009.

70. Pellegrini, E.; Field, M. J., A Generalized-Born Solvation Model for Macromolecular Hybrid-Potential Calculations. *J. Phys.Chem. A* **2002,** *106*, 1316-1326.

71. Korth, M., Third-Generation Hydrogen-Bonding Corrections for Semiempirical Qm Methods and Force Fields. *J. Chem. Theory Comput.* **2010,** *6*, 3808-3816.

72. Becke, A. D., Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988,** *38*, 3098-3100.

73. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988,** *37*, 785-789.

74. *Terachem V 1.9*, PetaChem, LLC See http://www.petachem.com: (2009,2015).

75. Ufimtsev, I. S.; Martinez, T. J., Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009,** *5*, 2619-2628.

76. Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C., Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016,** *116*, 5105-5154.

77. Boys, S. F.; Bernardi, F., The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970,** *19*, 553-566.

78. Kongsted, J.; Ryde, U., An Improved Method to Predict the Entropy Term with the Mm/Pbsa Approach. *J. Comput. Aided Mol. Des.* **2009,** *23*, 63-71.

79. Macke, T. J.; Case, D. A., Modeling Unusual Nucleic Acid Structures. In *Molecular Modeling of Nucleic Acids*, American Chemical Society: 1997; Vol. 682, pp 379-393.

80. Suárez, E.; Díaz, N.; Méndez, J.; Suárez, D., Cencalc: A Computational Tool for Conformational Entropy Calculations from Molecular Simulations. *J. Comput. Chem.* **2013,** *34*, 2041-2054.

81. Suárez, E.; Suárez, D., Multibody Local Approximation: Application to Conformational Entropy Calculations on Biomolecules. *J. Chem. Phys.* **2012,** *137*, 084115.

82. King, B. M.; Tidor, B., Mist: Maximum Information Spanning Trees for Dimension Reduction of Biological Data Sets. *Bioinformatics* **2009,** *25*, 1165-1172.

83. Kolossváry, I.; Guida, W. C., . Low-Mode Conformational Search Elucidated: Application to $C_{39}H_{80}$ and Flexible Docking of 9-Deazaguanine Inhibitors into Pnp. *J. Comput. Chem.* **1999,** *20*, 1671-1684.

84. Grossfield, A.; Zuckerman, D. M., Chapter 2 Quantifying Uncertainty and Sampling Quality in Biomolecular Simulations. In *Annu. Rep. Comput. Chem.*, Ralph, A. W., Ed. Elsevier: 2009; Vol. 5, pp 23-48.

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

86. Yilmazer, N. D.; Korth, M., Enhanced Semiempirical Qm Methods for Biomolecular Interactions. *Comput. Struct. Biotechnol. J.* **2015,** *13*, 169-175.

87. Chodera, J. D.; Mobley, D. L., Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annu. Rev. Biophys.* **2013,** *42*, 121-142.

88. Kohut, G.; Liwo, A.; Bősze, S.; Beke-Somfai, T.; Samsonov, S. A., Protein–Ligand Interaction Energy-Based Entropy Calculations: Fundamental Challenges for Flexible Systems. *J. Phys.Chem. B* **2018,** *122*, 7821-7827.

89. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011,** *51*, 69-82.

90. 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.

91. Suarez, D.; Diaz, N., Sampling Assessment for Molecular Simulations Using Conformational Entropy Calculations. *J Chem Theory Comput* **2014,** *10*, 4718-4729.

Table 1. Statistical measurements comparing the average values of the calculated end-point energies with the experimental binding free energies: The determination coefficient (R^2), the Spearman correlation coefficient (r_S), the root mean square error (RMSE) and the mean signed error (MSE) in kcal/mol, the slope and intercept of linear regression.

Ligands	R^2	r_S	**RMSE**	**MUE**	**Slope**	**Intercept**
	MM/PBSA ΔG_{bind}					
Full set	0.469	0.609	3.761	-3.106	1.714	-1.169
Neutrals	0.646	0.756	3.990	-3.485	1.879	-0.908
Alcohols	0.815	0.857	3.418	-2.880	2.048	0.124
Charged	0.134	0.228	3.422	-2.585	1.046	-2.474
	MM/PBSA ΔG_{int} (1 trajectory approx.)					
Full set	0.391	0.506	6.072	-5.681	1.529	-4.245
Neutrals	0.711	0.789	5.757	-5.461	1.932	-2.729
Alcohols	0.852	0.905	5.210	-4.886	2.102	-1.724
Charged	0.066	-0.002	6.479	-5.983	0.779	-6.517
	MM/PBSA ΔG_{int} (single-point calcs. on MD cluster representatives)					
Full set	0.259	0.494	6.237	-5.029	1.662	-2.491
Neutrals	0.563	0.741	5.950	-4.939	1.989	-0.209
Alcohols	0.521	0.673	5.819	-4.459	2.195	0.016
Charged	0.007	0.110	6.613	-5.154	1.039	-6.597
	MM/PBSA ΔG_{int} (single-point calcs. on LMOD docking models)					
Full set	0.208	0.442	5.403	-2.357	2.214	0.940
Neutrals	0.338	0.571	5.980	-4.520	2.206	-0.981
Alcohols	0.569	0.795	5.157	-3.519	2.725	1.426
Charged	0.004	-0.070	4.489	0.617	0.341	-0.975
	MM/PBSA ΔG_{int} (single-point calcs. on randomly chosen MD structures)					
Full set	0.073	0.313	8.370	-7.439	0.991	-7.462

Neutrals	0.426	0.666	6.686	-6.221	1.685	-4.211
Alcohols	0.444	0.592	6.337	-5.769	1.710	-3.734
Charged	0.004	0.165	10.244	-9.113	0.336	-10.717
	MM/PBSA ΔG_{int} (1 trajectory approx.; using last 10% of MD sampling)					
Full set	0.384	0.507	5.924	-5.485	1.570	-3.936
Neutrals	0.698	0.766	5.554	-5.225	1.948	-2.443
Alcohols	0.854	0.891	5.004	-4.648	2.140	-1.378
Charged	0.084	0.114	6.399	-5.843	0.936	-5.997
	MM/PBSA ΔG_{bind} (using last 10% of MD sampling)					
Full set	0.413	0.534	3.719	-3.067	1.559	-1.548
Neutrals	0.549	0.685	3.903	-3.271	1.789	-0.958
Alcohols	0.748	0.831	3.292	-2.639	1.999	0.226
Charged	0.113	0.245	3.450	-2.786	0.860	-3.124
	MM/PBSA (ε_{in}=4.00) ΔG_{bind}					
Full set	0.115	0.267	22.779	-21.814	2.144	0.115
Neutrals	0.494	0.678	19.869	-19.212	3.468	0.494
Alcohols	0.683	0.851	18.359	-17.720	3.737	0.683
Charged	0.031	-0.052	26.259	-25.392	1.423	0.031
	MM/GBSA ΔG_{bind}					
Full set	0.245	0.422	8.017	-7.701	1.167	-7.248
Neutrals	0.439	0.678	7.417	-7.039	1.652	-5.125
Alcohols	0.677	0.861	6.495	-6.220	1.780	-3.984
Charged	0.076	0.324	8.774	-8.612	0.557	-9.682
	MM/PB(vdW) ΔG_{bind}					
Full set	0.406	0.577	5.067	-4.513	1.668	-2.699
Neutrals	0.677	0.786	4.834	-4.395	1.999	-1.465
Alcohols	0.862	0.929	4.376	-3.933	2.205	-0.477
Charged	0.100	0.223	5.371	-4.676	1.041	-4.577

