

A Computational Study of the Mechanism of the Selective Crystallization of α - and β -Glycine from Water and Methanol–Water Mixture

Jie Chen and Bernhardt L. Trout*

Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E19-502b, Cambridge, Massachusetts 02139, United States

Received: April 30, 2010; Revised Manuscript Received: August 17, 2010

Understanding the control of polymorphism in organic crystals is of paramount importance to the pharmaceutical, chemical, and food industries. In this work, we investigated two mechanisms described in the literature about the selective crystallization of α - and β -glycine from water and from mixtures of water and methanol using molecular simulations. The link hypothesis (*J. Phys. Chem. B* 2008, 112, 7794; *Cryst. Growth Des.* 2006, 6, 1788; *J. Inclusion Phenom. Mol. Recognit. Chem.* 1990, 8, 395; *J. Am. Chem. Soc.* 1986, 108, 5871.), which tries to relate the structure of the polymorph obtained from crystallization to the structure of the prenucleation aggregates in the solutions, says the abundance of glycine cyclic dimers in aqueous solutions leads to the crystallization of α -glycine, the polymorph using cyclic dimers as the packing units. This hypothesis was studied first. We revisited the self-assembly of glycine molecules in solution using molecular dynamics to address the debate (*Phys. Rev. Lett.* 2007 99, 115702; *J. Phys. Chem. B* 2008 112, 7280; *J. Am. Chem. Soc.* 2008 130, 13973.) about which is the dominating species in the glycine aqueous solutions and whether there is a link between the solution chemistry and the polymorphic outcome of crystallization. The structures of the glycine clusters were characterized using a structural parameter called cyclic dimer fraction. The glycine clusters in methanol–water mixtures have higher cyclic dimer compositions than those in the pure aqueous solutions. Moreover, the glycine open-chain dimer is more stable than the cyclic dimer regardless of the presence of methanol. All these suggest that the link hypothesis does not work for the polymorphic system of glycine, and the selective crystallization of α - and β -glycine from water and methanol–water mixture, respectively, is not due to the abundance of glycine aggregates in the solution phase with a similar structure to the crystallizing solid form. The hypothesis of the methanol inhibition on the growth of α -glycine {010} and {01̄0} faces, proposed by Weissbuch (*Angew. Chem., Int. Ed.* 2005, 44, 3226.), was also studied. The interfaces between the {010} and {01̄0} faces of both crystal forms (α and β) and both solvents (water and methanol–water 3:7 mixture) were studied using molecular simulation. No strong binding of methanol onto the {010} and {01̄0} faces of both crystal forms was observed, and the addition of methanol dilutes the crystal–solvent interactions on all faces. Therefore, the selective crystallization of β and α -glycine with and without methanol does not follow either of the two mechanisms in the literature.

1. Introduction

Polymorphism^{1–8} is the ability of a solid material to exist in more than one form or crystal structure, while retaining the same chemical composition. This phenomenon leads to significant variability in the properties of products and continues to pose a challenge to scientists and engineers in producing crystal products of consistent quality. Understanding the control of polymorphism in organic crystals is of paramount importance to pharmaceutical, chemical, and food industries and is also of theoretical relevance in solid-state chemistry and physics. For example, the existence of polymorphism in the case of antiviral drug Ritonavir had a dramatic commercial effect on pharmaceuticals.^{9–11} In this study, we focused on the effects of solvent on polymorph selection of glycine.^{2,8,12,13}

1.1. Polymorphism of Glycine. Glycine, as a simple organic molecule with only 10 atoms, is of great interest to researchers for the study of polymorphism both experimentally and computationally. It has three forms, α , β , and γ , with thermodynamic stability in the order $\gamma > \alpha > \beta$.¹⁴ α -Glycine is packed in

centrosymmetric dimers and crystallizes from aqueous solutions,¹⁵ β -glycine is packed through a 2-fold screw symmetry axis and crystallizes from alcoholic aqueous solutions,¹⁶ and γ -glycine is packed in helical chains around the 3-fold screw axes parallel to the *c*-axis and crystallizes from acidic aqueous solutions and sometimes from aqueous solutions as well with a rate about 500 times slower than α -glycine.¹⁷ In this work, our focus was on understanding the selective crystallization of α - and β -glycine from water and water–methanol mixtures, respectively.

1.2. Link Hypothesis and the Controversy. The link hypothesis suggests that molecules self-assemble into various structures in the solution phase and the most stable/populated one gets carried out in the crystallization process.^{1–4} This hypothesis successfully explained the polymorphism of tetrolic acid in previous studies.^{1,2} Both experimental and molecular simulation approaches have been applied to study the formation of glycine clusters in aqueous solutions and to explore the relationship between the structures of the clusters in solution formed through self-assembly and those of the crystals obtained from solution crystallization. However, this leads to a debate of whether such a link exists. Myerson and the co-workers

* To whom correspondence should be addressed. E-mail: trout@mit.edu.
Tel: (617) 258-5021. Fax: (617) 253-2272.

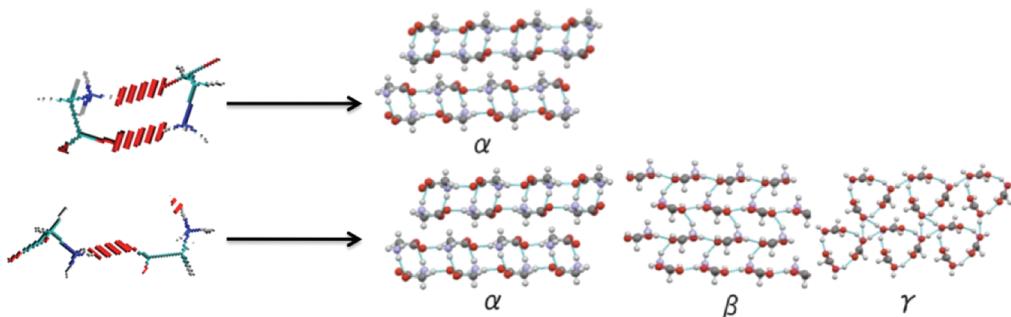


Figure 1. Link hypothesis for the polymorphic system of glycine. An open-chain dimer structure can be found in all three forms, while the cyclic dimer is only seen in the α -glycine. Therefore, only the cyclic dimer is the corresponding growth synthon in solutions for α -glycine by the argument of the link hypothesis.

