This is a peer-reviewed, accepted author manuscript of the following article: Warzecha, M., Verma, L., Johnston, B., Palmer, J., Florence, A., & Vekilov, P. G. (Accepted/In press). Olanzapine crystal symmetry originates in preformed centrosymmetric solute dimers. *Nature Chemistry*. https://doi.org/10.1038/s41557-020-0542-0

Olanzapine crystal symmetry originates in preformed centrosymmetric solute dimers

Monika Warzecha,[1] Lakshmanji Verma,[2] Blair F. Johnston,[1,3] Jeremy C. Palmer,[2,*] Alastair J. Florence,[1,*] and Peter G. Vekilov[2,4,*]

[1]*EPSRC CMAC Future Manufacturing Research Hub, c/o Strathclyde Institute of Pharmacy and Biomedical Sciences, Technology and Innovation Centre, 99 George Street, Glasgow, G1 1RD, U.K.*

[2]*Department of Chemical and Biomolecular Engineering, University of Houston, 4726 Calhoun Rd., Houston, TX 77204-4004, USA.*

[3]*National Physical Laboratory, Hampton Road, Teddington, Middlesex TW11 0LW, U.K*

[4]*Department of Chemistry, University of Houston, 3585 Cullen Blvd., Houston, TX 77204-5003, USA*

[*]Correspondence and request for material should be addressed to P.G.V. (vekilov@uh.edu), A.J.F (alastair.florence@strath.ac.uk), and J.C.P. (jpalmer2@central.uh.edu)

Abstract

The symmetries of a crystal are notoriously uncorrelated to those of its constituent molecules; this symmetry breaking is typically thought to occur during crystallization. Here we demonstrate that one of the two symmetry elements of olanzapine crystals, an inversion center, emerges in solute dimers extant in solution prior to crystallization. We combine time-resolved *in situ* scanning probe microscopy to monitor the crystal growth processes with all-atom molecular dynamics simulations. We show that crystals grow non-classically, predominantly by incorporation of centrosymmetric dimers. The growth rate of crystal layers exhibits a quadratic dependence on the solute concentration, characteristic of the second-order kinetics of incorporation of dimers, which exist in equilibrium with a majority of monomers. We show that growth by dimers is preferred owing to overwhelming accumulation of adsorbed dimers, which is complemented by dimerization on the surface and expedites dimer incorporation into growth sites.

The symmetry of crystals is a central quality that determines their function and allure [1]. Given the uncorrelated symmetries of a crystal and its constituent molecules [2,3], it was assumed that the symmetry breaking that distinguishes a crystal from its mother phase occurs during nucleation [4-6]. This thesis emanates from the presumption that crystallization proceeds as sequential addition of solute monomers, which, however, is not a universal law, by far. Whether crystals grow by association of monomers [7] or the alignment of preformed units [8] was vigorously debated at the advent of crystal growth science. Bolstered by analytical models [9] and careful experimentation [10], the molecular viewpoint ascended to classical status. Recent observations in complex biological, geological, and synthetic environments have exposed cases of nonclassical crystallization employing crystalline, liquid, and amorphous precursors [11-13]. The agency of mesoscopic precursors only marginally modifies our understanding of the origin of crystal symmetry since the ordered structures of nanocrystals originate during their own nucleation [12]; amorphous, liquid, or partially ordered [14] particles attain symmetry upon conversion to crystals or association to a host crystal.

To explore the origin of symmetry of crystals growing by incorporation of microscopic building blocks we monitor the growth of olanzapine (OZPN, Figure 1a) dihydrate ethanoate crystals (Figure 1 b and c) [15]. Olanzapine is an antipsychotic drug used to treat schizophrenia and bipolar disorder [16]. More than 60 solvated and non-solvated forms of OZPN have been identified to date; in all of them, but one grown in a polymer dispersion [17], OZPN molecules are arranged in a centrosymmetric dimer motif SC_0, comprised of two conformational enantiomers (Figure 1c) [18]. The dominance of dimeric structures has spurred speculation that the dimers preform in the solution where they capture the majority of the solute [19,20]. The symmetry group of $2OZPN{\cdot}EtOH{\cdot}2H_2O$ crystals, $P2_1/c$, contains two independent symmetry elements, an inversion center and a two-fold screw axis. Crystal growth by incorporation of centrosymmetric dimers signifies that one of the two crystal symmetry elements, the inversion center, emerges prior to crystallization, leaving only the 2_1 axis to arise in the course of crystallization.

Results

Crystal growth by association of solute dimers

We monitor the growth of 2OZPN·EtOH·2H_2O crystals (Figure 1b) from a 1/1 (v/v) ethanol/water mixture (where the EtOH mole fraction is 0.24) using time-resolved *in situ* atomic force microcopy (AFM) [20]. The crystal habit is dominated by the basal {002} faces that grow by incorporation of solute into steps generated by screw dislocations (Figure 1d). The OZPN steps grew at steady rates over extended periods and the evolutions of the step displacement (Figure 1e inset) afforded determinations of the step velocity v as the slope of the displacement – time correlations. Classical crystal growth theory assumes that steps grow by incorporation of solute monomers [9]. The implied monomolecular reaction should give rise to a linear correlation between v and the solute concentration C, $v = \beta\Omega(C - C_e)$, where subtracting the solubility C_e accounts for the reversibility of molecular attachment, Ω is the molecular volume in the crystal, and β is an effective kinetic coefficient, which includes the kinetic parameters for the selected growth mechanism, direct incorporation or *via* adsorption on the terraces [21,22]. Linear $v(C)$ correlations have been recorded for numerous solution grown crystals [23]. Surprisingly, OZPN exhibits a superlinear $v(C)$ dependence, which extends to concentrations more than twice higher than the solubility C_e (Figure 1e).