	MM/RISM ΔG_{bind}					
Full set	0.364	0.529	8.533	-8.119	1.747	-6.089
Neutrals	0.600	0.735	8.506	-8.170	2.073	-5.022
Alcohols	0.768	0.887	7.978	-7.661	2.228	-4.141
Charged	0.073	-0.108	8.569	-8.048	0.981	-8.094
	DFTB3/PBSA ΔG_{bind}					
Full set	0.541	0.637	13.012	-12.617	2.645	-8.151
Neutrals	0.798	0.853	12.821	-12.355	3.356	-5.444
Alcohols	0.906	0.938	12.171	-11.631	3.545	-4.333
Charged	0.145	0.202	13.271	-12.977	1.347	-12.138
	DFTB3/COSMO ΔG_{bind}					
Full set	0.517	0.635	12.328	-11.770	2.902	-6.604
Neutrals	0.767	0.832	12.454	-11.936	3.368	-4.992
Alcohols	0.892	0.916	12.053	-11.446	3.652	-3.841
Charged	0.155	0.106	12.152	-11.540	1.896	-9.375
	DFTB3/GBSA ΔG_{bind}					
Full set	0.333	0.490	16.839	-16.467	2.147	-13.353
Neutrals	0.655	0.776	15.836	-15.404	3.129	-9.161
Alcohols	0.812	0.899	14.794	-14.363	3.267	-7.863
Charged	0.062	0.242	18.126	-17.928	0.812	-18.381
	DFTB3/PBSA ΔG_{bind} (unrelaxed solute geometries)					
Full set	0.512	0.572	9.756	-9.414	2.128	-6.352
Neutrals	0.828	0.853	9.187	-8.797	2.834	-3.419
Alcohols	0.931	0.927	8.883	-8.361	3.170	-2.138
Charged	0.153	0.221	10.487	-10.262	1.087	-10.051
	DFTB3/PBSA ΔG_{int} (1 trajectory approx.)					
Full set	0.097	0.199	9.044	-8.561	0.884	-8.875
Neutrals	0.580	0.703	7.250	-7.001	1.691	-4.975

Alcohols	0.810	0.884	6.450	-6.255	1.839	-3.849
Charged	0.006	-0.115	11.044	-10.705	0.244	-12.532
PM6-DH+/PBSA ΔG_{bind} (unrelaxed solute geometries)						
Full set	0.424	0.574	7.347	-5.229	3.488	1.526
Neutrals	0.662	0.752	8.462	-7.590	3.195	-1.154
Alcohols	0.860	0.879	9.220	-8.293	3.745	-0.420
Charged	0.151	0.262	5.454	-1.982	2.460	1.546

Table 2. Statistical measurements comparing the average values of the MM/PBSA and DFTB3/PBSA energies including entropy corrections with the experimental binding free energies: The determination coefficient (R^2), the Spearman correlation coefficient (r_S), the root mean square error (RMSE) and the mean signed error (MSE) in kcal/mol, the slope and intercept of linear regression.

Ligands	R^2	r_S	**RMSE**	**MUE**	**Slope**	**Intercept**
	MM/PBSA + $S_{RRHO}^{MM/GBSA}$ ΔG_{bind}					
Full set	0.482	0.623	7.792	7.324	2.111	10.342
Neutrals	0.674	0.768	6.093	5.773	1.947	8.550
Alcohols	0.768	0.803	6.296	5.933	2.142	9.209
Charged	0.233	0.374	9.651	9.458	1.247	10.055
	MM/PBSA + S_{RRHO}^{MM} ΔG_{bind}					
Full set	0.482	0.643	7.757	7.410	1.865	9.760
Neutrals	0.653	0.796	6.706	6.439	1.843	8.912
Alcohols	0.767	0.848	7.237	6.977	2.006	9.862
Charged	0.169	0.240	9.005	8.746	1.143	9.093
	MM/PBSA + $S_{RRHO}^{MM/GBSA} + S_{conform}$ ΔG_{bind}					
Full set	0.359	0.503	9.643	9.441	1.336	10.353
Neutrals	0.542	0.654	8.494	8.390	1.203	8.985
Alcohols	0.696	0.684	8.454	8.349	1.427	9.573
Charged	0.106	0.201	11.029	10.885	0.720	10.208
	MM/GBSA + $S_{RRHO}^{MM/GBSA}$ ΔG_{bind}					
Full set	0.524	0.684	3.230	2.729	1.565	0.524
Neutrals	0.681	0.784	2.754	2.219	1.721	0.681
Alcohols	0.837	0.881	3.020	2.594	1.875	0.837
Charged	0.139	0.400	3.789	3.431	0.758	0.139

	MM/GBSA + S_{RRHO}^{MM} ΔG_{bind}					
Full set	0.361	0.546	3.415	2.815	1.319	3.681
Neutrals	0.507	0.698	3.520	2.885	1.617	4.695
Alcohols	0.717	0.839	4.012	3.638	1.738	5.754
Charged	0.088	0.342	3.264	2.720	0.654	1.884
	MM/GBSA + $S_{RRHO}^{MM/GBSA} + S_{conform}$ ΔG_{bind}					
Full set	0.304	0.490	5.021	4.846	0.790	4.274
Neutrals	0.623	0.708	4.919	4.836	0.977	4.768
Alcohols	0.835	0.877	5.057	5.009	1.159	5.465
Charged	0.015	0.115	5.158	4.859	0.231	3.000
	MM/PBSA + ΔS_{IE} ΔG_{int}					
Full set	0.256	0.575	13.157	8.316	4.998	19.171
Neutrals	0.707	0.808	2.715	0.218	2.579	4.849
Alcohols	0.743	0.819	3.063	0.494	2.725	5.44
Charged	0.468	0.527	20.025	19.450	4.298	27.419
	DFTB3/PBSA + $S_{RRHO}^{MM/GBSA}$ ΔG_{bind}					
Full set	0.683	0.752	3.835	-2.186	3.042	3.359
Neutrals	0.855	0.894	4.535	-3.097	3.424	4.013
Alcohols	0.918	0.931	4.627	-2.818	3.640	4.752
Charged	0.235	0.354	2.578	-0.934	1.548	0.391
	DFTB3/PBSA + S_{RRHO}^{MM} ΔG_{bind}					
Full set	0.625	0.709	3.701	-2.100	2.796	2.778
Neutrals	0.826	0.896	4.082	-2.431	3.321	4.376
Alcohols	0.907	0.942	3.947	-1.774	3.503	5.405
Charged	0.180	0.260	3.102	-1.646	1.445	-0.571
	DFTB3/PBSA + $S_{RRHO}^{MM/GBSA} + S_{conform}$ ΔG_{bind}					

