Evidence for a Size Dependent Nucleation Mechanism in Solid State Polymorph Transformations

Gregg T. Beckham,[†,§] Baron Peters,[‡,||] and Bernhardt L. Trout*[,†]

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and CECAM, Ecole Normale Supérieure, 69364 Lyon Cedex 7, France

Received: October 20, 2007; Revised Manuscript Received: March 26, 2008

This study applies aimless shooting and likelihood maximization to determine the molecular mechanism in the solid state polymorph transformation in terephthalic acid from over 500 candidate order parameters. The crystals examined here extend the range of crystal sizes considered in our previous work (*J. Amer. Chem. Soc.* 2007, *129*, 4714) and reveal a change in the mechanism with increasing system size. As the crystal size increases beyond that studied in our previous work, the polymorph transformation mechanism changes from a global distortion of the crystal to a local corner nucleation mechanism. In the corner nucleation mechanism, the interfacial area between the two polymorphs is minimized for a given nucleus size. However, this mechanism differs from classical nucleation theory in that the molecular level details are essential to describe the nucleation process, which involves nonspherical domains at the corner of the crystal. These new findings suggest that there is a range of sizes for which corner nucleation is the dominant mechanism of polymorph transitions, thus implying that different mechanistic regimes exist for nucleation based on crystal size. From a computational standpoint, this study demonstrates the utility of aimless shooting and likelihood maximization to identify nonintuitive reaction coordinates in complex systems.

Introduction

A technologically important case of nucleation is the interconversion between different crystal forms, or polymorphs.[1,2] Polymorphic systems are an ideal framework for the study of structure−function relationships, as polymorphs typically exhibit different physical and chemical properties. In the pharmaceutical industry, understanding and controlling polymorphism is often of great importance as different crystal forms lead to variations in both processing and in pharmacokinetics. A particularly high profile case of polymorphism in the pharmaceutical industry was highlighted when Abbott Laboratories' AIDS drug, Ritonavir converted to a more stable polymorph in the manufacturing phase, causing significant production setbacks and financial losses.[3] Further interest in polymorphism has stemmed from recent attempts to patent specific polymorphs.[4] Thus, understanding mechanisms leading to transformation between polymorphs would be of substantial benefit to controlling polymorphism, and would provide complementary kinetic information to the thermodynamic problem of polymorph prediction.[5–7] Solid state polymorph transformations in molecular crystals have received significant attention from the experimental community.[2,8–14] At present, however, gaining direct molecular level insight to the dynamical events occurring during solid state polymorph transformations is outside of the scope of experimental capabilities, and thus there is little definitive evidence for any particular mechanism. As remarked in an excellent review by Herbstein: "there are relatively few papers about the actual transition directly viewed by microscopic techniques in order to infer the mechanism, and not many about changes in crystal structure...*as the system passes through the transition*."[14]

Molecular simulation approaches alternatively offer the ability to track individual molecules in space and time, and can be utilized to study crystallization processes such as solid state polymorph transformations. To date, many computational approaches have proved fruitful for investigating nucleation processes for a wide range of complexity from model systems[15–26] to atomic and molecular crystals.[27–35] Several studies either utilize biasing methods, which assume a mechanism a priori and force the system from one state to another,[25,26,28,32,36] or alternatively run many trajectories at conditions where the metastable state is highly unstable so that the "rare" event occurs on accessible timescales.[16,27,33] While these methods provide valuable insight, a more ideal methodology would examine nucleation without assuming a mechanism a priori and simulate conditions comparable to experimental supersaturations and subcooling. For rough, high dimensional free energy landscapes, which are inherent in nucleation processes, transition path sampling (TPS)[37–40] enables the collection of unbiased pathways that connect stable end states, such as two polymorphs, by efficiently focusing the computational time on the dynamical bottleneck. Several research groups have successfully utilized the TPS method to study crystal nucleation.[19,21–23,30,31,35] In addition, recent developments have led to methods that systematically determine the mechanism or reaction coordinate from path sampling results.[22,41–43] The Genetic Neural Network approach of Ma and Dinner[42] and the subsequently developed likelihood maximization technique of Peters et al.,[22,43] both use an informatics approach to approximate the reaction coordinate as a function of physically relevant variables. These two methods have recently been compared, and it was shown that likelihood maximization requires less trajectory information to obtain an accurate reaction coordinate.[43] Accurate reaction coordinates are

* Corresponding author. E-mail: trout@mit.edu.
† Massachusetts Institute of Technology.
‡ CECAM.
§ Current address: National Bioenergy Center, National Renewable Energy Laboratory, Golden, Colorado 80401.
|| Current address: Department of Chemical Engineering, University of California at Santa Barbara, Santa Barbara, California 93106.

