Published in final edited form as:

J Comput Aided Mol Des. 2018 October ; 32(10): 983–999. doi:10.1007/s10822-018-0165-3.

Force Matching as a stepping stone to QM/MM CB[8] host/guest binding free energies: A SAMPL6 Cautionary Tale

Phillip S. Hudson^{1,2}, Kyungreem Han¹, H. Lee Woodcock², Bernard R. Brooks¹

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

²Department of Chemistry, University of South Florida, Tampa, Florida 33620, USA

Abstract

Use of Quantum Mechanical/Molecular Mechanical (QM/MM) methods in binding free energy calculations, particularly in the SAMPL challenge, often fail to achieve improvement over standard additive (MM) force fields. Frequently, the implementation is through use of reference potentials, or the so-called “indirect approach”, and inherently relies on sufficient overlap existing between MM and QM/MM configurational spaces. This overlap is generally poor, particularly for the use of free energy perturbation to perform the MM to QM/MM free energy correction at the end states of interest (e.g., bound and unbound states). However, by utilizing MM parameters that best reproduce forces obtained at the desired QM level of theory, it is possible to lessen the configurational disparity between MM and QM/MM. To this end, we sought to use force matching to generate MM parameters for the SAMPL6 CB[8] host-guest binding challenge, classically compute binding free energies, and apply energetic end state corrections to obtain QM/MM binding free energy differences. For the standard set of 11 molecules and the bonus set (including 3 additional challenge molecules), error statistics, such as the root mean square deviation (RMSE) were moderately poor (5.5 and 5.4 kcal/mol). Correlation statistics, however, were in the top 2 for both standard and bonus set submissions (R^2 of 0.42 and 0.26, τ of 0.64 and 0.47 respectively). High RMSE and moderate correlation strongly indicated the presence of systematic error. Identifiable issues were ameliorated for 2 of the guest molecules, resulting in a reduction of error and pointing to strong prospects for future use of this methodology.

Keywords

Host-Guest; Force Matching; Indirect Free Energy; SAMPL6

1 Introduction

Correctly assessing molecular thermodynamic properties is pivotal to rational drug design. Finding the likelihood of a drug to cross the blood brain barrier relies on properly computing its permeability across a membrane. Determining whether a medication can maintain efficacy under acidic conditions (e.g., stomach acid) necessitates calculating acidity

constants. These are just a few examples of chemical phenomena that require careful consideration. Thus, it is of great importance to gauge, in an unbiased manner, the status of computational methods in determining thermodynamic properties. This is the rationale behind the Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) challenge [1–5].

Thermodynamic free energy is a ubiquitous influence throughout chemistry. Hence, accurately computing free energy differences remains a highly sought, albeit elusive, goal. The theory underlying classical Free Energy Simulations (FES) is well established, with various tricks and techniques (such as alchemical approaches and soft-core potentials [6, 7]) that forsake physical intuition and detour through non-physical routes to connect thermodynamic states of interest (for a comprehensive review of the numerous methods to free energy calculation, we point readers to Ref. [8]). Unfortunately, the associated level of theory in which these methods are tenable, such as Molecular Mechanical (MM) additive force fields, is often insufficient concerning energetic rigor (e.g., not accurately describing polarization or charge transfer effects [9]). This is not to say that MM approaches cannot achieve accurate results, more that they have some unavoidable limitations. For example, the CHARMM force field, upon discovery of problems in the proper description of peptide backbone dihedrals, required reparameterizing the entire protein force field, as well as the addition of a spline-based energy correction term [10]. However, as they note, accurately describing conformational degrees of freedom in solution came at the expense of reproducing gas phase properties and pointed to limitations of a fixed charged force field. Many look to polarizable models as a way to overcome inadequacies of classical MM force fields (e.g., having the ability to intrinsically respond to changes in environment/media or lack thereof) to more accurately model chemical systems of interest in a variety of situations, but questions of best practices remain contentious [11].

Ideally, all FES calculations would be performed at a high level (e.g., Quantum Mechanical, QM, or mixed QM and MM, QM/MM) of accuracy, with an appropriate Hamiltonian that can properly describe the desired inter/intramolecular interactions as well as generate sampling in the relevant sections of the high level configurational landscape. This, unfortunately, remains impractical, not only in terms of computational effort, but in regards to forfeiting classical FES (e.g., alchemical) tricks. In principle, it is possible to setup FES calculations between states of interest (e.g., states X and Y) in such a way that, for the most part, is executed classically, but with free energy corrections at the end states of interest as shown in Fig 1.

This methodology, often referred to as the “indirect approach” [12–17], permits practitioners to use whatever methods they desire for the classical free energy difference between states X and Y, $\Delta A_{X \rightarrow Y}^{\text{MM}}$ (i.e., Fig. 1[iii]). Following this, two calculations of $\Delta A^{\text{MM} \rightarrow \text{QM/MM}}$ occur at each end state X and Y (Fig. 1[ii] and Fig. 1[iv] respectively). Then, by taking advantage of the fact free energy is a state function, we arrive at

$$\Delta A_{X \rightarrow Y}^{\text{QM/MM}[i]} = \Delta A_{X \rightarrow Y}^{\text{MM}[iii]} + \Delta A_Y^{\text{MM} \rightarrow \text{QM/MM}[iv]} - \Delta A_X^{\text{MM} \rightarrow \text{QM/MM}[ii]}. \quad (1)$$

The caveat, however, is that the free energy difference between levels of theory requires a single step calculation (e.g., without employing intermediate states) using a computationally inexpensive onesided method, such as Free Energy Perturbation [18] (FEP, Eq. 2),

$$\Delta A^{\text{MM} \rightarrow \text{QM/MM}} = -\frac{1}{\beta} \log \left\langle e^{-\beta \Delta U^{\text{MM} \rightarrow \text{QM/MM}}} \right\rangle_{\text{MM}}. \quad (2)$$

Although this formula is deceptively simple, requiring only low level MM sampling (i.e., $\langle \dots \rangle_{\text{MM}}$ denotes the ensemble average is over configurations collected during an MM simulation) then computing U^{MM} and $U^{\text{QM/MM}}$ of each collected MM configuration, there needs to be sufficient overlap between the MM and QM/MM configurational spaces. Even for small systems, such as alanine and serine dipeptide in gas-phase, configurational mismatch can be rather significant depending on the choice for the low level of theory (e.g., choosing semi-empirical QM over MM can improve overlap) [19–24]. It is possible to employ fast switching non-equilibrium work instead of potential energy differences to overcome large configurational disparity (i.e., use of Jarzynski's equation), but this can also become rather unwieldy if “soft” conformational degrees of freedom are highly dissimilar [21, 23, 25].

Alternatively, instead of traversing long paths in phase space, one can bridge the gap by picking the classical level of theory in such a way that configurational mismatch with the higher level of theory is mitigated. As configurations are governed by forces, it stands to reason that choosing low level MM energetic parameters as to best reproduce forces observed at the target QM/MM level of theory would facilitate convergence of one sided FEP calculations. Often referred to as simply “Force Matching” (FM) [26–34], this method utilizes a configurational training set (e.g., generated at either the desired high level of theory or with classical dynamics) to obtain parameters that give “QM-like” ensembles [24]. Clearly, parameter transferability is forsaken, but the improvement in configurational sampling which in turn allows for facilitating convergence of free energies between levels of theory or reproducing ensemble properties through reweighing justifies the sacrifice [30, 34].

Following this logic, FM was used to generate parameter for host and guest systems for the CB[8] SAMPL6 challenge based on work done using the ForceSolve [35, 36] program [24]. The parameterization approach was modeled similarly to how many popular force fields are implemented (specifically, the CHARMM force field [37]). First, charges are determined based on interaction with solvent (QM implicit solvent for our purposes), Leonnard Jones parameters obtained by analogous assignment of the CHARMM General Force Field [38] atom types, and intramolecular parameters fit using classical gas phase simulations matched to QM forces (with the exception of G13, discussed later). Once a complete parameter set was obtained for all guest and host of interests, we sought to answer two questions: how

well do the parameters perform obtained directly from FM, and does attempting to correct to QM/MM result in an improvement? Free energies computed without applying a correction scheme (i.e., purely classical calculations), were performed using the double decoupling method (DDM) [39, 40], and are detailed in the work published in Ref [41]. However, pursuant to the latter goal of utilizing a correction scheme in the host-guest calculation, the work herein focuses solely on the correction from free energies obtained with a force matched parameter set (i.e., the results of Ref [41]) to QM/MM in a similar manner to Fig 1.

2 Methods

The protocol used follows the flow chart laid out in Fig. 2, for which discussion of each step will follow. The procedure laid out is directly based on an approach utilized for small gas-phase systems [24] (e.g., butane and serine dipeptide), but a similar approach has yielded good results when applied to the AMBER protein force field [34]. Parameterizations were applied individually for each host and guest in the challenge (see Fig. 3), so transferability between systems with similar atom types is no longer applicable (which will be discussed in more detail later in text).

2.1 Non-Bonded (External, $\hat{F}_{\text{EXTE}}^{\text{MM[RESP / CGenFF]}}$) Parameter Assignment

Again, as shown in Fig. 2, first a series of QM calculations were performed in Gaussian16 [42]. For guest G0-G12 and host CB8, an initial B3LYP [43,44] Density Functional Theory (DFT) geometry optimization was obtained, followed by an MP2 [45] geometry optimization on the resulting structure from the DFT geometry optimization (both using the 6-31G(d) basis set [46, 47] and proceeded by frequency calculations to confirm steady state configurations were obtained). All geometry optimization were performed in SMD [48] implicit solvent. Following these optimizations, single point calculations based on the MP2/6-31G(d) densities were performed to obtain the Merz-Kollman electrostatic potential charges, which were then processed through the restrained electrostatic fitting procedure (RESP) to symmetrize charges on equivalent atoms [49–51]. A similar treatment was performed for the G13 molecule using a pseudopotential due to the presence of Pt(II) (see Fig. 3), for which details can be found in Ref. [41]. Once charge determination was completed, LJ parameters were used directly from CGenFF via the ParamChem server [38, 52].