We now eliminate four potential scenarios of apparent growth acceleration at high supersaturation. A first straightforward explanation would be an inaccurate value of the solubility C_e, interpolated for the temperature of the AFM measurements, 21°C, using the concentrations of equilibrium between 2OZPN·EtOH·2H_2O crystals and solution in 1/1 EtOH/H_2O at five disparate temperatures (Figure 2a) [24]. For an alternative determination, we continuously monitored a step site in a solution with $C = C_e$ at 21°C and observed that the attachment of solute from the solution is balanced by the detachment of solute from the step (Figure 2b). These dynamics indicate equilibrium between the solution and the step and ratify the value of C_e. A second hypothesis is

that mesoscopic solute-rich clusters that exist in OZPN water-ethanol mixtures [24]—and facilitate certain polymorph transformations [20]—may contribute to faster growth at higher supersaturation by delivering high OZPN concentrations to the steps (Figure 2c) [11,13,14,25]. Applying increased force to the AFM tip reveals that clusters deposited on the surface (Figure 2d) retain their distinct structure and size (Figure 2e) over several hours. In a complementary experiment, we removed the clusters from the solution by filtration and the step velocities remained unchanged. These two results suggest that the cluster contribution to growth is minimal. The third tested hypothetical scenario assumes that the density of kinks generated by thermal fluctuations of steps at equilibrium [9] is lower than the thermodynamic limit of about 0.3 [26,27] and increases at higher supersaturations. Since the kinks are the only sites where solute molecules can incorporate into crystals, increasing kink density driven by higher OZPN concentration may lead to superlinear $v(C)$ [28]. High-resolution AFM images of OZPN crystals surfaces and steps at equilibrium with the solution (Figure 2 f and g) reveal that the kink density is high and suggest that increasing supersaturation would not induce higher kink density and accelerated growth. Lastly, step pinning by impurities adsorbed on the terraces between steps may manifest as superlinear $v(C)$ [23]. Previous observations and theory, however, demonstrate that steps squeezing between adsorbed impurities transition from no growth to uninhibited growth within a narrow supersaturation range, where the 2D critical diameter is commensurate with the average impurity spacing [29,30]. This behavior is in contrast to the recorded $v(C)$ with OZPN (Figure 1e), in which the superlinear dependence extends from $(C - C_e)$ = 0.21 to 2.5 mM.

We propose that the superlinear $v(C)$ characterizes crystal growth occurring predominantly by incorporation of dimers, which form in a pool of OZPN monomers (Figure 3). Elevated OZPN concentrations shift the dimerization equilibrium towards dimers and nonlinearly enhance the dimer concentration. A kinetic scheme presented in Methods reveals that if crystals grow by incorporation of solute dimers that exist in equilibrium with monomers the step velocity would depend on the analytical concentration of the solute C as $v = \beta_D K_D \Omega_D (C^2 - C_e^2)$, where the

subscript D denotes dimer and K_D is the dimerization equilibrium constant. The $\upsilon(C)$ data are consistent with this functional relation (Figure 1 e and f).

Formation of dimers in the growth solution

Supporting the mechanism of growth by incorporation of dimers coexisting with solute monomers, Raman spectroscopy reveals OZPN dimers in equilibrium with OZPN monomers in a EtOH/H_2O solution and in chloroform ($CHCl_3$). We compare OZPN spectra from diluted solutions (0.005 M in 1/1 EtOH/H_2O and 0.009 M in $CHCl_3$) to spectra at relatively high concentrations (0.043 M 1/1 in EtOH/H_2O and 0.206 M in $CHCl_3$) and to spectra of the respective solid forms, 2OZPN·EtOH·2H_2O and 2OZPN·$CHCl_3$, in both of which OZPN is dimerized as SC_0 (Figure 4a). To understand the origins of the observed peaks, we model spectra for an isolated OZPN monomer and an isolated OZPN SC_0 dimer (Figure 4a). The model spectra for the OZPN monomer show several Raman peaks between 1200 and 1500 cm^{-1} (Figure 4a), whereas the model spectra for the dimer exhibit peaks around 1000 cm^{-1} and in the range 1500-1600 cm^{-1}. Analysis in SI relates these peaks to stretching and bending vibrations of the OZPN monomer and dimer.

Dimer peaks in both wave-number ranges are prominent in the solid-state spectra of both solvates, in accordance with the dimer structures of the solid forms, and at elevated concentrations in both solvents (Figure 4a). The low-concentration solution spectra reveal strong monomer peaks and lack of dimer peaks. The accumulation of dimers upon concentrating OZPN solution in $CHCl_3$ from 0.023 to 0.206 M is observed by NMR (Extended Data Figure 2 b-d); the upper limit on the OZPN concentration in 1/1 EtOH/H_2O prevented similar NMR characterization (Extended Data Figure 2a). Raman spectra at increasing intermediate concentrations in both solvents display gradual intensity decrease of the monomer peaks and growth of the dimer peaks (Extended Data Figure 1). The concentration responses of the monomer and dimer Raman peak intensities enable estimation of the OZPN dimerization constant K_D in the two solvents: 2.7±0.1 M^{-1} in EtOH/H_2O and 0.13±0.02 M^{-1} in $CHCl_3$. Mass balance calculations inform that the dimer

concentration C_D in the growth solution in EtOH/H_2O is 0.01 mM at the solubility 2.05 mM; C_D increases to 0.05 mM at the highest tested total OZPN concentration in the growth studies, C = 4.45 mM (Figure 1 e and f). Analyses in Methods imply that the superlinear increase of C_D dictates the quadratic $v(C)$ correlation (Figure 1 e and f, Extended Data Figure 4 and Supplementary Table 2). Notably, the low fraction of OZPN captured in solute dimers is a crucial prerequisite for quadratic $v(C)$, which represents the signature manifestation of crystal growth by association of dimers.

All-atom classical molecular dynamics evaluation of the potentials of mean force F between two OZPN monomers in EtOH/H_2O and $CHCl_3$ reveal deep minima at center-of-mass separations of 0.43 nm (Figure 4 b and c). In both solvents, the configurations of the two monomers occupying the minima fluctuate around the structure of the SC_o dimer observed in OZPN crystal forms (Figure 4 b and c, insets). At larger separations in both solvents, the SC_o dimers fall apart, however, F exhibits a shallow secondary minimum that corresponds to loose dimers with variable configuration (Figure 4 b and c). We evaluate the free energy of OZPN dimerization in the two solvents from the potential value at the respective deep minima: –8.1 kJ mol^{-1} in EtOH/H_2O and +2.5 kJ mol^{-1} in $CHCl_3$. The relative destabilization of the dimer in $CHCl_3$ conforms to more favorable interaction between OZPN and hydrophobic $CHCl_3$ than between two OZPN monomers. The dimerization constants K_D were evaluated as the ratio of integrals of F over the closest range minimum and the unbound state (Methods) and are 2.6±0.8 M^{-1} in EtOH/H_2O and 0.2±0.2 M^{-1} in $CHCl_3$. The similarity of the computed K_Ds to the values determined from the concentration responses of the monomer and dimer Raman peaks presents an independent validation of the simulations.