10.1021/jp710192u CCC: $40.75
Published on Web 06/04/2008

important because they enable calculations of kinetically meaningful free energy barriers[44–50] and accurate rate constants.[51–57]

We recently presented a computational study that illustrated the likelihood maximization approach by elucidating the mechanism of the solid state polymorph transformation in small crystals of terephthalic acid.[35] Terephthalic acid (TPA) is chosen as the model system for in-depth computational studies because the crystal structures of the Form I and Form II polymorphs are known[58] and because experiments show that the polymorph transformation exhibits an induction time for transformation unless significantly subcooled or mechanically perturbed.[10,59] Moreover, the TPA structure represents an ubiquitous motif in organic crystalline pharmaceuticals: carboxylic acid dimers that form long supramolecular synthons or molecular chains.[60] In the previous study, two system sizes were investigated: a $6 \times 6 \times 6$ (216 molecules) and a $7 \times 7 \times 7$ (343 molecules) crystal, both with free boundaries. Simulations of the Form I - Form II transition in those systems show a nucleation step, but the free energy barrier to transformation increases as a function of system size. The mechanism, or reaction coordinate, in these two small systems is shown to be a global distortion of the crystal as measured by the global **b** lattice parameter and the global **b/c** lattice parameter ratio for the $6 \times 6 \times 6$ and $7 \times 7 \times 7$ systems, respectively. The term "global" in this sense denotes an average over the entire crystal, which is measured on the nanometer length scale.

While global order parameters (OPs) serve as a good approximation to the reaction coordinate for the $6 \times 6 \times 6$ and $7 \times 7 \times 7$ crystals, one can imagine that on a larger length scale, an average lattice parameter measured over the entire crystal will not be able to capture the fluctuations on a specific side or corner of the crystal. Therefore, we proposed that these same global OPs will not be sufficient reaction coordinates in larger systems.[35] Instead, we hypothesize that for larger crystals, localized fluctuations will lead to nucleation events rather than the global distortion seen in the $6 \times 6 \times 6$ and $7 \times 7 \times 7$ systems. The goal of this study is to investigate whether local OPs become important for larger TPA crystal sizes and, if so, to determine the nature of the important local OPs so as to gain more definitive insight into the phenomenon of solid state polymorph transformations. Several types of local coordinates, such as new, local **b/c** lattice parameters, and the nucleus size of the Form II polymorph are tested as possible reaction coordinates.

Methodology

The molecular force field developed and validated previously is used for this study.[35] Two system sizes are examined in this study: an $8 \times 8 \times 8$ TPA crystal and a $10 \times 10 \times 10$ crystal, which correspond to 512 and 1000 molecules, respectively. For the $8 \times 8 \times 8$ system, the same scheme as described previously[35] is employed here: (1) an initial trajectory is collected, (2) aimless shooting is used to harvest reactive trajectories between the reactant and product, (3) likelihood maximization is applied to determine the reaction coordinate, (4) a histogram test is used to validate the reaction coordinate approximation, and (5) a kinetically relevant free energy barrier is determined using the correct reaction coordinate. Only steps 1−3 are conducted for the $10 \times 10 \times 10$ system due to the significant increase in computational expense.

Obtaining Initial Trajectories. Initial trajectories for the $7 \times 7 \times 7$ crystal (343 molecules) were previously obtained using umbrella sampling over the global **b/c** lattice parameter ratio.[35] This approach over the global **b/c** lattice parameter ratio, however, does not yield a set of configurations connecting the Form I and Form II crystal structures for the $8 \times 8 \times 8$ crystal. This preliminary result suggests that the global **b/c** lattice parameter is not an accurate reaction coordinate for the $8 \times 8 \times 8$ system. Therefore, umbrella sampling over a newly defined local **b/c** lattice parameter ratio is applied between minimized crystals from Form I to Form II. Figure 1 shows the outline of the new local **b/c** lattice parameter ratio that is used for both the Form I and Form II crystals. This corner is chosen because it is crystallographically equivalent to the corner that undergoes significant structural changes during the transformation in the smaller systems.

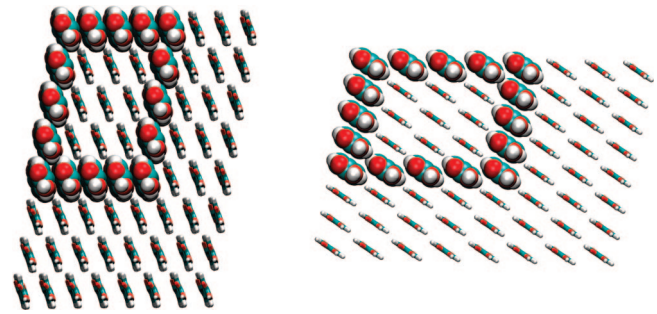

Figure 1. Form I and Form II TPA crystals from the **a**-axis view with the outline of the local **b/c** lattice parameter ratio used to obtain the initial trajectory for input to the aimless shooting method. This particular local OP was chosen because the same crystallographically equivalent corner in the $6 \times 6 \times 6$ and $7 \times 7 \times 7$ molecule systems undergoes significant changes during the nucleation event.

The initial trajectory for the $10 \times 10 \times 10$ crystal is generated using targeted molecular dynamics (TMD).[61] These two different methods, umbrella sampling and TMD, are applied to harvest initial trajectories for the $8 \times 8 \times 8$ and $10 \times 10 \times 10$ crystals, respectively, to examine the effect of the biasing method to obtain the initial aimless shooting point. To obtain an initial trajectory with TMD that retains some physical relevance, the Δ(rmsd) per time step is set to 2.5 Å/fs. The total trajectory length is therefore 2.3 ns with the initial rmsd of approximately 5.8 Å between the minimized end states. No difference is found in the methods in terms of generating an initial aimless shooting point as long as the Δ(rmsd) per time step is set low enough, as expected.