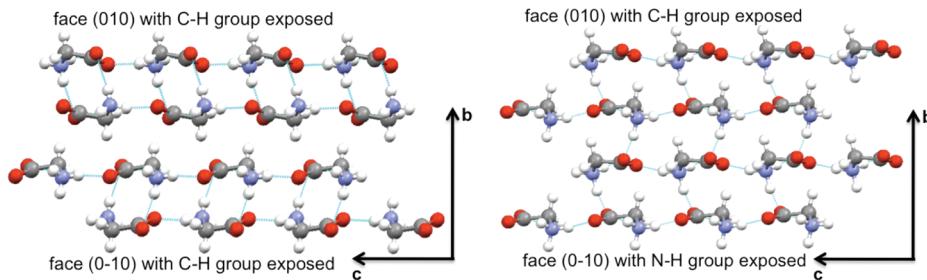


Figure 2. $\{010\}$ and $\{0\bar{1}0\}$ faces of α - and β -glycine, reproduced from Weissbuch's work:⁸ (a, left) α form exposing C–H bonds to the solution on both faces; (b, right) β form exposing C–H bonds to the solution on $\{010\}$ face and N–H bonds to the solution on $\{0\bar{1}0\}$ face. Reproduced with permission from ref 8. Copyright 2005 John Wiley & Sons, Inc.

performed diffusion-coefficient measurements¹⁸ and small-angle X-ray scattering (SAXS) experiments^{19,20} of the supersaturated aqueous solutions of glycine and suggested that the majority of glycine molecules exist as dimers in the aqueous solutions. Therefore, they concluded, glycine crystallizes as the cyclic dimer based α form. In contrast, Hamad et al. performed molecular dynamics study of various glycine aqueous solutions (from undersaturated to supersaturated) using the AMBER force field and suggested that the glycine monomer is the dominating species in glycine aqueous solutions near room temperature over a wide concentration range.^{6,21} Yu and co-workers performed freezing-point depression studies of glycine aqueous solutions and suggested that glycine exists mainly as monomers, not dimers, near 0 °C.⁷ Their results showed that approximately 25% of glycine molecules exist as dimers in 2.92 mol/(kg of H₂O) glycine solution if the osmotic abnormality is entirely attributed to dimerization. In contrast to Myerson and co-workers' report of progressively slowing diffusion in saturated glycine solutions, Yu and co-workers found no such slow down using PGSE NMR. Although the debate about what is the dominating species in glycine aqueous solution is still going on, there is a common shortfall in all these studies. They all focused only on the size of the clusters but not on structural information, with no differentiation between a cyclic and an open-chain structure when defining a dimer. From the link hypothesis, only the formation of more cyclic dimers in the solution leads to the crystallization of α -glycine which utilizes them as the packing units, while the open-chain dimers can be the building units for any of the α -, β -, and γ -glycine (as shown in Figure 1). Although Hamad et al. mentioned that the main type of interaction between glycine molecules was the single N–H...O–C hydrogen bonds and that the double hydrogen bonded dimers were only observed infrequently, they lumped these two structures together when characterizing the glycine clusters and categorized the clusters by their sizes: monomers, dimers, and trimers as clusters of one, two, and three molecules, respectively. In this work, we focus on the structural charac-

teristics of these glycine clusters. We compared the stability of the open-chain dimer and the cyclic dimer quantitatively and characterized the cyclic dimer composition of the overall glycine clusters using a structural parameter, dimer fraction. Moreover, very little information about the glycine clusters in the water–alcohol mixtures was gathered, which could, from another perspective, provide us additional knowledge of the applicability of the link hypothesis in understanding the polymorphism of glycine.

1.3. Hypothesis of the Methanol Inhibition on the α -Glycine $\{010\}$ and $\{0\bar{1}0\}$ Faces. α -Glycine shows symmetric $\{010\}$ and $\{0\bar{1}0\}$ faces with C–H bonds exposed, while β -glycine has asymmetric $\{010\}$ and $\{0\bar{1}0\}$ faces with C–H bonds exposed on the former and N–H bonds exposed on the latter, as shown in Figure 2. On the basis of this difference in crystal structures, Weissbuch et al.⁸ suggested that methanol molecules would bind strongly onto both the $\{010\}$ and $\{0\bar{1}0\}$ faces of α -glycine through the interactions between C–H bonds of glycine and the alkyl chains of the alcohol; they would only strongly bind onto the $\{010\}$ face of β -glycine with C–H bonds exposed but not the $\{0\bar{1}0\}$ face with the N–H bonds exposed. Therefore, they proposed that methanol inhibits the growth of both faces of the α form but only one face of the β form and thus leads to the crystallization of the latter polymorph. However, there has been no direct evidence of this hypothesis reported in the literature so far.

In this work, we studied these two mechanisms about the selective crystallization of α - and β -glycine with and without methanol using molecular simulations. We tested the link hypothesis first. We revisited the self-assembly of glycine in solutions using molecular dynamics and studied the structure of glycine clusters formed in both water and water–methanol mixtures. We also studied the impact of the presence of the α -glycine $\{010\}$ face on the structure of glycine clusters in its vicinity. This surface is composed of centrosymmetric double layers that are held firmly intact by hydrogen bonds between cyclic molecular pairs and is also of morphological significance.

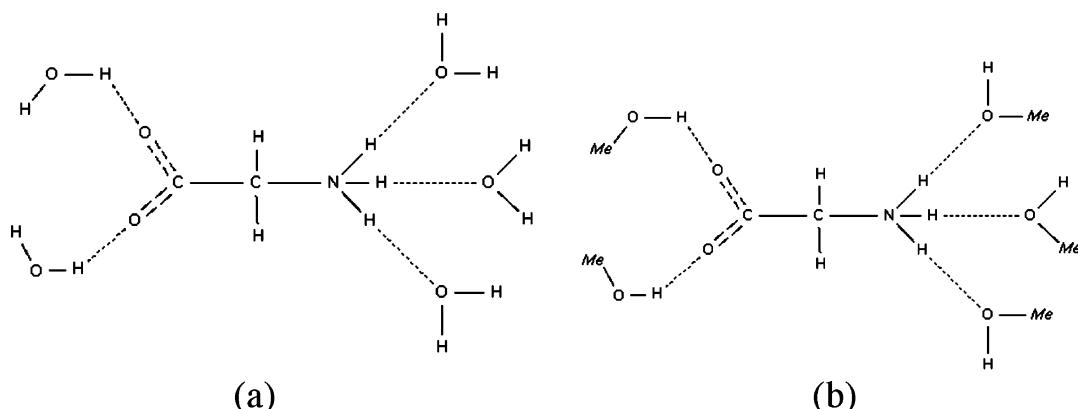


Figure 3. Schematic demonstration of the calculations of the interaction energies between the hydrogen-bonding sites of (a) glycine–water and (b) glycine–methanol.

After that, we studied the solvent–crystal interactions for the {010} and {010} faces of both glycine forms in both water and methanol–water mixture to test the selection mechanism proposed by Weissbuch: different level of methanol inhibition in the *b*-direction leads to the selective crystallization of α - and β -glycine. Our goal is purely to test this hypothesis, instead of studying whether the crystallization of the β form over the α form in methanol containing solutions is a growth-controlled process, in which focusing on the fast growing faces probably makes more sense.

