

# Diabat method for polymorph free energies: Extension to molecular crystals

Cite as: J. Chem. Phys. **153**, 244105 (2020); <https://doi.org/10.1063/5.0024727>

Submitted: 10 August 2020 . Accepted: 02 December 2020 . Published Online: 23 December 2020

 Kartik Kamat,  Rui Guo,  Susan M. Reutzler-Edens,  Sarah L. Price, and  Baron Peters



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Confronting pitfalls of AI-augmented molecular dynamics using statistical physics](#)

The Journal of Chemical Physics **153**, 234118 (2020); <https://doi.org/10.1063/5.0030931>

[Reflections on electron transfer theory](#)

The Journal of Chemical Physics **153**, 210401 (2020); <https://doi.org/10.1063/5.0035434>

[Excited state diabaticization on the cheap using DFT: Photoinduced electron and hole transfer](#)

The Journal of Chemical Physics **153**, 244111 (2020); <https://doi.org/10.1063/5.0035593>



**SHFQA**  
Quantum Analyzer  
8.5GHz

Zurich Instruments

**Your Qubits. Measured.**

Meet the next generation of quantum analyzers

- Readout for up to 64 qubits
- Operation at up to 8.5 GHz, mixer-calibration-free
- Signal optimization with minimal latency

Find out more

 Zurich Instruments

# Diabat method for polymorph free energies: Extension to molecular crystals

Cite as: *J. Chem. Phys.* **153**, 244105 (2020); doi: [10.1063/5.0024727](https://doi.org/10.1063/5.0024727)

Submitted: 10 August 2020 • Accepted: 2 December 2020 •

Published Online: 23 December 2020



View Online



Export Citation



CrossMark

Kartik Kamat,<sup>1</sup>  Rui Guo,<sup>2</sup>  Susan M. Reutzel-Edens,<sup>3</sup>  Sarah L. Price,<sup>2</sup>  and Baron Peters<sup>4,5,a)</sup> 

## AFFILIATIONS

<sup>1</sup>Department of Chemical Engineering, University of California–Santa Barbara, Santa Barbara, California 93106, USA

<sup>2</sup>Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

<sup>3</sup>Small Molecule Design and Development, Eli Lilly and Company, Indianapolis, Indiana 46285, USA

<sup>4</sup>Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

<sup>5</sup>Department of Chemistry and Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

<sup>a)</sup> Author to whom correspondence should be addressed: [baronp@illinois.edu](mailto:baronp@illinois.edu)

## ABSTRACT

Lattice-switch Monte Carlo and the related diabat methods have emerged as efficient and accurate ways to compute free energy differences between polymorphs. In this work, we introduce a one-to-one mapping from the reference positions and displacements in one molecular crystal to the positions and displacements in another. Two features of the mapping facilitate lattice-switch Monte Carlo and related diabat methods for computing polymorph free energy differences. First, the mapping is unitary so that its Jacobian does not complicate the free energy calculations. Second, the mapping is easily implemented for molecular crystals of arbitrary complexity. We demonstrate the mapping by computing free energy differences between polymorphs of benzene and carbamazepine. Free energy calculations for thermodynamic cycles, each involving three independently computed polymorph free energy differences, all return to the starting free energy with a high degree of precision. The calculations thus provide a force field independent validation of the method and allow us to estimate the precision of the individual free energy differences.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0024727>

## I. INTRODUCTION

Once regarded as a rare curiosity, polymorphism is now recognized as a ubiquitous and important phenomenon from metals<sup>1</sup> to minerals<sup>2</sup> to molecular crystals.<sup>3</sup> Polymorphs are different crystal structures for the same molecule or compound. Different polymorphs have different properties, e.g., different catalytic activities,<sup>4–7</sup> different photophysical properties,<sup>8–10</sup> different mechanical properties,<sup>11,12</sup> and different thermodynamic stabilities. For active pharmaceutical ingredients (APIs), which are often formulated as molecular crystals, thermodynamic stability is of paramount importance. The discovery and characterization of API polymorphs, especially the most stable forms, is important for ensuring reliable crystallization processes, for making stable products, and for optimizing solubility and bioavailability.<sup>13,14</sup> Polymorph

screening has become integral to pharmaceutical product development,<sup>15,16</sup> and computational studies are already making important contributions.<sup>17</sup>

Ongoing efforts in computational polymorph screening are directed at three different aspects of the problem. First, crystal structure prediction seeks to discover all local energy minima in the enormous space of periodic crystal structures.<sup>18–21</sup> Second, thermodynamic property calculations seek to determine relative stabilities and absolute solubilities for the most stable structures.<sup>22,23</sup> Third, studies of nucleation<sup>24–27</sup> and growth kinetics<sup>28–30</sup> seek to predict and control polymorph appearance and crystal growth habits.

This paper focuses on thermodynamic property calculations, specifically on the relative free energies of different polymorphs.<sup>31</sup> The importance of free energies beyond lattice energies is underscored by the work of Nyman and Day, who studied 508 organic

molecules and found that 9% of their polymorphic pairs have their relative stability reversed (at or below room temperature) on accounting for entropy.<sup>32</sup> Moreover, the *observed* polymorphs have free energies that typically differ by only a few kJ/mol, so extremely accurate models and methods are required to make quantitative conclusions.<sup>31,32</sup>

Several methods for computing polymorph free energies have already been developed. *Ab initio* calculations can predict free energies from 0 K lattice energies, zero-point energies, and harmonic vibration/phonon mode contributions. *Ab initio* calculations can be performed for any new compound with no force field development, but the predictions are sensitive to the model chemistry, i.e., to the choice of density functional, basis set, etc.<sup>33–35</sup> Recent advances have targeted high accuracy with state-of-the-art density functionals, dispersion corrections, and anharmonic vibrational corrections.<sup>36–39</sup> For these methods, entropic contributions beyond harmonic vibrational analyses are challenging because the *ab initio* calculations are (at present) too costly for advanced sampling methods.<sup>40,41</sup> However, quasi-harmonic approximations can also be quite accurate as long as the anticipated structural minimum does not restructure during molecular dynamics (MD) simulations at/above the temperature of interest.<sup>42</sup>

Advanced sampling techniques in molecular simulation are (at present) limited by force field availability and accuracy. However, they can fully capture entropic contributions,<sup>43–46</sup> even for anharmonic modes and potentially for hindered rotors as in plastic crystals. Molecular simulations with path integral techniques can, in principle, account for nuclear quantum effects such as zero-point contributions. Ultimately, one hopes to combine the strengths of *ab initio* calculations and advanced sampling methods by sampling with *ab initio* molecular dynamics instead of empirical force fields. Of course, achieving this goal will require extremely efficient sampling methods.

The first advanced sampling methods for crystal free energies were Frenkel–Ladd type methods,<sup>47</sup> which transform an Einstein crystal (or an Einstein molecule)<sup>48</sup> to the “real” crystal.<sup>49</sup> These calculations require many stages of thermodynamic integration or free energy perturbations, but they are powerful routes to the absolute free energy. Frenkel–Ladd calculations can also yield polymorph free energy differences by subtracting the results of two Frenkel–Ladd calculations. This strategy is prone to errors because two typically large free energies are being subtracted to obtain a typically small free energy difference.

Alternative advanced sampling methods drive the transformation directly from one polymorph to the other by biasing along collective variables. One such method, successfully used by Martoňák *et al.*<sup>50</sup> and others,<sup>45,51</sup> biases the crystal lattice parameters to drive the polymorph transformation. This method has not been used (and seems unlikely to work) for typical polymorphs of organic molecules that have to undergo large scale conformational changes or repack with entirely different relative orientations and so lose translational symmetry during the transformation.

A clever alternative known as lattice-switch Monte Carlo (LSMC) biases an energy gap order parameter to drive the polymorph transformation.<sup>52,53</sup> The LSMC transformation path is not easily visualized: one maps thermal displacements atom-by-atom from the 0 K structure for one polymorph onto the 0 K structure for the other and then computes the energy difference between

polymorphs with the same displacements to obtain the energy gap. The energy gap varies from negative values at equilibrium for the starting polymorph to positive values at equilibrium for the final polymorph. In previous work, we showed that the LSMC free energy profiles are superpositions of free energy diabats—one for the starting polymorph and one for the final polymorph with the two diabats crossing where the energy gap is zero.<sup>54</sup> We used (and generalized<sup>55</sup>) the Zwanzig–Bennett relation<sup>56,57</sup> to calculate Helmholtz and Gibbs free energy differences ( $\Delta F$  and  $\Delta G$ , respectively) between the two polymorphs.

The LSMC and diabat methods show significant promise, but nearly all applications have focused on allotropes, i.e., polymorphs of elemental solids.<sup>53,58,59</sup> The one exception is an LSMC calculation for polymorphs of a united-atom butane model by Bridgewater and Quigley.<sup>60,61</sup> They successfully implemented the LSMC approach by mapping centers of mass, orientations, and internal coordinate displacements from one polymorph to another, but their procedures must be implemented on a case-by-case basis for different molecules. Moreover, the Jacobian from Cartesian displacements (differential volume in 3 N dimensions) depends on the internal coordinates and therefore on the polymorph. Here, we introduce a general mapping that extends LSMC and diabat methods to enable calculations in the same manner for polymorphs of any organic molecular crystals. The procedures are illustrated with calculations for benzene and carbamazepine polymorphs.