Aimless Shooting. Aimless shooting pathways are collected with the two-point version of aimless shooting. All MD trajectories for aimless shooting are obtained using the NAMD simulation package[62] in the *NVE* ensemble at 300 K. A time step of 1 fs is used with a cutoff for nonbonded interactions of 14 Å. The slight differences in the nonbonded energy switching functions between CHARMM and NAMD do not cause noticeable differences in the relevant quantities such as average lattice parameters, most likely because of the generous nonbonded energy cutoff distance of 14 Å, so comparison of the results from our previous work[35] should not be affected by the different simulation packages. The total trajectory length for aimless shooting is 60 ps, which maintains the fraction of inconclusive trajectories below 1%. The time displacement between shooting point candidates is $\Delta t = 600$ fs, which achieves 38% acceptance. The first 100 trajectories are discarded to remove effects from the initial pathway. Following path equilibration, 1500 paths are collected for the $8 \times 8 \times 8$ system and 500 paths are collected for the $10 \times 10 \times 10$ system for use in likelihood maximization. The basin definitions are based on the global **b/c** lattice parameter. An end point with **b/c** greater than 1.55 is characterized as Form I and those with **b/c** less

than 1.40 are assigned to the post nucleation basin. These basin definitions are based on MD trajectories in the respective basins, as in the previous study.[35]

Likelihood Maximization. The new half-trajectory likelihood maximization scheme[43] is applied to extract the reaction coordinate. The reaction coordinate, r, is modeled as a linear combination of candidate OPs denoted as q, with α_0 through α_m as adjustable coefficients given by:

$$r(q) = \alpha_0 + \sum_{k=1}^{m} \alpha_k q_k \quad (1)$$

The half-trajectory likelihood method yields a more accurate approximation to the transition state isosurface and applies whether the dynamics are ballistic or diffusive.[43] In this formulation of likelihood maximization, the reaction coordinate is related to probability of committing to basin B as given by:

$$p_B(r) = \frac{1}{2}(1 + \tanh(r)) \quad (2)$$

The corresponding likelihood function is given as:

$$L = \prod_{x_k \to B} p_B(r(x_k)) \prod_{x_k \to A} (1 - p_B(r(x_k)) \quad (3)$$

Expression 3 is maximized over all coefficients and possible combinations of candidate OPs, yielding a likelihood score for each potential reaction coordinate. The Bayesian information criterion (BIC) is calculated to determine the statistical significance of different reaction coordinate models.

Several hundred OPs are tested for the increased system sizes. All possible combinations of the local **b/c** axis ratios are screened as candidate variables such as the one shown in Figure 1. Also, the global lattice vectors, local lattice angles, local twisting along supramolecular chains, and directional variation of the supramolecular chains relative to one another are screened.

An intuitive candidate for the reaction coordinate from classical nucleation theory is the size of the cluster of the stable phase.[63] Therefore, a cluster size analysis is performed to identify nuclei of the stable Form II crystal based on the unit cell environment around each molecule as measured by the **b** and **c** lattice parameters around each molecule. The **b** lattice parameter is measured as the distance to the adjacent molecules along the **b** axis, and the **c** lattice parameter is measured similarly. If the **b/c** ratio for a molecule is greater than 1.55, the molecule is labeled as Form I, and if the **b/c** ratio is less than 1.15, it is labeled as Form II. This local classification is referred to as a "unit cell OP". These unit cell OPs are used to quantify the cluster size for both polymorphs. Molecules with lattice parameters between those values are tagged as interface molecules.

Committor Probability Distribution. The histogram test is applied to quantify the error in the reaction coordinate from likelihood maximization.[64–67] The histogram should be peaked at 1/2, implying that points on the $r = 0$ surface are transition states with $p_B = 1/2$. For the $8 \times 8 \times 8$ crystal, 20 trajectories of 30 ps each are fired randomly from 50 independent configurations generated by a method inspired by the equilibrium path sampling method,[49] as described in our previous study. The same conditions are used as for aimless shooting. The mean and standard deviation of the resulting histogram is used to obtain the intrinsic committor distribution as described by Peters[67] for a quantitative measure of the reaction coordinate error.

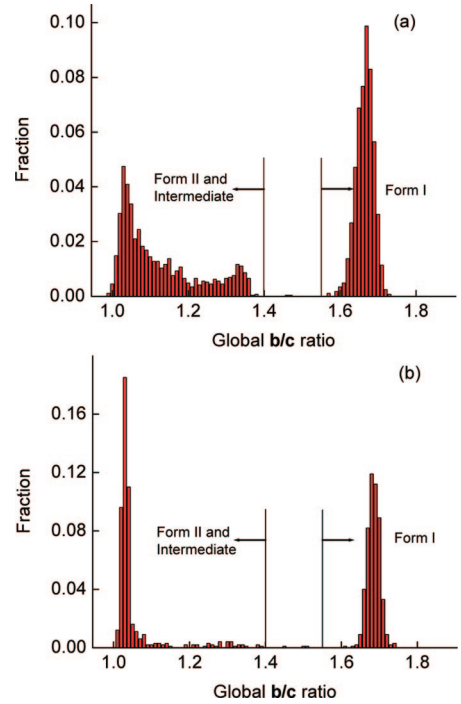

Figure 2. End points of aimless shooting simulations as a function of the global **b/c** axis ratio for the (a) $8 \times 8 \times 8$ and (b) $10 \times 10 \times 10$ crystal. The global **b/c** lattice parameter ratio is used for the quantitative basin definitions in the aimless shooting procedure.