2. Force Field Development

In order to investigate the influence of solvent on the polymorphic outcome of glycine, a potential that can not only reproduce three glycine polymorphs at ambient temperature but also describe glycine–solvent interactions equally well is required. Methanol molecules are modeled using the OPLS-AA potential,²² which has a good library of organic solvents and is optimized for liquid simulation. Water molecules are modeled using the TIP3P potential²³ with the LJ parameters obtained from the OPLS-AA force field. To maintain the consistency among all the potentials, it would be ideal to model glycine molecules using the OPLS-AA potential as well, if it satisfied the two criteria mentioned above. As noticed in many previous studies, the partial charges are the most influential parameters in modeling crystals.²⁴ We carefully tested two charge sets for the glycine zwitterion. The first charge set uses the assembled OPLS-AA charges, with the charges for the NH₃⁺ and COO[−] taken directly from glycine terminal groups, the charges for the CH₂ group hydrogen atoms taken from glycine amino acid residue, and the charge of C_α adjusted to give a neutral zwitterion. This approach was previously employed by Price in their study of glycine crystals using the AMBER potential.²⁵ The second charge set uses the CHELPG charges of glycine zwitterion obtained directly from ab initio calculations. Both of these two charge sets are directly available in the OPLS-AA force field. Although the OPLS-AA potential was preferred for glycine to maintain consistency with those of methanol and water molecules, we also investigated another candidate potential for glycine, the AMBER potential, which was shown previously to be able to reproduce the crystal structures of the three polymorphs of glycine and to correctly predict their relative energy ranking.²⁵

2.1. Force Field Validation Methods. We performed two tests to compare the OPLS-AA potential with assembled charges, the OPLS-AA potential with CHELPG charges and the AMBER potential. In the first test, the ability of these three

potentials to reproduce the crystal structure of all three polymorphic forms of glycine at ambient temperature was investigated. Molecular dynamics was applied to simulate crystal supercells in a NPT ensemble (298 K and 1 atm) with periodic boundary conditions. All MD simulations in this work were performed using CHARMM.^{26,27} Simulation boxes of 6a × 3b × 6c unit cells for α (432 molecules), 7a × 5b × 7c unit cells for β (490 molecules), and 5a × 5b × 6c unit cells for γ (450 molecules) were built to ensure that all simulation box edges were at least 28 Å, 2 times the cutoff distance for nonbonded interactions. The particle mesh Ewald summation method was used to correct for the long-range electrostatic interactions. When the AMBER potential was used, an arithmetic average combination rule was employed, the charge–charge interactions between atoms separated by three bonds (1–4 interactions) were scaled by 1.2 and the corresponding van der Waals interactions were scaled by 2. When the OPLS-AA potential was used, a geometric average combination rule was employed and the nonbonded 1–4 interactions were scaled by 2. Monoclinic crystal type was used for both α and β form which has 4 degrees of freedom (*a*, *b*, *c*, and β). Hexagonal crystal type was used for the γ form which has 2 degrees of freedom (*a*, *c*). All simulation boxes were sampled every 1 ps for 1 ns (with 1 fs time step), after a 1 ns equilibration run. The percentage change of the lattice parameters (PCLPs) and root mean squared difference (rmsd), compared to the experimental X-ray structure, were calculated using the average structure of the 1000 frames sampled in the production run. The relative energies of three glycine forms were calculated by averaging the energies sampled every 1 ps, normalizing using the number of molecules in the simulation box and resetting the lowest value to zero.

In the second test, the interaction energies of glycine–water and glycine–methanol were investigated. Hydrogen-bonding energies between glycine and water/methanol molecules were calculated following the approach developed by MacKerell.²⁸ In this approach, a water/methanol molecule was placed in the vicinity of a hydrogen-bonding site of a glycine molecule with fixed orientation. The configurations of glycine and water molecules were obtained from the α -glycine X-ray structure and MacKerell's work, respectively. The configuration of methanol was optimized using MP2/6-31G* in vacuum. The potential energy of the pair was optimized with respect to the distance between the water/methanol molecule and the glycine molecule with all the other internal coordinates fixed, as shown in Figure 3. The hydrogen-bonding energy was calculated using the optimized configuration. This procedure was performed for all the hydrogen-bonding sites of glycine including the two oxygen

TABLE 1: Percentage Change of Lattice Parameters (PCLPs), the Root Mean Squared Difference (Rmsd), and the Relative Energies of the Three Polymorphs of Glycine, Simulated Using MD at 298 K and 1 atm

| | | PCLP <i>a</i> (%) | PCLP <i>b</i> (%) | PCLP <i>c</i> (%) | PCLP β (%) | rmsd (\AA) | rel energies (kJ/mol) |
|-----------------------------|---------------|-------------------|-------------------|-------------------|------------------|-----------------------|-----------------------|
| OPLS with assembled charges | α form | −2.1 | 2.6 | −3.7 | 0.6 | 0.5 | 0.0 |
| | β form | −3.3 | 4.8 | −4.8 | 0.9 | 0.8 | 4.2 |
| | γ form | −0.3 | −0.3 | −1.6 | | 1.3 | 5.0 |
| OPLS with CHELPG charges | α form | −1.2 | 4.0 | −2.4 | 0.3 | 0.5 | 0.0 |
| | β form | −2.1 | 6.0 | −3.4 | 0.6 | 0.7 | 3.0 |
| | γ form | 0.5 | 0.5 | −0.8 | | 0.5 | 3.3 |
| AMBER | α form | −3.6 | 7.8 | −2.9 | 0.7 | 1.1 | 0.0 |
| | β form | 0.5 | 3.3 | −3.9 | 0.4 | 0.6 | 0.5 |
| | γ form | −1.7 | −1.7 | −1.3 | | 0.3 | 0.9 |

