PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Toward Polarizable AMOEBA Thermodynamics at Fixed Charge Efficiency Using a Dual Force Field Approach: Application to Organic Crystals

Ian J. Nessler[1], Jacob M. Litman[2]

and Michael J. Schnieders[2,3,*]

[1]Department of Chemical Engineering,

[2]Department of Biochemistry,

and

[3]Department of Biomedical Engineering,

University of Iowa, Iowa City, IA 52242

*Corresponding Author:

Email: michael-schnieders@uiowa.edu

Abstract

First principles prediction of the structure, thermodynamics and solubility of organic molecular crystals, which play a central role in chemical, material, pharmaceutical and engineering sciences, challenges both potential energy functions and sampling methodologies. Here we calculate absolute crystal deposition thermodynamics using a novel dual force field approach whose goal is to maintain the accuracy of advanced multipole force fields (e.g. the polarizable AMOEBA model) while performing more than 95% of the sampling in an inexpensive fixed charge (FC) force field (e.g. OPLS-AA). Absolute crystal sublimation/deposition phase transition free energies were determined using an alchemical path that grows the crystalline state from a vapor reference state based on sampling with the OPLS-AA force field, followed by dual force field thermodynamic corrections to change between FC and AMOEBA resolutions at both end states (we denote the three step path as AMOEBA/FC). Importantly, whereas the phase transition requires on the order of 200 nsec of sampling per compound, only 5 nsec of sampling was needed for the dual force field thermodynamic corrections to reach a mean statistical uncertainty of 0.05 kcal/mol. For five organic compounds, the mean unsigned error between direct use of AMOEBA and the AMOEBA/FC dual force field path was only 0.2 kcal/mol and not statistically significant. Compared to experimental deposition thermodynamics, the mean unsigned error for AMOEBA/FC (1.4 kcal/mol) was more than a factor of two smaller than uncorrected OPLS-AA (3.2 kcal/mol). Overall, the dual force field thermodynamic corrections reduced condensed phase sampling in the expensive force field by a factor of 40, and may prove useful for protein stability or binding thermodynamics in the future.

Introduction

Organic molecular crystals play a central role in chemical, material, pharmaceutical and engineering sciences.[1-4] First principles prediction of their structure[5-9], thermodynamics[10, 11] and solubility[12, 13] is a challenge for both potential energy functions and sampling methodologies.[14, 15] For example, it has been shown that classical force fields based on fixed atomic partial charges (FC), such as Amber[16], CHARMM[17] and OPLS-AA[18], lack the accuracy needed to correctly rank the relative stability of alternative polymorphs.[19, 20] To achieve sufficient accuracy for crystal structure prediction, atomic multipole expansions can be used to systematically reproduce the electrostatic potential outside the van der Waals surface of a rigid molecule, as defined by electronic structure calculations.[21, 22] However, molecular charge distributions are sensitive to both conformation and the electric field across the molecule, due in part to electronic polarization.[23] Thus, to apply the configurational sampling algorithms needed to quantify thermodynamics, multipolar force fields must address transferability of their multipole moments as a function of molecular conformation, while also consistently treating both intra- and intermolecular polarization.[23, 24] A few examples of multipolar energy functions include AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications)[24-26], GMM (Gaussian Multipolar Model)[27, 28], SIBFA (Sum of Interactions Between Fragments *Ab Initio* Computed)[29, 30] and NEMO (Non-empirical Molecular Orbital)[31], which are described in recent reviews[32-34]. The accuracy and transferability improvements of AMOEBA relative to FC force fields have been demonstrated in the context of water[25, 35], ion solvation[36], the properties of small organic molecules[24, 26, 36-38] and for protein energetics.[34, 39]

The increased domain of applicability of advanced multipolar force fields, however, comes at a price of greater computational expense relative to FC force fields by a factor of 5-10 or more for energy and force evaluations. To ameliorate the expense of sampling advanced potential energy functions, previous work to reweight from molecular mechanics (MM) sampling has been explored in the context of determining thermodynamics for quantum mechanical (QM) or QM/MM potential energy surfaces.[40-49] Recent work includes the dual-topology alchemical Hamiltonian replica exchange method (DTA-HREM)[50-52], non-Boltzmann Bennett (NBB) reweighting[53-55], and the multistate Bennett acceptance ratio (MBAR)[56, 57] approaches. The emergence of increasingly sophisticated polarizable atomic multipole force fields for organic molecules and proteins further motivates approaches that either reweight FC trajectories or

define a path that smoothly connects FC states to those defined by a more advanced multipolar force field.

Perhaps the simplest approach to computing the free energy difference between FC and more advanced force fields, such as AMOEBA, is direct reweighting via the Zwanzig relationship

$$\Delta G_{\mathrm{FC}\rightarrow\mathrm{AMOEBA}} = -k_B T \cdot \ln \langle e^{-\frac{E_{\mathrm{AMOEBA}} - E_{\mathrm{FC}}}{k_{\mathrm{B}} T}} \rangle_{\mathrm{FC}}$$

Equation 1.

where k_{B} is Boltzmann's constant, T is temperature in degrees Kelvin and the angle brackets denote an ensemble average.[58] This approach evaluates the potential energy of the expensive model only at intermediate samples. The reverse perturbation offers no efficiency improvement due to requiring direct sampling of the more expensive ensemble

$$\Delta G_{\mathrm{AMOEBA}\rightarrow\mathrm{FC}} = -\mathrm{k_B} T \cdot \ln \langle e^{-\frac{E_{\mathrm{FC}} - E_{\mathrm{AMOEBA}}}{k_{\mathrm{B}} T}} \rangle_{\mathrm{AMOEBA}}$$

Equation 2.

However, convergence of reweighting as shown in Eqs. 1 or 2 may fail due to lack of phase space overlap between force fields, which arises to differences in bonded terms (i.e. equilibrium bond distances or bond angles) and/or intermolecular contact distances (i.e. due to the balance of van der Waals and Coulombic interactions)[59, 60]. Phase space overlap between resolutions can be improved by coordinating their design and parameterization[60-62].

Here we explore an approach that performs more than 95% of the sampling using an inexpensive fixed charge force field, followed by the addition of two rigorous corrections to recover thermodynamics consistent with the more advanced force field. The method mitigates nontrivial differences between force fields by defining a dual force field (DFF) potential that enables explicit sampling of the transition between resolutions

$$U_{\mathrm{DFF}}(\lambda, \mathbf{X}) = \lambda \cdot U_{\mathrm{AMOEBA}}(\mathbf{X}) + (1 - \lambda) \cdot U_{\mathrm{FC}}(\mathbf{X})$$

Equation 3.

where U_{DFF} defines a smooth transition between the fixed charge energy function at $\lambda = 0$ and the polarizable AMOEBA energy function at $\lambda = 1$. Calculation of free energy differences due to force field resolution changes at both vapor and crystalline end states permits the computationally demanding sublimation/deposition phase transition to be sampled with the

inexpensive force field (Figure 1). The approach has similarities to "dual topology" style potential energy functions, which smoothly interpolate between chemical functional groups[63].

[[Figure 1 About Here]]

Using five organic compounds, we show that the DFF approach defines a thermodynamic path that is more efficient than direct simulation in AMOEBA, but maintains its agreement with experiment in the context of calculating absolute sublimation/deposition phase transitions. OPLS-AA[18, 64] and AMOEBA[24-26] were used as the fixed charge (i.e. cheap) and polarizable multipole (i.e. expensive) force fields, respectively, as implemented in the open source Force Field X (FFX) software package (http://ffx.biochem.uiowa.edu)[65-67] version 1.0.0-beta. Overall, the DFF interpolation between resolutions exhibits rapid convergence relative to the phase transition and allows decreased condensed phase sampling in the expensive potential by a factor of 40 for a thermodynamic path characterized by Growth of the Asymmetric Unit into a Crystal via alCHemy (GAUCHE) and described previously[14, 15]. Thus, polarizable atomic multipole AMOEBA thermodynamics are reproduced with an expense approaching that of fixed charge models.