Free Energy Calculation. After validating the reaction coordinate with the histogram test, the free energy is calculated with MD umbrella sampling for the $8 \times 8 \times 8$ crystal. Windows are 1.05 ns long with 50 ps of equilibration at the start of each reaction coordinate window. The weighted histogram analysis method is applied to construct the free energy curve.[46]

Results

Snapshots along a reactive aimless shooting trajectory for the $8 \times 8 \times 8$ crystal are given in the Supporting Information. The configurations at first visually seem similar to the mechanism previously obtained for the $6 \times 6 \times 6$ and $7 \times 7 \times 7$ crystals[35] as well as the $10 \times 10 \times 10$ crystal in this study. The global **b/c** values at the end points of the aimless shooting trajectories for both the $8 \times 8 \times 8$ and $10 \times 10 \times 10$ systems are shown in Figure 2. The distributions of global **b/c** values are split into two regions in the ranges of **b/c** < 1.4 and **b/c** > 1.55, indicating that the basin definitions are correct and that the trajectory length is sufficient.

Several sets of results from likelihood maximization are presented for comparison in Table 1. The likelihood maximization results for a one-dimensional reaction coordinate in the $8 \times 8 \times 8$ system are extracted from 1500 aimless shooting trajectories. The BIC for these data is 3.7. Three OPs are shown for the one-dimensional case: the nucleus size for the Form II cluster, the global **b/c** lattice parameter for comparison to results in our previous study,[35] and one of the local **b/c** OPs for a specific corner of the crystal. Both the Form II nucleus size and the global **b/c** axis ratio are inadequate 1-dimensional approximations to the reaction coordinate when compared to the local OP. The molecules used to measure the local **b/c** OP are highlighted in Figure 3 from the two **a**-axis views for a configuration on the transition state isosurface. The reaction coordinate is six layers in both the **b** and **c** dimensions, and is located in a different corner than the OP used to obtain an initial

TABLE 1: Likelihood Maximization Results for the Reaction Coordinate in the 8 × 8 × 8 and 10 × 10 × 10 Crystals[a]

system size	OP	likelihood score	transition state isosurface ($r = 0$)
8 × 8 × 8[b]	Form II nucleus size	−751.35	84 molecules
	global **b/c**	−714.73	1.50
	b/cfor corner 4, size = 7 × 6 × 6	−685.55	1.39
10 × 10 × 10[c]	Form II nucleus size	−308.57	81 molecules
	global **b/c**	−298.98	1.48
	b/cfor corner 4, size = 7 × 5 × 5	−255.88	1.35
7 × 7 × 7[d]	Form II nucleus size	−570.41	105 molecules
	global **b/c**[35]	−503.82	1.43

[a] In addition, the likelihood scores are included for the 7 × 7 × 7 optimal reaction coordinate from the previous study[35] along with the Form II nucleus size for comparison. [b] BIC = 3.66. [c] BIC = 3.11. [d] BIC = 4.08.

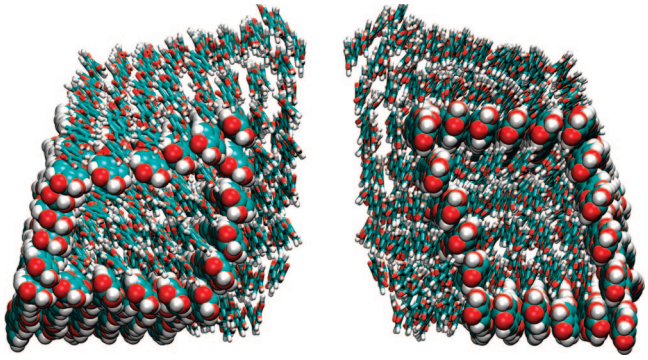

Figure 3. Snapshots from the **a**-axis view of the 8 × 8 × 8 system near the transition state. The molecules in space-fill format outline the local **b/c** axis ratio that is the best reaction coordinate approximation as determined by likelihood maximization.

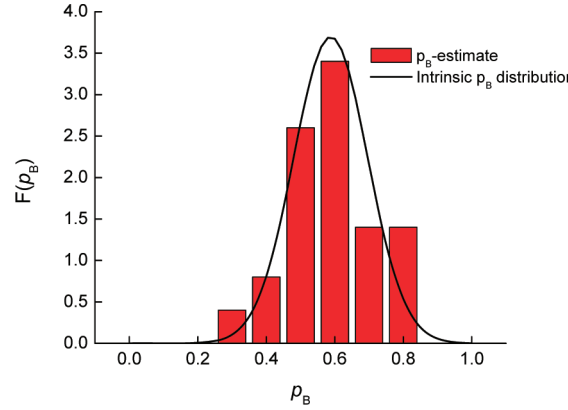

Figure 4. Committor probability histogram for the 8 × 8 × 8 system with a local $\mathbf{b/c}^{\ddagger} = 1.387$, as predicted from the $p_B(r)$ model with a one-dimensional reaction coordinate. $\mu_h = 0.588$ and $\sigma_h = 0.127$. The intrinsic distribution, with $\mu = 0.588$ and $\sigma = 0.108$, is plotted for comparison.

trajectory. The reason for this will be examined in the Discussion. Approximately 500 trajectories are collected for the 10 × 10 × 10 crystal. The optimal reaction coordinate for this system is also a local **b/c** lattice parameter ratio. In this case, it is five layers in the **b** and **c** directions. The reaction coordinate identifies a crystallographically equivalent corner, denoted in Table 1 as "corner 4", as found in the 8 × 8 × 8 system, suggesting a similar mechanism. In addition, the difference in the likelihood scores for the 8 × 8 × 8 reaction coordinate in the 10 × 10 × 10 system is less than the BIC, and vice versa, suggesting that these reaction coordinates will suffice for both systems.

