JCTC Journal of Chemical Theory and Computation

Article

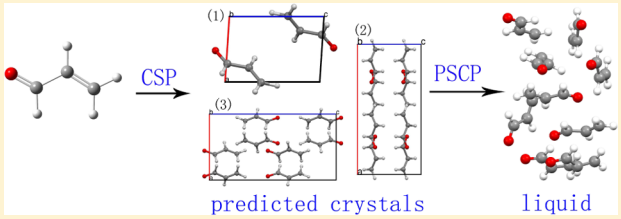
pubs.acs.org/JCTC

Toward Fully in Silico Melting Point Prediction Using Molecular Simulations

Yong Zhang and Edward J. Maginn*

Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, Indiana 46556, United States

ABSTRACT: Melting point is one of the most fundamental and practically important properties of a compound. Molecular simulation methods have been developed for the accurate computation of melting points. However, all of these methods need an experimental crystal structure as input, which means that such calculations are not really predictive since the melting point can be measured easily in experiments once a crystal structure is known. On the other hand, crystal structure prediction (CSP) has become an active field and significant progress has been made, although challenges still exist. One of the main challenges is the existence of many crystal structures (polymorphs) that are very close in energy. Thermal effects and kinetic factors make the situation even more complicated, such that it is still not trivial to predict experimental crystal structures. In this work, we exploit the fact that free energy differences are often small between crystal structures. We show that accurate melting point predictions can be made by using a reasonable crystal structure from CSP as a starting point for a free energy-based melting point calculation. The key is that most crystal structures predicted by CSP have free energies that are close to that of the experimental structure. The proposed method was tested on two rigid molecules and the results suggest that a fully in silico melting point prediction method is possible.



1. INTRODUCTION

Melting point (T_m) is one of the most fundamental properties for a compound. Efforts have been made to understand melting points and attempts have been carried out to predict them, ideally before the compounds have been synthesized.[1] The ability to accurately predict the melting point of a compound based only on its molecular formula can significantly accelerate the design of new materials. A particular example in this regard is ionic liquids (ILs), which have drawn extensive attention over the past decade due to their unique properties such as negligible vapor pressure, nonflammability, good ionic conductivity, etc.[2−6] By definition, ILs are a group of salts that have melting points below 100 °C, which clearly indicates that melting point is a key feature of this group of compounds.

Many methods have been developed for the computation of melting points, which were briefly reviewed in a recent paper.[1] Although correlations and group contribution methods have been used, to get accurate results for classes of compounds where little experimental data exist, one needs to utilize physics-based atomistic methods. These methods can be roughly categorized into two groups, the so-called "direct" methods and free-energy-based methods. For simple molecules such as argon, the results using all these methods tend to be consistent. For more complicated multiatomic molecules, however, free-energy-based methods are generally preferred.[1] Unfortunately, all these molecular simulation methods require the experimental crystal structure as input, which means the calculation is not really predictive as the melting point can be measured easily once the crystal structure is known.

In addition to melting points, molecular simulations are also used to predict various macroscopic and microscopic properties of molecules. "Are crystal structures predictable?" People has been asking this question for a long time. In 1994, Gavezzotti replied "No" and listed the challenges that one has to face in order to predict the crystal structure.[7] Since then, effort has been made to make crystal structure prediction possible. To encourage advances in the field, the Cambridge Crystallographic Data Centre (CCDC) launched the first crystal structure prediction (CSP) blind test in 1999.[8] Following that, blind tests were organized every 2−3 years, and the latest one, the fifth CSP blind test, was in 2010.[9] Significant progress can be effectively seen from the results of these blind tests[8−12] as well as recent reviews.[13−16] It is now believed that the "dawn of a new era" for crystal structure prediction is coming.[17]

In spite of these successes, challenges in CSP still exist, which lie mainly in two aspects. The first aspect is how to generate all possible crystal structures. This is not trivial especially for molecules with internal degrees of freedom. The other is to evaluate and rank the generated structures based on certain criteria to identify the most likely one.[12] Lattice energy is the widely used and most successful criterion.[14] These challenges and the fact that there are usually many possible crystal structures that are very close in energy[7,12,18−20] are stimulating further improvement in the field.[21]

The challenges in CSP can actually be exploited for melting point calculation or prediction. In thermodynamics, melting point is the temperature that the liquid phase and crystal phase share the same free energy. Crystal structures that are close in free energy should have similar melting points. So, in principle,

Received: December 13, 2012

   dx.doi.org/10.1021/ct301095j | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

if a melting point calculation is based on a crystal structure that is close to the experimental one in terms of free energy, but not necessarily the actual one or even close to it in structure, the calculated melting point should be a good estimate as long as the calculation procedure itself is reliable.

This idea is tested in this work on two rigid molecules for which both experimental crystal structures and the melting temperature were reported for the same structure. The first molecule is 2-propenal, target molecule XII in the fourth CSP blind test.[12] The second molecule is 4-cyano-pyridine N-oxide, an analog to target molecule XVI in the fifth CSP blind test[9] (see Figure 1). The two molecules will be referred to as mol I and mol II, respectively. To the best of our knowledge, this is the first time simulations have been used to both predict a crystal structure and subsequently compute a melting point. The proposed procedure and simulation details are described in the next section, followed by results and a discussion. A summary is provided in the last section.

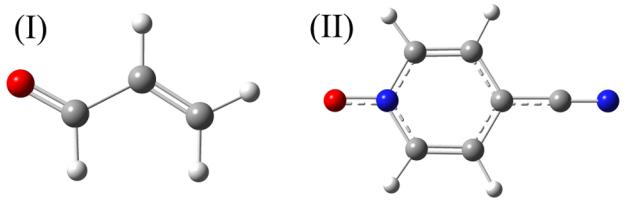


Figure 1. Optimized structure of the target molecules, 2-propenal (mol I) and 4-cyano-pyridine N-oxide (mol II), respectively, studied in the current work.