2.2 Bonded (Internal, $\hat{F}_{\text{INTE}}^{\text{MM[CGenFF]}}$) Parameter Assignment for CB8 and G0-G12

Initial internal parameters are also obtained directly from CGenFF via the ParamChem server [38, 52]. Ten ns of gas phase Langevin Dynamics (LD) was performed in CHARMM [37], with a coordinate saving frequency of 1 ps (i.e., saving 10,000 coordinate snapshots total), timestep of 1 fs, temperature of 298.15K, a collision frequency of 5 ps⁻¹, and no tapering of non-bonded interactions occurred. The molecule G13 underwent a slightly different treatment in terms of generation of configurations.

2.3 Parameterization of G13

The molecule G13 presented an extremely unique challenge, as classical parameters for square planar complexes are generally not readily available, particular for a species containing the transition metal platinum (with the exception of literature LJ parameters utilized throughout all calculations) [53]. Charges were assigned using a procedure outlined in Ref [41] as mentioned prior. Instead of attempting to create an initial set of parameters for G13 in the standard CHARMM manner [37] (for which the question of what is the appropriate high level treatment to preserve consistency with the CHARMM forcefield is not a straight forward one), we identified a semi-empirical QM suite that contained parameters for palladium in the CHARMM MNDO97 module [54], using MNDO extended to d orbitals, MNDO/(d) [55]. The fundamental assumption was that the degree of similarity between the palladium and platinum valence structures (i.e., they are both square planar metal complexes and are in the same column of the periodic table) would yield moderate configurational space overlap between the two molecules. This laid the groundwork for generating configurational sampling to be force matched to a more rigorous level of theory for G13. The simulation setup used to create the gas phase configurations for force matching was identical to that used for CB[8] and G0-G12, with the exception of the gas-phase LD simulation length being set to 5 ns, and configurations saved every 0.25 ps (i.e., 20,000 configurations generated in total). The caveat being that for the DFT force calculations of the configurations generated from the palladium substituted G13 system, the Pd would be replaced with Pt.

2.4 Intramolecular ($\hat{F}_{\text{INTE}}^{\text{MM[FM]}}$) Force Matching Procedure

For each of the configurations saved during the LD simulations (10,000 configurations for CB[8]/G0-G12 and 20,000 configurations for G13), gas-phase B3LYP/6-31G(d) force calculations (with G13 being the exception, requiring use of the B3LYP/LANL2DZ [56–58]) were performed via Q-Chem [59]. Again, as laid out in Fig. 2 from each of the DFT forces collected, classical electrostatic and van-der-Waals forces were subtracted out based on the corresponding configuration so that only the bonded degrees of freedom (i.e., bonds, angles, dihedrals, and so on) were fitted. All force matching was performed using the ForceSolve software, which utilizes a Bayesian formalism that acts to minimize a negative log-likelihood, and therefore the associated force residuals [24, 35, 36, 60, 61].

2.5 Correction Scheme

Given the parameters resulting from FM, Hamiltonian replica exchange [62] in conjunction with Bennett's Acceptance Ratio (BAR) [63] was used to compute binding free energies (ΔA_{bind}) via DDM (see Ref. [41] for details on the classical binding calculations). As mentioned prior, the goal of this work is to implement a "correction" scheme, that allows for indirect calculation of the binding free energy at a QM/MM level.¹ Since the desired level of

¹Herein, MM[FM] (which denotes MM with bonded parameters obtained from force matching, charges from RESP(SMD), and LJ from CGenFF) will be referred to simply as FM, and QM will be a placeholder for either B3LYP/LANL2DZ (for guest G13) or B3LYP/6-31G(d) for all other guest/host

theory is DFT, we consider only one sided approaches. More specifically, FEP is used to compute all calculations between FM/MM and QM/MM levels of theory.

Following the logic of Fig. 1 allows us to arrive at the cycle illustrated in Fig. 4. Because iv is determined (see Ref. [41] for details on this), only two free energy differences between levels of theory are required: the FM/MM to QM/MM correction of the guest ligand (Lig) in solution (Fig. 4[ii]),

$$\Delta A(\mathbf{Lig}^{FM \rightarrow QM})_{sys}^{MM} = -\frac{1}{\beta} \log \langle \exp[-\beta \Delta U(\mathbf{Lig}^{FM \rightarrow QM})_{sys}^{MM}] \rangle_{FM/MM} \quad (3)$$

and the FM/MM to QM/MM correction of the ligand bound to CB[8] (Fig. 4[iii])

$$\begin{aligned} \Delta A(\mathbf{Lig}^{FM \rightarrow QM}, \mathbf{Host}^{FM})_{sys}^{MM} = \\ -\frac{1}{\beta} \log \langle \exp[-\beta \Delta U(\mathbf{Lig}^{FM \rightarrow QM}, \mathbf{Host}^{FM})_{sys}^{MM}] \rangle_{FM/MM}, \end{aligned} \quad (4)$$

where only the ligand changes energetic description (i.e., switching from a classical to QM Hamiltonian with only the guest ligand in the QM region).

Although it is expected that the FEP correction calculations will converge moderately well (given the criteria described in the following section), many groups report difficulty in these calculations being compounded with growing QM system size [9, 19–24, 64–71]. Thus, the decision not to include an additional FEP correction for the CB[8] host between FM/MM and QM/MM was based on the fact that (1) computational overhead associated with computing a host system of about 144 atoms is rather expensive for QM/MM calculations and (2) $U^{(FM \rightarrow QM)/MM}$ fluctuations between FM/MM and QM/MM will be even greater, prohibiting convergence (see discussion below).

While performing the correction calculations, it is often informative to collect the standard deviation of potential energy differences for the FM configurations, $\sigma_{\Delta U^{(FM \rightarrow QM)/MM}}^{FM/MM}$ [22, 24]. Generally speaking, the larger this quantity is, the poorer the resulting FEP calculation will be, as large U spread implies that the exponential average will be dominated by a few snapshots that dip into the lower energy range, and so having $\sigma_{\Delta U^{(FM \rightarrow QM)/MM}}^{FM/MM} \leq 4k_B T$ (~ 2.4 kcal/mol) is desirable. Another telling quantity of interest is the so called bias measure [72, 73], which given the assumption that the spread of energy differences between end states is nearly identical, can be approximated [22, 24] as,

$$\begin{aligned} \Pi_{\Delta U^{(FM \rightarrow QM)/MM}}^{FM/MM} = \sqrt{\mathbf{W}_L \left[\frac{1}{2\pi} (N-1)^2 \right]} \\ - \sqrt{2\beta \left(\langle \Delta U^{(FM \rightarrow QM)/MM} \rangle_{FM/MM} - \Delta A^{(FM \rightarrow QM)/MM} \right)}, \end{aligned} \quad (5)$$

with $\mathbf{W}_L$ as the Lambert function, and $A^{(FM \rightarrow QM)/MM}$ obtained with FEP using a configurational sample size of N . The standard deviation of the potential energy difference in conjunction with the bias measure act as a metric for the “unreliability” of an FEP

calculation. Specifically, following the recommendations of Ref. [22], whenever $\Pi_U < 0.5$ or $\sigma_U > 4k_B T$ kcal/mol, convergence of the associated FEP calculation is questionable. However, even if $\Pi_U = 0.5$ and $\sigma_U = 4k_B T$, this is not a guarantee that the FEP correction is converged (i.e. necessary, but not sufficient) [22].

2.6 Simulation Setup

Using the GPU powered variant of DOMDEC [74] in CHARMM, host-guest and guest only system were initially equilibrated in the isobaric-isothermal ensemble with a Nosé Hoover thermostat [75] for about 500 ps to 298.15K at 1 atm, which was then followed up with 500 ps of constant volume/temperature equilibration. All systems employed particle mesh ewalds during production time, with an image and non-bond cut off of 14 Å, a kappa value of 0.36, a time step of 1 fs, and a switching function to taper non-bonded contributions starting at 10 Å. All TIP3 waters were constrained via SHAKE, and an additional 200 ps of equilibration was performed in which a center of mass (COM) restraint was gently applied (on either the host-guest system or guest only system) through gradually increasing the associated COM force constant every 50 ps from zero up to 5.0 kcal/mol·Å². This was then proceeded by a final 250 ps equilibration period and immediately followed up by 10 ns of production, in which coordinate snapshots were collected every 500 fs. Implementing a similar methodology to Ref. [19], QM/MM calculations were performed without cutoff and without the use of QM-MM ewalds. Instead, FM energies were also collected without cutoff, and subtracted from energies obtained with FM using periodic boundary conditions. This “PBC” energy was then added to the QM/MM energy obtained to account for periodic effects.

2.7 Parameter Interpolation Approach

Shortly before the SAMPL6 deadline, it came to our attention that a few software issues (explained in greater detail in the Results and Discussion section) prevented CHARMM from properly utilizing all the parameters provided from the FM procedure in the DDM calculations of Ref. [41]. Although for the classical DDM results were “reasonable”, the configurational mismatched caused by the missing parameters of this “Reduced” Force Match (RFM) set made use of corrections with FEP difficult.