Likelihood maximization is also reapplied to the 7 × 7 × 7 system from the previous study[35] to test the Form II cluster size as a potential reaction coordinate for that system. Table 1 shows that the global **b/c** axis ratio remains the optimal reaction coordinate and that the average Form II cluster size is approximately 100 molecules at the transition state.

The committor probability histogram for the local **b/c** OP for the 8 × 8 × 8 crystal is shown in Figure 4. It is peaked at $\mu_h = 0.588$ and the standard deviation is $\sigma_h = 0.127$. The intrinsic mean and standard deviation are calculated as described by Peters[67] and are $\mu = 0.588 \pm 0.01$ and $\sigma = 0.108 \pm 0.02$, respectively. For reference, a Gaussian distribution with moments from the intrinsic distribution is plotted along with the histogram data. The sharply peaked distribution quantitatively shows that the local **b/c** OP is an accurate approximation to the reaction coordinate.

Figure 5 shows the free energy along the local **b/c** OP in the 8 × 8 × 8 crystal found in likelihood maximization. As the system transforms from the Form I polymorph to Form II, the system overcomes a free energy barrier to nucleation and passes through a metastable intermediate prior to the growth phase. As in our previous study, the barrier to nucleation scales with system size, whereas the barrier to growth remains relatively small at 1−3

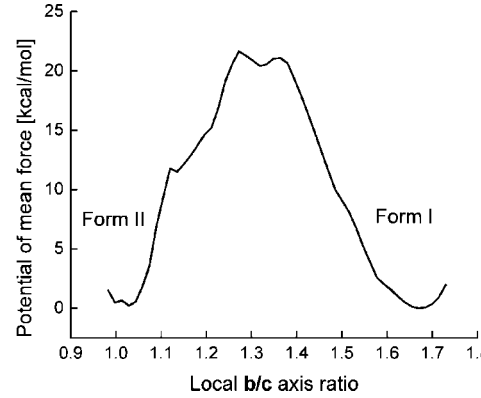

Figure 5. Potential of mean force calculated from MD umbrella sampling as a function of the local **b/c** reaction coordinate from likelihood maximization for the 8 × 8 × 8 system. The weighted histogram analysis method[46] is applied to combine the umbrella sampling windows.

kcal/mol. The nucleation barrier for the 6 × 6 × 6 crystal and for the 7 × 7 × 7 crystal is approximately 4 and 12 kcal/mol, respectively.[35] For the 8 × 8 × 8 system, the free energy barrier to nucleation is approximately 22 kcal/mol.

Discussion

Our results indicate that the global **b/c** lattice parameter is not a good approximation to the reaction coordinate relative to the local **b/c** OPs for the larger systems, even though it is a natural first choice for a reaction coordinate when studying this crystal transformation in vacuum. In searching for an initial trajectory, for instance, MD umbrella sampling over this global **b/c** axis ratio does not yield a realistic pathway for the 8 × 8 × 8 crystal.

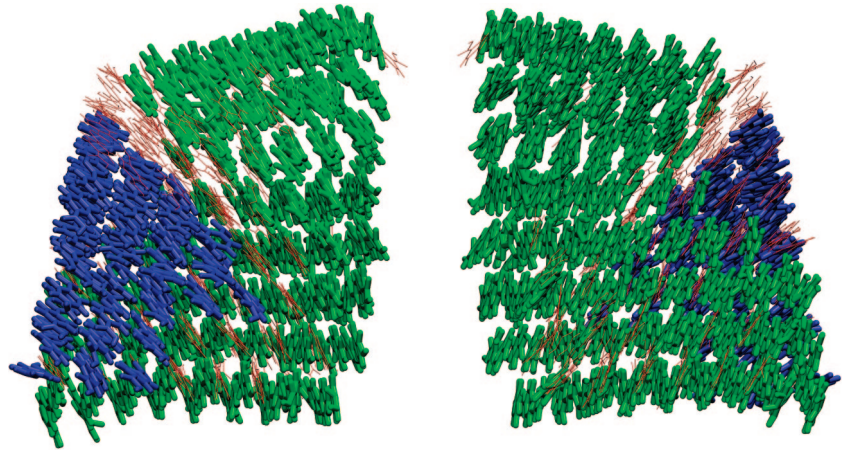

Figure 6. Snapshots from the **a**-axis view of the 8 × 8 × 8 system near the transition state. The molecules in blue are the molecules in the Form II lattice, whereas the green molecules are in the Form I lattice. The wireframe molecules in red form the interface between the two polymorphs.

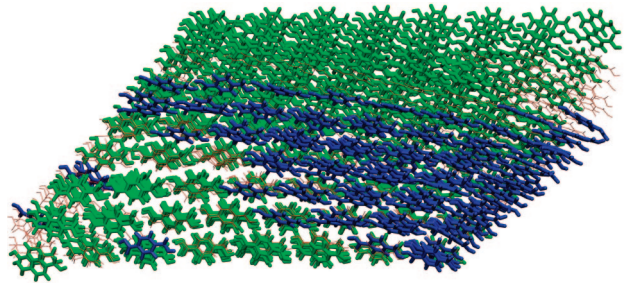

Figure 7. Snapshot from the **b**-axis view of the 8 × 8 × 8 system near the transition state. The molecules in blue are the molecules in the Form II lattice, whereas the green molecules are in the Form I lattice. The wireframe molecules in red form the interface between the two polymorphs.