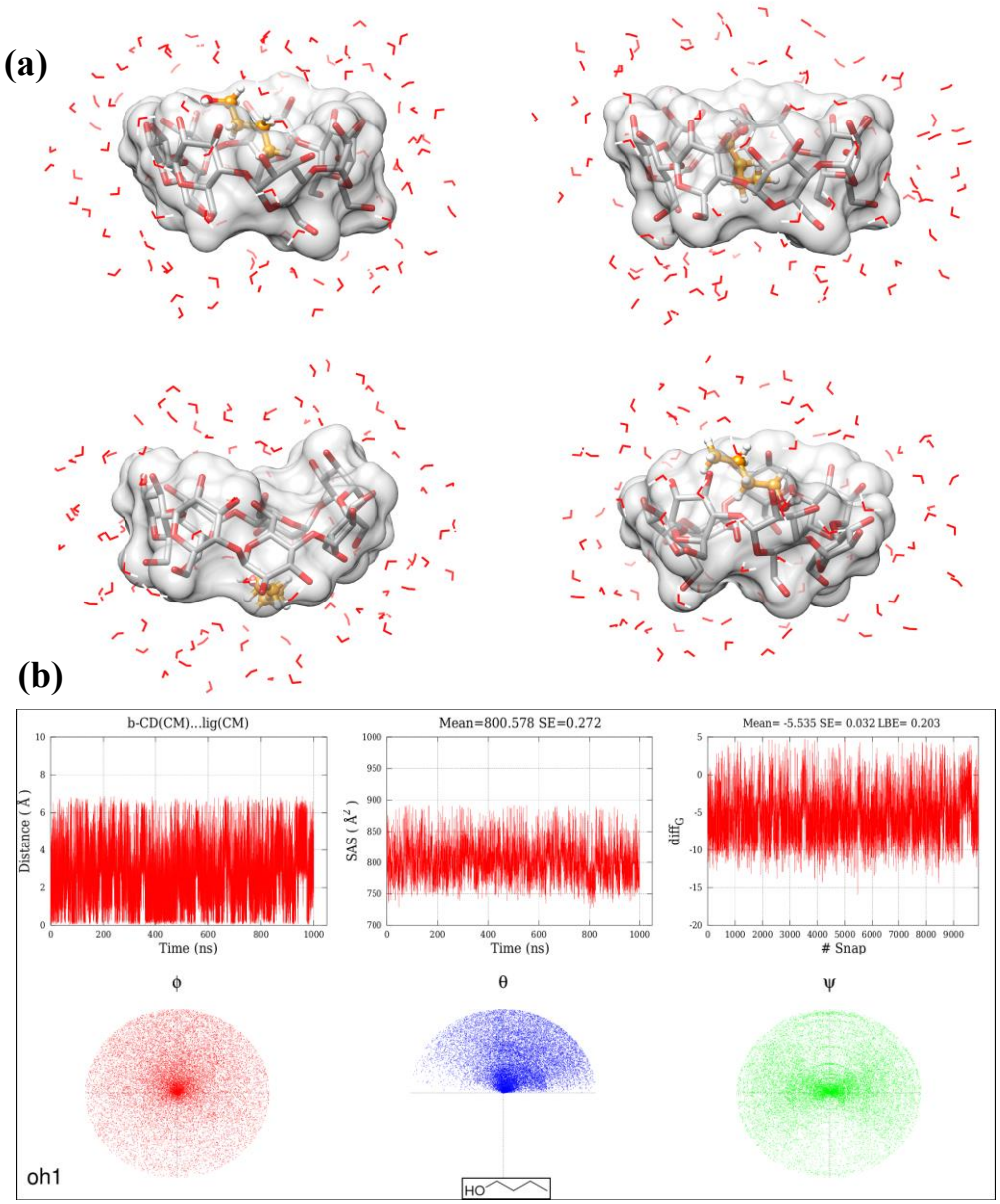
Full set	0.661	0.725	2.231	-0.070	2.267	3.371
Neutrals	0.883	0.927	2.343	-0.480	2.680	4.449
Alcohols	0.946	0.957	2.660	-0.403	2.924	5.115
Charged	0.156	0.223	2.067	0.493	1.021	0.545
DFTB3/GBSA + $S_{RRHO}^{MM/GBSA}$ ΔG_{bind}						
Full set	0.559	0.701	6.725	-6.036	2.544	-1.843
Neutrals	0.799	0.885	6.934	-6.146	3.197	0.297
Alcohols	0.894	0.945	6.498	-5.550	3.362	1.223
Charged	0.099	0.345	6.425	-5.885	1.014	-5.852
DFTB3/GBSA + S_{RRHO}^{MM} ΔG_{bind}						
Full set	0.429	0.598	6.755	-5.950	2.299	-2.424
Neutrals	0.706	0.802	6.459	-5.480	3.093	0.660
Alcohols	0.839	0.908	5.637	-4.506	3.225	1.875
Charged	0.073	0.338	7.143	-6.596	0.910	-6.814
DFTB3/GBSA + $S_{RRHO}^{MM/GBSA} + S_{conform}$ ΔG_{bind}						
Full set	0.436	0.601	4.562	-3.920	1.769	0.436
Neutrals	0.826	0.884	4.144	-3.529	2.453	0.826
Alcohols	0.907	0.943	3.934	-3.135	2.646	0.907
Charged	0.028	0.195	5.082	-4.458	0.486	0.028

Table 3. Statistical measurements comparing the MM and DFTB3 interaction energy in the gas-phase with the B3LYP-D3/6-31+G** experimental binding free energies. The determination coefficient (R^2), the Spearman correlation coefficient (r_S), the root mean square error (RMSE) and the mean signed error (MSE) in kcal/mol, the slope and intercept of linear regression.

Ligands	R^2	r_S	**RMSE**	**MUE**	**Slope**	**Intercept**
MM vs D3-B3YP/6-31+G**						
Full set	0.909	0.915	3.898	3.373	0.891	-3.590
Amines	0.928	0.958	4.372	3.829	1.049	4.473
Alcohols	0.965	0.936	3.496	3.177	1.091	-1.588
DFTB3 vs D3-B3YP/6-31+G**						
Full set	0.902	0.902	18.88	18.277	0.707	10.579
Amines	0.923	0.953	20.809	-20.548	0.838	15.141
Alcohols	0.843	0.842	15.637	14.997	0.333	3.308