Methods

Lattice Potential Energies

We analyzed five molecules from a prior study on deposition thermodynamics.[15] These compounds are shown Figure 2 and include acetanilide,[68] paracetamol (polymorph I),[69] methyl paraben (polymorph II),[70] ethyl paraben,[71] and phenacetin.[72] Experimental lattice parameters and the space group for each compound are given in Table 1, along with asymmetric unit composition and unit cell volume in Table 2. The deposition free energy values from the prior study, computed using AMOEBA directly, will be compared to the three step DFF thermodynamic path that combines FC phase transitions and DFF corrections. Analogous to previous work[15], the five molecules were optimized in the crystalline and vapor states using both the AMOEBA and OPLS-AA force fields within the FFX program. Space group and lattice parameters were constrained to their experimental values (Table 1) for both AMOEBA and OPLS-AA optimizations. AMOEBA parameters were obtained from Poltype[38], while OPLS-AA 2005 parameters were obtained from Schrödinger[73]. To calculate lattice energies, the minimized vapor energy was subtracted from the minimized crystal energy on a per molecule basis.

$$U_{\text{lattice}} = U_{\text{cryst}} - U_{\text{vac}}$$

Equation 4.

A van der Waals cutoff of 12.0 Å was used, with a multiplicative switch tapering the interaction energy to zero starting at 10.8 Å, which is consistent with parameterization of AMOEBA the model for water, organic molecules and the protein force field[24, 25, 39]. Polarizable electrostatic evaluations were conducted using a smooth[74] particle-mesh Ewald[75] (PME) algorithm for multipoles[76] which maintained the self-consistent field with a convergence criterion of 10^{-5} RMS Debye and supported space group symmetry.[65] For this work, PME parameters included a real-space cutoff of 9 Å, a mesh density of 2.0 grid points per Å, eighth order B-splines, and an Ewald Parameter of 0.42. The molecules were minimized to a tight RMS gradient convergence criterion of 10^{-4} kcal/mol/Å.

[[Figure 2 About Here]]

[[Table 1 About Here]]

[[Table 2 About Here]]

Deposition Free Energy and Dual Force Field Corrections

The DFF method, as employed in this work, adds a free energy correction composed of two terms to the deposition free energy calculated using the inexpensive OPLS-AA force field to recover direct AMOEBA thermodynamics, at approximately OPLS-AA efficiency. Consistent with our earlier work on these systems[15], the NVT ensemble was sampled at 298 degrees Kelvin using stochastic dynamics. For each of the three simulation legs, five independent trajectories were collected beginning from different random velocity vectors. The first simulation leg calculates the free energy to change from AMOEBA resolution into OPLS-AA resolution in vapor $\Delta G^{\text{Vapor}}_{\text{AMOEBA}\rightarrow\text{FC}}$. Next, the inexpensive potential $U_{\text{FC}}(\mathbf{X})$ is sampled to determine the FC deposition/sublimation phase transition free energy $\Delta G^{\text{Dep}}_{\text{FC}}$. Finally, the last simulation leg calculates the free energy to convert back from OPLS-AA resolution to AMOEBA resolution in the crystal state $\Delta G^{\text{Crystal}}_{\text{FC}\rightarrow\text{AMOEBA}}$. Summing the three simulation legs yields the AMOEBA/FC deposition free energy $\Delta G^{\text{Dep}}_{\text{AMOEBA/FC}}$

$$\Delta G^{\text{Dep}}_{\text{AMOEBA/FC}} = \Delta G^{\text{Vapor}}_{\text{AMOEBA}\rightarrow\text{FC}} + \Delta G^{\text{Dep}}_{\text{FC}} + \Delta G^{\text{Crystal}}_{\text{FC}\rightarrow\text{AMOEBA}}$$

Equation 5.

Orthogonal Space Sampling of the Thermodynamic Paths

The Orthogonal Space Random Walk (OSRW) method builds up a time-dependent bias by depositing two-dimensional Gaussian-shaped repulsive potentials as a function of the state variable λ and the derivative of the potential energy with respect to λ ($F_\lambda = \partial U/\partial\lambda$)[77, 78]. The total potential energy is then given by

$$U_m = U_{DFF}(\lambda, \mathbf{X}) + f_m(\lambda) + g_m(\lambda, F_\lambda)$$

Equation 6.

where $g_m(\lambda, F_\lambda)$ is the sum of the repulsive potentials (i.e. hills) centered at states given by $[\lambda(t_i), F_\lambda(t_i)]$[14]:

$$g_m(\lambda, F_\lambda) = \sum_{t_i} h \cdot e^{\left(\frac{|\lambda-\lambda(t_i)|^2}{2w_1^2} \times \frac{|F_\lambda - F_\lambda(t_i)|^2}{2w_2^2}\right)}$$

Equation 7.

The additional biasing dimension promotes crossing of hidden barriers relative to the simpler one-dimensional bias of original metadynamics approaches.[77, 78] We note that the gradient of U_{m} (i.e. the partial derivatives with respect to all atomic coordinates), which is needed to integrate equations of motion during OSRW dynamics, requires partial derivatives of the target function $U_{\mathrm{DFF}}(\lambda, \mathbf{X})$ that include $\partial U_{DFF}(\lambda, \mathbf{X})/\partial\lambda$, $\partial^2 U_{DFF}(\lambda, \mathbf{X})/\partial\lambda^2$ and $\partial^2 U_{DFF}(\lambda, \mathbf{X})/\partial\lambda\partial X$. These are given by

$$\partial U_{DFF}(\lambda, \mathbf{X})/\partial\lambda = U_{\mathrm{AMOEBA}}(\mathbf{X}) - U_{\mathrm{FC}}(\mathbf{X})$$

Equation 8.

$$\partial^2 U_{DFF}(\lambda, \mathbf{X})/\partial\lambda^2 = 0$$

Equation 9.

and

$$\partial^2 U_{DFF}(\lambda, \mathbf{X})/\partial\lambda\partial X = \partial U_{\mathrm{AMOEBA}}(\mathbf{X})/\partial X - \partial U_{\mathrm{FC}}(\mathbf{X})/\partial X$$

Equation 10.

The first two results above (Equations 8 and 9) are the 1st and 2nd partial derivatives of the dual force field potential energy (Equation 3) with respect to λ, while the last result (Equation 10) is equivalent to the difference in the partial derivative for each force field with respect to atomic

coordinate *X*. As shown in Figure 3, the ensemble average thermodynamic force $\langle \partial U/\partial \lambda \rangle$ is smooth and well-behaved as resolutions between the OPLS-AA and AMOEBA force fields are sampled (i.e. for $0 \leq \lambda \leq 1$). However, if this had not been the case (i.e. for force fields that are more dramatically different), higher powers of λ could be explored (i.e. in Equation 3 apply the substitution $\lambda \rightarrow \lambda^2$).

[[Figure 3 About Here]]

Simulations with Non-Crystallographic Symmetry

In cases where more than one molecule is present in the asymmetric unit (i.e. non-crystallographic symmetry), intermolecular interactions are smoothly turned off as the simulation transitions from the crystalline state to the vapor state. This allows each molecule to be independent in the vapor state (i.e. they can pass through each other). For example, due to non-crystallographic symmetry in the ethyl paraben crystal (Table 2), two molecules were simulated and computed deposition values normalized by a factor of two.

Results

Lattice Potential Energies

Displayed in Table 3 is the lattice energy for each compound using both OPLS-AA and AMOEBA force fields. The mean absolute difference for OPLS-AA relative to AMOEBA of 3.6 kcal/mol is significant relative to the goal of achieving chemical accuracy (i.e. ~1.0 kcal/mol). In particular, the three amide containing compounds (acetanilide, paracetamol and phenacetin) show increased crystalline stability of 4.0 to 6.6 kcal/mol under OPLS-AA relative to AMOEBA. The differences for the ester containing methyl and ethyl paraben compounds of 0.5 and 1.2 kcal/mol, respectively, are more modest.