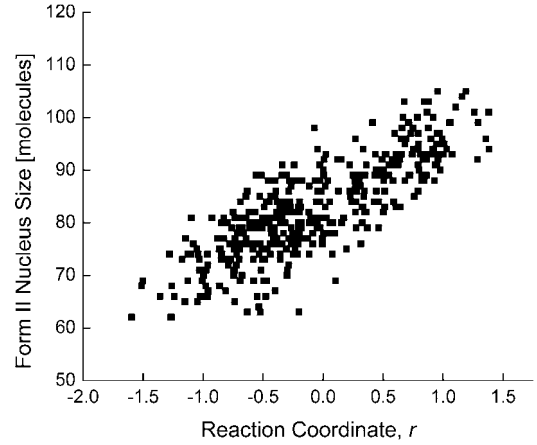

Figure 8. Form II nucleus size as a function of the local **b/c** reaction coordinate for the 8 × 8 × 8 crystal. The product basin is associated with large, positive *r*, and the reactant basin corresponds to large, negative *r*. Therefore, the more stable Form II nucleus decreases in size as the system approaches the reactant, or Form I crystal, and increases as the system nears the product, or Form II crystal. This plot demonstrates that even though the nucleus size is not the optimal reaction coordinate, it remains a useful illustration to explain the location of the reaction coordinate determined from likelihood maximization.

It is initially surprising that the local **b/c** OP in the corner of the crystal, as shown in Figure 3, is the reaction coordinate, rather than an OP similar to the one chosen to obtain an initial trajectory, as shown in Figure 1. The reaction coordinate is surprising because the corner chosen to obtain the initial trajectory appears to undergo significant structural changes near the transition state region. To investigate this finding further, it is useful to label the Form II cluster according to the unit cell OPs for configurations that the best reaction coordinate identified as transition states. Figures 6 and 7 show **a**-axis and **b**-axis views, respectively, of the transition state configuration with the molecules in blue for the Form II unit cell OP and in green for the Form I unit cell OP. The red wireframe molecules have local environments that are not indicative of either the Form I or Form II polymorph. Figures 6 and 7 clearly show that the Form II cluster exists in the corner of the crystal identified by the local **b/c** reaction coordinate. Although the cluster size is not the best reaction coordinate identified by likelihood maximization, it is still useful to illustrate that the Form II cluster size parameter coincides with the local reaction coordinate identified by likelihood maximization. To further verify that the Form II cluster size exhibits correlation with the reaction coordinate, the nucleus size is shown for the 8 × 8 × 8 system as a function of the best reaction coordinate according to likelihood maximization in Figure 8. As expected, the Form II nucleus size increases with increasing *r*, which is concomitant with the product basin, and the nucleus size decreases closer to the reactant basin.

The results presented in Figures 6–8 demonstrate that the global distortion mechanism previously found for the 6 × 6 × 6 and 7 × 7 × 7 systems is not valid for larger TPA crystals. Instead, the reaction coordinate for the 8 × 8 × 8 and 10 × 10 × 10 crystals closely correlates with the nucleus size in the corner of the crystal, thus indicating a localized nucleation mechanism as the system size increases, as hypothesized in our earlier work.[35] Borrowing ideas from classical nucleation theory,[68] the corner nucleation mechanism is reasonable for small crystals because it leads to the smallest possible interfacial area between the two domains for a given nucleus size. Conversely, in the interior of a perfect crystal, the interfacial area for a nucleus of the Form II crystal would be larger for an equivalent nucleus size, and would thus most likely exhibit a substantially higher free energy barrier to nucleation. To our knowledge, the findings in this study represent the first definitive identification of a molecular mechanism for nucleation in a solid state polymorph transformation.

However, we anticipate that as the crystal size increases to experimentally and industrially reasonable crystal sizes, corner nucleation will compete with bulk nucleation because of the disparate scaling in the nucleation rate laws of the surface and bulk mechanisms, i.e., as the crystal size increases, the number

of bulk nucleation sites will increase linearly while the number of available corners remains approximately constant. This scaling difference implies that different mechanistic regimes exist for nucleation based on crystal size. The present work demonstrated a localized nucleation event at a corner for small crystals, but the size range for the change to a bulk nucleation mechanism may be beyond the scope of simulation. However, experiments could investigate our mechanistic predictions. For instance, corners and flat surfaces could be tapped with an atomic force microscope to determine which region is most susceptible to nucleation. Alternatively, functionalized nanoparticles could be designed to bind to specific crystal surfaces such as the exposed carboxylic acid groups of TPA. The functionalized nanoparticles could stabilize or stress the crystal surfaces, which in the case of surface nucleation, could modify nucleation rates as compared to the free crystal.

Single-crystal-to-single-crystal, single-crystal-to-polycrystal, and polycrystal-to-polycrystal polymorph transitions are all likely to begin with isolated nucleation sites. The primary observation of this paper is that nucleation leading to a polymorph transition in small TPA crystals occurs at the edge of a domain. The corner nucleation mechanism can be justified by surface area arguments akin to the molecule independent arguments of classical nucleation theory, so the mechanism presented here for small crystals may apply qualitatively to other small molecular crystals, even those whose specific molecular rearrangements are different.