Factors favoring growth by dimers

The selection of crystal growth by incorporation of dimers from a solution, in which the dimers are a minority component, implies that dimer incorporation into kinks is significantly

faster than monomer association. To understand the thermodynamic and kinetic mechanisms that motivate faster growth by dimers, we note that the lower concentration and slower diffusion of the dimers overcompensate for their larger molecular volume; in consequence, the combination of these three factors should contribute to slower growth by dimers. Further analysis in the Supplementary Text indicates that the translational and rotational entropies of the dimer in the solution are greater than those of the monomer and, accordingly, contribute to higher dimer transition-state free energy and slower association to the kinks. To clarify why, despite the unfavorable activation entropy contribution, dimer ingress is faster, we consider whether the crystal building blocks reach the kinks directly from the solution or after adsorption on the terraces between steps and diffusion towards the steps (pathways *i* and *ii* in Figure 5a). Two observations indicate that the surface diffusion pathway is preferred.

First, steps with broad front terraces and narrow rear terraces (such as the step highlighted in green in the central and right parts of Figure 5b) grow faster than the step immediately behind them (e.g., the pink step), defined by a narrow front terrace and a broad rear terrace (Figure 5b and c). In the left parts of Figures 5 b and c, the blue step escaping from the green step provides an analogous example. The accelerated step velocity in the presence of closely positioned (≤ 150 nm) rear steps suggests that the steps mostly feed from the front. The observed directional asymmetry of solute supply, the Ehrlich-Schwoebel effect [31], is impossible if solute reaches the steps directly from the solution, but often observed during step growth *via* adsorption on the terraces and surface diffusion [32]. Second, the velocity of steps separated by l < 250 nm is significantly slower than v of isolated steps (Figure 5d and Extended Data Figure 5, see analyses in Methods). During direct incorporation into steps, the solute supply fields are three-dimensional, extending into the bulk solution, and the competition for supply only marginally impacts the velocity of closely spaced steps [21]. By contrast, closely spaced steps that feed from the surface compete for supply from a 2D space of the terrace between them and exhibit a strong $v(l)$ dependence [22,33].

We find that the dimers reach the kinks faster not because the activation barriers that they encounter along the surface diffusion pathway are lower than for the monomers. The effective transition-state enthalpy for molecular incorporation for the surface diffusion pathway $\Delta H^{\neq} = \Delta H^{\neq}_{ads} + \Delta H^{\neq}_{SD} - \Delta H^{\neq}_{des} + \Delta H^{\neq}_{kink}$, where the addends are the barriers for adsorption, surface diffusion, desorption, and incorporation into kinks, respectively (Figure5e). Arguably, $\Delta H^{\neq}_{ads}$ and $\Delta H^{\neq}_{kink}$ are the dominant contributors to higher $\Delta H^{\neq}$ [33]. These barriers have been ascribed to the consecutive stripping, along the incorporation pathway, of solvent molecules bound to the solute [21,34]. The diversity of polar structural groups in OZPN suggests that water molecules bind to it with disparate energies. The energies of binding of nine water molecules to the OZPN monomer, computed using an optimized intermolecular force field for hydrogen-bonded organic molecular crystals [20], suggest that water molecules removed upon dimer formation bind to OZPN with energy higher than –20 kJ mol^{-1} (ref. [20], Figure 5f). Remarkably, the five strongly bound waters, with binding energies lower than –28 kJ mol^{-1}, localize on the outward surface of the OZPN molecule (Figure 5f) [20]. They persist in the dimer so that the four of them that shed upon incorporation (leaving one H_2O molecule per OZPN in the lattice) would contribute equally to the kinetic barriers for dimer and monomer adsorption and association to the kinks (Figure 5f).

We attribute the fast growth by dimers to their stronger adsorption on the terraces between steps and supplementary dimerization of monomers occurring in the adsorbed state. All-atom molecular dynamics simulations reveal that the adsorption enthalpy of the monomers $\Delta H^{o}_{ads,M}$ = –4 ± 2 kJ mol^{-1}, computed as the difference between their enthalpies at the (002) face of 2OZPN·EtOH·2H_2O crystals and in the solution bulk, is higher than $\Delta H^{o}_{ads,D}$ = –10 ± 4 kJ mol^{-1} of the dimers (Figure 6). The stronger adsorption of dimers, suggested by the MD simulations, boosts their surface concentration. The crystal surface acts as a reservoir that concentrates the dimers and promotes their incorporation into the kinks. The stronger dimer adsorption is complemented by additional dimerization occurring on the crystal surface. The thermodynamic

cycle that includes dimerization in the solution bulk and on the surface and adsorption of monomers and dimers (Extended Data Figure 7a) suggests that the enthalpy of dimerization on the surface $\Delta H^o_{D,surf} = -2\Delta H^o_{ads,M} + \Delta H^o_{D,bulk} + \Delta H^o_{ads,D}$, where $\Delta H^o_{D,bulk}$= −11 kJ mol^{-1} is the bulk dimerization enthalpy, and $\Delta H^o_{D,surf}$ = −13 kJ mol^{-1}. Surface dimerization further boosts the dimer concentration at the terraces and the rate of dimer incorporation into steps. From the point of view of the above expression for $\Delta H^{\neq}$ and the relation $\Delta H^o_{ads} = \Delta H^{\neq}_{ads} - \Delta H^{\neq}_{des}$ (Figure 5e), lower $\Delta H^o_{ads,D}$ reduces the dimers' effective barrier for incorporation into kinks, $\Delta H^{\neq}$, in which ΔH^o_{ads} is a major constituent, and contributes to significantly faster kinetics of incorporation.

Discussion

We establish a deviation from the classical mechanisms of crystallization, which assume that crystals grow by sequential association of single solute molecules. We show that a preformed centrosymmetric solute dimer is the preferred growth unit for olanzapine crystals even though dimers comprise a minority of the solute population in the solution bulk. We identify the reason for faster growth by dimers as their stronger adsorption on the crystal surface supplemented by additional dimerization on the surface, which creates a reservoir for ready dimer incorporation into steps. We put forth a quadratic correlation between the step velocity and the solute concentration as a general criterion to identify the self-assembled solute molecular form—dimer or higher oligomer—that associates to kinks, constituting growth. These findings dictate critical modification in crystallization process models that assume that the step velocity scales linearly with solute concentration [35].

Growth of OZPN crystals by centrosymmetric dimers presents an example of a crystal symmetry element emerging in the solution and on the crystal surface prior to crystallization, in contrast to the accepted paradigm that symmetry breaking occurs exclusively during crystallization. An intriguing open question is whether dissolution obeys or violates the principle of microscopic reversibility, which, for OZPN, requires dissolution by detachment of dimers.