Fortunately, for two of the guest molecules, G0 and G1, the missing terms were readily identifiable (39 and 20 unique parameters, respectively), and thus a two tier correction scheme, based on Fig. 4, was devised in which the results from the RFM binding free energy calculations were corrected to FM, followed by a second correction to QM/MM (see Fig. 5). The free energy difference between classical end states of interest (e.g., X and Y), $\Delta A^{X \rightarrow Y}$ can be computed with more rigorous approaches, such as BAR [63] or the respective multistate extension, MBAR [76]. Pursuant to this, a series of intermediate states are constructed between states X and Y. Generally, this would imply the use of a hybrid potential energy function,

$$U_\lambda = (1 - \lambda)U^X - \lambda U^Y, \quad (6)$$

where λ is set incrementally between 0 and 1 (e.g., $\lambda=0.0, 0.1, \dots, 0.9, 1.0$, giving 11 intermediate simulations). At $\lambda=0.0$ or 1.0 , we have U^X and U^Y respectively, and the free energy difference between the two can be calculated as

$$\Delta A^{X \rightarrow Y} = \sum_{i=0}^9 \Delta A^{\lambda_i \rightarrow \lambda_{i+1}}, \lambda_i = i / 10 \quad (7)$$

However, the CHARMM dedicated free energy module, PERT [77], is not GPU supported. Thus, in place of defining intermediate states as linear combinations of potential energy functions, intermediate states were defined by directly modifying force field terms as a linear combination of end state parameters. As an example, consider the bond energy terms for states X and Y,

$$E_{BOND}^X = \sum_{\text{bonds}} K_{BOND}^X (b - b_0^X)^2 \quad (8)$$

$$E_{BOND}^Y = \sum_{\text{bonds}} K_{BOND}^Y (b - b_0^Y)^2 \quad (9)$$

where K_{BOND} is the force constant and b_0 is the equilibrium offset (in units of kcal/mol·Å² and Å, respectively), which have different values for states X and Y. Due to the end states being topologically similar (i.e., preserving the same connectivity), we can manufacture a hybrid bond energy term similar to how the hybrid potential energy function is defined,

$$E_{BOND}^\lambda = \sum_{\text{bonds}} K_{BOND}^\lambda (b - b_0^\lambda)^2 \quad (10)$$

$$K_{BOND}^\lambda = (1 - \lambda)K_{BOND}^X + \lambda K_{BOND}^Y \quad (11)$$

$$b_0^\lambda = (1 - \lambda)b_0^X + \lambda b_0^Y \quad (12)$$

Similar setups can be defined for angle, dihedral, and improper dihedral terms as well, and is known as parameter interpolation [78]. However, trying to alter all terms simultaneously can result in major problems in configurational overlap between intermediate states, so it is advised to modify each collection of energy terms (i.e., Bonds, Angles, Dihedrals, Improvers) one at a time. Therefore, instead of a single λ , there is a λ associated with each term (i.e., $\lambda_B, \lambda_A, \lambda_D, \lambda_I$ for Bonds, Angles, Dihedral, and Improvers, see Eq. 13), and thus a new expression can be found for the free energy difference between states X and Y:

$$\begin{aligned} \Delta A^{X \rightarrow Y} = & \sum_{i=0}^9 \Delta A(\lambda_B^{i \rightarrow i+1}, \lambda_A^0, \lambda_D^0, \lambda_I^0) + \Delta A(\lambda_B^{10}, \lambda_A^{i \rightarrow i+1}, \lambda_D^0, \lambda_I^0) \\ & + \Delta A(\lambda_B^{10}, \lambda_A^{10}, \lambda_D^{i \rightarrow i+1}, \lambda_I^0) + \Delta A(\lambda_B^{10}, \lambda_A^{10}, \lambda_D^{10}, \lambda_I^{i \rightarrow i+1}). \end{aligned} \quad (13)$$

Returning to Fig. 5, $A^{(RFM \rightarrow FM)/MM}$ corrections were computed for the host only system (Fig. 5[v]), the guest only system (Fig. 5[vi]), and host-guest system (Fig. 5[vii]). Every $A^{(RFM \rightarrow FM)/MM}$ was obtained using 40 intermediate λ simulations and the parameter interpolation scheme (i.e., 11 λ states for bonds, 11 λ states for angles, etc.), with each simulation setup in the same aforementioned fashion, having a 250 ps equilibration period, a NVT production of 5 ns, and a coordinate saving frequency of 1 ps (i.e., 5,000 coordinate snapshots per lambda simulation). Following the collection of the requisite energies for all λ states, MBAR calculations were performed to get $A^{(RFM \rightarrow FM)/MM}$.

3 Results and Discussion

3.1 Overall Performance

As mentioned prior, results for correcting directly from RFM/MM to QM/(RFM)MM were used as submission in the SAMPL6 challenge. A statistical summary of the “correction” scheme submission can be found in Table 1, and each host-guest result is displayed in Table 2, along with a linear regression plot of the submitted calculations versus experimentally obtained binding free energies (Fig. 6). Two sets of statistics were provided in Table 1. First, there is the statistics generated from using the “Standard” (i.e., required) set of host-guest calculations (i.e., the binding of guest G0 through G10). Thirty six submissions were obtained for the standard set, with just six of those submissions also attempting the “Bonus” G11, G12, and G13 molecules. The correction scheme’s performance on the standard set in regards to error metrics (e.g., RMSE, MAE, and ME) was overall moderately poor, with a ranking for RMSE (5.50 kcal/mol) and MAE (5.06 kcal/mol) in the bottom half of submissions (about 22nd out of 36). Surprisingly, ME (2.28 kcal/mol) was in the top 25% (8th out of 36), which is more than likely a function of gratuitous error cancellation.

The two correlation measures, R^2 and τ , however, proved to be quite excellent for the standard set. Specifically, with an R^2 of 0.46 (the second highest reported) and τ of 0.64 (the highest of all standard set submissions), findings based on the correction scheme seem to correlate decently with qualitative trends observed (at least within the scope of all submissions). In terms of the bonus set findings, it appears that although RMSE and MAE improve very slightly (5.43 and 4.77 kcal/mol, respectively), the additional 3 challenge molecules cause the quality of correlation to deteriorate ($R^2=0.26$ and $\tau=0.47$). Regardless, the performance amongst the 6 submissions attempting the bonus set ranked the correction results 3rd out of 6 for RMSE and MAE, 1st out of 6 for ME, and 2nd out of 6 for R^2 , m , and τ .

3.2 Overview of Results

Unfortunately, based on observations from Table 2 and Fig. 6, the correction scheme exacerbated deviation from experiment in a number of the RFM results. In particular, for the binding of guests G0, G4, and G12, corrections were on the order of about +11, -7, and +13 kcal/mol for each respectively. However, based on the bias measure Π and standard deviation σ_U metrics discussed earlier, we see large fluctuations in potential energy ($\sigma_U > 5$ kcal/mol) and $\Pi < -3.0$, indicating problems with converging FEP corrections for these systems. In fact, we see markers of convergence problems (e.g., small Π and large σ_U) for G0, G1, G2,

G3, G4, G8, G11, G12, and G13. This led into an inquiry as to why the FEP calculations would not converge properly and exposed a multitude of factors that exacerbated configurational mismatch between FM/MM and QM/MM levels of theory.

3.3 Exclusion of Parameters

The largest determining factor in the problems found in the “correction” scheme are almost exclusively related to software difficulties regarding the various parameters produced in the FM procedures. Specifically, not all of the parameters from FM were utilized, and a summary of how many are missing per system is shown in Table 3, and an explanation as to the cause is provided for each below. Take note, the analysis to follow excludes the guest molecule G13, as the complications involving that binding calculation are rather unique and will be discussed in detail later.

“Additional” Improper Dihedrals.—One of the “features” of the original ForceSolve program (which was employed during the SAMPL6 challenge) was the ability to detect molecular topology based on bond connectivity. The program then assigns parameter type based on the topological graph generated. Unfortunately this also included the automatic creation of improper dihedrals for all atomic centers with exactly 3 bonds. Improper dihedrals facilitate fits by ensuring planarity amongst a group of atoms, but this is often achieved by fine tuning standard dihedral terms. Unfortunately, CHARMM requires the explicit identification of improper dihedral terms (unlike dihedrals and angles, which are generated automatically), and the use of the same topological setup generated prior to the fitting scheme (as only the parameters should change, not the topology) resulted in the “new” improper dihedrals not being utilized.

Not catching these additions corresponded to the exclusions of some improper dihedral terms, which while not necessarily required in the original force field (CGenFF), are an integral part of the new FM fit. Updates in the FM code have been made to only fit terms prescribed by the user. Regardless, numerous molecules (specifically, G0, G1, G2, G3, G4, G12, and CB[8]) had these additional improper dihedral terms missing, and thus exhibited unnatural flexibility (e.g. the bending of aromatic rings). Although in terms of classical potentials, the bending of an aromatic ring doesn't incur a massive energetic penalty, for QM methods this is quite energetically unfavorable, and thus the magnitude of QM energy fluctuations becomes makes single step free energy perturbation untenable (as seen in Table 2 by the values of σ_U).

Missing Dihedrals.—Another function of the original ForceSolve program was the fitting of both angle terms $K_{Ang}(\theta - \theta_0)^2$ and 1-3 Urey-Bradley (UB) terms $K_{UB}(b^{1,3} - b_0^{1,3})^2$. As a rule of thumb, the more terms, the better the force fit, but an interesting “feature” of the CHARMM program involves the parsing of angle parameters and the corresponding influence on dihedral list generation. Essentially, CHARMM automatically detects whether the equilibrium value of an angle term is arbitrarily close to values of 0° or 180°, and if so, does not generate any dihedrals containing the angle parameters in question, as torsions are ill defined for systems with completely linear geometry.