Although not directly addressed here, nucleation events may also initiate from internal defects within the bulk of the crystal. The number of these internal defects will also scale with the crystal size which may also favor other nucleation mechanisms for large crystals. Investigation of these alternative nucleation mechanisms would be a valuable complement to this study. In addition, further computational studies on polymorph transformations in other molecular crystals, such as those that undergo intramolecular conformational changes during a polymorph transformation, would provide a worthwhile complement to this study to determine if corner nucleation mechanisms are possible transformation pathways in other systems.

Conclusions

This study reveals that as the crystal size of TPA increases beyond 500 molecules, the mechanism changes from a global distortion to a localized nucleation mechanism at a corner of the crystal. Transition states along the reaction coordinate are nonspherical domains of the Form II polymorph in the corner of the crystal. To our knowledge, this represents the first definitive identification of a molecular mechanism for nucleation in a molecular crystal. A cluster analysis shows that the average size of the Form II nucleus at the transition state is approximately 80 molecules and that an interface exists between the two polymorphs. In addition, we show that although the number of molecules in the Form II nucleus is not the best reaction coordinate for the system sizes explored, a strong correlation exists between the best reaction coordinate and nucleus size.

The corner nucleation mechanism for the solid state polymorph transformation in TPA yields the smallest possible interfacial area between the two polymorphs at a given nucleus size, which reduces the surface free energy penalty required to form a new polymorphic domain. Based on these qualitative considerations borrowed from classical nucleation theory, corner nucleation could be a common polymorph transition mechanism in small crystals, but further study on other systems is clearly required. Additionally, we hypothesize that as crystals become significantly larger, corner nucleation may give way to bulk nucleation because bulk nucleation rate laws increase linearly with crystal volume while surface and corner nucleation will be weaker functions of size. Thus, our results imply that different nucleation mechanisms exist at different crystal sizes.

Finally, the recognition of a nonintuitive, but accurate, reaction coordinate highlights the merits of likelihood maximization. Because incorrect reaction coordinates can yield misleading results for quantities of interest, aimless shooting and likelihood maximization are effective ways to obtain correct reaction coordinates as input for free energy calculations[44–50] and pathway optimizations[41,69] that require a priori specification of the reaction coordinate or reaction coordinates components.

Acknowledgment. We thank Merck & Co., Inc. and the Singapore-MIT Alliance for financial support. GTB thanks the National Science Foundation for a Graduate Research Fellowship. BP is supported at the CECAM by the EC through the Marie Curie EXT project MEXTCT-2005-023311. We thank Jane Rempel for reading this manuscript.

Supporting Information Available: Snapshots along a reactive aimless shooting trajectory for the $8 \times 8 \times 8$ crystal. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Mullin, J. W. *Crystallization*, 4th ed.; Butterworth-Heinemann: London, 2001.
(2) Vippagunta, S. R.; Brittain, H. G.; Grant, D. J. W. *Adv. Drug Delivery Rev.* **2001**, *48*, 3–26.
(3) Morissette, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2180–2184.
(4) Thayer, A. M. *Chem. Eng. News* **2007**, *85*, 31–34.
(5) Leusen, F. J. J. *J. Cryst. Growth* **1996**, *166*, 900–903.
(6) Price, S. L. *Adv. Drug Delivery Rev.* **2004**, *56*, 301–319.
(7) Raiteri, P.; Martonak, R.; Parrinello, M. *Angew. Chem., Int. Ed.* **2005**, *44*, 3769–3773.
(8) Mnyukh, Y. V. , *Fundamentals of solid-state phase transitions, ferromagnetism and ferroelectricity*. 1st ed.; Springer: New York, 1998.
(9) Cardew, P. T.; Davey, R. J.; Ruddick, A. J. *J. Chem. Soc., Faraday Trans.* **1984**, *80*, 659–668.
(10) Davey, R. J.; et al. *Nature* **1993**, *366*, 248–250.
(11) Morris, K. R.; et al. *Adv. Drug Delivery Rev.* **2001**, *48*, 91–114.
(12) Zhang, G. G. Z.; et al. *J. Pharm. Sci.* **2002**, *91*, 1089–1100.
(13) Kim, Y. S.; Paskow, H. C.; Rousseau, R. W. *Cryst. Growth Des.* **2005**, *5*, 1623–1632.
(14) Herbstein, F. H. *Acta Crystallogr.* **2006**, *B62*, 341–383.
(15) ten Wolde, P. R.; Ruiz-Montero, M. J.; Frenkel, D. *Faraday Discuss.* **1996**, *104*, 93–110.
(16) Anwar, J.; Boateng, P. K. *J. Am. Chem. Soc.* **1998**, *120*, 9600–9604.
(17) Auer, S.; Frenkel, D. *Nature* **2001**, *409*, 1020–1023.
(18) Berim, G. O.; Ruckenstein, E. *J. Chem. Phys.* **2002**, *117*, 7732–7737.
(19) Moroni, D.; ten Wolde, P. R.; Bolhuis, P. G. *Phys. Rev. Lett.* **2005**, *94*, 235703.
(20) Brendel, K.; Barkema, G. T.; van Beijeren, H. *Phys. Rev. E* **2005**, *71*, 031601.
(21) Pan, A. C. *J. Phys. Chem. B* **2006**, *110*, 3692–3696.
(22) Peters, B.; Trout, B. L. *J. Chem. Phys.* **2006**, *125*, 054108.
(23) Trudu, F.; Donadio, D.; Parrinello, M. *Phys. Rev. Lett.* **2006**, *97*, 105701.
(24) Punnathanam, S.; Monson, P. A. *J. Chem. Phys.* **2006**, *125*, 024508.
(25) Desgranges, C.; Delhommelle, J. *J. Am. Chem. Soc.* **2006**, *128*, 15104–15105.
(26) Desgranges, C.; Delhommelle, J. *J. Chem. Phys.* **2007**, *126*, 054501.
(27) Matsumoto, M. ; Saito, S.; Ohmine, I. *Nature* **2002**, *416*, 409−413.
(28) Radhakrishnan, R.; Trout, B. L. *J. Chem. Phys.* **2002**, *117*, 1786–1796.
(29) Radhakrishnan, R.; Trout, B. L. *J. Am. Chem. Soc.* **2003**, *125*, 7743–7747.
(30) Zahn, D. *Phys. Rev. Lett.* **2004**, *92*, 040801.

