JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Absolute Organic Crystal Thermodynamics: Growth of the Asymmetric Unit into a Crystal via Alchemy

Jooyeon Park,[†] Ian Nessler,[†] Brian McClain,[§] Dainius Macikenas,[§] Jonas Baltrusaitis,[∥] and Michael J. Schnieders*[,†,‡]

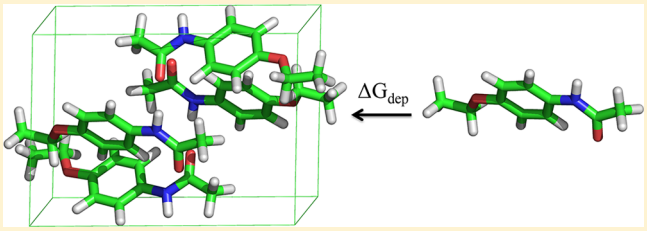
[†]Department of Biomedical Engineering, and [‡]Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242, United States

[§]Vertex Pharmaceuticals Incorporated, Cambridge Massachusetts 02139, United States

[∥]University of Twente, 7500 AE Enschede, The Netherlands

Supporting Information

ABSTRACT: The solubility of organic molecules is of critical importance to the pharmaceutical industry; however, robust computational methods to predict this quantity from first-principles are lacking. Solubility can be computed from a thermodynamic cycle that decomposes standard state solubility into the sum of solid−vapor sublimation and vapor−liquid solvation free energies $\Delta G^{\circ}_{\text{solubility}} = \Delta G^{\circ}_{\text{sub}} + \Delta G^{\circ}_{\text{solv}}$. Over the past few decades, alchemical simulation methods to compute solvation free energy using classical force fields have become widely used. However, analogous methods for determining the free energy of the sublimation/deposition phase transition are currently limited by the necessity of a priori knowledge of the atomic coordinates of the crystal. Here, we describe progress toward an alternative scheme based on growth of the asymmetric unit into a crystal via alchemy (GAUCHE). GAUCHE computes deposition free energy $\Delta G^{\circ}_{\text{dep}} = -\Delta G^{\circ}_{\text{sub}} = -k_B T \ln(V_c/V_g) + \Delta G_{\text{AU}} + \Delta G_{\text{AU}\rightarrow\text{UC}}$ as the sum of an entropic term to account for compressing a vapor at 1 M standard state (V_g) into the molar volume of the crystal (V_c), where k_B is Boltzmann's constant and T is temperature in degrees Kelvin, plus two simulation steps. In the first simulation step, the deposition free energy ΔG_{AU} for a system composed of only N_{AU} asymmetric unit (AU) molecule(s) is computed beginning from an arbitrary conformation in vacuum. In the second simulation step, the change in free energy $\Delta G_{\text{AU}\rightarrow\text{UC}}$ to expand the asymmetric unit degrees of freedom into a unit cell (UC) composed of N_{UC} independent molecules is computed. This latter step accounts for the favorable free energy of removing the constraint that every symmetry mate of the asymmetric unit has an identical conformation and intermolecular interactions. The current work is based on NVT simulations, which requires knowledge of the crystal space group and unit cell parameters from experiment, but not a priori knowledge of crystalline atomic coordinates. GAUCHE was applied to 5 organic molecules whose sublimation free energy has been measured experimentally, based on the polarizable AMOEBA force field and more than a microsecond of sampling per compound in the program Force Field X. The mean unsigned and RMS errors were only 1.6 and 1.7 kcal/mol, respectively, which indicates that GAUCHE is capable of accurate prediction of absolute sublimation thermodynamics.



INTRODUCTION

The bioavailability and pharmacokinetics of a drug depend in part on its formulation, which often must overcome the low solubility of pharmaceutical compounds. Therefore, accurate computational methods to predict solubility thermodynamics could aid drug development and promote optimal formulations. To this end, the solubility equilibrium can be decomposed using a thermodynamic cycle that defines standard state solubility via the sum of solid−vapor sublimation $\Delta G^{\circ}_{\text{sub}}$ and vapor−liquid solvation free energy $\Delta G^{\circ}_{\text{solv}}$ terms[1]

$$\Delta G^{\circ}_{\text{solubility}} = \Delta G^{\circ}_{\text{sub}} + \Delta G^{\circ}_{\text{solv}} \quad (1)$$

Alchemical simulation methods to predict $\Delta G^{\circ}_{\text{solv}}$ described over the last few decades[2] can now be performed efficiently[3] and with high statistical precision,[4] although work to overcome force field limitations is ongoing.[5]

On the other hand, current methods to predict $\Delta G^{\circ}_{\text{sub}}$ rely on a priori knowledge of the crystalline atomic coordinates,[6] on direct experimental measurements of crystalline properties such as fugacity,[7] or on various heuristic approaches.[8] Although crystal structure prediction (CSP) has made significant progress over the past decade, as indicated by the successes and increasing difficulty of the first blind CSP test[9] to the fifth,[10] a unifying limitation of prediction methods has been their reliance on ranking structural candidates by potential energy, rather than on free energy. This striking approximation, the neglect of entropic differences between putative polymorphs,

Special Issue: Free Energy Calculations: Three Decades of Adventure in Chemistry and Biophysics

Received: February 28, 2014

   dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

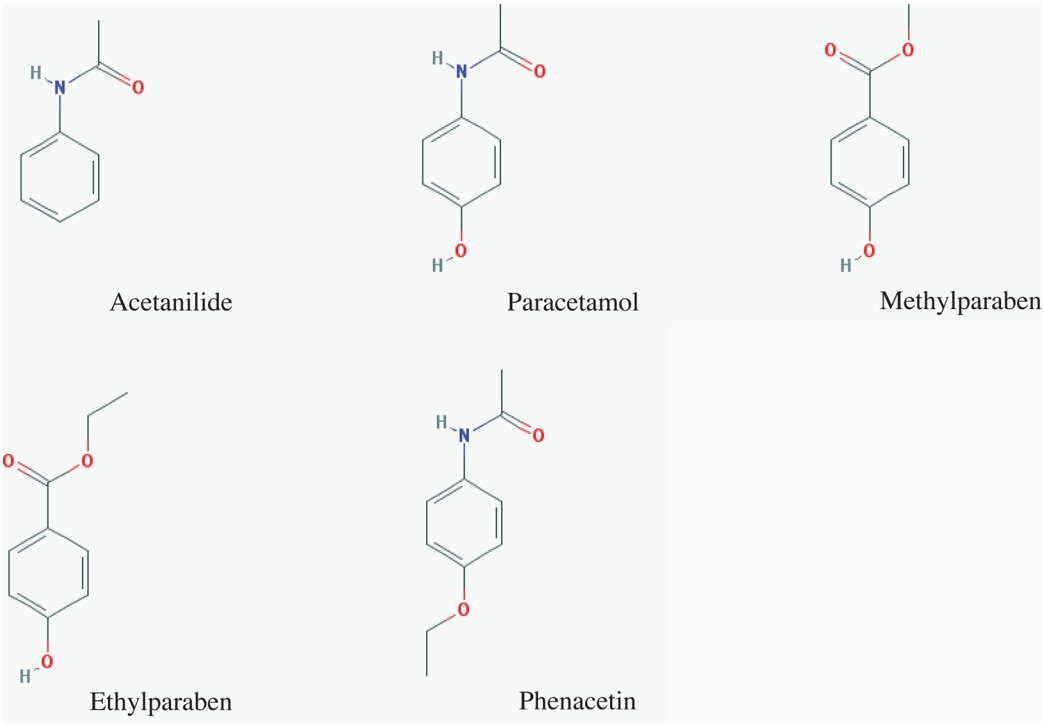


Figure 1. Chemical structures for the five molecules studied: acetanilide, paracetamol, methyl paraben, ethyl paraben, and phenacetin.

may limit CSP accuracy to the same degree as limitations in potential energy functions. For example, there is experimental evidence that the polymorphs of paracetamol differ in their entropic contribution to sublimation free energy at 298 K by 0.5 kcal/mol.[11] As the size and complexity of organic molecules increase toward large pharmaceutical compounds, peptides,[12] or even proteins, neglect of entropy between polymorphs becomes unreasonable.