[[Table 3 About Here]]

Deposition Free Energy and Dual Force Field Corrections

As shown in Figure 1 above, the accelerated thermodynamic pathway consists of three steps. First, the transition to the cheap force field is completed in the vapor state $\Delta G^{\text{Vapor}}_{\text{AMOEBA}\rightarrow\text{FC}}$, followed by the deposition/sublimation phase transition for the asymmetric unit using the cheap force field $\Delta G^{\text{AU Dep}}_{\text{FC}}$, and finally the transition back to the expensive force field in the crystalline state $\Delta G^{\text{Crystal}}_{\text{FC}\rightarrow\text{AMOEBA}}$. Convergence of the independent deposition/sublimation phase transition

trajectories for the asymmetric unit in both the expensive (i.e. AMOEBA) and cheap (i.e. OPLS-AA) force fields is presented in Figure 4 for both acetanilide and paracetamol. Convergence of the independent DFF transition trajectories in vapor and crystalline phases is shown in Figure 5 for both acetanilide and paracetamol. Free energy differences were computed by collecting the ensemble average thermodynamic force $\langle \partial U/\partial \lambda \rangle$ for 200 equally sized bins along the λ parameterized thermodynamic path, followed by numerical thermodynamic integration. Simulation legs were considered to be converged once the mean free energy difference for the five independent trajectories changed by less than 0.1 kcal/mol for the last ¼ the trajectory.

[[Figure 4 About Here]]

[[Figure 5 About Here]]

The total DFF corrections in Table 4 follow the trend of the lattice potential energy differences shown in Table 3, albeit slightly smaller in magnitude in all cases. After 5 nsec of sampling for both of the dual force field legs, the mean standard deviation of the free energy difference had fallen to 0.05 kcal/mol (Table 4). The direct deposition/sublimation phase transitions for the asymmetric unit using OPLS-AA and AMOEBA[15] are shown in Table 5. The differences between AMOEBA $\Delta G_{\mathrm{AMOEBA}}^{\mathrm{AU\ Dep}}$ and OPLS-AA $\Delta G_{\mathrm{FC}}^{\mathrm{AU\ Dep}}$ follow the trends seen in Table 3 for lattice potential energy differences, with OPLS-AA showing greater stabilization than AMOEBA. For example, the mean unsigned error of 3.1 kcal/mol is only slightly smaller than the mean lattice potential energy difference of 3.6 kcal/mol. The standard deviation for the phase transition free energy after 200 nsec of sampling was for 0.41 kcal/mol for AMOEBA and 0.27 kcal/mol for OPLS-AA (Table 5). The DFF corrections from Table 4 were added to the pure OPLS-AA deposition values $\Delta G_{\mathrm{FC}}^{\mathrm{AU\ Dep}}$ in Table 5 to yield corrected values denoted $\Delta \mathrm{G}_{\mathrm{AMOEBA/FC}}^{\mathrm{AU\ Dep}}$. Simulation legs were considered to be converged once the mean free energy difference for the five independent trajectories remained approximately constant for the last ¼ or more of the trajectory.

[[Table 4 About Here]]

[[Table 5 About Here]]

We note that the statistical uncertainty foxr the DFF AMOEBA/FC deposition is almost completely due to the OPLS-AA sublimation/deposition phase transition ($\Delta G_{\mathrm{FC}}^{\mathrm{AU\ Dep}}$) step of the thermodynamic cycle, and not from the DFF corrections. To further reduce the statistical

uncertainty, this analysis suggests the focus should be on the OPLS-AA deposition/sublimation phase transition, which is further considered in the Conclusions below. Comparison of pure AMOEBA to the DFF AMOEBA/FC results show a mean unsigned error of only 0.2 kcal/mol, indicating successful application of the DFF thermodynamic path (Figure 1 and Eq. 5). Overall, the DFF method reduced the amount of condensed phase sampling in the more expensive AMOEBA force field by a factor of 40 (i.e. 200 nsec per trial per compound was reduced to only 5 nsec). Although the current version of FFX (1.0.0-beta) does not include optimized code for FC electrostatics, the wall clock time saved using OPLS-AA relative to AMOEBA was a factor of ~2 in this work due to elimination of the self-consistent field calculation (i.e. fixed multipole interactions with zero dipole and quadrupole components are computed for OPLS-AA). In the future, it is reasonable to expect the speed-up of the DFF AMOEBA/FC approach should reach a factor of ~5, based on codes that implement relatively optimized code paths for both fixed partial charge (i.e. OPLS-AA) and polarizable atomic multipole (i.e. AMOEBA) force fields such as TINKER[79].

Absolute Deposition Thermodynamics vs. Experiment

The AMOEBA and AMOEBA/FC asymmetric unit deposition free energy values given in Table 5 can be compared to experiment after addition of 1) an ideal gas correction to account for compressing a 1 molar vapor into the volume of the crystal and 2) a correction to account for removal of the perfect symmetry constraint applied during the sublimation/deposition phase transition simulations. This later correction, a part of the GAUCHE path, has been described previously.[15] As shown in Table 6, both AMOEBA and AMOEBA/FC approaches produce absolute deposition free energy values that compare favorably to experiment, with mean unsigned errors of 1.6 and 1.4 kcal/mol, respectively. For all compounds, the difference between AMOEBA and the accelerated AMOEBA/FC DFF result is not significant based on Student's t-test. While OPLS-AA and AMOEBA deposition thermodynamics are clearly different for some crystals, the results from the AMOEBA/FC dual force field path are not distinguishable from the direct AMOEBA path (Figure 5).

[[Table 6 About Here]]

[[Figure 5 About Here]]

Decomposition into Enthalpic and Entropic Contributions

Crystal structure prediction and the ranking of polymorphs is often based on direct use of potential energy rather than thermodynamic stability (i.e. free energy)[80]. Although efficient, methods that neglect entropic contributions are unable to describe changes in polymorph stability as a function of temperature. To overcome this common approximation, the GAUCHE procedure was developed to efficiently calculate absolute deposition free energy[15]. Insights into the origin of crystal stability differences can sometimes be obtained by decomposing free energy differences into enthalpic and entropic contributions using the relationships

$$\Delta \mathrm{H}^{\mathrm{Dep}} = \langle U_{\mathrm{crystal}} \rangle - \langle U_{\mathrm{vapor}} \rangle$$

Equation 11.

and

$$-T\Delta \mathrm{S}^{\mathrm{Dep}} = \Delta \mathrm{G}^{\mathrm{Dep}} - \Delta \mathrm{H}^{\mathrm{Dep}}$$

Equation 12.

where temperature (T) is 298 degrees Kelvin for the current work and the NVT ensemble was sampled using stochastic dynamics with experimental unit cell parameters (Table 1). The importance of entropic contributions is shown by comparing acetanilide to methyl and ethyl paraben in Table 7; while methyl and ethyl paraben have lower enthalpy of deposition under AMOEBA, acetanilide's lesser entropic penalty results in a more overall favorable deposition free energy.

[[Table 7 Here]]

Conclusions

The DFF approach combines the strengths of both advanced polarizable atomic multipole force fields and efficient fixed partial charge models for organic crystal thermodynamics, while mitigating their primary weaknesses (i.e. FC accuracy limitations and the increased computational cost of AMOEBA). The AMOEBA/FC thermodynamic path was both accurate and cost effective for acetanilide, phenacetin, methyl parben, ethyl paraben and paracetamol crystals with a MUE of 1.4 kcal/mol relative to experiment, which is substantially less than the OPLS-AA MUE of 3.0 kcal/mol. Furthermore, the AMOEBA/OPLS-AA DFF method was statistically indistinguishable from using AMOEBA directly, with a MUE of only 0.2 kcal/mol

relative to AMOEBA. Finally, the DFF protocol enabled sampling a path between energy functions with large inherent differences in both their bonded (i.e. equilibrium bond and angle values) and non-bonded functional forms (i.e. van der Waals, permanent electrostatics and explicit polarization). This serves to overcome limitations in reweighting procedures (e.g. the Zwanzig relationship) that require significant phase space overlap.

