CrystEngComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/crystengcomm

On the Study of Crystal Growth via Interfacial Analysis and String Optimization

Adam Idu Jion[#] and Raj Rajagopalan[#,*]

[#]*Department of Chemical and Biomolecular Engineering*
National University of Singapore, Singapore 117585

[#,*]*Skolkovo Institute of Science and Technology*
Moscow Region, The Russian Federation 143025

Abstract

A systematic approach to the study of crystal growth is presented. This approach is based on finding the minimum energy path(s) for crystal growth units docking at the interface. Here, we demonstrate the string-optimizing technique of the Finite-Temperature String (FTS) method with the theoretical approach of Liu *et al*. to study the configuration free energies of molecules in higher dimensional space. Using the molecular dynamics simulation and interfacial analysis of morphologically important (010) and (011) faces of α-glycine crystals in aqueous solution as an example, we extend previous works and show that the FTS method can be used to calculate fraction of growth units and activation energies of flexible molecules present at the crystal-solution interface. We then discuss our work on the (010) surface of β-glycine grown in mixed-solvent, and show that the presence of methanol molecules can inhibit crystal growth.

1. Importance and Difficulties of Studying Crystal Growth in Solutions

The molecular mechanism of crystal growth in solution is an essential step towards predicting crystal morphology. Since the shape of a crystal influences its physical and chemical properties, the study of crystal growth is of great importance to the chemical and pharmaceutical industry [1]. However, the molecular mechanism induced by the solution environment (*i.e.* level of supersaturation, solvent used and solution purity) at the crystal interface is not fully understood [2]. Thus crystal growth in solutions is an active area of research.

Although there are many experimental techniques [1, 3, 4] to study crystal growth in solutions, most are unable to provide an atomic scale of resolution at the crystal-solution interface, and/or have problems investigating fast growing and high-energy surfaces [5]. Whilst purely atomistic computer simulations can in principle be applied for such purposes, and have indeed complemented experimental results, they are extremely time consuming and demand large computational resources. A way forward is to use a *multi-scale approach* [6-8] that combines computer simulation with thermodynamic analysis. However, such an approach requires the development of a novel technique if flexible or complicated molecules were to be investigated. In this paper, we present such a technique. This technique is computationally cheap, reliable and robust. It can be easily extended to arbitrarily complicated and flexible molecules, and enables the results of desktop computer simulations to be extrapolated towards the thermodynamic limit.

2. FTS Method and String Optimization

The Finite-Temperature String method (FTS) has been used in the past to study the transition mechanisms and transition rates between metastable states in systems with complex energy landscapes [9-11]. Examples of such phenomena include the conformation changes in molecules, chemical reactions and phase transitions. In such cases, the FTS method is used to the generate free energy surfaces *and* calculate the minimum energy paths (in the implementation by Vanden-Eijnden and Venturoli [10], for example, the free energy space is generated via sampling within the *Voronoi cell* associated with discretized points lying on a string). Because the FTS method samples the free energy surface in locally important regions, it is able to detect the *metadynamics* of the system, enabling the study of transition states in great detail. However, the FTS method has never been used in the study of crystal growth. This is because in crystal growth, reaction pathways are highly-dimensional, and are often degenerate with many barrier-crossing events, making the generation of free energy surface via local importance sampling computationally expensive. Nevertheless, we will show that if the free energy surface can be generated via alternate means, it can be combined with the string optimization component of the Finite-Temperature String method to *approximate* the *generalised* reaction coordinate for crystal growth. Therefore it can be used in cases where the free energy, and thus activation energy, reside on a *hypersurface* (*i.e.* where the free energy is a function of n-dimensions, and $n \geq 3$).

In the present paper, we use the string-optimizing component of the FTS method together with the theoretical approach of Liu *et al.* [7, 8] to create a novel technique that can be used to calculate fraction of growth units and activation energies of flexible molecules at the crystal-solution interface. As an example, we extend and formalize an earlier work on α-glycine crystals grown from solutions [6]. We then discuss our work on the (010) surface of β-glycine grown in mixed-solvent.

The premise of the string-optimization technique of the FTS method lies in finding the minimum energy path(s) (MEP) between two points '*a*' and '*b*' in a given energy landscape by repeatedly iterating an arbitrary curve called a *string* (Fig. 1). A string is a path, φ, that will connect an arbitrary initial configuration *'a'* to a final desired configuration *'b'*. Each iteration of the string method brings the string closer towards the minimum energy path, φ_{MEP} (Fig. 2):

$$\varphi_{MEP} = \lim_{n\to\infty} \varphi_n \tag{1}$$

where n is the number of iterations. Mathematically speaking, given an energy landscape $V(\mathbf{x})$, defined on a suitable vector space $\mathbf{x}$, the FTS method seeks to find a string φ_{MEP} between points *'a'* and '*b*' such that the normal force experienced by the string $\mathbf{f}^{\perp}(\varphi_{MEP})$ is zero everywhere:

$$\mathbf{f}^{\perp}(\varphi_{MEP}) = -\nabla V^{\perp} = 0 \tag{2}$$

This is akin to finding the global minimum of the functional

$$F[\varphi] = \int_{\sigma(a)}^{\sigma(b)} d\sigma \mathbf{f}^{\perp}(\varphi(\sigma))^{T} \mathbf{f}^{\perp}(\varphi(\sigma)) \tag{3}$$

over all paths connecting $\varphi(\sigma(a))$ to $\varphi(\sigma(b))$. Here σ is a parameterization variable that increases monotonically from one end of the string such that $\varphi(\sigma(a))$ denotes the start of the string and $\varphi(\sigma(b))$ denotes the end of the string.

3. Generation of the Free Energy Surface

In the case of crystal growth, the string-optimization component of the FTS method must be applied to an energy landscape that represents the crystal-solution interface. Although the FTS method is able generate free energy surfaces [10], it is computationally expensive and may not be feasible for the highly dimensional space found in crystallization. As such, we generate the energy landscape by using standard molecular dynamics simulations at the crystal-solution interface [6-8], and by using a statistical thermodynamics relationship [12] that scales the microscopic orientation distributions of molecules towards the thermodynamic limit:

$$G = -k_B T \ln[W] + const \tag{4}$$

Here G represents the Gibbs free energy, k_B is the Boltzmann constant, T is the absolute temperature, and W is the multiplicity of the microscopic state. The multiplicity of the microscopic state, W, is the number of possible permutations between the molecules of the system, and hence, measures the total number of system configurations (*i.e.* system *ensemble*).

In an earlier work, for example, Gnanasambandam and Rajagopalan [6] used glycine zwitterions as a model to study crystal growth at the (010) and (011) surfaces of α-glycine crystals in aqueous solution. They represented the orientation of glycine zwitterions by two dipole vectors, $C_\alpha \to C$ and $C_\alpha \to N$ (Fig. 3) with reference to the surface normal. This resulted in glycine zwitterions being specified by three coordinates – the angles made by the dipole vectors with respect to the surface normal, θ_{CC}, θ_{CN}, and the azimuth angle of the $C_\alpha \to C$ dipole vector, ϕ_{CC}. Hence equation (4) now becomes

$$G(\theta_{CC}, \theta_{CN}, \phi_{CC}) = -k_B T \ln[p(\theta_{CC}, \theta_{CN}, \phi_{CC})] + \text{ constant} \tag{5}$$

where $p(\theta_{CC}, \theta_{CN}, \phi_{CC})$ is the probability density of finding a glycine zwitterion of orientation $\theta_{CC}, \theta_{CN}, \phi_{CC}$. Thus the free energy surface of glycine zwitterions at the crystal-solution interface is a function of these three variables (*i.e.* $\theta_{CC}, \theta_{CN}, \phi_{CC}$).