TABLE 2: Hydrogen Bond Energies between Glycine and Water/Methanol, Calculated Using Both the Ab Initio Method and the Empirical Force Field Method

| ab initio (MP2) energy (kJ/mol) | OPLS-AA w/assembled charges | | OPLS-AA w/CHELPG charges | | AMBER | | |
|------------------------------------|-----------------------------|--------------------------|--------------------------|--------------------------|-----------------|--------------------------|-------|
| | energy (kJ/mol) | dev from MP2 (kJ/mol) | energy (kJ/mol) | dev from MP2 (kJ/mol) | energy (kJ/mol) | dev from MP2 (kJ/mol) | |
| O1(gly)–H(wat) | −46.5 | −59.4 | −13.0 | −49.8 | −3.8 | −56.5 | −10.5 |
| O2(gly)–H(wat) | −42.3 | −56.1 | −13.8 | −46.9 | −4.6 | −53.6 | −11.7 |
| H1(N,gly)–O(wat) | −53.2 | −62.0 | −8.4 | −52.7 | 0.4 | −54.0 | −0.4 |
| H2(N,gly)–O(wat) | −49.4 | −49.0 | 0.4 | −41.9 | 7.5 | −44.4 | 5.0 |
| H3(N,gly)–O(wat) | −52.3 | −52.3 | 0.0 | −44.4 | 7.5 | −47.3 | 5.0 |
| O1(gly)–H(mth) | −48.1 | −57.8 | −10.0 | −49.4 | −1.7 | −56.1 | −8.0 |
| O2(gly)–H(mth) | −44.8 | −54.8 | −10.5 | −46.9 | −2.1 | −53.6 | −8.8 |
| H1(N,gly)–O(mth) | −53.6 | −61.1 | −7.5 | −52.7 | 0.8 | −54.4 | −0.8 |
| H2(N,gly)–O(mth) | −49.8 | −51.1 | −0.8 | −44.0 | 5.9 | −46.9 | 2.9 |
| H3(N,gly)–O(mth) | −49.8 | −52.7 | −2.9 | −46.0 | 4.2 | −49.0 | 0.8 |

sites of the COO^- group and the three hydrogen sites of the NH_3^+ group, using all the three empirical potentials. An ab initio method (MP2/6-31G*) was used as a reference.

2.2. Force Field Validation Results. The MD-calculated structures of the three polymorphs of glycine clearly depend on the choice of charges and the choice of potential models, as suggested by the PCLP and rmsd values listed in Table 1. Although it succeeded in reproducing the stability order of the three polymorphs in Price's fixed cell angle simulations, the AMBER potential failed in our simulations with adjustable angle β for α - and β -glycine. This additional degree of freedom in angle β , which monoclinic cell type should have, gave lower energy values for α - and β - than for γ -glycine, which is inconsistent with their experimental stability order. A similar trend was also obtained when the other two OPLS potentials were used. Therefore, we could hardly draw any conclusion about which potential is superior on the basis of their ability to predict the relative stability of the three glycine forms. A close examination of the PCPL values shows that all the three potentials gave similar deviations in the lattice parameters. The biggest deviation was the cell parameter *b* of α -glycine, which increased by 7.8% when simulated using the AMBER force field. Rmsd is usually a good indicator of whether the ensemble-average structure in MD can preserve the local interactions in the experimental X-ray structure, such as hydrogen bonds. The OPLS-AA potential with CHELPG charges gave reasonable rmsd values for all three forms of glycine, while the OPLS-AA potential with assembled charges and the AMBER potential gave significantly higher rmsd values for γ - and α -glycine, respectively. With that, we conclude that the three potentials tested in this work perform similarly in simulating glycine crystals with the OPLS-AA potential with CHELPG charges being slightly better in preserving local structures.

Hydrogen bond energies between glycine and water/methanol molecules calculated using both ab initio method and the empirical force field method are listed in Table 2. When glycine

is modeled using OPLS-AA potential with CHELPG charges, the differences between the hydrogen bond energies calculated using the empirical force field method and the ab initio MP2 method are the smallest on average for both water and methanol.

Combining the results of the solid-phase simulations of glycine crystals and the hydrogen bond energies between glycine and water/methanol, we decided to use the OPLS-AA potential with CHELPG charges for glycine.

3. Link Hypothesis

3.1. Computational Details. The thermodynamic stabilities of a cyclic dimer, an open-chain dimer, and two fully solvated glycine monomers in both aqueous and methanol–water solutions were compared by performing the free energy calculations using the umbrella sampling technique which directs two fully apart glycine molecules to move closer along a particular axis (order parameter) by adding a harmonic biasing potential. The order parameter (OP) used to direct the sampling process is defined as the average of two distances between the carboxyl carbon of one glycine molecule and the nitrogen of the other one. For each simulation, two glycine molecules were inserted into a pre-equilibrated solvent box. To adequately sample, one glycine molecule is confined in the center of the simulation box, using a harmonic restraint with a force constant of 836.8 kJ/(mol· \AA^2). A harmonic functional form of an umbrella potential is used, as shown below

$$U = k_u(\delta - \delta_0)^2$$

where k_u is the harmonic force constant with units kJ/(mol· \AA^2) and δ_0 is the equilibrium point of the sampling window.

Molecular dynamics simulations for both glycine aqueous and 30% (mole fraction) methanol–water solutions at various concentrations (Table 3) were conducted at 298 K and 1 atm to

TABLE 3: Details of the Glycine Solutions Used in the MD Simulations

| glycine–water solutions | | | glycine–methanol–water solutions | | | |
|-------------------------|-----------|----------------|----------------------------------|-----------|--------------|----------------|
| no. glycine | no. water | concen (mol/L) | no. glycine | no. water | no. methanol | concen (mol/L) |
| 81 | 4500 | 1.0 | 81 | 3150 | 1350 | 0.7 |
| 144 | 4000 | 1.9 | 144 | 2800 | 1200 | 1.4 |
| 216 | 4000 | 2.7 | 216 | 2800 | 1200 | 2.0 |
| 256 | 4000 | 3.1 | 256 | 2800 | 1200 | 2.4 |
| 324 | 4000 | 3.8 | 324 | 2800 | 1200 | 2.9 |

study the structure of the glycine clusters formed in solutions. The reported experimental solubility of α -glycine is 25 g/100 mL in water at 298 K and reduces to 2.9 g/100 mL in 50% (v/v) methanol–water mixture. The concentrations used in this study range from undersaturation to supersaturation. The cyclic dimer fraction of the hydrogen bond network among all the glycine molecules was used to characterize the cluster structure. It was defined as

$$\text{cyclic dimer fraction} = \frac{\text{total no. of H bonds in cyclic dimers}}{\text{total no. of H bonds formed among all glycine molecules}}$$

For example, if there are only four glycine molecules in the system and they exist in a conformation shown in Figure 4, there are two hydrogen bonds in a cyclic dimer and four hydrogen bonds in all glycine molecules. Thus, the cyclic dimer fraction is 0.5. The cyclic dimer fractions of α -, β -, and γ -glycine crystals are 0.33, 0.0, and 0.0, respectively. The hydrogen bonds between glycine molecules are defined as the (C)O–H(N) distance no larger than 2.2 Å and the N–H–O(C) angle no less than 140°, which is same as the criteria used in the previous computational study performed by Hamad. Each simulation box was equilibrated for 2 ns and sampled for 14 ns to calculate cyclic dimer fractions with periodic boundary conditions and the particle mesh Ewald summation method to correct for the long-range electrostatic interactions. The time step used in simulation was 1 fs and the samples were taken every 1000 steps (1 ps interval).