2. METHODOLOGY AND SIMULATION DETAILS

2.1. Crystal Structure Prediction. The crystal structures of the two target molecules were predicted using the packages MOLPAK[22] and PMIN by closely following the corresponding procedures. The structure of each target molecule was optimized at the B3LYP/6-31G(d) level. To mimic the electronic field due to the bulk environment, an implicit dimethylsulfoxide (DMSO) solvent was included with the polarizable continuum solvation model (PCM) using the integral equation formalism embedded in Gaussian 09.[23] The optimized structure of each target molecule was kept rigid throughout the crystal prediction procedure. The optimized structure was fed to MOLPAK and possible crystal structures were generated in a systematic manner. For each molecule, around 7000 packing structures (19^3 with 10 degree step search in each of the three directions) were considered in each of 31 common coordination types, which cover 13 of the most commonly observed space groups (including the known experimental crystal structures for both target molecules).

The 500 highest density MOLPAK packing structures were then passed to PMIN, which refines the crude structures by minimizing the lattice energy. No internal degrees of freedom are considered in PMIN and the intermolecular Coulombic interaction is based on atom-centered charge. The atomic partial charges were derived using the CHELPG method at the HF/6-31G(d) level. The 500 structures were ranked based on the lattice energy, and the one with the lowest lattice energy was taken as the most likely crystal structure.

It has been found that thermal effects at finite temperature, including the entropy contribution, can be important in the ranking of crystal structures generated in this manner.[18,24,25] So ideally, free energy at finite temperature should be used to evaluate the generated structures. However, practically, the accuracy of the calculated lattice energy highly depends on the applied force field.[21] For classical force fields such as the one used in the current work, the uncertainty may be of the same order of magnitude as the contribution from thermal effects.[26] For the molecules included in the blind tests, even with more accurate force fields, it has been found that the importance of the dynamic contributions is very limited and the potential energies calculated at various levels of theory have been widely used.[12] Kinetics of the crystallization process can play an important role as well, but it is usually not considered in the prediction methods. In the current work, no thermal effects or kinetic factors were considered in evaluating the generated crystal structures. Note that the term "predicted" is being only used to denote computationally generated low energy structures and not to imply that they are predicted polymorphs.

2.2. Melting Point Calculation. For each target molecule, the melting point calculation was carried out starting from the experimental crystal structure (exp.) and three most likely predicted crystal structures (pred.1, pred.2, and pred.3 in the order of increasing energy), respectively. Note that none of the predicted crystal structures were the experimental structure. The revised pseudosupercritical path (PSCP) method was used to calculate the melting point.[1,27]

Classical molecular dynamics simulations in the isothermal–isobaric (NPT) and canonical (NVT) ensembles were carried out using the LAMMPS package.[28] A time step of 1 fs was used in all simulations with periodic boundary conditions. The all-atom general Amber force field (GAFF)[29] was used to describe both intra- and intermolecular interactions. The atomic charges were derived by fitting the electrostatic potential surface of the optimized structure at the B3LYP/6-311++G(d,p) level using the RESP method.[30] Long range electrostatic interactions were calculated using the Ewald/n method.[31] A real space cutoff of 12 Å was applied. The temperature and pressure were controlled by the Nosé–Hoover thermostat[32] and the extended Lagrangian approach,[33] respectively. The pressure was fixed to be one atmosphere in all constant pressure simulations with isotropic volume fluctuations for the liquid phase and anisotropic cell for the crystalline phase.

The crystal phases of the simulation systems were set up by reproducing the corresponding crystal unit cell, in the X, Y, and Z directions, respectively. To ensure reasonable parallel simulation efficiency, the simulation box was chosen in such a way that the final simulation box is a close approximation to a cubic cell in each case. For mol I, the simulation box includes 768 (exp.), 864 (pred.1), 864 (pred.2), and 840 (pred.3) molecules, respectively. In all cases, 672 molecules were used in the mol II simulations. The liquid phase systems were prepared by putting the same number of molecules randomly in a cubic box as was used for the crystal phase. For both crystal phase and liquid phase simulations, the initial boxes were equilibrated for a time period of 2 ns in the NPT ensemble. NVT ensemble simulations were then carried out under the equilibrium densities calculated during the final 1 ns of the 2 ns NPT trajectories.

In the PSCP method, the free energy curves for the single phase systems were calculated based on the Gibbs–Helmholtz equation

$$\frac{G}{RT} - \left(\frac{G}{RT}\right)_{\mathrm{ref}} = \frac{G^{\mathrm{NPT}}}{RT} = \int_{T_{\mathrm{ref}}}^{T} -\frac{H^{\mathrm{NPT}}}{RT^2}\mathrm{d}T \qquad (1)$$

dx.doi.org/10.1021/ct301095j | J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

where T_{ref} is an arbitrary reference temperature and H^{NPT} is the enthalpy from NPT ensemble simulations. The absolute free energy difference between the liquid (L) and crystal (C) phases is

$$\Delta G(T) = G_L(T) - G_C(T) \quad (2)$$

$$=(G_L^{NPT}(T) + G_L(T_{ref})) - (G_C^{NPT}(T) + G_C(T_{ref})) \quad (3)$$

$$=(G_L^{NPT}(T) - G_C^{NPT}(T)) + \Delta G(T_{ref}) \quad (4)$$

where the first term in parentheses is obtained from application of eq 1 to the liquid and crystal phases and $\Delta G(T_{ref}) = G_L(T_{ref}) - G_C(T_{ref})$ is calculated along the PSCP cycle.[34] In the current work, the integration in eq 1 was carried out for a temperature range of 120−220 K for mol I and 450−550 K for mol II. The reference temperatures were chosen to be 200 and 540 K, respectively. Note that the choice of the reference temperature does not affect the calculation results.[35] Other parameters in the PSCP method closely followed those in our previous work.[34]