In future work, we plan to incorporate transition-tempering into the orthogonal space sampling algorithm (i.e. transition-tempered OSRW) to further reduce statistical uncertainty[81], especially for the sublimation/deposition phase transition step. We also plan to replace the three discrete simulations that form the thermodynamic cycle described in Figure 1 with a single simulation that "on-the-fly" turns AMOEBA into OPLS-AA at the beginning of the thermodynamic path (i.e. in vacuum) and then back into AMOEBA at the end (i.e. in the crystalline state). This will serve to avoid any discrepancy in the optimal unit cell parameters or coordinates between force field resolutions for NPT ensembles. We also plan to explore the domain of applicability of DFF thermodynamic paths for applications beyond crystal thermodynamics, including acceleration of small molecule solvation thermodynamics[82], protein/ligand binding[83] and protein folding stability.[84]

Acknowledgements

We thank Dr. Ed Harder (Schrödinger, LLC) for providing the OPLS-AA parameters used in this work. All computations were performed on The University of Iowa Neon cluster with support and guidance from Glenn Johnson and Ben Rogers. NSF Award CHE-1404147 and NIH Award R01 DC002842 from NIDCD supported MJS. JML acknowledges an NIH fellowship from Award T32 GM008365 and a University of Iowa Presidential Fellowship. IJN was partially supported by fellowships from The University of Iowa Center for Research by Undergraduates. The authors declare that no competing interests exist.

Tables

Table 1. Compounds studied and their associated CSD reference codes, space groups and unit cell parameters. Roman numerals following paracetamol and methyl paraben correspond to polymorph.

Compound	CSD Code	Space Group	*a*	*b*	*c*	*α*	*β*	*γ*
Acetanilide	ACANIL	Pbca	19.64	9.48	7.98	90	90.0	90
Paracetamol I	HXACAN01	P21/c	11.72	9.40	12.93	90	147.0	90
Methyl Paraben II	CEBGOF03	P21/c	4.82	14.63	10.24	90	99.8	90
Ethyl Paraben	FEGLEI	P21/c	13.76	13.18	11.58	90	125.5	90
Phenacetin	PYRAZB10	P21/c	13.25	9.65	7.81	90	104.9	90

Table 2. Molecular weight, number of molecules per asymmetric unit (AU), unit cell (UC) volume, number of unit cell molecules, volume per molecule and experimental temperature for each crystal studied.

Compound	Mol. Weight (g/mol)	AU Molecules	UC Vol. (Å^3)	Z	Vol./Z (Å^3)	Temp. (K)
Acetanilide	135.16	1	1486.1	8	185.8	297
Paracetamol I	151.16	1	776.3	4	194.1	297
Methyl Paraben II	152.15	1	711.3	4	177.8	100
Ethly Paraben	166.17	2	1710.2	8	213.8	297
Phenacetin	179.22	1	965.0	4	241.3	297

Table 3. OPLS-AA and AMOEBA lattice potential energies for each compound (kcal/mol).

	OPLS-AA			AMOEBA			
Compound	**Crystal**	**Vacuum**	**Lattice**	**Crystal**	**Vacuum**	**Lattice**	**Diff.**
Acetanilide	-28.73	-2.01	-26.72	-38.05	-15.32	-22.73	3.99
Paracetamol	-44.45	-10.12	-34.32	-41.04	-13.35	-27.70	6.63
Methyl Paraben	-23.58	-0.13	-23.45	-30.77	-8.48	-22.29	1.16
Ethyl Paraben	-24.33	0.78	-25.11	-28.56	-3.98	-24.58	0.53
Phenacetin	-37.49	-5.70	-31.80	-36.05	-9.92	-26.13	5.67
Mean			-28.28			-24.68	3.60

Table 4. DFF thermodynamic corrections for each compound in vapor $\Delta G^{Vapor}_{AMOEBA \rightarrow FC}$ and crystalline $\Delta G^{Crystal}_{FC \rightarrow AMOEBA}$ phases after 5 nsec of sampling (kcal/mol). Each value is the mean of 5 trials ± the standard deviation.

Compound	$\Delta G^{Vapor}_{AMOEBA \rightarrow FC}$	$\Delta G^{Crystal}_{FC \rightarrow AMOEBA}$	Total
Acetanilide	14.00±0.02	-10.12±0.02	3.87±0.03
Paracetamol	2.72 ±0.03	2.42±0.02	5.14±0.04
Methyl Paraben	8.48±0.01	-7.57±0.02	0.91±0.02
Ethyl Paraben	4.32±0.01	-4.14±0.03	0.18±0.03
Phenacetin	3.66±0.03	1.16±0.14	4.81±0.15

Table 5. Shown are asymmetric unit absolute deposition free energy values for AMOEBA ($\Delta G^{AU\ Dep}_{AMOEBA}$), OPLS-AA ($\Delta G^{AU\ Dep}_{FC}$) and OPLS-AA corrected to AMOEBA ($\Delta G^{AU\ Dep}_{AMOEBA/FC}$). In the latter case, the total DFF corrections (Table 4) have been added the OPLS-AA deposition values (kcal/mol).

Compound	$\Delta G^{AU\ Dep}_{AMOEBA}$	$\Delta G^{AU\ Dep}_{FC}$		$\Delta G^{AU\ Dep}_{AMOEBA/FC}$	
		Deposition	UE	Deposition	UE
Acetanilide	-12.42±0.29	-16.06±0.12	3.64	-12.19±0.12	0.23
Paracetamol	-13.98±0.57	-19.43±0.29	5.45	-14.30±0.29	0.32
Methyl Paraben	-9.81±0.29	-10.80±0.16	0.99	-9.89±0.16	0.08
Ethyl Paraben	-10.31±0.71	-10.86±0.37	0.55	-10.68±0.37	0.37
Phenacetin	-14.47±0.21	-19.10±0.39	4.63	-14.29±0.40	0.18
Mean	-12.20±0.41	-15.25±0.27	3.05	-12.27±0.27	0.23

Table 6. Shown are absolute deposition free energy values for AMOEBA and AMOEBA/FC after including ideal gas and GAUCHE corrections (kcal/mol). The unsigned errors (UE) are relative to experiment.

Compound	Expt.	Ideal Gas Correction	GAUCHE Correction	ΔG^{Dep}_{AMOEBA}	UE	$\Delta G^{Dep}_{AMOEBA/FC}$	UE
Acetanilide	-11.57	1.30	-0.25±0.24	-11.37±0.38	0.20	-11.14±0.27	0.43
Paracetamol	-16.23	1.27	-1.67±0.18	-14.38±0.60	1.85	-14.70±0.34	1.53
Methyl Paraben	-11.98	1.32	-1.74±0.14	-10.23±0.32	1.75	-10.31±0.21	1.67
Ethyl Paraben	-12.27	1.21	-1.19±0.18	-10.29±0.73	1.98	-10.66±0.41	1.61
Phenacetin	-14.39	1.14	-3.08±0.36	-16.41±0.42	2.02	-16.23±0.54	1.84
Mean					1.56		1.42

Table 7. The decomposition of AMOEBA absolute deposition free energy values for each compound into enthalpic and entropic contributions at 298 degrees Kelvin.

Compound	ΔG^{Dep}_{AMOEBA}	ΔH^{Dep}_{AMOEBA}	$-T\Delta S^{Dep}_{AMOEBA}$
Acetanilide	-11.37±0.38	-20.81±0.15	9.44±0.40
Paracetamol	-14.38±0.60	-25.93±0.18	11.55±0.63
Methyl Paraben	-10.23±0.32	-21.19±0.23	10.96±0.39
Ethyl Paraben	-10.29±0.73	-23.30±0.24	13.01±0.77
Phenacetin	-16.41±0.42	-22.96±0.28	6.55±0.50
Mean	-12.54	-22.84	10.30