However, in the case of using both the UB and angle parameters, it is possible to define a more complex angle form. Occasionally this will manifest in a situation where the individual angle component takes on an equilibrium angle of 0° , but the UB term is non trivial, and the combination gives an energetic minima at a non-zero angle value (as shown in Fig 7), but regardless CHARMM will throw out any associated dihedrals containing these 0° angle terms.

3.4 Mixing of Parameter sets

Overlap in Atom Typing between Host and Guest—A major oversight that occurred as a result of conditioning to working with transferable force fields is that there was some overlap in atom types between host and guest species. In some cases where two or more atom types are shared between host and guest and correspond to a bonded atom pair, this results in replacement of the bonded term, utilizing either the host or guest bonded parameter depending on the order in which parameters are read in. Since parameters are uniquely generated for each guest and host system, overwriting of terms can cause significant changes to the conformational sampling for each system. In the case of the calculations run with the double decoupling scheme, the host parameters were read in prior to guest. Thus for each host/guest pairing, the host system had slight to significant difference (i.e. either 1 parameter is different, as is the case for G5, or up to 5 as seen for G9) in description based on the degree of overlap. Table 4 gives the number of parameters overlapping between host and guest systems, and the RMSD between gas phase minimized host sets. What this parameter overlap illustrates is that for each host-guest system, the host is “subtly” different case by case, and not using the parameter set generated for CB[8] based on the FM procedure, rather a combination with some of the FM parameters from the CB[8] and some FM parameters of the guest inhabiting the host.

Dihedral Over-writing Parser Bug—This is a bug associated with reading in the CGenFF forcefield prior to the FM parameter set, resulting in dihedral terms that contain multiple dihedral multiplicities being a combination of both CGenFF and FM parameter sets. The one main observation regarding this bug, is that it is not consistent. Initially, it was presumed to only occur when reading in systems that have dihedrals with varying multiplicities (such as $n=2, n=3, n=4$). In one test case, for a dihedral with 3 different multiplicities ($n=2, n=3, n=5$), the dihedral set stored in CHARMM when reading in CGenFF parameters, followed by reading in the desired FM parameters, resulted in a term that had $n=2$ and $n=3$ from CGenFF and $n=5$ from the FM set. However, it is hard to identify the exact cause, as not all multiply defined dihedrals cause this problem, but some do, and require creation of an entirely separate, self contained, topology and parameter file for the system of interest. To illustrate, Table 5 contains the RMSD between minimization of identical gas phase guest structures with CGenFF parameters read in prior versus reading in only FM parameters in a self contained topology and parameter set, along with differences in dihedral energies of equilibrated guest structures (excluding CB[8]) prepared with and without reading of CGenFF parameters prior.

Out of all the issues discovered, this one is particularly unsettling, due to the fact that it seems unpredictable in nature. Based solely off gas phase RMSD, one could feasibly assume

that G2 would be a system for which the parameter interpolation scheme would be applicable. However, energetically the difference, albeit small, is observable. This is excluding the fact that the impact of the effect is difficult to measure, as looking at G7, one would assume that the energetic consequence would be smaller than that of G3, for which has a much larger gas-phase RMSD, but the observed energetic effect is over three times greater for G7 in magnitude.

3.5 Quantifying the Effects of the Various Issues

A variety of problems have been presented as potential causes to the failure of the FEP calculations (i.e., based on the bias metric Π and standard deviation of the potential energy metrics laid out prior). To determine the most significant underlying cause of failure, a Pearsons correlation matrix (Table. 6) was constructed between the deviation of calculated results (obtained with the correction scheme) from experiment ϵ_{err} and various quantities of interest. The major observations to be made, is that the number of missing parameters has a strong correlation with the accuracy of the FEP “correction” calculation (the total number of missing parameters has a correlation value of +0.82 with the deviation from experimental value for guest G0-G4 and G12, and +0.58 amongst G0-G12). Surprisingly, the amount of parameter overlap between guest and host had a correlation of -0.11 with deviation of calculated result from experiment.

3.6 Parameter Interpolation Results

Using the parameter interpolation scheme garnered exclusively better results when applied to G0 and G1 binding calculations. A summary of the relevant findings are displayed in Table 7.

G0.—Almost immediately we see that the correction calculation results of G0 significantly improve, with a deviation from experimental of almost 10 kcal/mol, down to ~ 1.5 kcal/mol. However, going from RFM/MM to FM/MM, the classical binding free energy jumps by about +3.3 kcal/mol, and goes from an underestimation of -1.3 kcal/mol to an overestimation of +2 kcal/mol. In terms of our convergence metrics, we see the standard deviation of U drop dramatically (e.g., -7.3 kcal/mol for the G0 only system, and -9.3 kcal/mol for CB[8]+G0), indicating a move in the right direction, but still just about 1 kcal/mol outside of acceptable range for both, indicating a need for more sampling. We see a similar trend in the bias metric Π when using FEP from RFM versus FM, with improvements of +3.3 and +6.4 for G0 and CB[8]+G0 Π values respectively. However, these Π values are off from acceptable by ~ 2.5 .

G1.—Improvements obtained by FEP corrections from FM/MM over RFM/MM are less dramatic as G0, but still give a net increase in accuracy of about 0.5 kcal/mol. In terms of the purely classical result we don't see a meaningful change (a gain in accuracy of about 0.08 kcal/mol for FM/MM over RFM/MM). With respect to the metrics, standard deviations for U are significantly better, almost falling in satisfactory range (below 2.4 kcal/mol). However, the bias measure Π is still negative, indicating the need to do more sampling in order to properly converge the calculation.

3.7 Challenges with G13

Guest molecule G13 presented a unique challenge, quite possibly the most difficult of all molecules, due to the presence of the Pt(II) metal center, for which parameters are not readily available. Although a work around was implemented for the generation of configurational sampling via MNDO(d) dynamics with palladium, the system was plagued with persistent problems. A variety of methods were used to best tackle this problem. First, the initial choice of theory for G13, B3LYP/LANL2DZ, came under scrutiny when a simple geometry optimization using this level of theory failed to converge with any degree of geometric stability (the carbonyl substituents would bend in and out of planarity and never reach a stable state). Alternatives to this level of theory were explored, including using semi-empirical PM6 available through MOPAC [79], and the CHARMM VALBOND module [80], which facilitates potentials based on the universal force field (UFF). However, the PM6 forces on the generated configurations were too high in magnitude (~ 1000 kcal/molÅ), and the VALBOND results caused problems with regards to the hydrogens on the distal nitrogens. Use of DFT was then revisited, but instead of using a pseudopotential, the all atom DZP [81] basis set was utilized, which was able to arrive at a stable geometric minima.

The second major problem, which again was not determined till shortly before the SAMPL6 deadline, was an issue regarding the atom type assignment for G13. (see Fig. 8). As seen on the left side of Fig. 8, the initial atom type assignment gave identical typings to both the oxygens coordinated to the Pt as well as the nitrogens. That is to say, a single angle parameter was used to simultaneously describe the cis O-Pt-N angle ($\approx 90^\circ$) and the trans O-Pt-N angles ($\approx 180^\circ$). A simple addition of new atom types, based on whether the O-Pt-N angle was cis or trans, dramatically improved FM fits.

4 Concluding Remarks and Prospects

The use of FM with a correction scheme in this work is very much a “proof of concept”, in the sense that the level of theory chosen for matching wasn’t a particularly high one (e.g., a more complex density functional with a larger basis set), and quite a few glaring considerations need addressed before this methodology is ready for universally large scale application. However, regardless of all aforementioned problems, the results obtained were consistently amongst the top two in correlation (either for the correction scheme results or the classical DDM results). It was also demonstrated that for two cases where software/implementation problems were fairly dominant, proper implementation resulted in a net gain in correction scheme accuracy (i.e., using the parameter interpolation scheme for G0 and G1). Findings from the correlation values of Table 6 showed that missing parameters had a major impact on the accuracy, depending on the amount of parameters omitted.

Potential energy fluctuations and bias measures strongly correlate with σ_A obtained with the correction scheme, indicating that these quantities inherently provide details about the precision of a calculation, but as indicators of accuracy, the relationship is less pronounced. However, although its a bit unexpected, the overlap between guest and host parameters does not seem to have a large influence on the accuracy of the correction calculations. Though, this might not be the case if QM/MM correction were applied to the host as well.

Since parameter extrapolation is not possible for every case where parameters are missing, and problems with the dihedral parser of CHARMM remain, it would be of great interest to validate results with a second program. Utilizing the expedient (e.g., GPU supported) and dedicated free energy code through the OpenMM/YANK interface [82–84] or with the OpenMM tools [85] alchemy module have been of great interest. Currently, YANK does not support CHARMM format, and it appears as though at present that parsing CHARMM objects through the alchemy module is not quite compatible. This was verified through a series of sanity checks through the annihilation of charges and sterics (i.e., van-der Waals terms). When guest charges are manually set to zero in the CHARMM protein structure and topology files, the energies in OpenMM/alchemy agree with those obtained when the electrostatics control variable λ_{elec} is set to zero (i.e., guest system charged interactions are annihilated). However, following the same logic when applied to the annihilation of sterics (e.g., setting the ϵ_i values of the guest LJ terms to zero), energies are not in agreement with those obtained when the steric control variable λ_{steric} is set to zero (i.e., sterically annihilated guest). Also, energies continue to change when ϵ_i values are scaled by orders of magnitude, again indicating the annihilation of sterics is not truly occurring as necessary. Developing routes to correct this are largely underway.