3. RESULTS

The experimental[12,36] and predicted crystal structures of mol I and mol II are shown in Figures 2 and 3, respectively. The crystal unit cell parameters as well as densities at experimental temperature or 0 K (for the predicted structures), space group, the Z parameter (number of molecules in the unit cell), and the lattice energies for the predicted structures are summarized in Table 1. For mol I, all the predicted crystal structures have slightly lower densities than that measured experimentally at finite temperature.[12] The opposite trend was observed for mol II, however, in which slightly higher densities were observed for the predicted structures as compared to the experimental crystal.[36] For both molecules, at least one predicted crystal structure belongs to the same space group that the experimental structure is in. However, it is obvious from Figures 2 and 3, as well as the crystal unit cell parameters in Table 1, that the predicted most possible crystal structures failed to match the known experimental structure for both molecules.

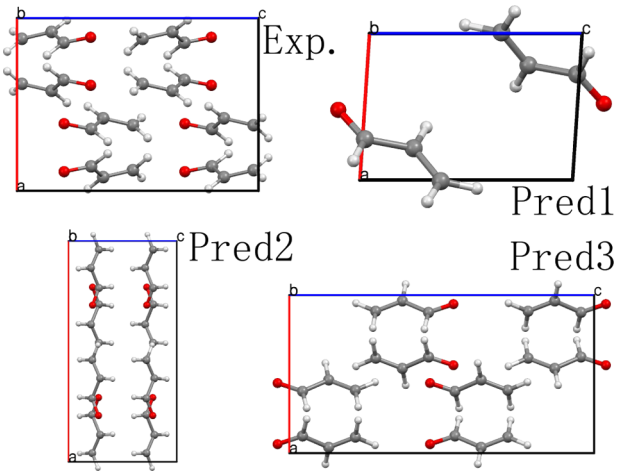


Figure 2. Experimental crystal structure of mol I together with the three most likely predictions obtained from CSP. The predicted structures do not match the experimental structure for this molecule.

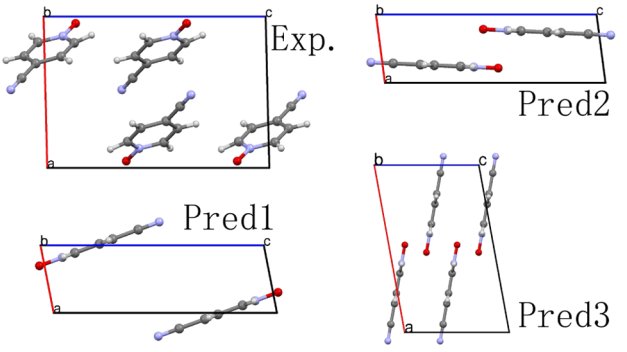


Figure 3. Experimental crystal structure of mol II together with the three most likely predictions from CSP. The predicted structures do not match the experimental structure for this molecule.

When evaluating the lattice energy for each structure, PMIN ignores all the intramolecular degrees of freedom (rigid molecule model) and only considers intermolecular interactions. Using the flexible all-atom GAFF force field, the crystal structures were minimized with the cell parameter fixed and the lattice energies achieved this way are provided as the last column in Table 1. The relative energies changed slightly comparing to the PMIN lattice energy. When the structures are described with the GAFF force field in the PSCP method, all crystals were found to be stable and the structural parameters changed only slightly from the PMIN structures. The equilibrated lattice parameters at 150 K for mol I and 295 K for mol II, the temperatures at which the crystal structures were measured in experiments,[12,36] are summarized in Table 2. For both molecules, the GAFF force field tends to underestimate the experimental densities slightly (by 5.3% for mol I and 3.6% for mol II). For mol I, the computed density starting from the experimentally measured crystal structure still has the highest density among the four structures, but the difference with the three predicted structures is small (<1.3%). For mol II, the experimental crystal was also found to have the highest density among the four structures. The densities of the three predicted structures are 2.0%, 1.8%, and 2.2% lower than the experimental value, respectively. The ranking of the predicted structures based on their potential energy also changed when the GAFF force field was used. For mol I, pred.1 is still the most stable structure. However, the order of pred.2 and pred.3 is switched. The potential energy of the experimental crystal was found to be higher than all three predicted structures. For mol II, the potential energy of pred.1 is only slightly lower than pred.3, and both are higher than pred.2. The experimental crystal was found to be the most stable structure among the four, but the difference in energy is still within ∼1 kJ/mol of each other.

For each crystal structure, the PSCP method was applied to calculate the melting point. The results are summarized in Table 3. When the experimental crystal structure was used, the melting point was calculated to be 168 K for mol I, only 18 K lower than the experimental value.[12] For the three predicted structures, the melting points were predicted to be 167 K, 178 K, and 165 K, respectively, consistent with the results when the experimental crystal structure was used. Besides the statistical error from the PSCP calculation, which was estimated to be less than 1 K in each case, the slight variation in the calculated melting points are likely due to the different inter- and intramolecular interaction in each crystal structure. For example, the O···O distances are found to differ in the four

dx.doi.org/10.1021/ct301095j | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Summary of Experimental and Predicted Crystal Parameters (Length in Ångstrom and Angle in Degree) as Well as the calculated Lattice Energies (Relative to the One with the Lowest Lattice Energy, Denoted as E_{lat}) Obtained from PMIN of Mol I and Mol II[a]

