

A Comparison of Methods for Computing Relative Anhydrous–Hydrate Stability with Molecular Simulation

Eric C. Dybeck,* Andrew Thiel, Michael J. Schnieders, Frank C. Pickard, IV, Geoffrey P.F. Wood, Joseph F. Krzyzaniak, and Bruno C. Hancock



Cite This: *Cryst. Growth Des.* 2023, 23, 142–167



Read Online

ACCESS |



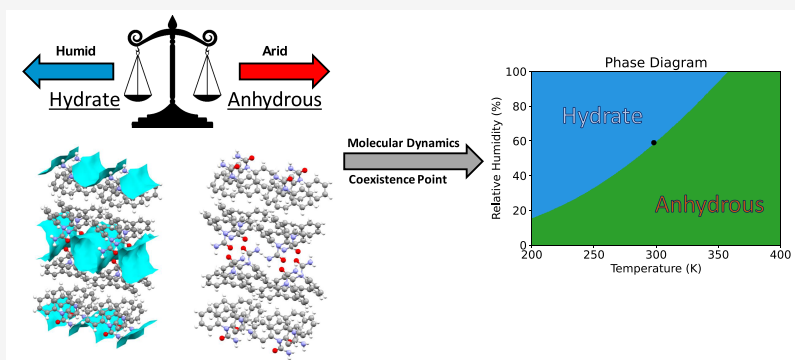
Metrics & More



Article Recommendations



Supporting Information



ABSTRACT: The transformation of a pharmaceutical solid from an anhydrous crystal into a hydrated form during drug development represents a risk to a product's safety and efficacy due to the potential impact on stability, bioavailability, and manufacturability. In this work, we examine 10 classical free energy simulation protocols to evaluate the thermodynamic stability of hydrated crystals relative to their anhydrous forms. Molecular dynamics simulations are used to compute the Gibbs free energies of the crystals of three pharmaceutically relevant systems using two fixed-charge potentials, GAFF and OPLS, as well as the polarizable AMOEBA model. In addition, we explore a variety of water models, including TIP3P, TIP4P, and AMOEBA, for both the interstitial water and the effects of ambient humidity. The AMOEBA model predicts free energy values most consistent with experimental measurements among the models examined. The benefits of a fully polarizable water model relative to fixed-charged models appear to derive predominantly from a better treatment of water's dipole moment in the crystalline phase. Despite this improved physical treatment, we find that no single model produces reliable predictions of the phase boundary between hydrated and anhydrous crystals from theory alone. However, we show that accurate phase diagrams can be constructed from the simulations by introducing a single experimentally determined coexistence point. With this single experimental data point as input, the phase boundary is correctly predicted within 10% relative humidity on the temperature range of 15 to 75 °C for all three systems examined. Furthermore, we demonstrate that with this known coexistence point as an input, the differences between the various potentials and the water models become insignificant, as all models yield accurate phase boundaries regardless of whether polarization is included due to significant temperature-dependent error cancellation between models.

INTRODUCTION

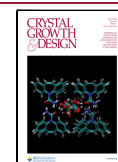
A key milestone in the drug development pipeline is the solid form selection of the Active Pharmaceutical Ingredient (API) for preclinical, clinical, and commercial development. Appropriately selecting the optimal form for development is of paramount importance because of its link to bioperformance, stability, and manufacturability, among other key properties.^{1–8} Ideally, a thermodynamically stable solid form is chosen so that it will not undergo a spontaneous form change during manufacture, storage, and dosing. However, this is not always possible, particularly when there is a thermodynamic phase boundary close to ambient conditions.

Research into polymorphism and other solid form changes increased significantly in the mid-1990s following the high-profile late-stage polymorphic transition and subsequent recall of ritonavir.^{9,10} However, despite this increased focus, a number of notable unintended form changes have subsequently occurred both before and after commercializa-

Received: July 22, 2022

Revised: October 24, 2022

Published: November 29, 2022



ACS Publications

© 2022 American Chemical Society

142

<https://doi.org/10.1021/acs.cgd.2c00832>
Cryst. Growth Des. 2023, 23, 142–167

tion.^{11–16} These unintended form changes are a natural consequence of the challenges in experimentally enumerating, isolating, and rigorously evaluating the stability of all possible crystal structures of a given API, especially if solid-state transformation kinetics are slow. Appropriately formulated computational approaches for crystal structure enumeration do not suffer from these kinetic limitations and can act alongside experimental techniques to aid in solid form selection. As a result, physics-based computational models are regularly applied by the pharmaceutical industry.^{17–19}

One of the most prominent development opportunities for computational crystallographic modeling efforts is the inclusion of the effects of temperature and chemical potential (*i.e.*, humidity) on solid form stability. In the pharmaceutical industry, a drug product can experience a wide range of environmental conditions over the course of manufacturing, transportation, and storage that can lead to unintended physical and chemical changes to the drug substance. Thus, accurate predictions of solid form phase diagrams as a function of thermodynamic quantities such as temperature and humidity are of immediate practical importance.

A significant amount of recent effort has focused on incorporating the effects of temperature into solid form stability, first to predict the melting point of anhydrous solids^{20–25} and later extended to predict the enantiotropic transition between two forms.^{19,26–47} Comparably fewer molecular model development efforts have focused on predicting the transition between a *hydrate* and an anhydrous crystal form.^{48–56} This is despite the fact that transitions between an anhydrous and hydrated form are especially alarming in the context of drug product development, as they can decrease aqueous dissolution and solubility by over an order of magnitude¹ and in some cases lead to clinical failures and commercial drug product lot recalls.^{7,57–59} In addition to the added risk factors of a change in hydration state, the thermodynamic phase boundaries between hydrates and anhydrous forms are also notoriously difficult to accurately determine experimentally.⁶⁰ These difficulties in characterizing anhydrous–hydrate transitions, combined with the relatively high risk associated with unintended transitions, make this an ideal application for predictive modeling.^{7,57,61}

Anhydrous–hydrate transitions between molecular crystals have been previously modeled using a quantum quasi-harmonic approximation (QHA).^{53,62,63} In these investigations, the authors found that with first-principles-based molecular modeling, the transition temperature of hydrated and anhydrous metal oxides could be predicted to within 20 °C of the experimental values in all three systems examined.^{53,62,63}

In crystals of pharmaceutically relevant small molecules, the (quasi-)harmonic approximation can break down for numerous reasons. First, the low-frequency vibrations in organic small molecule crystals often come from weak intermolecular forces such as dispersion or hydrogen bonding, which are highly anharmonic. These anharmonic vibrations result in nonlinear temperature-dependent free energies, particularly for larger flexible molecules.^{19,27,36,47} The low-frequency modes also can come from torsional rotations, which make explicit sampling of the conformational space especially important. Pharmaceutical small molecules also frequently crystallize into solid forms with dynamic disorder, which has been previously shown to cause deviations in the predicted chemical potentials computed with the harmonic approximation.³⁶ Finally, in the

case of hydrates, lattice water molecules undergo librational motions (*i.e.*, hindered translations and rotations) that are fundamentally anharmonic in nature.^{64–68} For example, in the extreme case of a channel hydrate, the water molecules have translational freedom in at least one direction, making decomposition of the motion into a set of harmonic vibrational modes impossible.

In principle, molecular dynamics (MD) simulations can overcome these difficulties by sampling phase space, capturing both the harmonic and the anharmonic motions that contribute to the overall stability. Therefore, in this work, we explore the efficacy of using MD-based approaches to explore the thermodynamics of anhydrous–hydrate transitions as a function of ambient temperature and relative humidity. To the best of our knowledge, no molecular-dynamics-based free energy approaches incorporating both temperature and relative humidity have been applied to predict anhydrous–hydrate transitions in crystals of pharmaceutically relevant molecules.

For the reasons stated above, modeling hydrates with MD presents some unique challenges relative to modeling anhydrous structures. First, when modeling any organic crystal with molecular dynamics, it is important to select a suitable potential (*i.e.*, force field) to represent the inter- and intramolecular interactions of the organic compound of interest. Often, bespoke point-charge potentials developed from experimental and/or quantum mechanical data are used in this area.^{17,19,69} More expensive multipole^{70–73} or polarizable^{72,74,75} models are also regularly used. However, generalized point-charge models are more convenient, especially if there is appropriate error cancellation.^{30,31,36} While the effect of applying more physically rigorous force fields on anhydrous polymorphs has been explored previously,⁷⁶ their efficacy in modeling hydrated crystals remains an open question.

When modeling hydrates, there are additional considerations, specifically the appropriate treatment of water molecules within the hydrated crystal. TIP3P⁷⁷ is a popular water model used for a variety of solution-phase chemical and biological systems. However, the TIP3P water model has the well-known limitation of having isotropic point-charges, which may not be suitable for modeling anisotropic systems with long-range order, such as periodic crystals. More sophisticated models, such as TIP4P or TIP5P,⁷⁷ incorporate off-site charges, which could potentially provide additional physical accuracy when modeling organic crystals containing water. However, any point-charge model inherently neglects electrostatic polarization, which may be a significant source of error when modeling the anisotropic environment of organic crystals. Finally, and perhaps most importantly, many commonly used water models, such as TIP3P, are explicitly parameterized to reproduce bulk aqueous phase properties, a molecular environment that is fundamentally dissimilar to the small molecule organic crystals under consideration in this study. Therefore, we also explore models that account for polarization effects, specifically AMOEBA.^{74,75,78–85}

A final consideration when modeling the transition from a hydrate to an anhydrous form is how to represent the humid background environment. When considering the anhydrous crystal, the excess water molecules appear in the form of a thermodynamic reservoir of water with a corresponding reference chemical potential. An illustration of this stoichiometric balance and the water reservoir is shown in Figure 1.

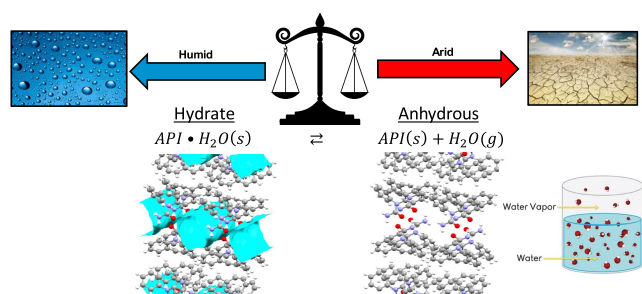


Figure 1. An illustration of the thermodynamic equilibrium between a hydrated and an anhydrous crystal, including the effect of humidity. Under humid conditions, the hydrated form is more favorable, while under arid conditions, the anhydrous form is more favorable. To balance the stoichiometry of the crystalline water in the hydrated form, a water “reservoir” must be included on the same side of the equation as the anhydrous crystal. These additional water molecules can be represented either as liquid or vaporous water. Modeling these two distinct water phases leads to two slightly different chemical potentials in practice, and the effect of this difference is one of the main focuses of this work.

There are multiple ways to simulate 100% ambient relative humidity within the context of a molecular mechanics model, and we consider two in the present study. The first approach involves simulating a box of liquid water. The chemical potential of water in the liquid box can be determined numerically using standard solvation free energy perturbation protocols at each temperature of interest.^{86–88} An alternate approach is to analytically compute the chemical potential of a gas-phase water molecule at the vapor pressure corresponding to each temperature of interest.⁵³ The saturated vapor pressure of water can be taken from empirical steam tables or from approximations such as the Antoine equation. In both the liquid and vaporous reference states, the effect on the chemical potential of shifting to a specific relative humidity is added analytically.^{50,89,90} The competing approaches of vaporous and liquid phase reservoirs are illustrated in Figure 1.

If the modeled vapor and liquid phase reservoirs were both perfect representations of real saturated water, then their chemical potentials would also be equal to each other, in accordance with the zeroth law of thermodynamics, and thus, there would be no difference in the model between these two representations. In practice, the *in silico* boiling point of water varies between the different approximate water models, as we discuss later, and none perfectly match the experimental value due to the simplifications inherent in the model parametrization. The consequence of this discrepancy is that the vapor pressure returned by the Antoine equation does not actually correspond with the (T,P) state at which the vaporous and liquid phases are in thermodynamic equilibrium *in a specific water model*.

Using an empirical value for the saturated vapor pressure would initially appear to provide more accuracy and circumvent the limitations of the water model in capturing the boiling point. However, the approximate water model is still needed to treat the interstitial waters in the simulation of the hydrate regardless of the representation of the saturated water. Therefore, it is more theoretically rigorous to use the explicit simulation of the liquid box with the same approximate water model as this maximizes the self-consistency across phases and benefits from the most cancellation of errors.

Nevertheless, the analytical vapor approach with the Antoine equation has the advantage that it does not require any molecular simulations. Therefore, this approach should be thought of as a traditional “engineering approximation” that attempts to capture the central essence of a chemical system at a significantly reduced computational complexity. The impact of this simplified approximation on computed anhydrous–hydrate relative stability is demonstrated as part of this work.

Ultimately, the above considerations lead to a large combinatorial “model-space” of anhydrous–hydrate transitions with molecular mechanics that spans three dimensions: API potential, water model, and reference water representation. In this work, we present a systematic comparison of these 10 modeling approaches (enumerated in Supporting Information Table S3b) on real anhydrous–hydrate transitions, and we seek to answer whether any choice yields a notably superior prediction when compared with measured experimental observables.

For the API potential, we explore two common point-charge models: OPLS and GAFF, as well as the polarizable AMOEBA model. For the lattice water in the hydrates, we use the TIP3P point-charge model, the off-site point-charge TIP4Pew model,⁹¹ and the polarizable AMOEBA water model. Finally, for the humidity model, we employ both the numerical/liquid chemical potential and the vaporous/empirical chemical potential approaches described above.

The specific questions that we seek to answer in this study are as follows:

1. How sensitive are computed transition enthalpy, entropy, and free energy estimates to the choice of API potential and interstitial water model?
2. How well does a box of liquid water with the various water models mimic the chemical potential of an analytical vapor water at empirically tabulated vapor pressures?
3. Does any combination of models give good enough agreement to be practically useful for modeling the experimental phase transition of pharmaceutical hydrates?

To answer the above questions, we generate the temperature-dependent Gibbs free energy curves for anhydrous and hydrate crystals of three pharmaceutically relevant small molecule systems and compare them at different relative humidities. We then use these values to compute the following observables: transition temperature, critical water activity, enthalpy, entropy, and free energy of dehydration. The small molecule systems examined in this work include two rigid molecules: carbamazepine and theophylline, as well as gandotinib, which contains multiple rotatable bonds (Figure 2). These systems were chosen because they all have

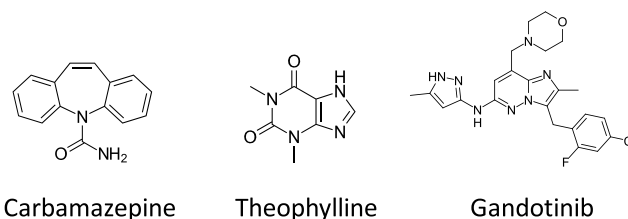


Figure 2. The chemical structures of the three compounds examined in this work.

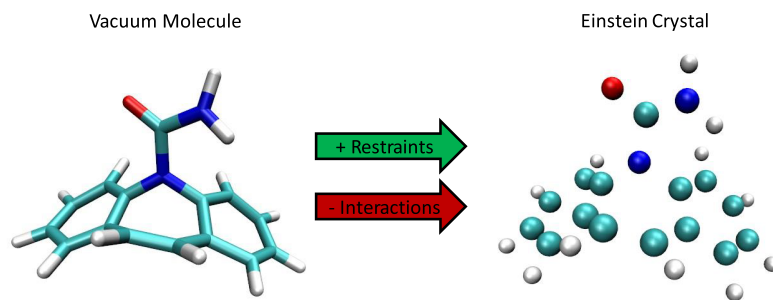


Figure 3. The Einstein crystal method (ECM) for computing the free energy of a molecule in a vacuum relative to the ideal Einstein crystal state. On the left, a molecule of carbamazepine is shown in the vacuum state. On the right, the Einstein crystal state is shown where all atoms have been restrained and all interatomic interactions have been removed. The ideal Einstein crystal state is a convenient reference state for crystalline phases because the physical state can be connected to the reference state without passing through a first-order phase transition.

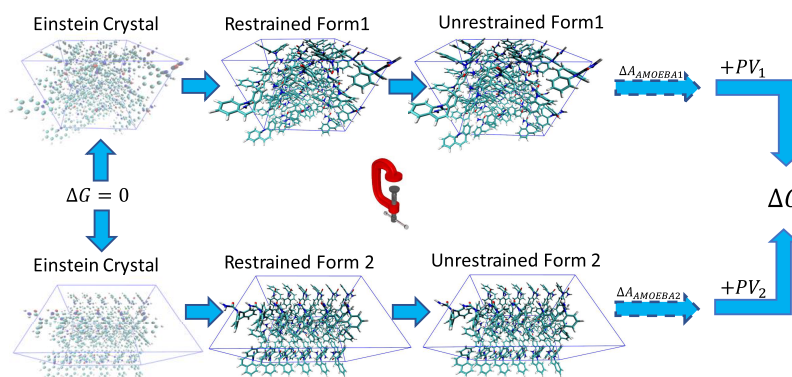


Figure 4. The Einstein crystal method (ECM) uses the above thermodynamic cycle to compute the relative free energy between two polymorphs. With the ECM, the two structures both begin in an ideal Einstein crystal state, where all inter- and intramolecular interactions are removed and all atoms are harmonically restrained. In the next step, the interactions in the crystal are gradually reintroduced followed by restraint removal. Finally, the pressure–volume terms of each structure are added, yielding the Gibbs free energy difference between the structures.

pharmaceutical relevance, the hydrates span multiple water stoichiometries ranging from 1:1 to 4:1, and they are well studied systems with a significant amount of published experimental data.^{50,60,92–94} In all cases, we compare the computed thermodynamic quantities to their experimentally measured counterparts.

In addition to modeling the anhydrous–hydrate transitions in these systems, we also evaluate the efficacy of the same potentials to estimate the thermodynamics of enantiotropic polymorph–polymorph transitions, as well as the conversion from the crystal into the melted liquid state. Improvements in the computed anhydrous–hydrate transitions with polarizable potentials could come from either an improved treatment of the API molecules or the water molecules in either the interstitial or liquid state. The anhydrous–anhydrous and solid–liquid transitions in these systems are a useful control, allowing us to probe the effect of polarizable force fields on the API molecules separately from the effect on the water molecules.

Finally, in industrial modeling settings, it is common to construct a model that combines both theoretical and experimental data rather than relying on theory alone. Indeed, preliminary experimental data are often available prior to model building, and thus there is no additional cost to include them. These hybrid models are typically used to interpolate or even extrapolate from known conditions to other conditions, where experimental measurements have not been conducted, due to resource limitations.⁹⁵ Therefore, we also explore the applicability of combining the purely theoretical free energies

from molecular simulation with limited experimental data to construct temperature–humidity phase diagrams over the entire temperature range of interest.

In the following section, we provide a detailed explanation of the method for computing the free energy of hydrated and anhydrous crystals with molecular dynamics, including a review of previous work and derivations of the underlying theory from first principles. In the subsequent section, we present the computed free energies of the hydrated and anhydrous crystals of the three systems in this work, contrasting the computed results among the various potentials and comparing them to experimental measurements. Finally, we present a hybrid approach that “anchors” the molecular simulation results using a single experimental data point that produces highly accurate phase diagrams. We then discuss the potential applications for this technology and how it can support future drug product development efforts.

METHODS

In this work, we compute Gibbs free energy values of all systems using the Einstein crystal method (ECM) with atomistic molecular dynamics simulations (Figures 3 and 4). The fundamentals of the ECM methodology have been described in detail previously.^{32–34,40,96–99} The evaluation of hydrate stabilities relative to anhydrous forms in this work requires several modifications to the original ECM, most notably to account for a difference in stoichiometry between the forms being compared. In addition, the molecules in this work are treated as fully flexible, rather than the more traditional rigid molecule treatment, to account for the important role of conformational sampling to the crystal stability,

particularly for the dynamically disordered gandotinib crystals. Finally, the humidity effects are treated using both a liquid and vapor phase model, which ultimately involves connecting solid, liquid, and vapor phase systems through the same ECM protocol. Previously, the fully flexible ECM was used to compare anhydrous polymorphs,^{19,36,76} as well as to compare solid and liquid phases.^{99,100} However, to the best of our knowledge, this is the first study to apply the fully flexible ECM to organic small molecule crystals of different stoichiometry, as well as the first study connecting solid, liquid, and vapor phase systems.