Although using standard molecular dynamics to study activated processes such as crystallization may result in reduced sampling efficiency, this drawback can be mitigated by using the theoretical framework of Liu *et al.* (to be discussed in section 4). The morphology of crystals such as urea and glycine grown in solutions, for example, has been successfully predicted using standard molecular dynamics and a thermodynamic model involving interface structure analysis [6-8, 10, 13], even though they are activated processes.

4. Interface Structure Analysis and Glycine Polymorphism

In the original work by Liu *et al.* [7, 8], interface structure analysis was used for a 1-dimensional reaction coordinate system – that is, the orientation of a solute molecule was assumed to be described fully by a single angle θ (defined below). The fraction of molecules at the interface that eventually docked onto the crystal surface to become the bulk crystal was given by

$$\delta = \frac{1}{2}\int p \times \exp[(1-\Delta G)/2] \times sech[(1-\Delta G)/2]d\theta \qquad (6)$$

where θ is the angle between a single dipole vector (*i.e.* C→O of the urea molecule) and the surface normal, ΔG is the energy barrier (in units of $k_B T$) a molecule at the interface must overcome to be part of the bulk crystal, and p is the probability density of the molecule's dipole vector being at θ. Once the fraction of favourable growth units is known, the growth rates for the morphologically important faces can be determined. For a one-dimensional reaction coordinate, calculating ΔG is trivial, and can be obtained from a plot (Fig. 4). However, for flexible molecules like glycine that require a higher number of dimensional coordinates in order to be specified, the free energy landscape is a hypersurface (*i.e.* a plot involving four coordinates - $G, \theta_{CC}, \theta_{CN}, \phi_{CC}$), and calculating ΔG is non-trivial (Fig. 5).

In a previous work on "extended interface structure analysis" for α-glycine [6], ΔG used in equation (6) was not able to be properly calculated due to a highly-dimensional energy landscape. Instead the fraction of favourable growth units, δ, was estimated by Gnanasambandam and Rajagopalan via looking at slices along the θ_{CN}-coordinate and by manual-searching of coordinates for energy basins in which orientations were considered favourable (Fig. 5). This approach is not only subjective, but it is tedious and error-prone, as

it is difficult to map out all the possible paths a molecule could take, let alone estimate its energy barrier. As such, in this paper, we extend and formalize previous works with a more mathematical approach. We will make use of the FTS method to find minimum energy paths, calculate the height of their respective energy barriers, and even compute the overall activation energies for α-glycine crystal growth at the (010) and (011) surfaces. We then compare our computed results with previous work, and extend it further by examining the (010) surface of β-polymorph grown in mixed solvent.

The α-polymorph of glycine crystals is more stable at ambient temperatures than the β-polymorph [14], and is the predominant polymorph when grown in aqueous solutions [15]. However, when grown in alcoholic solutions, the β-polymorph predominates [16]. The mechanism behind this is a subject of intense debate [3, 5, 17, 18]. We will not attempt to resolve the debate in this paper. However, we will show using the method described above, that the presence of alcoholic molecules such as methanol inhibit the proper docking of glycine zwitterions onto the crystal surface. This inhibition is reflected by a slightly higher activation energy at the crystal surface.

4.1 Simulation Details

Crystal slabs for the α-polymorph were created by the *Materials Studio* software [19]. They were constructed in two stages. First, structures of the unit cell were obtained from the Cambridge Structural Database and cut along the (010) and (011) planes respectively. The orientations of glycine zwitterions in its α-polymorph unit cell when cleaved along the (010) and (011) faces are given in Table 1. The (010) cut was made such that the $-CH_2$ group of the glycine zwitterion was exposed to the bulk solution (Fig. 6). This was needed so as to conform to experimental evidence [3]. The cleaved unit cells were then tessellated to give large supercell slabs. Additional vacuum slabs were then placed on top of these supercells so as to accommodate the supersaturated bulk solutions ($\sigma = 0.60$). The total size and contents of the simulation boxes are given in Table 2.

In a similar manner, crystal slabs for the β-polymorph (Fig. 7) were obtained from the Cambridge Structural Database and cut along the (010) plane using *Materials Studio* software. Orientations of glycine zwitterions in its β-polymorph unit cell when cleaved along the (010) face are given in Table 3. Due to the reduced solubility of glycine in alcoholic

solutions [20-22], however, crystal supercells for the (010) surface were made much larger so as to provide better sampling statistics within the interface (Fig 8). Additional vacuum slabs were then placed on top of these supercells so as to accommodate the supersaturated bulk solutions (σ = 0.60). The total size and contents of the simulation boxes are given in Table 4.

Following the protocols and methodology of Gnanasambandam and Rajagopalan [6] so as to allow for detailed comparison, we use the Amber ff03 force-field [23] with Mulliken partial charges (see Table 5) and the SPC/E model [24] of water for our simulation as was done previously [6]. We do this for both systems of α- and β-polymorphs. These simulations were then conducted using the *GROMACS* software [25] for a period of 10 ns in the canonical (NVT) ensemble. Sampling was done in the last 7 ns within the interfacial layer. Periodic boundary conditions were applied in the three coordinate directions. Simulation boxes were coupled to a Berendsen thermostat at 298 K with a relaxation time of 1 ps. The time step for each simulation was 1 fs. The particle-mesh Ewald (PME) method was used for treating long-range electrostatic interactions, and a cut-off radius of 0.9 nm was chosen. For Lennard-Jones interactions, the cut-off was 1 nm.

5. Results and Discussion

The key step in employing the FTS method is to decompose the energy state space Ω into a collection of strings φ_i (Fig. 9). Each string starts from a different state point i in Ω but ends at the same, final state point f. Hence for systems in higher dimensions, equation (6) becomes

$$\delta_i = \frac{1}{2}\int_0^1 p \times \exp[(1-\Delta G_i)/2] \times sech[(1-\Delta G_i)/2]ds \tag{7}$$

and

$$\delta_{tot} = \sum_{i=1}^{m} \delta_i \tag{8}$$

In practice, molecular dynamics simulation is represented by particles in a series of discrete configurations. Therefore, m is the total number of points in discretized Ω, and G_i is the re-

parameterized curve of φ_i based on a normalised, monotonically increasing variable *s*. That is, G_i is a string in *s*-space the maps directly onto the string φ_i in Ω-space. Mathematically, this is represented by

$$\varphi_i = \{G_i(s) : s \in [0,1]\} \tag{9}$$

for some normalized reaction coordinate s so that

$$G_i : [0,1] \to \Omega \tag{10}$$

The symbol δ_{tot} denotes the total fraction of molecules at the interface that will eventually become a bulk crystal, and ΔG_i is the maximum energy barrier obtained in string i between points $[s,1]$.