	a	b	c	α	β	γ	density (g/cm^3)	space group	Z	E_{lat} (kJ/mol) PMIN	E_{lat} (kJ/mol) GAFF
Mol I (150 or 0 K)											
exp.[b]	6.970	9.514	9.752	90.00	90.00	90.00	1.152	*Pbca*	8		0.439
pred.1	4.1292	6.9830	5.9875	90.00	93.75	90.00	1.080	P21	2	0.000	0.000
pred.2	14.0778	7.1540	6.9168	90.00	90.00	90.00	1.069	*Pbca*	8	0.167	1.177
pred.3	7.2139	6.9535	13.9385	90.00	90.00	90.00	1.065	*Pbca*	8	0.251	1.196
Mol II (295 or 0 K)											
exp.[c]	7.880	6.100	11.610	90.00	88.60	90.00	1.430	*P*21/*c*	4		−3.477
pred.1	3.5799	6.6342	11.6780	90.02	78.94	90.02	1.465	$P\bar{1}$	2	0.000	0.000
pred.2	3.5858	6.6339	11.5323	90.00	82.55	90.00	1.466	*P*21	2	0.042	−0.132
pred.3	11.6235	6.6340	7.1718	90.00	79.652	90.00	1.466	*P*21/*c*	4	0.042	−0.214

[a]The lattice energies evaluated using the GAFF force field are provided in the last column. [b]Experimental value taken from ref 12. [c]Experimental value taken from ref 36.

Table 2. Summary of Crystal Parameters (Length in Ångstroms and Angle in Degrees) as Well as the Average Densities and Potential Energies (Relative to the One with the Lowest Lattice Energy, Denoted as E_P) of Mol I and Mol II after NPT Ensemble Equilibrations Using Molecular Dynamics Simulations and the GAFF Force Field at Finite Temperature

	a	b	c	α	β	γ	density (g/cm^3)	E_p(kJ/mol)
Mol I (150 K)								
exp.	7.0165	9.7292	10.0039	90.00	90.00	90.00	1.091	0.887 ± 0.004
pred.1	4.0146	6.9884	6.0976	90.00	93.68	90.00	1.091	0.000 ± 0.004
pred.2	14.1302	7.0966	6.8988	90.00	90.00	90.00	1.077	0.765 ± 0.004
pred.3	7.1518	6.9148	13.9835	90.00	90.00	90.00	1.077	0.259 ± 0.004
Mol II (295 K)								
exp.	7.7830	6.3517	11.716	90.00	88.61	90.00	1.379	−1.033 ± 0.013
pred.1	3.6343	6.8274	12.1163	90.02	79.34	90.01	1.351	0.000 ± 0.013
pred.2	3.6428	6.8506	11.9050	90.00	82.78	90.00	1.354	−0.234 ± 0.013
pred.3	12.0418	6.8364	7.3020	90.00	79.84	90.00	1.349	0.025 ± 0.013

Table 3. Summary of Calculated Melting Points (T_m), Fusion Enthalpy (ΔH_f), and Fusion Entropy (ΔS_f) Using the PSCP Method for the Experimental and Three Predicted Crystal Structures of Each Molecule

crystal structure	T_m (K)	ΔH_f (kJ/mol)	ΔS_f (J/mol/K)
Mol I (T_m (Exp.) = 186 K[a])			
exp.	168	5.774 ± 0.126	34.309 ± 0.753
pred.1	167	7.113 ± 0.126	42.593 ± 0.753
pred.2	178	7.364 ± 0.126	41.380 ± 0.711
pred.3	165	6.569 ± 0.126	39.915 ± 0.753
Mol II (T_m (Exp.) = 494 K[b])			
exp.	544	33.807 ± 0.460	62.132 ± 0.837
pred.1	519	30.878 ± 0.416	59.496 ± 0.795
pred.2	518	30.752 ± 0.416	59.413 ± 0.795
pred.3	521	31.045 ± 0.460	59.580 ± 0.879

[a]Experimental value taken from ref 12. [b]Experimental value taken from ref 36.

crystal structures. The associated difference in entropy should play a role as well (see below). The free energy as a function of temperature calculated for each crystal structure using the Gibbs−Helmholtz relation[34] is shown in Figure 4. The upper panel shows the results when the lowest temperature (T_0) was taken to be the reference temperature (T_0 = 120 K in this case) and that assumes $(G/RT)_{ref}$ = 0 so all curves start from zero at 120 K. At 220 K, as shown in the inset, the largest difference is about 0.56 (unitless), only ∼2% of the absolute values. Since lattice energy was used in ranking the predicted crystal structures, the above results suggested that the entropy contribution to the free energy is similar for all the structures. The relative free energies of the crystal and liquid phases, computed by shifting G^{NPT} based on the calculated $\Delta G(T_{ref})$ along the PSCP cycle, are shown in the lower panel of Figure 4. Due to the small differences in G between each crystal structure, the free energy curves differ from one another by a small amount, resulting in very similar melting points.

These results support the idea that crystal structures with similar energies will end up with similar melting points. It is worth noting that the predicted melting points do not seem to have a clear correlation with the lattice energy or the equilibrated potential energy, indicating that entropy also plays a role in determining the melting points.

Similarly, for mol II, the simulation based on the experimental crystal structure resulted in a melting point of 544 K, 50 K (or 10%) higher than the experimental value of 494 K.[36] With the predicted structures, the melting points were calculated to be 519 K, 518 K, and 521 K, respectively, all within 5% of that calculated using the experimental crystal structure. The free energy curves calculated for each crystal structure are shown in Figure 5. Similar to mol I, the free energy curves are very close to each other within the 450−550 K range and the deviation was found to be less than 2%.