## II. OVERVIEW OF THE DIABAT METHOD

We consider two polymorphs, **A** and **B**. A finite-temperature molecular dynamics (MD) simulation of polymorph **A** samples configurations near, but displaced from, the 0 K crystal structure of **A**. Each configuration visited in state **A** is decomposed into a reference 0 K configuration and a displacement vector  $\tilde{\mathbf{u}}^N$ . The displacements are imposed via a one-to-one mapping onto a reference 0 K configuration in **B**. The energy gap is computed by subtracting the energy of configurations that have exactly the same displacements. Specifically, the energy gap between the two polymorphs (**A** and **B**) is given by

$$\Delta E(\tilde{\mathbf{u}}^N) = E_A(\tilde{\mathbf{u}}^N) - E_B(\tilde{\mathbf{u}}^N). \quad (1)$$

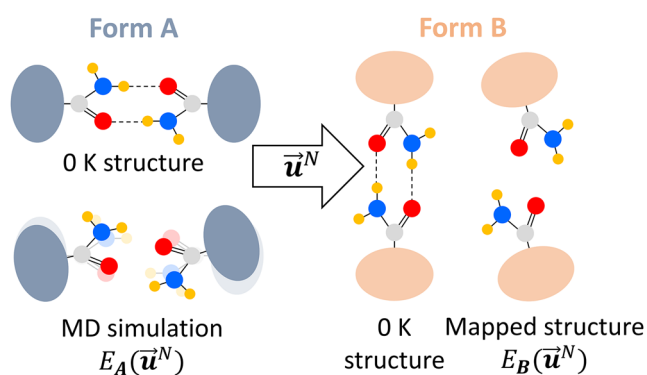
An observed configuration in **A** is mapped onto a configuration in **B** (see Fig. 1) by means of atomic displacements ( $\tilde{\mathbf{u}}^N$ ).

LSMC uses regular Monte Carlo moves with a special lattice-switch move.<sup>53,58</sup> The lattice-switch, if accepted, switches the active system from one polymorph to the other. The Metropolis criteria used to accept/reject the lattice-switch move (from phase **A** to **B**) in the canonical ensemble<sup>58,62</sup> is

$$P_{\mathbf{A} \rightarrow \mathbf{B}}^{\text{acc}} = \min\left\{1, e^{\beta \Delta E}\right\}, \quad (2)$$

where  $\beta$  is  $1/k_B T$ . The method samples the free energy barrier between the two polymorphs along the energy gap order parameter in a single simulation.

The lattice-switch move is always rejected in the minima of the free energy basins of the two polymorphs. This is because a



**FIG. 1.** A schematic for the carbamazepine molecule where only the carboxamide atoms have been shown for clarity. A configuration from an MD simulation is compared to the 0 K reference structure to calculate the atomic displacements. The atomic displacements are then one-to-one mapped and applied to the 0 K form B structure to yield a mapped configuration in B. The mapping procedure is described in detail in Sec. III (main text) and Sec. S.1 (supplementary material).

typical equilibrium configuration in **A** is a highly unfavorable configuration in **B** (and vice versa). Thus, the value of the energy gap in the basin of **A** is highly negative. Clearly, from Eq. (2), lattice-switch moves are mostly accepted if the system is near states of  $\Delta E \approx 0$ . LSMC typically uses a bias sampling strategy (e.g., transition matrix Monte Carlo<sup>63,64</sup>) to bias the simulation to these “gateway” states.

In seemingly unrelated research on electron transfer,<sup>65,66</sup> free energy diabats as functions of the energy gap obey the Zwanzig–Bennett relation,<sup>56,57,67–69</sup>

$$F_A(\Delta E) - F_B(\Delta E) = \Delta E. \quad (3)$$

In electron transfer, the diabats  $F_A(\Delta E)$  and  $F_B(\Delta E)$  are free energies of the system restricted to one electronic state. In the context of polymorphs, the diabats are free energies of the system restricted to one of the two polymorphs. The Zwanzig–Bennett relation facilitates the diabat calculations. Upon computing  $F_A(\Delta E)$  at  $\Delta E$ , the value of  $F_B(\Delta E)$  is exactly known.

The diabat methods use the Zwanzig–Bennett relation to compute  $F_A(\Delta E)$  and  $F_B(\Delta E)$  for the two polymorphs. Diabat interpolation yields tremendous efficiency advantages over other methods in cases where the diabats are parabolic. Kamat and Peters used just two unbiased simulations to compute Helmholtz free energy differences of the Gaussian core solid. The results were in good agreement with an independent study that used the multi-stage Frenkel–Ladd Einstein crystal method.<sup>54</sup> However, diabat interpolation fails when the diabats are not parabolic. The parabolic model for the diabat is an approximation, and there is a theoretical basis to anticipate parabolic diabats for typical atomic solids. Additionally, the Zwanzig–Bennett relation in Eq. (3) is limited to crystals of the same density (volume), whereas most practical applications require the isothermal–isobaric (NPT) ensemble.

Kamat and Peters showed that Gibbs free energy diabats using the order parameter<sup>35</sup>

$$\Delta Q = E_A(\vec{u}^N) - E_B(\vec{u}^N) + P(V_A - V_B) + \beta^{-1} \ln \frac{V_B}{V_A} \quad (4)$$

obey a Zwanzig–Bennett-like relation in the NPT ensemble,

$$G_A(\Delta Q) - G_B(\Delta Q) = \Delta Q. \quad (5)$$

As with the Helmholtz version of the Zwanzig–Bennett relation, Eq. (5) allows us to compute one Gibbs free energy diabat and obtain the other for free.

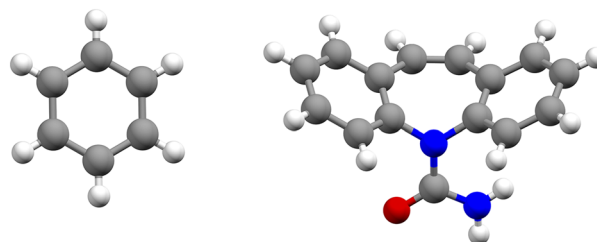
Kamat and Peters used short unbiased simulations to construct preliminary bias potentials to reduce the barrier of the free energy surfaces. The free energy surfaces for the polymorphs are typically anharmonic. A parabolic fit to the basin minima (obtained from unbiased simulations) is used to generate the preliminary bias potentials. If the true free energy surface were perfectly parabolic, the preliminary biases flatten the free energy surface, so a single simulation can sample the complete range of the order parameter. For most cases, the preliminary bias potentials only partially flatten the landscape, but they still reduce the computational effort required to sample the full pathway from one polymorph to the other. Kamat and Peters demonstrated moderate efficiency advantages compared to the standard LSMC framework for the body-centered cubic and hexagonal close-packed phases of zirconium.<sup>55</sup> In this work, we extend the diabat method to non-orthorhombic crystals of flexible molecules. The mapping procedure we propose is system independent and easy to implement. We demonstrate the method for benzene and carbamazepine molecular crystals. We also discuss the efficiency advantages of the method.

### III. MOLECULAR CRYSTALS

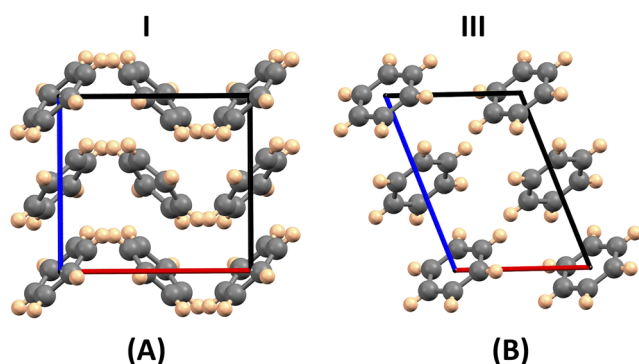
In this section, we demonstrate the method for molecular crystals of the two molecules shown in Fig. 2. Benzene was selected to illustrate the method for a plastic crystal with symmetric and/or degenerate configurations. Carbamazepine was selected to illustrate that the method can be applied to complex pharmaceutical molecules.

#### A. Benzene

Molecular crystals pose new challenges to the diabat approach. In particular, symmetric molecules that can rotate or jump between degenerate configurations within the crystal pose difficulties for the atom-to-atom mapping procedure. In principle, any one-to-one



**FIG. 2.** The planar benzene molecule (left) is symmetric about the center of mass. The carbamazepine molecule (right) has no symmetry apart from the flipping of the amide group as the aromatic rings are not co-planar.