The impact of the presence of α -glycine crystal face {010} on the glycine cluster structures in the aqueous solutions was also investigated. This crystal surface is composed of double layers of centrosymmetric dimers, similar to the double layer structure of membrane. The aim is to check if there is any increase in the cyclic dimer fraction of the glycine clusters in solution when contracting with an ordered crystal surface with bilayer cyclic dimer structure. A crystal–solution–crystal sandwich was set up, with the α -glycine {010} face contacting with the glycine aqueous solution (Figure 5). Two 10a × 2b × 10c α -glycine crystals were placed on each side of the solution box. Five glycine aqueous solutions were used here, and the number of glycine and water molecules was kept exactly same

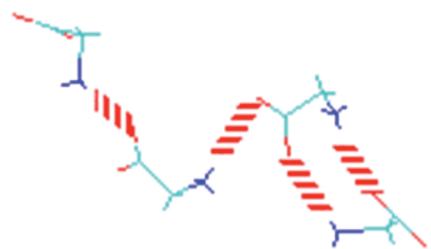


Figure 4. An example of the calculation of the cyclic dimer fraction (cyclic dimer fraction = 0.5). Hydrogen bonds among glycine molecules are in red lines.

as the ones in the previous solution studies for direct comparison. The study of the methanol–water solutions was not performed, since α -glycine does not grow from methanol–water mixtures. To prevent the glycine molecules on the crystal surface from diffusing into the solution phase, a harmonic restraint with force constant 8.36 kJ/mol was applied to the rmsd of the non-hydrogen atoms in the crystal from a reference structure which was calculated by pre-equilibrating the crystal super cell at 298 K and 1 atm and extracting the average. A 16 ns MD simulation was performed for the simulation box in a NPT ensemble (298 K and 1 atm) with 3D periodic boundary conditions and the particle mesh Ewald summation method. After equilibration (the first 2 ns simulation), the thickness of the solution phase in the y direction was around 56 Å, which is larger than twice the cutoff distance for the nonbonded interactions (14 Å). The dimer fraction of the glycine molecules in the solution phase was calculated using the frames taken from the last 14 ns run with 1 ps sampling intervals.

3.2. Results and Discussion. Free energy profiles of the formation of a cyclic dimer from two fully solvated monomers in both water and methanol–water 3:7 mixture are plotted in Figure 6. On both curves, there is a dip near OP = 3.7 Å, which corresponds to the cyclic dimer basin. When the OP is greater than 6.0 Å, the free energy curve is almost flat, which corresponds to two completely separated (fully solvated) glycine molecules. The dip near OP = 5.2 Å corresponds to the intermediate open-chain dimer, where only one hydrogen bond is formed between two glycine molecules. The basin positions

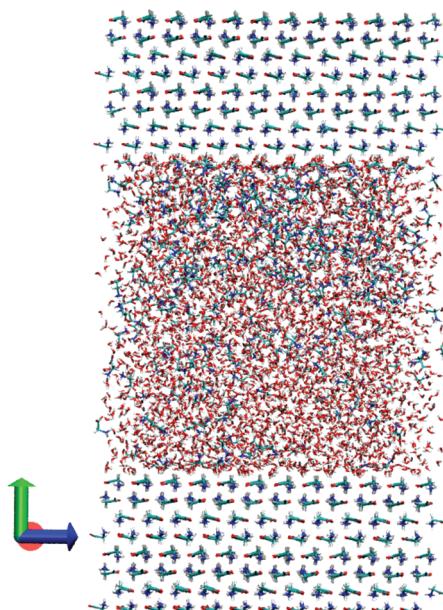


Figure 5. Simulation box of the 3.8 mol/L glycine aqueous solution contacting with α -glycine {010} face, viewed along the a-axis of the crystal. There are 324 glycine molecules and 4000 water molecules in the center solution phase, 800 glycine molecules in the upper crystal phase, and 800 glycine molecules in the lower crystal phase.

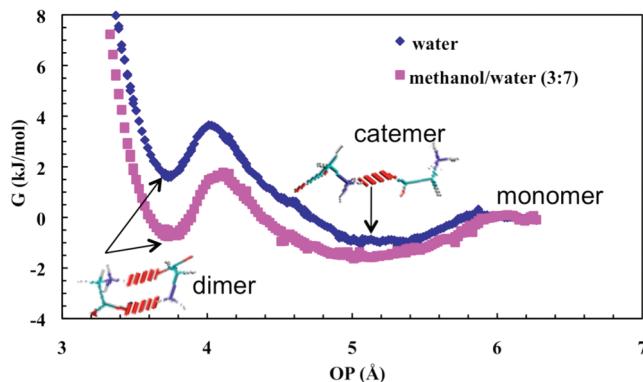


Figure 6. Free energy profiles of the glycine dimerization reaction in water and methanol–water 3:7 mixture at 298 K and 1 atm obtained by using MD umbrella sampling.

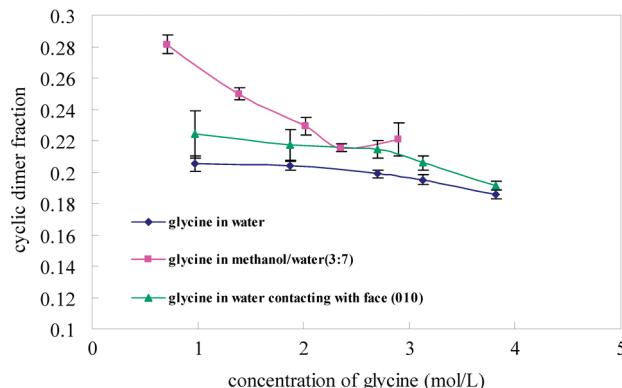


Figure 7. Cyclic dimer fractions of the glycine hydrogen bond network in solution at 298 K and 1 atm with and without contacting with the {010} face of α -glycine. The distance cutoff for hydrogen bond is $(\text{N})\text{H}-\text{O}(\text{C}) < 2.2 \text{ \AA}$ and the angle cutoff for hydrogen bond is $\text{N}-\text{H}-\text{O}(\text{C}) > 140^\circ$.

do not shift significantly when solvents are switched. In both water and methanol–water 3:7 mixture, the open-chain dimer is more stable than both the cyclic dimer and the monomer and the free energy gain for the formation of an open-chain dimer from two monomers is slightly larger in methanol–water mixture than in pure water. The free energy difference between the cyclic dimer and open-chain dimer is about 0.9 and 2.5 kJ/mol in methanol–water mixture and water, respectively, which implies that the cyclic dimer is more ready to form in the former. This contradicts the link hypothesis, by which the cyclic dimers should be destabilized with the addition of methanol.