The fusion enthalpy (ΔH_f) and fusion entropy (ΔS_f) associated with each crystal structure was also calculated and are given in Table 3. For mol I, the largest ΔH_f was found for

dx.doi.org/10.1021/ct301095j | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

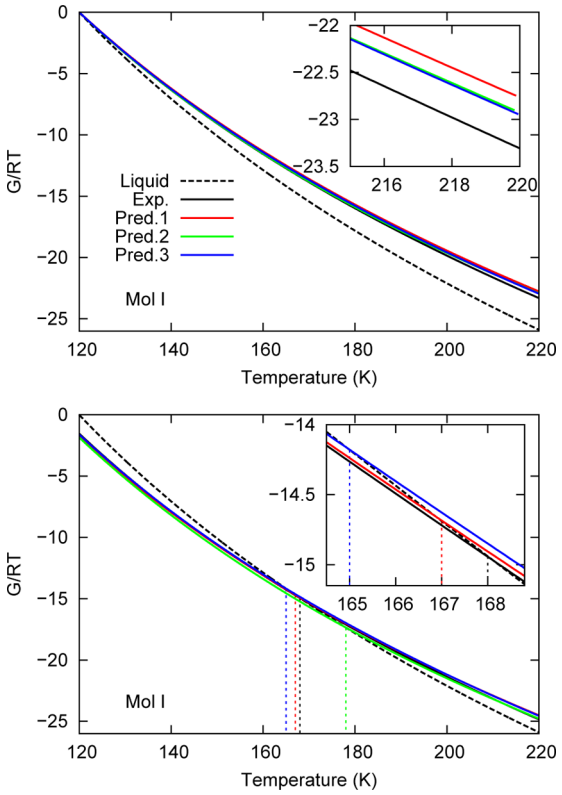


Figure 4. Free energy as a function of temperature calculated for the liquid and crystal phases of mol I based on NPT ensemble molecular dynamics simulations using the Gibbs−Helmholtz relation. Upper panel: the free energy curves with the assumption that $\Delta G(T_{ref}) = 0$ at $T_{ref} = T_0 = 120$ K, the crystal phase curves at high temperature are detailed in the inset. Lower panel: the free energy curves shifted from G^{NPT} based on $\Delta G(T_{ref})$ calculated along the PSCP cycle at $T_{ref} = 200$ K. The melting points are indicated by vertical dotted lines and detailed in the inset. The free energy curves calculated based on different crystal structures are very close to each other and similar melting points were obtained.

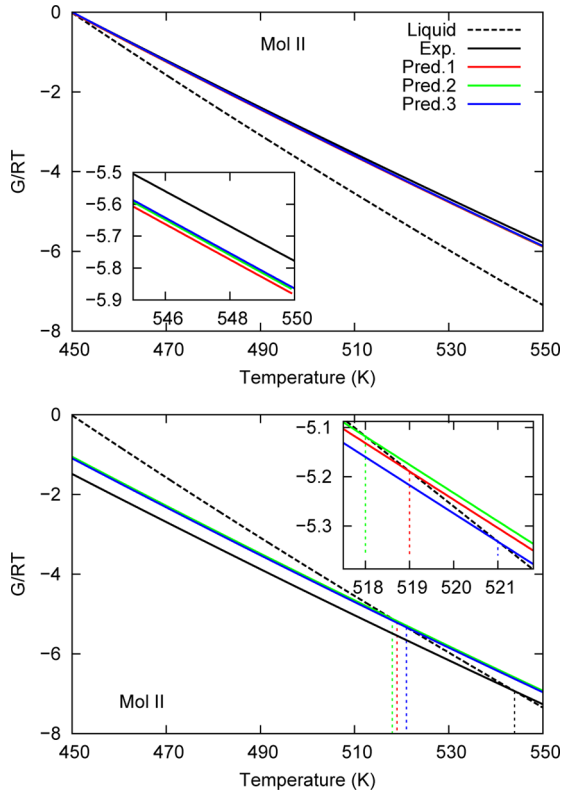


Figure 5. Free energy as a function of temperature calculated for the liquid and crystal phase of mol II based on NPT ensemble molecular dynamics simulations using the Gibbs−Helmholtz relation. Upper panel: the free energy curves with the assumption that $\Delta G(T_{ref}) = 0$ at $T_{ref} = T_0 = 450$ K, the crystal phase curves at high temperature are detailed in the inset. Lower panel: the free energy curves shifted from G^{NPT} based on $\Delta G(T_{ref})$ calculated along the PSCP cycle at $T_{ref} = 540$ K. The melting points are indicated by vertical dotted lines and detailed in the inset. The free energy curves calculated based on different crystal structures are very close to each other and similar melting points were obtained.

pred.2 of 7.364 kJ/mol and the smallest for the experimental crystal structure of 5.775 kJ/mol. The difference (1.590 kJ/mol) is about the level of thermal energy at the melting temperature. The fusion entropies were also calculated and found to be similar to each other. For mol II, the relative differences in the calculated fusion enthalpy and fusion entropy are even smaller (within 9% for ΔH_f and within 5% for ΔS_f). These results as well as the computed melting points suggest that the predicted crystal structures for both molecules have very similar free energies even though the correct space group was not captured.

4. DISCUSSION

In the current work, two relatively simple molecules were chosen to demonstrate a procedure for complete in silico prediction of melting points. Both molecules are relatively rigid and share many similarities such as their elemental composition. However, the two target molecules present differences as well. For example, the size of mol II is about twice that of mol I (in terms of numbers of atoms). The two molecules include different structural groups (double bonds in mol I and a ring in mol II), which would likely affect the packing pattern in their crystal structures. The crystal structures of the two molecules also belong to different space groups, which is the result of different inter- and intramolecular interactions in the two molecules. In addition, the two molecules have quite different melting points (186 K for mol I and 494 K for mol II).[12,36] The success of the proposed method in predicting the melting points of two molecules with such different characters indicates that the method is reliable and robust.