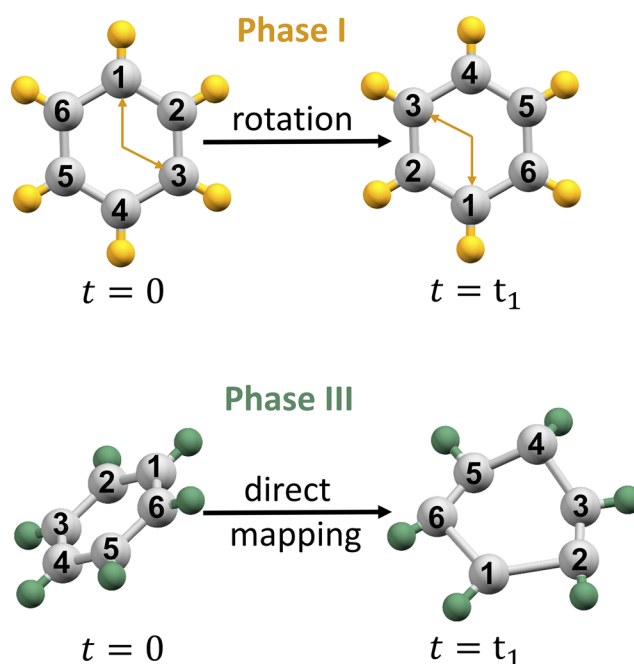
**FIG. 3.** (a) Benzene polymorph I (Pbca,  $Z = 4$ ,  $Z' = 0.5$ ) with lattice parameters (in Å)  $a = 7.28$ ,  $b = 9.20$ ,  $c = 6.68$ , and  $\alpha = \beta = \gamma = 90^\circ$ .<sup>75</sup> (b) Benzene polymorph III (P2<sub>1</sub>/c,  $Z = 2$ ,  $Z' = 0.5$ ) with lattice parameters  $a = 5.41$ ,  $b = 5.37$ ,  $c = 7.53$ ,  $\alpha = \gamma = 90^\circ$ , and  $\beta = 110^\circ$ .<sup>76</sup> Figures are rendered in Mercury.<sup>77</sup> The red and blue axes represent lattice edges  $a$  and  $c$ , respectively.

mapping should work, but, in practice, the Zwanzig–Bennett relation is a perturbation result and its accuracy is best when small  $\Delta Q$  changes in one phase correspond to small  $\Delta Q$  changes in the other. Molecular symmetry needs special handling to give the smallest  $\Delta Q$  changes for the polymorphs involved.

Benzene is one of the simplest molecules that present all the challenges. It has five experimentally observed polymorphs,<sup>70–72</sup> and up to seven possible stable phases have been reported at finite temperature using a classical forcefield.<sup>51</sup> Here, we compute the free energy difference between polymorphs I and III (Fig. 3) with the OPLS-AA (optimized potentials for liquid simulations all atom) forcefield<sup>73,74</sup> (see the [supplementary material](#) for computational details).

For atomic solids,  $\Delta Q$  was computed by directly mapping atom displacements from one phase to the other. Directly mapped Cartesian displacements, however, work poorly for crystals of symmetric molecules that can switch between degenerate configurations in the crystal lattice. In such cases, the displacements mapped directly from one phase to another distort the molecular structure, as shown in Fig. 4. Bridgwater and Quigley encountered the same problem, and they used an internal coordinate mapping in their LSMC study of solid butane phases.<sup>60</sup>

To overcome the problem in Fig. 4, we define a local coordinate system for each benzene molecule. First, we construct the vectors from the center of mass of the molecule to carbon atoms 1 and 3,  $\vec{w}^1, \vec{w}^2$  in reference phase I and  $\vec{W}^1, \vec{W}^2$  in reference phase III, as shown in Fig. 5. The linearly independent set of vectors ( $\vec{w}^1, \vec{w}^2, \vec{w}^2 \times \vec{w}^1$ ) is used to construct an orthonormal basis set ( $\vec{b}^1, \vec{b}^2, \vec{b}^3$ ) with Gram–Schmidt orthogonalization.<sup>78</sup> For all other atoms in the molecule, displacements from the reference position in polymorph I are resolved with respect to the new orthonormal basis. These displacements are then imposed on an analogously defined set of orthonormal basis vectors for a corresponding molecule in polymorph III (see the [supplementary material](#) for details). The procedure is repeated for all the molecules in phases I and III. A rotation of the molecule  $i$  in its local plane in phase I maps as a local rotation of the molecule  $i$  in phase III. Details on the construction of the local



**FIG. 4.** The  $t = 0$  structures are reference structures of I and III. An unbiased simulation in phase I can lead to a configuration shown at  $t_1$ . A  $180^\circ$  rotation of the benzene molecule in phase I places the atoms into a perfectly acceptable configuration. The same Cartesian displacements applied to a benzene molecule in phase III cause severe distortions because the benzenes in phases I and III are oriented differently.

coordinate system and the mapping procedure are available in the [supplementary material](#).

Collective translational motion, however, can still lead to a high energy configuration in the mapped phase, as shown in Fig. 6, where displacements that are easy in one phase map to high energy atom overlaps in the other phase. We prevent the atom overlaps by adding a system-independent and exactly removable bias. Specifically, we use harmonic restraints on all the atoms. The lattice-switch and/or diabatic calculations are performed with these harmonic restraints turned on so that the lattice-switch moves involve only small displacements. We must then turn the restraints off at the ends of the paths. Accordingly, our thermodynamic path, shown in Fig. 7, mixes the diabatic approach with the Frenkel–Ladd approach. Additionally, unlike the Frenkel–Ladd approach, the Lennard–Jones and electrostatic interactions need not be turned off.

In the calculations, we start with the average volume crystal structures for the two polymorphs as obtained from equilibration simulation at 250 K and 1 bar (see the [supplementary material](#) for details). We ignore the contribution from volume fluctuations but account for the change in average volumes. See Sec. S.9 of the [supplementary material](#) for evidence that the fluctuations are negligible compared to the change in average volumes.

First, form I is reversibly transformed to form I<sub>S</sub> where all the atoms of the molecule have harmonic restraints (subscript S denotes the springs). Similarly, form III is reversibly transformed to III<sub>S</sub>. Each benzene molecule can rotate among six degenerate minimum

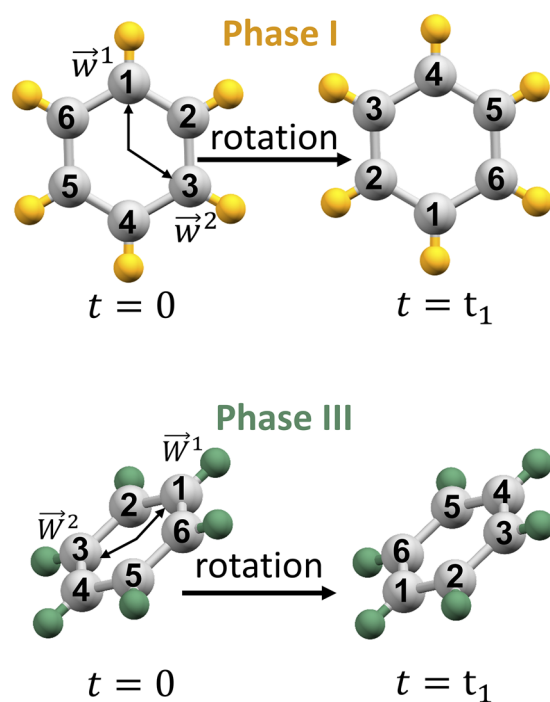


FIG. 5. The local coordinate system for each molecule helps to map local changes in phase I to phase III with similar distortions in both phases.

energy configurations in states I and III. The harmonic restraints bias sampling toward just one of these minima. However, the other minima are accounted for upon reversible growth and removal of the springs in the paths from I to  $I_S$  and  $III_S$  to III. For crystals with large barriers between local minima, it may be necessary to use replica exchange methods to enhance sampling while growing and dissolving the springs.