As a first approximation, we define

$$\Delta G_i = \begin{cases} G_{\max\, peak,i}(s_{peak}) - G_i(s) & \text{for } s < s_{peak} \\ 0 & \text{for } s \geq s_{peak} \end{cases} \tag{11}$$

where $G_{\max\, peak,i}(s_{peak})$ is the maximum peak in free energy for string i at position s_{peak} and $G_i(s)$ is the free energy for string i at position s. Here we implicitly assume that a string will have only one significant peak (Fig. 10). We are aware of algorithms that can handle multiple peaks [26, 27] but these require additional, user-defined parameters and might be considered in future work if warranted. In the present context, our approximation is sufficient to yield good results (see Table 6).

Once the activation energy, $\Delta G_{i,}$, associated with each string is obtained, it can be reconciled with the physical reality of the molecules which generated the free energy in the first place. The overall activation energy will then be the weighted-average of all the energy barriers, where the weights are simply the density of the configuration multiplicity of the

molecules. For a system of N molecules, for example, where each molecule is specified by a vector $\mathbf{x}$ (for glycine, $\mathbf{x}=[\theta_{CC},\theta_{CN},\phi_{CC}]$) , and n_i molecules are in the state $\mathbf{x}_i$, the overall activation energy, $\Delta\overline{G}$, is given by

$$\Delta\overline{G} = \frac{\sum_{i}^{m} n_i \Delta G_i}{N} \tag{12}$$

Here, m is the number of grid points used to approximate the state space, Ω, and N is the total number of molecules, and is related to n_i by the following summation:

$$N = \sum_{i}^{m} n_i \tag{13}$$

5.1 Interfacial Analysis of α-glycine Revisited

We conducted molecular dynamics simulation at the (010) and (011) crystal-solution interfaces of the α-polymorph (see section 4.1 for details). We then examined the orientation distribution of the molecules at the interface and generated the free energy hypersurface according to equation (5). Plots of our results can be seen in Figure 5. Subsequently, we analysed our results using the FTS method and interfacial analysis of Liu *et al.* [7, 8].

The FTS method was carried out as described earlier. The state space Ω was discretized into 8000 points and the fraction of favourable growth units δ for the morphologically important (010) and (011) faces of α-glycine crystals grown in aqueous solutions were computed. We chose 8000 points because this choice corresponded to the number of bins (*i.e.* 20 x 20 x 20 for the three angles, $\theta_{CC},\theta_{CN},\phi_{CC}$, that defined the configuration orientation of glycine molecules) we used to create our dataset from molecular dynamics simulations. We implemented an algorithmic version of the FTS method by Eijnden and Venturolli [10] in the *MATLAB* package [28], and ran our code on a single node of a Quad-Core Intel Xeon processor over a period of 2.5 days. The results obtained were in good agreement with previous work (Table 6).

We then extend the previous work by calculating the activation energy for crystal growth at the (010) and (011) surfaces. Our results are tabulated in Table 7. We notice that the activation energy of the (011) surface is much higher (> 10 kJ mol^{-1}) than the (010) surface. As suggested by Gnanasambandam and Rajagopalan [6], this could be due to H-bonding interference by a Type 2 molecule at the bulk crystal surface (see Fig. 6), resulting in a misorientation of glycine molecules docking from the bulk solution onto the interface. Because H-bonds are very strong, a considerable amount of energy is required for the interfacial glycine molecules to re-orientate into the correct configuration (*i.e.* Type 4). This is reflected by the higher generalized activation energy of the (011) surface. However, it is important to note that the (011) surface of the α-polymorph is more 'corrugated' than the (010) surface, allowing for more glycine molecules to reside at the interface (Figs. 11-12). As a result, the growth rate of the (011) surface is still faster than the (010) surface (see reference [6] for detailed calculations) despite its higher activation energy. This is consistent with the experimental studies of α-glycine crystal growth in solutions [29].

5.2 β-glycine and Surface Inhibition by Methanol

It is speculated that if the polymorphism between α- and β-glycine could be attributed to a surface phenomenon, that surface would be the (010) plane [18]. As such, we use the technique and methodology described above to investigate the growth at the (010) surface of the β-polymorph in 50% v/v water-methanol solution.

Our calculations show that the β-polymorph has a slightly higher activation energy than the α-polymorph for a similar bulk supersaturation of $\sigma = 0.60$ (Table 8). This difference in activation energy can be attributed to surface inhibition by methanol molecules. That is, the presence of methanol molecules at the interface somehow makes it difficult for the proper docking of glycine molecules onto the crystal surface. Hence, more energy is required for glycine molecules to dock with the correct configuration. Our results show that the presence of methanol molecules at the interface increases the activation energy by about 4.2 kJ mol^{-1}. However, the mechanism behind this is unclear. It could be due to some form of stearic hindrance by methanol molecules, and/or the fact that the reduced solubility of glycine in methanol-water solution causes a higher degree of solvation of glycine monomers which in turn creates an additional energy barrier of desolvation for docking to take place.

5.3. Concluding Remarks

One of the advantages of using the string-optimization technique of the FTS method is that it can be easily extended to more complicated molecules such as proteins in n-dimensional space. As seen by the work of Gnanasambandam and Rajagopalan [6], even a small molecule like glycine requires three dimensions (*i.e.* $\theta_{CC}, \theta_{CN}, \phi_{CC}$) to be minimally specified. This resulted in a 4-dimensional energy landscape which required a considerable effort to analyse. Although they managed to get accurate results via careful slicing of the hyperspace into pseudo 3-dimensional energy landscapes (from which the fraction of favourable orientations could be estimated), such ad-hoc methods will not work for larger molecules. This is because larger and more flexible molecules require an even higher number of coordinates to be completely specified, and would result in an energy hyperspace that can never be visually plotted. As such, any attempts to study the energy hyperspace via slicing and visual inspection carries the risk of missing out on energy channels or energy funnelling between and within slices. With the string-optimization component of the FTS method, however, the energy hyperspace can be transformed into a vector space and be treated analytically. As a result, it is *scalable* and allows the study of molecules in highly-dimensional spaces. The only caveat in the current approach is that standard molecular dynamics is used to generate the free energy surface. As mentioned in section 3, this means that for highly activated processes, there will be reduced efficiency in the sampling of the state space, resulting in a possible overestimate in the activation energies calculated.

In this paper we have described the use of the FTS method [9-11] with the interfacial structure analysis of Liu *et al.* [7, 8]. We demonstrated the technique on the morphologically important (010) and (011) faces of α-glycine crystals grown in aqueous solution, and was able to calculate the fraction of favourable growth units at the interface as well as the generalised activation energies of the (010) and (011) surfaces. We then demonstrated the technique on the (010) surface of β-glycine crystal grown in 50% v/v water-methanol solution, and showed that the presence of methanol at the interface inhibits crystal growth by about 4.2 kJ mol^{-1}. We also highlighted that the FTS method can be used to study more complicated and flexible organic molecules such as proteins in higher dimensional space. Thus it is a useful tool for crystal engineering.

Acknowledgments: RR would like to thank the Singapore-MIT Alliance for financial support for parts of the work described here. Adam would like to thank Prof. Srini for useful discussions.