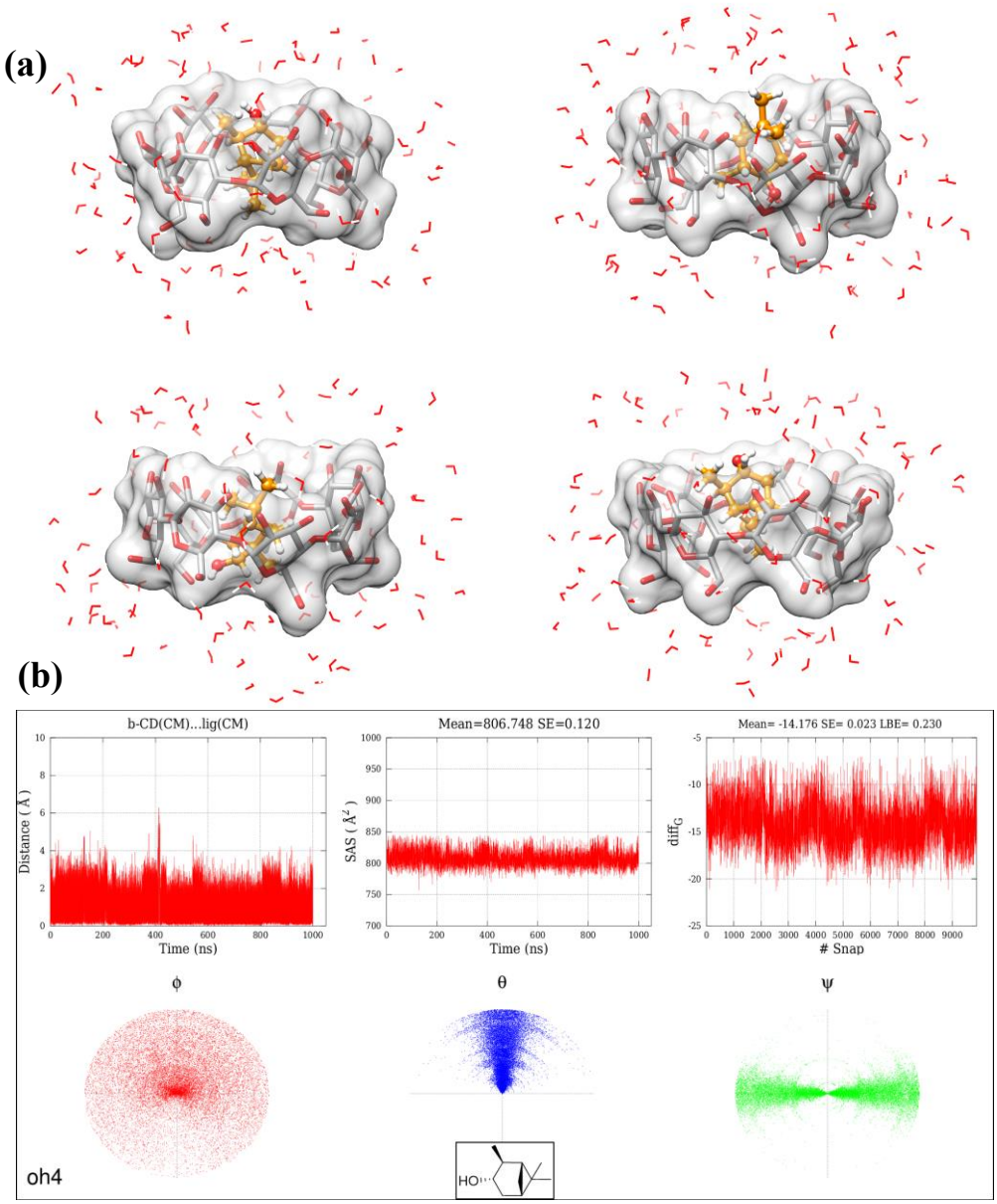
Figure 1 (a) Views of snapshots along the 1.0 μs MD simulations of selected β-CD complexes (β-CD/**am01** snapshots were taken from the second half of the trajectory). Ligands are shown in ball-and-stick models while the β-CD host is shown in stick model and translucent molecular surface. All the frames are superposed to a common reference structure showing the cyclodextrin having its wide entrance upwards. (b) Time evolution of structural and energetic properties. The Cartesian plots display the β-CD···ligand CM separation, molecular surface, and MM/PBSA interaction energy while, in the dial plots of the Euler rotational angles, the radius is the time axis. The Euler angles (*xyx* convention) characterize the relative orientation of two rigid coordinate systems placed at the center of mass of the β-CD and ligand fragments.

Figure 1 (cont.)



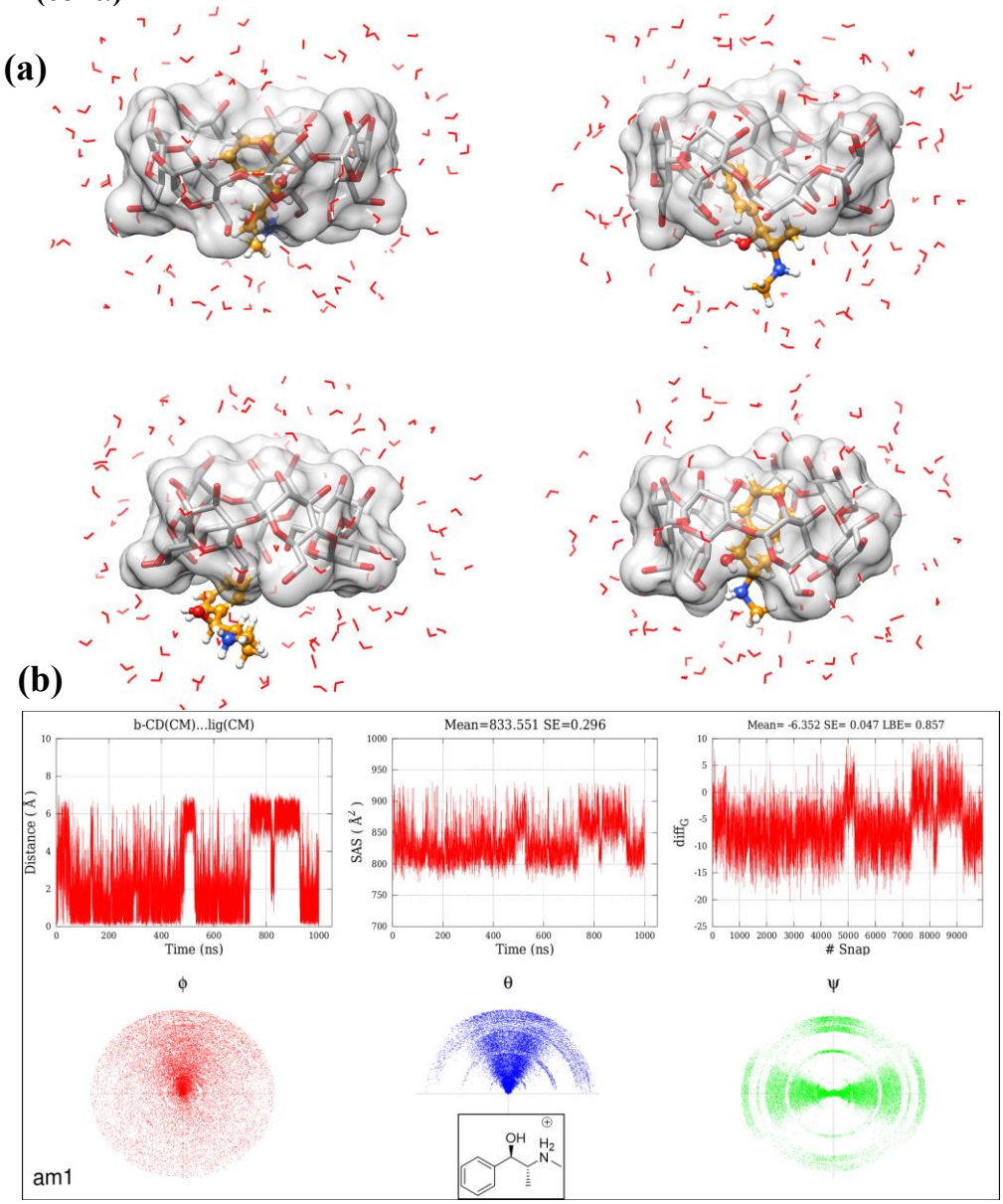
β-CD/butanol (**oh1**)

Figure 1 (cont.)