Given the presence of dimers in the growth solution and the dimeric structure of the 2OZPN·EtOH·2H_2O crystals, it appears reasonable to expect that the solute dimers are selected during crystal nucleation, i.e., OZPN crystal nucleation by assembly of preexisting dimers is faster than nucleation of the same crystal structure by assembly of monomers. This reasoning may be flawed, however, if nucleation follows the classical pathway of sequential association of monomers or dimers. Indeed, classical nucleation theory identifies two powerful governing parameters of the nucleation rate: the surface free energy of the interface between the incipient nucleus and the host solution and the supersaturation. The surface tension cannot bias the selection between the two nucleation scenarios because it is prescribed by the structures of the respective nuclei, which are identical. Straightforward mass balance calculations, based on the value of the dimerization constant K_D = 2.7 M^{-1}, indicate that the concentration of dimers would exceed that of monomers if the total OZPN concentration is greater than 1.1 M. Crystals nucleate at concentrations several-fold greater than the solubility C_e = 2.05 mM, i.e., of order 10 mM, at which the monomers are in multifold excess. Thus, supersaturation favors classical nucleation by monomer assembly.

On the other hand, OZPN crystals do not necessarily follow classical nucleation pathways. Careful AFM observation revealed that the nucleation of OZPN hydrate, a distinct crystal form, is hosted and facilitated by mesoscopic solute-rich clusters [20]. Such clusters arise in a line of OZPN solutions in EtOH/H_2O, including 1/1 EtOH/H_2O [24]; they associate to the surface of large 2OZPN·EtOH·2H_2O crystals, where, notably, they do not contribute to crystal growth (Figure 2 d and e). The unique cluster properties (their size is steady in time and independent of the OZPN concentration; the OZPN fraction captured in the clusters is dictated by the solution thermodynamics) distinguish them from other condensed phases; these behaviors have been ascribed to the formation, diffusion, and decay of transient OZPN dimers that instigate cluster formation and accumulate in the clusters (Figure 2c, callout) [24]. The elevated concentration in the clusters of transient dimers, which are akin to the crystallographic centrosymmetric dimers, may

facilitate their restructuring and assembly into a crystal nucleus comprised of dimers. This scenario presents a feasible pathway whereby the creation of a symmetry element in the dimers precedes crystal nucleation and illuminates strategies for polymorph control that promote or eliminate preformed crystal units [36,37].

References

1 Vainshtein, B. K. *Fundamentals of Crystals: Fundamentals of Crystals. Symmetry, and Methods of Structural Crystallography*. Second edn, (Springer, 1994).

2 Barker, T. V. Molecular and Crystal Symmetry. *Nature* **111**, 632, (1923).

3 Gavezzotti, A. Are crystal strutures predictable? *Accounts of chemical research* **27**, 309-314, (1994).

4 Landau, L. D. & Lifschitz, L. *Statistical Physics, Part 1*. 3rd edn, Vol. 5 (Pergamon Press, 1977).

5 Braga, D., Grepioni, F. & Desiraju, G. R. Crystal Engineering and Organometallic Architecture. *Chemical Reviews* **98**, 1375-1406, (1998).

6 Bernstein, J. *Polymorphism in Molecular Crystals*. 424 (Oxford University Press, 2007).

7 Stranski, I. N. Zur Theorie des Kristallwachstums. *Z. Phys. Chem.* **136**, 259-278, (1928).

8 Blüh, O. Einige bei der Untersuchung von Kolloiden im Wechselfeld auftretende Erscheinungen. *Kolloid-Zeitschrift* **37**, 267-270, (1925).

9 Burton, W. K., Cabrera, N. &. Frank, F.C. The growth of crystals and equilibrium structure of their surfaces. *Phil. Trans. Roy. Soc. London Ser. A* **243**, 299- 360, (1951).

10 Bennema, P. Analysis of crystal growth models for slightly supersaturated solutions. *J. Crystal Growth* **1**, 278-286, (1967).

11 Gliko, O. *et al.* A metastable prerequisite for the growth of lumazine synthase crystals. *J. Amer. Chem. Soc.* **127**, 3433-3438, (2005).

12 Li, D. *et al.* Direction-Specific Interactions Control Crystal Growth by Oriented Attachment. *Science* **336**, 1014-1018, (2012).

13 De Yoreo, J. J. *et al.* Crystallization by particle attachment in synthetic, biogenic, and geologic environments. *Science* **349**, aaa6760, (2015).

14 Lupulescu, A. I. & Rimer, J. D. In Situ Imaging of Silicalite-1 Surface Growth Reveals the Mechanism of Crystallization. *Science* **344**, 729-732, (2014).

15 Wawrzycka-Gorczyca, I., Borowski, P., Osypiuk-Tomasik, J., Mazur, L. & Koziol, A. Crystal structure of olanzapine and its solvates. Part 3. Two and three-component solvates with water, ethanol, butan-2-ol and dichloromethane. *Journal of Molecular Structure* **830**, 188-197, (2007).

16 Fulton, B. & Goa, K. L. Olanzapine. *Drugs* **53**, 281-298, (1997).

17 Askin, S. *et al.* Olanzapine Form IV: Discovery of a New Polymorphic Form Enabled by Computed Crystal Energy Landscapes. *Crystal Growth & Design*, (2019).

18 Bhardwaj, R. M. *et al.* Exploring the Experimental and Computed Crystal Energy Landscape of Olanzapine. *Crystal Growth & Design* **13**, 1602-1617, (2013).