Figures

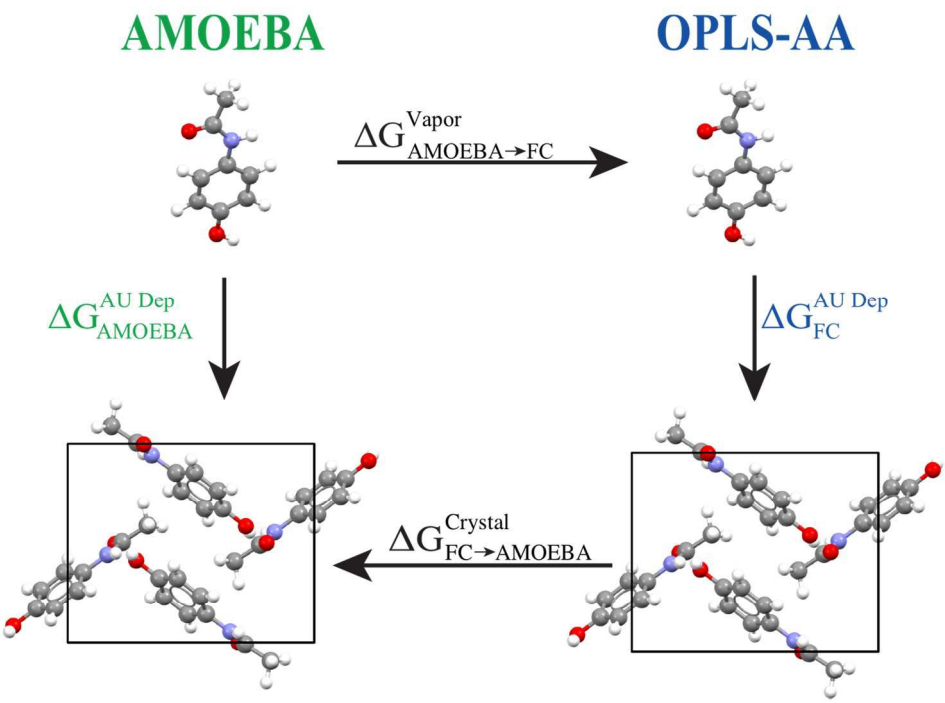

Figure 1. This diagram summarizes the thermodynamic cycle for computing absolute deposition free energy at the accuracy of the polarizable AMOEBA force field, but with an efficiency approaching the fixed atomic partial charge OPLS-AA force field. The vertical sublimation/deposition phase transition steps each require ~200 nsec of sampling while the horizontal steps to change resolution converge in only 5 nsec. This equates to a factor of 40 reduction in the amount condensed phase AMOEBA sampling required.

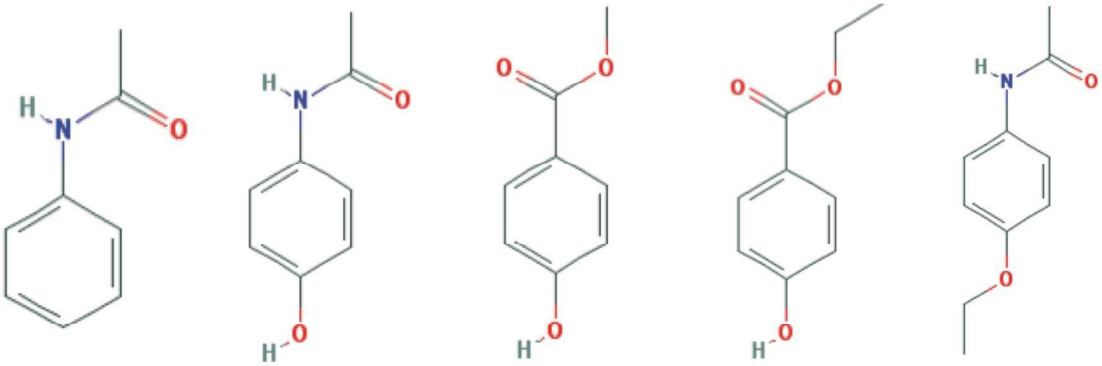

Figure 2. The structure of each organic compound.

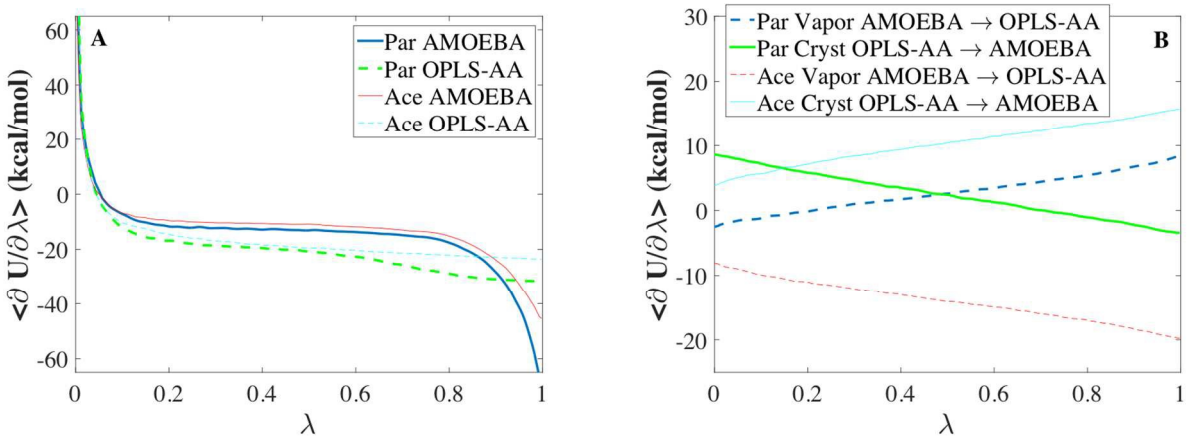

Figure 3. The ensemble average thermodynamic force $\langle \partial U/\partial \lambda \rangle$ along the depositon path (Panel A) and dual force field paths (Panel B) are shown for paracetamol (Par) and acetanilide (Ace).

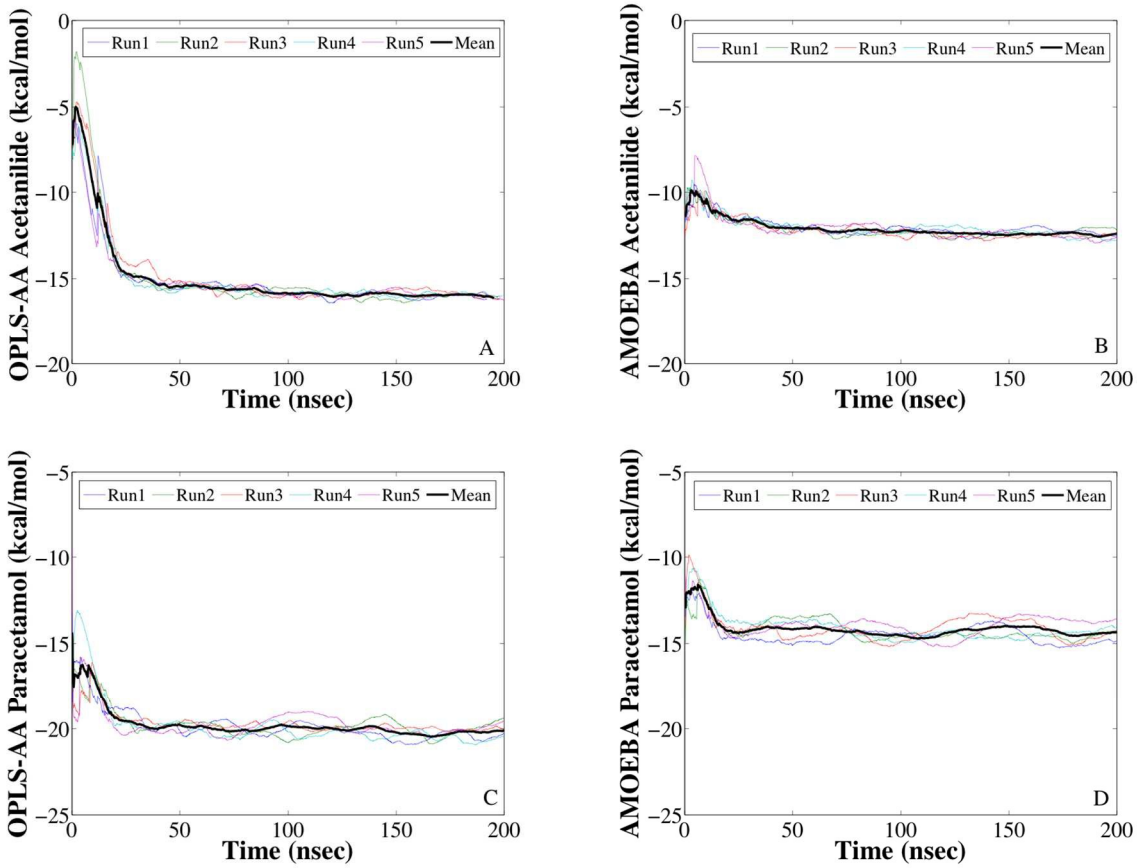

Figure 4. Convergence of the deposition free energy is shown for alchemical simulations of the asymmetric unit of acetanilide (Panels A & B) and paracetamol (Panels C & D). Panels A & C use the OPLS-AA force field while Panels B & D use the AMOEBA force fields.