Other considerations are question of whether RESP charges based on MP2/6–31(d) with SMD being the most effective representative of Coloumbic interactions, if B3LYP/6-31G(d) is a sufficient level of theory, and the rather pervasive question of needing better LJ parameters. However, as observed in the results of Ref. [41], classical approaches can give decent free energies even when the configurational mismatch between the classical MM and higher level QM is moderate. Unfortunately, when using the same classical MM approach in a correction scheme, if the configurational mismatch is significant (i.e., as demonstrated in Table 6 regarding the missing parameters), converging these calculations is dubious at best, and care must be taken to make sure that the configurational overlap remain as large and as faithful to the desired QM surface as possible.

Acknowledgements

The authors would like to thank Richard Venable, John Legato, Qiao Zheng, and Michael R. Jones for technical assistance. We extend our gratitude to Stefan Boresch, Florentina Tofoleanu, and Andrea Rizzi for helpful comments on the manuscript and general insights. This work was partially supported by the intramural research program of the National Heart, Lung and Blood Institute (NHLBI) of the National Institutes of Health and employed the high-performance computational capabilities of the LoBoS and Biowulf Linux clusters at the National Institutes of Health. (<http://www.lobos.nih.gov> and <http://biowulf.nih.gov>). PSH acknowledges funding support from the Intramural Research Program of the NIH, NHLBI. HLW would like to highlight that this material is based upon work supported by the National Science Foundation under CHE–1464946.

References

1. Guthrie JP (2009) A blind challenge for computational solvation free energies: Introduction and overview. *J Phys Chem B* 113(14):4501–4507, DOI 10.1021/jp806724u, URL 10.1021/jp806724u [PubMed: 19338360]
2. Geballe MT, Skillman AG, Nicholls A, Guthrie JP, Taylor PJ (2010) The SAMPL2 blind prediction challenge: Introduction and overview. *J Comput Aided Mol Des* 24(4):259–279, DOI 10.1007/s10822-010-9350-8, URL 10.1007/s10822-010-9350-8 [PubMed: 20455007]
3. Muddana HS, Daniel Varnado C, Bielawski CW, Urbach AR, Isaacs L, Geballe MT, Gilson MK (2012) Blind prediction of host-guest binding affinities: A new SAMPL3 challenge. *J Comput*

Aided Mol Des 26(5):475–487, DOI 10.1007/s10822-012-9554-1, URL 10.1007/s10822-012-9554-1 [PubMed: 22366955]

4. Muddana HS, Fenley AT, Mobley DL, Gilson MK (2014) The SAMPL4 host-guest blind prediction challenge: An overview. *J Comput Aided Mol Des* 28(4):305–317, DOI 10.1007/s10822-014-9735-1, URL 10.1007/s10822-014-9735-1 [PubMed: 24599514]
5. Yin J, Henriksen NM, Slochow DR, Shirts MR, Chiu MW, Mobley DL, Gilson MK (2016) Overview of the SAMPL5 host-guest challenge: Are we doing better? *J Comput Aided Mol Des* 31(1):1–19, DOI 10.1007/s10822-016-9974-4, URL 10.1007/s10822-016-9974-4 [PubMed: 27658802]
6. Zacharias M, Straatsma TP, McCammon JA (1994) Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J Chem Phys* 100(12):9025–9031, DOI 10.1063/1.466707, URL 10.1063/L466707
7. Beutler TC, Mark AE, van Schaik RC, Gerber PR, van Gunsteren WF (1994) Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett* 222(6):529–539, DOI 10.1016/0009-2614(94)00397-1, URL 10.1016/0009-2614(94)00397-1
8. Chipot C, Pohorille A (eds) (2007) *Free Energy Calculations*. Springer Berlin Heidelberg, DOI 10.1007/978-3-540-38448-9, URL 10.1007/978-3-540-38448-9
9. Ryde U, Söderhjelm P (2016) Ligand-binding affinity estimates supported by quantum-mechanical methods. *Chem Rev* 116(9):5520–5566, DOI 10.1021/acs.chemrev.5b00630, URL 10.1021/acs.chemrev.5b00630 [PubMed: 27077817]
10. Mackerell AD, Feig M, Brooks CL (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25(11):1400–1415, DOI 10.1002/jcc.20065, URL 10.1002/jcc.20065 [PubMed: 15185334]
11. Manzoni F, Söderhjelm P (2014) Prediction of hydration free energies for the SAMPL4 data set with the AMOEBA polarizable force field. *J Comput Aided Mol Des* 28(3):235–244, DOI 10.1007/s10822-014-9733-3, URL 10.1007/s10822-014-9733-3 [PubMed: 24577872]
12. Gao J, Xia X (1992) A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science* 258(5082):631–635, DOI 10.1126/science.1411573, URL 10.1126/science.1411573 [PubMed: 1411573]
13. Gao J, Luque FJ, Orozco M (1993) Induced dipole moment and atomic charges based on average electrostatic potentials in aqueous solution. *J Chem Phys* 98(4):2975–2982, DOI 10.1063/1.464126, URL 10.1063/L464126
14. Luzhkov V, Warshel A (1992) Microscopic models for quantum mechanical calculations of chemical processes in solutions: Ld/ampac and SCAAS/AMPAC calculations of solvation energies. *J Comput Chem* 13(2):199–213, DOI 10.1002/jcc.540130212, URL 10.1002/jcc.540130212
15. Wesolowski T, Warshel A (1994) Ab initio free energy perturbation calculations of solvation free energy using the frozen density functional approach. *J Phys Chem* 98(20):5183–5187, DOI 10.1021/j100071a003, URL 10.1021/j100071a003
16. Gao J, Freindorf M (1997) Hybrid ab initio QM/MM simulation of N-methylacetamide in aqueous solution. *J Phys Chem A* 101(17):3182–3188, DOI 10.1021/jp970041q, URL 10.1021/jp970041q
17. Zheng YJ, Merz KM (1992) Mechanism of the human carbonic anhydrase II-catalyzed hydration of carbon dioxide. *J Am Chem Soc* 114(26):10498–10507, DOI 10.1021/ja00052a054, URL 10.1021/ja00052a054
18. Zwanzig RW (1954) High-temperature equation of state by a perturbation method. i. nonpolar gases. *J Chem Phys* 22(8):1420–1426, DOI 10.1063/1.1740409, URL 10.1063/1.1740409
19. König G, Hudson PS, Boresch S, Woodcock HL (2014) Multiscale free energy simulations: An efficient method for connecting classical MD simulations to QM or QM/MM free energies using non-Boltzmann Bennett reweighting schemes. *J Chem Theory Comput* 10(4):1406–1419, DOI 10.1021/ct401118k, URL 10.1021/ct401118k [PubMed: 24803863]
20. Hudson PS, White JK, Kearns FL, Hodoscek M, Boresch S, Lee Woodcock H (2015) Efficiently computing pathway free energies: New approaches based on chain-of-replica and non-Boltzmann

- Bennett reweighting schemes. *Biochim Biophys Acta, Gen Subj* 1850(5):944–953, DOI 10.1016/j.bbagen.2014.09.016, URL 10.1016/j.bbagen.2014.09.016
21. Hudson PS, Woodcock HL, Boresch S (2015) Use of nonequilibrium work methods to compute free energy differences between molecular mechanical and quantum mechanical representations of molecular systems. *J Phys Chem Lett* 6(23):4850–4856, DOI 10.1021/acs.jpclett.5b02164, URL 10.1021/acs.jpclett.5b02164 [PubMed: 26539729]
22. Boresch S, Woodcock HL (2016) Convergence of single-step free energy perturbation. *Mol Phys* 115(9-12):1200–1213, DOI 10.1080/00268976.2016.1269960, URL 10.1080/00268976.2016.1269960
23. Kearns FL, Hudson PS, Woodcock HL, Boresch S (2017) Computing converged free energy differences between levels of theory via nonequilibrium work methods: Challenges and opportunities. *J Comput Chem* 38(16):1376–1388, DOI 10.1002/jcc.24706, URL 10.1002/jcc.24706 [PubMed: 28272811]
24. Hudson PS, Boresch S, Rogers DM, Woodcock HL (2018) Accelerating QM/MM free energy computations via intramolecular force matching. *J Chem Theory Comput* (in revision)
25. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78(14):2690–2693, DOI 10.1103/physrevlett.78.2690, URL 10.1103/physrevlett.78.2690
26. Csányi G, Albaret T, Payne MC, De Vita A (2004) “Learn on the fly”: A hybrid classical and quantum-mechanical molecular dynamics simulation. *Phys Rev Lett* 93(17), DOI 10.1103/physrevlett.93.175503, URL 10.1103/physrevlett.93.175503
27. Akin-Ojo O, Song Y, Wang F (2008) Developing ab initio quality force fields from condensed phase quantum-mechanics/molecular-mechanics calculations through the adaptive force matching method. *J Chem Phys* 129(6):064108, DOI 10.1063/1.2965882, URL 10.1063/L2965882 [PubMed: 18715052]
28. Akin-Ojo O, Wang F (2010) The quest for the best nonpolarizable water model from the adaptive force matching method. *J Comput Chem* 32(3):453–462, DOI 10.1002/jcc.21634, URL 10.1002/jcc.21634 [PubMed: 20730778]
29. Wang F, Akin-Ojo O, Pinnick E, Song Y (2011) Approaching post-hartree-fock quality potential energy surfaces with simple pair-wise expressions: Parameterising point-charge-based force fields for liquid water using the adaptive force matching method. *Mol Simul* 37(7):591–605, DOI 10.1080/08927022.2011.565759, URL 10.1080/08927022.2011.565759
30. Pinnick ER, Calderon CE, Rusnak AJ, Wang F (2012) Achieving fast convergence of ab initio free energy perturbation calculations with the adaptive force-matching method. *Theor Chem Acc* 131(3), DOI 10.1007/s00214-012-1146-6, URL 10.1007/s00214-012-1146-6
31. Li J, Wang F (2015) Pairwise-additive force fields for selected aqueous monovalent ions from adaptive force matching. *J Chem Phys* 143(19):194505, DOI 10.1063/1.4935599, URL 10.1063/L4935599 [PubMed: 26590540]
32. Wang LP, Van Voorhis T (2010) Communication: Hybrid ensembles for improved force matching. *J Chem Phys* 133(23):231101, DOI 10.1063/1.3519043, URL 10.1063/L3519043 [PubMed: 21186847]
33. Wang LP, Chen J, Van Voorhis T (2012) Systematic parametrization of polarizable force fields from quantum chemistry data. *J Chem Theory Comput* 9(1):452–460, DOI 10.1021/ct300826t, URL 10.1021/ct300826t [PubMed: 26589047]
34. Wang LP, McKiernan KA, Gomes J, Beauchamp KA, Head-Gordon T, Rice JE, Swope WC, Martinez TJ, Pande VS (2017) Building a more predictive protein force field: A systematic and reproducible route to AMBER-FB15. *J Phys Chem B* 121(16):4023–4039, DOI 10.1021/acs.jpcb.7b02320, URL 10.1021/acs.jpcb.7b02320 [PubMed: 28306259]
35. Rogers DM (2016) ForceSolve. <https://github.com/frobnitzem/forcesolve>
36. Rogers DM (2018) ChemParam. <https://github.com/frobnitzem/chemparam>
37. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoseck M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M