The diabatic approach is used to compute the free energy difference between  $I_S$  and  $III_S$ . We use the Zwanzig–Bennett relation to calculate the Landau free-energies of  $I_S$  and  $III_S$ . The potential energy functions of  $I_S$  and  $III_S$  have an additional spring energy term relative to the potential energy functions of I and III. The Gram–Schmidt based mapping procedure, however, causes the spring energies to exactly cancel out while computing  $\Delta Q$  for forms  $I_S$  and  $III_S$  (see Sec. S.2 in the [supplementary material](#)). Thus, the Zwanzig–Bennett equation is equally valid for forms  $I_S$  and  $III_S$ ,

$$G_{I_S}(\Delta Q) - G_{III_S}(\Delta Q) = \Delta Q. \quad (6)$$

We now describe the procedure for calculating the free energy difference along path 2 in our thermodynamic path (Fig. 7). Unbiased simulations are run for both phases ( $I_S$  and  $III_S$ ). Estimates for the local free energy surface are used to construct bias potentials for umbrella sampling the free energy surfaces (see Sec. S.4 in the [supplementary material](#)). The order parameter  $\Delta Q$  requires periodic interruption of molecular dynamics to simulate the two phases simultaneously. The Langevin thermostat in LAMMPS equilibrates to an incorrect temperature when combined with dynamics that is restarted every single time step. We fix this problem

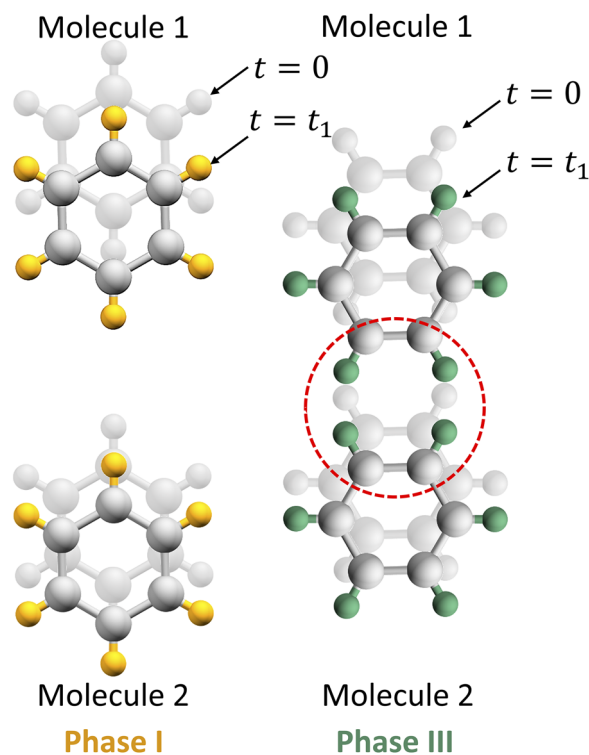
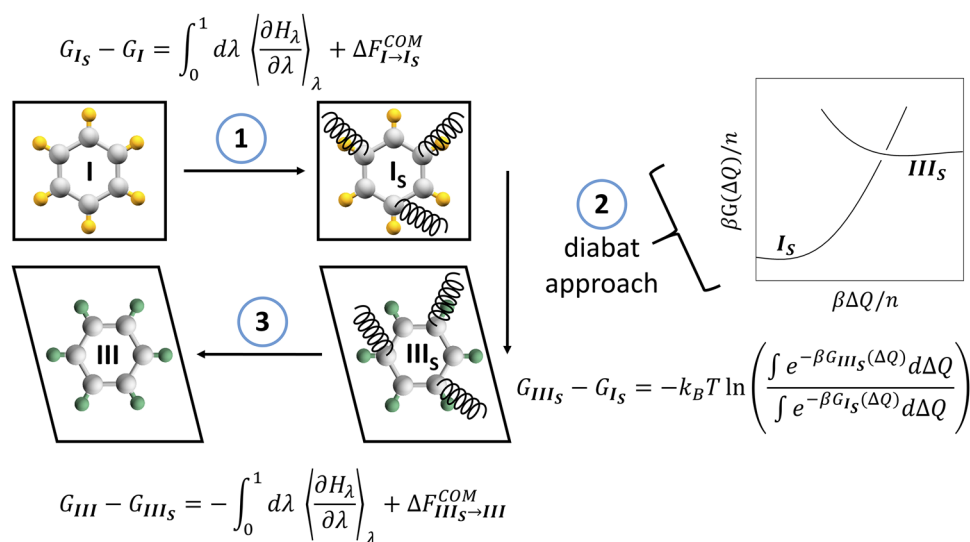


FIG. 6. Two molecules in phase I are one-to-one mapped to generate the corresponding lattice-switch configuration in phase III. The shift in the center of mass of the molecules in phase I is applied into phase III, as described in Sec. S.1 of the [supplementary material](#). The collective translation motion in phase I is an energetically unfavorable configuration in phase III.

by using a custom implementation of the Langevin algorithm (see Sec. S.5 for a discussion on the cause and a fix for the problem). The histograms obtained from umbrella sampling are stitched using WHAM<sup>79–81</sup> followed by unweighting to remove the effect of the preliminary bias. As shown in Fig. 8, the solid points are obtained from using the Zwanzig–Bennett relation on the umbrella sampled solid lines. Clearly, near the intersection region, the solid lines pass through the solid points, validating the method. Boltzmann exponentiation of the free energy curves followed by numerical integration over the order parameter and taking the logarithm yields the free energy difference. The calculated free energy difference between the crystal structures  $I_S$  and  $III_S$  (at 250 K and 1 bar) is  $0.1931 \pm 0.0002$  kcal/mol. We perform two very different consistency checks for the obtained result. We umbrella sample the Landau free energies [ $G_A(\Delta Q)$  and  $G_B(\Delta Q)$ ] without using preliminary weights (which requires more windows) to obtain a free energy difference of  $0.19299 \pm 0.00007$  kcal/mol. We also dissolved the non-bonded and bonded interactions in the structures  $I_S$  and  $III_S$  to obtain an Einstein crystal from both phases. The estimated value of the free energy difference from this route is  $0.197 \pm 0.004$  kcal/mol (see the [supplementary material](#) for details). Clearly, the free energy differences calculated in three different ways are in good agreement with each other. The contributions from paths 1 and 3 must be added to this result to get the total Gibbs free energy difference between



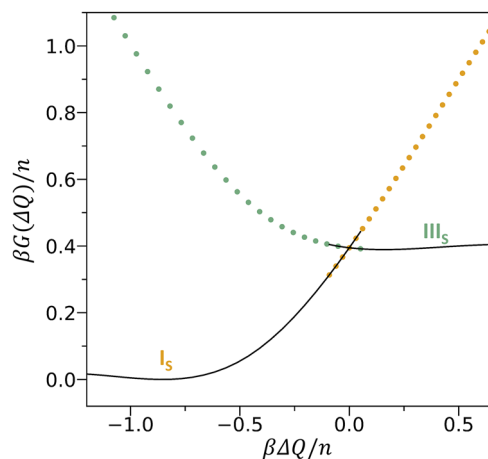
**FIG. 7.** The harmonic restraints are turned on in path 1. The diabatic method is used to calculate the free energy differences along path 2. The harmonic restraints are then reversibly removed. We only show springs on some atoms for clarity. Relevant equations for the individual paths are summarized above (see the [supplementary material](#) for details). The diabatic approach yields umbrella sampled free energy surfaces [ $G(\Delta Q)$  vs  $\Delta Q$  for  $I_s$  and  $III_s$ ], as discussed in [Fig. 8](#). The free energy difference between crystal structures  $I_s$  and  $III_s$  is obtained by using numerical integration to calculate the quantity inside the logarithm.

polymorphs **I** and **III** (see [Fig. 7](#) for a summary of equations used in paths 1–3). The values are  $0.153 \pm 0.005$  kcal/mol and  $0.157 \pm 0.006$  kcal/mol calculated from the diabatic method and Frenkel–Ladd method, respectively (benzene form **I** is more stable than **III**). Paths 1 and 3 are common to the diabatic method and the Frenkel–Ladd method and used a combined simulation time of 80 ns. The simulation time used for path 2 in the diabatic method is 5–7 ns and 115 ns using the Einstein crystal method (see Sec. S.7 in the [supplementary material](#) for simulation details). The total simulation time (on adding paths 1–3) is 87 ns for the diabatic method compared

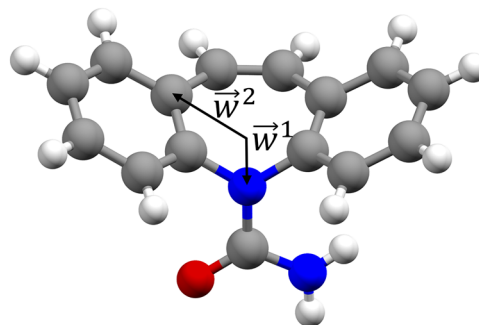
to the 195 ns used in the Einstein crystal method. Clearly, the simulation time required for path 2 is significantly lowered but the creation and dissolution of springs in paths 1 and 3 are bottlenecks for further efficiency gains.