To address this challenge, we propose a simulation procedure to compute absolute organic crystal deposition free energy based on growth of the asymmetric unit into a crystal via alchemy (GAUCHE). Beginning from the asymmetric unit (AU) in vapor offers the appealing advantages of (1) inexpensive force evaluations on N_{AU} molecules relative to a simulation system composed of one or more copies of the N_{UC} unit cell (UC) molecules, (2) reduced degrees of freedom during the alchemical phase transition to accelerate sampling convergence, and (3) avoidance of the need for a priori knowledge of crystalline atomic coordinates. GAUCHE decomposes standard state deposition free energy ΔG°_{dep} into a sum of two simulation terms and an entropic correction ΔG_{Vol} to move from the volume of a 1 M standard state vapor (V_g) to the molar volume of the crystal (V_c)

$$\Delta G^{\circ}_{dep} = -\Delta G^{\circ}_{sub} = \Delta G_{Vol} + \Delta G_{AU} + \Delta G_{AU\rightarrow UC} \quad (2)$$

where

$$\Delta G_{Vol} = -k_B T \ln(V_c/V_g) \quad (3)$$

and $k_B T$ is Boltzmann's constant multiplied by temperature in degrees Kelvin. First, the deposition free energy for a system composed of only N_{AU} asymmetric unit molecule(s) ΔG_{AU} is computed, beginning from an arbitrary conformation in vacuum.[13] Although the current work is based on NVT simulations and requires knowledge of the crystal space group and unit cell parameters, no a priori knowledge of crystalline atomic coordinates is needed. Second, GAUCHE computes the change in free energy $\Delta G_{AU\rightarrow UC}$ to expand the asymmetric unit into a larger system composed of all N_{UC} molecules in the unit cell. This latter step accounts for the favorable free energy of removing the constraint that at every dynamics step, each symmetry mate of the asymmetric unit adopts an identical conformation and has identical intermolecular interactions. For example, a unit cell end state permits favorable coupled motions between asymmetric units.

GAUCHE builds on our previous work to compute of the solubility of the *n*-alkylamide series.[13] Our approach requires explicit support for space group symmetry operators in the implementation of the underlying potential energy function,[13] which in this work is the polarizable AMOEBA force field[14] implemented in the Force Field X (FFX) program.[13,15] It also depends on enhanced sampling techniques to overcome large free energy barriers between alternative crystalline conformations, which is addressed using the Orthogonal Space Random Walk (OSRW)[16] variant of metadynamics.[17] Here, we focus on five compounds that include a wider range of organic functional groups than in our previous work on *n*-alkylamides.[13] For each compound, the absolute sublimation free energy has been experimentally measured.[11,18] Finally, we compare our rigorous alchemical approach to approximate, but much faster, lattice potential energy and end-state estimates of crystal stability.

METHODS

Organic Compounds and AMOEBA Parameterization.

Five molecules were selected from approximately 30 compounds with experimentally measured sublimation free energy based on the following criteria: (1) molecular weight larger than 130 g/mol, (2) all atoms from the set [H, C, N, O], and (3) the absence of bicyclic or polycyclic ring systems. The compounds include acetanilide,[19] paracetamol (polymorph I),[20] methyl paraben (polymorph II),[21] ethyl paraben,[22] and

dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Compounds Studied and Their Associated CSD Reference Codes, Space Groups, and Unit Cell Parameters[a]

compd.	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	*P*21/*c*	11.72	9.40	12.93	90	147.0	90
methyl paraben II	CEBGOF03	*P*21/*c*	4.82	14.63	10.24	90	99.8	90
ethyl paraben	FEGLEI	*P*21/*c*	13.76	13.18	11.58	90	125.5	90
penacetin	PYRAZB10	*P*21/*c*	13.25	9.65	7.81	90	104.9	90

[a]Roman numerals following paracetamol and methyl paraben correspond to polymorph.

Table 2. Molecular Weight, Number of Molecules Per Asymmetric Unit, Unit Cell Volume, Number of Unit Cell Molecules, Volume Per Molecule, and Experimental Temperature for Each Crystal Studied

compd.	mol. wt (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

phenacetin[23] and are shown in Figure 1. Some functional groups, such as halogens[24] and polycyclic rings, are a current focus of AMOEBA development efforts and will be examined in future work on organic crystal thermodynamics. Crystallographic space group and unit cell information was obtained from the Cambridge Structural Database (CSD) and is provided in Tables 1 and 2. AMOEBA force field parameters were prepared for the five molecules using the PolType program,[14c] which in turn relies on Gaussian,[25] GDMA,[26] and TINKER.[27]

Crystal Lattice Potential Energy. Beginning from the experimental coordinates, each molecule was optimized in FFX[13,15] to compare with dispersion corrected density functional theory (DFT-D) lattice energies. During minimization, space group symmetry was enforced. The van der Waals cutoff was set to 12.0 Å and smoothly tapered to zero interaction energy using a multiplicative switch beginning at 10.8 Å. Polarizable AMOEBA electrostatics were evaluated using a version of the smooth[28] particle-mesh Ewald[29] (PME) algorithm for multipoles[30] that explicitly supports space group symmetry[15a] and the self-consistent field calculation was converged to 10^{-5} RMS Debye. High precision PME parameters included a real-space cutoff of 9.0 Å, a mesh density of 2.0 grid points per Å, eighth order B-splines and an Ewald parameter of 0.45. This allowed minimization to reach an RMS gradient convergence criterion of 10^{-4} (kcal·mol^{-1}·Å^{-1}). A second minimization was performed in vacuum, beginning from the crystal minimum, to provide a reference energy E_{vac}, which is subtracted from the potential energy per molecule at the crystal minimum E_{cryst} to give the lattice potential energy

$$E_{lattice} = E_{cryst} - E_{vac} \tag{4}$$

Periodic solid-state ab initio calculations were performed in the program suite CRYSTAL'09,[31] which uses functions localized at atoms as the basis for expansion of the crystalline orbitals via a linear combination of atomic orbitals (LCAO) technique. All-electron Gaussian type basis sets, the hybrid B3LYP[32] Hamiltonian and a dispersion correction were used.[33] The DFT exchange−correlation contribution is evaluated by numerical integration over the unit cell volume. Radial and angular points of the grid were generated through Gauss−Legendre radial quadrature and Lebedev two-dimensional angular point distributions with a pruned grid of 75 radial and 974 angular points. The level of accuracy in evaluating the Coulomb and Hartree−Fock exchange series was controlled by five parameters[31b] and values of 7, 7, 7, 7, and 16 were used. The reciprocal space integration was performed by sampling the Brillouin zone with the 6 × 6 × 6 Pack−Monkhorst net.[34] Structure optimizations were performed using analytical energy gradients with respect to atomic coordinates with cell parameters fixed,[35] within a quasi-Newton approach based on the Broyden−Fletcher−Goldfarb−Shanno scheme for Hessian updating.[36] Convergence was checked on both gradient components and nuclear displacements and was signaled when the RMS gradient was 0.00015 hartree/Bohr and RMS displacement was 0.0006 Bohr. The 6-31G*[37] basis set was used, and condensed phase energies were corrected for BSSE via the counterpoise method.[38]