Cyclic dimer fractions of the hydrogen bond network of glycine in various solutions are plotted in Figure 7. Cyclic dimers exist in both pure and methanol containing solutions over the concentration range studied here. The choice of solvent clearly affects the dimer composition of the hydrogen bond network of glycine. The cyclic dimer fractions are higher in the methanol–water solutions than those in the aqueous solutions with the same concentration, which suggests that the glycine clusters in the methanol–water mixtures assemble the structure of α -glycine more than those formed in pure water. This is consistent with the previous results of the free energy calculations: cyclic dimers are more stable in water–methanol mixtures. Moreover, the dimer fraction decreases as the concentration of glycine increases, while the number of hydrogen bonds formed per glycine molecule increases with concentration, as shown in Figure 8. This indicates that the hydrogen bond network among all glycine molecules tends to propagate faster in a form other than the cyclic dimers and the

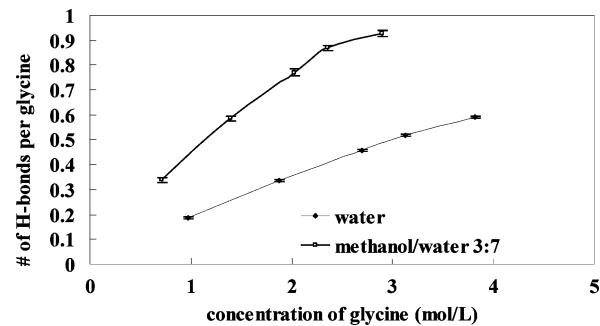


Figure 8. Number of hydrogen bonds per glycine molecule in the simulation box, calculated at 298 K and 1 atm.

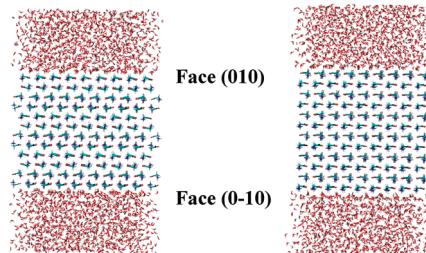


Figure 9. Simulation boxes of the {010} and ({01̄} faces of α -glycine (left) and β -glycine (right) contacting with water.

cyclic dimer structure is less favored in the supersaturated solutions. The fact that the number of hydrogen bonds formed per glycine molecule is larger in the methanol–water solutions than in the pure aqueous solutions again suggests that glycine molecules tend to aggregate more easily in the former. All these facts strongly point to that cyclic dimers are more favored in the methanol–water mixtures than in the pure water and therefore that the link hypothesis does not work here.

The cyclic dimer fractions of glycine molecules in aqueous solutions contacting with α -glycine crystal face {010} are plotted in Figure 5 as well. It is clear that the presence of the crystal surface with the cyclic dimer double layer increases the cyclic dimer composition in the solution phase. The increase is consistent for all the concentrations studied here. Although this increase can probably facilitate the growth of α -glycine, it is not large enough to strongly support the link hypothesis or to explain the selective crystallization of α - and β -glycine since the cyclic dimer fractions of glycine clusters in methanol–water 3:7 mixtures are still higher.

4. Hypothesis of the Methanol Inhibition on the α -Glycine {010} and {01̄} Faces

4.1. Computational Details. Molecular dynamics was applied to study the solvent–crystal interactions on the {010} and {01̄} faces of α - and β -glycine in both water and methanol–water 3:7 mixtures. Each simulation cell consists of a solvent–crystal–solvent slab with solvent molecules contacting the {010} and {01̄} faces, as shown in Figure 9. The crystal in the center consists of 768 molecules and is constructed from an $8a \times 3b \times 8c$ and $8a \times 6b \times 8c$ supercell for α and β , respectively. The thickness of the crystal in the b direction is about 37 Å, which is more than twice the cutoff for nonbonded interactions (14 Å) to ensure that the solvent molecules near the {010} face do not interact with the solvent molecules near the {01̄} face. The total number of solvent molecules in the system is 2439 for water simulation and 2000 for methanol–water simulation with 30% as methanol. Solvent molecules are evenly distributed on both sides of the crystal in the initial configuration. A 2 ns

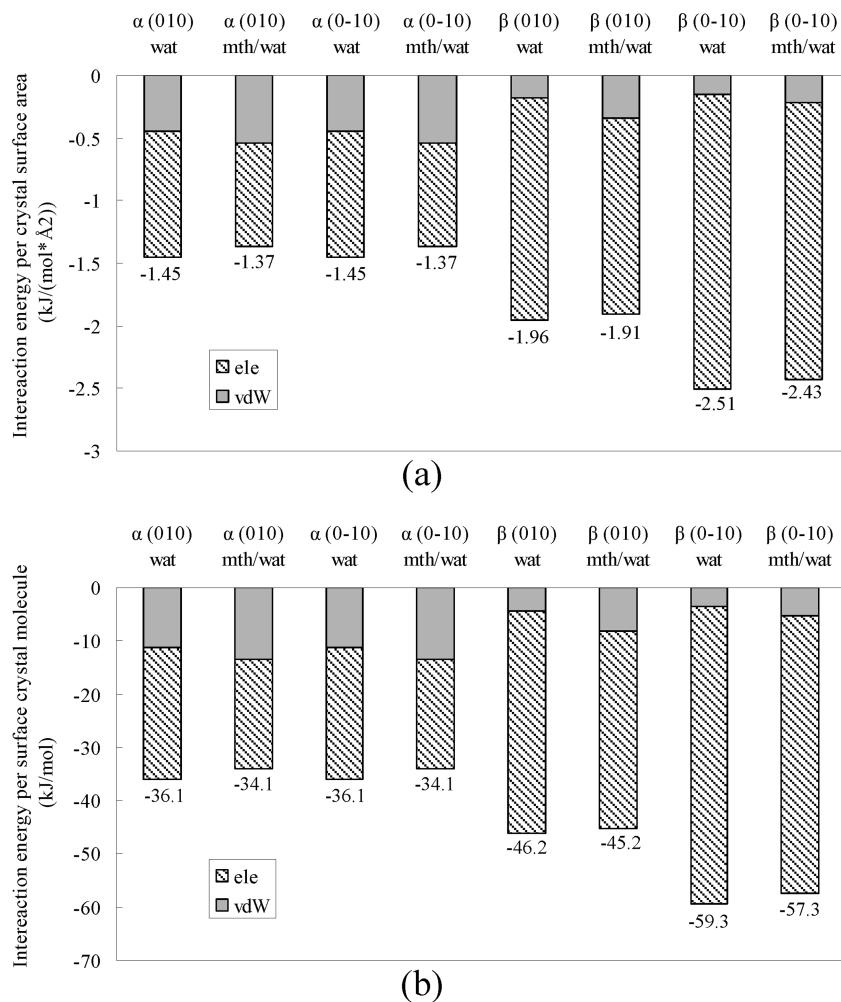


Figure 10. Solvent–crystal interaction energies on face {010} and face {0̄10} of both α - and β -glycine for water and methanol–water mixtures calculated using MD simulation at 298 K and 1 atm (a) normalized based on surface area, 1594.05 \AA^2 for α -glycine and 1511.84 \AA^2 for β -glycine, and (b) normalized based on the number of glycine molecules on crystal surface, 64 for both α - and β -glycine. The numbers labeled on the plots are the normalized total interaction energies.