## B. Carbamazepine

Carbamazepine ([Fig. 9](#)) has five reported polymorphs in the Cambridge Structural Database.<sup>82,83</sup> In forms **I–IV**, carbamazepine assembles in hydrogen-bonded dimer motifs, whereas form **V** is a catemeric<sup>84,85</sup> polymorph where the carboxamide group hydrogen bonds to two different neighbors. The novel form **V** polymorph has only been produced by templating carbamazepine crystal growth on an isostructural carbamazepine analog crystal.<sup>86</sup> The stability order of the well-studied forms **I–IV** at high temperatures follows the rank order of melting: **I** > **IV** > **III** > **II**.<sup>87</sup> This differs from the stability order at 0 K, which is approximated from melting enthalpy data: **III** > **I** > **IV** > **II**.<sup>87</sup> Thus, form **III** is thermodynamically the most stable polymorph at lower temperatures in agreement with lattice



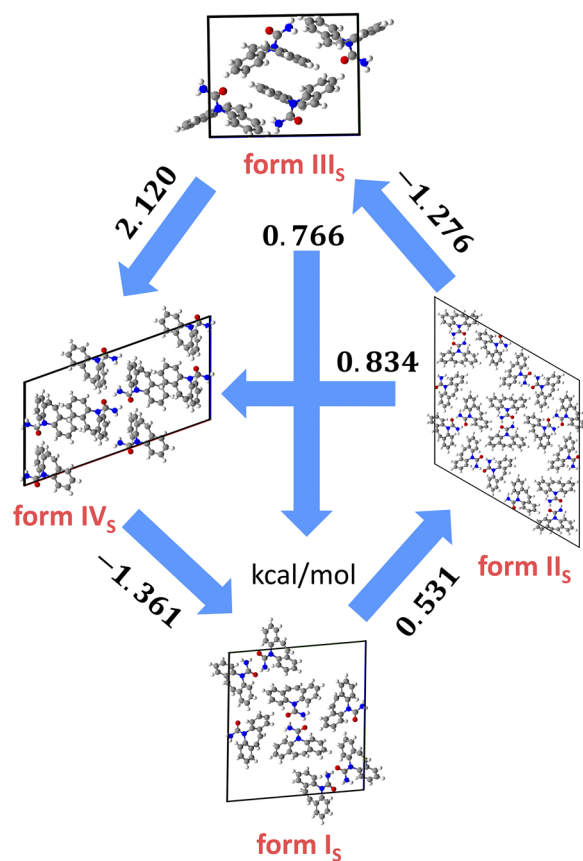
**FIG. 8.** The free energy curves (diabats) as a function of the order parameter for benzene crystals  $I_s$  and  $III_s$ . The axes are dimensionless. The black solid lines are obtained from umbrella sampling along  $\Delta Q$ . The solid points (shown as yellow and green dots) are obtained by using the Zwanzig–Bennett relation [Eq. (6)], i.e., by subtracting/adding the  $\Delta Q$ -coordinate to the points on the solid line. The figure clearly shows that near the intersection of the two free energy surfaces, the solid lines perfectly pass through the points theoretically predicted.  $n$  represents number of molecules (400 for benzene). The temperature is 250 K.



**FIG. 9.** The two vectors from the center of mass of the carbamazepine molecule used to construct the local coordinate system (using the same procedure as described previously).

energy calculations at 0 K.<sup>88</sup> Form **III** is enantiotropically related to form **I** (and also form **IV**) with the second most stable polymorph at 0 K, form **I**, being the most stable form above 351 K.<sup>89</sup>

We study all the forms at 300 K and 1 atm using an AMBER-GAFF forcefield.<sup>90,91</sup> We use a simple, force-field independent thermodynamic test to estimate error bars associated with our sampling procedures. We first calculate pairwise free energy differences between the polymorphs.  $N$  polymorphs have  $N!/((2!(N-2)!))$  diabat pairs. The combinatorial costs grow quickly, so we completed this analysis for only the first four polymorphs to reduce the computational effort (see the [supplementary material](#) for a partial set of calculations that include form **V**). From the pairwise free energy differences, we construct closed paths (see [Fig. 10](#)) starting/ending at **I<sub>S</sub>** with three forms in each path. For four forms, there are a total of three such paths that start/end at **I<sub>S</sub>**. The free energy differences are state functions and theoretically must be zero in a closed loop, regardless of the force field used. This allows us to use the residuals to estimate the error bars associated with sampling during the diabat calculations. The theoretical mean for the distribution of residual error in free energy changes is 0.0 kcal/mol. The total free energy

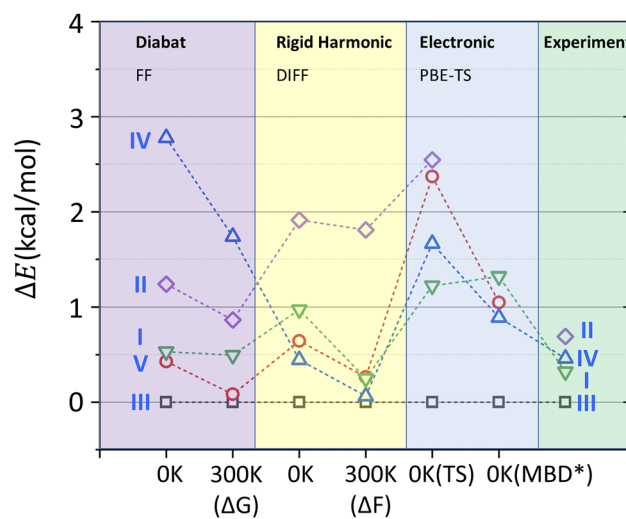


**FIG. 10.** Pairwise free energy differences in kcal/mol for polymorphs **I<sub>S</sub>**–**IV<sub>S</sub>** of carbamazepine calculated at 300 K and 1 atm using the diabat method (path 2 in [Fig. 7](#)). The harmonic spring restraints acting on all the atoms in the crystals are omitted for clarity, and their thermodynamic effects counteracted in paths 1 and 3 (see [Fig. 7](#)).

changes calculated for the three closed loops are 0.021 kcal/mol, 0.004 kcal/mol, and  $-0.007$  kcal/mol, respectively.

We first calculate the standard deviation ( $s_x$ ) for the three residual free energies (using a mean 0) obtained after completing the circuits. The standard deviation associated with the mean is  $s_x/\sqrt{3}$ , which yields 0.007 kcal/mol. This is the standard error associated with our pairwise free energy differences.

The harmonic restraints on the crystals are then dissolved (see the [supplementary material](#)) and free energies for carbamazepine polymorphs reported with respect to form **III**. The resulting relative Gibbs free energies for carbamazepine polymorphs **I**, **II**, **IV**, and **V** at  $P = 1$  atm and  $T = 300$  K are 0.494 kcal/mol, 0.862 kcal/mol, 1.738 kcal/mol, and 0.082 kcal/mol, respectively, with a standard error of  $\pm 0.007$  kcal/mol. This error estimate does not include error from the AMBER-GAFF force field. We predict that at 300 K and 1 atm, form **III** is the most stable, followed by **V**, **I**, **II**, and **IV**. The relative free energy for form **I** estimated from experimental solubility data<sup>89,97</sup> is  $0.08 \pm 0.02$  kcal/mol (see Sec. S.10 in the [supplementary material](#) for details). The relative polymorphic energy differences for a static lattice (0 K) and at 300 K are compared with other estimates in [Fig. 11](#) and Sec. S.11 in the [supplementary material](#). Temperature reduces the metastability of all forms relative to form **III** in both the diabat method and when estimated by rigid-molecule lattice dynamics but to different extents. More accurate estimates of the underlying potential energy surface, either based on the molecular charge density or periodic dispersion-corrected density functional electronic structure calculations, suggests that this is the main



**FIG. 11.** Stability of carbamazepine polymorphs from static (0 K) lattice energy differences and at 300 K calculated via different methods. The diabat method (this work) uses the AMBER-GAFF force field. The rigid molecule distributed intermolecular force-field calculations (DIFF) uses the rigid-molecule harmonic lattice phonons<sup>92</sup> and lattice energy calculated using DMACRYS.<sup>93</sup> The electronic lattice energy calculations correspond to the crystal structures optimized using CASTEP<sup>94</sup> with the PBE functional and TS dispersion correction,<sup>95</sup> with the energy reevaluated with the MBD\* dispersion correction.<sup>96</sup> The last data points are the relative experimental stability inferred from enthalpy of melting data.<sup>87</sup>



source of error in the estimated free energy differences. Additionally, stability inferred for crystals at finite temperatures from energy minimized 0 K structures depends on the value of the lattice energy at those minima, and it is common for organic molecules to have multiple lattice energy minima that correspond to the same finite temperature structure.<sup>98,99</sup> The experimental study provides  $4!/(2!(4 - 2!)) = 6$  pairwise free energy rankings, e.g., **III** more stable than **I**, **IV** more stable than **II**, etc. The AMBER-GAFF force field in combination with our diabatic approach correctly ranks the stability in five of these six pairs. One pair is qualitatively wrong (we mistakenly predict that **II** is more stable than **IV**). The diabatic method proposed here can estimate the free energy differences using the dynamics of the molecule, as given by the potential energy surface (force field). For many molecular crystals, the dynamical motion can be very anharmonic, the thermal expansion is usually anisotropic, and the isolated molecule modes are modified by the intermolecular forces in a very polymorph specific manner. Thus, the diabatic method should give considerably more realistic free energies for many pharmaceutical crystals than harmonic lattice dynamics methods. However, in practice, the improvement in the physical basis of the calculation methodology may be less important than the accuracy of potential energy surfaces that can be used. This implementation of the diabatic method is sufficiently computationally feasible that it should be extendable to quite large, flexible pharmaceutical molecules, given the availability of a suitable force field. Additionally, it may not always be necessary to include paths 1 and 3 to avoid intermolecular overlaps. Indeed, these steps (the largest source of error in our calculation) were not needed in the previous calculations of Bridgwater and Quigley.<sup>60</sup> As computer power improves, the diabatic method could be applied to estimate relative free energies of many observed and hypothetical polymorphs generated by crystal structure prediction methods.<sup>20,100</sup> Although the LSMC and diabatic approaches are conceptually different from established methods, they are not intrinsically difficult to implement. With a few minor modifications to existing simulation packages, they might rival Frenkel–Ladd methods on accuracy, efficiency, and ease of implementation.