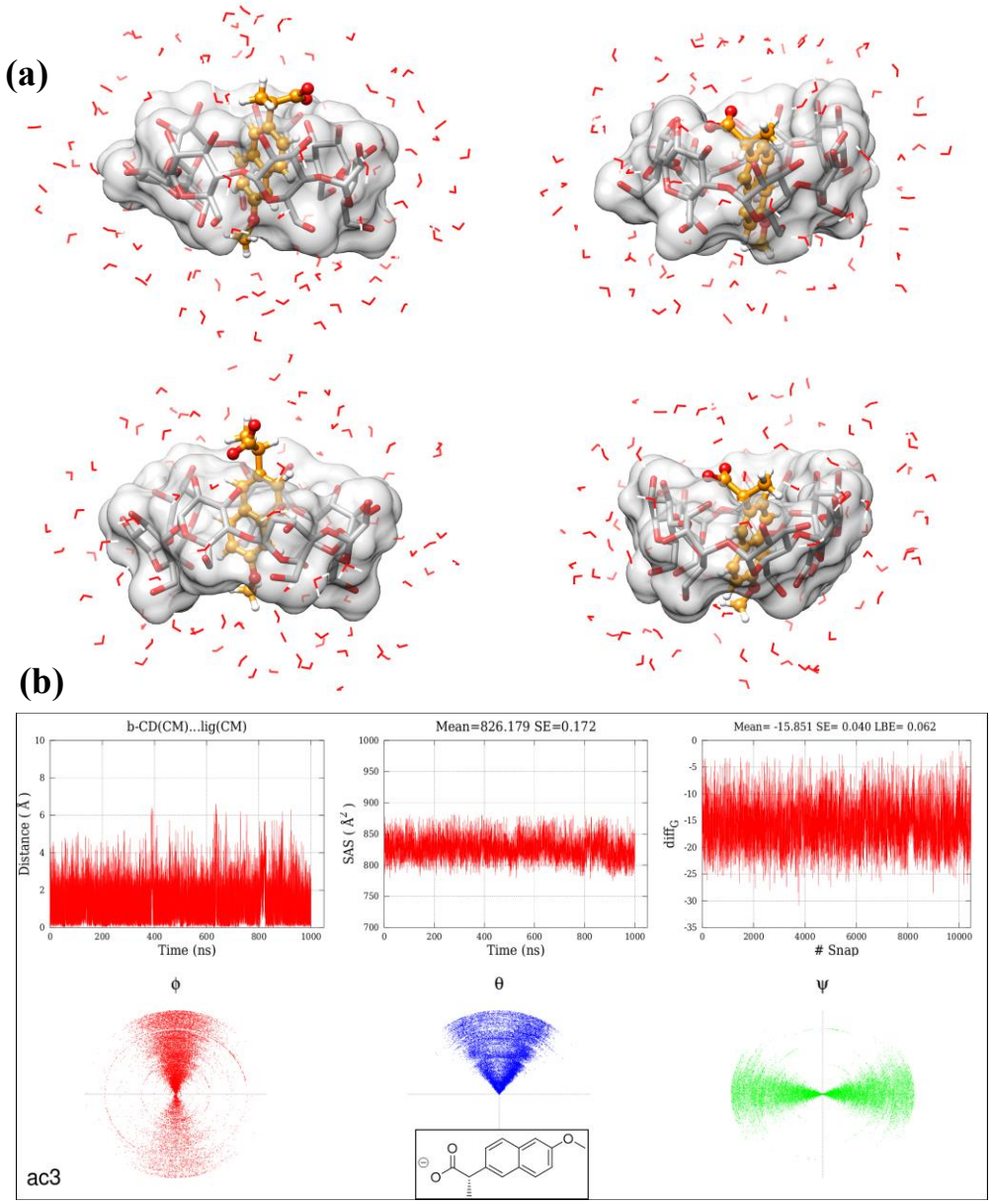
β-CD/1*R*,2*R*,3*S*,5*R*-pinanediol (**oh4**)

Figure 1 (cont.)



β-CD/1*R*,2*R*-pseudoephedrine (**am01**)

Figure 1 (cont.)



β-CD/naproxen (**ac3**)

Figure 2. Comparison between the average end-point relative energies (ΔG_{calc} in kcal/mol) computed with various protocols and the corresponding experimental free energies (ΔG_{exp}). The determination coefficient (R^2), the Spearman correlation coefficient (r_S) and the root mean square error of the relative differences (RMS, in kcal/mol) are also indicated for the whole data. The blue dashed line is the least squared fit line between the calculated and the reference data. The coloring and labels of the data points refer to the category and identity of the ligand bound to β-CD. Vertical bars represent the maximum statistical uncertainty of the ΔG_{calc} values as estimated by block averaging.

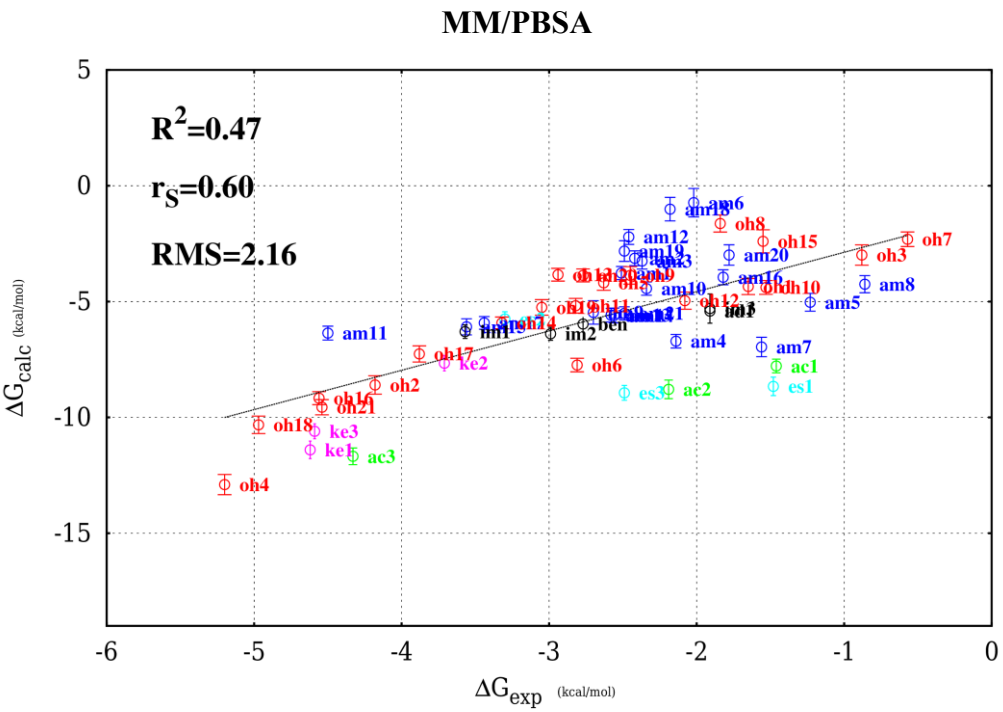
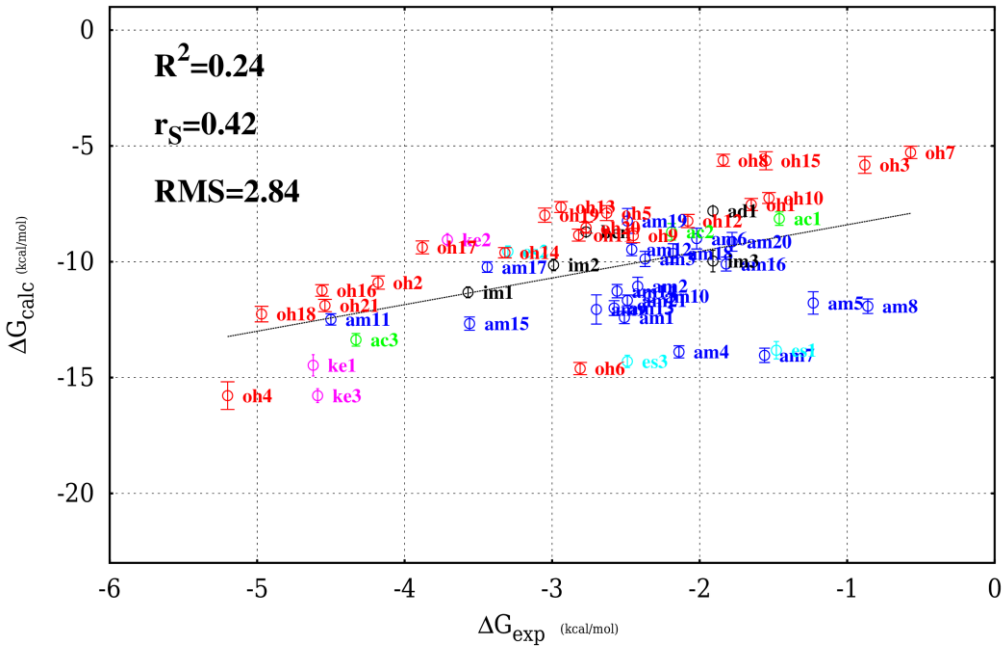