- (2009) CHARMM: The biomolecular simulation program. *J Comput Chem* 30(10):1545–1614, DOI 10.1002/jcc.21287, URL 10.1002/jcc.21287 [PubMed: 19444816]
38. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD (2009) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* 31(4):NA–NA, DOI 10.1002/jcc.21367, URL 10.1002/jcc.21367
39. Gilson M, Given J, Bush B, McCammon J (1997) The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophys J* 72(3):1047–1069, DOI 10.1016/s0006-3495(97)78756-3, URL 10.1016/s0006-3495(97)78756-3 [PubMed: 9138555]
40. Boresch S, Tettinger F, Leitgeb M, Karplus M (2003) Absolute binding free energies:~ a quantitative approach for their calculation. *J Phys Chem B* 107(35):9535–9551, DOI 10.1021/jp0217839, URL 10.1021/jp0217839
41. Han K, Hudson PS, Jones MR, Nishikawa N, Tofoleanu F, Brooks BR (2018) Prediction of cb[8] host–guest binding free energies in sampl6 using the double-decoupling method. *J Comput Aided Mol Des* DOI 10.1007/s10822-018-0144-8, URL 10.1007/s10822-018-0144-8
42. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2016) Gaussian16 Revision B.01. Gaussian Inc Wallingford CT
43. Becke AD (1993) Density-functional thermochemistry. III. the role of exact exchange. *J Chem Phys* 98(7):5648–5652, DOI 10.1063/1.464913, URL 10.1063/1.464913
44. Lee C, Yang W, Parr RG (1988) Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B* 37(2):785–789, DOI 10.1103/physrevb.37.785, URL 10.1103/physrevb.37.785
45. Møller C, Plesset MS (1934) Note on an approximation treatment for many-electron systems. *Phys Rev* 46(7):618–622, DOI 10.1103/physrev.46.618, URL 10.1103/physrev.46.618
46. Hariharan PC, Pople JA (1973) The influence of polarization functions on molecular orbital hydrogenation energies. *Theor Chimica Acta* 28(3):213–222, DOI 10.1007/bf00533485, URL 10.1007/bf00533485
47. Francel MM, Pietro WJ, Hehre WJ, Binkley JS, Gordon MS, DeFrees DJ, Pople JA (1982) Self-consistent molecular orbital methods. XXIII. a polarization-type basis set for second-row elements. *J Chem Phys* 77(7):3654–3665, DOI 10.1063/1.444267, URL 10.1063/L444267
48. Marenich AV, Cramer CJ, Truhlar DG (2009) Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J Phys Chem B* 113(18):6378–6396, DOI 10.1021/jp810292n, URL 10.1021/jp810292n [PubMed: 19366259]
49. Bayly CI, Cieplak P, Cornell W, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J Phys Chem* 97(40):10269–10280, DOI 10.1021/j100142a004, URL 10.1021/j100142a004
50. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graphics Modell* 25(2):247–260, DOI 10.1016/j.jmgm.2005.12.005, URL 10.1016/j.jmgm.2005.12.005
51. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comp Chem* 25(9):1157–1174, DOI 10.1002/jcc.20035, URL 10.1002/2Fjcc.20035 [PubMed: 15116359]
52. Vanommeslaeghe K, Raman EP, MacKerell AD (2012) Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J Chem Inf Model* 52(12):3155–3168, DOI 10.1021/ci3003649, URL <http://pubs.acs.org/doi/10.1021/ci3003649> [PubMed: 23145473]

53. Yao S, Plastaras JP, Marzilli LG (1994) A molecular mechanics AMBER-type force field for modeling platinum complexes of guanine derivatives. *Inorg Chem* 33(26):6061–6077, DOI 10.1021/ic00104a015, URL 10.1021/ic00104a015
54. Ojeda-May P, Nam K (2017) Acceleration of semiempirical QM/MM methods through message passage interface (MPI), hybrid MPI/Open multiprocessing, and self-consistent field accelerator implementations. *J Chem Theory Comput* 13(8):3525–3536, DOI 10.1021/acs.jctc.7b00322, URL 10.1021/acs.jctc.7b00322 [PubMed: 28628742]
55. Thiel W, Voityuk AA (1996) Extension of MNDO to d orbitals:~ parameters and results for the second-row elements and for the zinc group. *J Phys Chem* 100(2):616–626, DOI 10.1021/jp952148o, URL 10.1021/jp952148o
56. Hay PJ, Wadt WR (1985) *Ab initio* effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J Chem Phys* 82(1):299–310, DOI 10.1063/1.448975, URL <http://aip.scitation.org/doi/10.1063/L448975>
57. Wadt WR, Hay PJ (1985) *Ab initio* effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. *J Chem Phys* 82(1):284–298, DOI 10.1063/1.448800, URL <http://aip.scitation.org/doi/10.1063/L448800>
58. Hay PJ, Wadt WR (1985) *Ab initio* effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J Chem Phys* 82(1):270–283, DOI 10.1063/1.448799, URL <http://aip.scitation.org/doi/10.1063/L448799>
59. Shao Y, Gan Z, Epifanovsky E, Gilbert AT, Wormit M, Kussmann J, Lange AW, Behn A, Deng J, Feng X, Ghosh D, Goldey M, Horn PR, Jacobson LD, Kaliman I, Khaliullin RZ, Ku T, Landau A, Liu J, Proynov EI, Rhee YM, Richard RM, Rohrdanz MA, Steele RP, Sundstrom EJ, Woodcock HL, Zimmerman PM, Zuev D, Albrecht B, Alguire E, Austin B, Beran GJO, Bernard YA, Berquist E, Brandhorst K, Bravaya KB, Brown ST, Casanova D, Chang CM, Chen Y, Chien SH, Closser KD, Crittenden DL, Diedenhofen M, DiStasio RA, Do H, Dutoi AD, Edgar RG, Fatehi S, Fusti-Molnar L, Ghysels A, Golubeva-Zadorozhnaya A, Gomes J, Hanson-Heine MW, Harbach PH, Hauser AW, Hohenstein EG, Holden ZC, Jagau TC, Ji H, Kaduk B, Khistyayev K, Kim J, Kim J, King RA, Klunzinger P, Kosenkov D, Kowalczyk T, Krauter CM, Lao KU, Laurent AD, Lawler KV, Levchenko SV, Lin CY, Liu F, Livshits E, Lochan RC, Luenser A, Manohar P, Manzer SF, Mao SP, Mardirossian N, Marenich AV, Maurer SA, Mayhall NJ, Neuscamman E, Oana CM, Olivares-Amaya R, O'Neill DP, Parkhill JA, Perrine TM, Peverati R, Prociuk A, Rehn DR, Rosta E, Russ NJ, Sharada SM, Sharma S, Small DW, Sodt A, Stein T, Stuck D, Su YC, Thom AJ, Tsuchimochi T, Vanovschi V, Vogt L, Vydrov O, Wang T, Watson MA, Wenzel J, White A, Williams CF, Yang J, Yeganeh S, Yost SR, You ZQ, Zhang IY, Zhang X, Zhao Y, Brooks BR, Chan GK, Chipman DM, Cramer CJ, Goddard WA, Gordon MS, Hehre WJ, Klamt A, Schaefer HF, Schmidt MW, Sherrill CD, Truhlar DG, Warshel A, Xu X, Aspuru-Guzik A, Baer R, Bell AT, Besley NA, Chai JD, Dreuw A, Dunietz BD, Furlani TR, Gwaltney SR, Hsu CP, Jung Y, Kong J, Lambrecht DS, Liang W, Ochsenfeld C, Rassolov VA, Slipchenko LV, Subotnik JE, Van Voorhis T, Herbert JM, Krylov AI, Gill PM, Head-Gordon M (2014) Advances in molecular quantum chemistry contained in the q-chem 4 program package. *Mol Phys* 113(2):184–215, DOI 10.1080/00268976.2014.952696, URL 10.1080/00268976.2014.952696
60. Rogers DM, Beck TL (2010) Resolution and scale independent function matching using a string energy penalized spline prior. ArXiv e-print URL <http://arxiv.org/abs/1003.4741>, 1003.4741
61. Rogers DM, Beck TL (2008) ForceSolve. URL <http://forcesolve.sourceforge.net>
62. Sugita Y, Kitao A, Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. *J Chem Phys* 113(15):6042–6051, DOI 10.1063/1.1308516, URL 10.1063/L1308516
63. Bennett CH (1976) Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys* 22(2):245–268, DOI 10.1016/0021-9991(76)90078-4, URL 10.1016/0021-9991(76)90078-4
64. Ryde U (2017) How Many Conformations Need To Be Sampled To Obtain Converged QM/MM Energies? The Curse of Exponential Averaging. *J Chem Theory Comput* 13(11):5745–5752, DOI 10.1021/acs.jctc.7b00826, URL <http://pubs.acs.org/doi/10.1021/acs.jctc.7b00826> [PubMed: 29024586]
65. Olsson MA, Söderhjelm P, Ryde U (2016) Converging ligand-binding free energies obtained with free-energy perturbations at the quantum mechanical level: FULL PAPER. *J Comp Chem* 37(17):