#### IV. CONCLUSION

We extended the diabatic method to enable the calculation of Gibbs free energy differences between polymorphs for molecular crystals. The diabatic method uses lattice-switch energy gaps and a Zwanzig–Bennett relation to compute free energy diabat between the two polymorphs. The new extensions for molecular crystals use a one-to-one Gram–Schmidt based mapping of displacements and harmonic restraints to limit the size of atom displacements. We used an unbiased simulation of each polymorph to create a preliminary bias potential. The free energy on the biased (and partly flattened) surface was then umbrella sampled to efficiently obtain the two free energy diabat. The same sampling and mapping procedures are readily applied to polymorphs of any molecular crystal.

The Gibbs free energy differences were calculated for polymorphs **I** and **III** of benzene and for a nearly complete set of paths between polymorphs of carbamazepine. For benzene, we compared the results to Frenkel–Ladd calculations. For carbamazepine, we

constructed multiple thermodynamic paths from form **I** to two other polymorphs and back to form **I** so that the accumulated free energy change should be zero. These calculations allowed us to assess the accuracy of the diabatic method and advanced sampling procedures with no contributions from force field error. Using 5 ns of the total simulation time for each pair of polymorphs, the error in computed free energies per polymorph pair is of the order 0.005 kcal/mol–0.007 kcal/mol for both the benzene and carbamazepine systems. Additionally, the procedure can be repeated at different temperatures to calculate relative free energy differences between a polymorphic pair to establish stability relationships (as a function of temperature).

For carbamazepine, the AMBER-GAFF force field results also correctly ranked the stability (at 300 K and 1 atm) in five of six polymorph pairs. Through continued development of diabatic free energy tools, we hope to make the sampling aspects of polymorph free energy calculations so efficient that they can (in the future) be performed with state-of-the-art *ab initio* molecular dynamics<sup>101–103</sup> and path integrals for nuclear quantum effects.<sup>104–107</sup>

#### SUPPLEMENTARY MATERIAL

The [supplementary material](#) contains details on the construction of the local coordinate system, a discussion on the spring energies in the mapped configurations, computational details, a discussion on umbrella sampling using biasing forces, Langevin thermostatting combined with periodic restarts, diabat from pairwise calculations for benzene and carbamazepine polymorphs, thermodynamic integration calculations, relative lattice energies for carbamazepine polymorphs, a discussion on the contribution to the free energy from volume fluctuations, calculation of relative stability of carbamazepine polymorphs (forms **I** and **III**) from experimental solubility data, and computational details for the different *ab initio* methods discussed in [Fig. 11](#).

#### ACKNOWLEDGMENTS

We thank Eli Lilly and Company for funding this work through the Digital Design 2020 research grant. We thank Mike Doherty and Matteo Salvalaglio for helpful discussions and Jamshed Anwar and James Carruthers at Lancaster University for their help with the carbamazepine force field.

#### DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

#### REFERENCES

- W. E. Addison, *The Allotropy of the Elements* (American Elsevier Publishing Company, 1964).
- A. Putnis, *An Introduction to Mineral Sciences* (Cambridge University Press, 1992).
- J. Bernstein, *Polymorphism in Molecular Crystals* (Oxford University Press, 2002), Vol. 14.

- <sup>4</sup>W. Yan, B. Chen, S. M. Mahurin, V. Schwartz, D. R. Mullins, A. R. Lupini, S. J. Pennycook, S. Dai, and S. H. Overbury, "Preparation and comparison of supported gold nanocatalysts on anatase, brookite, rutile, and P<sub>25</sub> polymorphs of TiO<sub>2</sub> for catalytic oxidation of CO," *J. Phys. Chem. B* **109**(21), 10676–10685 (2005).
- <sup>5</sup>D. M. Robinson, Y. B. Go, M. Mui, G. Gardner, Z. Zhang, D. Mastrogianni, E. Garfunkel, J. Li, M. Greenblatt, and G. C. Dismukes, "Photochemical water oxidation by crystalline polymorphs of manganese oxides: Structural requirements for catalysis," *J. Am. Chem. Soc.* **135**(9), 3494–3501 (2013).
- <sup>6</sup>W. L. Kwong, C. C. Lee, A. Shchukarev, E. Björn, and J. Messinger, "High-performance iron (III) oxide electrocatalyst for water oxidation in strongly acidic media," *J. Catal.* **365**, 29–35 (2018).
- <sup>7</sup>B. Zheng, W. Hua, Y. Yue, and Z. Gao, "Dehydrogenation of propane to propene over different polymorphs of gallium oxide," *J. Catal.* **232**(1), 143–151 (2005).
- <sup>8</sup>X. Chen, Y. Luo, J. Zhang, K. Jiang, J. B. Pendry, and S. Zhang, "Macroscopic invisibility cloaking of visible light," *Nat. Commun.* **2**(1), 176 (2011).
- <sup>9</sup>T. Wonglakhon and D. Zahn, "Interaction potentials for modelling GaN precipitation and solid state polymorphism," *J. Phys.: Condens. Matter* **32**(20), 205401 (2020).
- <sup>10</sup>K. Wang, H. Zhang, S. Chen, G. Yang, J. Zhang, W. Tian, Z. Su, and Y. Wang, "Organic polymorphs: One-compound-based crystals with molecular-conformation- and packing-dependent luminescent properties," *Adv. Mater.* **26**(35), 6168–6173 (2014).
- <sup>11</sup>T. Beyer, G. M. Day, and S. L. Price, "The prediction, morphology, and mechanical properties of the polymorphs of paracetamol," *J. Am. Chem. Soc.* **123**(21), 5086–5094 (2001).
- <sup>12</sup>C. Sun and D. J. Grant, "Influence of crystal structure on the tableting properties of sulfamerazine polymorphs," *Pharm. Res.* **18**(3), 274–280 (2001).
- <sup>13</sup>R. Censi and P. Di Martino, "Polymorph impact on the bioavailability and stability of poorly soluble drugs," *Molecules* **20**(10), 18759–18776 (2015).
- <sup>14</sup>C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings," *Adv. Drug Delivery Rev.* **23**(1), 3–25 (1997).
- <sup>15</sup>J. Aaltonen, M. Alleso, S. Mirza, V. Koradia, K. Gordon, and J. Rantanen, "Solid form screening—A review," *Eur. J. Pharm. Biopharm.* **71**(1), 23–37 (2009).
- <sup>16</sup>G. R. Desiraju, "Crystal engineering: A brief overview," *J. Chem. Sci.* **122**(5), 667–675 (2010).
- <sup>17</sup>S. Price, "The computational prediction of pharmaceutical crystal structures and polymorphism," *Adv. Drug Delivery Rev.* **56**(3), 301–319 (2004).
- <sup>18</sup>S. L. Price, D. E. Braun, and S. M. Reutzel-Edens, "Can computed crystal energy landscapes help understand pharmaceutical solids?," *Chem. Commun.* **52**(44), 7065–7077 (2016).
- <sup>19</sup>J. Kendrick, F. J. J. Leusen, M. A. Neumann, and J. van de Streek, "Progress in crystal structure prediction," *Chem. Eur. J.* **17**(38), 10736–10744 (2011).
- <sup>20</sup>G. M. Day, "Current approaches to predicting molecular organic crystal structures," *Crystallogr. Rev.* **17**(1), 3–52 (2011).
- <sup>21</sup>C. C. Pantelides, C. S. Adjiman, and A. V. Kazantsev, "General computational algorithms for *ab initio* crystal structure prediction for organic molecules, in *Prediction and Calculation of Crystal Structures* (Springer, 2014), pp. 25–58.
- <sup>22</sup>G. J. O. Beran, "Modeling polymorphic molecular crystals with electronic structure theory," *Chem. Rev.* **116**(9), 5567–5613 (2016).
- <sup>23</sup>T.-Q. Yu and M. E. Tuckerman, "Temperature-accelerated method for exploring polymorphism in molecular crystals based on free energy," *Phys. Rev. Lett.* **107**(1), 015701 (2011).
- <sup>24</sup>B. Peters, "Competing nucleation pathways in a mixture of oppositely charged colloids: Out-of-equilibrium nucleation revisited," *J. Chem. Phys.* **131**(24), 244103 (2009).
- <sup>25</sup>N. Duff, Y. R. Dahal, J. D. Schmit, and B. Peters, "Salting out the polar polymorph: Analysis by alchemical solvent transformation," *J. Chem. Phys.* **140**(1), 014501 (2014).
- <sup>26</sup>M. Salvalaglio, C. Perego, F. Giberti, M. Mazzotti, and M. Parrinello, "Molecular-dynamics simulations of urea nucleation from aqueous solution," *Proc. Natl. Acad. Sci. U. S. A.* **112**(1), E6–E14 (2015).
- <sup>27</sup>G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen, and A. Michaelides, "Crystal nucleation in liquids: Open questions and future challenges in molecular dynamics simulations," *Chem. Rev.* **116**(12), 7078–7116 (2016).
- <sup>28</sup>M. N. Joswiak, B. Peters, and M. F. Doherty, "*In silico* crystal growth rate prediction for NaCl from aqueous solution," *Cryst. Growth Des.* **18**(10), 6302–6306 (2018).
- <sup>29</sup>D. Han, T. Karmakar, Z. Bjelobrk, J. Gong, and M. Parrinello, "Solvent-mediated morphology selection of the active pharmaceutical ingredient isoniazid: Experimental and simulation studies," *Chem. Eng. Sci.* **204**, 320–328 (2019).
- <sup>30</sup>M. Salvalaglio, T. Vetter, M. Mazzotti, and M. Parrinello, "Controlling and predicting crystal shapes: The case of urea," *Angew. Chem., Int. Ed.* **52**(50), 13369–13372 (2013).
- <sup>31</sup>A. J. Cruz-Cabeza, S. M. Reutzel-Edens, and J. Bernstein, "Facts and fictions about polymorphism," *Chem. Soc. Rev.* **44**(23), 8619–8635 (2015).
- <sup>32</sup>J. Nyman and G. M. Day, "Static and lattice vibrational energy differences between polymorphs," *CrystEngComm* **17**(28), 5154–5165 (2015).
- <sup>33</sup>S. R. Whittleton, A. Otero-de-la-Roza, and E. R. Johnson, "Exchange-hole dipole dispersion model for accurate energy ranking in molecular crystal structure prediction," *J. Chem. Theory Comput.* **13**(2), 441–450 (2017).
- <sup>34</sup>J. G. Brandenburg, T. Maas, and S. Grimme, "Benchmarking DFT and semiempirical methods on structures and lattice energies for ten ice polymorphs," *J. Chem. Phys.* **142**(12), 124104 (2015).
- <sup>35</sup>K. Hongo, M. A. Watson, R. S. Sánchez-Carrera, T. Iitaka, and A. Aspuru-Guzik, "Failure of conventional density functionals for the prediction of molecular crystal polymorphism: A quantum Monte Carlo study," *J. Phys. Chem. Lett.* **1**(12), 1789–1794 (2010).
- <sup>36</sup>J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, and G. K.-L. Chan, "*Ab initio* determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy," *Science* **345**(6197), 640–643 (2014).
- <sup>37</sup>G. J. O. Beran, "A new era for *ab initio* molecular crystal lattice energy prediction," *Angew. Chem., Int. Ed.* **54**(2), 396–398 (2015).
- <sup>38</sup>J. Hoja, H.-Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio, and A. Tkatchenko, "Reliable and practical computational description of molecular crystal polymorphs," *Sci. Adv.* **5**(1), eaau3338 (2019).
- <sup>39</sup>A. M. Reilly and A. Tkatchenko, "Role of dispersion interactions in the polymorphism and entropic stabilization of the aspirin crystal," *Phys. Rev. Lett.* **113**(5), 055701 (2014).
- <sup>40</sup>J. G. Brandenburg, J. Potticary, H. A. Sparkes, S. L. Price, and S. R. Hall, "Thermal expansion of carbamazepine: Systematic crystallographic measurements challenge quantum chemical calculations," *J. Phys. Chem. Lett.* **8**(17), 4319–4324 (2017).
- <sup>41</sup>J. L. McKinley and G. J. O. Beran, "Identifying pragmatic quasi-harmonic electronic structure approaches for modeling molecular crystal thermal expansion," *Faraday Discuss.* **211**, 181–207 (2018).
- <sup>42</sup>N. S. Abraham and M. R. Shirts, "Statistical mechanical approximations to more efficiently determine polymorph free energy differences for small organic molecules," *J. Chem. Theory Comput.* **16**(10), 6503–6512 (2020).
- <sup>43</sup>D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Elsevier, 2001).
- <sup>44</sup>I. J. Nessler, J. M. Litman, and M. J. Schnieders, "Toward polarizable AMOEBA thermodynamics at fixed charge efficiency using a dual force field approach: Application to organic crystals," *Phys. Chem. Chem. Phys.* **18**(44), 30313–30322 (2016).
- <sup>45</sup>E. Schneider, L. Vogt, and M. E. Tuckerman, "Exploring polymorphism of benzene and naphthalene with free energy based enhanced molecular dynamics," *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **72**(4), 542–550 (2016).
- <sup>46</sup>E. C. Dybeck, N. S. Abraham, N. P. Schieber, and M. R. Shirts, "Capturing entropic contributions to temperature-mediated polymorphic transformations through molecular modeling," *Cryst. Growth Des.* **17**(4), 1775–1787 (2017).
- <sup>47</sup>D. Frenkel and A. J. C. Ladd, "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.* **81**(7), 3188–3193 (1984).
- <sup>48</sup>C. Vega and E. G. Noya, "Revisiting the Frenkel–Ladd method to compute the free energy of solids: The Einstein molecule approach," *J. Chem. Phys.* **127**(15), 154113 (2007).