End-State Approximation to Crystal Stability. Using the optimized vapor phase structure, normal modes were computed in TINKER.[27] The normal modes are converted to a vibrational free energy using the Einstein model of independent quantum harmonic oscillators

$$G_{vib} = \sum_{3N-6}^{i=1} \left(\frac{h\nu_i}{2}\right)\coth\left(\frac{h\nu_i}{2k_BT}\right) - \sum_{3N-6}^{i=1}\left[\left(\frac{h\nu_i}{2}\right)\coth\left(\frac{h\nu_i}{2k_BT}\right) - 2k_BT\ln\left(\coth\left(\frac{h\nu_i}{2k_BT}\right)\right)\right] \tag{5}$$

where ν_i is the ith vibrational frequency, h is Planck's constant, N is the number of atoms in the molecule and $3N-6$ is the total number of vibrational modes. For the crystalline phase of the molecule, the unit cell parameters as well as the conformation of the molecules were optimized. During crystal minimization in TINKER, symmetry constraints were applied to the unit cell parameters but not the molecules. A convergence criterion of 0.5 kcal·mol^{-1}·Å^{-1} was used. The same Einstein model was used for deriving the crystal vibrational free energy using normal modes calculated from the entire unit cell instead of a single molecule. Therefore, the number of molecules in the unit cell was used to normalize the crystalline vibrational

dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

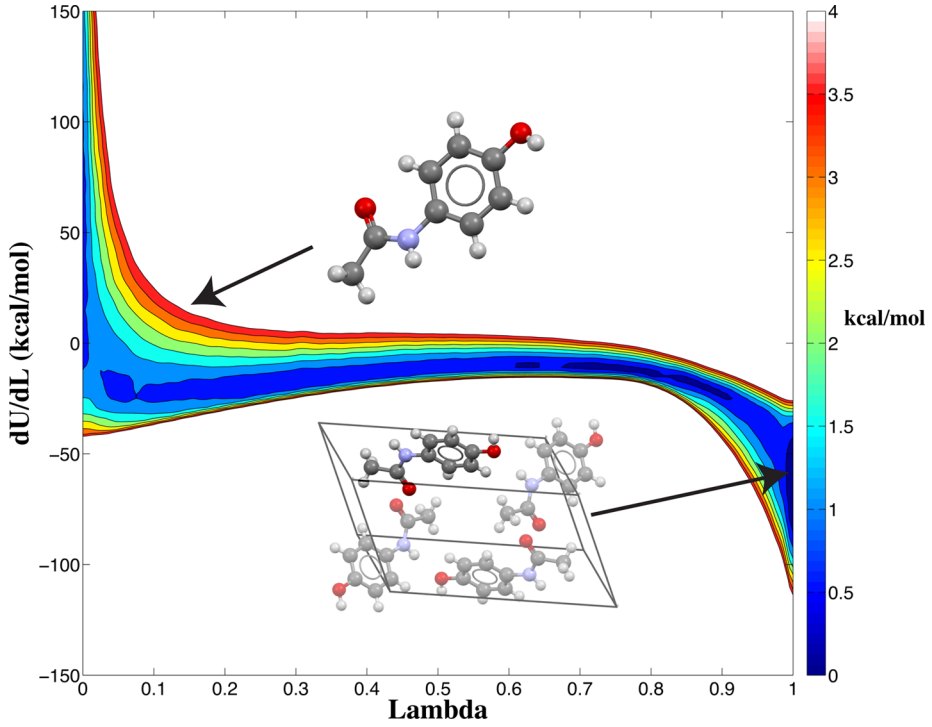


Figure 2. Shown is a potential of mean force for the alchemical path between vapor and crystalline states for paracetamol based on the OSRW method. Enhanced sampling with OSRW speeds convergence of numerical thermodynamic integration that is used to compute ΔG_{AU}.

contribution. Lattice energy and the change in vibrational free energy were combined with translational and rotational terms[39] to yield an end-state estimate of deposition free energy ΔG°_{ES} as

$$\Delta G^{\circ}_{ES} = \Delta G_{trans} + \Delta G_{rot} + \Delta G_{vib} + E_{lattice} \quad (6)$$

Growth of the Asymmetric Unit into a Crystal via Alchemy (GAUCHE). The first step of GAUCHE is simulation of the sublimation-deposition phase transition based on a system consisting of only the molecule(s) in the asymmetric unit, which smoothly samples an alchemical path between vapor and crystalline phases to yield ΔG_{AU}. This alchemical thermodynamic path provides a smooth transition between vacuum and crystalline states by slowly turning on intermolecular interactions with symmetry mates. The overall behavior of the alchemical transformation for paracetamol under the OSRW sampling methodology can be observed in Figure 2, which shows a two-dimensional potential of mean force that is a function of λ and $\partial U/\partial\lambda$. OSRW applies a time-dependent bias that flattens the surface along λ and $\partial U/\partial\lambda$ until a random walk is achieved, thereby accelerating convergence of thermodynamic integration. In vapor (λ = 0), many isoenergetic rigid body coordinates would result in high-energy steric clashes in the crystalline phase (λ = 1), which explains the unfavorable (large positive) values of $\partial U/\partial\lambda$ sampled in the vapor region. Once the alchemical simulation reached values of λ near 1, only conformations consistent with reasonable crystal packing are sampled, which have favorable (negative) values of $\partial U/\partial\lambda$. We note that each simulation sampled the AMOEBA crystalline free energy minimum multiple times.

Constraining a crystal to have perfect symmetry within the unit cell eliminates the possibility of coupled motions between symmetry mates, which is a favorable free energy contribution. The second simulation step of GAUCHE uses a thermodynamic cycle (Figure 3) to compute the change in free energy $\Delta G_{AU\rightarrow UC}$ along a path between the asymmetric unit simulation system and one that includes the entire unit cell. The first step of this thermodynamic cycle computes the free energy $\Delta G_{AU\rightarrow RAU}$ to add a harmonic restraint to each atom i in the asymmetric unit between its current simulation coordinates $\mathbf{r}_i$ and those of the energy minimized crystal $\mathbf{r}_{min,i}$

$$E_{restraint} = \sum_{N_{atoms}}^{i=1} \frac{k}{2} |\mathbf{r}_i - \mathbf{r}_{min,i}|^2 \quad (7)$$

with k chosen as 20 kcal·mol^{-1}·Å^{-1}. The second step of the thermodynamic cycle computes the free energy change $\Delta G_{RAU\rightarrow Vac}$ to transfer the restrained asymmetric unit (RAU) molecules into vapor by an alchemical annihilation method that turns off both inter- and intramolecular nonbonded interactions. The third step computes the free energy change $\Delta G_{Vac\rightarrow RUC}$ to move from a restrained unit cell in vapor into the crystalline phase by alchemically restoring all nonbonded interactions. The final step of the thermodynamic cycle computes the free energy to remove the harmonic restraints from all atoms in the unit cell $\Delta G_{RUC\rightarrow UC}$, where the UC restraints are analogous to the AU restraints defined in eq 7. For the final two steps, the simulation system is composed of N_{symm} copies of the asymmetric unit, where N_{symm} is the number of symmetry operators of the space group. Therefore, the calculated free energy of these steps is normalized by N_{symm} to achieve the energy per asymmetric unit. Overall, the free energy change in moving from an asymmetric unit simulation to a unit cell simulation is given by

$$\Delta G_{AU\rightarrow UC} = \Delta G_{AU\rightarrow RAU} + \Delta G_{RAU\rightarrow Vac} + \Delta G_{Vac\rightarrow RUC} + \Delta G_{RUC\rightarrow UC} \quad (8)$$