- 1589–1600, DOI 10.1002/jcc.24375, URL <http://doi.wiley.com/10.1002/jcc.24375> [PubMed: 27117350]
66. Olsson MA, Ryde U (2017) Comparison of QM/MM Methods To Obtain Ligand-Binding Free Energies. *J Chem Theory Comput* 13(5):2245–2253, DOI 10.1021/acs.jctc.6b01217, URL <http://pubs.acs.org/doi/10.1021/acs.jctc.6b01217> [PubMed: 28355487]
67. Heimdal J, Ryde U (2012) Convergence of QM/MM free-energy perturbations based on molecular-mechanics or semiempirical simulations. *Phys Chem Chem Phys* 14(36):12592, DOI 10.1039/c2cp41005b, URL <http://xlink.rsc.org/?DOI=c2cp41005b> [PubMed: 22797613]
68. Genheden S, Cabedo Martinez AI, Criddle MP, Essex JW (2014) Extensive all-atom Monte Carlo sampling and QM/MM corrections in the SAMPL4 hydration free energy challenge. *J Comput Aided Mol Des* 28(3):187–200, DOI 10.1007/s10822-014-9717-3, URL <http://link.springer.com/10.1007/s10822-014-9717-3>
69. Cave-Ayland C, Skylaris CK, Essex JW (2015) Direct Validation of the Single Step Classical to Quantum Free Energy Perturbation. *J Phys Chem B* 119(3):1017–1025, DOI 10.1021/jp506459v, URL <http://pubs.acs.org/doi/10.1021/jp506459v> [PubMed: 25238649]
70. Sampson C, Fox T, Tautermann CS, Woods C, Skylaris CK (2015) A “Stepping Stone” Approach for Obtaining Quantum Free Energies of Hydration. *J Phys Chem B* 119(23):7030–7040, DOI 10.1021/acs.jpcc.5b01625, URL <http://pubs.acs.org/doi/10.1021/acs.jpcc.5b01625> [PubMed: 25985723]
71. Genheden S, Ryde U, Söderhjelm P (2015) Binding affinities by alchemical perturbation using QM/MM with a large QM system and polarizable MM model. *J Comp Chem* 36(28):2114–2124, DOI 10.1002/jcc.24048, URL <http://doi.wiley.com/10.1002/jcc.24048> [PubMed: 26280564]
72. Wu D, Kofke DA (2004) Model for small-sample bias of free-energy calculations applied to gaussian-distributed nonequilibrium work measurements. *J Chem Phys* 121(18):8742–8747, DOI 10.1063/1.1806413, URL 10.1063/1.1806413 [PubMed: 15527337]
73. Wu D, Kofke DA (2005) Phase-space overlap measures. i. fail-safe bias detection in free energies calculated by molecular simulation. *J Chem Phys* 123(5):054103, DOI 10.1063/1.1992483, URL 10.1063/1.1992483 [PubMed: 16108627]
74. Hynninen AP, Crowley MF (2013) New faster CHARMM molecular dynamics engine. *J Comput Chem* 35(5):406–413, DOI 10.1002/jcc.23501, URL 10.1002/jcc.23501 [PubMed: 24302199]
75. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* 31(3):1695–1697, DOI 10.1103/physrev.31.1695, URL 10.1103/physrev.31.1695
76. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129(12):124105, DOI 10.1063/1.2978177, URL 10.1063/L2978177 [PubMed: 19045004]
77. Boresch S, Karplus M (1999) The Role of Bonded Terms in Free Energy Simulations: 1. Theoretical Analysis. *J Phys Chem A* 103(1):103–118, DOI 10.1021/jp981628n, URL <http://pubs.acs.org/doi/abs/10.1021/jp981628n>
78. Giese TJ, York DM (2018) A GPU-accelerated parameter interpolation thermodynamic integration free energy method. *J Chem Theory Comput* 14(3):1564–1582, DOI 10.1021/acs.jctc.7b01175, URL 10.1021/acs.jctc.7b01175 [PubMed: 29357243]
79. Stewart JJP (2007) Optimization of parameters for semiempirical methods v: Modification of NDDO approximations and application to 70 elements. *J Mol Model* 13(12):1173–1213, DOI 10.1007/s00894-007-0233-4, URL 10.1007/s00894-007-0233-4 [PubMed: 17828561]
80. Cleveland T, Landis CR (1996) Valence bond concepts applied to the molecular mechanics description of molecular shapes. 2. applications to hypervalent molecules of the p-block. *J Am Chem Soc* 118(25):6020–6030, DOI 10.1021/ja950652i, URL 10.1021/ja950652i
81. Berrêdo RCd, Jorge FE (2010) All-electron double zeta basis sets for platinum: Estimating scalar relativistic effects on platinum(II) anticancer drugs. *J Mol Struct: THEOCHEM* 961(1):107–112, DOI 10.1016/j.theochem.2010.09.007, URL <http://www.sciencedirect.com/science/article/pii/S0166128010005786>
82. Wang K, Chodera JD, Yang Y, Shirts MR (2013) Identifying ligand binding sites and poses using GPU-accelerated Hamiltonian replica exchange molecular dynamics. *J Comput Aided Mol Des*

27(12):989–1007, DOI 10.1007/s10822-013-9689-8, URL 10.1007/s10822-013-9689-8 [PubMed: 24297454]

83. Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang LP, Simmonett AC, Harrigan MP, Stern CD, Wiewiora RP, Brooks BR, Pande VS (2017) OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS Comput Biol 13(7):e1005659, DOI 10.1371/journal.pcbi.1005659, URL 10.1371/journal.pcbi.1005659 [PubMed: 28746339]
84. Rizzi A, Chodera J, Naden L, Beauchamp K, Grinaway P, Rustenburg B, Albanese S, Saladi S (2018) Choder-alab/Yank: More post-sams bugfixes. DOI 10.5281/zenodo.1226361, URL <https://zenodo.org/record/1226361>
85. Chodera J, Rizzi A, Naden L, Beauchamp K, Grinaway P, Fass J, Rustenburg B, Ross GA, Simmonett A, Swenson DW (2018) Choderalab/Openmmtools: 0.15.0 - restraint forces. DOI 10.5281/zenodo.1205753, URL <https://zenodo.org/record/1205753>
86. Rizzi A, Murkli S, McNeill JN, Yao W, Sullivan M, Gilson MK, Chiu MW, Isaacs L, Gibb BC, Mobley DL, Chodera JD (2018) Overview of the SAMPL6 host-guest binding affinity prediction challenge. J Comput Aid Mol Des DOI 10.1101/371724, URL 10.1101/371724GoogleScholar
87. Murkli S, McNeill JN, Isaacs L (2018) Cucurbit[8]uril-guest complexes: blinded dataset for the sampl6 challenge. Supramol Chem (submitted)

Fig. 1:
The indirect scheme for computing QM/MM free energies

Fig. 2:

A flow chart illustrating the process of gathering the information required for the FM procedure. The only deviation from this approach occurs for guest G13, which is described further in text

Fig. 3:
The guest molecules featured in the SAMPL6 Cucurbit[8]uril (CB[8]) host-guest challenge.
The standard set includes guest G0-G10, whereas the bonus molecules are G11-G13.

Fig. 4:

The thermodynamics cycle utilized to compute QM/(FM)/MM binding energies through FEP corrections performed at [ii] and [iii].

Fig. 5:
The extended cycle required to connect RFM/MM to QM/MM, with FM/MM acting as an intermediate state.

Fig. 6:

Plot of calculated binding values versus the experimentally determined numbers. Image and statistics provided courtesy of Andrea Rizzi.[86]

(i)

(ii)

Fig. 7:

(i) An example of an angle energy from G0 described as a combination of both a harmonic angle parameter set and UB parameter set. Although the angle parameter has an equilibrium value of zero, it is clearly observed that the minimum energy value occurs close to $\approx 130^\circ$. The dependence of $b^{1,3}$ on θ is easily obtained as a function of the CG2R61-CG2R61 bond lengths (in this case assumed to be constant and set to their minimum value of $b_0=1.3287 \text{ \AA}$) using the law of cosines, giving $b^{1,3} = b_0\sqrt{2(1 - \cos(\theta))}$. Atoms with type CG2R61 are labeled in (ii) with a dark violet color.