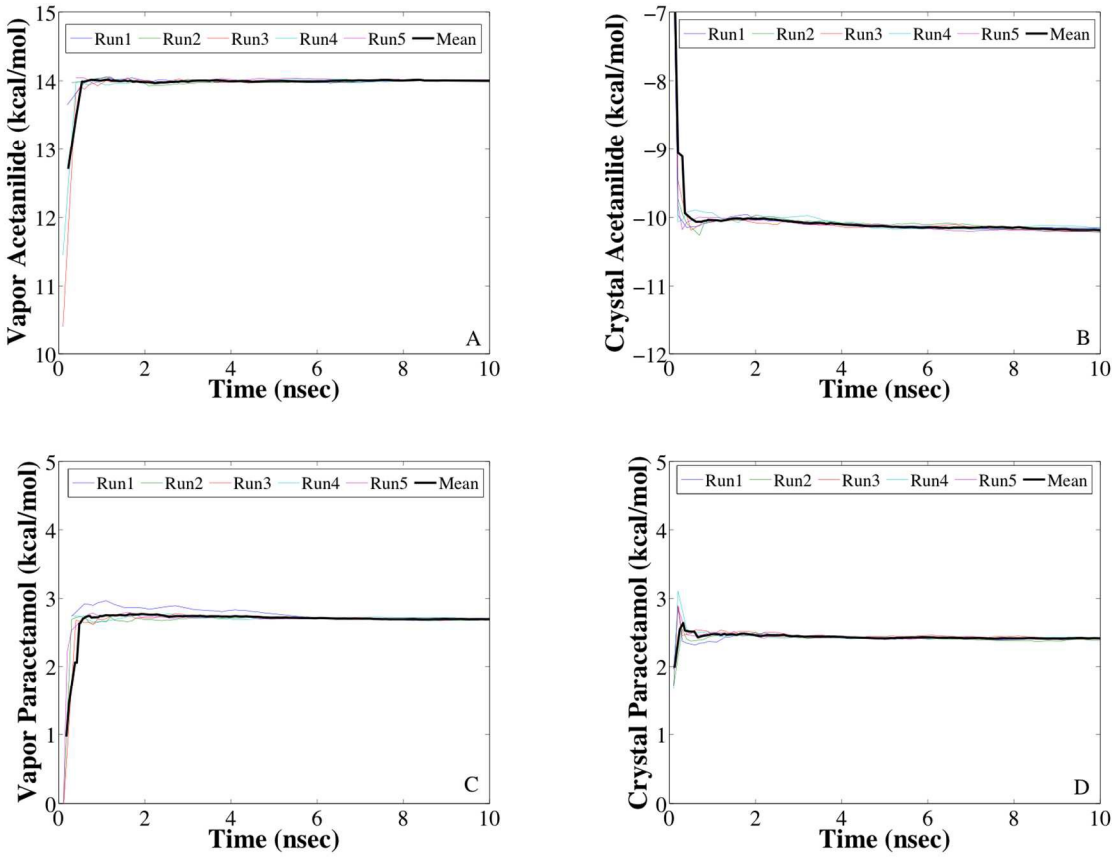

Figure 5. Convergence of the free energy change for acetanilide (Panels A & B) and paracetamol (Panels C & D) dual force field simulations are shown. In Panels A & C, the AMOEBA force field is transformed into the OPLS-AA force field in vapor. In Panels B & D, the OPLS-AA force field is transformed into the AMOEBA force field in the crystalline environment.

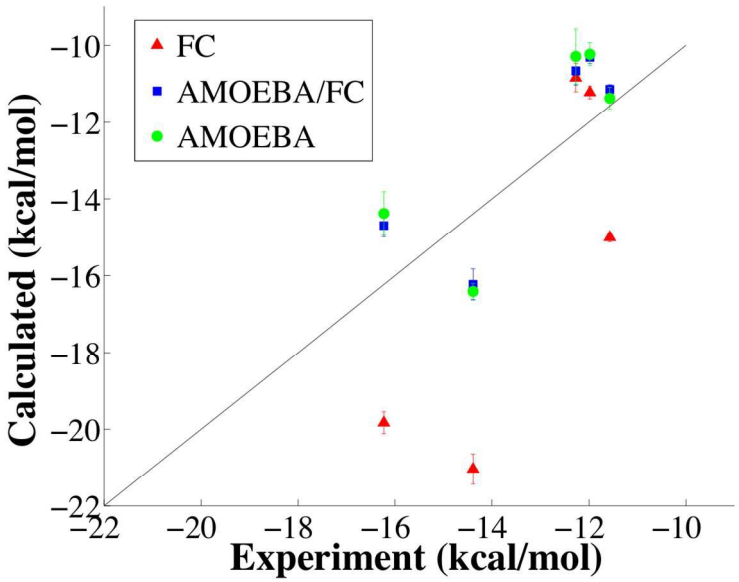

Figure 6. Absolute crystal deposition free energy values from FC (OPLS-AA), AMOEBA/FC dual force field approach and direct AMOEBA calculations are compared to experiment.

References

1. Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P., and Williams, D. E. (2000) A test of crystal structure prediction of small organic molecules, *Acta Crystallographica Section B 56*, 697-714.

2. Price, S. L. (2004) The computational prediction of pharmaceutical crystal structures and polymorphism, *Adv. Drug Deliver. Rev. 56*, 301-319.

3. Furukawa, H., Cordova, K. E., O'Keeffe, M., and Yaghi, O. M. (2013) The Chemistry and Applications of Metal-Organic Frameworks, *Science 341*.

4. Desiraju, G. R. (2013) Crystal Engineering: From Molecule to Crystal, *J. Am. Chem. Soc. 135*, 9952-9967.

5. Price, S. L. (2008) From crystal structure prediction to polymorph prediction: interpreting the crystal energy landscape, *Phys. Chem. Chem. Phys. 10*, 1996-2009.

6. Bardwell, D. A., Adjiman, C. S., Arnautova, Y. A., Bartashevich, E., Boerrigter, S. X. M., Braun, D. E., Cruz-Cabeza, A. J., Day, G. M., Della Valle, R. G., Desiraju, G. R., van Eijck, B. P., Facelli, J. C., Ferraro, M. B., Grillo, D., Habgood, M., Hofmann, D. W. M., Hofmann, F., Jose, K. V. J., Karamertzanis, P. G., Kazantsev, A. V., Kendrick, J., Kuleshova, L. N., Leusen, F. J. J., Maleev, A. V., Misquitta, A. J., Mohamed, S., Needs, R. J., Neumann, M. A., Nikylov, D., Orendt, A. M., Pal, R., Pantelides, C. C., Pickard, C. J., Price, L. S., Price, S. L., Scheraga, H. A., van de Streek, J., Thakur, T. S., Tiwari, S., Venuti, E., and Zhitkov, I. K. (2011) Towards crystal structure prediction of complex organic compounds - a report on the fifth blind test, *Acta Crystallogr. B 67*, 535-551.

7. Price, S. L. (2014) Predicting crystal structures of organic compounds, *Chem. Soc. Rev. 43*, 2098-2111.

8. Yang, J., Hu, W., Usvyat, D., Matthews, D., Schütz, M., and Chan, G. K.-L. (2014) Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy, *Science 345*, 640-643.

9. Thompson, H. P. G., and Day, G. M. (2014) Which conformations make stable crystal structures? Mapping crystalline molecular geometries to the conformational energy landscape, *Chemical Science 5*, 3173-3182.

10. Hylton, R. K., Tizzard, G. J., Threlfall, T. L., Ellis, A. L., Coles, S. J., Seaton, C. C., Schulze, E., Lorenz, H., Seidel-Morgenstern, A., Stein, M., and Price, S. L. (2015) Are the Crystal Structures of Enantiopure and Racemic Mandelic Acids Determined by Kinetics or Thermodynamics?, *J. Am. Chem. Soc. 137*, 11095-11104.

11. Heit, Y. N., Nanda, K. D., and Beran, G. J. O. (2016) Predicting finite-temperature properties of crystalline carbon dioxide from first principles with quantitative accuracy, *Chemical Science 7*, 246-255.