Instead of choosing the simulation system to be a single unit cell, a system with arbitrarily many copies of the unit cell would

dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

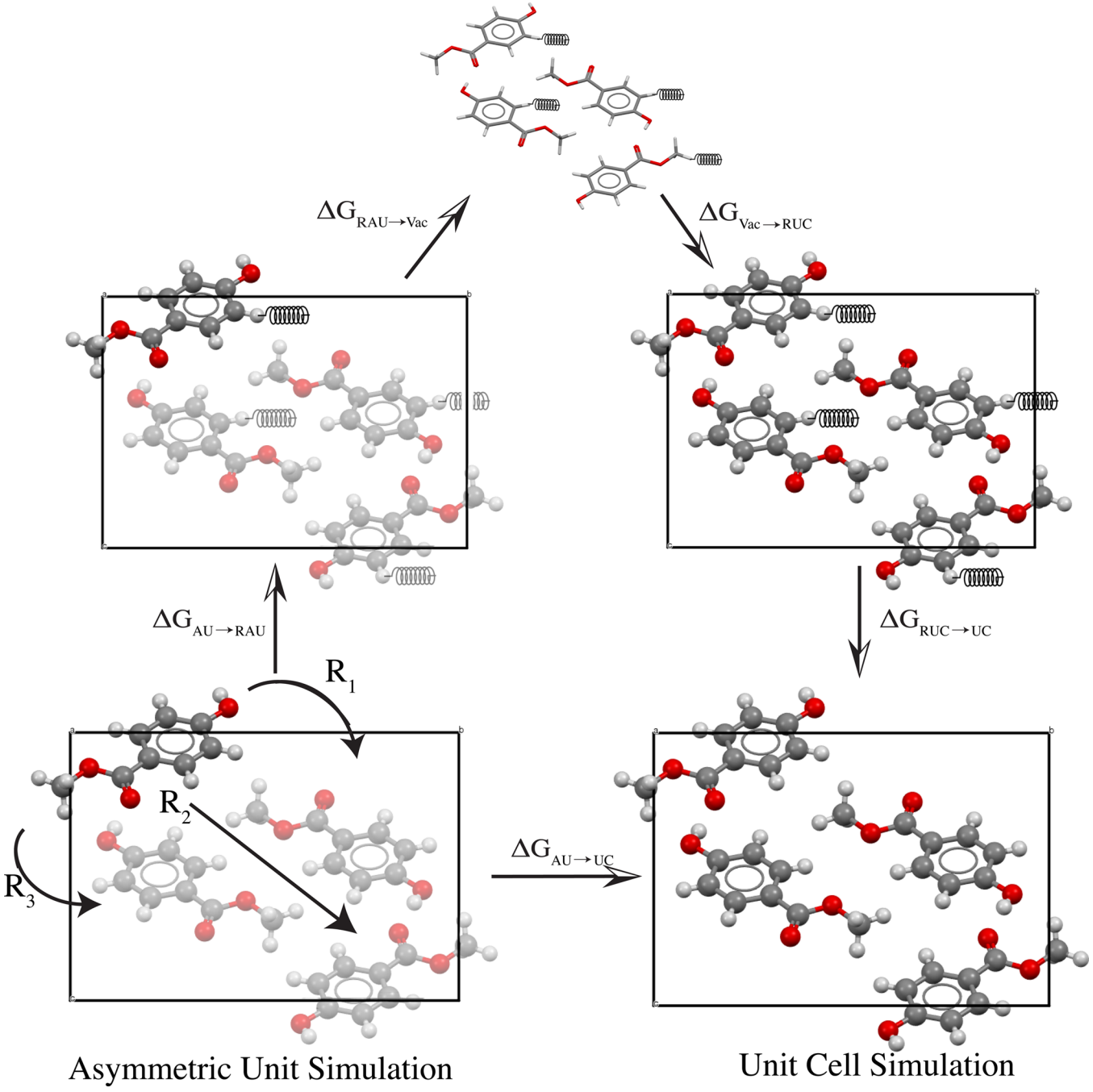


Figure 3. Shown is a thermodynamic cycle for computing the free energy change $\Delta G_{AU \rightarrow UC}$ between simulation systems composed of (1) an asymmetric unit and (2) a unit cell. In this case, the $P2_1/c$ methyl paraben II asymmetric unit is expanded into a P1 unit cell.

allow for coupled motions between unit cells. However, this is not explored in the current work.

For all steps, the free energy change was determined as the mean of five independent canonical ensemble (NVT) simulations per compound.[13] For each individual trajectory, stochastic dynamics (SD) was performed using a 1.0 fs time step. Asymmetric unit deposition simulations ΔG_{AU} were 200 ns in length, except for ethyl paraben, which was simulated for 300 ns, yielding at least a microsecond of sampling per compound in aggregate. The components of the $\Delta G_{AU \rightarrow UC}$ thermodynamic cycle converged after 10−50 ns. For all simulations, the volume was held constant during the simulations based on fixing the unit cell parameters at their experimental values. SD provided temperature fluctuations around a target of 298.15 K. Each ΔG_{AU} alchemical simulation began at the experimental coordinates of the molecule in the crystalline state ($\lambda = 1$). We note that equivalent results can be obtained by starting each simulation from an arbitrary conformation in vacuum ($\lambda = 0$), although sampling convergence is superior in the former case.

Multiple Molecules in the Asymmetric Unit. Our approach also handles the case of multiple molecules within an asymmetric unit, which is necessary for salts, cocrystals, and cases of noncrystallographic symmetry (NCS). This is accomplished by smoothly turning off all intermolecular interactions between molecules in the asymmetric unit as

dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

they transition from the crystalline state ($\lambda = 1$) into the vapor state ($\lambda = 0$) during alchemical simulations of ΔG_{AU}. The effect in vapor is that the asymmetric unit molecules can pass through each other, while softcore repulsion gently removes molecular overlaps for $\lambda > 0$. For example, because there are two copies of ethyl paraben in the asymmetric unit due to NCS, the computed values of ΔG_{AU} and $\Delta G_{AU \to UC}$ must be normalized by a factor of 2.

RESULTS

Lattice Energies for AMOEBA and Dispersion Corrected DFT. Currently, ab initio dynamics are much too expensive to reach the microsecond time scale required for convergence of GAUCHE, as described here. On the other hand, our implementation of the polarizable AMOEBA force field, which models permanent electrostatics on each atom using an ideal point multipole through quadrupole order and an ideal point induced dipole,[40] achieves 200 ns of sampling on a single compute node in about a week. Here, we compare AMOEBA lattice potential energies to dispersion corrected DFT. Ideally comparisons would be based on MP2 electron correlation and the basis sets typically used to parametrize AMOEBA in vacuum;[14c] however, these methods are not yet well established for periodic systems of large organic molecules. Shown in Table 3 are lattice energies for the five compounds studied here based on AMOEBA and dispersion corrected B3LYP/6-31G* with and without zero point energy (ZPE). AMOEBA lattice energies were systematically less favorable by 1.3 kcal/mol than dispersion corrected B3LYP/6-31G* when the ZPE is not considered. After inclusion of ZPE changes between vacuum and condensed phase, AMOEBA lattice energies become systematically too favorable by 2.2 kcal/mol. This behavior is reasonable for a polarizable force field such as AMOEBA, whose empirical repulsion-dispersion energy term was fit to reproduce experimental liquid densities and heats of vaporization and thereby implicitly includes ZPE changes between vapor and condensed phase to some degree.[5a,14a,41] On the other hand, substantial relative differences in ZPE energy between polymorphs[42] may be a limiting factor to the accuracy of classical multipolar force fields[43] for some crystal structure prediction applications.