Appendix: Code Validation

In this section, we validate our implementation of the string-optimization technique of the FTS. As a trial function, the following analytic equation [30] is used:

$$V(x, y) = (1 - x^2 - y^2)^2 + \frac{y^2}{x^2 + y^2} \tag{14}$$

The three-dimensional and contour plots of V are illustrated in Figure 13. As observed, the energy landscape has two minima at A = (-1, 0) and B = (1, 0). The exact minimum energy pathway φ_{MEP} between points A and B is given by the arc of the unit circle: $x^2 + y^2 = 1$. That is, the relationship between the variables x and y connecting points A = (-1, 0) to B = (1, 0) resulting in the lowest energy barrier ΔV is given by the equation

$$y = \left(1 - x^2\right)^{1/2} \tag{15}$$

When the Finite-Temperature String method is used to approximate the analytical solution given in equation (15), the results are a good fit (Fig. 14).

References

1. A. S. Myerson, *Handbook of Industrial Crystallization*, Butterworth-Heinemann, 2002.
2. E. Vlieg, M. Deij, D. Kaminski, H. Meekes and W. van Enckevort, *Faraday Discussions*, 2007, **136**, 57-69.
3. D. Gidalevitz, R. Feidenhans'l, S. Matlis, D.-M. Smilgies, M. J. Christensen and L. Leiserowitz, *Angewandte Chemie International Edition in English*, 1997, **36**, 955-959.
4. C. E. Hughes, S. Hamad, K. D. M. Harris, C. R. A. Catlow and P. C. Griffiths, *Faraday Discussions*, 2007, **136**, 71-89.
5. J. Huang, T. Stringfellow and L. Yu, *Journal of the American Chemical Society*, 2008, **130**, 13973-13980.
6. S. Gnanasambandam and R. Rajagopalan, *CrystEngComm*, 2010, **12**, 1740-1749.
7. X. Y. Liu, E. S. Boek, W. J. Briels and P. Bennema, *The Journal of Chemical Physics*, 1995, **103**, 3747-3754.
8. X. Y. Liu, E. S. Boek, W. J. Briels and P. Bennema, *Nature*, 1995, **374**, 342-345.
9. W. Ren, E. Vanden-Eijnden, P. Maragakis and W. E, *The Journal of Chemical Physics*, 2005, **123**, 134109-134112.
10. E. Vanden-Eijnden and M. Venturoli, *The Journal of Chemical Physics*, 2009, **130**, 194103-194117.
11. W. E, W. Ren and E. Vanden-Eijnden, *The Journal of Physical Chemistry B*, 2005, **109**, 6688-6693.
12. S. R. Fowler and E. A. Guggenheim, *Statistical Thermodynamics*, Cambridge University, London, 1960.
13. S. Gnanasambandam, Z. Hu, J. Jiang and R. Rajagopalan, *The Journal of Physical Chemistry B*, 2008, **113**, 752-758.
14. G. Perlovich, L. Hansen and A. Bauer-Brandl, *Journal of Thermal Analysis and Calorimetry*, 2001, **66**, 699-715.
15. G. Albrecht and R. B. Corey, *Journal of the American Chemical Society*, 1939, **61**, 1087-1103.
16. Y. Iitaka, *Acta Cryst.*, 1960, **13**, 34-45.
17. S. Hamad, C. E. Hughes, C. R. A. Catlow and K. D. M. Harris, *The Journal of Physical Chemistry B*, 2008, **112**, 7280-7288.
18. I. Weissbuch, V. Y. Torbeev, L. Leiserowitz and M. Lahav, *Angewandte Chemie International Edition*, 2005, **44**, 3226-3229.
19. Accelrys, *Materials Studio*, (2001-2011) Accelrys Software Inc.
20. J. B. Dalton and C. L. A. Schmidt, *Journal of Biological Chemistry*, 1933, **103**, 549-578.
21. K. GEKKO, *Journal of Biochemistry*, 1981, **90**, 1633-1641.
22. K. Park, J. M. B. Evans and A. S. Myerson, *Crystal Growth & Design*, 2003, **3**, 991-995.
23. Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang and P. Kollman, *Journal of Computational Chemistry*, 2003, **24**, 1999-2012.
24. H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *The Journal of Physical Chemistry*, 1987, **91**, 6269-6271.
25. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *Journal of Chemical Theory and Computation*, 2008, **4**, 435-447.
26. A. V. Oppenheim, R. W. Schafer and J. R. Buck, *Discrete-time signal processing*, Second edn., Prentice Hall, 1999.
27. J. Pan and W. J. Tompkins, *Biomedical Engineering, IEEE Transactions on*, 1985, **BME-32**, 230-236.
28. *MATLAB - The Language of Technical Computing*, (1984-2010) The MathWorks, Inc.
29. L. Li and N. Rodríguez-Hornedo, *Journal of Crystal Growth*, 1992, **121**, 33-38.

30. W. E, W. Ren and E. Vanden-Eijnden, *The Journal of Chemical Physics*, 2007, **126**, 164103-164108.

List of Tables

Plane	Molecule Type	θ_{CC} / Degrees	Φ_{CC} /Degrees	θ_{CN} / Degrees
010	1	81	290	64
	2	81	110	64
	3	99	110	116
	4	99	290	116
011	1	143	25	71
	2	24	228	88
	3	37	204	109
	4	156	48	92

Table 1: Orientations of glycine zwitterions in its α-polymorph unit cell when cleaved along the (010) and (011) faces. θ_{CC} is the angle between the $C_\alpha \rightarrow C$ dipole vector and the surface normal, θ_{CN} is the angle between the $C_\alpha \rightarrow N$ dipole vector and the surface normal. Φ_{CC} is the azimuthal angle of the $C_\alpha \rightarrow C$ dipole vector

Plane	(010)	(011)
No. of glycine molecules in the bulk crystal	560	672
No. of glycine molecules in the solution	206	373
No. of water molecules	1931	3544
Dimensions of the simulation box (Å)	x: 40.84 y: 85.90 z: 35.54	x: 52.63 y: 102.08 z: 35.31

Table 2: Glycine/water mixtures for the crystal/bulk solution of the α-polymorph. Bulk glycine supersaturation is $\sigma = 0.60$ in pure water.

Plane	Molecule Type	θ_{CC} / Degrees	φ_{CC} /Degrees	θ_{CN} / Degrees
010	1	81	290	64
	2	81	110	64

Table 3: Orientations of glycine zwitterions in its β-polymorph unit cell when cleaved along the (010) face. θ_{CC} is the angle between the $C_\alpha \rightarrow C$ dipole vector and the surface normal, θ_{CN} is the angle between the $C_\alpha \rightarrow N$ dipole vector and the surface normal. Φ_{CC} is the azimuthal angle of the $C_\alpha \rightarrow C$ dipole vector

Plane	(010)
Number of glycine molecules in the bulk crystal	3888
Number of glycine molecules in solution	200
Number of methanol molecules	4106
Number of water molecules	9582
Dimensions of the simulation box (Å)	x: 91.70 y: 105.40 z: 96.90

Table 4: Glycine/water-methanol mixtures for the crystal/bulk solution of the β-polymorph. Bulk glycine supersaturation is σ = 0.60 in 50% v/v water-methanol solution.

Atom index	Atom	Charge
1	N	–0.127
2	H1	0.199
3	H2	0.218
4	H3	0.224
5	CA	0.007
6	HA1	0.064
7	HA2	0.061
8	C	0.483
9	OC1	–0.578
10	OC2	–0.552

Table 5: Mulliken partial charges for zwitterionic glycine. Adapted from [6].