Because of the significant technical complexities enumerated above, we provide a full review of the ECM methodology, noting the additions here to account for different phases, flexible molecules, crystals with disorder, and crystals of different stoichiometries. Emphasis is placed on defining the analytical thermodynamic reference state because the analytical free energy values between crystals of different stoichiometry do not trivially cancel in contrast to previous work, which focused solely on anhydrous polymorphs. In the following sections, we provide a rigorous definition of the thermodynamic reference state followed by a detailed description of how systems of different phases and molecular compositions can be connected to this reference state within the ECM framework. An extended discussion of the foundational statistical mechanics underlying this method can be found in the [Supporting Information](#).

Ideal Frozen Particle Reference State. The ECM involves a series of simulations where a physical system is converted into an ideal reference state of noninteracting stationary particles corresponding to zero free energy ($G \equiv 0$). While there are numerous possible thermodynamic reference states that can be used in molecular simulations, such as "all elements in their standard state at 298 K and 1 atm" (STP) or the Ben-Naim reference with all molecules in solution at 1 M concentration,¹⁰¹ the "ideal frozen particle" is a natural choice for crystalline systems, as we will discuss later in this section. Because this reference state is less common, we will briefly define it from statistical mechanics first principles. An extended discussion of the statistical mechanics is also presented in the [Supporting Information](#) and in Vega and Noya.³³

Consider an arbitrary system containing a series of translationally fixed ideal gas particles. A visual depiction of this frozen particle state is shown in [Figures 3](#) and [4](#). The total partition function for this system consists of a product of partition functions corresponding to the nuclear, electronic, vibrational, rotational, and translational contributions:

$$Q_{\text{tot}} = Q_{\text{nuc}} Q_{\text{elec}} Q_{\text{vib}} Q_{\text{rot}} Q_{\text{trans}} \quad (1)$$

For classical simulations, no nuclear effects are included in the Hamiltonian ($Q_{\text{nuc}} = 1$). For ideal, noninteracting particles in a vacuum with fixed positions, [eq 1](#) reduces to unity since $Q_{\text{nuc}} = Q_{\text{elec}} = Q_{\text{vib}} = Q_{\text{rot}} = Q_{\text{trans}} = 1$. This leads naturally to the defined "ideal frozen particle" reference state used within this work:

$$G = -k_{\text{B}}T \ln Q_{\text{tot}} = 0 \quad (2)$$

From the above equation, we observe that this relation holds regardless of thermodynamic state parameters such as temperature, pressure, volume, and number of particles. We show later that this robust independence of the reference state, particularly to temperature and number of particles, is quite useful, as it enables comparisons between different phases and different stoichiometries.

The objective of the Einstein crystal method is to construct a smooth thermodynamic pathway connecting the physical system of interest with the frozen particle reference state. In practice, molecular dynamics or Monte Carlo (MC) simulations are used to numerically estimate the free energy difference between the physical systems and the Einstein crystal, where all particles are noninteracting and restrained within a harmonic potential. From this ideal Einstein crystal state, the conversion to the frozen particle reference state is an analytical expression, and therefore, it is a convenient reference state for the ECM. The analytical expression to compute the free energy of the Einstein crystal relative to the frozen particle reference state is provided later in [eq 3](#) and derived in more detail in [Supporting](#)

[Information Section 4](#). In the following section, we describe in general terms how solid, liquid, and vacuum phase systems of arbitrary size and chemical composition can be transformed into the ideal Einstein crystal state using the ECM. In the [Simulation Details](#) section, specific implementation details are given for constructing the thermodynamic pathway via alchemical simulation in GROMACS.¹⁰²

Einstein Crystal Method for Vacuum Phase Compounds. As an illustrative example of the Einstein crystal method for converting a physical system into the ideal Einstein crystal state, consider a single carbamazepine molecule in a vacuum. The process to connect this compound to the ideal Einstein crystal is shown in [Figure 3](#). On the left pane of [Figure 3](#), the atoms in the molecule are fully connected through the usual valence bond potentials, and the molecule can rotate freely. On the right pane, harmonic restraints have been added to all atoms in the molecule. In addition, all interatomic interactions have been alchemically removed from the system, resulting in a set of noninteracting particles in a 3D harmonic potential. These two transformations are performed numerically using standard alchemical free energy perturbation techniques.^{103–105} The final step, to compute the free energy of the set of noninteracting harmonic oscillators relative to the translationally frozen particle state, is computed analytically. The Helmholtz free energy of a vacuum molecule is:

$$\begin{aligned} \Delta A_{\text{mol}} &= \Delta A_{\text{ECM}} + \Delta A_{\text{EC}} = \Delta A_{\text{ECM}} - k_{\text{B}}T \sum_{i=1}^N \ln Q_{\text{EC},i} \\ &= \Delta A_{\text{ECM}} - \frac{3}{2} k_{\text{B}}T \sum_{i=1}^N \ln \frac{4\pi^2 m_i k_{\text{B}}^2 T^2}{k_i h^2} \end{aligned} \quad (3)$$

where ΔA_{mol} is the free energy of the molecule; ΔA_{ECM} is the free energy to convert to the physical molecule from the Einstein crystal state, which is the negative of the free energy that comes from applying harmonic restraints and decoupling physical interactions from the vacuum molecule; ΔA_{EC} is the free energy of the Einstein crystal relative to the frozen particle reference state; m_i , k_i , and $Q_{\text{EC},i}$ are, respectively, the mass, harmonic spring constant, and partition function of particle i in the resulting Einstein crystal; and the summation runs over all particles in the molecule $i = 1 \dots N$ except for a single reference atom at index j that has the trivial value $Q_{\text{EC},j} = 1$ for reasons that we describe in the following paragraph.

A subtle, yet important, factor to consider in the above system is how to define the all-atom harmonic restraints. Specifically, the equilibrium positions of the harmonic restraints must be defined using internal coordinates, rather than absolute coordinates, since the molecule is in a vacuum and there is, by definition, nothing else in the system upon which to tether the molecule. In the Einstein crystal method, the harmonic restraints are defined relative to a single reference atom in the molecule. As an implementation detail, this internal-coordinate-harmonic-restraint is accomplished by holding a single reference atom translationally fixed throughout the duration of the simulation, at which point the "system coordinates" (sometimes referred to as the "lab frame") become indistinguishable from the coordinates relative to the reference atom (the "molecular frame"). Because the reference atom is translationally fixed, this atom is often referred to as the "frozen atom".³⁴

Translationally constraining a single atom in the system is equivalent to defining this atom as the "origin" of the Cartesian coordinate system. In all vacuum and periodic systems, the three degrees of freedom corresponding with translations of the entire system in all three spatial dimensions are extraneous. Indeed, the full-system translations relative to the origin are typically explicitly removed from the simulation to produce more numerically stable trajectories (see [Supporting Information Section 5](#)). Because of this translational invariance, any particle (or collection of particles) can be used to define the origin of the coordinate system without influencing the underlying phase space. The most widely used choice of origin for molecular simulations is the center of mass of the system. For various reasons that have been discussed previously, a more natural formulation of the Einstein crystal method involves using a single atom in the system as the origin rather than the center of mass.³⁵

The distinction between using the center of mass as the origin or using a single atom close to the center of mass as the origin may appear negligible (and in the absence of atomic restraints, it is negligible); however, in free energy perturbation simulations where atoms are being harmonically restrained to the absolute system coordinates (*i.e.*, lab frame), the difference in the computed free energy can become significant, particularly for small systems with only a few particles.^{40,97,106} See [Supporting Information Section 5](#) for an extended discussion of this difference. It is important to reiterate here that in any simulation where an atom is translationally fixed, the translational center of mass motion should *not* be removed, as is otherwise common practice. If the center of mass of the system is also held fixed, the translational constraints on the reference atom would violate the equipartition theorem and lead to spurious simulations.³⁴

In this work, a single frozen atom is used as the origin for all simulations rather than using the center of mass. For vacuum phase simulations, the carbon atom closest to the center of mass is used. For periodic crystals, the carbon atom closest to the center of mass of the first molecule in the system is used. Ultimately, the choice of reference atom is arbitrary, and in [Supporting Information Section 5](#), we show this equivalence by computing the free energy of gas phase carbamazepine with three different frozen reference atoms with the results being statistically indistinguishable.

Einstein Crystal Method for Periodic Systems. The Einstein crystal method to convert a periodic crystal or liquid phase into an Einstein crystal follows a similar process to the one described above for vacuum phase compounds. [Figure 4](#) depicts this process for two forms of carbamazepine. First, atomic restraints are applied to atoms in the crystal. Next, the physical intermolecular and intramolecular interactions in the system are removed from all atoms within the restrained crystal, leaving a set of ideal particles within 3D harmonic potentials. Finally, the particles in the harmonic potentials are converted to the frozen reference state, where all particles are translationally fixed. As with the vacuum ECM, the first two transformations are performed numerically using alchemical free energy perturbation, and the final transformation, from the harmonic potentials to the frozen reference state, is performed analytically.

The thermodynamic pathway, to convert the physical crystal into the Einstein crystal, is split into multiple independent steps and multiple lambda windows within each step (more details are provided later in the [Simulation Details](#) section). Importantly, the application of harmonic restraints to all the atoms in the crystal is broken into two steps, with some atoms being restrained before the decoupling of nonbonded interactions and the other atoms being restrained after nonbonded interactions have been removed. This two-step restraint process is done so that hydrogens, disordered atoms, and symmetry-related atoms can sample all equivalent positions before being restrained. The initial restraints are applied on nondisordered heavy atoms and act as the order parameters defining the polymorph. Restraining the crystal atoms relative to these order parameters prevents a first-order phase transition into a fluid phase during the decoupling of intermolecular interactions. One can, in principle, choose to apply biasing restraints on any number of alternative order parameters, such as center-of-mass or orientation distribution functions.^{107,108} Ultimately, the different simulations with varying restraint schemes should converge to the same free energy estimate given a sufficient sampling time. The efficiency of different order parameter choices with respect to the convergence of the free energy estimate is outside the scope of the present study. The intermolecular interactions are removed from the partially restrained crystal, which allow the remaining unrestrained atoms to sample all equivalent positions that may have been kinetically hindered in the physical crystal. Finally, the remaining hydrogen and disordered heavy atoms are restrained. The effect on the free energy of neglecting this restraint partitioning for disordered/symmetrical atoms is discussed in more detail and demonstrated numerically for the disordered chlorofluorophenyl ring of gandotinib in [Supporting Information Section 6](#).

Finally, another type of system considered within this work, using the ECM formalism, is isotropic liquids. It has been noted previously that it can be challenging to directly convert a liquid into an Einstein

crystal because the alchemical perturbation passes through a first-order phase transition from the disordered liquid state to the ordered solid.^{109,110} One approach to avoid this first-order phase transition is to artificially strengthen the intermolecular interactions in the liquid to arrive at a translationally ordered glassy state, which can then be converted into the Einstein crystal.¹¹⁰

An alternative approach to avoid the first-order phase transition, adopted here, is to weaken the intermolecular interactions to pass through the ideal gas state on the way to the Einstein crystal. Specifically, the intermolecular interactions in the liquid system are removed without adding any harmonic restraints. After this perturbation is complete, the molecules no longer interact and move independently throughout the simulation box effectively as an ideal gas. Next, the translational motion of a single atom in each molecule is (analytically) removed from the system (see [Supporting Information Sections 4 and 5](#) for more details). Finally, the noninteracting and translationally fixed molecules are converted into the ideal Einstein crystal state. This final step to convert the translationally fixed molecules to the Einstein crystal is computed numerically through simulation. However, this thermodynamic transformation is identical to the one performed on a vacuum compound. Therefore, in practice, this transformation only needs to be performed once for each chemical compound in the vacuum state, and then the value can be reused in all subsequent liquid ECM computations where that compound is present.

Channel hydrates represent a "hybrid" case between the solid and liquid ECM pathways described above. In the case of a channel hydrate, the pharmaceutical compounds remain translationally fixed as in a solid, while the water molecules in the crystal are translationally free to move in at least one direction, causing them to resemble the liquid state more closely than the solid state.^{111–113} In these situations, it is formally incorrect to restrain the water molecules as described in the solid phase pathway above. This is because the water molecules visit multiple equivalent/disorder positions within the channel during the simulation and adding all-atom harmonic restraints artificially breaks this indistinguishability (see an extended discussion of disorder and indistinguishability in [Supporting Information Section 6](#)). Channel hydrate systems are connected to the ideal Einstein crystal using a combination of the above ECM pathways for solid and liquid systems. Namely, the primary pharmaceutical compounds are atomically restrained and annihilated, while the water molecules are decoupled without atomic restraints followed by analytically removing translations and transforming the resulting pinned ideal gas molecules into the ideal Einstein crystal. Although this special treatment of co-formers is described in the context of water in channel hydrates, this same treatment could also be applied to other solvent molecules to compute the free energy of solvates using the ECM.

The final free energy estimate for a given periodic system relative to the translationally frozen ideal particle reference state is a summation of the numerical perturbations to convert the physical system into the ideal Einstein crystal as well as the analytical conversion to the ideal frozen particle reference state. This process is shown earlier for a vacuum molecule in [eq 3](#). For a periodic system, the free energy estimate is:

$$\begin{aligned}\Delta A_{\text{system}} &= \Delta A_{\text{ECM}} + \Delta A_{\text{EC}} \\ &= \Delta A_{\text{ECM}} - k_{\text{B}}T \ln \frac{Q_{\text{tr}}}{N_{\text{f}}} - k_{\text{B}}T \sum_{i=1}^N \ln Q_{\text{EC},i} \\ &\quad - k_{\text{B}}T \sum_{m=1}^M N_m \left(\ln \left| \frac{Q_{\text{tr},m}}{N_m} \right| + 1 - \frac{\Delta A_{\text{mol},m}}{k_{\text{B}}T} \right)\end{aligned}\quad (4)$$

where ΔA_{system} is the free energy of the periodic system, ΔA_{ECM} is the negative of the free energy to convert the physical system into the decoupled and restrained Einstein crystal state through applying harmonic restraints and decoupling interactions, Q_{tr} is the translational partition function to integrate the frozen atom over the entire system volume, and N_{f} is the number of molecules in the system

indistinguishable from the one containing the frozen atom. The first summation runs over all $i = 1 \dots N$ particles in the crystal that have been converted into harmonic oscillators, and the second summation runs over all $m = 1 \dots M$ molecules that have been decoupled without adding harmonic restraints (liquid and channel hydrate molecules). In the second summation, $Q_{tr,m}$ is the translational partition function to integrate the "pinned" atom in molecule m over the entire system volume, $\Delta A_{\text{mol},m}$ is the free energy of molecule m pinned at a single atom as described in eq 3, and N_m is the total number of molecules in the system indistinguishable from molecule m . For a more thorough derivation and discussion of the above expressions, see [Supporting Information Sections 4–6](#), as well as previous publications involving the ECM.^{33,34}

It is worth emphasizing here that for anhydrous systems with the same number of molecules, the same harmonic restraints, and (essentially) the same volume, all the terms in eq 4 cancel except for ΔA_{ECM} . For this reason, the additional terms pertaining to the reference state did not need to be explicitly computed in previous studies utilizing the ECM to compare exclusively forms of identical size and stoichiometry. For the comparison of hydrates with anhydrous crystals, the volumes are no longer equivalent, and the numbers of particles in the system are different, which are the reasons that we explicitly define and compute these terms in this work. Furthermore, translationally free molecules in the liquid phase (or channel-hydrate phase) must be treated differently in the ECM to account for the molecular indistinguishability, in contrast to the translationally confined molecules in the solid phase that are inherently distinguishable. Equation 4 represents a universal expression for self-consistently computing the free energy of a system with the ECM regardless of the stoichiometry and regardless of whether the molecules are free in the case of a liquid, confined in the case of a solid, or a combination of the two in the case of a channel hydrate.

Converting from Helmholtz to Gibbs Free Energy. The alchemical perturbations in the Einstein crystal method for periodic systems are generally performed in the canonical ensemble (NVT) to avoid complications related to adding harmonic restraints with pressure coupling. The ECM therefore computes the *Helmholtz* free energy, $A(V)$, of the system at a particular system volume V . To compute the Gibbs free energy (within the NPT ensemble) of the system at finite temperature and pressure, the volume-dependent Helmholtz free energy must be integrated over all possible system volumes:¹¹⁴

$$G(P, T) = -k_B T \frac{1}{V_0} \int_0^\infty \exp \left[-\frac{A(V, T) + PV}{k_B T} \right] dV \quad (5)$$

where $G(P, T)$ is the Gibbs free energy of the system at a particular pressure, P , and temperature, T , and the integral is taken over the space of all possible volumes with the normalization constant V_0 . Conveniently, if the ECM calculation is performed at the equilibrium volume $V(P, T)$ at a particular temperature and pressure, the integral above can be collapsed to:

$$G(P, T) = A(V(P, T), T) + PV(P, T) - k_B T \ln \frac{P\sigma(V, T)}{k_B T} \quad (6)$$

where $A(V(P, T), T)$ is the Helmholtz free energy of the system at a particular temperature and volume and $\sigma(V, T)$ is the standard deviation of the system volume at equilibrium. In the thermodynamic limit, the final "fluctuation" term above vanishes to produce the familiar macroscopic expression $G = A + PV$. In practice, this fluctuation term is typically negligible even for system sizes used in molecular simulations and is generally ignored in ECM calculations. For a full derivation of the above equation and discussion of the negligible contribution of the fluctuation term, see [Supporting Information Section 7](#).

Computing Free Energies across a Range of Temperatures. The final eq 6 from the previous section yields the Gibbs free energy

of a system at a single temperature. Frequently, we are interested in the free energy over a range of temperatures when running ECM calculations (e.g., to compute the transition temperature between two polymorphs or between a solid and liquid phase). To determine the free energy across a range of temperatures, one could, in principle, run separate ECM calculations for each temperature of interest. However, a more computationally efficient approach is to run a series of NPT simulations over the desired range of temperatures (referred to subsequently as the "temperature scan") and then use these simulations to compute the change in free energy as the system moves away from the reference temperature of the ECM calculation to other temperatures of interest.

The reduced free energy of a system across a set of temperatures can be computed using a variety of standard free energy perturbation techniques including Thermodynamic Integration (TI)^{100,109} or the Multistate Bennett Acceptance Ratio (MBAR).^{19,36,76,115} The dimensionalized Gibbs free energy is then recovered from the reduced free energies according to the following equation:

$$G(T) = k_B T \Delta f(T) + \frac{T}{T_{\text{ECM}}} G(T_{\text{ECM}}) - k_B T \ln \frac{V(T)}{V(T_{\text{ECM}})} + \frac{3}{2} k_B T \ln \frac{T}{T_{\text{ECM}}} \quad (7)$$

where T_{ECM} is the temperature at which the ECM calculation is run, $\Delta f(T)$ is the reduced free energy difference between the system at temperatures T and T_{ECM} , and $V(T)$ is the system volume. The final two terms in the above expression are a small correction for center of mass motion removal effects in the temperature scans (see [Supporting Information Section 8](#) for more details). It is important to note here that when computing the stability of crystals of different stoichiometry (or simply a different number of atoms), the Gibbs free energies in eq 7 above must be evaluated in MBAR using the total energy, rather than just the potential energy as in previous works, because the contribution of kinetic energy to the total partition function will not completely cancel.

In previous studies using the ECM for polymorph free energies, the *relative* Gibbs free energies between two polymorphs across a range of temperatures were estimated from the *relative* reduced free energy.^{19,36,76} Here, we compute the *absolute* free energies of individual solid and liquid phases. Absolute free energies in this context refer to the fact that the free energy values are computed up to a common analytical reference state rather than being computed up to another physical polymorph. In this work, the analytical reference state is the ideal translationally frozen particle state defined earlier. It is important to stress that this "absolute" free energy of an individual phase relative to the analytical frozen particle reference state does not have any physical significance in isolation, and it does not correspond to any experimental observable. Nevertheless, computing this absolute Gibbs free energy of individual phases has numerous advantages. First, the results no longer depend on an arbitrary choice of a reference polymorph; therefore, the statistical uncertainty of each system can be assessed individually without being correlated to an arbitrary reference phase. Second, computing the absolute free energy of individual phases allows us to simulate each phase independently without prior knowledge of the other systems it will be later compared against. For example, the phases can be driven to Einstein crystals at different temperatures or with different harmonic restraint values because there is no requirement that the analytical reference state of the forms precisely cancels.