12. Jorgensen, W. L., and Duffy, E. M. (2002) Prediction of drug solubility from structure, *Adv. Drug Deliver. Rev. 54*, 355-366.

13. Chu, K. A., and Yalkowsky, S. H. (2009) Predicting Aqueous Solubility: The Role of Crystallinity, *Current Drug Metabolism 10*, 1184-1191.

14. Schnieders, M. J., Baltrusaitis, J., Shi, Y., Chattree, G., Zheng, L., Yang, W., and Ren, P. (2012) The structure, thermodynamics, and solubility of organic crystals from simulation with a polarizable force field, *J. Chem. Theory Comput. 8*, 1721-1736.

15. Park, J., Nessler, I., McClain, B., Macikenas, D., Baltrusaitis, J., and Schnieders, M. J. (2014) Absolute Organic Crystal Thermodynamics: Growth of the Asymmetric Unit into a Crystal via Alchemy, *J. Chem. Theory Comput. 10*, 2781-2791.

16. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1996) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc. 118*, 2309.

17. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586-3616.

18. Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, *J. Am. Chem. Soc. 118*, 11225-11236.

19. Price, S. L. (1996) Applications of realistic electrostatic modeling to molecules in complexes, solids and proteins, *Journal of the Chemical Society, Faraday Transactions 92*, 2997-3008.

20. Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P., and Williams, D. E. (2000) A test of crystal structure prediction of small organic molecules, *Acta Crystallogr. B 56*, 697-714.

21. Stone, A. J., and Alderton, M. (1985) Distributed multipole analysis - methods and applications, *Mol. Phys. 56*, 1047-1064.

22. Stone, A. J. (2005) Distributed multipole analysis: Stability for large basis sets, *J. Chem. Theory Comput. 1*, 1128-1132.

23. Ren, P., and Ponder, J. W. (2002) Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations, *J. Comput. Chem. 23*, 1497-1506.

24. Ren, P., Wu, C., and Ponder, J. W. (2011) Polarizable atomic multipole-based molecular mechanics for organic molecules, *J. Chem. Theory Comput. 7*, 3143-3161.

25. Ren, P., and Ponder, J. W. (2003) Polarizable atomic multipole water model for molecular mechanics simulation, *J. Phys. Chem. B 107*, 5933-5947.

26. Ponder, J. W., Wu, C., Ren, P., Pande, V. S., Chodera, J. D., Schnieders, M. J., Haque, I., Mobley, D. L., Lambrecht, D. S., DiStasio, R. A., Head-Gordon, M., Clark, G. N. I., Johnson, M. E., and Head-Gordon, T. (2010) Current status of the AMOEBA polarizable force field, *The Journal of Physical Chemistry B 114*, 2549-2564.

27. Piquemal, J. P., Cisneros, G. A., Reinhardt, P., Gresh, N., and Darden, T. A. (2006) Towards a force field based on density fitting, *J. Chem. Phys. 124*, 104101.

28. Elking, D. M., Cisneros, G. A., Piquemal, J.-P., Darden, T. A., and Pedersen, L. G. (2009) Gaussian Multipole Model (GMM), *J. Chem. Theory Comput. 6*, 190-202.

29. Gresh, N. (2006) Development, validation, and applications of anisotropic polarizable molecular mechanics to study ligand and drug-receptor interactions, *Curr. Pharm. Des. 12*, 2121-2158.

30. Gresh, N., Cisneros, G. A., Darden, T. A., and Piquemal, J. P. (2007) Anisotropic, polarizable molecular mechanics studies of inter- and intramoecular interactions and ligand-macromolecule complexes. A bottom-up strategy, *J. Chem. Theory Comput. 3*, 1960-1986.

31. Holt, A., Bostrom, J., Karlstrom, G., and Lindh, R. (2010) A NEMO Potential that Includes the Dipole-Quadrupole and Quadrupole-Quadrupole Polarizability, *J. Comput. Chem. 31*, 1583-1591.

32. Ponder, J. W., and Case, D. A. (2003) Force fields for protein simulations, In *Adv. Protein Chem.*, pp 27-85, Academic Press.

33. Ren, P., Chun, J., Thomas, D. G., Schnieders, M. J., Marucho, M., Zhang, J., and Baker, N. A. (2012) Biomolecular electrostatics and solvation: a computational perspective, *Q. Rev. Biophys. 45*, 427-491.

34. Shi, Y., Ren, P., Schnieders, M., and Piquemal, J.-P. (2015) Polarizable Force Fields for Biomolecular Modeling, In *Reviews in Computational Chemistry Volume 28*, pp 51-86, John Wiley & Sons, Inc.

35. Ren, P., and Ponder, J. W. (2004) Temperature and pressure dependence of the AMOEBA water model, *J. Phys. Chem. B 108*, 13427-13437.

36. Grossfield, A., Ren, P. Y., and Ponder, J. W. (2003) Ion solvation thermodynamics from simulation with a polarizable force field, *J. Am. Chem. Soc. 125*, 15671-15682.

37. Shi, Y., Wu, C., Ponder, J. W., and Ren, P. (2011) Multipole electrostatics in hydration free energy calculations, *J. Comput. Chem. 32*, 967-977.

38. Wu, J. C., Chattree, G., and Ren, P. (2012) Automation of AMOEBA polarizable force field parameterization for small molecules, *Theor. Chem. Acc. 131*, 1-11.

39. Shi, Y., Xia, Z., Zhang, J., Best, R., Wu, C., Ponder, J. W., and Ren, P. (2013) Polarizable atomic multipole-based AMOEBA force field for proteins, *J. Chem. Theory Comput. 9*, 4046-4063.

40. Gao, J., and Xia, X. (1992) A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations, *Science 258*, 631-635.

41. Luzhkov, V., and 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*, 199-213.

42. Wesolowski, T., and Warshel, A. (1994) Ab Initio Free Energy Perturbation Calculations of Solvation Free Energy Using the Frozen Density Functional Approach, *The Journal of Physical Chemistry 98*, 5183-5187.

43. Stanton, R. V., Little, L. R., and Merz, K. M., Jr. (1995) Quantum Free Energy Perturbation Study within a PM3/MM Coupled Potential, *The Journal of Physical Chemistry 99*, 483-486.

44. Li, G., Zhang, X., and Cui, Q. (2003) Free Energy Perturbation Calculations with Combined QM/MM Potentials Complications, Simplifications, and Applications to Redox Potential Calculations, *The Journal of Physical Chemistry B 107*, 8643-8653.

45. Li, G., and Cui, Q. (2003) pKa Calculations with QM/MM Free Energy Perturbations, *The Journal of Physical Chemistry B 107*, 14521-14528.

46. Olsson, M. H. M., Hong, G., and Warshel, A. (2003) Frozen Density Functional Free Energy Simulations of Redox Proteins: Computational Studies of the Reduction Potential of Plastocyanin and Rusticyanin, *J. Am. Chem. Soc. 125*, 5025-5039.

47. Hu, H., and Yang, W. (2005) Dual-topology/dual-coordinate free-energy simulation using QM/MM force field, *The Journal of Chemical Physics 123*, 041102.

48. Riccardi, D., Schaefer, P., and Cui, Q. (2005) pKa Calculations in Solution and Proteins with QM/MM Free Energy Perturbation Simulations: A Quantitative Test of QM/MM Protocols, *The Journal of Physical Chemistry B 109*, 17715-17733.

49. Woods, C. J., Manby, F. R., and Mulholland, A. J. (2008) An efficient method for the calculation of quantum mechanics/molecular mechanics free energies, *The Journal of Chemical Physics 128*, 014109.

50. Li, H., and Yang, W. (2007) Sampling enhancement for the quantum mechanical potential based molecular dynamics simulations: A general algorithm and its extension for free energy calculation on rugged energy surface, *The Journal of Chemical Physics 126*, 114104.

51. Min, D., Zheng, L., Harris, W., Chen, M., Lv, C., and Yang, W. (2010) Practically Efficient QM/MM Alchemical Free Energy Simulations: The Orthogonal Space Random Walk Strategy, *J. Chem. Theory Comput. 6*, 2253-2266.