19 Sun, Y. *et al.* Modeling Olanzapine Solution Growth Morphologies. *Crystal Growth & Design* **18**, 905-911, (2018).
20 Warzecha, M. *et al.* Direct Observation of Templated Two-Step Nucleation Mechanism during Olanzapine Hydrate Formation. *Crystal Growth & Design* **17**, 6382-6393, (2017).
21 Chernov, A. A. *Modern Crystallography III, Crystal Growth,*. (Springer, 1984).
22 Gilmer, G. H., Ghez, R. & Cabrera, N. An analysis of combined volume and surface diffusion processes in crystal growth. *J. Crystal Growth* **8**, 79-93, (1971).
23 De Yoreo, J. J. & Vekilov, P. G. Principles of crystal nucleation and growth. in *Biomineralization* Vol. 54 *Reviews in Mineralogy & Geochemistry* 57-93 (2003).
24 Warzecha, M., Safari, M. S., Florence, A. J. & Vekilov, P. G. Mesoscopic Solute-Rich Clusters in Olanzapine Solutions. *Crystal Growth & Design* **17**, 6668-6676, (2017).
25 Jiang, Y. *et al.* Growth of organic crystals via attachment and transformation of nanoscopic precursors. *Nature Communications* **8**, 15933, (2017).
26 Yau, S.-T., Thomas, B. R. & Vekilov, P. G. Molecular mechanisms of crystallization and defect formation. *Phys. Rev. Lett.* **85**, 353-356, (2000).
27 Lovette, M. A. & Doherty, M. F. Multisite models to determine the distribution of kink sites adjacent to low-energy edges. *Physical Review E* **85**, 021604, (2012).
28 Joswiak, M. N., Doherty, M. F. & Peters, B. Ion dissolution mechanism and kinetics at kink sites on NaCl surfaces. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 656-661, (2018).
29 Shtukenberg, A. G., Ward, M. D. & Kahr, B. Crystal Growth with Macromolecular Additives. *Chemical Reviews* **117**, 14042-14090, (2017).
30 Ma, W., Lutsko, J. F., Rimer, J. D. & Vekilov, P. G. Antagonistic cooperativity between crystal growth modifiers. *Nature* **577**, 497-501, (2020).
31 Ehrlich, G. & Hudda, F. G. Asymetric capture at steps. *J. Chem. Phys.* **44**, 1039-1052, (1966).
32 Olafson, K. N., Rimer, J. D. & Vekilov, P. G. Early Onset of Kinetic Roughening due to a Finite Step Width in Hematin Crystallization. *Physical Review Letters* **119**, 198101, (2017).
33 Vekilov, P. G., Kuznetsov, Y. G. & Chernov., A. A. The effect of temperature on step motion; (101) ADP face. *J. Crystal Growth* **121**, 44-52, (1992).
34 Vekilov, P. G. What determines the rate of growth of crystals from solution? . *Crystal Growth and Design,* **7**, 2796-2810, (2007).
35 Dandekar, P., Kuvadia, Z. B. & Doherty, M. F. Engineering Crystal Morphology. *Annual Review of Materials Research* **43**, 359-386, (2013).
36 Erdemir, D. *et al.* Relationship between self-association of glycine molecules in supersaturated solutions and solid state outcome. *Physical review letters* **99**, 115702, (2007).
37 Parveen, S., Davey, R. J., Dent, G. & Pritchard, R. G. Linking solution chemistry to crystal nucleation: the case of tetrolic acid. *Chemical Communications*, 1531-1533, (2005).

Acknowledgments

We thank Michael Doherty, Sarah Price, Elias Vlieg, Dominique Maes, and Jeffrey Rimer for invaluable discussions and suggestions. This work was supported by the National Science Foundation (Award No. DMR-1710354), and NASA (Award Nos. NNX14AD68G and NNX14AE79G). The authors acknowledge the CMAC National Facility, housed within the University of Strathclyde's Technology and Innovation Centre, and funded with a UK Research Partnership Investment Fund (UKRPIF) capital award, SFC ref H13054, from the Higher Education Funding Council for England (HEFCE).

Author contributions

P.G.V., M.W., and A.J.F. conceived this work, P.G.V. and M.W. designed experiments, M.W. performed all experiments and analyzed data, B.J. and M.W. modelled the Raman spectra and adsorption on the OZPN crystal surface, P.G.V. developed kinetic and thermodynamic models, L.V. and J.C.P. carried out MD simulations, and P.G.V, M.W., L.V., J.C.P., and A.J.F. wrote the paper. All authors discussed the results and commented on the manuscript.

Competing interest statement

The authors declare no competing financial interest.

Figure Legends

Figure 1 | The structure and growth of OZPN crystals. a. The OZPN molecule: Ph, benzene ring; DZP, diazepine ring; TPh, thiophene ring; PZ, piperazine ring. Atom numbering follows Bhardwaj, *et al.*, [18]. **b and c.** Optical micrograph, with the (002) face upwards in b, and the crystal structure of the dihydrate ethanoate mixed solvate 2OZPN·EtOH·2H_2O in space group $P2_1/c$ (Cambridge Structural Database REFCODE WEXQEW [15]) in c. In **c** one unit cell is shown; one centrosymmetric OZPN SC_0 dimer is highlighted in turquoise. In the remaining OZPN molecules, C atoms are shown in charcoal, H atoms, in silver, N, in blue and S, in yellow. In the EtOH molecules, the C atoms are drawn in purple, O in red and H, in red. In the water molecules, O is drawn in red and H, in gold. **d.** Generation of new crystal layers by a screw dislocation outcropping on the (002) face at C = 3.87 mM. Inset: height profile along the horizontal line extending left of the dislocation outcrop point. Height jumps indicate steps generated by the dislocation; average step height, ca. 10 Å, approximately equals half of the c lattice parameter illustrated in **c**. **e.** Evolutions of the step displacement in the [110] direction with respect to a fixed reference point measured from sequential AFM images, as illustrated in Supplementary Fig. 1, at the indicated $C - C_e$; the step velocity v is determined as the slope of these correlations; error bars indicate the standard deviation from the average of 17 measurements for all points at $C - C_e$= 0.21 mM, 15 at 0.33 mM and 0.43 nM, 20 at 0.70 mM, 17 at 1.95 mM and 2.03 mM, 20 at 2.55 mM, and 15 at 2.55. **f.** The velocity v of steps growing in the [110] direction separated by distances greater than 250 nm as a function of OZPN concentration C in 1/1 (v/v) EtOH/H_2O. The solubility of 2OZPN·EtOH·2H_2O crystals in this solvent, C_e = 2.05 mM, at 21°C. **g.** The linear correlation between v and $(C^2 - C_e^2)$ has R = 0.95. Dotted lines in **e** and **f** depict the relation $v = \beta_D K_D \Omega_D (C^2 - C_e^2)$, where β_D is the kinetic coefficient for growth by dimer incorporation, $\Omega_D = 2\Omega_M$ = 0.94 nm^3 is the volume occupied by a dimer in the crystal, and K_D is the dimerization equilibrium constant. Error bars in **e** and **f** indicate the standard deviation from the average v determined as the slope of the displacement-time correlations and are smaller than the symbol size for most data points.

Figure 2 | Potential mechanisms of observed accelerating step growth. a. The solubility of 2OZPN·EtOH·2H_2O crystals determined as the highest concentration reached by dissolution of excess crystals in bulk experiments,[24] green spheres, and AFM measurements illustrated in b, blue diamond. The standard deviation determined from three independent measurements is smaller than the symbol size. **b.** *In situ* AFM kymograph collected with disabled scanning along the y-axis at 21°C. In this imaging mode, the vertical axis represents time. **c.** Schematic of the potential contribution of mesoscopic solute-rich clusters (brown) to the growth an OZPN crystal (gold, steps are shown in blue and kinks, in red). Callout: Schematic highlighting dimer-enriched composition of the cluster phase. **d.** An AFM image of clusters (indicated with yellow arrow) formed in OZPN EtOH/H_2O solution land on OZPN crystal surface at C = 2.68 mM. **e.** An AFM image after scanning with increased tip force, which reveals that clusters partially incorporated into the crystal (indicated with green arrow) etch faster than the underlying crystal revealing that they do not integrate into the crystal lattice. **f.** AFM image of a step at $C = C_e$. **g.** A lattice resolution AFM image of a (002) face. A model of the structure of one 2OZPN·EtOH·2H_2O crystal layer, comprised of SC_0 dimers, is overlaid (C, blue; H, silver; N, blue; S, yellow). Inset: Fourier transform of the AFM image illustrates rhombic symmetry of the molecular arrangement on (002) face; detectable fourth-order peaks correspond to 3.78 Å resolution.