(31) Zahn, D. *J. Phys. Chem. B* **2007**, *111*, 5249–5253.
(32) Leyssale, J. M.; Delhommelle, J.; Millot, C. *J. Am. Chem. Soc.* **2004**, *126*, 12286–12287.
(33) Leyssale, J. M.; Delhommelle, J.; Millot, C. *Chem. Phys.* **2005**, *122*, 184518.
(34) Beaucage, P.; Mousseau, N. *Phys. Rev. B* **2005**, *71*, 094102.
(35) Beckham, G. T. *J. Am. Chem. Soc.* **2007**, *129*, 4714–4724.
(36) ten Wolde, P. R.; Ruiz-Montero, M. J.; Frenkel, D. *J. Chem. Phys.* **1996**, *104*, 9932–9947.
(37) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Faraday Discuss.* **1998**, *110*, 421–436.
(38) Dellago, C. *J. Chem. Phys.* **1998**, *108*, 1964–1977.
(39) Dellago, C.; Bolhuis, P. G.; Chandler, D. *J. Chem. Phys.* **1999**, *110*, 6617–6625.
(40) Dellago, C.; Bolhuis, P. G.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 9236–9245.
(41) Maragliano, L. *J. Chem. Phys.* **2006**, *125*, 024106.
(42) Ma, A.; Dinner, A. R. *J. Phys. Chem. B* **2005**, *109*, 6769–6779.
(43) Peters, B.; Beckham, G. T.; Trout, B. L. *J. Chem. Phys*. **2007**, *127*, 1.
(44) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
(45) Kottalam, J.; Case, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 7690–7697.
(46) Kumar, S.; et al. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(47) Kong, X.; Brooks, C. L. *J. Chem. Phys.* **1996**, *106*, 2414–2423.
(48) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.
(49) Radhakrishnan, R.; Schlick, T. *J. Chem. Phys.* **2004**, *121*, 2436–2444.
(50) Ensing, B.; et al. *Acc. Chem. Res.* **2006**, *39*, 73–81.
(51) Garrett, B.C.; Truhlar, D. G. *J. Am. Chem. Soc.* **1979**, *101*, 4534–4548.
(52) Frenkel, D. Smit, B. *Understanding molecular simulations: from algorithms to applications*, 2nd ed.;Computational science series 1; Academic Press: San Diego, 2002.
(53) Peters, B.; Bell, A. T.; Chakraborty, A. *J. Chem. Phys.* **2004**, *121*, 4453–4460.
(54) Moroni, D.; Bolhuis, P. G.; van Erp, T. S. *J. Chem. Phys.* **2004**, *120*, 4055–4065.
(55) van Erp, T. S.; Bolhuis, P. G. *J. Comp. Phys.* **2005**, *205*, 157–181.
(56) Best, R. B.; Hummer, G. *Proc. Natl. Acad. Sci. U.S.A* **2005**, *102*, 6732–6737.
(57) Allen, R. J.; Frenkel, D.; ten Wolde, P. R. *J. Chem. Phys.* **2006**, *124*, 194111.
(58) Bailey, M.; Brown, C. J. *Acta Crystallogr.* **1967**, *22*, 387–391.
(59) Davey, R. J.; et al. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1003–1009.
(60) Desiraju, G. R. *Angew. Chem., Int. Ed.* **1995**, *34*, 2311–2327.
(61) Schlitter, J.; et al. *Mol. Sim.* **1993**, *10*, 291–308.
(62) Phillips, J. C. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
(63) Turnbull, D.; Fisher, J. C. *J. Chem. Phys.* **1949**, *17*, 71–73.
(64) Du, R.; et al. *J. Chem. Phys.* **1998**, *108*, 334–350.
(65) Geissler, P. G.; Dellago, C.; Chandler, D. *J. Phys Chem. B* **1999**, *103*, 3706–3710.
(66) Truhlar, D.G.; Garrett, B. C. *J. Phys. Chem. B* **2000**, *104*, 1069–1072.
(67) Peters, B. *J. Chem. Phys.* **2006**, *125*, 241101.
(68) Debenedetti, P. G. *Metastable liquids: Concepts and principles*, 1st ed.; Princeton University Press; Princeton, N. J., 1996.
(69) Miller, T. F.; Vanden-Eijnden, E.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14559–14564.

JP710192U