Fig. 8:
A figure illustrating the initial atom typing assignment versus the altered assignment of guest molecule G13.

Table 1:

Performance statistics, provided by the SAMPL6 organizers [86] based on using the correction scheme between RFM/MM and QM/MM without connecting to the full FM/MM parameter set.

	RMSE ^a	MAE	ME	R ²	<i>m</i> ^b	τ ^c
G0-G10	5.50	5.06	2.28	0.42	2.12	0.64
Standard ^d	(²² / ₃₆)	(²² / ₃₆)	(⁸ / ₃₆)	(² / ₃₆)	(³⁴ / ₃₆)	(¹ / ₃₆)
G0-G13	5.43	4.77	1.13	0.26	1.76	0.47
Bonus	(³ / ₆)	(³ / ₆)	(¹ / ₆)	(² / ₆)	(² / ₆)	(² / ₆)

^aRMSE is calculated with respect to experiment, and is in units of kcal/mol along with the Mean Absolute Error (MAE) and the Mean Error (ME)

^bThe coefficient of determination, R^2 and the slope of the best fit line between prediction and experiment, m , are obtained from linear regression

^c τ is the Kendall rank correlation coefficient.

^dRankings are provided with respect to the other submissions, with 36 for the standard set and only 6 also including the bonus set. For RMSE, MAE, and ME, the smaller they are the better the associated ranking, whereas for correlation based metrics (i.e., R^2 and τ) larger values are better.

Table 2:

Table summarizing results obtained with just the RFM potential and the QM/(RFM)MM free energy difference found after “correction”. Experimental results (exp) were provided by the SAMPL6 organizers [86, 87]. Standard deviations of potential energies and bias measure Π values are metrics to indicate potential problems. All units (excluding the unitless Π) are provided in kcal/mol

ID	$\Delta A_{bind}^{RFM/MM}$	$\Delta A_{bind}^{QM/(RFM)MM}$	exp.	$\sigma_{\Delta U}^G$	$\sigma_{\Delta U}^{HG}$	$\Pi_{\Delta U}^G$	$\Pi_{\Delta U}^{HG}$
G0	-8.05 ± 0.88	2.98 ± 1.56	-6.69 ± 0.05	10.76	12.95	-5.16	-7.74
G1	-1.43 ± 1.51	-2.01 ± 1.55	-7.65 ± 0.04	5.86	6.25	-2.26	-2.72
G2	-10.25 ± 0.20	-10.66 ± 1.05	-7.66 ± 0.05	5.68	5.87	-2.92	-3.09
G3	-10.82 ± 0.85	-10.60 ± 1.35	-6.45 ± 0.06	6.88	6.75	-3.86	-3.69
G4	-5.42 ± 0.51	-12.42 ± 1.29	-7.80 ± 0.04	6.12	6.06	-3.21	-4.26
G5	-16.29 ± 0.34	-16.32 ± 0.36	-8.18 ± 0.05	1.81	1.78	0.91	0.86
G6	-11.73 ± 0.23	-10.95 ± 0.26	-8.34 ± 0.05	1.42	1.45	1.26	1.46
G7	-13.34 ± 0.14	-12.73 ± 0.15	-10.00 ± 0.10	1.54	1.47	1.24	1.30
G8	-19.67 ± 0.30	-19.84 ± 0.35	-13.5 ± 0.04	2.33	2.26	0.17	0.08
G9	-13.87 ± 0.13	-13.72 ± 0.16	-8.68 ± 0.08	1.58	1.57	1.06	1.18
G10	-11.76 ± 0.20	-11.98 ± 0.21	-8.22 ± 0.07	1.57	1.59	1.26	1.12
G11	-7.70 ± 0.21	-8.64 ± 0.68	-7.77 ± 0.05	5.67	4.8	-2.76	-3.06
G12	-13.90 ± 2.40	-1.10 ± 2.41	-9.86 ± 0.03	12.96	12.95	-5.54	-5.49
G13	-4.45 ± 0.19	-5.68 ± 0.47	-7.11 ± 0.03	4.32	4.36	-2.32	-2.42

Table 3:

A table summarizing the number of terms (i.e., dihedral and improper) missing during the double decoupling scheme (i.e., RFM/MM calculations).

	Missing FM Dihedrals	Missing FM Improper Dihedrals	Total Missing FM Terms
G0	36	12	48
G1	36	12	48
G2	18	7	25
G3	3	11	14
G4	18	6	24
G5	0	0	0
G6	0	0	0
G7	0	0	0
G8	0	0	0
G9	0	0	0
G10	0	0	0
G11	0	0	0
G12	45	13	58
CB[8]	64	32	96

Table 4:

A summary of the parameter overlap, and a glimpse into the disparity caused via RMSD differences from minimization. All RMSD are in comparison to host minimizations with all terms present (full FM), however to provide a comparison between the FM with missing terms (RFM) is also provided. All RMSD values are in units of Å

Guest-Host Set	Parameters Overwritten	Gas Phase CB[8] RMSD
CB[8]+G0	2	0.081
CB[8]+G1	2	0.075
CB[8]+G2	2	0.090
CB[8]+G3	2	0.092
CB[8]+G4	2	0.059
CB[8]+G5	1	0.097
CB[8]+G6	2	0.082
CB[8]+G7	2	0.080
CB[8]+G8	2	0.082
CB[8]+G9	5	0.195
CB[8]+G10	2	0.101
CB[8]+G11	2	0.100
CB[8]+G12	2	0.082
CB[8]-RFM	-	0.050

Table 5:

RMSD between gas phase minimizations prepared with and without loading CGenFF parameter sets, and ΔE_{dihe}^{equil} of equilibrated structures (generated with prior loading of CGenFF) evaluated energetically with and without CGenFF parameters read in. Energies and RMSD are in units of kcal/mol and Å, respectively.

Guest	Gas Phase RMSD	ΔE_{dihe}^{equil}
G0	0.00	0.00
G1	0.00	0.00
G2	0.00	-0.01
G3	0.74	1.85
G4	0.18	2.44
G5	0.00	0.00
G6	0.02	-2.32
G7	0.05	-6.67
G8	0.19	-6.60
G9	0.00	0.00
G10	0.00	0.00
G11	0.00	0.00
G12	0.01	3.63
CB[8]	0.00	-

Table 6:

A table demonstrating the Pearson's correlation coefficients between quantities of interest. Specifically, the deviation from experimental (exp) results $|\epsilon_{err}| = |\text{exp.} - A^{calc}|$ based on correction from RFM, the standard deviations of $U^{RFM \rightarrow QM/MM}$ for the guest system and the host-guest complex ($\sigma_{\Delta U}^G$ and $\sigma_{\Delta U}^{HG}$, respectively), the uncertainty σ_A^{calc} obtained via block averaging, as well as the Π values for the guest and host-guest systems. For the missing terms (Dihe, Impr, and Total), the quantity in parenthesis is the correlation when comparing all the guest molecules (G0-G12), versus just the molecules that have omitted terms (G0, G1, G2, G3, G4, and G12 without parenthesis), and refer to the number of missing dihedrals, impropers, and the combined totals. Last, the amount of parameters in common between the guest system and CB[8] molecule is considered based on having atom type overlap.

	$ \epsilon_{err} $	$\sigma_{\Delta U}^G$	$\sigma_{\Delta U}^{HG}$	σ_A	$\Pi_{\Delta U}^G$	$\Pi_{\Delta U}^{HG}$	Missing			Par Overlap
							Dihe	Impr	Total	
$ \epsilon_A $	1.00	0.49	0.56	0.48	-0.35	-0.41	0.78(0.59)	0.71(0.48)	0.82(0.58)	-0.11
$\sigma_{\Delta U}^G$		1.00	0.99	0.95	-0.96	-0.94	0.67(0.85)	0.66(0.87)	0.71(0.88)	-0.17
$\sigma_{\Delta U}^{HG}$			1.00	0.93	-0.94	-0.95	0.69(0.87)	0.66(0.88)	0.73(0.90)	-0.16
σ_A				1.00	-0.92	-0.88	0.74(0.90)	0.72(0.94)	0.79(0.93)	-0.21
Π_G					1.00	0.98	-0.41(-0.77)	-0.52(-0.86)	-0.46(-0.81)	0.19
Π_{HG}						1.00	-0.45(-0.78)	-0.38(-0.84)	-0.47(-0.81)	0.19

Table 7:

Correction results based on utilizing the full FM parameter set via the parameter interpolation scheme, as opposed to the reduced (RFM) set, along with the corresponding σ_U and Π_U of $U^{FM \rightarrow QM/MM}$ for both the host-guest complex and guest only system, as well as experimental results (exp). All values, with the exception of Π , are provided in kcal/mol

ID	$\Delta A_{bind}^{RFM/MM}$	$\Delta A_{bind}^{QM/(RFM)MM}$	$\Delta A_{bind}^{FM/MM}$	$\Delta A_{bind}^{QM/(FM)MM}$	exp	$\sigma_{\Delta U}^G$	$\sigma_{\Delta U}^{HG}$	$\Pi_{\Delta U}^G$	$\Pi_{\Delta U}^{HG}$
G0	-8.05 ± 0.88	2.98 ± 1.56	-4.72 ± 0.88	-5.17 ± 1.55	-6.69	3.43	3.69	-1.89	-1.34
G1	-1.43 ± 1.51	-2.01 ± 1.55	-1.51 ± 1.51	-2.49 ± 1.60	-7.65	2.50	2.39	-0.26	-0.72