Table 3. Comparison of the Lattice Energy between AMOEBA and Dispersion Corrected DFT (D-DFT) Evaluated Using the B3LYP Functional and 6-31G* Basis Set (kcal/mol)[a]

compd.	AMOEBA	DFT-D	abs. ΔE	DFT-D + ZPE	abs. ΔE
acetanilide	−22.56	−24.20	1.64	−20.79	1.77
paracetamol I	−27.69	−31.34	3.65	−27.19	0.50
methyl paraben II	−22.05	−23.10	1.05	−19.56	2.49
ethyl paraben	−24.58	−21.44	3.14	−18.64	5.94
phenacetin	−25.12	−28.34	3.22	−24.92	0.20
mean	−24.40	−25.68	2.54	−22.22	2.18

[a]Energies were computed after minimization of the experimental coordinates in the respective potential.

Alchemical Stochastic Dynamics. The deposition free energy was computed from two sets of simulations. First, the free energy difference ΔG_{AU} between the vapor and crystalline states was computed from a simulation composed of only asymmetric unit molecule(s). The results for five independent simulations per crystal are given in Table 4. The largest statistical uncertainty is for ethyl paraben, which has the largest asymmetric unit as it contains two NCS copies of ethyl paraben. Overall, the average statistical uncertainty is only 0.42 kcal/mol, which is less than 5% of the average deposition free energy. Convergence of the five independent simulations for methyl paraben II and ethyl paraben can be observed in Figure 4 and Figure 5, respectively.

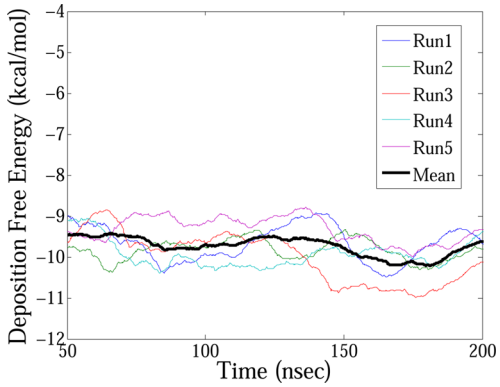


Figure 4. Shown is the convergence of five independent simulations of the methyl paraben II asymmetric unit deposition free energy (ΔG_{AU}).

The second set of simulations are combined to compute the favorable free energy $\Delta G_{AU \to UC}$ to expand the crystalline state from a simulation system composed of only the N_{AU} asymmetric unit molecule(s) into a simulation system composed of all N_{UC} molecules in the unit cell. The latter unit cell simulation system permits entropic degrees of freedom, such as coupled motions between symmetry mates, which are impossible in the former. The results for each simulation needed to complete the thermodynamic cycle given

Table 4. Free Energy Values ΔG_{AU} (kcal/mol) Computed from Five Independent 200 ns Simulations of the Asymmetric Unit along an Alchemical Path between Vapor and Crystalline States[a]

	simulation						
compd.	1	2	3	4	5	mean	σ
acetanilide	−12.32	−12.64	−12.25	−12.78	−12.09	−12.42	0.29
paracetamol I	−14.57	−14.64	−13.52	−13.53	−13.66	−13.98	0.57
methyl paraben II	−9.68	−9.83	−9.39	−10.12	−10.05	−9.81	0.30
ethyl paraben	−10.26	−10.39	−9.18	−10.65	−11.09	−10.31	0.71
phenacetin	−14.68	−14.14	−14.6	−14.46	−14.45	−14.47	0.21
mean							0.42

[a]The mean and standard deviation (σ) are given for each compound.

dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

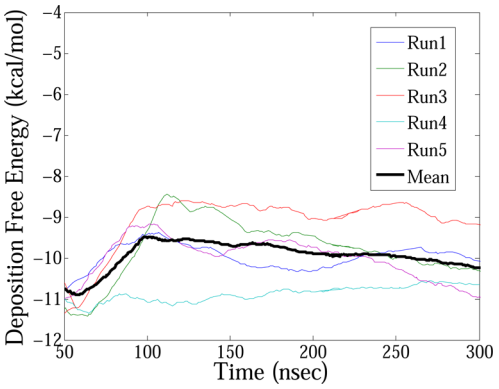


Figure 5. Shown is the convergence of five independent simulations of the ethyl paraben asymmetric unit deposition free energy (ΔG_{AU}).

in eq 8 and diagrammed in Figure 3 are given in Table 5. The first three contributions, including $\Delta G_{AU\rightarrow RAU}$, $\Delta G_{RAU\rightarrow Vac}$, and $\Delta G_{Vac\rightarrow RUC}$, converge in approximately 20 ns, which is an order of magnitude faster than the alchemical vapor−solid phase transition. Their statistical uncertainty is generally less than 0.25 kcal/mol. However, the free energy to remove symmetry restraints from the unit cell molecules $\Delta G_{RUC\rightarrow UC}$ proved to require sampling of at least 50 ns, although this is still considerably less than ΔG_{AU}.

Summing volume entropy ΔG_{Vol}, asymmetric unit deposition ΔG_{AU} and asymmetric unit expansion into the unit cell $\Delta G_{AU\rightarrow UC}$ yields the standard state deposition free energy ΔG°_{dep} (eq 2) based on GAUCHE. Results for each compound are given in Table 6, along with a comparison to experimental sublimation free energy ($\Delta G^{\circ}_{dep} = -\Delta G^{\circ}_{sub}$). Overall, the sublimation free energy values from GAUCHE exhibit a mean signed error of −0.75 kcal/mol relative to experiment, which indicates the simulations slightly underestimate crystal stability. Although the favorable free energy to further expand the P1 unit cell end state to a replicated super cell end state (i.e., a system built from 2 × 2 × 2 = 8 unit cells) may correct this, such a minor gain in accuracy did not justify the added computational expense in this case. The mean unsigned error of 1.56 kcal/mol and the root-mean-square error (RMSE) of 1.70 kcal/mol are larger than the AMOEBA solvation free energy RMSE for small organic molecules of 0.68 kcal/mol.[5a,14b,44] However, the small molecules in those solvation free energy studies have an average molecular weight that is only ∼1/3 that of the compounds studied here. It is reasonable to suspect that larger, more complicated molecules would exhibit larger errors. For the solvation free energy of OPLS_2005, the overall MUE for 239 small molecules is 1.10 kcal/mol. However, for amide and ester functional groups that have similar chemistry to those molecules studied here, the OPLS_2005 MUEs are 2.4 and 1.6 kcal/mol, respectively.[44a] Therefore, the overall accuracy of the sublimation free energies obtained through the GAUCHE method is comparable to that of AMOEBA and OPLS_2005 solvation free energies. The average total statistical uncertainty of the deposition free energy calculation is 0.93 kcal/mol, suggesting that further sampling improvements may decrease the GAUCHE RMSE even further. Overall, GAUCHE offers an efficient alchemical simulation protocol to complement existing solvation free energy methods and complete the solubility thermodynamic cycle.