- <sup>49</sup>M. Yang, E. Dybeck, G. Sun, C. Peng, B. Samas, V. M. Burger, Q. Zeng, Y. Jin, M. A. Bellucci, Y. Liu *et al.*, "Prediction of the relative free energies of drug polymorphs above zero Kelvin," *Cryst. Growth Des.* **20**(8), 5211–5224 (2020).
- <sup>50</sup>R. Martoňák, D. Donadio, A. R. Oganov, and M. Parrinello, "Crystal structure transformations in SiO<sub>2</sub> from classical and *ab initio* metadynamics," *Nat. Mater.* **5**(8), 623–626 (2006).
- <sup>51</sup>P. Raiteri, R. Martoňák, and M. Parrinello, "Exploring polymorphism: The case of benzene," *Angew. Chem., Int. Ed.* **44**(24), 3769–3773 (2005).
- <sup>52</sup>A. D. Bruce, N. B. Wilding, and G. J. Ackland, "Free energy of crystalline solids: A lattice-switch Monte Carlo method," *Phys. Rev. Lett.* **79**(16), 3002 (1997).
- <sup>53</sup>A. D. Bruce, A. N. Jackson, G. J. Ackland, and N. B. Wilding, "Lattice-switch Monte Carlo method," *Phys. Rev. E* **61**(1), 906 (2000).
- <sup>54</sup>K. Kamat and B. Peters, "Diabat interpolation for polymorph free-energy differences," *J. Phys. Chem. Lett.* **8**(3), 655–660 (2017).
- <sup>55</sup>K. Kamat and B. Peters, "Gibbs free-energy differences between polymorphs via a diabat approach," *J. Chem. Phys.* **149**(21), 214106 (2018).
- <sup>56</sup>R. W. Zwanzig, "High-temperature equation of state by a perturbation method. I. Nonpolar gases," *J. Chem. Phys.* **22**(8), 1420–1426 (1954).
- <sup>57</sup>C. H. Bennett, "Efficient estimation of free energy differences from Monte Carlo data," *J. Comput. Phys.* **22**(2), 245–268 (1976).
- <sup>58</sup>A. N. Jackson, A. D. Bruce, and G. J. Ackland, "Lattice-switch Monte Carlo method: Application to soft potentials," *Phys. Rev. E* **65**(3), 036710 (2002).
- <sup>59</sup>T. L. Underwood and G. J. Ackland, "Monteswitch: A package for evaluating solid–solid free energy differences via lattice-switch Monte Carlo," *Comput. Phys. Commun.* **215**, 204–222 (2017).
- <sup>60</sup>S. Bridgwater and D. Quigley, "Lattice-switching Monte Carlo method for crystals of flexible molecules," *Phys. Rev. E* **90**(6), 063313 (2014).
- <sup>61</sup>S. Bridgwater, "Accurate free energy methods for model organic solids," Ph.D. thesis, University of Warwick, 2014.
- <sup>62</sup>N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.* **21**(6), 1087–1092 (1953).
- <sup>63</sup>G. R. Smith and A. D. Bruce, "A study of the multi-canonical Monte Carlo method," *J. Phys. Math. Gen.* **28**(23), 6623 (1995).
- <sup>64</sup>M. Fitzgerald, R. R. Picard, and R. N. Silver, "Canonical transition probabilities for adaptive metropolis simulation," *Europhys. Lett.* **46**(3), 282 (1999).
- <sup>65</sup>J. K. Hwang and A. Warshel, "Microscopic examination of free-energy relationships for electron transfer in polar solvents," *J. Am. Chem. Soc.* **109**(3), 715–720 (1987).
- <sup>66</sup>J. Blumberger, I. Tavernelli, M. L. Klein, and M. Sprik, "Diabatic free energy curves and coordination fluctuations for the aqueous Ag<sup>+</sup>/Ag<sup>2+</sup> redox couple: A biased Born–Oppenheimer molecular dynamics investigation," *J. Chem. Phys.* **124**(6), 064507 (2006).
- <sup>67</sup>M. Tachiya, "Relation between the electron-transfer rate and the free energy change of reaction," *J. Phys. Chem.* **93**(20), 7050–7052 (1989).
- <sup>68</sup>D. W. Small, D. V. Matyushov, and G. A. Voth, "The theory of electron transfer reactions: What may be missing?," *J. Am. Chem. Soc.* **125**(24), 7470–7478 (2003).
- <sup>69</sup>Y. Tateyama, J. Blumberger, M. Sprik, and I. Tavernelli, "Density-functional molecular-dynamics study of the redox reactions of two anionic, aqueous transition-metal complexes," *J. Chem. Phys.* **122**(23), 234505 (2005).
- <sup>70</sup>M. M. Thiéry and J. M. Léger, "High pressure solid phases of benzene. I. Raman and X-ray studies of C<sub>6</sub>H<sub>6</sub> at 294 K up to 25 GPa," *J. Chem. Phys.* **89**(7), 4255–4271 (1988).
- <sup>71</sup>T. Shoda, K. Yamahara, K. Okazaki, and D. E. Williams, "Molecular packing analysis of benzene crystals. Part 2. Prediction of experimental crystal structure polymorphs at low and high pressure," *J. Mol. Struct.: THEOCHEM* **333**(3), 267–274 (1995).
- <sup>72</sup>F. Cansell, D. Fabre, and J. P. Petit, "Phase transitions and chemical transformations of benzene up to 550 °C and 30 GPa," *J. Chem. Phys.* **99**(10), 7300–7304 (1993).
- <sup>73</sup>W. L. Jorgensen and J. Tirado-Rives, "The OPLS (optimized potentials for liquid simulations) potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin," *J. Am. Chem. Soc.* **110**(6), 1657–1666 (1988).
- <sup>74</sup>L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives, and W. L. Jorgensen, "Lig-ParGen web server: An automatic OPLS-AA parameter generator for organic ligands," *Nucleic Acids Res.* **45**(W1), W331–W336 (2017).
- <sup>75</sup>A. Budzianowski and A. Katrusiak, "Pressure-frozen benzene I revisited," *Acta Crystallogr., Sect. B: Struct. Sci.* **62**(1), 94–101 (2006).
- <sup>76</sup>G. J. Piermarini, A. D. Mighell, C. E. Weir, and S. Block, "Crystal structure of benzene II at 25 kilobars," *Science* **165**(3899), 1250–1255 (1969).
- <sup>77</sup>C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, and P. A. Wood, "Mercury CSD 2.0—New features for the visualization and investigation of crystal structures," *J. Appl. Crystallogr.* **41**(2), 466–470 (2008).
- <sup>78</sup>S. J. Leon, I. Bica, and T. Hohn, *Linear Algebra with Applications* (Pearson Prentice Hall Upper, Saddle River, NJ, USA, 2006).
- <sup>79</sup>S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method," *J. Comput. Chem.* **13**(8), 1011–1021 (1992).
- <sup>80</sup>B. Roux, "The calculation of the potential of mean force using computer simulations," *Comput. Phys. Commun.* **91**(1–3), 275–282 (1995).
- <sup>81</sup>A. Grossfield, "WHAM: The weighted histogram analysis method," Version 2(9), 06 (2012).
- <sup>82</sup>F. H. Allen, "The Cambridge structural Database: A quarter of a million crystal structures and rising," *Acta Crystallogr., Sect. B: Struct. Sci.* **58**(3), 380–388 (2002).
- <sup>83</sup>C. R. Groom and F. H. Allen, "The Cambridge structural database in retrospect and prospect," *Angew. Chem., Int. Ed.* **53**(3), 662–671 (2014).
- <sup>84</sup>L. D'Ascenzo and P. Auffinger, "A comprehensive classification and nomenclature of carboxyl–carboxyl(Ate) supramolecular motifs and related catemers: Implications for biomolecular systems," *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **71**(Pt 2), 164–175 (2015).
- <sup>85</sup>T. Beyer and S. L. Price, "Dimer or catemer? Low-energy crystal packings for small carboxylic acids," *J. Phys. Chem. B* **104**(12), 2647–2655 (2000).
- <sup>86</sup>J.-B. Arlin, L. S. Price, S. L. Price, and A. J. Florence, "A strategy for producing predicted polymorphs: Catemeric carbamazepine form V," *Chem. Commun.* **47**(25), 7074–7076 (2011).
- <sup>87</sup>A. L. Grzesiak, M. Lang, K. Kim, and A. J. Matzger, "Comparison of the four anhydrous polymorphs of carbamazepine and the crystal structure of form I," *J. Pharm. Sci.* **92**(11), 2260–2271 (2003).
- <sup>88</sup>A. J. Florence, A. Johnston, S. L. Price, H. Nowell, A. R. Kennedy, and N. Shankland, "An automated parallel crystallisation search for predicted crystal structures and packing motifs of carbamazepine," *J. Pharm. Sci.* **95**(9), 1918–1930 (2006).
- <sup>89</sup>K. Park, J. M. B. Evans, and A. S. Myerson, "Determination of solubility of polymorphs using differential scanning calorimetry," *Cryst. Growth Des.* **3**(6), 991–995 (2003).
- <sup>90</sup>J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *J. Comput. Chem.* **25**(9), 1157–1174 (2004).
- <sup>91</sup>J. Wang, W. Wang, P. A. Kollman, and D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations," *J. Mol. Graphics Modell.* **25**(2), 247–260 (2006).
- <sup>92</sup>J. Nyman and G. M. Day, "Modelling temperature-dependent properties of polymorphic organic molecular crystals," *Phys. Chem. Chem. Phys.* **18**(45), 31132–31143 (2016).
- <sup>93</sup>S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis, and G. M. Day, "Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials," *Phys. Chem. Chem. Phys.* **12**(30), 8478–8490 (2010).
- <sup>94</sup>S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. Probert, K. Refson, and M. C. Payne, "First principles methods using CASTEP," *Z. Kristallogr.-Cryst. Mater.* **220**(5–6), 567–570 (2005).
- <sup>95</sup>A. Tkatchenko and M. Scheffler, "Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data," *Phys. Rev. Lett.* **102**(7), 073005 (2009).
- <sup>96</sup>A. Ambrosetti, A. M. Reilly, R. A. DiStasio, Jr., and A. Tkatchenko, "Long-range correlation energy calculated from coupled atomic response functions," *J. Chem. Phys.* **140**(18), 18A508 (2014).