For both molecules, the CSP calculation used in the current work failed to obtain the experimental structure as one of the three most possible solutions. This is likely due to the fact that a relatively simple CSP method was used. It is possible that better predictions are obtained using MOLPAK/PMIN by tuning the options in the package (for example, finer grid size in generating possible structures or a more accurate force field in evaluating the lattice energies), or by using a more advanced method. Such CSP refinements are beyond the scope of the current work. Actually, the excellent melting point prediction results based on the use of incorrect crystal structures strongly supports the idea that the actual experimental crystal structure is not necessary in the melting point calculations as long as the predicted structure has a similar lattice energy or free energy as the experimental crystal.

In the fourth blind test, the correct crystal structure of mol I (which was referred to as mol XII there) was submitted as the

dx.doi.org/10.1021/ct301095j | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 4. Summary of Experimental and Equilibrated (Using GAFF Force Field) Crystal Parameters (Length in Ångstrom and Angle in Degrees) as Well as the Average Densities and Potential Energies (Relative to the One with the Lowest Energy, Denoted as E_p) of Target Molecule VI from the 2nd CSP Blind Test[a]

	a	b	c	α	β	γ	density (g/cm^3)	space group	Z	E_p (kJ/mol)
Experimental Crystal Parameters										
form I (183 K)[b]	8.479	8.958	14.894	90.00	91.86	90.00	1.464	$P21/c$	4	
form II (203 K)[b]	12.110	10.792	17.464	90.00	97.32	90.00	1.463	$P21/c$	8	
form III (95 K)[c]	10.618	9.318	23.056	90.00	90.00	90.00	1.452	$Pbca$	8	
Equilibrated Crystal Parameters										
form I (183 K)	8.618	9.068	14.728	90.00	91.88	90.00	1.348			1.51 ± 0.04
form II (203 K)	11.979	11.128	17.271	90.00	97.40	90.00	1.358			0.00 ± 0.04
form III (95 K)	11.044	9.188	22.753	90.00	90.00	90.00	1.343			7.24 ± 0.04

[a]The experimental or simulation temperatures are noted in the bracket. [b]Experimental value taken from ref 46. [c]Experimental value taken from ref 47.

preferred solution by many participant groups using MOLPAK with advanced energy evaluation methods or different CSP procedures.[12] Crystal structures of even more challenging target molecules have been successfully predicted in past blind tests and methods have significantly improved since the first blind test in 1999. Particularly, the handling of the internal degrees of freedom of molecules has been the focus of many developments and great progress has been made.[37−41] In the latest (fifth) blind test in 2010, target molecule XX contains eight flexible torsional angles. For such a challenging molecule, at least three groups succeeded in identifying its correct crystal structure including two groups who submitted the experimental crystal structure as their first prediction.[9] The other direction that significant improvement has been made over the years is the more and more accurate evaluation of the lattice energy. The package PMIN used in the current work applies a classical force field with static atomic partial charges in calculating the lattice energies. The distributed multipole electrostatic model was proposed almost 20 years ago[42] and has been shown to improve the accuracy of the prediction significantly.[20] It is now widely applied in CSP. With increases in computational power, fully ab initio calculation of crystal structures has become possible.[43,44] Other methods such as fragment-based electronic structure methods have also been proposed.[45] With these developments, the success of CSP has been significantly improved. In the fourth blind test carried out in 2007, for the first time, one participant successfully predicted the crystal structures of all four target molecules with increasing complexity.[15] These advances suggest that CSP methods can yield good predictions of crystal structures, though as shown here only "reasonable" structures are required for accurate melting point predictions.

It is worth mentioning that accurate prediction of melting point for a given molecule also depends on the melting point calculation method. Recently, we compared several popular melting point calculation methods and found the free energy based PSCP method to be more reliable than other "direct" approaches.[1]

The other concern is the force field applied with the melting point calculation methods. Melting point calculation on 6-amino-2-(phenylsulfonylimino)-1,2-dihydrophridine, also known as the target molecule VI from the second blind test, serves as a good example (Table 4). At least three structures had been crystallized in experiments for this molecule.[46,47] The three crystal structures as well as the optimized molecular structure are shown in Figure 6. Similar melting points in the range 498−510 K have been reported for all three crystal

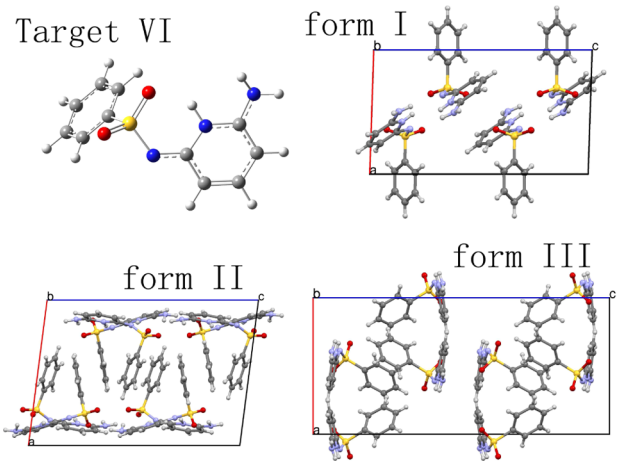


Figure 6. Optimized molecular structure of target molecule VI from the second CSP blind test and its three experimental crystal structures.

structures.[47] Using the PSCP method and GAFF force field, melting point calculations were carried out for each of the three experimental crystal structures, and the results are summarized in Table 5. Based on the simulations, a melting point of 486 K

Table 5. Summary of Calculated Melting Points (T_m), Fusion Enthalpy (ΔH_f), and Fusion Entropy (ΔS_f) Using the PSCP Method for the Experimental Crystal Structures of Target Molecule VI from the 2nd Blind Test[a]