Most importantly for this work, computing the absolute free energy enables comparing phases of different size (i.e., number of particles) and chemical composition (e.g., anhydrous polymorphs with hydrates). This reference-phase-independent free energy across temperature is ultimately made possible because the ideal frozen particle reference state has no temperature or size dependence in the classical limit. Therefore, there is no need to rely on the cancellation of the temperature dependence of a physical target and reference polymorph of equal size and stoichiometry, as was done in previous work.^{19,36,76}

The above expression for the Gibbs free energy of a system is computed using an ECM calculation at a single temperature, T_{ECM} , to supply the reference free energy $G(T_{\text{ECM}})$. The final Gibbs free energy across temperature should theoretically be independent of the choice of this temperature if sufficient sampling has been achieved at each simulated state. Therefore, running the ECM calculation multiple times with different choices of T_{ECM} provides a straightforward method to assess the convergence of the calculations through this thermodynamic cycle closure. Specifically, the ECM calculation can be run at any number of temperatures, and the final Gibbs free energy and uncertainty can be estimated as the average and standard deviation of the mean of the multiple calculations:

$$G(T) = \frac{1}{N} \sum_{i=1}^N G_{T_{\text{ECM},i}}(T) \quad (8)$$

$$\delta G(T) = \sqrt{\frac{1}{N} \sum_{i=1}^N (G(T) - G_{T_{\text{ECM},i}}(T))^2} \quad (9)$$

where $G(T)$ and $\delta G(T)$ are the mean and uncertainty in the computed free energy at temperature T , $G_{T_{\text{ECM},i}}(T)$ is the Gibbs free energy estimate using the ECM calculation at $T_{\text{ECM},i}$, and the summations run over all $i = 1 \dots N$ of the independent ECM calculations. As a best practice, it is always recommended to run ECM calculations of each system with at least two different values for T_{ECM} so that this thermodynamic cycle closure and convergence can be assessed.

Pressure and Humidity Effects. The previous sections and eq 3 describe how to compute the Helmholtz free energy of a translationally constrained molecule in a vacuum at a particular temperature. Equation 7 extends this Helmholtz free energy to other temperatures of interest. However, in almost all situations of practical interest, a "gas phase" molecule (e.g., in a stoichiometric equation) corresponds with the chemical potential of a bath of molecules at nonzero pressure and volume. The thermochemical properties of gas-phase molecules are computed from the simulated vacuum phase ($P = 0$) in the following way:⁵³

$$U(T, P) = U(T, 0) + \frac{3}{2}k_B T \quad (10)$$

$$H(T, P) = U(T, 0) + \frac{3}{2}k_B T + PV = U(T, 0) + \frac{5}{2}k_B T \quad (11)$$

$$\begin{aligned} G(T, P) &= A(T, 0) - k_B T \ln q_{t,v} + PV - k_B T \\ &= A(T, 0) - k_B T \ln q_{t,v} \end{aligned} \quad (12)$$

$$S(T, P) = \frac{H(T, P) - G(T, P)}{T} \quad (13)$$

$$V(T, P) = \frac{k_B T}{P} \quad (14)$$

where the quantities U , H , G , A , S , and V correspond with the energy, enthalpy, Gibbs free energy, Helmholtz free energy, entropy, and volume, respectively, and $q_{t,v}$ is the translational partition function for the reference atom integrated over a volume corresponding with the molar volume of an ideal gas at the given temperature and pressure. The values of $U(T, 0)$ come directly from each vacuum phase simulation, and $A(T, 0)$ comes from eq 7. In the above equations, the term $\frac{3}{2}k_B T$ added to the energy and enthalpy corresponds with the additional kinetic energy from the three degrees of freedom when transitioning from a translationally fixed vacuum compound to a translationally free molecule in a box. The enthalpy is additionally modulated by PV , which for an ideal gas is equivalent to $k_B T$. For the Gibbs free energy, this transition from a translationally restrained state to a freely moving particle is equivalent to adding the

translational partition function as well as the PV term. The additional factor of $k_B T$ comes from the transition from a system of size $N=1$ molecule to a system of $N \rightarrow \infty$ indistinguishable molecules in a macroscopically large thermodynamic bath (see Supporting Information Section 6 for more details). The Gibbs free energies for the gas phase molecules are only needed when computing transitions between the gas phase and another condensed phase. The use of the ECM for predicting boiling points or sublimation points for pharmaceutical compounds is outside the scope of the present study. However, chemical potential plots for the computed gas–liquid transition of the three pharmaceutical compounds in this work are provided in Supporting Information Section 9.

Rigid water molecules are treated separately from the above formulation for flexible molecules because the free energy cannot be computed using the ECM since the "bonds" in a rigid molecule cannot be decoupled. Instead, the energy, enthalpy, and Helmholtz free energy for rigid water molecules are computed analytically rather than through numerical simulation.^{116,117} The expressions for this analytical calculation are provided in Supporting Information Section 14.

The effect of water activity in the comparison of hydrates and anhydrous forms is modeled in this work through both a vapor phase and liquid phase reference state for the water. For the vapor phase reference, the state corresponding with 100% humidity is modeled by setting the water partial pressure equal to the vapor pressure at each temperature. The vapor pressure of water is computed using the Antoine equation:

$$\begin{aligned} \log_{10} P^*(T) &= A - \frac{B}{C + (T - 273.15)} \\ &= 8.07131 - \frac{1730.63}{233.426 + (T - 273.15)} \end{aligned} \quad (15)$$

where P^* is the saturated vapor pressure of water at temperature T . For specific saturation levels (relative humidity) between 0 and 100%, the pressure is modulated by the relative humidity, RH:

$$P(T) = \left(\frac{\text{RH}}{100} \right) P^*(T) \quad (16)$$

and the water activity, humidity, and partial pressure are related through:^{50,89,90,118}

$$a_w = \frac{P(T)}{P^*(T)} = \frac{\text{RH}}{100} \quad (17)$$

The above estimates for the temperature and pressure dependent chemical potential of water assume ideal gas behavior, which will break down at sufficiently high pressures and low temperatures. However, for the temperatures and pressures considered in this work (i.e., "ambient conditions"), this approximation is reasonable (see Supporting Information Section 14 for an additional quantification of this approximation).¹¹⁹

For the liquid phase water reference, the chemical potential of water in the liquid system corresponds with 100% humidity ($a_w = 1$) at each temperature. The chemical potential and entropy for a given water activity and temperature are computed by

$$\mu(T) = \mu^*(T) + k_B T \ln a_w \quad (18)$$

$$S(T) = S^*(T) + k_B \ln a_w \quad (19)$$

where $\mu^*(T)$ and $S^*(T)$ are, respectively, the chemical potential and entropy of the water computed from the liquid phase.

Computing Gibbs Free Energies in AMOEBA. The Gibbs free energy of systems in the AMOEBA potential is computed by first calculating the free energy in the OPLS potential and then adding on a correction term to switch from OPLS to AMOEBA at the physical end state (the so-called bookend correction).^{35,76,104,120} This final switching term at the end state is shown as ΔA_{AMOEBA} in Figure 4.

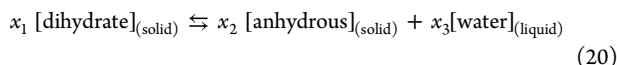
The Helmholtz free energy necessary to switch from OPLS to AMOEBA is computed using the dual force field (DFF) approach as

implemented in ForceFieldX version 1.0.0-beta (<https://ffx.biochem.uiowa.edu/>). The theoretical details about this approach have been described elsewhere.^{82,121} Briefly, the system is simulated with the parameters from both potentials, along with a coupling parameter, λ , that controls whether the forces acting on the particles in the system at each time step come from AMOEBA ($\lambda = 0.0$), from OPLS ($\lambda = 1.0$), or from a weighted-average of the two Hamiltonians for $\lambda \in (0.0, 1.0)$. A two-dimensional Gaussian biasing potential is applied to the λ and $\left(\frac{\partial U}{\partial \lambda}\right)$ variables over time using the Orthogonal Space Random Walk (OSRW) method¹²² until the entire λ -coordinate is traversed. The dependence of the bias on $\left(\frac{\partial U}{\partial \lambda}\right)$ encourages the simulation to overcome hidden barriers orthogonal to the reaction coordinate. Finally, the collection of biasing potentials is integrated to compute the free energy difference between the states.

The DFF switching term is computed in the NVT ensemble using the structure equilibrated in the AMOEBA potential. The system in the OPLS potential is then converted into the Einstein crystal using the ECM as described above. The sum of the ΔA_{AMOEBA} term from the DFF calculation and the ΔA_{system} term from eq 4 for the OPLS system yields the Helmholtz free energy difference between the AMOEBA system and the ideal frozen particle reference state. The temperature scans for each system are performed in the AMOEBA potential, and the Gibbs free energies for the systems across temperature are computed with eqs 6–7.

Stoichiometry Balancing, Normalizing, and Shifting Free Energies. In all free energy-based stability comparisons, it is essential to have an equal number of each compound on the left-hand and right-hand side of the thermodynamic equation. Furthermore, free energies are typically reported on a per-mole basis of a particular compound (*i.e.*, as chemical potentials) to produce more transferable values between studies since free energy is extensive. However, the actual systems that are simulated in the molecular simulations often do not have an identical number of compounds. Therefore, satisfying these two constraints and producing a sensible free energy difference involve both a stoichiometric balancing and a normalization of the raw free energies of each system before being combined.

Consider an equilibrium between carbamazepine dihydrate and an anhydrous polymorph, both with 48 carbamazepine molecules, along with a 500-molecule box of liquid water:



The terms x_i in the above expression are the coefficients for each system i needed to balance the number of carbamazepine and water on the left- and right-hand side and then normalize them per-mole of carbamazepine. In the general case, the coefficients needed to satisfy the constraints can be expressed as a linear series of equations in the form $A\vec{x} = b$ where each element of the matrix A_{ij} contains the total number of compound j in system i :

$$\begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ N_{31} & 0 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 48 & 48 & 0 \\ 96 & 0 & 500 \\ 48 & 0 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (21)$$

The final row of the matrix A contains the total number of the designated "normalization compound" in each system on the left-hand side of the expression, and the corresponding final element of the b -vector is set equal to unity. In the above example, the resulting coefficient vector is $\vec{x} = \left[\frac{1}{48}, -\frac{1}{48}, -\frac{2}{500} \right]$. The final balanced and normalized free energy difference for the thermodynamic equilibrium, ΔG , is then computed as the dot product of the vector of system Gibbs free energies, $\vec{G}$, with the coefficient vector such that $\Delta G = \vec{G} \cdot \vec{x}$.

In solid form stability comparisons like the ones presented in this work, often some melting points or other transition temperatures are already known from experiments. These measured transition temperatures effectively designate points at which $\Delta G = 0$, and

these points can be introduced to the model to give more accurate relative free energies across temperature.

In previous work, we showed that the chemical potentials of crystal structures computed with lower-level point-charge molecular mechanics force fields can be made more accurate by applying a constant "enthalpic shift" to the free energies ($\Delta G(T) \approx \Delta G_{\text{MM}}(T) + \Delta H_{\text{shift}}$). This enthalpic shift effectively invokes a single approximation that the molecular mechanics force field has complete cancellation of errors when the entropies of the systems are referenced against those of the chemical constituents in their standard state (see [Supporting Information Section 10](#) for more details).

In our previous work, the enthalpic shift for each structure was computed using lattice energies derived from quantum mechanics in the limit of zero Kelvin.¹⁹ However, the shift for each system can also be computed using experimentally measured transition temperatures. Consider the general case of N measured transition temperatures along with the N corresponding thermodynamic equilibrium expressions and the value of ΔG_{MM} for each of those transitions measured from the original molecular mechanics force field. The enthalpic shift for each system, i , which is necessary to satisfy all known transition points, can be computed as a system of linear equations in the form $A\vec{x} = b$:

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} \\ x_{21} & x_{22} & \cdots & x_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ x_{i2} & x_{i2} & \cdots & x_{ij} \end{bmatrix} \begin{bmatrix} \Delta H_1 \\ \Delta H_2 \\ \vdots \\ \Delta H_i \end{bmatrix} = \begin{bmatrix} \Delta G_{\text{MM},1} \\ \Delta G_{\text{MM},2} \\ \cdots \\ \Delta G_{\text{MM},i} \end{bmatrix} \quad (22)$$

where each matrix element A_{ij} is the coefficient x_{ij} for system j in equilibrium expression i and the b -vector contains the values $\Delta G_{\text{MM}, i}$ for each transition expression computed in the unmodified molecular mechanics potential. Coefficients for systems that are not present in the equilibrium expression are set to 0. The resulting solution for $\vec{x}$ contains the appropriate shift to apply to each system to minimize the residual to the experimental measurements. Because the thermodynamic equilibria produce *relative* free energies, there are numerous equivalent solutions to the ΔH -vector. More generally, for a given set of N measured transition temperatures, consisting of S systems, and coefficient matrix A with rank R , there are $(S - R)$ systems that have an arbitrary shift, and these shifts are set to 0 by adding $\Delta G_{\text{MM}, i} = 0$ and a trivial row to the A matrix:

$$x_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (23)$$

Extracting Free Energy Curves from Experimental Data. The computed relative enthalpy, entropy, and Gibbs free energy between systems in this work are compared against the corresponding experimental measurements. It is rare to directly measure the Gibbs free energy of competing phases as a function of temperature. However, the transition temperature (T_{trans}) is frequently measured experimentally, which corresponds with $\Delta G = 0$. Additionally, when the transition temperature is determined through differential scanning calorimetry (DSC), the enthalpy of transition (ΔH_{trans}) can also be estimated. The entropy of transition can then be computed using $\Delta S_{\text{trans}} = \Delta H_{\text{trans}}/T_{\text{trans}}$. Finally, the relative free energy near the transition temperature can be approximated using a Taylor series approximation:

$$\begin{aligned} \Delta G(T) &= \Delta G(T_{\text{trans}}) + \left(\frac{d\Delta G}{dT} \right)_p (T - T_{\text{trans}}) + O(3) \dots \\ &\approx \Delta G(T_{\text{trans}}) - \Delta S_{\text{trans}}(T - T_{\text{trans}}) \end{aligned} \quad (24)$$

For the anhydrous–hydrate systems, the transition temperature is no longer a single value and is instead dependent on the ambient water activity. For a given temperature, the water activity at which the hydrate and anhydrous systems are equally stable is called the critical water activity, a_w . In the three systems in this work, the critical water activity has been measured at multiple temperatures of interest. This

function essentially corresponds with the water activity at a constant free energy ($\Delta G = 0$). The temperature-dependent Gibbs free energy difference for a given constant water activity can be approximated from the temperature-dependent critical water activity by¹²³

$$\Delta G(T) = nk_B T \ln \left| \frac{a_w(T)}{a_w^*} \right| \quad (25)$$

where $a_w(T)$ is the measured critical water activity, a_w^* is the desired reference water activity, $\Delta G(T)$ is the molar Gibbs free energy of the hydrate relative to the anhydrous form, and n is the stoichiometric amount of water in the hydrate relative to the number of the pharmaceutical compounds. The transition entropy and enthalpy can then be computed from the free energy using $\Delta S = \left(\frac{d\Delta G}{dT} \right)_p$ and $\Delta H = T\Delta S$. The measured transition enthalpy, entropy, and Gibbs free energy for the compounds in this work computed in this manner are listed in Tables 1–4.

Table 1. Experimental Transition Thermodynamics for the Crystal Structures of Carbamazepine

	form III–form I ⁹²	form I–liquid ⁹³	dihydrate–form III ⁵⁰
transition temperature (°C)	81	194	27 at 64% RH
transition enthalpy (kcal/mol)	0.70	6.10	3.34
transition entropy (cal/mol-K)	1.98	13.08	11.22

Table 2. Experimental Transition Thermodynamics for the Crystal Structures of Theophylline

	form II–form I ⁹⁴	form I–liquid ⁹⁴	monohydrate–form II ⁵⁰
transition temperature (°C)	264	273	27 at 60% RH
transition enthalpy (kcal/mol)	0.48	7.02	1.79
transition entropy (cal/mol-K)	0.89	12.84	5.99

Table 3. Experimental Transition Thermodynamics for the Crystal Structures of Gandotinib

	form I–liquid ⁶⁰	tetrahydrate–form I ⁶⁰
transition temperature (°C)	235	27 at 48% RH
transition enthalpy (kcal/mol)	13.26	12.22
transition entropy (cal/mol-K)	26.12	41.02

Table 4. Experimental Transition Thermodynamics for Water

	ice I _h –liquid ¹⁵⁵	liquid–vapor ¹⁵⁵
transition temperature (°C)	0	100
transition enthalpy (kcal/mol)	1.44	9.72
transition entropy (cal/mol-K)	5.26	26.05

The predicted enthalpies, entropies, and Gibbs free energies for the transitions in this work are compared against the experimental values derived above. The actual predicted transition temperature from simulation is *highly* sensitive to small changes in the enthalpy and free energy.¹²⁴ In some cases, the free energies derived from the simulation are sufficiently erroneous that enantiotropically related polymorphs are predicted to have a monotropic relationship and do not have any predicted transition at any temperature. Therefore, to facilitate a uniform comparison between theory and experiments for all systems, we compare the predicted and measured Gibbs free

energy at the experimental transition temperature ($\Delta G_{\text{pred}}(T_{\text{trans, exp}})$ vs $\Delta G_{\text{exp}}(T_{\text{trans, exp}}) = 0$) rather than directly comparing the predicted and measured transition temperature.

The predicted relative Gibbs free energy at the experimental transition temperature is composed of both an enthalpic and entropic contribution, and it is useful to assess the accuracy of each of these two terms individually to determine the more significant source of error and variability across models. Therefore, we also compare the predicted relative enthalpy and entropy at the experimental transition temperature to the measured enthalpy and entropy of transition. The enthalpy and free energy have the same units, and the model residuals for these two terms can be plotted on the same scale. However, the entropy and enthalpy have different units, and the magnitude of the residuals for these two terms cannot be directly compared. Therefore, we elect to compare the combined temperature-entropy term $T\Delta S$ at the transition temperature rather than ΔS for the purposes of a more direct and meaningful comparison of residuals in the enthalpic and entropic contributions in each model. For simplicity, all mentions of relative entropy in the remainder of the study refer to the combined term $T\Delta S$ unless otherwise noted.

SIMULATION DETAILS

Initial Structures and Force-Field Parameters. Initial structures for all simulations were obtained from the Cambridge Structural Database (CSD). The crystal structure unit cells were expanded into supercells so that all forms had the same number of API molecules, and each principal axis was greater than twice the nonbonded interaction cutoff of 0.8 nm. CSD reference codes and supercell expansion ratios for all systems are shown in Section 1 of the Supporting Information.

Force-field parameters for the general Amber force field (GAFF) were generated with the Antechamber module of Amber,¹²⁵ with partial charges fit to the electrostatic potential generated with HF/6-31G(d) by RESP.¹²⁶ The charges were calculated according to the Merz–Singh–Kollman scheme using Gaussian 09.¹²⁷ Parameters for all molecules in OPLS were generated using LigParGen.^{128,129} AMOEBA parameters for all molecules were generated using the procedure described in a previous work.⁸⁰ We used the TIP3P,⁷⁷ TIP4Pew,⁹¹ and AMOEBA⁷⁸ water models.

Alchemical Conversion into an Einstein Crystal. The free energy for all systems, relative to the analytical Einstein crystal reference state, was computed using free energy perturbation (FEP). Free energy estimates for all FEP steps were computed from the simulation data using MBAR.¹¹⁵ The alchemical conversion of physical systems into the Einstein crystal was performed in GROMACS version 2020.4. A detailed description of how to implement the ECM in GROMACS, including treatment of the frozen reference atom, can be found in Aragonés *et al.*³⁴ Additionally, example input files necessary to run all steps of the ECM for the three chemical systems in the vacuum, liquid, and crystal phases are included in the Supporting Information.