Figure 2 (cont.)

MM/GBSA



DFTB3/PBSA

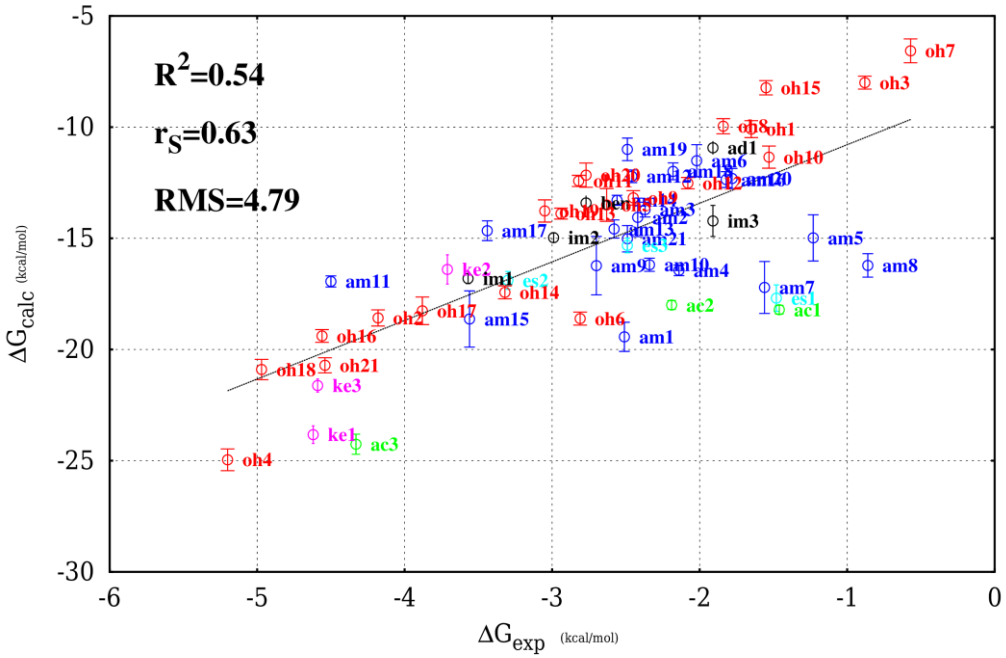


Figure 2 (cont.)

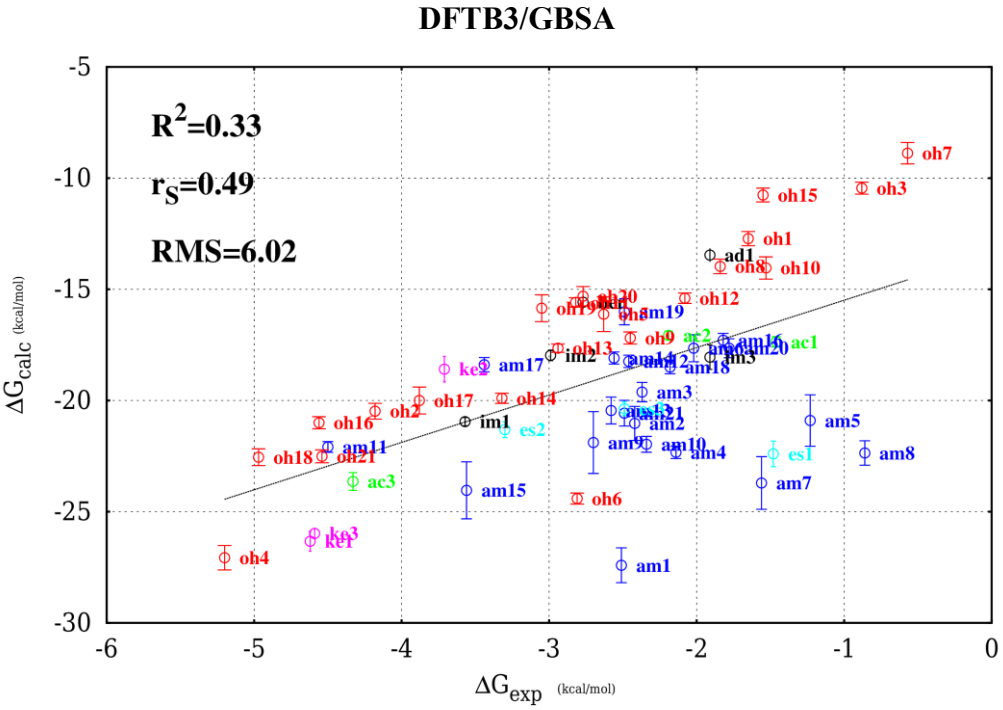
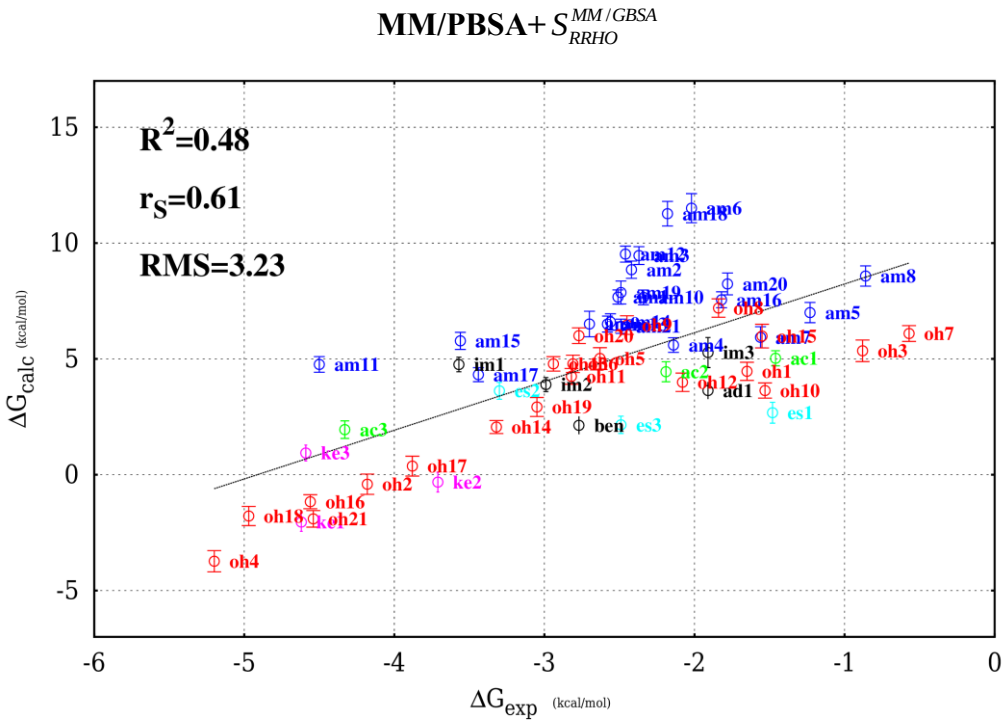
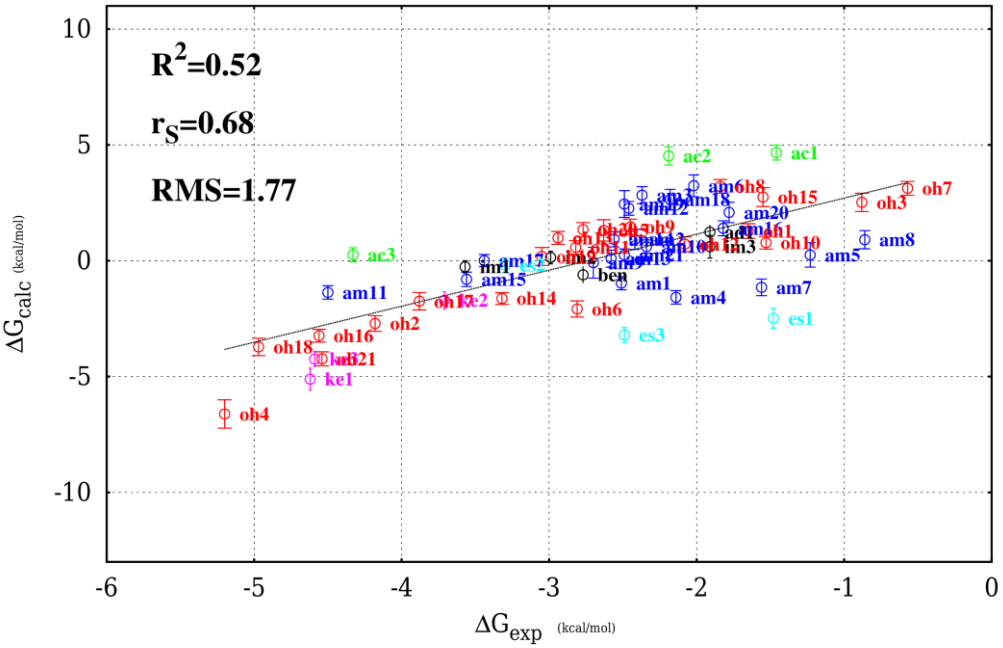