Figure 3 | Schematic of two alternative growth mechanisms. Upper path: two OZPN monomers (space filling model, C, cyan; H, silver; N, blue; S, yellow) form a dimer in the solution in step one, which incorporates into the crystal (gold, steps are shown in blue and kinks, in red) as a whole in step two. Lower

path: the two monomers incorporate sequentially in steps one and two, respectively; the SC_0 dimer forms in step two, in the crystal.

Figure 4 | OZPN dimers in solution. a. Raman spectra of OZPN dissolved in 1/1 EtOH/H_2O and $CHCl_3$ at listed concentrations and of two solid solvates are compared to model spectra for OZPN monomer and dimer. **b. and c.** All-atom molecular dynamics calculation of the potential of mean force ΔF between two OZPN monomers in 1/1 EtOH/H_2O in **b** and $CHCl_3$ in **c.** The line width represents the computational uncertainty. Ten representative snapshots of the configuration of the dimer (the constituent monomers are shown in red and blue, respectively) occupying the deep ΔF minima for each solvent are shown; one of the ten is highlighted for clarity. Four representative configurations of the two monomers at separations longer than the deep minima are shown for each solvent. Insets: The root-mean-squared deviations of the positions of all OZPN atoms in the dimers occupying the deep minima from the SC_0 dimer structure in the 2OZPN·EtOH·$2H_2O$ crystals as functions of time are small, indicating that the dimers assembled in the solution are nearly identical to the SC_0 dimer found in the crystal structure.

Figure 5 | Why is growth by dimers faster? a. Schematic of two pathways of a growth unit, monomer or dimer, from solution to kinks: direct incorporation, *i*, and *via* adsorption on the terraces followed by diffusion towards the steps, *ii*. Thick red arrow indicates direction of step growth. Thin brown arrows indicate alternative supply fluxes towards the steps. **b. and c.** Evolution of the step configuration over 30 min during growth at C = 2.38 mM. Thick red arrows indicate direction of step growth. Steps are highlighted in yellow, blue, green and pink for easy reference; original images are in Supplementary Figure 2. **d.** The dependence of the step velocity v on the step separation l at C = 2.38 mM. **e.** Enthalpy variation along the reaction coordinate for the surface diffusion mechanism. $\Delta H_{ads}^{\neq}$, $\Delta H_{des}^{\neq}$, $\Delta H_{SD}^{\neq}$, $\Delta H_{kink}^{\neq}$ are the activation barriers for, respectively, adsorption, desorption, surface diffusion, and incorporation into kinks from the surface. The equilibrium enthalpy of adsorption $\Delta H_{ads}^{o} = \Delta H_{ads}^{\neq} - \Delta H_{des}^{\neq}$. **f.** Water molecules associated with an OZPN molecule along the reaction pathway of dimer incorporation; binding energies lower than –20 kJ mol^{-1} (ref. [20]) are shown. Waters bound with energies higher than –20 kJ mol^{-1} (pink) are released when OZPN dimer is formed in the solution; waters bound strongly with listed energies (orange) are released when a dimer enters the crystal leaving two waters (red) in the unit cell.

Figure 6 | The adsorption of monomers and dimers on the (002) OZPN face. MD simulation results. OZPN molecules in the crystal are represented with their solvent accessible surface, drawn in blue, and the lattice water molecules, with their van der Waals surfaces, where red beads represent O and white beads, H. In the solution, EtOH molecules are shown as thin sticks; solvent waters are omitted for clarity. C atoms are shown in cyan, O in red, and N in blue. OZPN monomers and dimers in the solution and at the surface are shown as thick sticks. Only two of the numerous surface configurations of the adsorbed monomers and dimers are shown. ΔH_{ads}^{o} are evaluated as the difference between the averaged enthalpies of monomers and dimers on the crystal surface and in the solution bulk (Methods).

Methods

Materials.

Most solvents were analytical grade and purchased from Fisher Scientific, UK. Deuterated solvents for nuclear magnetic resonance (NMR) characterization were purchased from Sigma-Aldrich, UK. Olanzapine (OZPN) was purchased from Molekula Ltd., UK. The commercial preparation was confirmed as form I by powder X-ray diffraction (PXRD) and was used without further purification.

Crystallization.

2OZPN·EtOH·2H_2O crystals with space group $P2_1/c$ for AFM analysis were grown to lengths of 0.2-2 mm by dissolving excess OZPN I in 5 mL of 1/1 (v/v) EtOH/H_2O mixture. The solution was heated up to 50°C and stirred for around 15 min. The solution was then filtered through 0.22 μm PES filters to clean 5 mL vials containing coverslips scratched near the center. This procedure ensured that the crystals were immobilized on the coverslips without glue. The coverslip with OZPN crystals was removed from the vials and rinsed with 1/1 (v/v) EtOH/H_2O to remove loosely attached crystals. The presence of the desired 2OZPN·EtOH·2H_2O was confirmed by powder and single-crystal XRD (Extended Data Fig. 8). The coverslips with the attached crystals were mounted on the AFM scanner.

Crystal morphology.

OZPN hydrate ethanoate (HE) has been reported to crystallize in two polymorphic forms: A, with space group $P2_1/c$ [38] and B, with space group $C2/c$ [39,40]. OZPN HE form A is structured as parallel dimers of olanzapine molecules, whereas form B shows orthogonal arrangement of dimers with ethanol and water molecules occupying the cavity [40]. OZPN HE form A crystals were obtained by cooling crystallization from mixtures of EtOH (between 10 – 90% v/v) in H_2O. The unit cell dimensions are summarized in Warzecha *et al.* [41]. The crystals are pale yellow in color

and exhibit a rhombohedral morphology with large {001} faces and smaller {110} faces (Extended Data Fig. 8).

Single crystal face indexing.