The alchemical transformations for periodic crystals are split into six steps, similar to the process described in Yang *et al.*¹⁹ In the first step, all nondisordered heavy atoms in the crystal are harmonically restrained to their equilibrium lattice positions with a force-constant $k_1 = 1000$ kJ/mol-nm². In the second step, all nonbonded interactions in the crystal are removed from the system, and the charges in the system are fully decoupled prior to decoupling the van der Waals (vdW) interactions. This nonbonded decoupling step is connected through Hamiltonian replica exchange, with an exchange attempt frequency of 5 ps^{−1}, to facilitate the sampling of symmetrical/disorder positions in all states. In the third step,

all dihedral interactions are removed from the system. In the fourth step, all remaining unrestrained atoms are restrained to a restraint value of k_1 . In the fifth step, all atoms are further restrained from k_1 to $k_2 = 200,000$ kJ/mol-nm². In the sixth step, all remaining bonded interactions are removed, yielding an ideal Einstein crystal. The total Helmholtz free energy associated with this transformation from the physical state to the Einstein crystal state is computed as the sum of the individual perturbations:

$$\Delta A_{\text{ECM}} = \Delta A_{\text{restraint1}} + \Delta A_{\text{nonbonded}} + \Delta A_{\text{dihedral}} + \Delta A_{\text{restraint2}} + \Delta A_{\text{restraint3}} + \Delta A_{\text{bonded}} \quad (26)$$

For vacuum phase simulations, the nonbonded interactions are removed in the first step without applying restraints. In the second step, the dihedral interactions are removed. In the third step, all atoms are harmonically restrained to k_2 . In the final step, all remaining bonded interactions are removed to yield an ideal Einstein crystal. For liquid phase systems, all nonbonded interactions are removed without restraints, and then the noninteracting (ideal gas) compounds are all converted into an ideal Einstein crystal as described above for vacuum compounds.

The lambda scheduling function employed for each of the various perturbation types described above is summarized in Table S11a of the Supporting Information. A full justification for all scheduling functions, as well as the formula used to determine the total number of lambda windows for a given system and step, is provided in Supporting Information Section 11 for each perturbation type in Table S11a. The nonbonded interactions are decoupled in this work using a soft-core potential¹³⁰ as implemented in GROMACS with $\alpha_{\text{SC}} = 0.5$ and $\sigma_{\text{SC}} = 0.3$ and using a linear lambda-scheduling function with the charges being completely removed prior to decoupling the van der Waals interactions.

Because Coulombic and van der Waals interactions are decoupled separately, an additional hyperparameter arises in terms of the fraction of the lambda windows that are dedicated to the decoupling of each of the two interaction types. The optimal partition is system-dependent and difficult to set *a priori*. Furthermore, in our experience, the soft-core decoupling scheme with linear spacing experiences disproportionately poor overlap in the first few windows near the fully decoupled state. To alleviate the above issues, we dynamically adjust the lambda spacing between the equilibration and production runs to ensure an effective distribution of states to the decoupling of both Coulombic and van der Waals interactions to maximize the tri-diagonal of the overlap matrix. Complete details of this dynamic optimization scheme are provided in Supporting Information Section 11.

Briefly, normal distribution functions are fitted to each row of the overlap matrix generated from MBAR during the equilibration run. Then, the fitted standard deviations of the overlap at each state are interpolated with a spline fit to generate a single function characterizing the overlap between any two hypothetical states along the alchemical coordinate $O(\lambda_1, \lambda_2)$. Finally, we self-consistently solve for a new lambda vector $\lambda = [\lambda_0, \lambda_1, \dots, \lambda_L]$ such that the vector contains exactly $L + 1$ points monotonically spanning the interval $[0, 1]$ where the overlap between $\lambda[i]$ and $\lambda[i + 1]$ is constant for all $i = 0 \dots L - 1$. This refinement procedure can, in principle, be applied iteratively on another equilibration run and repeated until the lambda vector converges within a sufficient tolerance.

However, in our experience, the scheme produces an acceptable vector after a single iteration. As a proof of concept, overlap matrices are provided in Supporting Information Section 11 for decoupling the interactions in a box of liquid carbamazepine for both the default linear spacing and the adjusted spacing after a single iteration of optimization. A Python script for performing this lambda schedule optimization is also included in the Supporting Information.

Physical Sampling across a Range of Temperatures.

The NPT simulations across the temperature range of interest were run for 1 ns of equilibration followed by 1 ns of equilibration with replica exchange and 5 ns of production simulation. The nonbonded cutoff distance was set to 0.8 nm. Long-range Coulombic and dispersion interactions were calculated using particle mesh Ewald summation.^{131,132} Temperature coupling was achieved via the stochastic integrator¹³³ (Langevin dynamics) as implemented in GROMACS with a time constant $\tau_t = 0.2$ ps⁻¹. Pressure coupling in NPT simulations was achieved using the isotropic Parrinello–Rahman barostat¹³⁴ with $\tau_p = 1$ ps⁻¹ for liquid simulations and the anisotropic Parrinello–Rahman barostat with $\tau_p = 100$ ps⁻¹ for anisotropic crystalline simulations. The temperature scans were connected through temperature replica exchange with an exchange frequency of 5 ps⁻¹ between neighboring states. The simulations were also annealed during equilibration, from the maximum temperature sampled by the system down to the target temperature, to ensure complete equilibration.

Liquid phase and vacuum phase systems were simulated from $T_{\text{min}} = 200$ K to $T_{\text{max}} = 800$ K, while solid phase systems were simulated from 100 to 600 K, except hydrates and ice I_h , which were simulated up to 400 and 300 K, respectively, due to physical instability at higher temperatures. The total number of temperature states, as well as their values along the interval $[T_{\text{min}}, T_{\text{max}}]$, was chosen to optimize the overlap and exchange frequency between adjacent states (see Supporting Information Section 12 for more details). Thermochemical properties, for all systems at all other temperatures of interest within the interval $[T_{\text{min}}, T_{\text{max}}]$, were interpolated from the simulated states using cubic splines as implemented in SciPy version 1.3.2.¹³⁵ The free energies for systems across the temperature range were computed with eq 7, using two independent ECM calculations at two different temperatures to empirically estimate the uncertainty in the free energies with eq 9. A table of all systems, along with their simulated temperature range and temperatures for the ECM calculation, is included in Supporting Information Section 2.

The NPT AMOEBA simulations across the temperature range were performed using ForceFieldX and were run for 0.1 ns of equilibration followed by 1 ns of production. Coulombic interactions were treated with a 0.8 nm cutoff for real-space Ewald interactions, and the particle mesh Ewald algorithm was employed for periodic reciprocal-space interactions. Van der Waals interactions were treated with a 1.5 nm cutoff. Temperature coupling was achieved with a stochastic integrator, and pressure coupling was achieved using a Monte Carlo barostat.¹³⁶ Dual force-field (DFF) calculations for vacuum compounds were run with an initial histogram bias height of 0.005 kcal/mol, and DFF calculations for crystals and liquids were run with initial bias heights of 0.2 kcal/mol. Biasing potential heights were tempered during simulation to ensure smooth convergence using a tempering parameter $\Delta T = 8k_B T$ as described in Dama *et al.*¹³⁷

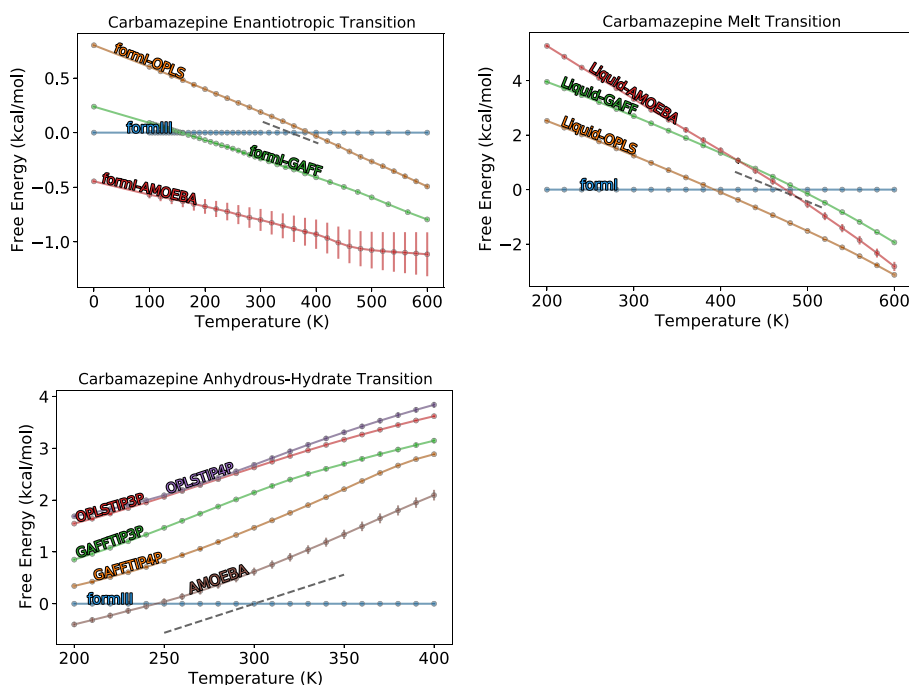


Figure 5. Computed relative chemical potentials for carbamazepine in all simulated potentials: (a) the enantiotropic transition of form I and form III, (b) the melt transition of form I, and (c) the dehydration of carbamazepine dihydrate into form III at 64% relative humidity. The approximate experimental curve in all transitions is shown as a dashed line. The relative hydrate free energy curves in all potentials are essentially parallel to each other and to the experimental approximation, indicating a consistent estimation of the dehydration entropy.

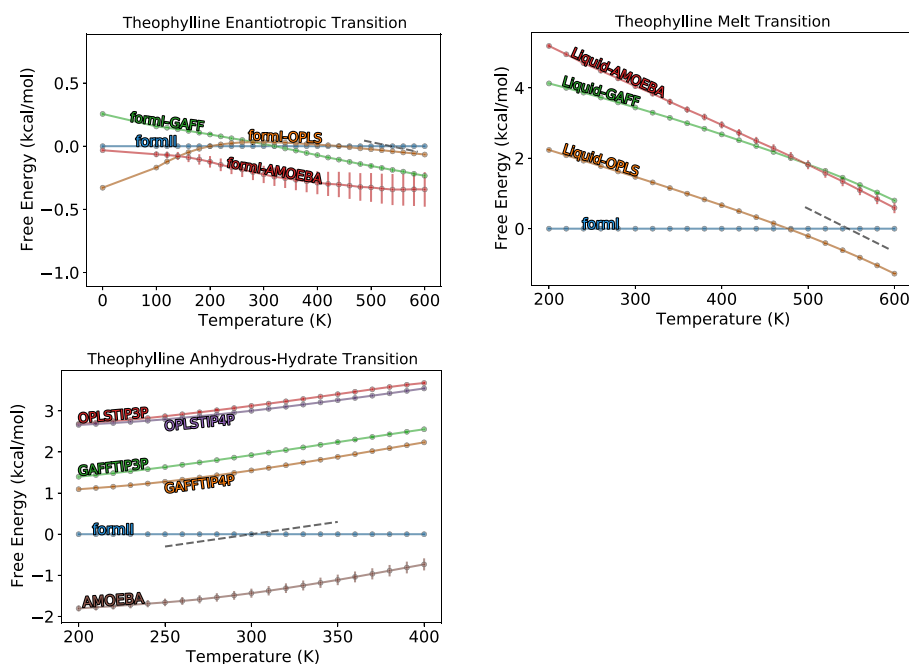


Figure 6. Computed relative chemical potentials for theophylline in all simulated potentials for (a) the enantiotropic transition of form I and form II, (b) the melt transition of form I, and (c) the dehydration of theophylline monohydrate into form II at 60% relative humidity. The approximate experimental curve in all transitions is shown as a dashed line. The relative hydrate free energy curves in all potentials are essentially parallel to each other and to the experimental approximation, indicating a consistent estimation of the dehydration entropy.

A complete set of GROMACS and ForceFieldX simulation input files is included in the [Supporting Information](#) for each of the above steps described above for all three molecular systems. The input files include starting structures, topologies, and the MD simulation parameters necessary to facilitate all the free energy perturbations described in this work. The total computational expense necessary to generate a phase diagram

like the ones shown in this work will depend on numerous factors including the size of the system, the short-range interaction cutoff, and the simulation hardware among other factors. Nevertheless, to provide a rough approximation of the cost, we list the total computational expense of each thermodynamic step of the workflow in [Supporting Information Section 11](#) using carbamazepine dihydrate as a

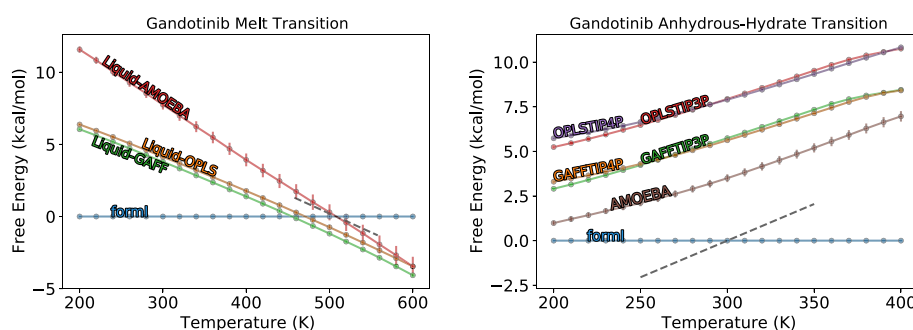


Figure 7. Computed relative chemical potentials for gandotinib in all simulated potentials for (a) the melt transition of form I and (b) the dehydration of gandotinib tetrahydrate into form I at 48% relative humidity. The approximate experimental curve in all transitions is shown as a dashed line. The relative hydrate free energy curves in all potentials are essentially parallel to each other and to the experimental approximation, indicating a consistent estimation of the dehydration entropy.

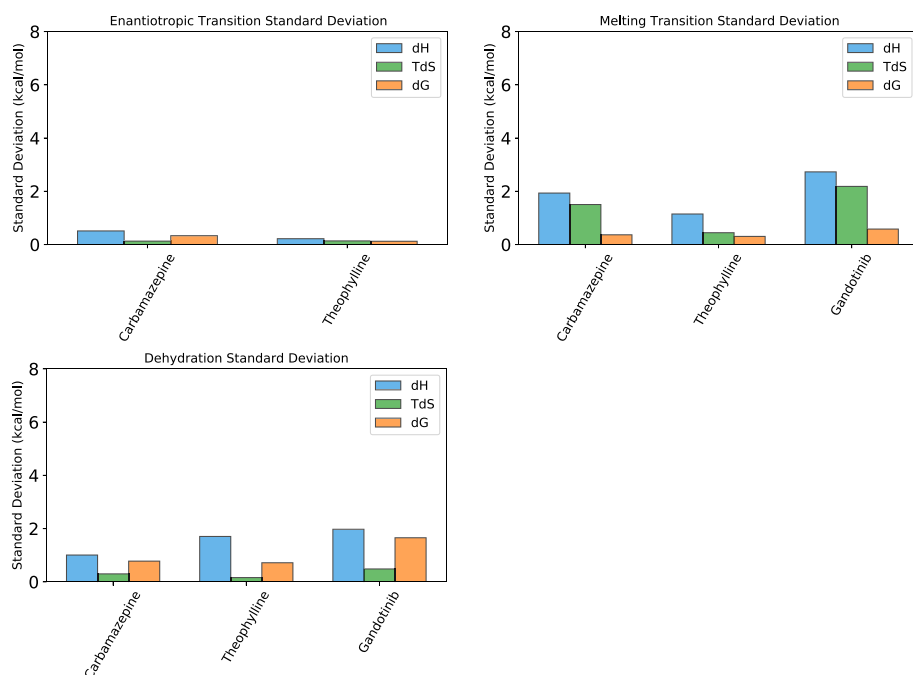


Figure 8. Standard deviations of the predicted transition enthalpy, entropy, and Gibbs free energy across all simulated potentials for (a) enantiotropic transitions, (b) melt transitions, and (c) anhydrous–hydrate transitions. The variance in the entropy across potentials is smaller than the variance in the enthalpy in all cases.

prototypical system. In the point-charge potentials, computing the temperature-dependent free energy of a single form requires approximately 4,000 CPU-hours of simulation, and in AMOEBA, it requires approximately 22,000 CPU-hours. Thus, constructing a phase diagram consisting of two forms would take approximately 8,000 and 44,000 CPU-hours for GAFF/OPLS and AMOEBA, respectively.

RESULTS AND DISCUSSION

Computed Free Energies and the Relative Effect of API Potential. The computed chemical potentials, as a function of temperature, for the crystal structures and liquid phases of carbamazepine, theophylline, and gandotinib, are shown in Figures 5–7 for all simulation potentials examined. In all plots, a single phase is chosen as the reference phase ($G(T) \equiv 0, \forall T$), and the predicted relative free energy of the second phase is displayed for each of the different simulation potentials. Positive values indicate that the reference phase is predicted to be thermodynamically more stable at that

temperature, and negative values indicate that the reference phase is predicted to be less stable. The temperature at which two phases' respective traces intersect indicates a predicted transition temperature. The experimentally measured transition temperature and the approximate temperature-dependent free energy near the transition, computed with eq 24 and eq 25, is represented by a dashed line in each plot. The reference phase for the plots of enantiotropic transition, melt transition, and anhydrous–hydrate transition is taken to be the experimentally stable form at low temperature, the solid phase, and the anhydrous crystal, respectively.

In Figures 5–7, we observe that the three simulation potentials, GAFF, OPLS, and AMOEBA, produce roughly parallel free energy traces within each system and transition type. The similar slopes in the free energy traces indicate that the predicted temperature dependence of the relative free energy (*i.e.*, the relative entropy) is quite similar regardless of the simulation potential. By contrast, the relative enthalpy has

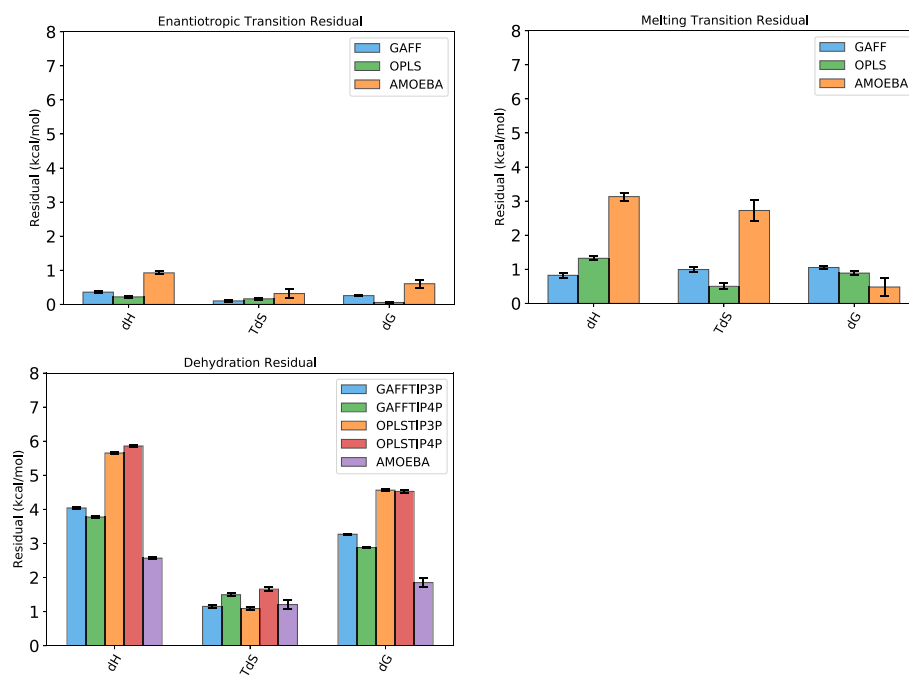


Figure 9. Residual of computed enthalpies, entropies, and Gibbs free energies of transition relative to experimental measurements for each potential averaged across the three systems examined in this work for (a) enantiotropic transitions, (b) melt transitions, and (c) anhydrous–hydrate transitions.

a large variation across potentials leading to the observed difference in the overall free energy.