Table 5. Shown Are the Free Energy Values (kcal/mol) for Each Simulation Step in the Thermodynamic Cycle (Figure 3) Used to Compute the Free Energy Change $\Delta G_{AU\rightarrow UC}$ of Moving from a Simulation of the Asymmetric Unit to a Simulation of the Unit Cell

		simulation						
compd.	component	1	2	3	4	5	mean	σ
acetanilide	$\Delta G_{AU\rightarrow RAU}$	9.44	10.05	10.66	9.57	10.32	10.01	0.51
	$\Delta G_{RAU\rightarrow Vac}$	29.48	29.46	29.47	29.47	29.47	29.47	0.01
	$\Delta G_{Vac\rightarrow RUC}$	−29.48	−29.48	−29.48	−29.48	−29.47	−29.48	0.00
	$\Delta G_{RUC\rightarrow UC}$	−9.88	−10.14	−10.88	−9.95	−10.43	−10.26	0.41
	$\Delta G_{AU\rightarrow UC}$						−0.25	0.65
paracetamol I	$\Delta G_{AU\rightarrow RAU}$	8.78	8.71	8.69	8.67	8.73	8.72	0.04
	$\Delta G_{RAU\rightarrow Vac}$	32.90	32.92	32.91	32.92	32.91	32.91	0.01
	$\Delta G_{Vac\rightarrow RUC}$	−33.01	−33.01	−33.01	−33.02	−33.00	−33.01	0.01
	$\Delta G_{RUC\rightarrow UC}$	−9.89	−10.39	−11.38	−9.95	−9.85	−10.29	0.64
	$\Delta G_{AU\rightarrow UC}$						−1.67	0.64
methyl paraben II	$\Delta G_{AU\rightarrow RAU}$	8.28	8.44	8.40	8.33	8.35	8.36	0.06
	$\Delta G_{RAU\rightarrow Vac}$	26.14	26.16	26.14	26.14	26.15	26.14	0.01
	$\Delta G_{Vac\rightarrow RUC}$	−26.22	−26.22	−26.22	−26.22	−26.22	−26.22	0.00
	$\Delta G_{RUC\rightarrow UC}$	−9.76	−9.47	−9.92	−11.58	−9.41	−10.03	0.89
	$\Delta G_{AU\rightarrow UC}$						−1.74	0.89
ethyl paraben	$\Delta G_{AU\rightarrow RAU}$	9.73	9.99	10.18	10.06	10.06	10.00	0.17
	$\Delta G_{RAU\rightarrow Vac}$	24.14	24.13	24.13	24.13	24.14	24.13	0.01
	$\Delta G_{Vac\rightarrow RUC}$	−24.14	−24.15	−24.14	−24.15	−24.14	−24.14	0.00
	$\Delta G_{RUC\rightarrow UC}$	−11.59	−10.53	−11.73	−11.27	−10.82	−11.19	0.51
	$\Delta G_{AU\rightarrow UC}$						−1.19	0.54
phenacetin	$\Delta G_{AU\rightarrow RAU}$	12.51	12.36	12.83	12.24	12.67	12.52	0.23
	$\Delta G_{RAU\rightarrow Vac}$	30.20	30.21	30.21	30.18	30.20	30.20	0.01
	$\Delta G_{Vac\rightarrow RUC}$	−30.20	−30.20	−30.20	−30.20	−30.20	−30.20	0.00
	$\Delta G_{RUC\rightarrow UC}$	−15.93	−16.27	−13.56	−15.72	−16.51	−15.60	1.18
	$\Delta G_{AU\rightarrow UC}$						−3.08	1.20

dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 6. Calculated and Experimental Absolute Sublimation Free Energies for Each Compound (kcal/mol)

compd.	ΔG_{Vol}	ΔG_{AU}	$\Delta G_{AU\rightarrow UC}$	calc. ΔG°_{sub}	expt. ΔG°_{sub}	abs. ΔE	σ
acetanilide	1.30	−12.42	−0.25	11.37	11.57	0.20	0.72
paracetamol I	1.27	−13.98	−1.67	14.38	16.23	1.85	0.86
methyl paraben II	1.32	−9.81	−1.74	10.23	11.98	1.75	0.94
ethyl paraben	1.21	−10.31	−1.19	10.29	12.27	1.98	0.89
phenacetin	1.14	−14.47	−3.08	16.41	14.39	2.02	1.22
mean						1.56	0.93

AMOEBA Structures and Hydrogen-Bonding Compared to Experiment. Optimization of stochastic dynamics snapshots from the OSRW/AMOEBA simulations produced conformations that had slightly more favorable potential energies relative to local optimization of experimental crystal structures. As shown in Table 7, the potential energy difference between the optimized experimental structures and the lowest potential energy found from optimizing stochastic dynamics snapshots was quite small (0.17 kcal/mol or less), except for phenacetin. In this case, the AMOEBA crystal minimum exhibits a 180° torsional flip relative to the experimental crystal structure for an aromatic ether functional group. Although ether groups have been studied during development of the AMOEBA force field,[5a] further systematic work on aromatic ethers may be needed.

Table 7. Lattice Energies of the AMOEBA Potential from (1) the Minimized Experimental Coordinates and (2) the Lowest Potential Energy Found via Minimization of Simulation Snapshots (kcal/mol)

compd.	expt. minimized	snapshot minimized	ΔE
acetanilide	−22.556	−22.725	−0.169
paracetamol I	−27.694	−27.696	−0.002
methyl paraben II	−22.063	−22.096	−0.033
ethyl paraben	−24.580	−24.598	−0.018
phenacetin	−26.124	−27.247	−1.123

Overall, hydrogen bond distances for the AMOEBA minimum energy structures are nearly identical to those of the experimental crystal structures, as shown in Table 8. The mean N−H··O distance for the experimental crystal structures and AMOEBA minima were 2.943 and 2.946 Å, respectively. The mean O−H··O distance for the experimental crystal structures and AMOEBA minima were 2.738 and 2.754 Å, respectively. The discrepancies for N−H··O and O−H··O of 0.003 and 0.012 Å, respectively, are of similar magnitude to experimental uncertainty and refinement artifacts.

Table 8. Hydrogen-Bond Distances (Å) for the Experimental Crystal Structure, for AMOEBA Minimization of the Experimental Structure, and for AMOEBA Minimization of the Most Favorable Snapshot from the Alchemical Simulations

compd.	H-bond	expt.	expt. minimized	snapshot minimized
acetanilide	N−H··O	2.943	3.030	2.901
paracetamol I	N−H··O	2.934	2.971	2.972
	O−H··O	2.663	2.728	2.728
methyl paraben II	O−H··O	2.687	2.727	2.727
ethyl paraben	O−H··O	2.864	2.799	2.808
phenacetin	N−H··O	2.953	3.086	2.964
mean	N−H··O	2.943	3.029	2.946
mean	O−H··O	2.738	2.751	2.754

Comparison of Lattice Energy, End-State, and Alchemical Approaches. We conclude the Results section with a brief comparison of lattice energy, end-state, and alchemical approaches for predicting the stability of organic crystals as shown in Table 9. Lattice potential energy is the simplest metric and is often used for polymorph prediction; however, it completely neglects entropic effects and the temperature dependence of stability. On average, this approach overestimates stability by 11.3 kcal/mol at room temperature. An end-state approach based on single structure estimates of vibrational entropies offers an inexpensive, but approximate framework to begin accounting for entropic changes between vapor and crystalline end states. In this case, the stability is underestimated on average by 8.2 kcal/mol. Finally, GAUCHE offers a rigorous way to account for entropic contributions to the absolute stability of organic crystals as a function of temperature. Although more expensive than lattice potential energy and end-state approaches by a factor of ∼10^9, the mean signed error is below 1.0 kcal/mol.