Figure 3 Comparison between the average end-point relative energies including entropy corrections (ΔG_{calc} in kcal/mol) and the experimental free energies (ΔG_{exp}). The determination coefficient (R^2), the Spearman correlation coefficient (r_S) and the root mean square error of the relative differences (RMS, in kcal/mol) are also indicated for the whole data. The blue dashed line is the least squared fit line between the calculated and the reference data. The coloring and labels of the data points refer to the category and identity of the ligand bound to β-CD. Vertical bars represent the maximum statistical uncertainty of the ΔG_{calc} values as estimated by block averaging.



MM/GBSA+ $S_{RRHO}^{MM/GBSA}$



MM/PBSA+ ΔS_{IE}

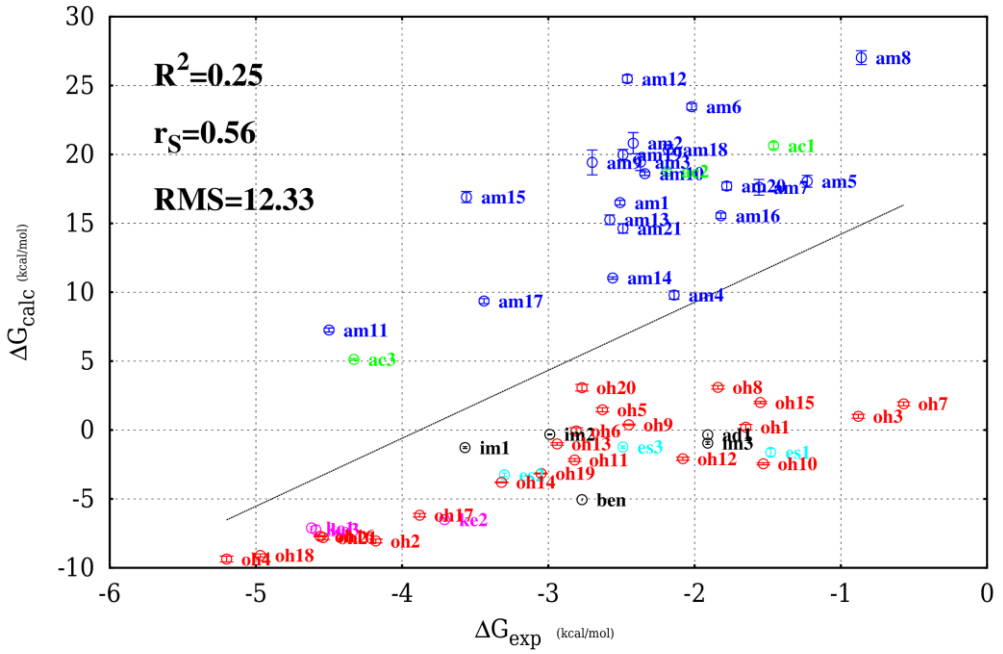


Figure 3 (cont.)