To quantify the sensitivity of the simulation potential on the various thermodynamic observables, the standard deviations in the predicted relative enthalpy, entropy, and free energy across all simulated potentials are shown in Figure 8 for each system and transition type. In all cases, the standard deviation of the relative entropy is lower than the standard deviation in the relative enthalpy. These results are consistent with previous results, where it was observed that the computed entropy in crystalline systems is less sensitive to the choice of force field than the corresponding enthalpy differences.³⁶ In essence, this would be consistent with the notion that the potentials have more pronounced differences about the exact energetic favorability of the phases near their local energy minima (the depth of the minima). However, there is more unanimous agreement across potentials on how much freedom the particles in the system have to fluctuate away from their most energetically favorable position (the breadth of the minima). It is worth noting that the short 5 ns simulations in this work do not provide an exhaustive search of the crystallographic configuration space and only explore the local basin around the starting structure. It is always possible that longer simulations would find new areas of the minima that have more pronounced differences between the potentials and thus a greater difference in the entropy.

The computed free energy traces for each potential in Figures 5–7 are generally parallel to the experimental lines in addition to being parallel to each other. In Figure 9, the residual between the computed and experimental values for the transformation enthalpy, entropy, and free energy is given for each potential and each transformation type. Overall, the residuals for the anhydrous polymorph–polymorph transformations are the lowest of the three transformation types. This is likely because the cancellation of errors will be best in this type of transformation, where both systems are in the solid

state and have identical stoichiometries. The standard deviations of the predicted transition thermodynamics across the various potentials (Figure 8) are also smaller for the purely anhydrous polymorph transformations than for the hydrated and liquid transitions, further supporting the error cancellation hypothesis. The anhydrous–hydrate transitions have the largest residuals to experiments, while the liquid transformation residuals generally fall between those of the anhydrous–hydrate and anhydrous polymorph–polymorph transitions.

When comparing the performance of the different potentials, we examine the predicted relative free energy at the experimental transition temperature in each potential. At the experimental transition temperature, the two phases are in equilibrium ($\Delta G = 0$), and thus, the absolute value of the computed relative free energy at this temperature provides a simple measure of the accuracy of the potential. The two point-charge potentials, GAFF and OPLS, produce virtually identical residuals with respect to the experimental values (Figure 9) for all three transition types. Surprisingly, in the anhydrous polymorph transitions, the polarizable AMOEBA force field performs slightly worse than the general point-charge GAFF and OPLS potentials (0.61 ± 0.12 kcal/mol versus 0.16 ± 0.01 kcal/mol) based on the computed relative free energy at the transition temperature. For the solid-to-liquid melt transitions, the AMOEBA potential was found to have the smallest average residual with respect to experiment for the free energy (0.48 ± 0.28 kcal/mol versus 0.97 ± 0.05 kcal/mol for GAFF and OPLS), and therefore, AMOEBA produces the closest predicted melting points among the three potentials. However, the residuals for the enthalpy and entropy of transition in AMOEBA are both significantly worse when compared to the point-charge GAFF and OPLS force fields, suggesting that the favorable melting point predictions in AMOEBA may be a result of fortuitous enthalpy–entropy compensation in these systems.

In the anhydrous–hydrate transitions AMOEBA consistently outperforms the point-charge GAFF and OPLS force fields when predicting both the enthalpy and Gibbs free energy of transformation (Figure 9). At room temperature and the experimental transition humidity, the mean absolute AMOEBA free energy is 1.85 ± 0.14 kcal/mol compared with 3.81 ± 0.03 kcal/mol for the point-charge potentials when averaged across the four possible model combinations. In addition to having lower residuals, the AMOEBA potential for anhydrous–hydrate transitions also produces free energies more evenly distributed around the experimental values, in contrast with the point-charge potentials, which exhibit systematic bias. For the anhydrous and melt transitions in Figures 5–7, the experimental transition falls somewhere between the predictions of GAFF and OPLS in the majority of transitions. This suggests that the residuals to experiments are a product of standard numerical variations in the parametrization schemes for these force fields rather than a systematic error of the model itself. In contrast, the stabilities of the hydrated crystals relative to the anhydrous forms are systematically underpredicted by both GAFF and OPLS in all three systems examined. With AMOEBA, not only are the predicted stabilities of the hydrates closer to the true value in all three systems, but also the residuals to experiments are more evenly distributed around zero, with carbamazepine and gandotinib hydrates having slight underpredictions and the theophylline hydrate stability being overpredicted. This again suggests that there is a more systematic effect in the anhydrous–hydrate transitions that is better captured with AMOEBA than the point-charge GAFF and OPLS force fields. However, more studies should be performed on additional systems to probe the generalizability of these observations.

The lower free energy residuals for AMOEBA in the anhydrous–hydrate transitions occur even though AMOEBA has noticeably higher residuals in the anhydrous polymorphic transitions (Figure 9a). The lower residuals with AMOEBA for the anhydrous–hydrate transitions combined with the relatively higher residuals in the anhydrous polymorphic transitions for the same compounds suggest that the improved accuracy of the former results specifically from the polarizable treatment of the water molecules within the hydrated crystals. When an anhydrous structure is converted into a hydrated crystal, the water molecules are taken out of an aqueous liquid environment and transferred to an environment of primarily organic small molecules. In the point-charge potentials with the fixed-charge water models, the water molecules are constrained to adopt the same overpolarized charges in both the liquid and crystalline state. By contrast, in AMOEBA, the water is polarizable and can adapt to satisfy both the aqueous and crystalline environments independently. The lower enthalpy and free energy residuals in AMOEBA suggest that the energy cost to repolarize the water molecules represents a significant contribution to the overall anhydrous–hydrate transition and the polarizable AMOEBA model is particularly well suited to capture this change relative to the point-charge models. At the same time, the predicted entropy of dehydration is essentially equally accurate in all potentials and water models (Figure 9c), implying that polarization does not significantly change the temperature dependence of the relative stability even though the predicted relative enthalpy improves. This is in contrast with studies that evaluated the performance of polarizable force fields when applied to biological phenomena such as protein folding, where polar-

ization notably improved the accuracy of modeling temperature-dependent effects.^{138–140}

To further explore the polarization effects of the water molecules in the hydrates, we computed the average dipole moment of the water within each system over the full trajectory at each simulated temperature (Figure 10a). The

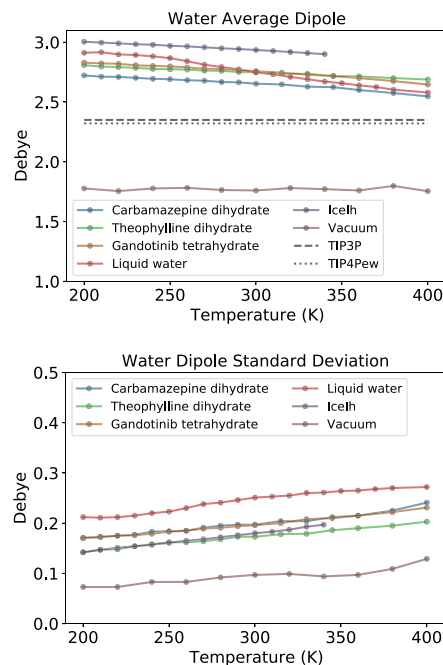


Figure 10. (a) Average and (b) standard deviation in the dipole moment for water within the solid, liquid, and vacuum systems in this work simulated with the AMOEBA potential. The static dipole moments for the TIP3P and TIP4Pew water models are also shown. The effects of temperature on the average polarity and fluctuation in the polarity of water in the three hydrated crystals are nearly identical to the effects in liquid water.

corresponding standard deviation in the water dipole moment for each system at each temperature is also reported in Figure 10b. The corresponding average and standard deviation of the dipole moment for the API molecules in each system are included in Supporting Information Section 13. The computed dipole moment for water in AMOEBA is consistent with previously reported values for the vacuum, liquid, and ice I_h states.⁷⁸ The average dipole for liquid water in AMOEBA at 300 K is also consistent with recent *ab initio* molecular dynamics simulations that estimated the value of 2.91 ± 0.28 D.¹⁴¹

At room temperature, the average dipole moment for crystalline water molecules is roughly 10% higher with AMOEBA (2.65–2.75 D) than with the fixed-charge TIP3P (2.35 D) and TIP4Pew (2.32 D) water models (Figure 10a). However, a somewhat surprising result is that the average dipole moment of the water molecules inside liquid water at room temperature (2.75 D) is roughly equivalent to that of the water in the hydrates, suggesting that there may be significant cancellation of errors when using the nonpolarizable water models to predict hydrate stability. The fluctuation in the dipole moment with liquid water in AMOEBA shows a more pronounced difference to the fluctuation of water in the three hydrates (Figure 11b), which seems more likely to account for some of the observed improvement in the predicted free

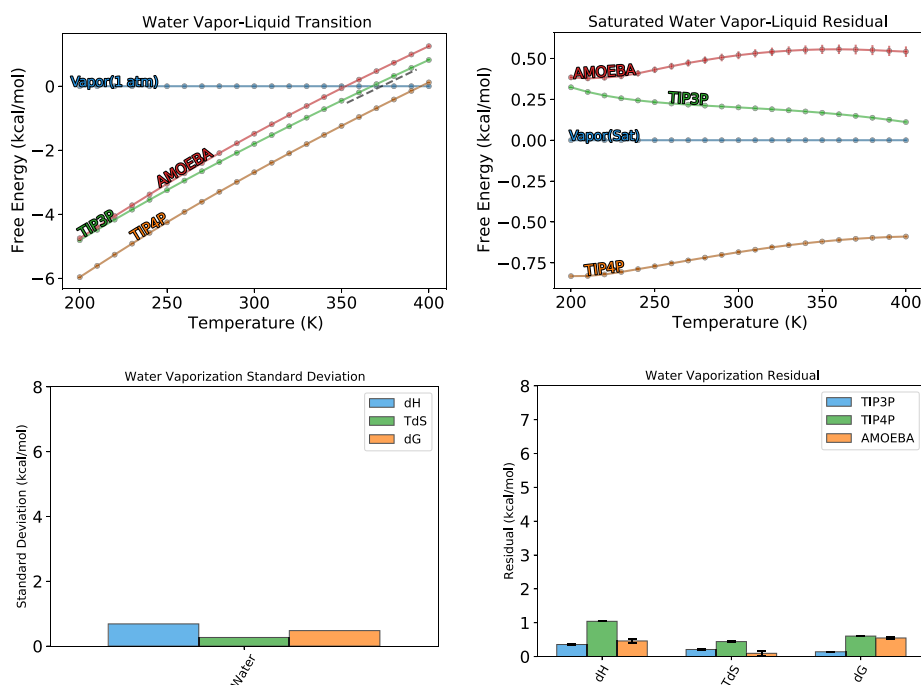


Figure 11. Computed chemical potentials for liquid water with different water models relative to an analytical vapor phase water at (a) 1 atm of pressure and (b) the saturated vapor pressure according to the Antoine equation. The approximate experimental chemical potential curve is shown as a dashed line using the measured boiling point and enthalpy of vaporization. The computed (c) standard deviation and (d) residuals to experiments across potentials for the enthalpy, entropy, and free energy of vaporization of water are also shown for the various water models examined in this work.

energy difference when comparing hydrates to anhydrous crystals using AMOEBA rather than nonpolarizable point-charge potentials (Figure 9).

Another clear result from Figure 10 is the remarkable similarity in the temperature dependence of the mean and standard deviation of the water dipoles across all the various systems. Specifically, as temperature increases, the average dipole moment of water decreases and the variance in the dipole moment increases, and the rate of change of both is nearly identical regardless of whether the water molecules are in the liquid state, in the solid ice I_h state, or contained within one of the hydrates. The result of this similar temperature dependence in the dipoles is that the effect of polarization will almost entirely cancel when comparing the stability of hydrates to anhydrous crystals, which at least partially explains why the predicted entropy of dehydration is similar regardless of whether a polarizable or nonpolarizable water model is being used (Figure 10).

Finally, we observe that for these anhydrous–hydrate pairs, the entropic contribution to the relative free energy ($T\Delta S_{\text{trans}} \approx 5.8$ kcal/mol) is at least an order of magnitude larger than that of the average anhydrous polymorph pair ($T\Delta S_{\text{trans}} \approx 0.1 - 0.5$ kcal/mol).⁴³ Indeed, this entropic contribution is similar in magnitude to the three solid–liquid melt transitions in this work (4.8 kcal/mol), which is intuitively consistent with the theoretical picture that the water molecules in the hydrate are transferred from a solid-like state into a liquid state upon dehydration.^{142,143} This large entropy change upon dehydration also underscores the fact that temperature and entropy are *not* negligible when comparing hydrated crystals to anhydrous forms, in contrast to comparisons of purely anhydrous polymorphs where thermal corrections are frequently small enough that zero Kelvin lattice energies are a sufficient

surrogate for the true ambient temperature stability. Furthermore, the results of this work suggest that molecular dynamics simulations are well suited to capture this large entropy difference between hydrated and anhydrous forms based on the low residuals to experiments for the dehydration entropy (Figure 9c). It remains an interesting open research question whether harmonic vibration approximations for hydrated crystals would also accurately capture this entropy difference. However, this question is outside the scope of the present study.

Given the residuals with respect to the experimentally measured enthalpy, entropy, and free energy of transition shown in Figure 9, it is natural to ask whether the actual transition temperature is predictable with these molecular mechanics models. Unfortunately, in all systems examined in this work, the answer appears to be “no”. This is due to the high sensitivity of the predicted transition temperature to both the entropy and enthalpy difference of the phases.¹²⁴ For example, in the case of the anhydrous polymorphic transition of carbamazepine form III to form I (Figure 5), the predicted ΔS in GAFF and OPLS at the experimental transition temperature are 2.2 ± 0.1 and 2.3 ± 0.1 cal mol^{−1} K^{−1}, respectively, which are both within 15% of the experimentally measured value of 1.98 cal mol^{−1} K^{−1} and arguably as good as could be expected from generalized force fields. However, the predicted transition temperature for this set of polymorphs in GAFF and OPLS is nevertheless starkly different at 155 and 385 K, respectively, due to the sensitivity of the transition temperature to the enthalpy and free energy difference.

For similar reasons, predicted values of melt transition and dehydration transition temperatures are also highly sensitive and difficult to predict. In the melt transition of carbamazepine, the estimated melting point differs by roughly 100 K

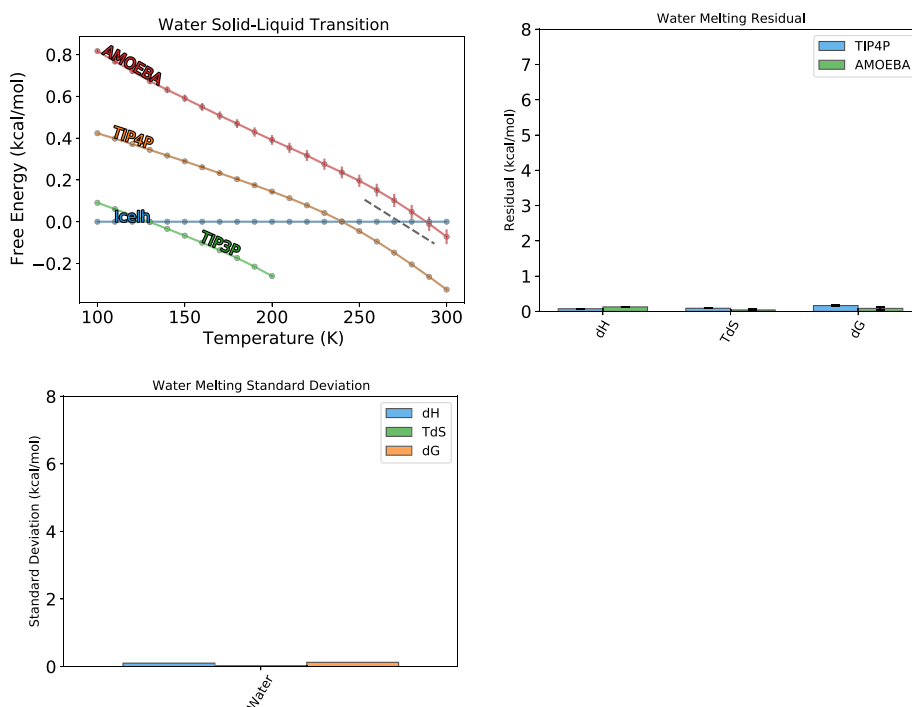


Figure 12. Computed (a) chemical potentials for ice I_h relative to liquid water for different water models. The approximate experimental chemical potential curve is shown as a dashed line using the measured melting point and enthalpy of fusion. The computed (b) residuals to experiments and (c) standard deviation for the enthalpy, entropy, and free energy of vaporization of water are also shown for TIP4P and AMOEBA.

between OPLS and GAFF (Figure 5b). The predicted dehydration temperature of carbamazepine dihydrate to form III also varies widely between the different force-field models. These results demonstrate that generalized molecular mechanics models are unlikely to be able to produce useful, fully theoretical predictions of a dehydration temperature between a specific hydrate and anhydrous form. Nevertheless, we observed that these generalized models perform quite well in estimating the entropy of dehydration when compared with experimental measurements (Figure 9c). Furthermore, the predicted dehydration entropy is not sensitive to the choice of potential (Figure 8c). Together, these two insights enable the opportunity to derive highly accurate anhydrous–hydrate phase diagrams when the molecular modeling results are combined with a single experimental data point, an idea that we will explore further later in this work.

Effect of Water Model and Water Reference State.

The computed free energies for the hydrated crystals relative to the anhydrous forms are impacted by the choice of model for the interstitial water molecules in addition to the choice of force field for the parent API, as described above. From Figures 5–7 we observe that the different water models TIP3P, TIP4Pew, and AMOEBA produce noticeably different (although roughly parallel) free energy curves. The residuals shown in Figure 9c indicate that the choice of TIP3P and TIP4P has a relatively small impact on the accuracy of the computed dehydration free energy. The AMOEBA model produces the lowest residuals among the three water models examined, although this potential also necessarily affects the parent API molecules in addition to the interstitial water, and this effect cannot be separated. The improved residuals with AMOEBA relative to the fixed-charge TIP3P and TIP4P models come from the enthalpic contributions, with relatively

less change in the predicted entropy, as mentioned previously (Figure 9c).

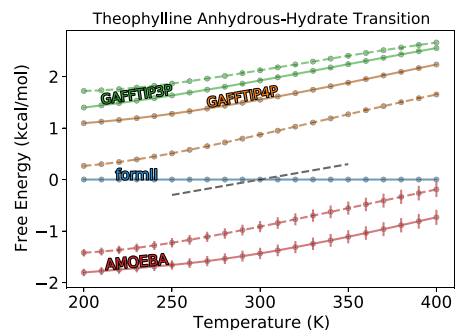


Figure 13. Computed chemical potentials for theophylline monohydrate relative to anhydrous form II in various potentials using a reference water chemical potential derived numerically from a liquid box (solid line) and derived empirically/analytically from water vapor and the Antoine equation (dashed line).

The reference state for excess water upon dehydration can be approximated using either a liquid box of water or saturated vapor water, and this choice has a downstream effect on the predicted stability of a hydrated and anhydrous crystal. Figure 11b shows the computed relative free energy difference between liquid water and vapor water set to a pressure equal to the vapor pressure of water at that temperature, as computed by the Antoine equation. This plot delineates the exact difference in free energy (per stoichiometric water) that will occur on the predicted anhydrous–hydrate free energy depending on whether the empirical/vapor water or liquid water is used as the reference. The impact that this difference of reference state water has on the overall predicted anhydrous–hydrate free energy curve is demonstrated in

Figure 14 for theophylline monohydrate relative to anhydrous form II for the various potentials and water models. Similar

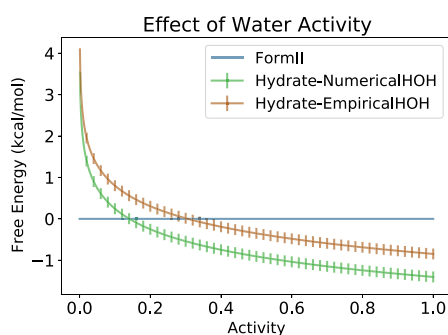


Figure 14. Computed chemical potential for theophylline monohydrate relative to anhydrous form II at $T = 360$ K as a function of the reference water activity in AMOEBA using a reference water chemical potential derived numerically from a liquid box and derived empirically/analytically with the Antoine equation. The monohydrate form becomes more stable relative to the anhydrous form as the water activity increases. The effect of water activity on the relative free energy of the hydrate is computed using eq 18 and is identical with both the numerical and empirical reference states.

plots for all systems and all potentials are included in Supporting Information Section 14, though it is useful to highlight that the (per water) numerical effect on the free energy of this reference state change from liquid to saturated vapor is a fixed quantity, which depends only on the water model and does not depend on the API or crystal structure.