Fraction of favourable orientation, δ	Values obtained by manual counting and estimation (See reference [6])	Values computed by Finite-Temperature String Method
δ_{010}	0.240	0.216
δ_{011}	0.130	0.108

Table 6: Comparison of favourable growth units obtained from FTS and visual inspection for the α-polymorph. δ_{010} is the fraction of favourable orientation in the (010) plane and δ_{011} is the fraction of favourable orientation in the (011) plane. The FTS is able to calculate these values to reasonable agreement with previous work [6]. Note that bulk glycine supersaturation is σ = 0.60 in pure water.

Plane	Activation energy / kJ mol^{-1}
010	17.3 ± 0.5
011	27.7 ± 0.5

Table 7: Generalized activation energies for the crystal faces of the α-polymorph. Bulk glycine supersaturation is σ = 0.60 in pure water. The (011) surface has a higher activation energy due to interference from a Type 2 molecule at the interface (see Table 1, Fig. 6 and ref. [6] for detailed discussion)

Plane	Fraction of favourable orientation, δ	Activation Energy / KJ mol^{-1}
(010)	0.13	21.5 ± 0.5

Table 8: Favourable growth-units and generalized activation energy for the (010) surface of the β-polymorph. Bulk glycine supersaturation is σ = 0.60 in 50% v/v water-methanol solution.

List of Figures

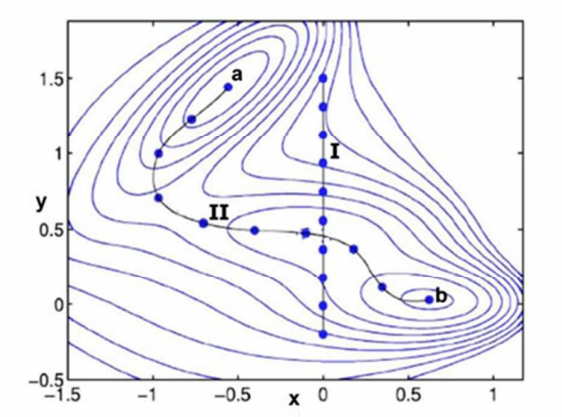


Figure 1: Hypothetical two-dimensional energy landscape with energy minima, 'a' and 'b'. String I is an initial arbitrary string that is repeatedly iterated in order to form String II. String II is the minimum energy pathway between points 'a' and 'b'.

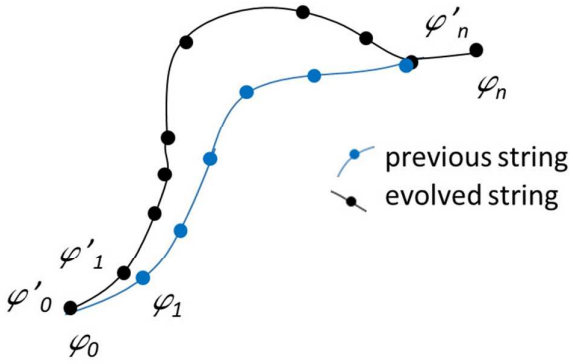


Figure 2: Each string will undergo repeated iteration and reparameterization. The final evolved string will represent the minimum energy path.

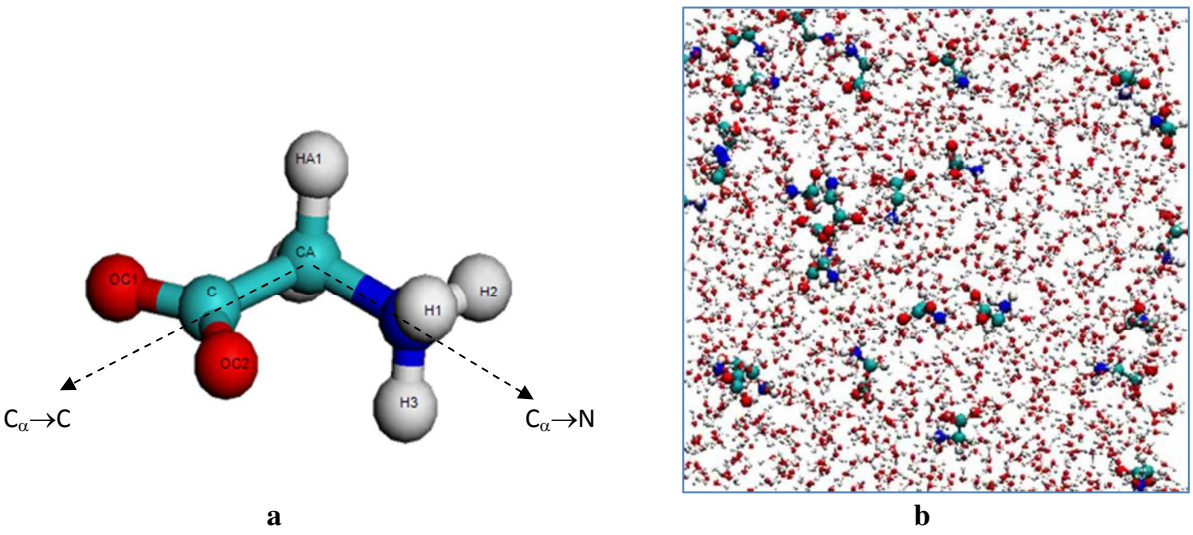


Figure 3: (a) Dipole vectors of glycine molecule. (b) Snapshot of a molecular dynamics simulation for glycine in water.

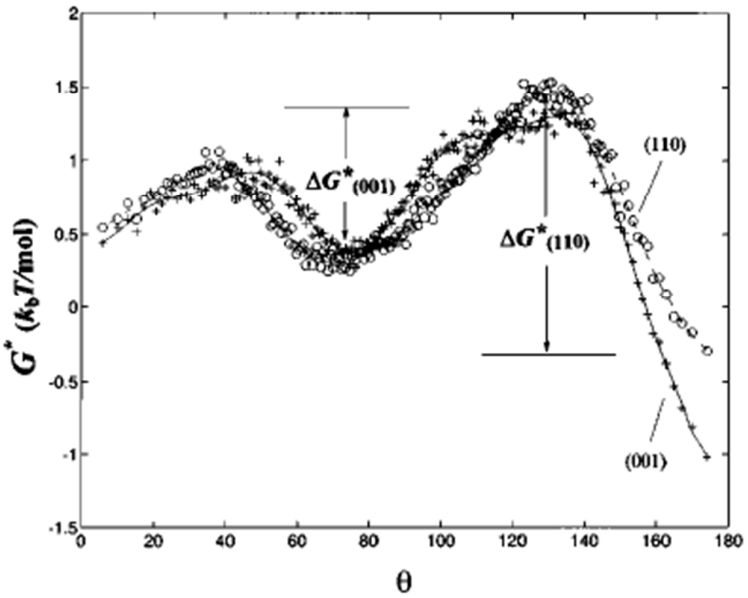


Figure 4: Gibbs free energy diagram for urea crystals grown in solution [7]. The reaction coordinate, θ, is the angle the dipole vector C→O of urea molecules make with the surface normal of the crystal slab. For such a simple, one-dimensional reaction coordinate, calculating the energy barrier (*i.e.* activation energy), ΔG, is quite trivial.