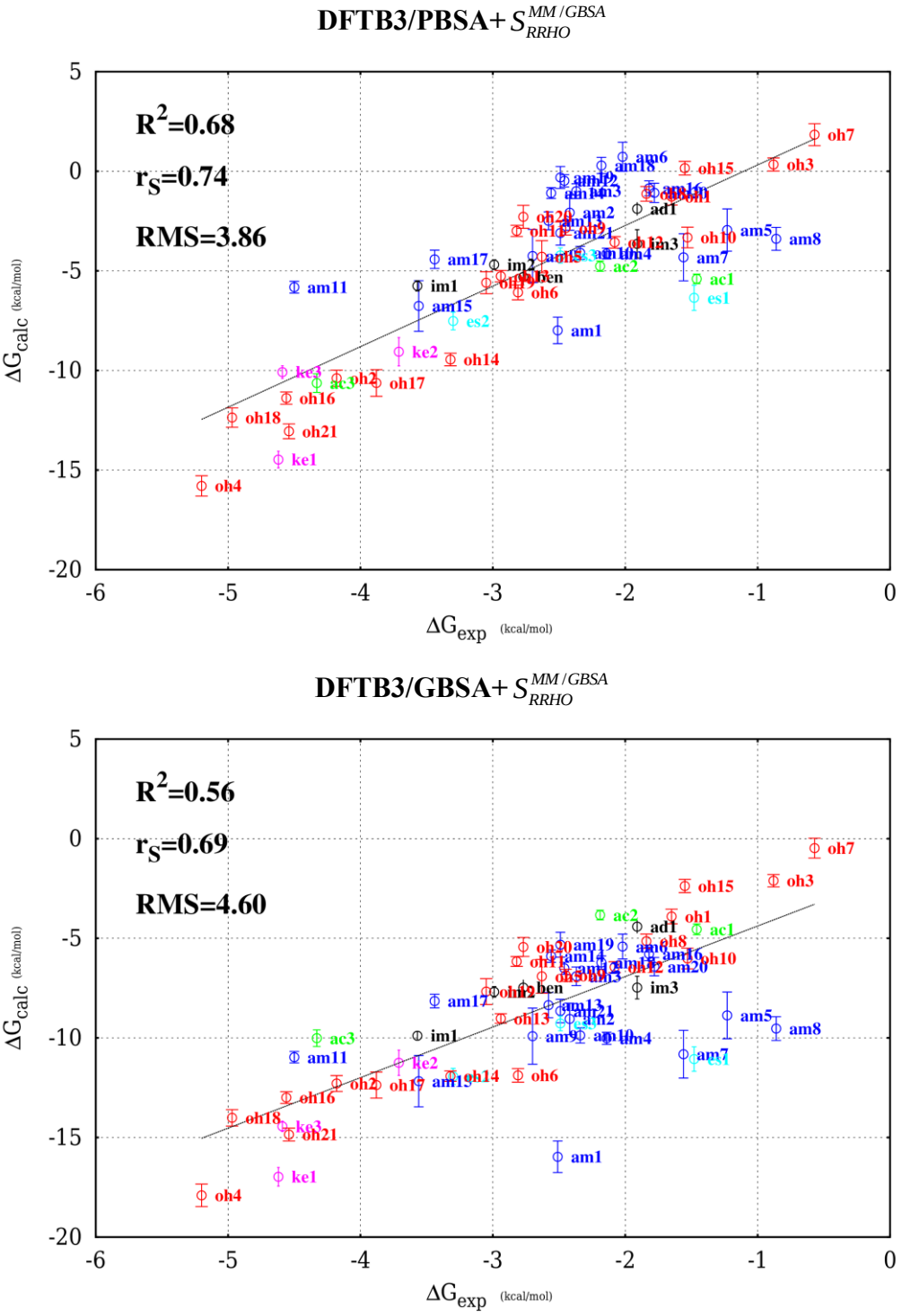


Figure 4. Comparison between the MM and DFTB3 gas-phase interaction energies (in kcal/mol) and the equivalent B3LYP-D3/6-31+G** values (ΔE_{DFT}). The determination coefficient (R^2), the Spearman correlation coefficient (r_S) and the root mean square error of the relative differences (RMS, in kcal/mol) are also indicated.

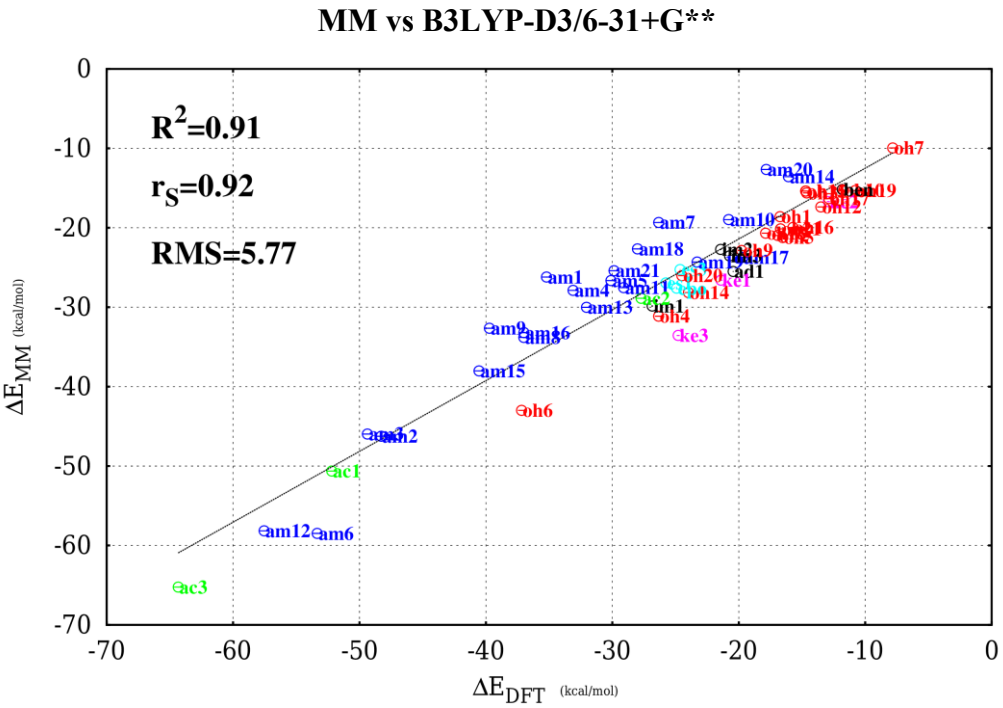
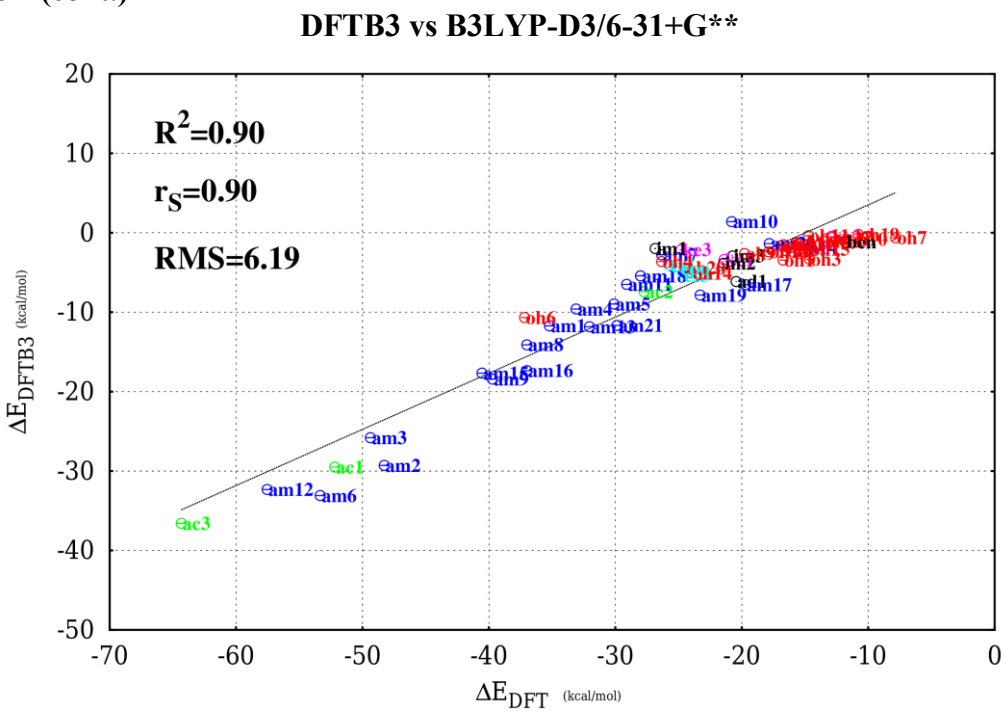


Figure 4 (cont.)



For Table of Contents use only

Affinity Calculations of Cyclodextrin Host-Guest Complexes: Assessment of Strengths and Weaknesses of End-Point Free Energy Methods

Dimas Suárez and Natalia Díaz