In the previous section, we noted that the choice of API potential has a significantly smaller effect on the computed entropy of transition than on the enthalpy of transition. We observe here that this same trend occurs for water, where the water model and the choice of water reference state have a less pronounced impact on the relative entropy than on the relative enthalpy. Figure 12a shows the computed free energy difference between liquid water and water vapor held at 1 atm of pressure for the various water models used in this study. For the water melt transition, a rigorous assessment of the solid–liquid transition thermodynamics for the myriad of ice polymorphs in various different force fields is outside the scope of the present study and has been explored previously by other investigators.^{22,81,144–147} However, we show in Figure 12a the computed transition between ice I_h and liquid water for the three water models studied in this work. In both water transitions, the variation across potentials and the average residual to experiments are lower for the computed entropy of transition as compared to the computed enthalpy of transition. A similar effect is also observed for the choice of liquid versus saturated vapor for the excess water reference state (Figure 11b). At 300 K, the free energy penalty to add or remove water from a hydrate will depend on whether a liquid or saturated vapor reference state is being used, and the difference varies between -0.69 and 0.52 kcal/mol depending on the water model used. In contrast, the entropy penalty for the two reference states varies on the narrower range of -0.47 to 0.16 kcal/mol.

In summary, we find that the most significant factor influencing the predicted anhydrous–hydrate relative free energy curve is the choice of API potential. The choice of water model and the choice of liquid or vapor reference state have relatively minor effects. Figure 15 shows the change in the

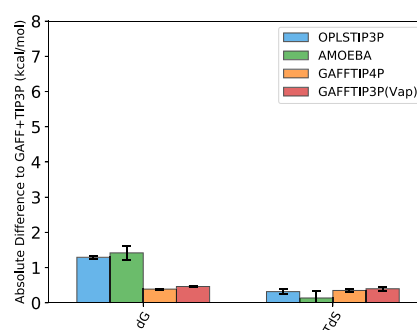


Figure 15. Relative effect of potential, water model, and water reference treatment on the predicted free energy and entropy of transition averaged across the three systems examined in this work. In all cases, the predicted transition values with the various models are shown relative to the predictions with GAFF and TIP3P to indicate the (unsigned) magnitude of the change for each model design choice. The choice of API potential has the largest overall effect on the predicted free energy of transition. The effect of API potential is smaller for the predicted relative entropy than for the predicted free energy.

predicted relative Gibbs free energy and relative entropy of transition when moving from GAFF and TIP3P and a liquid water reference state to various other possible models averaged across the three systems investigated. The impact of the API potential has the most significant effect, accounting for a roughly 1.5 kcal/mol change on average in the predicted free energy of the hydrate relative to the anhydrous form. Changing the water model from TIP3P to TIP4P and changing from a liquid to a vapor reference state both have a smaller average impact at roughly 0.5 kcal/mol. Although this effect of water treatment on the free energy is smaller, it is worth reiterating that a change of this size still has a significant impact on the predicted dehydration temperature because of the sharp sensitivity as can be seen in Figures 5–Figure 7 as well as Figure 13. However, we again note that the predicted entropy of dehydration is insensitive to all model combinations.

Predicted Phase Diagrams Using Room Temperature Critical Water Activity. The accurate prediction of the entropy difference between hydrated and anhydrous crystals opens the possibility for useful models of dehydration when combined with some experimental data. In Figures 5–7, we show that the exact Gibbs free energy differences between the hydrated and anhydrous crystals in this work vary significantly between the different possible potentials and water representations. This varying Gibbs free energy, in turn, translates into wildly different predictions for the dehydration temperature from the pure-theory chemical potential curves.

This large variation in the free energy and predicted dehydration temperature can be eliminated if the dehydration temperature from experiments is provided as an input to the model (or equivalently, if a measured critical water activity at a single temperature is provided). With this single experimental coexistence point as an input, the relative Gibbs free energy curves from our fully theoretical molecular dynamics simulations can be shifted by a constant value such that the free energy is exactly zero at the coexistence point regardless of the size of the initial residual from the model. An example of this constant “enthalpic shift” applied to a free energy curve is shown pictorially in Figure 16, and the procedure for computing the appropriate enthalpic shift is shown in eq 22.

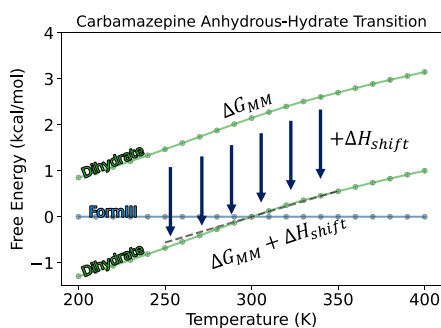


Figure 16. Illustration of applying an enthalpic shift to the predicted chemical potentials for carbamazepine dihydrate relative to form III to match experimental data at room temperature.

With the empirical enthalpic shift, the free energy at the coexistence temperature will be correct by construction. However, relative free energies at temperatures near this

coexistence point will also be reasonably accurate because they are determined by the relative entropy, as shown in the Taylor approximation in eq 24. The entropy of transition is observed to be both consistently less variable across potentials (Figure 8) and closer to the true experimental values (Figure 9) than the predicted Gibbs free energy of transition.

One potential application for these hybrid experimental–theoretical free energy curves is to construct temperature–humidity phase diagrams for hydrated and anhydrous crystals. The computed phase diagrams for the hydrated and anhydrous crystals of the compounds in this work are shown in Figure 17 using the experimentally measured critical water activity at room temperature to apply an enthalpic shift. At low temperatures and high humidities, the hydrated forms are favored, whereas at high temperatures and low humidities, the anhydrous forms are predicted to be more stable, consistent with experimental observations. In Figure 17a, the experimentally measured critical water activities at all measured temperatures are superimposed onto the phase diagrams

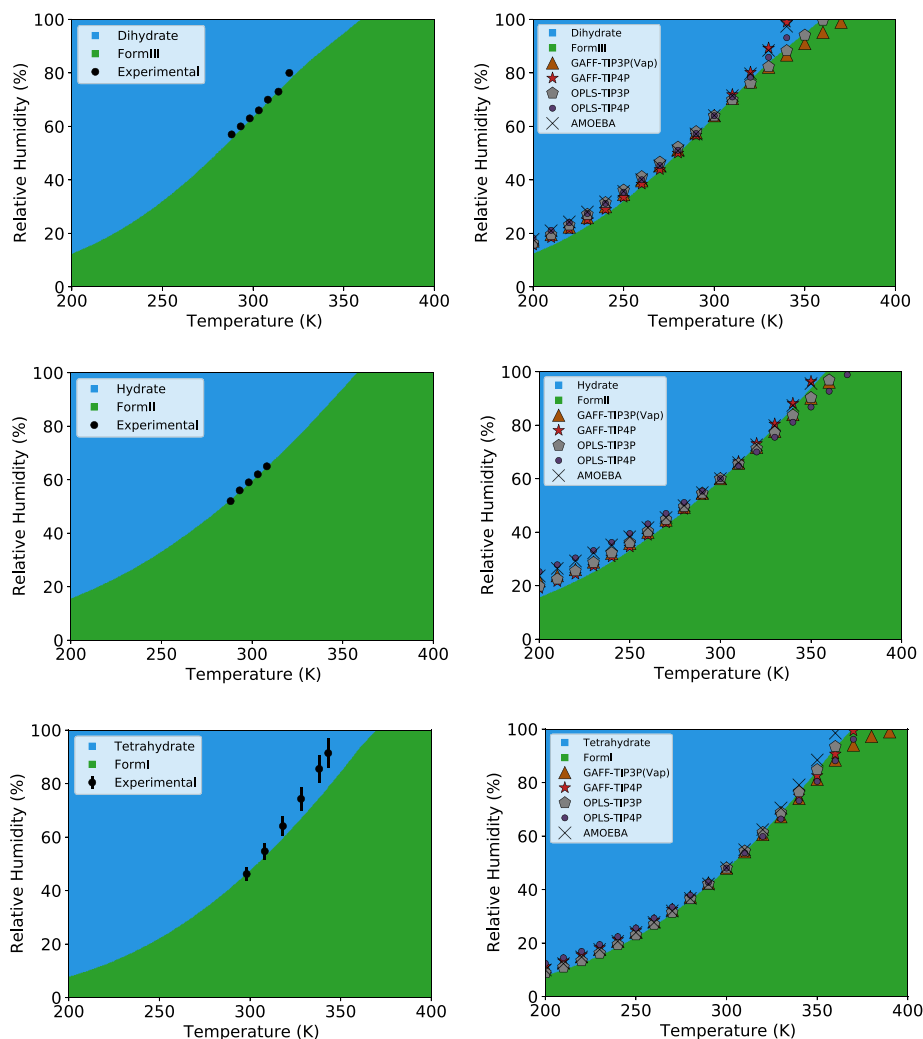


Figure 17. Computed phase diagrams for hydrated and anhydrous forms of carbamazepine (top), theophylline (middle), and gandotinib (bottom) over a range of temperatures and humidities using the GAFF potential and the TIP3P water model fitted to the experimentally measured coexistence point at room temperature. Data points are included on the phase diagram representing the coexistence points derived from (left) experimental measurements and (right) various other potentials and water models. Models with "(Vap)" indicate that the coexistence points were determined using water chemical potentials derived analytically in the vapor phase with the Antoine equation rather than numerically from a liquid box. In all systems, the predicted coexistence points (critical water activities) in the temperature range of -25 to 75 °C are within 10% regardless of the choice of potential, water model, or reference water treatment.

computed with GAFF and TIP3P. In all three systems, the predicted phase diagram has remarkable agreement with the experimental data points at all temperatures. The largest deviation between the experimental measurements and the model occurs for the tetrahydrated crystal of gandotinib relative to anhydrous form I at 75 °C, with a roughly 10% humidity deviation between the prediction (80%) and the experiments ($90 \pm 5\%$).

An important result that emerges from Figure 17 is that the predicted temperature–humidity phase diagrams are ultimately quite similar regardless of the choice of API potential, the water model, or the reference water treatment once the fit to the room temperature coexistence point has been applied. In Figure 17b, the phase boundaries predicted using various potentials and water treatments are superimposed onto the phase diagrams generated with GAFF and TIP3P. The computed boundaries are essentially equivalent over a wide range of temperatures, with all predicted coexistence points between -25 and 75 °C being within 10% humidity of each other. The similarity in this predicted phase boundary across different models is a direct result of the relatively small variation in the predicted entropy of transformation (Figure 8). This insensitivity of the phase diagram with respect to including or excluding polarization is likely due to the cancellation of errors in the temperature dependence of the average water dipole and fluctuation in the water dipole between water molecules in the hydrates and water molecules in the liquid phase (Figure 10). Overall, the results shown in Figure 17 indicate that the choice of potential and/or water representation is not especially significant when constructing accurate *in silico* phase diagrams if a single experimental coexistence point is known.

It is worth noting that this experimental coexistence correction is not limited to the comparison of only two phases. As an example of this, the quadruple phase diagram for carbamazepine form I, form III, dihydrate, and liquid is shown in Figure 18 using the room temperature critical water activity between form III and the dihydrate as well as the melting points of form I and form III.

One application for this type of diagram is in predicting the boundary between a hydrate and an anhydrous form at temperatures where the form is not the globally stable

anhydrous structure. In Figure 18, a dashed line is included to indicate the computed boundary between carbamazepine form I and the dihydrate at temperatures where form III is more stable than form I. No critical water activity data from form I are used in Figure 18, so this is effectively generating the phase boundary between a hydrated and metastable anhydrous form using only water activity data measured with the globally stable anhydrous polymorph. Indeed, experimentally determining the critical water activity of a metastable form can be quite difficult in practice if the metastable form rapidly converts to the globally stable anhydrous polymorph in the water–solvent slurries used for these measurements.¹⁴⁸ This challenge is particularly noteworthy for systems with enantiotropically related polymorphs, such as the example of carbamazepine shown here, where both polymorphs will be metastable in at least some of the temperature range where the critical water activity measurements are performed. In such cases, theoretical models like the ones shown in this work can aid in constructing the complete phase boundaries for all forms over all temperatures even when the forms are metastable.

Figure 18 also highlights another application for computed multiphase diagrams in providing an indication of the transition temperature between two enantiotropically related polymorphs when only melt transition data are available. Numerous experimental techniques exist to rigorously determine enantiotropic transition points, such as solubility extrapolation^{11,90,149,150} and competitive cascade slurries at different temperatures.^{19,151} However, the melting points of anhydrous polymorphs are more easily accessible quantities with minimal material requirements. We demonstrate here that these measured quantities could be combined with simulation data to give a preliminary assessment of the solid form crossover region that could be used to guide the future experimental search for the transition temperature.

Finally, although the focus of this work is exclusively on known crystal structures with previously measured coexistence data, there are numerous ways in which these same molecular models can assist with prospective crystal structure prediction as well. CSP methods have been previously applied to hydrates to predict the most stable hydrated form.^{48,152–154} However, this stability ranking is rarely combined with the corresponding anhydrous CSP (or hydrated CSPs of higher stoichiometry) to determine the *globally* stable form at a given temperature and humidity among a plethora of possible hydration states. The molecular simulations described in this work could be used to answer this important question.

In truly prospective CSP studies, where no hydrated forms have yet been observed and the risk of a late-stage hydrate is being actively evaluated, there will not be measured coexistence points to fit against. However, CSPs are generally performed with highly accurate QM evaluations of the lattice energy at 0 K. Assuming that the stability at 0 K is well approximated by these QM methods, an accurate prediction of the relative stability at ambient conditions between theoretical anhydrous and hydrated structures would almost entirely depend on whether the classical molecular dynamics model can accurately capture the *relative* effects of temperature and humidity. The reasonably accurate representation of temperature and humidity effects in the three pharmaceutically relevant systems in this work is a preliminary validation that these effects can be captured by classical molecular simulations, at least in some cases, and gives reason for optimism that the

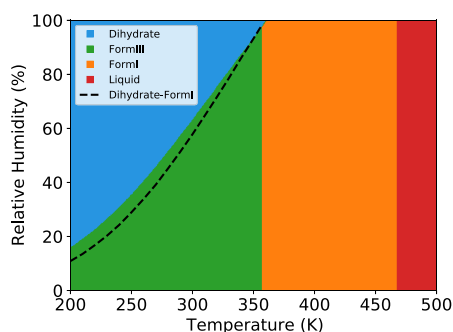


Figure 18. Computed phase diagram for carbamazepine dihydrate, form I, form III, and liquid using simulation results from OPLS +TIP3P along with the experimentally determined critical water activity of dihydrate–form III at room temperature and the melting points of form I and form III. The dashed line indicates the predicted phase boundary between the metastable form I and the dihydrate. The enantiotropic transition between form I and form III is also shown.

same temperature and humidity effects can be predicted on theoretical CSP structures in the future.

CONCLUSIONS

A series of classical free energy simulations were performed to systematically examine their ability to predict the transition thermodynamics of hydrated crystals relative to anhydrous forms in three pharmaceutically relevant small molecule systems. The simulations used two generic fixed-charge potentials, GAFF and OPLS, as well as the polarizable AMOEBA force field. Additionally, the TIP3P, TIP4Pew, and AMOEBA potentials described interstitial hydrate water. To model the chemical potential of water under various relative humidity conditions, we considered two different thermodynamic water baths: a simulated liquid system and an empirically derived vaporous phase. As a control, the nonwater transitions between anhydrous polymorphs and melting transitions for these three compounds were also computed. In all cases, the temperature-dependent free energies of all phases were computed with molecular dynamics, and the predicted enthalpy, entropy, and Gibbs free energy of transition were compared against the corresponding experimental coexistence point measurements.^{50,60,92–94}

Overall, the predicted anhydrous–hydrate transition temperatures are very sensitive to the force field, and realistically, no current molecular mechanics model can predict reliably accurate coexistence points to serve as a suitable experimental replacement. However, when the purely *in silico* results are anchored to a single experimental coexistence point, the computed free energies of the hydrated and anhydrous crystals are quite accurate. The resulting predictions can be used to construct phase diagrams that predict the coexistence line to within 10% relative humidity of experimental measurements across the range of 15 to 75 °C for the three systems in this work.

Additionally, this hybrid theoretical–experimental phase boundary is insensitive to the choice of force field, water model, or reference humidity representation, with all models agreeing on the boundary to within 10% relative humidity on the range of –25 to 75 °C for all three molecular systems. The ability to predict accurate phase diagrams regardless of the model appears to result from the cancellation of errors in the entropy of transition. A systematic comparison of the predicted transition entropy and enthalpy estimates revealed that the variation in the entropy term ($T\Delta S$) between potentials is indeed smaller than the corresponding variation in the predicted enthalpy in all cases. An analysis of the calculated water dipole moments with AMOEBA revealed that the water molecules in the hydrates have roughly the same dipole moment and temperature dependence as those in liquid water, leading to the fortuitous cancellation of errors when using a fixed-charge water model to estimate the entropy of hydrates.

When no experimental coexistence point is used to anchor the results, the most significant factor influencing the predicted anhydrous–hydrate free energy is the API potential, where the computed stability spans roughly 3 kcal/mol on average depending on the potential used. The choices for water representation between TIP3P and TIP4P as well as the representation of ambient water have comparatively smaller impact on the predicted free energy.

When these pure-theory results are compared to experimental measurements, the AMOEBA force field predicts relative Gibbs free energies with the lowest mean absolute

residuals to experiments for the three anhydrous–hydrate pairs examined in this work (~2 kcal/mol). This improved accuracy comes almost entirely from the predicted enthalpy of transition, whereas the residuals for the relative entropy are essentially the same across all the three potentials. The average dipole moment of water in the hydrates in AMOEBA is similar to liquid water, as mentioned earlier. However, the dipole fluctuation in liquid water is slightly higher than in the organic hydrates, which could account for some of the improvement when using AMOEBA to predict relative hydrate stability from pure theory. Despite this improvement, the results with AMOEBA must still be anchored to an experimental coexistence point to produce useful anhydrous–hydrate phase diagrams.

The anchored phase diagrams presented in this work can be extended to a comparison of more than two phases of interest. Several practical applications for these multiphase diagrams have been demonstrated, including generating anhydrous–hydrate phase boundaries for metastable anhydrous forms, as well as estimating the transition temperature for enantiotropically related polymorphs. The ability to capture relative temperature and humidity effects with molecular simulation, as shown in this work, opens potential new avenues for prospective structure prediction, such as combining theoretical structures of different hydration states together into a unified solid form landscape across temperature and humidity.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.cgd.2c00832>.