The crystal faces were indexed on a Bruker D8 Venture diffractometer equipped with a CCD detector using graphite-monochromated Cu Kα radiation (λ=1.54056 Å) and an APEX-3 face indexing plug-in.

Time resolved *in situ* atomic force microscopy.

Washed crystals attached to a glass cover slip were imaged with AFM FastScan (Bruker) using FastScan D SS probe (force constant 0.25 N/m and tip radius 1 nm) in OZPN supersaturated solutions with concentration C between 2.25 and 4.60 mM. Temperature in the AFM liquid cell equaled room temperature, 21°C. All AFM data were analyzed using NanoScope Analysis 1.5 software (Bruker) [42]. Height images were corrected by first order flattening and plane fitting. No further image processing was performed.

The solubility measurements using AFM involve the observation of the behavior of the step on the surface of OZPN HE. At the solubility, the steps neither grow nor dissolve, although the step edge fluctuates owing to the attachment and detachment of individual molecules to it. To obtain such an image, during data collection the y-axis is disabled and the AFM tip moves only in the x-direction while the y-coordinate now represents time, generating a pseudoimage (kymograph), in which the vertical dimension represents time (Figure 2b). Several molecular attachment and detachment events from and to the step are recorded, although no net growth or dissolution of the step take place.

Removal of the mesoscopic olanzapine-rich clusters.

AFM observations revealed that in some experiments, clusters similar to those discussed in Ref. [41] attached to growing OZPN crystal surfaces. The cluster number concentration was

significantly reduced after filtration through a 100 nm pore size PTFE filter (Whatman Puradisc). In the filtered solutions the clusters reappeared after several hours. During OZPN crystal growth from freshly filtered solutions clusters were not observed on the crystal surface.

Solution Raman spectroscopy.

We used a Kaiser RXN2 Raman spectrometer with PhAT probe to monitor the solution non-invasively and avoid crystallization on the probe. We used an Invictus diode laser operating at 350 mW and wavelength of 785 nm. We optically expanded the laser beam to a 3 mm spot size. This allows simultaneous analysis of a large sample area. This beam had a focal length of 12.5 mm. The beam was directed at the solution sample through a side of the vial. The scattered intensity was recorded using a CCD detector cooled to −40 °C by Peltier coolers. Before any measurements, the spectrometer was calibrated by measuring the relative peak intensity and peak position of a cyclohexane standard. Raman spectra were recorded using the iC Raman V4.1 software package with 2 scans and 5 s integration time for samples in chloroform and 4 scans and 5 s integration time for EtOH/H_2O samples. Solvent and glass spectra were subtracted.

Solid-state Raman spectroscopy.

Raman spectra for OZPN crystals were collected using a XplorA Microscope Horiba Scientific (Jobin Yvon Technology) Raman microscope with a 532 nm (diode-pumped, solid state) laser excitation source and CCD equipped with a high precision X, Y, Z motorized stage in the shift region of 50–3400 cm^{-1} over 10 scans using a resolution of 1 cm^{-1} and an exposure time of 1 s.

Modelling of Raman spectra.

The structures of the OZPN monomer and dimer were directly transferred from the OZPN I crystal structure. The structures were geometrically optimized using DMol^3 module in Materials Studio 7.0 (Accelrys Software Inc.). We used the GGA PBE functional with convergence tolerance for energy of 1.0 x 10^{-4} H, a maximum force of 0.02 H/Å, maximum displacement of 5.0 x 10^{-2} Å,

maximum iterations 50, maximum step size 0.3 Å, and 4.4 DN. Raman spectra of OZPN dissolved in ethanol, water and chloroform were generated based on frequency calculations using COSMO (COnductor-like Screening MOdel) [43] and the DMol[3] module in Materials Studio (Accelrys Software Inc.) COSMO treats each solvent as a continuum with a permittivity ε surrounding the solute molecules outside of a molecular cavity [43]. COSMO derives the polarization charges of the continuum, caused by the polarity of the solute, from a scaled-conductor approximation.

Characterization of the dimerization equilibrium.

Raman spectroscopy has become one of the fastest and most reliable techniques to identify chemical environment of species in solution and in a solid form. It has been applied for quantitative measurement of solute concentration [44], solid concentration in slurries [45], to monitor polymorphic transformation [46], and to determine dimerization constants [47].

We used Raman spectroscopy to determine the composition of OZPN monomers and dimers in the solutions at several concentrations in $ChCl_3$ and EtOH/H_2O (Figure S1 a, b). The collected spectra were compared with Raman spectra of OZPN solid forms containing the SC_o dimer (EtOH/H_2O and chloroform solvates) and the simulated spectra for the OZPN monomer and dimer models generated with DMol[3] COSMO (Figure S1 c, d). This comparison indicates whether OZPN dimers are present in solution.

To evaluate the dimerization constant, we used Raman spectra of OZPN at distinct concentrations in $CHCl_3$ and EtOH/H2O (Extended Data Figure 1 a and b). Detailed band assignation is described by Ayala *et al.* [48]. In both solvents, we observed an increase of the intensity of 1500-1600 cm^{-1} bands and a decrease of the bands in region 1150-1350 cm^{-1} (Extended Data Figure 1 a and b) with increasing OZPN concentration. Based on simulated Raman spectra for OZPN, the monomer bands at 1150-1350 cm^{-1} correspond to bending of both methyl groups δCH3 and methylene groups δCH2 of the piperazine ring. During SC_o dimer formation these groups are hindered, resulting in a reduction of the intensity of the 1150-1350 cm^{-1} bands. The SC_o dimer is

stabilized by multiple C–H···π interactions between the piperazine, thiophene, and phenyl rings. These interactions contribute to the high intensity in the 1500-1600 cm^{-1} region observed for OZPN solvate crystals that contain SC_o dimer, and are associated with three types of motions: 1) stretching νC=C of the benzene ring, 2) stretching νC=C of the thiophene ring, and, 3) stretching νC=N coupled with bending δNH deformations of diazepine ring and bending of thiophene δCH_3. Based on the observation that the simulated SC_o dimer spectrum and measured spectra for high OZPN concentrations in EtOH/H2O and $CHCl_3$ show similar high intensity peak in the 1500-1600 cm^{-1} region, we conclude that SC_o dimer is formed in solution.

The Raman intensity is directly related to the concentration of the species generating the signal [10]

$$I = JC \tag{1}$$

where J is the molar intensity coefficient, and C is the concentration of the respective species.