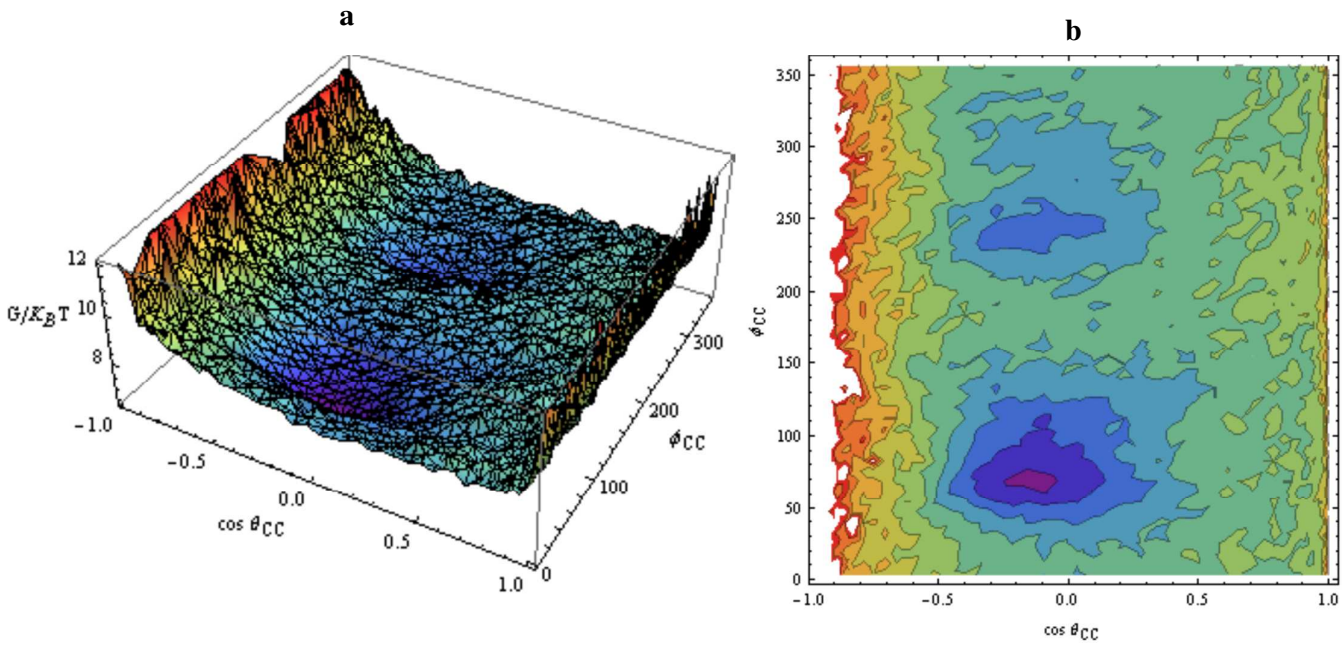


Figure 5: (a) Gibbs free energy distribution for the (010) surface as a function of θ_{CC} and ϕ_{CC} at $\theta_{CN} = 116^{o}$ (corresponding to a Type 3 growth unit) and its associated contour plot (b). In previous work, slices about 30° thick were examined in the θ_{CN} direction. The figures seen here are the juxtapositions of three slices centered around $\theta_{CN} = 116^{o}$. The fraction of molecules with favourable orientations is then counted based on visual inspection of the energy basin. See reference [6] for more details. By using the FTS method, however, we no longer need to cut arbitrary slices, and furthermore, we can extend our work to even higher dimensions.

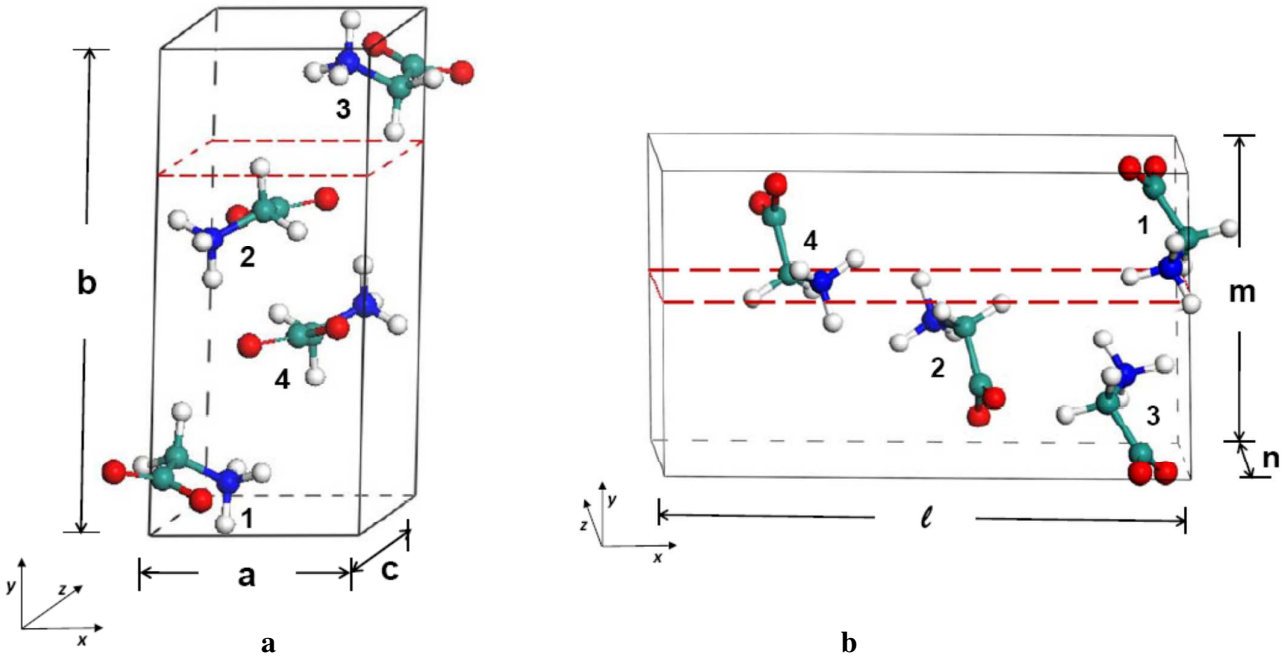


Figure 6: Arrangement of the glycine molecules of the α-polymorph relative to the (010) and (011) surfaces (a) (010) surface (b) (011) surface. Orientations of molecules above the broken red line are considered growth units. For the (010) surface the Type 3 molecule is the growth unit. For the (011) surface the Type 4 molecule is the growth unit. See Table 1 for the orientation configuration of the molecules and reference [6] for more details. Label: Red-Oxygen, Blue-Nitrogen, Cyan-Carbon and White-Hydrogen

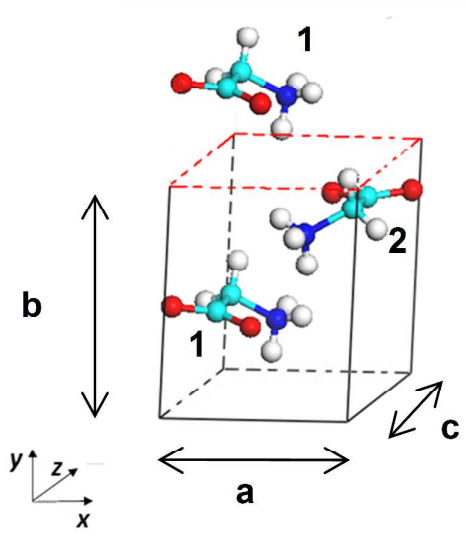


Figure 7: Arrangement of the glycine molecules in the β-polymorph. The unit cell is bound by the black and red lines. There are only two types of molecules in the unit cell – Type 1 and Type 2. They are anti-parallel to each other. The (010) surface is indicated by a broken red line. It exposes molecule of Type 2. Type 1 molecules are considered the growth-units for this crystal slab. See Table 3 for the orientation angles of Type 1 and Type 2 molecules. Label: Red-Oxygen, Blue-Nitrogen, Cyan-Carbon and White-Hydrogen