CSD reference codes and supercells for all crystals, additional derivations and implementation details of the computational methodology, gas–liquid relative free energy traces, thermalized dipole moments of the API molecules in various condensed-phase systems, and additional comparisons of the liquid and vapor water reference states (PDF)

Python script for executing the lambda schedule optimization routine and a complete set of input files for all GROMACS and ForceFieldX simulations (ZIP)

AUTHOR INFORMATION

Corresponding Author

Eric C. Dybeck – Pfizer Worldwide R&D, Groton, Connecticut 06340, United States; orcid.org/0000-0003-4058-9756; Email: eric.dybeck@gmail.com

Authors

Andrew Thiel – Roy J Carver Department of Biomedical Engineering, University of Iowa, Iowa City, Iowa 52242-1002, United States

Michael J. Schnieders – Roy J Carver Department of Biomedical Engineering, University of Iowa, Iowa City, Iowa 52242-1002, United States; orcid.org/0000-0003-1260-4592

Frank C. Pickard, IV – Pfizer Worldwide R&D, Groton, Connecticut 06340, United States; orcid.org/0000-0002-9608-3466

Geoffrey P.F. Wood – Pfizer Worldwide R&D, Groton, Connecticut 06340, United States

Joseph F. Krzyzaniak – Pfizer Worldwide R&D, Groton, Connecticut 06340, United States

Bruno C. Hancock – Pfizer Worldwide R&D, Groton, Connecticut 06340, United States; orcid.org/0000-0003-4387-2389

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.cgd.2c00832>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Brian Samas, Vishnu Sresht, Brajesh Rai, and Enoch Huang for helpful discussions. We would also like to thank Chris Ho for generating the AMOEBA parameter files for all compounds using the Automated AMOEBA Force Field Generator within Drug Design Methodologies, LLC. St. Louis, Missouri. We would also like to thank Ghazala Sadiq for obtaining structures for crystalline water from the Cambridge Structural Database. The authors acknowledge the High Performance Computing (HPC) Center at Pfizer for providing computational resources that have contributed to the research results reported within this paper. M.J.S. was supported by NSF CHE-1751688, and A.C.T. was supported by the NSF Graduate Research Fellowship Program DGE-1945994.

REFERENCES

- (1) Pudipeddi, M.; Serajuddin, A. T. M. Trends in Solubility of Polymorphs. *J. Pharm. Sci.* **2005**, *94*, 929–939.
- (2) De Villiers, M. M.; van der Watt, J. G.; Lötter, A. P. Kinetic Study of the Solid-State Photolytic Degradation of Two Polymorphic Forms of Furosemide. *Int. J. Pharm.* **1992**, *88*, 275–283.
- (3) Beyer, T.; Day, G. M.; Price, S. L. The Prediction, Morphology, and Mechanical Properties of the Polymorphs of Paracetamol. *J. Am. Chem. Soc.* **2001**, *123*, 5086–5094.
- (4) Singhal, D.; Curatolo, W. Drug Polymorphism and Dosage Form Design: A Practical Perspective. *Adv. Drug Delivery Rev.* **2004**, *56*, 335–347.
- (5) Censi, R.; Di Martino, P. Polymorph Impact on the Bioavailability and Stability of Poorly Soluble Drugs. *Molecules* **2015**, *20*, 18759–18776.
- (6) Lee, A. Y.; Erdemir, D.; Myerson, A. S. Crystal Polymorphism in Chemical Process Development. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *259*–280.
- (7) Meyer, M. C.; Straughn, A. B.; Jarvi, E. J.; Wood, G. C.; Pelsor, F. R.; Shah, V. P. The Bioinequivalence of Carbamazepine Tablets with a History of Clinical Failures. *Pharm. Res.* **1992**, *1612*–1616.
- (8) Zhang, G. G.; Law, D.; Schmitt, E. A.; Qiu, Y. Phase Transformation Considerations during Process Development and Manufacture of Solid Oral Dosage Forms. *Adv. Drug Delivery Rev.* **2004**, *56*, 371–390.
- (9) Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.; Spanton, S.; Dziki, W.; Porter, W.; Quick, J.; Bauer, P.; Donaubauber, J.; Narayanan, B. A.; Soldani, M.; Riley, D.; Mcfarland, K. Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development. *Org. Process Res. Dev.* **2000**, *4*, 413–417.
- (10) Bauer, J.; Spanton, S.; Henry, R.; Quick, J.; Dziki, W.; Porter, W.; Morris, J. Ritonavir: An Extraordinary Example of Conformational Polymorphism. *Pharm. Res.* **2001**, *18*, 859–866.
- (11) Chekal, B. P.; Campeta, A. M.; Abramov, Y. A.; Feeder, N.; Glynn, P. P.; McLaughlin, R. W.; Meenan, P. A.; Singer, R. A. The Challenges of Developing an API Crystallization Process for a Complex Polymorphic and Highly Solvating System. *Part I. Org. Process Res. Dev.* **2009**, *13*, 1327–1337.
- (12) Rietveld, I. B.; Céolin, R. Rotigotine: Unexpected Polymorphism with Predictable Overall Monotropic Behavior. *J. Pharm. Sci.* **2015**, *104*, 4117–4122.
- (13) Pan, D.; Crull, G.; Yin, S.; Grosso, J. Low Level Drug Product API Form Analysis - Avalide Tablet NIR Quantitative Method Development and Robustness Challenges. *J. Pharm. Biomed. Anal.* **2014**, *89*, 268–275.
- (14) Franc, A.; Muselík, J.; Zeman, J.; Lukášová, I.; Kurhájec, S.; Bartoníčková, E.; Galvánková, L.; Mika, F.; Dominik, M.; Vetchý, D. The Effect of Amorphous and Crystal Sodium Warfarin and Its Content Uniformity on Bioequivalence of Tablets. *Eur. J. Pharm. Sci.* **2018**, *125*, 120–129.
- (15) Wardrop, J.; Law, D.; Qiu, Y.; Engh, K.; Faitsch, L.; Ling, C. Influence of Solid Phase and Formulation Processing on Stability of Abbott-232 Tablet Formulations. *J. Pharm. Sci.* **2006**, *95*, 2380–2392.
- (16) Zhang, G. G.; Gu, C.; Zell, M. T.; Burkhardt, R. T.; Munson, E. J.; Grant, D. J. W. Crystallization and Transitions of Sulfamerazine Polymorphs. *J. Pharm. Sci.* **2002**, *91*, 1089–1100.
- (17) Zhang, P.; Wood, G. P. F.; Ma, J.; Yang, M.; Liu, Y.; Sun, G.; Jiang, Y. A.; Hancock, B. C.; Wen, S. Harnessing Cloud Architecture for Crystal Structure Prediction Calculations. **2018**, 6891, DOI: [10.1021/acs.cgd.8b01098](https://doi.org/10.1021/acs.cgd.8b01098).
- (18) Nyman, J.; Reutzel-Edens, S. M. Crystal Structure Prediction Is Changing from Basic Science to Applied Technology. *Faraday Discuss.* **2018**, 459.
- (19) Yang, M.; Dybeck, E.; Sun, G.; Peng, C.; Samas, B.; Burger, V. M.; Zeng, Q.; Jin, Y.; Bellucci, M. A.; Liu, Y.; Zhang, P.; Ma, J.; Jiang, Y. A.; Hancock, B. C.; Wen, S.; Wood, G. P. F. Prediction of the Relative Free Energies of Drug Polymorphs above Zero Kelvin. *Cryst. Growth Des.* **2020**, *20*, 5211–5224.
- (20) Eike, D. M.; Maginn, E. J. Atomistic Simulation of Solid-Liquid Coexistence for Molecular Systems: Application to Triazole and Benzene. *J. Chem. Phys.* **2006**, *124*, 164503.
- (21) Jayaraman, S.; Maginn, E. J. Computing the Melting Point and Thermodynamic Stability of the Orthorhombic and Monoclinic Crystalline Polymorphs of the Ionic Liquid 1-n-Butyl-3-Methylimidazolium Chloride. *J. Chem. Phys.* **2007**, *127*, 214504.
- (22) Vega, C.; Abascal, J. L. F.; Conde, M. M.; Aragonés, J. L. What Ice Can Teach Us about Water Interactions: A Critical Comparison of the Performance of Different Water Models. *Faraday Discuss.* **2009**, *2008*, 251–276.
- (23) Kofke, D. A. Gibbs-Duhem Integration: A New Method for Direct Evaluation of Phase Coexistence by Molecular Simulation. *Mol. Phys.* **1993**, *78*, 1331–1336.
- (24) Sweatman, M. B. Comparison of Absolute Free Energy Calculation Methods for Fluids and Solids. *Mol. Phys.* **2015**, *113*, 1206–1216.
- (25) Smoukov, S. K.; Burrows, S. A.; Korotkin, I.; Boek, E.; Karabasov, S. Benchmarking of Molecular Dynamics Force Fields for Solid-liquid and Solid-solid Phase Transitions in Alkanes. *J. Phys. Chem. B* **2021**, *125*, 32.
- (26) Beran, G. J. O.; Hartman, J. D.; Heit, Y. N. Predicting Molecular Crystal Properties from First Principles: Finite-Temperature Thermochemistry to NMR Crystallography. *Acc. Chem. Res.* **2016**, 2501.
- (27) Abraham, N. S.; Shirts, M. R. Statistical Mechanical Approximations to More Efficiently Determine Polymorph Free Energy Differences for Small Organic Molecules. *J. Chem. Theory Comput.* **2020**, *16*, 6503–6512.
- (28) Raiteri, P.; Martonák, R.; Parrinello, M. Exploring Polymorphism: The Case of Benzene. *Angew. Chem., Int. Ed.* **2005**, *44*, 3769–3773.
- (29) Zykova-Timan, T.; Raiteri, P.; Parrinello, M. Investigating the Polymorphism in PR179: A Combined Crystal Structure Prediction and Metadynamics Study. *J. Phys. Chem. B* **2008**, *112*, 13231–13237.
- (30) Yu, T.-Q.; Tuckerman, M. E. Temperature-Accelerated Method for Exploring Polymorphism in Molecular Crystals Based on Free Energy. *Phys. Rev. Lett.* **2011**, *107*, No. 015701.

- (31) Schneider, E.; Vogt, L.; Tuckerman, M. E. Exploring Polymorphism of Benzene and Naphthalene with Free Energy Based Enhanced Molecular Dynamics. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2016**, *72*, 542–550.
- (32) Frenkel, D.; Ladd, A. J. C. New Monte Carlo Method to Compute the Free Energy of Arbitrary Solids. Application to the Fcc and Hcp Phases of Hard Spheres. *J. Chem. Phys.* **1984**, *81*, 3188.
- (33) Vega, C.; Noya, E. G. Revisiting the Frenkel-Ladd Method to Compute the Free Energy of Solids: The Einstein Molecule Approach. *J. Chem. Phys.* **2007**, *127*, 154113.
- (34) Aragonés, J. L.; Noya, E. G.; Valeriani, C.; Vega, C. Free Energy Calculations for Molecular Solids Using GROMACS. *J. Chem. Phys.* **2013**, *139*, 034104.
- (35) Dybeck, E. C.; König, G.; Brooks, B. R.; Shirts, M. R. Comparison of Methods to Reweight from Classical Molecular Simulations to QM/MM Potentials. *J. Chem. Theory Comput.* **2016**, *12*, 1466–1480.
- (36) Dybeck, E. C.; Abraham, N. S.; Schieber, N. P.; Shirts, M. R. Capturing Entropic Contributions to Temperature-Mediated Polymorphic Transformations Through Molecular Modeling. *Cryst. Growth Des.* **2017**, *17*, 1775–1787.
- (37) Heit, Y. N.; Nanda, K. D.; Beran, G. J. O. Predicting Finite-Temperature Properties of Crystalline Carbon Dioxide from First Principles with Quantitative Accuracy. *Chem. Sci.* **2016**, *7*, 246–255.
- (38) Schieber, N. P.; Dybeck, E. C.; Shirts, M. R. Using Reweighting and Free Energy Surface Interpolation to Predict Solid-Solid Phase Diagrams. *J. Chem. Phys.* **2018**, *148*, 144104.
- (39) Duarte, G.; Matos, R.; Mobley, D. Challenges of the Use of Atomistic Simulations to Predict Solubilities of Drug-like Molecules. *FL1000Research* **2018**, *7*.
- (40) Khanna, V.; Anwar, J.; Frenkel, D.; Doherty, M. F.; Peters, B. Free Energies of Crystals Computed Using Einstein Crystal with Fixed Center of Mass and Differing Spring Constants. *J. Chem. Phys.* **2021**, *154*, 164509.
- (41) Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. Efficient Handling of Molecular Flexibility in Lattice Energy Minimization of Organic Crystals. *J. Chem. Theory Comput.* **2011**, *7*, 1998–2016.
- (42) Hoja, J.; Ko, H. Y.; Neumann, M. A.; Car, R.; DiStasio, R. A., Jr.; Tkatchenko, A. Reliable and Practical Computational Description of Molecular Crystal Polymorphs. *Acta Hortic. Sin.* **2019**, *5*, No. eaau3338.
- (43) Nyman, J.; Day, G. M. Static and Lattice Vibrational Energy Differences between Polymorphs. *CrystEngComm* **2015**, *17*, 5154–5165.
- (44) Nyman, J.; Day, G. M. Modelling Temperature-Dependent Properties of Polymorphic Organic Molecular Crystals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 31132–31143.
- (45) Ramírez, R.; Neuerburg, N.; Fernández-Serra, M.-V.; Herrero, C. P. Quasi-Harmonic Approximation of Thermodynamic Properties of Ice Ih, II, and III. *J. Chem. Phys.* **2012**, *137*, No. 044502.
- (46) Abraham, N. S.; Shirts, M. R. Thermal Gradient Approach for the Quasi-Harmonic Approximation and Its Application to Improved Treatment of Anisotropic Expansion. *J. Chem. Theory Comput.* **2018**, *14*, 5904–5919.
- (47) Abraham, N. S.; Shirts, M. R. Adding Anisotropy to the Standard Quasi-Harmonic Approximation Still Fails in Several Ways to Capture Organic Crystal Thermodynamics. *Cryst. Growth Des.* **2019**, *19*, 6911–6924.
- (48) Hulme, A. T.; Price, S. L. Toward the Prediction of Organic Hydrate Crystal Structures. *J. Chem. Theory Comput.* **2007**, *3*, 1597–1608.
- (49) Abramov, Y. A. Virtual Hydrate Screening and Coformer Selection for Improved Relative Humidity Stability. *CrystEngComm* **2015**, *17*, 5216–5224.
- (50) Liu, C.; Dang, L.; Bai, W.; Wei, H. Facile Method for the Prediction of Anhydrate/Hydrate Transformation Point. *Ind. Eng. Chem. Res.* **2013**, *52*, 16506–16512.
- (51) Larsen, A. S.; Rantanen, J.; Johansson, K. E. Computational Dehydration of Crystalline Hydrates Using Molecular Dynamics Simulations. *J. Pharm. Sci.* **2017**, *106*, 348–355.
- (52) Larsen, A. S.; Olsen, M. A.; Moustafa, H.; Larsen, F. H.; Sauer, S. P. A.; Rantanen, J.; Madsen, A. Ø. Determining Short-Lived Solid Forms during Phase Transformations Using Molecular Dynamics. *CrystEngComm* **2019**, *21*, 4020–4024.
- (53) Costa, D. G.; Rocha, A. B.; Souza, W. F.; Chiaro, S. S. X.; Leitão, A. A. Ab Initio Simulation of Changes in Geometry, Electronic Structure, and Gibbs Free Energy Caused by Dehydration of Hydrotalcites Containing Cl⁻ and CO₃²⁻ Counteranions. *J. Phys. Chem. B* **2011**, *115*, 3531–3537.
- (54) Tilbury, C. J.; Chen, J.; Mattei, A.; Chen, S.; Sheikh, A. Y. Combining Theoretical and Data-Driven Approaches to Predict Drug Substance Hydrate Formation. *Cryst. Growth Des.* **2018**, *18*, 57–67.
- (55) Basford, P. A.; Back, K. R.; Cram, M.; Docherty, R.; Davey, R. J.; Cruz-Cabeza, A. J. Impact of Crystal Structure and Molecular Conformation on the Hydration Kinetics of Fluconazole. *Cryst. Growth Des.* **2019**, *19*, 7193–7205.
- (56) Basford, P. A.; Cameron, C. A.; Cruz-Cabeza, A. J. Conformational Change Initiates Dehydration in Fluconazole Monohydrate. *Cryst. Growth Des.* **2020**, *20*, 6044–6056.
- (57) Hamad, M. L.; Engen, W.; Morris, K. R. Impact of Hydration State and Molecular Oxygen on the Chemical Stability of Levothyroxine Sodium. *Pharm. Dev. Technol.* **2015**, *20*, 314–319.
- (58) Shah, H. S.; Chaturvedi, K.; Hamad, M.; Bates, S.; Hussain, A.; Morris, K. New Insights on Solid-State Changes in the Levothyroxine Sodium Pentahydrate during Dehydration and Its Relationship to Chemical Instability. *AAPS PharmSciTech* **2019**, *20*, 39.
- (59) Kaur, N.; Young, V. G., Jr.; Su, Y.; Suryanarayanan, R. Partial Dehydration of Levothyroxine Sodium Pentahydrate in a Drug Product Environment: Structural Insights into Stability. *Mol. Pharmaceutics* **2020**, *17*, 3915–3929.
- (60) Braun, D. E.; McMahon, J. A.; Bhardwaj, R. M.; Nyman, J.; Neumann, M. A.; Van De Streek, J.; Reutzel-Edens, S. M. Inconvenient Truths about Solid Form Landscapes Revealed in the Polymorphs and Hydrates of Gandotinib. *Cryst. Growth Des.* **2019**, *19*, 2947–2962.
- (61) Bučar, D.-K.; Lancaster, R. W.; Bernstein, J. Disappearing Polymorphs Revisited. *Angew. Chem., Int. Ed.* **2015**, *54*, 6972–6993.
- (62) Tavares, S. R.; Vaiss, V. S.; Wypych, F.; Leitão, A. A. Similarities between Zinc Hydroxide Chloride Monohydrate and Its Dehydrated Form: A Theoretical Study of Their Structures and Anionic Exchange Properties. *J. Phys. Chem. C* **2014**, *118*, 19106–19113.
- (63) Tavares, S. R.; Vaiss, V. S.; Wypych, F.; Leitão, A. A. Theoretical Study of the Anion Exchange Properties and the Thermal Decomposition of Zn₅(OH)₈(NO₃)₂•2H₂O and Zn₅(OH)₈(NO₃)₂•2NH₃. **2015**, DOI: 10.1016/j.clay.2015.05.011.
- (64) Chihara, H.; Shinoda, T. Librational Motion in Diatomic Molecular Crystals. *J. Phys. Soc. Jpn.* **1962**, *17*, 1395–1397.
- (65) Descamps, M.; Correia, N. T.; Derollez, P.; Danede, F.; Capet, F. Plastic and Glassy Crystal States of Caffeine. *J. Phys. Chem. B* **2005**, *109*, 16092–16098.
- (66) Murugan, N. A.; Sayeed, A. Thermal Behavior of Disordered Phase of Caffeine Molecular Crystal: Insights from Monte Carlo Simulation Studies. *J. Chem. Phys.* **2009**, *130*, 204514.
- (67) Hore, S.; Dinnebier, R.; Wen, W.; Hanson, J.; Maier, J. Structure of Plastic Crystalline Succinonitrile: High-Resolution in Situ Powder Diffraction. *Z. Anorg. Allg. Chem.* **2009**, *635*, 88–93.
- (68) Shalaev, E.; Wu, K.; Shamblin, S.; Krzyzaniak, J. F.; Descamps, M. Crystalline Mesophases: Structure, Mobility, and Pharmaceutical Properties. *Adv. Drug Delivery Rev.* **2016**, *100*, 194–211.
- (69) Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. A Major Advance in Crystal Structure Prediction. *Angew. Chem., Int. Ed.* **2008**, *47*, 2427–2430.
- (70) Mooij, W. T.; van Eijck, B. P.; Price, S. L.; Verwer, P.; Kroon, J. Crystal Structure Predictions for Acetic Acid. *J. Comput. Chem.* **1998**, *19*, 459–474.