MD simulation in a NPT ensemble (298 K and 1 atm) with periodic boundary conditions and the particle mesh Ewald summation method were performed to equilibrate the system. The center of mass of the crystal was placed at the origin, and images were updated for all molecules around the origin before the simulation and only for solvent molecules during the simulation. To prevent the glycine molecules on the crystal surface from diffusing into the solution phase, a harmonic restraint with a force constant of 8.36 kJ/mol was again applied to the root mean squared difference of those non-hydrogen atoms in the crystal from a reference structure which was calculated by pre-equilibrating the crystal super cell at 298 K and 1 atm and extracting the average. The height of the solvent layer on both sides is at least 20 \AA for all simulations, which is again larger than the 14 \AA cutoff used for nonbonded interactions. After equilibration, another 2 ns production run was harvested to calculate the solvent–crystal interactions, with samples taken every 1 ps. Crystal–solvent interactions are calculated as the pairwise summation of the electrostatic and van der Waals interactions between the crystal molecules and the solvent molecules which are less than 14 \AA away (the cutoff for nonbonded interactions), since any solvent molecules farther away than that do not contribute to the interaction energies.

4.2. Results and Discussion. The interaction energies normalized based on both the total surface area and the number of glycine molecules in the first crystal surface layer are plotted

in Figure 10a,b. The contribution from van der Waals interactions and electrostatic interactions are marked out respectively. It is clear that both the {010} and {0̄10} faces of β -glycine interact with solvents (water and methanol–water 3:7 mixture) more strongly than those of the α form and all crystal surfaces interact more strongly with water than with methanol–water 3:7 mixture. The electrostatic interaction is the main contribution in all scenarios. With the presence of methanol, the van der Waals contribution increases slightly, but it does not make up for the loss in the electrostatic interactions. A close look at the β -glycine {010} face contacting with the methanol–water solution (Figure 11) shows that there are pockets on the surface which are preliminarily occupied by water molecules either due to their smaller size or stronger hydrogen-bonding ability, highlighted in those green circles. Methanol molecules near the glycine crystal surface usually have their hydroxyl group pointing toward the surface to form hydrogen bonds, highlighted in the orange circle, instead of forming van der Waals interactions with the exposed C–H bonds. We also see these contacts on the interface between the α -glycine {010} and {0̄10} faces and the methanol–water solution, which is not shown here for the simplicity. Opposite to the hypothesis proposed by Weissbuch, in our simulation study there is no strong binding of methanol molecules onto the {010} and {0̄10} faces of α -glycine and the {010} face of β -glycine (which all have the C–H bonds exposed) and the addition of methanol dilutes the

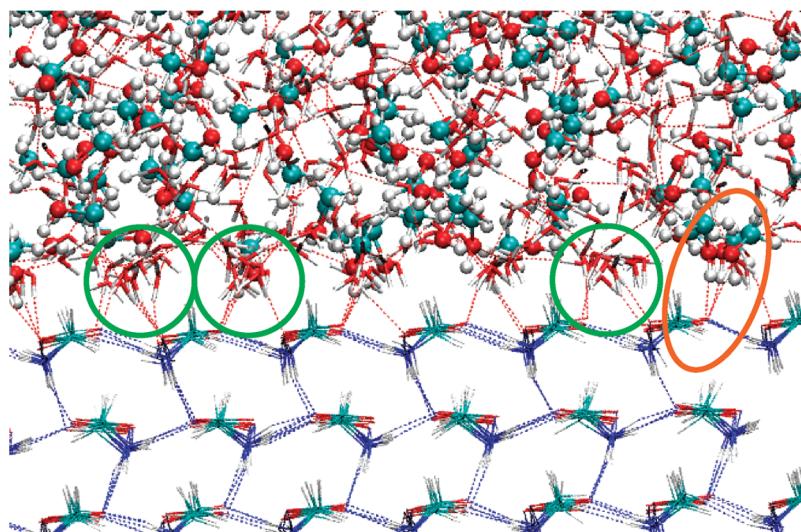


Figure 11. A snapshot of the β -glycine $\{010\}$ face interacting with methanol–water mixture, viewed along the c axis of β -glycine. Glycine and water molecules are in line representation, and methanol molecules are in ball and stick representation. Hydrogen bonds are in dashed lines. The green circles highlight the water molecules in a pocket on the $\{010\}$ face. The orange circle highlights the hydrogen bonds formed between methanol molecules and glycine.

crystal–solvent interactions on all interfaces. Weissbuch also used the fact that the addition of small amount of alcohol gives α crystals with larger $\{010\}$ and $\{\bar{0}\bar{1}0\}$ faces to support the methanol inhibition on these two faces which leads to their slower growth rates. It is worth to point out that the size of a face in the crystal morphology is decided by the relative growth rates of all faces. It is insufficient to draw conclusions about the methanol inhibition on the b -direction of glycine crystals based on the equilibrium crystal morphology, since the increase in the size of the $\{010\}$ and $\{\bar{0}\bar{1}0\}$ faces could also be the result of larger increases in the growth rates of other faces after the addition of methanol. Our simulation results at least show that methanol is energetically less favored than water on the $\{010\}$ and $\{\bar{0}\bar{1}0\}$ faces of both forms.