- <sup>97</sup>R. J. Behme and D. Brooke, "Heat of fusion measurement of a low melting polymorph of carbamazepine that undergoes multiple-phase changes during differential scanning calorimetry analysis," *J. Pharm. Sci.* **80**(10), 986–990 (1991).
- <sup>98</sup>E. C. Dybeck, D. P. McMahon, G. M. Day, and M. R. Shirts, "Exploring the multi-minima behavior of small molecule crystal polymorphs at finite temperature," *Cryst. Growth Des.* **19**(10), 5568–5580 (2019).
- <sup>99</sup>N. F. Francia, L. S. Price, J. Nyman, S. L. Price, and M. Salvalaglio, "Systematic finite-temperature reduction of crystal energy landscapes," *Cryst. Growth Des.* **20**(10), 6847–6862 (2020).
- <sup>100</sup>S. L. Price, "Is zeroth order crystal structure prediction (CSP\_0) coming to maturity? What should we aim for in an ideal crystal structure prediction code?," *Faraday Discuss.* **211**, 9–30 (2018).
- <sup>101</sup>L. R. Pestana, N. Mardirossian, M. Head-Gordon, and T. Head-Gordon, "Ab initio molecular dynamics simulations of liquid water using high quality meta-GGA functionals," *Chem. Sci.* **8**(5), 3554–3565 (2017).
- <sup>102</sup>T. Cheng, H. Xiao, and W. A. Goddard, "Full atomistic reaction mechanism with kinetics for CO reduction on Cu(100) from ab initio molecular dynamics free-energy calculations at 298 K," *Proc. Natl. Acad. Sci. U. S. A.* **114**(8), 1795–1800 (2017).
- <sup>103</sup>M. Ceriotti, J. More, and D. E. Manolopoulos, "A Python interface for ab initio path integral molecular dynamics simulations," *Comput. Phys. Commun.* **185**(3), 1019–1026 (2014).
- <sup>104</sup>M. Parrinello and A. Rahman, "Study of an F center in molten KCl," *J. Chem. Phys.* **80**(2), 860–867 (1984).
- <sup>105</sup>O. Marsalek and T. E. Markland, "Ab initio molecular dynamics with nuclear quantum effects at classical cost: Ring polymer contraction for density functional theory," *J. Chem. Phys.* **144**(5), 054112 (2016).
- <sup>106</sup>V. Kapil, J. VandeVondele, and M. Ceriotti, "Accurate molecular dynamics and nuclear quantum effects at low cost by multiple steps in real and imaginary time: Using density functional theory to accelerate wavefunction methods," *J. Chem. Phys.* **144**(5), 054111 (2016).
- <sup>107</sup>T. E. Markland and M. Ceriotti, "Nuclear quantum effects enter the mainstream," *Nat. Rev. Chem.* **2**(3), 0109 (2018).