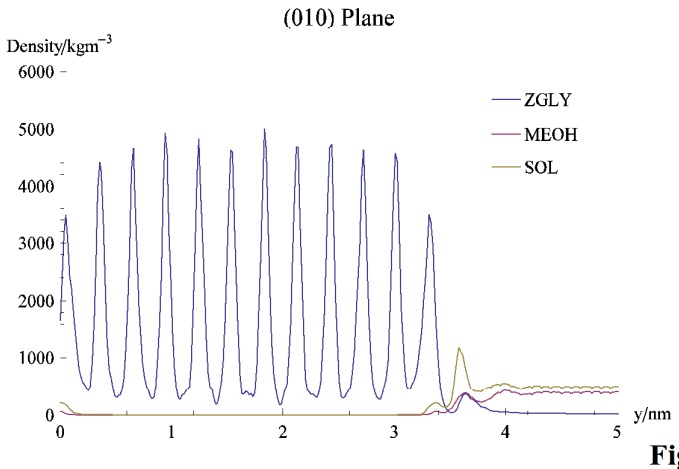


Figure 8: Density profile at the (010) surface of a β-glycine crystal slab in contact with its supersaturated solution of glycine zwitterions, methanol molecules and water molecules. The interface is from 3.5 nm to 4.0 nm. All sampling is done within the interface.

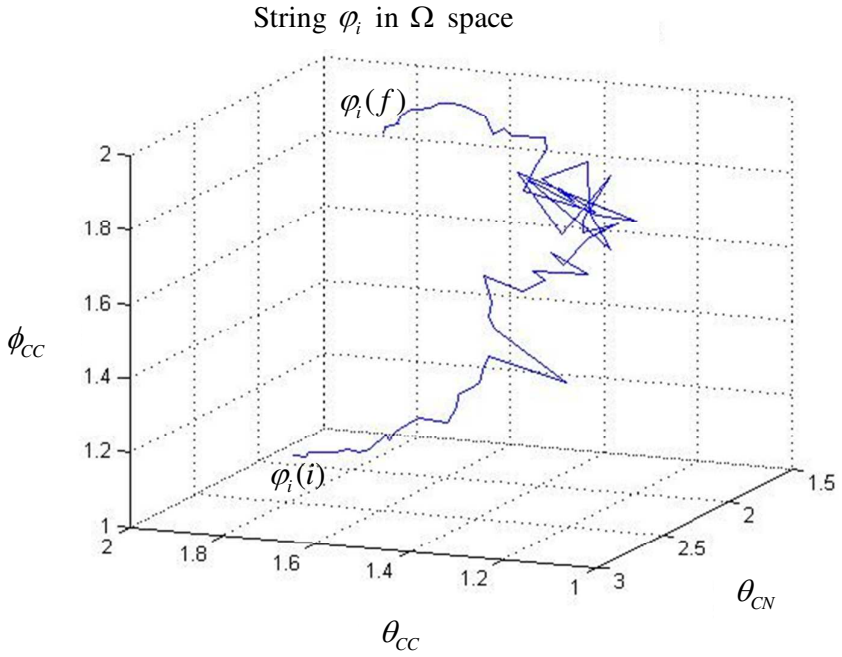


Figure 9: An example of a single string φ_i along a minimum energy path connecting arbitrary starting configuration point $\varphi_i(i)$ to final configuration point $\varphi_i(f)$. Here f is a point in Ω that corresponds to a Type 3 growth unit (See Figure 6 and Table 1). Ω is eventually decomposed into thousands of such strings.

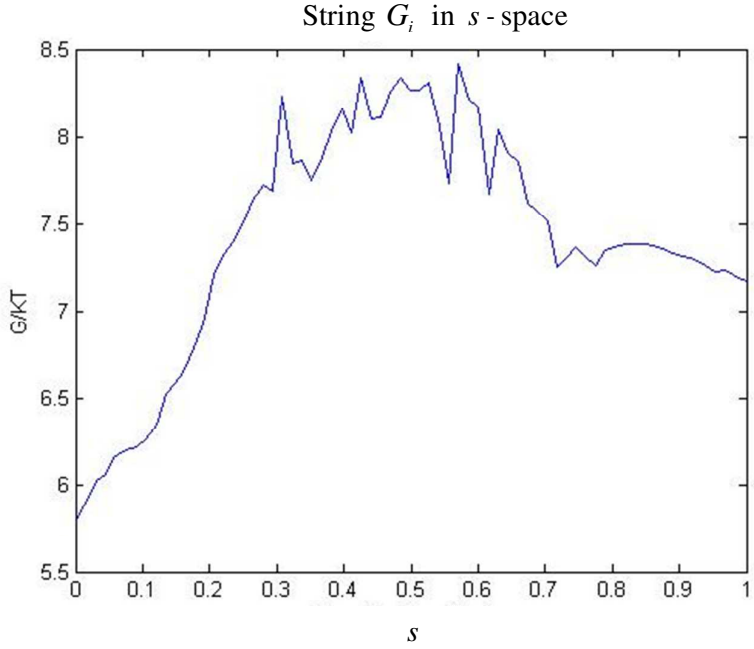


Figure 10: G_i is the parameterization of string φ_i along a normalized reaction coordinate s The maximum peak occurs at $s = 0.57$. Any molecular orientation to the right of this peak will face no energy barrier, and will be automatically considered a crystal growth unit.

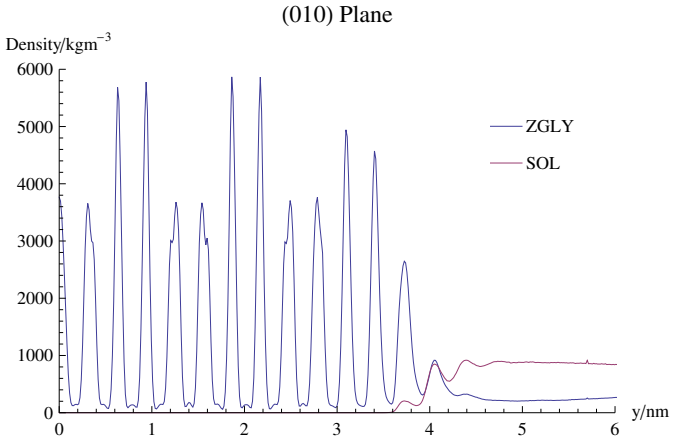


Figure 11: Density profile at the (010) surface of an α-glycine crystal slab in contact with its supersaturated solution of glycine zwitterions and water molecules. The interface is from 4.0 nm to 4.5 nm. All sampling is done within the interface.