Table 9. Negative of the AMOEBA Lattice Energy, the Negative of the End-State Approximation ΔG°_{ES} for Deposition Free Energy, the GAUCHE Prediction for Sublimation Free Energy $\Delta G^{\circ}_{GAUCHE}$ and the Experimental Sublimation Free Energy (kcal/mol)

compd.	$-E_{Lattice}$	$-\Delta G^{\circ}_{ES}$	$\Delta G^{\circ}_{GAUCHE}$	expt.
acetanilide	22.56	4.17	11.37	11.57
paracetamol I	27.69	7.58	14.38	16.23
methyl paraben II	22.06	4.58	10.23	11.98
ethyl paraben	24.58	3.20	10.29	12.27
phenacetin	26.12	6.11	16.41	14.39
RMSE	11.34	8.19	1.70	

CONCLUSIONS

The goal of this work was to describe an efficient alchemical method for computing absolute organic crystal thermodynamics that is analogous to solvation free energy methods. Five organic molecules were chosen whose experimental sublimation free energy has been measured (Tables 1 and 2). Our method begins by computing the sublimation-deposition phase transition free energy ΔG_{AU} based on a simulation system composed of only the asymmetric unit molecules (Table 4). Next, we demonstrated a novel strategy to compute the free energy change between a crystalline simulation system with only N_{AU} asymmetric unit molecules and one with N_{UC} unit cell molecules. The magnitude of the correction $\Delta G_{AU\rightarrow UC}$ (Table 5, eq 8) was the smallest for acetanilide, the smallest molecule studied, and the largest for phenacetin, the largest molecule

dx.doi.org/10.1021/ct500180m | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

studied. This is consistent with the premise that as the size of the asymmetric unit of the crystal increases so does the importance of entropic contributions to stability. Overall, the GAUCHE approach achieved a RMSE of 1.70 kcal/mol and mean signed and unsigned errors of −0.75 and 1.56 kcal/mol, respectively.

For our thermodynamic calculations, convergence of torsional degrees of freedom relies on long trajectories (i.e., 200 ns) and some degree of acceleration from the OSRW. However, as the number of torsional degrees of freedom increases, further sophistication of our sampling methodology may be required.[45] In future work, explicit support for enhanced sampling of rotatable torsions will be pursued to reduce statistical uncertainty and enable simulation of more complex molecules with less computation.

Compared to alchemical approaches that begin from the unit cell, GAUCHE has the appealing advantage of not requiring a priori knowledge of crystalline coordinates determined from X-ray crystallography experiments. However, this work focused on NVT simulations, which do require knowledge of the space group and unit cell parameters. Work is underway to extend GAUCHE to the NPT ensemble, and thereby eliminate reliance on all experimental information and enable crystal structure prediction. For example, the most favorable deposition free energy from NPT simulations of all reasonable space groups (i.e., those with mirror planes are eliminated for chiral molecules) would predict the space group of the most stable polymorph, its unit cell parameters, and its structural ensemble.

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Email: michael-schnieders@uiowa.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Jeff Blaney for the suggestion to apply our space group implementation of AMOEBA to organic crystal thermodynamics. This work was inspired by Christopher Bayly's insightful suggestion to compare our approach directly to experimental sublimation free energy, in addition to experimental solubility, to clearly separate out sublimation issues. J.P. acknowledges Vertex Pharmaceuticals for their support during an internship (Summer 2013). We thank Johnny Wu, Pengyu Ren, and Jay Ponder for helpful discussions regarding parametrization of the AMOEBA force field for organic molecules and Wei Yang for advice on using OSRW. All computations were performed on The University of Iowa Neon cluster and depended on support and guidance from Neon system administrators Glenn Johnson and Ben Rogers.

■ REFERENCES

(1) Palmer, D. S.; Llinàs, A.; Morao, I.; Day, G. M.; Goodman, J. M.; Glen, R. C.; Mitchell, J. B. O. Predicting intrinsic aqueous solubility by a thermodynamic cycle. *Mol. Pharmaceutics* **2008**, *5* (2), 266−279.

(2) (a) Jorgensen, W. Monte Carlo simulation of differences in free energies of hydration. *J. Chem. Phys.* **1985**, *83* (6), 3050−3054. (b) Lybrand, T. P.; Ghosh, I.; McCammon, J. A. Hydration of chloride and bromide anions: Determination of relative free energy by computer simulation. *J. Am. Chem. Soc.* **1985**, *107* (25), 7793−7794.

(3) Paluch, A. S.; Shah, J. K.; Maginn, E. J. Efficient solvation free energy calculations of amino acid analogs by expanded ensemble molecular simulation. *J. Chem. Theory Comput.* **2011**, *7* (5), 1394−1403.

(4) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys.* **2003**, *119* (11), 5740−5761.

(5) (a) Ren, P.; Wu, C.; Ponder, J. W. Polarizable atomic multipole-based molecular mechanics for organic molecules. *J. Chem. Theory Comput.* **2011**, *7* (10), 3143−3161. (b) Mobley, D.; Liu, S.; Cerutti, D.; Swope, W.; Rice, J. Alchemical prediction of hydration free energies for SAMPL. *J. Comput. Aided Mol. Des.* **2012**, *26* (5), 551−562.

(6) (a) Palmer, D. S.; McDonagh, J. L.; Mitchell, J. B. O.; van Mourik, T.; Fedorov, M. V. First-principles calculation of the intrinsic aqueous solubility of crystalline druglike molecules. *J. Chem. Theory Comput.* **2012**, *8* (9), 3322−3337. (b) Salahinejad, M.; Le, T. C.; Winkler, D. A. Aqueous solubility prediction: Do crystal lattice interactions help? *Mol. Pharmacol.* **2013**, *10* (7), 2757−2766.

(7) Paluch, A. S.; Maginn, E. J. Efficient estimation of the equilibrium solution-phase fugacity of soluble nonelectrolyte solids in binary solvents by molecular simulation. *Ind. Eng. Chem. Res.* **2013**, *52* (38), 13743−13760.

(8) Jorgensen, W. L.; Duffy, E. M. Prediction of drug solubility from structure. *Adv. Drug Deliver. Rev.* **2002**, *54* (3), 355−366.

(9) 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.; Williams, D. E. A test of crystal structure prediction of small organic molecules. *Acta Crystallogr. B* **2000**, *56* (4), 697−714.

(10) 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.; Zhitkov, I. K. Towards crystal structure prediction of complex organic compounds—A report on the fifth blind test. *Acta Crystallogr. B* **2011**, *67* (6), 535−551.

(11) Perlovich, G. L.; Volkova, T. V.; Bauer-Brandl, A. Polymorphism of paracetamol. *J. Therm. Anal. Calorim.* **2007**, *89* (3), 767−774.

(12) (a) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-β spine of amyloid-like fibrils. *Nature* **2005**, *435* (7043), 773−778. (b) Colletier, J.-P.; Laganowsky, A.; Landau, M.; Zhao, M.; Soriaga, A. B.; Goldschmidt, L.; Flot, D.; Cascio, D.; Sawaya, M. R.; Eisenberg, D. Molecular basis for amyloid-β polymorphism. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (41), 16938−16943.

(13) Schnieders, M. J.; Baltrusaitis, J.; Shi, Y.; Chattree, G.; Zheng, L.; Yang, W.; Ren, P. The structure, thermodynamics, and solubility of organic crystals from simulation with a polarizable force field. *J. Chem. Theory Comput.* **2012**, *8* (5), 1721−1736.