- (71) Habgood, M.; Sugden, I. J.; Kazantsev, A. V.; Adjiman, C. S.; Pantelides, C. C. Efficient Handling of Molecular Flexibility in Ab Initio Generation of Crystal Structures. *J. Chem. Theory Comput.* **2015**, *15*, 1957.
- (72) Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling Organic Crystal Structures Using Distributed Multipole and Polarizability-Based Model Intermolecular Potentials. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.
- (73) Vasileiadis, M.; Pantelides, C. C.; Adjiman, C. S. Prediction of the Crystal Structures of Axitinib, a Polymorphic Pharmaceutical Molecule. *Chem. Eng. Sci.* **2015**, *121*, 60–76.
- (74) 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*, 1721–1736.
- (75) Park, J.; Nessler, I.; McClain, B.; Macikenas, D.; Baltrusaitis, J.; Schnieders, M. J. Absolute Organic Crystal Thermodynamics: Growth of the Asymmetric Unit into a Crystal via Alchemy. *J. Chem. Theory Comput.* **2014**, *10*, 2781–2791.
- (76) Dybeck, E. C.; Schieber, N. P.; Shirts, M. R. Effects of a More Accurate Polarizable Hamiltonian on Polymorph Free Energies Computed Efficiently by Reweighting Point-Charge Potentials. *J. Chem. Theory Comput.* **2016**, *12*, 3491–3505.
- (77) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (78) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.
- (79) 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., Jr.; 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*, 2549–2564.
- (80) Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 3143–3161.
- (81) Wang, L. P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956–9972.
- (82) Nessler, I. J.; Litman, J. M.; Schnieders, M. J. Toward Polarizable AMOEBA Thermodynamics at Fixed Charge Efficiency Using a Dual Force Field Approach: Application to Organic Crystals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 30313–30322.
- (83) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J. P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput.* **2018**, *14*, 2084–2108.
- (84) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annu. Rev. Biophys.* **2019**, *48*, 371–394.
- (85) Shi, Y.; Laury, M. L.; Wang, Z.; Ponder, J. W. AMOEBA Binding Free Energies for the SAMPL7 TrimerTrip Host–Guest Challenge. *J. Comput.-Aided Mol. Des.* **2021**, *35*, 79–93.
- (86) Hummer, G.; Pratt, L. R.; García, A. E. Hydration Free Energy of Water. *J. Phys. Chem.* **1995**, *99*, 14188–14194.
- (87) Sakane, S.; Yezdimer, E. M.; Liu, W.; Barriocanal, J. A.; Doren, D. J.; Wood, R. H. Exploring the Ab Initio/Classical Free Energy Perturbation Method: The Hydration Free Energy of Water. *J. Chem. Phys.* **2000**, *113*, 2583–2593.
- (88) Milne, A. W.; Jorge, M. Polarization Corrections and the Hydration Free Energy of Water. *J. Chem. Theory Comput.* **2019**, *15*, 1065–1078.
- (89) Chen, L. R.; Grant, D. J. W. Extension of Clausius–Clapeyron Equation to Predict Hydrate Stability at Different Temperatures. *Pharm. Dev. Technol.* **1998**, *3*, 487–494.
- (90) Gu, C. H.; Grant, D. J. W. Estimating the Relative Stability of Polymorphs and Hydrates from Heats of Solution and Solubility Data. *J. Pharm. Sci.* **2001**, *90*, 1277–1287.
- (91) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.
- (92) Park, K.; Evans, J. M. B.; Myerson, A. S. Determination of Solubility of Polymorphs Using Differential Scanning Calorimetry Determination of Solubility of Polymorphs Using Differential Scanning Calorimetry. *Cryst. Growth Des.* **2003**, *3*, 991–995.
- (93) Grzesiak, A. L.; Lang, M.; Kim, K.; Matzger, A. J. Comparison of the Four Anhydrous Polymorphs of Carbamazepine and the Crystal Structure of Form I. *J. Pharm. Sci.* **2003**, *92*, 2260–2271.
- (94) Legendre, B.; Randzio, S. L. Transitionometric Analysis of Solid II/Solid I Transition in Anhydrous Theophylline. *Int. J. Pharm.* **2007**, *343*, 41–47.
- (95) Pickard, F. C.; König, G.; Tofoleanu, F.; Lee, J.; Simmonett, A. C.; Shao, Y.; Ponder, J. W.; Brooks, B. R. Blind Prediction of Distribution in the SAMPL5 Challenge with QM Based Protomer and PK a Corrections. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 1087–1100.
- (96) Li, L.; Totton, T.; Frenkel, D. Computational Methodology for Solubility Prediction: Application to the Sparingly Soluble Solutes. *J. Chem. Phys.* **2017**, *146*, 214110.
- (97) Noya, E. G.; Conde, M. M.; Vega, C. Computing the Free Energy of Molecular Solids by the Einstein Molecule Approach: Ices XIII and XIV, Hard-Dumbbells and a Patchy Model of Proteins. *J. Chem. Phys.* **2008**, *129*, 1–16.
- (98) Bellucci, M. A.; Gobbo, G.; Wijethunga, T. K.; Ciccotti, G.; Trout, B. L. Solubility of Paracetamol in Ethanol by Molecular Dynamics Using the Extended Einstein Crystal Method and Experiments. *J. Chem. Phys.* **2019**, *150*, 094107.
- (99) Sellers, M. S.; Li, M.; Bc, S.; Brennan, J. K. Free-Energy Calculations Using Classical Molecular Simulation: Application to the Determination of the Melting Point and Chemical Potential of a Flexible RDX Model. *Phys. Chem. Chem. Phys.* **2016**, *18*, 7841–7850.
- (100) Eike, D. M.; Brennecke, J. F.; Maginn, E. J. Toward a Robust and General Molecular Simulation Method for Computing Solid-Liquid Coexistence. *J. Chem. Phys.* **2005**, *122*, 14115.
- (101) Ben-Naim, A.; Marcus, Y. Solvation Thermodynamics of Nonionic Solutes. *J. Chem. Phys.* **1984**, *81*, 2016–2027.
- (102) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
- (103) Zwanzig, R. High - Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
- (104) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (105) Rickman, J. M.; LeSar, R. Free-Energy Calculations in Materials Research. *Annu. Rev. Mater. Res.* **2002**, *32*, 195–217.
- (106) Polson, J. M.; Trizac, E.; Pronk, S.; Frenkel, D. Finite-Size Corrections to the Free Energies of Crystalline Solids. *J. Chem. Phys.* **2000**, *112*, 5339–5342.
- (107) Santiso, E. E.; Trout, B. L. A General Set of Order Parameters for Molecular Crystals. *J. Chem. Phys.* **2011**, *134*, No. 064109.
- (108) Yu, T.-Q.; Chen, P.-Y.; Chen, M.; Samanta, A.; Vanden-Eijnden, E.; Tuckerman, M. Order-Parameter-Aided Temperature-Accelerated Sampling for the Exploration of Crystal Polymorphism and Solid-Liquid Phase Transitions. *J. Chem. Phys.* **2014**, *140*, 214109.
- (109) Zhang, Y.; Maginn, E. J. A Comparison of Methods for Melting Point Calculation Using Molecular Dynamics Simulations. *J. Chem. Phys.* **2012**, *136*, 144116.
- (110) Zhuang, L.; Wang, R.; Lindberg, G. E.; Hu, H.; Li, X. Z.; Wang, F. From a Liquid to a Crystal without Going through a First-Order Phase Transition: Determining the Free Energy of Melting with Glassy Intermediates. *J. Phys. Chem. B* **2019**, *123*, 7740–7747.

- (111) Granick, S. Motions and Relaxations of Confined Liquids. *Science* **1991**, 1374–1379.
- (112) Koga, K.; Gao, G. T.; Tanaka, H.; Zeng, X. C. Formation of Ordered Ice Nanotubes inside Carbon Nanotubes. *Nature* **2001**, 412, 802–805.
- (113) Alexiadis, A.; Kassinos, S. Molecular Simulation of Water in Carbon Nanotubes. *Chem. Rev.* **2008**, 108, 5014–5034.
- (114) Chong, S.-H.; Ham, S. Thermodynamic-Ensemble Independence of Solvation Free Energy. *J. Chem. Theory Comput.* **2015**, 11, 378–380.
- (115) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, 129, 124105.
- (116) McQuarrie, D. A. *Statistical Thermodynamics*; Harper & Row, 1973.
- (117) David, C. W. The Tensor Moment of Inertia. *Chem. Educ. Mater.* **2006**, 1–6.
- (118) Reutzel-Edens, S. M.; Braun, D. E.; Newman, A. W. Polymorphism in the Pharmaceutical Industry: Solid Form and Drug Development, 2. In *Polymorphism in the Pharmaceutical Industry: Solid Form and Drug Development 2*; John Wiley & Sons 2018; p 161.
- (119) Cengel, Y. A.; Boles, M. A. *Thermodynamics An Engineering Approach*; 5th Editio McGraw Hill Companies: New York, New York, 2002.
- (120) König, G.; Pickard, F.; Huang, J.; Thiel, W.; MacKerell, A.; Brooks, B.; York, D. A Comparison of QM/MM Simulations with and without the Drude Oscillator Model Based on Hydration Free Energies of Simple Solutes. *Molecules* **2018**, 23, 2695.
- (121) Litman, J.; Thiel, A. C.; Schnieders, M. J. Scalable Indirect Free Energy Method Applied to Divalent Cation-Metalloprotein Binding. *J. Chem. Theory Comput.* **2019**, 15, 4602–4614.
- (122) Zheng, L.; Chen, M.; Yang, W. Random Walk in Orthogonal Space to Achieve Efficient Free-Energy Simulation of Complex Systems. *Proc. Natl. Acad. Sci.* **2008**, 105, 20227.
- (123) Rager, T.; Geoffroy, A.; Hilfiker, R.; Storey, J. M. D. The Crystalline State of Methylene Blue: A Zoo of Hydrates. *Phys. Chem. Chem. Phys.* **2012**, 14, 8074–8082.
- (124) Abramov, Y.; Sun, G.; Zhou, Y.; Yang, M.; Zeng, Q.; Shen, Z. Solid-Form Transition Temperature Prediction from a Virtual Polymorph Screening: A Reality Check. *Cryst. Growth Des.* **2019**, 19, 7132–7137.
- (125) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, 25, 1157–1174.
- (126) Bayly, C. C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. ...* **1993**, 97, 10269–10280.
- (127) 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.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J. *Gaussian09*; Gaussian, Inc.: Wallingford, CT 2009.
- (128) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, 11, 3499–3509.
- (129) Dodda, L. S.; Cabeza De Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen Web Server: An Automatic OPLS-AA Parameter Generator for Organic Ligands. *Nucleic Acids Res.* **2017**, 45, W331–W336.
- (130) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem. Phys. Lett.* **1994**, 222, 529–539.
- (131) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, 1995, 8577–8593.
- (132) Wennberg, C. L.; Murtola, T.; Hess, B.; Lindahl, E. Lennard-Jones Lattice Summation in Bilayer Simulations Has Critical Effects on Surface Tension and Lipid Properties. *J. Chem. Theory Comput.* **2013**, 9, 3527–3537.
- (133) Goga, N.; Rzepiela, A. J.; de Vries, A. H.; Marrink, S. J.; Berendsen, H. J. C. Efficient Algorithms for Langevin and DPD Dynamics. *J. Chem. Theory Comput.* **2012**, 8, 3637–3649.
- (134) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, 52, 7182–7190.
- (135) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; Van Der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, I.; Feng, Y.; Moore, E. W.; Vanderplas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; Van Mulbregt, P.; Vijaykumar, A.; Bardelli, A. P.; Rothberg, A.; Hilboll, A.; Kloeckner, A.; Scopatz, A.; Lee, A.; Rokem, A.; Woods, C. N.; Fulton, C.; Masson, C.; Häggström, C.; Fitzgerald, C.; Nicholson, D. A.; Hagen, D. R.; Pasechnik, D. V.; Olivetti, E.; Martin, E.; Wieser, E.; Silva, F.; Lenders, F.; Wilhelm, F.; Young, G.; Price, G. A.; Ingold, G. L.; Allen, G. E.; Lee, G. R.; Audren, H.; Probst, I.; Dietrich, J. P.; Silterra, J.; Webber, J. T.; Slavič, J.; Nothman, A.; Buchner, J.; Kulick, J.; Schönberger, J. L.; De Miranda Cardoso, J. V.; Reimer, J.; Harrington, J.; Rodríguez, J. L. C.; Nunez-Iglesias, J.; Kuczynski, J.; Tritz, K.; Thoma, M.; Newville, M.; Kümmerer, M.; Bolingbroke, M.; Tarte, M.; Pak, M.; Smith, N. J.; Nowaczyk, N.; Shebanov, N.; Pavlyk, O.; Brodtkorb, P. A.; Lee, P.; McGibbon, R. T.; Feldbauer, R.; Lewis, S.; Tygier, S.; Sievert, S.; Vigna, S.; Peterson, S.; More, S.; Pudlik, T.; Oshima, T.; Pingel, T. J.; Robitaille, T. P.; Spura, T.; Jones, T. R.; Cera, T.; Leslie, T.; Zito, T.; Krauss, T.; Upadhyay, U.; Halchenko, Y. O.; Vázquez-Baeza, Y.; et al. *Nat. Methods* **2020**, 17, 261–272.
- (136) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Application*; Elsevier, 2002.
- (137) Dama, J. F.; Rotskoff, G.; Parrinello, M.; Voth, G. A. Transition-Tempered Metadynamics : Robust , Convergent Metadynamics via On-the-Fly Transition Barrier Estimation. *J. Chem. Theory Comput.* **2014**, 10, 3626–3633.
- (138) Huang, J.; Lopes, P. E.; Roux, B.; Mackerell, A. D. Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *J. Phys. Chem. Lett.* **2014**, 5, 3144–3150.
- (139) Huang, J.; Mackerell, A. D., Jr. Induction of Peptide Bond Dipoles Drives Cooperative Helix Formation in the (AAQAA)₃ Peptide. *Biophys. J.* **2014**, 107, 991–997.
- (140) Lemkul, J. A.; Huang, J.; MacKerell, A. D., Jr. Induced Dipole-Dipole Interactions Influence the Unfolding Pathways of Wild-Type and Mutant Amyloid β -Peptides. *J. Phys. Chem. B* **2015**, 119, 15574–15582.
- (141) DiStasio, R. A.; Santra, B.; Li, Z.; Wu, X.; Car, R. The Individual and Collective Effects of Exact Exchange and Dispersion Interactions on the Ab Initio Structure of Liquid Water. *J. Chem. Phys.* **2014**, 141, 1–18.
- (142) Cruz-Cabeza, A. J.; Wright, S. E.; Bacchi, A. On the Entropy Cost of Making Solvates. *Chem. Commun.* **2020**, 56, 5127–5130.
- (143) Wright, S. E.; Cole, J. C.; Cruz-Cabeza, A. J. Conformational Change in Molecular Crystals: Impact of Solvate Formation and Importance of Conformational Free Energies. *Cryst. Growth Des.* **2021**, 21, 6924–6936.
- (144) Aragoes, J. L.; Conde, M. M.; Noya, E. G.; Vega, C. The Phase Diagram of Water at High Pressures as Obtained by Computer Simulations of the TIP4P/2005 Model: The Appearance of a Plastic Crystal Phase. *Phys. Chem. Chem. Phys.* **2009**, 11, 543–555.
- (145) Gao, G. T.; Zeng, X. C.; Tanaka, H. The Melting Temperature of Proton-Disordered Hexagonal Ice: A Computer Simulation of 4-Site Transferable Intermolecular Potential Model of Water. *J. Chem. Phys.* **2000**, 112, 8534–8538.
- (146) Koyama, Y.; Tanaka, H.; Gao, G.; Zeng, X. C. Melting Points and Thermal Expansivities of Proton-Disordered Hexagonal Ice with Several Model Potentials. *J. Chem. Phys.* **2004**, 121, 7926–7931.

(147) Wang, L.-P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2013**, *9*, 452–460.

(148) Braun, D. E.; Gelbrich, T.; Kahlenberg, V.; Tessadri, R.; Wieser, J.; Griesser, U. J. Stability of Solvates and Packing Systematics of Nine Crystal Forms of the Antipsychotic Drug Aripiprazole. *Cryst. Growth Des.* **2009**, *9*, 1054–1065.

(149) Campeta, A. M.; Chekal, B. P.; Abramov, Y. A.; Meenan, P. A.; Henson, M. J.; Shi, B.; Singer, R. A.; Horspool, K. R. Development of a Targeted Polymorph Screening Approach for a Complex Polymorphic and Highly Solvating API. *J. Pharm. Sci.* **2010**, *99*, 3874–3886.

(150) Samas, B.; Seadeek, C.; Campeta, A. M.; Chekal, B. P. A Thermodynamic-Based Approach to Analyzing a Highly Solvating Polymorphic System: The Desolvation Window Method. *J. Pharm. Sci.* **2011**, *100*, 186–194.

(151) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; Case, D. H.; Chadha, R.; Cole, J. C.; Cosburn, K.; Cuppen, H. M.; Curtis, F.; Day, G. M.; Distasio, R. A., Jr.; Dzyabchenko, A.; Van Eijck, B. P.; Elking, D. M.; Can Den Ende, J. A.; Facelli, J. C.; Ferraro, M. B.; Fusti-Molnar, L.; Gatsiou, C.-A.; Gee, T. S.; De Gelder, R.; Ghiringhelli, L. M.; Goto, H.; Grimme, S.; Guo, R.; Hofmann, D. W. M.; Hoja, J.; Hylton, R. K.; Iuzzolino, L.; Jankiewicz, W.; De Jong, D. T.; Kendrick, J.; De Klerk, N. J. J.; Ko, H.-Y.; Kuleshova, L. N.; Li, X.; Lohani, S.; Leusen, F. J. J.; Lund, A. M.; Lv, J.; Ma, Y.; Marom, N.; Masunov, A. E.; McCabe, P.; McMahon, D. P.; Meekes, H.; Metz, M. P.; Misquitta, A. J.; Mohamed, S.; Monserrat, B.; Needs, R. J.; Neumann, M. A.; Nyman, J.; Obata, S.; Oberhofer, H.; Oganov, A. R.; Orendt, A. M.; Pagola, G. I.; Pantelides, C. C.; Pickard, C. J.; Podeszwa, R.; Price, L. S.; Price, S. L.; Pulido, A.; Read, M. G.; Reuter, K.; Schneider, E.; Schober, C.; Shields, G. P.; Singh, P.; Sugden, I. J.; Szaleqicz, K.; Taylor, C. R.; Tkatchenko, A.; Tuckerman, M. E.; Vacarro, F.; Vasileiadis, M.; Vazquez-Mayagoitia, A.; Vogt, L.; Wang, Y.; Watson, R. E.; De Wijs, G. A.; Yang, J.; Zhu, Q.; Groom, C. R. Report on the Sixth Blind Test of Organic Crystal-Structure Prediction. *Methods* **2016**, *72*, 439–459.

(152) Van Eijck, B. P.; Kroon, J. Structure Predictions Allowing More than One Molecule in the Asymmetric Unit. *Acta Crystallogr. Sect. B: Struct. Sci.* **2000**, *56*, 535–542.

(153) Braun, D. E.; Bhardwaj, R. M.; Florence, A. J.; Tocher, D. A.; Price, S. L. Complex Polymorphic System of Gallic Acid - Five Monohydrates, Three Anhydrides, and over 20 Solvates. *Cryst. Growth Des.* **2013**, *13*, 19–23.

(154) Braun, D. E.; Kahlenberg, V.; Griesser, U. J. Experimental and Computational Hydrate Screening: Cytosine, 5-Flucytosine, and Their Solid Solution. *Cryst. Growth Des.* **2017**, *17*, 4347–4364.

(155) Weast, R. C. *CRC Handbook of Chemistry and Physics*, 62nd Edition; 1981.

Recommended by ACS

Synergistic Control of Nonlinear Growth Kinetics and Nucleation Kinetics in the Concomitant Crystallization of Aripiprazole as Reflected by the Ostwald Ratio

Zhenjie Gao, Junbo Gong, *et al.*

NOVEMBER 14, 2022

INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH

READ 

Seeking Rules Governing Mixed Molecular Crystallization

Norbert M. Villeneuve, James D. Wuest, *et al.*

DECEMBER 15, 2022

CRYSTAL GROWTH & DESIGN

READ 

Kinetic Retraction at the Onset of Concomitant Crystallization and Implication on Polymorphic Formation

Weiwei Tang, Tonglei Li, *et al.*

MAY 27, 2022

MOLECULAR PHARMACEUTICS

READ 

Structural, Physicochemical, and Biopharmaceutical Properties of Cocrystals with *RS*- and *R*-Praziquantel—Generation and Prolongation of the Supers...

Christian Rodríguez-Ruiz, Herbert Höpfl, *et al.*

SEPTEMBER 02, 2022

CRYSTAL GROWTH & DESIGN

READ 

Get More Suggestions >