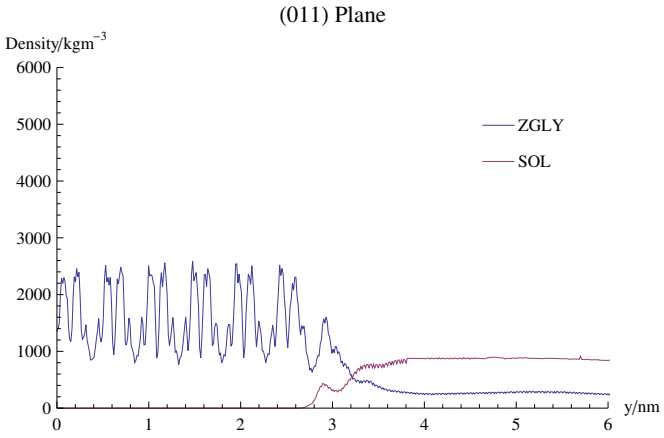
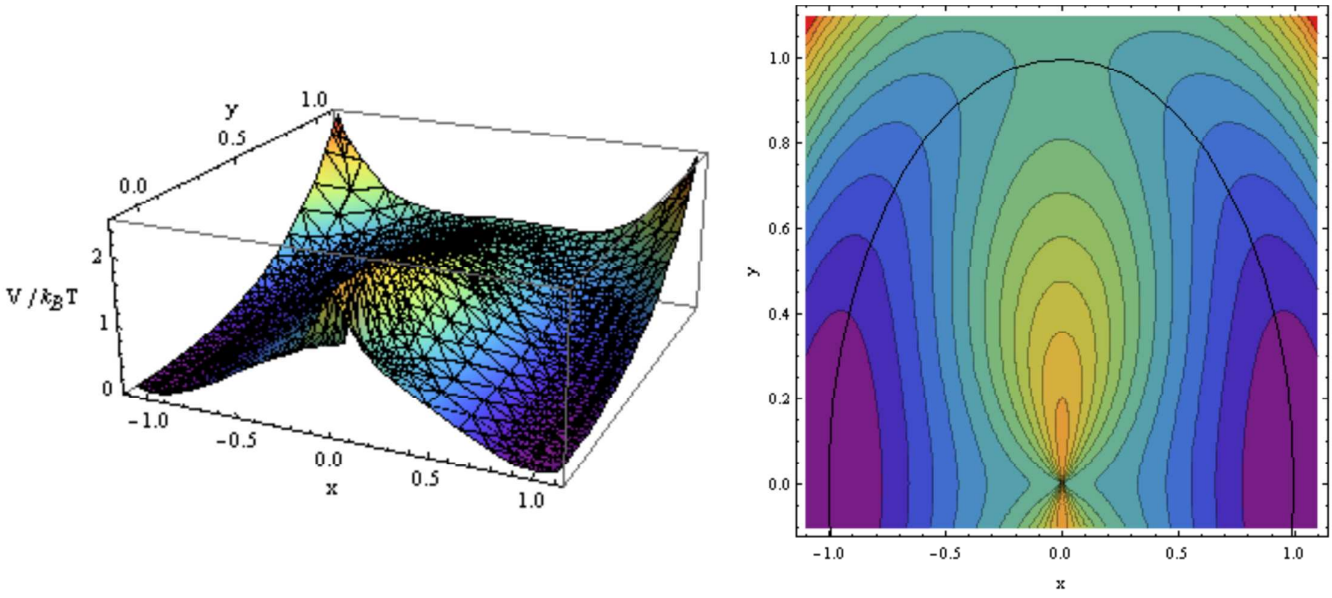


Figure 12: Density profile at the (011) surface of an α-glycine crystal slab in contact with its supersaturated solution of glycine zwitterions and water molecules. The interface is from 2.8 nm to 3.4 nm. All sampling is done within the interface. Note that the amount of glycine at the (011) interface is much higher than the (010) interface of Figure 11. This is probably due to the 'corrugated' surface of the (011) plane.



a **b**

Figure 13: (a) Energy landscape of the real-valued function $V(x, y)$. (b) Contour plot of $V(x, y)$. The black arc is the minimum energy pathway φ_{MEP} between the two energy minima A = (-1, 0) and B = (1, 0).

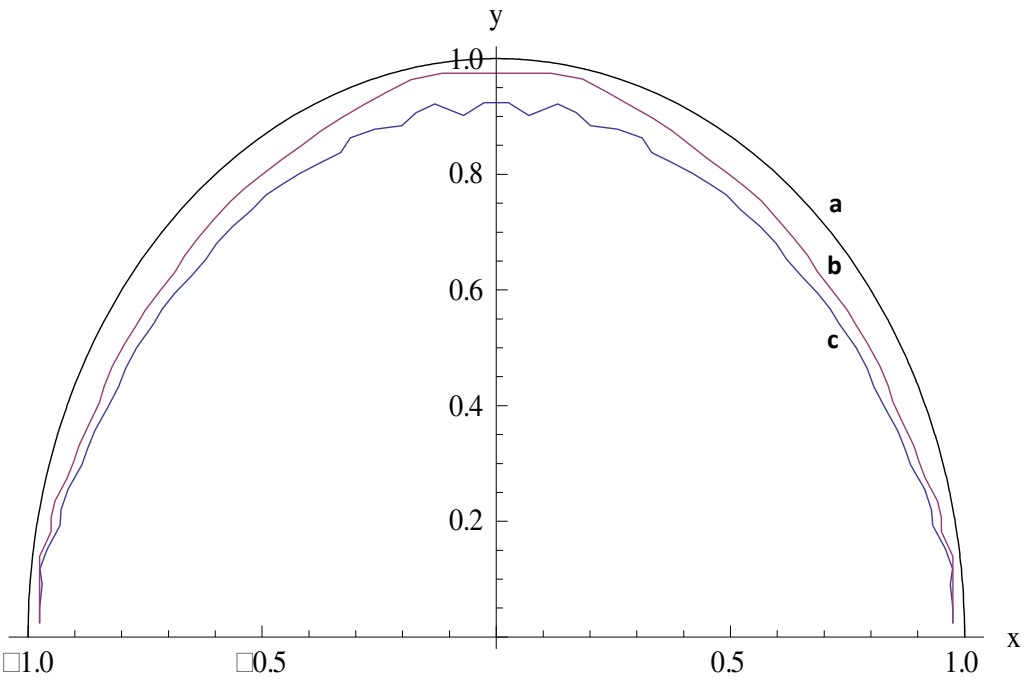
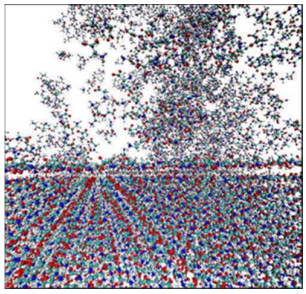


Figure 14: (a) Analytical solution of the minimum energy path in *V(x, y)* connecting points *A* = (-1, 0) to *B* = (1, 0) is given by the curve $y = (1 - x^2)^{1/2}$. (b) Numerical solution of the minimum energy path given by the FTS method with 10000 iterations, (c) Numerical solution of the minimum energy path given by the FTS method with 2000 iterations.

For Table of Contents Only



Mathematical ‘strings’ can be used with computer simulations and statistical mechanics to calculate fraction of growth units and activation energies of flexible molecules present at the crystal-solution interface.