(14) (a) Ren, P.; Ponder, J. W. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* **2003**, *107* (24), 5933−5947. (b) 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.; Head-Gordon, T. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **2010**, *114* (8), 2549−2564. (c) Wu, J. C.; Chattree, G.; Ren, P. Automation of AMOEBA polarizable force field

parameterization for small molecules. *Theor. Chem. Acc.* **2012**, *131* (3), 1−11.
(15) (a) Schnieders, M. J.; Fenn, T. D.; Pande, V. S. Polarizable atomic multipole X-ray refinement: Particle mesh Ewald electrostatics for macromolecular crystals. *J. Chem. Theory Comput.* **2011**, *7* (4), 1141−1156. (b) Fenn, T. D.; Schnieders, M. J. Polarizable atomic multipole X-ray refinement: Weighting schemes for macromolecular diffraction. *Acta Crystallogr. D* **2011**, *67* (11), 957−965.
(16) (a) Zheng, L.; Chen, M.; Yang, W. Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (51), 20227−20232. (b) Zheng, L.; Chen, M.; Yang, W. Simultaneous escaping of explicit and hidden free energy barriers: Application of the orthogonal space random walk strategy in generalized ensemble based conformational sampling. *J. Chem. Phys.* **2009**, *130* (23), 234105. (c) Zheng, L.; Yang, W. Practically efficient and robust free energy calculations: Double-integration orthogonal space tempering. *J. Chem. Theory Comput.* **2012**, *8* (3), 810−823.
(17) (a) Kong, X.; Brooks, C. L., III Lambda-dynamics: A new approach to free energy calculations. *J. Chem. Phys.* **1996**, *105* (6), 2414−2423. (b) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1* (5), 826−843.
(18) (a) Perlovich, G. L.; Volkova, T. V.; Bauer-Brandl, A. Towards an understanding of the molecular mechanism of solvation of drug molecules: A thermodynamic approach by crystal lattice energy, sublimation, and solubility exemplified by paracetamol, acetanilide, and phenacetin. *J. Pharm. Sci.* **2006**, *95* (10), 2158−2169. (b) Perlovich, G. L.; Volkova, T. V.; Manin, A. N.; Bauer-Brandl, A. Influence of position and size of substituents on the mechanism of partitioning: A thermodynamic study on acetaminophens, hydroxybenzoic acids, and parabens. *AAPS PharmSciTech* **2008**, *9* (1), 205−216.
(19) Brown, C. Further refinement of the crystal structure of acetanilide. *Acta Crystallogr.* **1966**, *21* (3), 442−445.
(20) Haisa, M.; Kashino, S.; Kawai, R.; Maeda, H. The monoclinic form of p-hydroxyacetanilide. *Acta Crystallogr. B* **1976**, *32* (4), 1283−1285.
(21) Nath, N. K.; Aggarwal, H.; Nangia, A. Crystal structure of methyl paraben polymorph II. *Cryst. Growth Des.* **2011**, *11* (4), 967−971.
(22) Lin, X. The structure of ethyl paraben. *Chin. J. Struct. Chem.* **1986**, *5*, 281.
(23) Patel, U.; Patel, T. C.; Singh, T. P. Structure of phenacetin, $C_{10}H_{13}NO_2$. *Acta Crystallogr. C* **1983**, *39* (10), 1445−1447.
(24) Mu, X.; Wang, Q.; Wang, L.-P.; Fried, S. D.; Piquemal, J.-P.; Dalby, K. N.; Ren, P. Modeling organochlorine compounds and the σ-hole effect using a polarizable multipole force field. *J. Phys. Chem. B* **2014**, submitted.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, A.02; Gaussian, Inc.: Wallingford, CT, 2009.
(26) (a) Stone, A. J.; Alderton, M. Distributed multipole analysis: Methods and applications. *Mol. Phys.* **1985**, *56* (5), 1047−1064. (b) Stone, A. J. Distributed multipole analysis: Stability for large basis sets. *J. Chem. Theory Comput.* **2005**, *1* (6), 1128−1132.
(27) Ponder, J. W. *TINKER: Software Tools for Molecular Design*, 5.0; J.W. Ponder: Saint Louis, MO, 2009.
(28) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle-mesh Ewald method. *J. Chem. Phys.* **1995**, *103* (19), 8577−8593.
(29) Darden, T.; York, D.; Pedersen, L. Particle-mesh Ewald: An *n* log(*n*) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.
(30) Sagui, C.; Pedersen, L. G.; Darden, T. A. Towards an accurate representation of electrostatics in classical force fields: Efficient implementation of multipolar interactions in biomolecular simulations. *J. Chem. Phys.* **2004**, *120* (1), 73−87.
(31) (a) Dovesi, R.; Orlando, R.; Civalleri, B.; Roetti, C.; Saunders Victor, R.; Zicovich-Wilson Claudio, M. CRYSTAL: A computational tool for the ab initio study of the electronic properties of crystals. *Zeitschrift für Kristallographie* **2005**, *220*, 571. (b) Dovesi, R.; Saunders, V. R.; Roetti, R.; Orlando, R.; Zicovich-Wilson, C. M.; Pascale, F.; Civalleri, B.; Doll, K.; Harrison, N. M.; Bush, I. J.; D'Arco, P.; Llunell, M. *CRYSTAL09 User's Manual*, University of Torino: Torino, 2009.
(32) (a) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle−Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37* (2), 785−789. (b) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648−5652.
(33) Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1* (2), 211−228.
(34) Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **1976**, *13* (12), 5188−5192.
(35) (a) Doll, K. Implementation of analytical Hartree−Fock gradients for periodic systems. *Comput. Phys. Commun.* **2001**, *137* (1), 74−88. (b) Doll, K.; Saunders, V. R.; Harrison, N. M. Analytical Hartree−Fock gradients for periodic systems. *Int. J. Quantum Chem.* **2001**, *82* (1), 1−13. (c) Civalleri, B.; D'Arco, P.; Orlando, R.; Saunders, V. R.; Dovesi, R. Hartree−Fock geometry optimization of periodic systems with the Crystal code. *Chem. Phys. Lett.* **2001**, *348* (1−2), 131−138.
(36) (a) Broyden, C. G. The convergence of a class of double-rank minimization algorithms 1. General considerations. *IMA J. Appl. Math* **1970**, *6* (1), 76−90. (b) Fletcher, R. A new approach to variable metric algorithms. *Computer J.* **1970**, *13* (3), 317−322. (c) Goldfarb, D. A family of variable-metric methods derived by variational means. *Math. Comput.* **1970**, *24* (109), 23−26. (d) Shanno, D. F. Conditioning of quasi-newton methods for function minimization. *Math. Comput.* **1970**, *24* (111), 647−656.
(37) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **1971**, *54* (2), 724−728.
(38) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. State of the art in counterpoise theory. *Chem. Rev.* **1994**, *94* (7), 1873−1885.
(39) Tidor, B.; Karplus, M. The contribution of vibrational entropy to molecular association: The dimerization of insulin. *J. Mol. Biol.* **1994**, *238* (3), 405−414.
(40) Ren, P.; Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comput. Chem.* **2002**, *23* (16), 1497−1506.
(41) Ren, P.; Ponder, J. W. Temperature and pressure dependence of the AMOEBA water model. *J. Phys. Chem. B* **2004**, *108* (35), 13427−13437.
(42) Rivera, S. A.; Allis, D. G.; Hudson, B. S. Importance of vibrational zero-point energy contribution to the relative polymorph energies of hydrogen-bonded species. *Cryst. Growth Des.* **2008**, *8* (11), 3905−3907.
(43) (a) Piquemal, J. P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. Towards a force field based on density fitting. *J. Chem. Phys.* **2006**, *124* (10), 104101. (b) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J. P. Anisotropic, polarizable molecular mechanics studies of inter- and intramoecular interactions and ligand−macro-

molecule complexes. A bottom-up strategy. *J. Chem. Theory Comput.* **2007**, *3* (6), 1960−1986.
(44) (a) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comput.* **2010**, *6* (5), 1509−1519. (b) Shi, Y.; Wu, C.; Ponder, J. W.; Ren, P. Multipole electrostatics in hydration free energy calculations. *J. Comput. Chem.* **2011**, *32* (5), 967−977.
(45) Mobley, D. L.; Chodera, J. D.; Dill, K. A. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **2006**, *125* (8), 084902.