5. Conclusions

The mechanisms by which particular polymorphs are formed require a thorough understanding of the molecular level interactions in solution and on crystal–solution interfaces. In this study, we report this understanding based on molecular simulations. In particular, we studied two proposed mechanisms for polymorph selection of glycine, (1) the link hypothesis and (2) the asymmetric methanol inhibition on the $\{010\}$ and $\{\bar{0}\bar{1}0\}$ faces of α - and β -glycine. Glycine clusters formed in methanol–water 3:7 mixtures have a higher cyclic dimer composition than those formed in pure water with the same glycine concentration. Glycine dimers (both open-chain and cyclic dimers) are slightly stabilized with respect to the two fully solvated glycine monomers with the addition of methanol. Moreover, the glycine open-chain dimer is always more stable than the cyclic dimer regardless the presence of methanol. All these suggest that the selective crystallization of α - and β -glycine from aqueous and methanol–water solutions, respectively, is not due to the abundance of glycine aggregates in the solution phase with a similar structure to the crystallizing solid form. The presence of α -glycine $\{010\}$ face, which consists of a double layer of cyclic dimers, induces an increase in the cyclic dimer composition of the glycine clusters in its vicinity. However, the increase does not significantly offset the cyclic dimer fraction values in the methanol–water mixtures and does not render support to the link hypothesis. Additionally, different levels of methanol

inhibition on the $\{010\}$ and $\{\bar{0}\bar{1}0\}$ faces of α - and β -glycine cannot be the reason for the polymorph selection by solvent either. There is no strong binding of methanol molecules to the $\{010\}$ and $\{\bar{0}\bar{1}0\}$ faces of α -glycine and the $\{010\}$ face of β -glycine which all have C–H bonds exposed. Instead, the addition of methanol dilutes the solvent–crystal interactions on all interfaces. Moreover, the $\{010\}$ and $\{\bar{0}\bar{1}0\}$ faces of β -glycine always interact with solvents more strongly than those of α -glycine regardless of the presence of methanol. Therefore, the selective crystallization of β - and α -glycine with and without methanol does not follow either of the two mechanisms in the literature. More sophisticated computational methods to study nucleation and crystal growth such as those in development,^{29–33} as well as more advanced experimental techniques, are required to fully understand the polymorph selection of glycine.

Acknowledgment. This work was supported by the Singapore-MIT Alliance.

References and Notes

- Chen, J.; Trout, B. L. *J. Phys. Chem. B* **2008**, *112*, 7794.
- Davey, R. J.; Dent, G.; Mughal, R. K.; Parveen, S. *Cryst. Growth Des.* **2006**, *6*, 1788.
- Etter, M. C.; Parker, D. L.; Ruberu, S. R.; Panunto, T. W.; Britton, D. J. *Inclusion Phenom. Mol. Recognit. Chem.* **1990**, *8*, 395.
- Etter, M. C.; Urbanczyklijkowska, Z.; Jahn, D. A.; Frye, J. S. *J. Am. Chem. Soc.* **1986**, *108*, 5871.
- Erdemir, D.; Chattopadhyay, S.; Guo, L.; Ilavsky, J.; Amenitsch, H.; Segre, C. U.; Myerson, A. S. *Phys. Rev. Lett.* **2007**, *99*, 115702.
- Hamad, S.; Hughes, C. E.; Catlow, C. R. A.; Harris, K. D. M. *J. Phys. Chem. B* **2008**, *112*, 7280.
- Huang, J.; Stringfellow, T. C.; Yu, L. *J. Am. Chem. Soc.* **2008**, *130*, 13973.
- Weissbuch, I.; Torbeev, V. Y.; Leiserowitz, L.; Lahav, M. *Angew. Chem., Int. Ed.* **2005**, *44*, 3226.
- Bauer, J.; Spanton, S.; Henry, R.; Quick, J.; Dziki, W.; Porter, W.; Morris, J. *Pharm. Res.* **2001**, *18*, 859.
- Kempf, D. J.; Marsh, K. C.; Denissen, J. F.; McDonald, E.; Vasavanonda, S.; Flentge, C. A.; Green, B. E.; Fino, L.; Park, C. H.; Kong, X. P.; Widesburg, N. E.; Saldivar, A.; Ruiz, L.; Kati, W. M.; Sham, H. L.; Robins, T.; Stewart, K. D.; Hsu, A.; Plattner, J. J.; Leonard, J. M.; Norbeck, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2484.
- Law, D.; Krill, S. L.; Schmitt, E. A.; Fort, J. J.; Qiu, Y. H.; Wang, W. L.; Porter, W. R. *J. Pharm. Sci.* **2001**, *90*, 1015.
- Davey, R. J.; Allen, K.; Blagden, N.; Cross, W. I.; Lieberman, H. F.; Quayle, M. J.; Righini, S.; Seton, L.; Tiddy, G. J. T. *CrystEngComm* **2002**, *257*.

- (13) Hamad, S.; Moon, C.; Catlow, C. R. A.; Hulme, A. T.; Price, S. L. *J. Phys. Chem. B* **2006**, *110*, 3323.
- (14) Perlovich, G. L.; Hansen, L. K.; Bauer-Brandl, A. *J. Therm. Anal. Calorim.* **2001**, *66*, 699.
- (15) Albrecht, G.; Corey, R. B. *J. Am. Chem. Soc.* **1939**, *61*, 1087.
- (16) Iitaka, Y. *Acta Crystallogr.* **1960**, *13*, 35.
- (17) Chew, J. W.; Black, S. N.; Chow, P. S.; Tan, R. B. H.; Carpenter, K. J. *CrystEngComm* **2007**, *9*, 128.
- (18) Myerson, A. S.; Lo, P. Y. *J. Cryst. Growth* **1990**, *99*, 1048.
- (19) Erdemir, D.; Chattopadhyay, S.; Guo, L.; Ilavsky, J.; Amenitsch, H.; Segre, C. U.; Myerson, A. S. *Phys. Rev. Lett.* **2007**, *99*.
- (20) Chattopadhyay, S.; Erdemir, D.; Evans, J. M. B.; Ilavsky, J.; Amenitsch, H.; Segre, C. U.; Myerson, A. S. *Cryst. Growth Des.* **2005**, *5*, 523.
- (21) Hughes, C. E.; Hamad, S.; Harris, K. D. M.; Catlow, C. R. A.; Griffiths, P. C. *Faraday Discuss.* **2007**, *136*, 71.
- (22) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- (23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (24) Chen, J.; Trout, B. L. Computer-aided Solvent Selection for Improving the Morphology of Needle-like Crystals. Manuscript in preparation, 2010.
- (25) Price, S. L.; Hamad, S.; Torrisi, A.; Karamertzanis, P. G.; Leslie, M.; Catlow, C. R. A. *Mol. Simul.* **2006**, *32*, 985.
- (26) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
- (27) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
- (28) Mackerell, A. D. http://mackerell.umaryland.edu/Empirical_FF_Dev.html.
- (29) Beckham, G. T.; Peters, B.; Starbuck, C.; Variankaval, N.; Trout, B. L. *J. Am. Chem. Soc.* **2007**, *129*, 4714.
- (30) Peters, B.; Trout, B. L. *J. Chem. Phys.* **2006**, *125*.
- (31) Weinan, E.; Ren, W. Q.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2005**, *413*, 242.
- (32) Weinan, E.; Ren, W. Q.; Vanden-Eijnden, E. *J. Phys. Chem. B* **2005**, *109*, 6688.
- (33) Santiso, E. E.; Trout, B. L. Manuscript in preparation, 2010.

JP1039496