52. Yang, W., Cui, Q., Min, D., and Li, H. (2010) Chapter 4 - QM/MM Alchemical Free Energy Simulations: Challenges and Recent Developments, In *Annu. Rep. Comput. Chem.* (Ralph, A. W., Ed.), pp 51-62, Elsevier.

53. Konig, G., Hudson, P. S., Boresch, S., and Woodcock, H. L. (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*, 1406-1419.

54. Konig, G., Pickard, F. C., Mei, Y., and Brooks, B. R. (2014) Predicting hydration free energies with a hybrid QM/MM approach: an evaluation of implicit and explicit solvation models in SAMPL4, *J. Comput. Aided Mol. Des. 28*, 245-257.

55. König, G., Mei, Y., Pickard, F. C., Simmonett, A. C., Miller, B. T., Herbert, J. M., Woodcock, H. L., Brooks, B. R., and Shao, Y. (2016) Computation of Hydration Free Energies Using the Multiple Environment Single System Quantum Mechanical/Molecular Mechanical Method, *J. Chem. Theory Comput. 12*, 332-344.

56. Dybeck, E. C., König, G., Brooks, B. R., and Shirts, M. R. (2016) Comparison of Methods To Reweight from Classical Molecular Simulations to QM/MM Potentials, *J. Chem. Theory Comput. 12*, 1466-1480.

57. Shirts, M. R., and Chodera, J. D. (2008) Statistically optimal analysis of samples from multiple equilibrium states, *J. Chem. Phys. 129*, 10.

58. Zwanzig, R. W. (1954) High-temperature equation of state by a perturbation method. I. Nonpolar gases, *The Journal of Chemical Physics 22*, 1420-1426.

59. Lyman, E., Ytreberg, F. M., and Zuckerman, D. M. (2006) Resolution exchange simulation, *Phys. Rev. Lett. 96*, 4.

60. Dybeck, E. C., Schieber, N. P., and Shirts, M. R. (2016) Effects of a More Accurate Polarizable Hamiltonian on Polymorph Free Energies Computed Efficiently by Reweighting Point-Charge Potentials, *J. Chem. Theory Comput.*

61. Akin-Ojo, O., Song, Y., and Wang, F. (2008) Developing ab initio quality force fields from condensed phase quantum-mechanics/molecular-mechanics calculations through the adaptive force matching method, *The Journal of Chemical Physics 129*, 064108.

62. Spiriti, J., and Zuckerman, D. M. (2014) Tunable Coarse Graining for Monte Carlo Simulations of Proteins via Smoothed Energy Tables: Direct and Exchange Simulations, *J. Chem. Theory Comput. 10*, 5161-5177.

63. Kollman, P. (1993) Free energy calculations: Applications to chemical and biochemical phenomena, *Chem. Rev. 93*, 2395-2417.

64. Jorgensen, W. L., and Tirado-Rives, J. (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, *J. Am. Chem. Soc. 110*, 1657-1666.

65. Schnieders, M. J., Fenn, T. D., and Pande, V. S. (2011) Polarizable atomic multipole X-ray refinement: Particle mesh Ewald electrostatics for macromolecular crystals, *J. Chem. Theory Comput. 7*, 1141-1156.

66. Fenn, T. D., and Schnieders, M. J. (2011) Polarizable atomic multipole X-ray refinement: Weighting schemes for macromolecular diffraction, *Acta Crystallogr. D 67*, 957-965.

67. LuCore, Stephen D., Litman, Jacob M., Powers, Kyle T., Gao, S., Lynn, Ava M., Tollefson, William T. A., Fenn, Timothy D., Washington, M. T., and Schnieders, Michael J. (2015) Dead-End Elimination with a Polarizable Force Field Repacks PCNA Structures, *Biophys. J. 109*, 816-826.

68. Brown, C. (1966) Further refinement of the crystal structure of acetanilide, *Acta Crystallogr. 21*, 442-445.

69. Haisa, M., Kashino, S., Kawai, R., and Maeda, H. (1976) The Monoclinic Form of p-Hydroxyacetanilide, *Acta Crystallographica Section B 32*, 1283-1285.

70. Nath, N. K., Aggarwal, H., and Nangia, A. (2011) Crystal structure of methyl paraben polymorph II, *Cryst. Growth Des. 11*, 967-971.

71. Lin, X. (1986) The structure of ethyl paraben, *Chin. J. Struct. Chem. 5*, 281.

72. Patel, U., Patel, T. C., and Singh, T. P. (1983) Structure of phenacetin, $C_{10}H_{13}NO_2$, *Acta Crystallographica Section C 39*, 1445-1447.

73. Banks, J. L., Beard, H. S., Cao, Y., Cho, A. E., Damm, W., Farid, R., Felts, A. K., Halgren, T. A., Mainz, D. T., Maple, J. R., Murphy, R., Philipp, D. M., Repasky, M. P., Zhang, L.

Y., Berne, B. J., Friesner, R. A., Gallicchio, E., and Levy, R. M. (2005) Integrated Modeling Program, Applied Chemical Theory (IMPACT), *J. Comput. Chem. 26*, 1752-1780.

74. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle-mesh Ewald method, *J. Chem. Phys. 103*, 8577-8593.

75. Darden, T., York, D., and Pedersen, L. (1993) Particle-mesh Ewald - An n log(n) method for Ewald sums in large systems, *J. Chem. Phys. 98*, 10089-10092.

76. Sagui, C., Pedersen, L. G., and Darden, T. A. (2004) Towards an accurate representation of electrostatics in classical force fields: Efficient implementation of multipolar interactions in biomolecular simulations, *J. Chem. Phys. 120*, 73-87.

77. Zheng, L., Chen, M., and Yang, W. (2008) Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems, *Proc. Natl. Acad. Sci. U.S.A. 105*, 20227-20232.

78. Lee, S., Chen, M., Yang, W., and Richards, N. G. J. (2010) Sampling Long Time Scale Protein Motions: OSRW Simulation of Active Site Loop Conformational Free Energies in Formyl-CoA:Oxalate CoA Transferase, *J. Am. Chem. Soc. 132*, 7252-7253.

79. Ponder, J. W. (2015) TINKER: Software Tools for Molecular Design, 7.0 ed., p TINKER: Software Tools for Molecular Design, Saint Louis, MO.

80. Bardwell, D. A., Adjiman, C. S., Arnautova, Y. A., Bartashevich, E., Boerrigter, S. X. M., Braun, D. E., Cruz-Cabeza, A. J., Day, G. M., Della Valle, R. G., Desiraju, G. R., van Eijck, B. P., Facelli, J. C., Ferraro, M. B., Grillo, D., Habgood, M., Hofmann, D. W. M., Hofmann, F., Jose, K. V. J., Karamertzanis, P. G., Kazantsev, A. V., Kendrick, J., Kuleshova, L. N., Leusen, F. J. J., Maleev, A. V., Misquitta, A. J., Mohamed, S., Needs, R. J., Neumann, M. A., Nikylov, D., Orendt, A. M., Pal, R., Pantelides, C. C., Pickard, C. J., Price, L. S., Price, S. L., Scheraga, H. A., van de Streek, J., Thakur, T. S., Tiwari, S., Venuti, E., and Zhitkov, I. K. (2011) Towards crystal structure prediction of complex

organic compounds - a report on the fifth blind test, *Acta Crystallographica Section B-Structural Science Crystal Engineering and Materials 67*, 535-551.

81. Dama, J. F., Rotskoff, G., Parrinello, M., and Voth, G. A. (2014) Transition-Tempered Metadynamics: Robust, Convergent Metadynamics via On-the-Fly Transition Barrier Estimation, *J. Chem. Theory Comput. 10*, 3626-3633.

82. Mobley, D. L., and Guthrie, J. P. (2014) FreeSolv: a database of experimental and calculated hydration free energies, with input files, *J. Comput. Aided Mol. Des. 28*, 711-720.

83. Jiao, D., Golubkov, P. A., Darden, T. A., and Ren, P. (2008) Calculation of protein-ligand binding free energy by using a polarizable potential, *Proc. Natl. Acad. Sci. U.S.A. 105*, 6290-6295.

84. Dill, K. A., and MacCallum, J. L. (2012) The Protein-Folding Problem, 50 Years On, *Science 338*, 1042-1046.