	T_m (K)			
crystal structure	calc.	exp.	ΔH_f (kJ/mol)	ΔS_f (J/mol/K)
form I	465	510[b]	18.12 ± 0.75	38.91 ± 1.63
form II	486	498[b]	21.63 ± 0.84	44.48 ± 1.72
form III	355	508[c]	8.62 ± 0.63	24.27 ± 1.76

[a]Experimental melting points are also provided for comparison. [b]Experimental value taken from ref 46. [c]Experimental value taken from ref 47.

was found for one of the crystal structures (form II), which is very close to the experimental value. However, a melting point as low as 355 K was found in the calculation for the crystal form III, which is about 150 K lower than experimental results. This molecule has actually raised serious difficulty for crystal structure prediction due to the lack of an accurate force field.[21] A better force field is likely needed to properly describe the different hydrogen bonding patterns in the different crystal

dx.doi.org/10.1021/ct301095j | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

structures in order to reliably evaluate their energies[19] and accurately calculate the melting points.

Obviously, like CSP or other types of molecular simulation methods, the performance of the PSCP method highly depends on the accuracy of the applied force field.[35] Ideally, advanced force fields or ab initio method can be applied, which will make such calculations more "predictive". Unfortunately, this is not possible based on current computing hardware resources. Effort are being made to develop more accurate and efficient force field for various types of molecules, however, and the proposed method will definitely benefit from these advances.

5. CONCLUDING REMARKS

We have presented a procedure that enables the accurate prediction of the melting point of a compound given only its molecular structure as input. This approach takes advantage of recent developments in both melting point calculation methods and crystal structure prediction techniques. Instead of using the experimental crystal structure as input, it was shown that a predicted structure can be used to obtain an accurate melting point, even one with a totally different space group from the experimental structure. The success of the approach hinges on the fact that the predicted crystal structures and the experimental crystal structure have similar free energies. This makes crystal polymorph prediction difficult but also means that melting temperatures are relatively insensitive to the actual crystal structure used in the calculation.

This method was tested on two rigid molecules for which both experimental crystal structures and melting points are available. The crystal structure prediction was carried out for both molecules using the packages MOLPAK and PMIN. The experimental structure was not found in the three most likely prediction results. Melting points were calculated using the pseudosupercritical path (PSCP) method starting from the experimental crystal and the three most likely predictions for each molecule. The calculated melting points agree with each other very well although the starting crystal structures are quite different, even belonging to different space groups in some cases. In addition, for each molecule, the calculated melting point also agrees reasonably well with the experimental results, especially considering the fact that the melting point calculation is still not a trivial task.[1] It should be noted, however, that it is probably still not possible to compute the melting point differences in crystal polymorphs with this approach.

The method was then applied to a more complex molecule with intramolecular degrees of freedom (6-amino-2-(phenylsulfonylimino)-1,2-dihydrophridine). CSP was not attempted, but instead, the PSCP method was applied directly to three different reported crystal structures. The results were mixed, with agreement ranging from excellent to poor, depending on the starting structure. This result showed that regardless of the methods applied, inaccurate force fields will lead to inaccurate melting points.

In the current work, a relatively simple crystal structure prediction method and two rigid molecules were used to demonstrate the method. Many advanced CSP methods have been developed and successfully applied to more complicated molecules as nicely summarized in the reports following each CSP blind test organized by the Cambridge Crystallographic Data Centre (CCDC).[8−12] The melting point calculation method (PSCP) has been also proven to be reliable and robust for complex molecules.[1] This suggests that it is now possible to make accurate fully in silico melting point predictions for a wide range of molecules.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ed@nd.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This material is based upon work supported by the Air Force Office of Scientific Research under AFOSR Award Number FA9550-10-1-0244 and by the Advanced Research Projects Agency—Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0000094. Computational resources were provided by the Center for Research Computing (CRC) at the University of Notre Dame. We thank Prof. Herman L. Ammon for helping us with the setup of MOLPAK/PMIN simulations and Prof. Sarah L. Price for helpful discussions.

REFERENCES

(1) Zhang, Y.; Maginn, E. J. *J. Chem. Phys.* **2012**, *136*, 144116.
(2) Plechkova, N.; Seddon, K. R. *Chem. Rev.* **2008**, *37*, 123−150.
(3) Zhang, S.; Lu, X.; Zhang, Y.; Zhou, Q.; Sun, J.; Han, L.; Yue, G.; Liu, X.; Cheng, W.; Li, S. In *Molecular Thermodynamics of Complex Systems: Structure and Bonding*; Lu, X., Hu, Y., Eds.; Springer Verlag: Berlin/Heidelberg, 2009; Vol. *131*, pp 143−191.
(4) Ma, Z.; Yu, J.; Dai, S. *Adv. Mater.* **2010**, *22* (2), 261−285.
(5) Hallett, J. P.; Welton, T. *Chem. Rev.* **2011**, *111*, 3058−3576.
(6) Wang, H.; Gurau, G.; Rogers, R. D. *Chem. Soc. Rev.* **2012**, *41*, 1519−1537.
(7) Gavezzotti, A. *Acc. Chem. Res.* **1994**, *27*, 309−314.
(8) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr.* **2000**, *B56*, 697−714.
(9) Bardwell, D. A.; Adjiman, C. S.; Arnautova, Y. A.; Bartashevich, E.; Boerrigter, S. X. M.; Braun, D. E.; Cruz-Cabeza, A. J.; Day, G. M.; Valle, R. G. D.; Desiraju, G. R.; van Eijck, B. P.; Facelli, J. C.; Ferraro, M. B.; Grillo, D.; Habgood, M.; Hofmann, D. W. M.; Hofmann, F.; Jose, K. V. J.; Karamertzanis, P. G.; Kazantsev, A. V.; Kendrick, J.; Kuleshova, L. N.; Leusen, F. J. J.; Maleev, A. V.; Misquitta, A. J.; Mohamed, S.; Needs, R. J.; Neumann, M. A.; Nikylov, D.; Orendt, A. M.; Pal, R.; Pantelides, C. C.; Pickard, C. J.; Price, L. S.; Price, S. L.; Scheraga, H. A.; van de Streek, J.; Thakur, T. S.; Tiwari, S.; Venuti, E.; Zhitkov, I. K. *Acta Crystallogr.* **2011**, *B67*, 535−551.
(10) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr.* **2002**, *B58*, 647−661.
(11) Day, G. M.; Motherwell, W. D. S.; Ammon, H. L.; Boerrigter, S. X. M.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.; Dunitz, J. D.; Schweizer, B.; van Eijck, B. P.; Erk, P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Leusen, F. J. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P. *Acta Crystallogr.* **2005**, *B61*, 511−527.
(12) Day, G. M.; Cooper, T. G.; Cruz-Cabeza, A. J.; Hejczyk, K. E.; Ammon, H.; Boerrigter, S. X.; Tan, J. S.; Della Valle, R. G.; Venuti, E.; Jose, J.; Gadre, S. R.; Desiraju, G. R.; Thakur, T. S.; van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Neumann, M. A.; Leusen, F. J. J.; Kendrick, J.; Price, S. L.; Misquitta, A. J.; Karamertzanis, P. G.; Welch, G. W. A.; Scheraga, H. A.;

Arnautova, Y. A.; Schmidt, M. U.; van de Streek, J.; Wolf, A. K.; Schweizer, B. *Acta Crystallogr.* **2009**, *B65*, 107−125.
(13) Price, S. L. *Int. Rev. Phys. Chem.* **2008**, *27*, 541−568.
(14) Price, S. L. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1996−2009.
(15) Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. *Angew. Chem., Int. Ed.* **2008**, *47*, 2427−2430.
(16) Woodley, S. M.; Catlow, R. *Nat. Mater.* **2008**, *7*, 937−946.
(17) Lehmann, C. W. *Angew. Chem., Int. Ed.* **2011**, *50*, 5616−5617.
(18) Gavezzotti, A.; Filippini, G. *J. Am. Chem. Soc.* **1995**, *117*, 12299−12305.
(19) Day, G. M.; Chisholm, J.; Shan, N.; Motherwell, W. D. S.; Jones, W. *Cryst. Growth Des.* **2004**, *4*, 1327−1340.
(20) Day, G. M.; Motherwell, W. D. S.; Jones, W. *Cryst. Growth Des.* **2005**, *5*, 1023−1033.
(21) Chan, H. C. S.; Kendrick, J.; Leusen, F. J. J. *Angew. Chem., Int. Ed.* **2011**, *50*, 2979−2981.
(22) Holden, J. R.; Du, Z.; Ammon, H. L. *J. Comput. Chem.* **1993**, *14*, 422−437.
(23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ã.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.*Gaussian 09* Revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
(24) van Eijck, B. P. *J. Comput. Chem.* **2001**, *22*, 816−826.
(25) Anghel, A. T.; Day, G. M.; Price, S. L. *Cryst. Eng. Comm* **2002**, *4*, 348−355.
(26) van Eijck, B. P.; Kroon, J. *J. Comput. Chem.* **1999**, *20*, 799−812.
(27) Zhang, Y.; Maginn, E. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12157−12164.
(28) Plimpton, S. *J. Comput. Phys.* **1995**, *117*, 1−19.
(29) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(30) Bayly, C.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(31) Veld, P. J. I.; Ismail, A. E.; Greest, G. S. *J. Chem. Phys.* **2007**, *127*, 144711.
(32) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(33) Shinoda, W.; Shiga, M.; Mikami, M. *Phys. Rev. B* **2004**, *69*, 134103.
(34) Zhang, Y.; Maginn, E. J. *J. Phys. Chem. B* **2012**, *116*, 10036−10048.
(35) Jayaraman, S.; Maginn, E. J. *J. Chem. Phys.* **2007**, *127* (21), 214504.
(36) Hardcastle, K. I.; Laing, M. J.; McGauley, T. J.; Lehner, C. F. *J. Cryst. Mol. Sruct.* **1974**, *4*, 305.
(37) Karamertzanis, P. G.; Price, S. L. *J. Chem. Theory Comput.* **2006**, *2*, 1184−1199.
(38) Day, G. M.; Motherwell, W. D. S.; Jones, W. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1693−1704.
(39) Karamertzanis, P. G.; Pantelides, C. C. *Mol. Phys.* **2007**, *105*, 273−291.
(40) Kim, S.; Orendt, A. M.; Ferraro, M. B.; Facelli, J. C. *J. Comput. Chem.* **2009**, *30*, 1973−1985.
(41) Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. *J. Chem. Theory Comput.* **2011**, *7*, 1998−2016.
(42) Willock, D. J.; Price, S. L.; Leslie, M.; Catlow, C. R. A. *J. Comput. Chem.* **1995**, *16*, 628−647.
(43) Neumann, M.; Perrin, M.-A. *J. Phys. Chem. B* **2005**, *109*, 15531−15541.
(44) Neumann, M. A. *J. Phys. Chem. B* **2008**, *112*, 9810−9829.
(45) Wen, S.; Nanda, K.; Huang, Y.; Beran, G. J. O. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7578−7590.
(46) Jetti, R. K. R.; Boese, R.; Sarma, J. A. R. P.; Reddy, L. S.; Vishweshwar, P.; Desiraju, G. R. *Angew. Chem., Int. Ed.* **2003**, *42*, 1963−1967.
(47) Roy, S.; Matzger, A. J. *Angew. Chem., Int. Ed.* **2009**, *48*, 8505−8508.