We assume that no higher aggregates than dimers form in OZPN solutions. Then, the molar analytical concentration of OZPN C and the concentrations of monomers C_M and dimers C_D are related

$$2C_D + C_M = C\ . \tag{2}$$

Assuming dimerization equilibrium,

$$2M \ \rightleftarrows \ D\ \text{, for which } K_D = C_D/C_M^2\ , \tag{3}$$

we obtain an expression for dimerization equilibrium constant K_D

$$\frac{C-C_M}{2C_M^2} = K_D\ , \tag{4}$$

and the following relations:

$$2K_D C_M^2 + C_M - C = 0\ , \tag{5}$$

$$C_M = \frac{\sqrt{1+8K_D C}-1}{4K_D}, \tag{6}$$

and $$C_D = K_D C_M^2 = \frac{1}{16K_D}\left[\sqrt{1 + 8K_D C} - 1\right]^2 . \quad (7)$$

Solving Eq. (2) for C_M and substituting the resulting relation in the equilibrium constant expression yields a result for C_D identical to Eq. (7).

The band at 1517 cm^{-1} is associated with the dimer and it preserves its shape at all OZPN concentrations. The intensity of this peak at the highest tested C was selected as an internal intensity standard I_{st}. We chose about 20 values of K_D and for each value K_D computed the C_D for each of the tested concentrations. We refer to computed C_D at the highest C as C_{Dst}. The expected intensity of the peak at 1517 cm^{-1} for each assumed K_D was calculated via

$$I_D = \frac{C_D I_{st}}{C_{Dst}} . \quad (8)$$

For each assumed K_D, the the root mean squared deviation (RMSD) of the computed I_D from the experimentally measured values at five different OZPN concentrations in both in 1/1 (v/v) $EtOH/H_2O$ and $CHCl_3$ (Extended Data Figure 3). The K_D that yielded the lowest RMSD was taken as the best estimate of the dimerization constant.

Nuclear Magnetic Resonance.

Stock solutions were prepared by dissolving the maximum amount of OZPN I soluble at room temperature in deuterated $CDCl_3$ and 1/1 (v/v) $EtOD/D_2O$. The mass of the dissolved OZPN was verified after the samples were completely dried and desolvated. ^{1}HNMR spectra were collected at various concentrations. For this, stock solutions in each solvent was diluted at ratios for chloroform-D, 0.80, 0.40, 0.10; for $EtOD/D_2O$, 0.95, 0.60, 0.25. All solutions were filtered through 0.22 μm PTFE filters and NMR spectra were collected up to 15 min after sample preparation using a Bruker Advance 3HD two-channel 500 MHz spectrometer. Sixteen scans were acquired for each sample at room temperature using a residual solvent peak as the reference. Spectra were analyzed using Top Spin 3.5 (Bruker) and the shifts were quoted in ppm on the δ scale.

The structure of OZPN self-associates was also examined by NMR spectroscopy in 1/1 EtOD/D_2O and $CDCl_3$. ^1H NMR concentration dependent studies of OZPN in 1/1 EtOD/D_2O did not show any changes in chemical shifts (Extended Data Figure 2a). The accessible concentration range is low due to low OZPN solubility in this solvent [39].

A significantly higher OZPN concentration is accessible in chloroform. ^1H NMR dilution data in $CDCl_3$ (Extended Data Figure 2, Supplementary Table 1) show concentration dependent changes in chemical shift suggesting association of OZPN molecules into a molecular complex. There is an up-field shift of all piperazine (PZ) protons (H12, 13, 14, 15, Figure 1a), the methyl group of thiophene (TPh) ring (H16), and the phenyl (Ph) protons (H8) with increasing OZPN concentration. Complexation induced shifts (CIS) are an indication of either single equilibrium of self-association of OZPN molecules into a dimer or a competition between multiple structures. The number of self-association species can be estimated from the shape of normalized dilution curves [49,50]. If there are many competitive complexes formed the shifts in the NMR spectrum will be a result of a fast exchange of protons between these complexes and dilution curves will not follow the same trend. The dilution curves for OZPN in $CDCl_3$ (Extended Data Figure 2 c) follow this trend and indicate that only one complexation process takes place in this solvent. The observed shifts of all piperazine protons, the methyl group of the thiophene ring and aromatic protons indicate multiple C–H···π interactions between OZPN molecules, corresponding to those observed in the structure of SC_o dimer (Extended Data Figure 2 d).

Molecular dynamics (MD) simulations.

We employed all-atom classical MD simulations to study the dimerization and adsorption of olanzapine in an ethanol-water mixture; as well as dimerization in pure chloroform. We used GROMACS 5.1.5 [51] and the TIP4P/Ew [52] model for water (H_2O). Olanzapine (OZPN), ethanol (EtOH), and chloroform ($CHCl_3$) were described using the all-atom variant of the optimized potential for liquid simulations force field (OPLS-AA), with parameters assigned by the

LigParGen server [53,54]. In all simulations, van der Waals and Coulomb forces were truncated at a 1 nm cutoff, and long-range contributions to the electrostatic interactions were handled using the particle-mesh Ewald method [55] with parameters chosen to ensure a relative error of less than 10^{-5} kJ/mol in the calculated energy. The MD trajectories were propagated using a leap-frog integration scheme with a 2 fs time step. To facilitate use of this large time step, degrees of freedom associated with fast bond vibrations were removed by simulating water molecules as rigid bodies using the SETTLE algorithm [56] and constraining bonds between hydrogen and heavy atoms in other molecular species using the LINCS algorithm [57]. Isothermal conditions were maintained using a Bussi-Parrinello velocity rescaling thermostat [58] with a 0.1 ps time constant; separate thermostats were used for molecules in liquid and solid phases in simulations where both were present. Simulations under isobaric and isostress conditions were performed using a Parrinello-Rahman barostat [59] with a 2 ps time constant.

The OPLS-AA force field accurately describes the thermophysical properties of the solvents (ethanol-water mixture [60,61] and chloroform [62]) examined in our study. We also confirmed that it provides a reasonable description of olanzapine by performing simulations of the crystalline 2OZPN·EtOH·$2H_2O$ solvate. Crystallographic information for the monoclinic unit cell (space group $P2_1/c$) was obtained from the x-ray diffraction (XRD) data reported in the Cambridge Structural Database CSD (REFCODE WEXQEW [38]). The initial configuration for the simulations was generated by replicating the unit cell to create a $4 \times 4 \times 2$ supercell. The supercell structure was simulated at 300 K and 1 bar for 6 ns using an anisotropic Parrinello-Rahman barostat [59] to impose isostress conditions. Structural parameters were computed by averaging over configurations saved every 10 ps over the last 5 ns of the simulation and compared against the XRD data (Supplementary Table 3). The computed structural parameters agreed well with XRD data, with less than 2% deviation in lattice parameters and volume (V), supporting the appropriateness of the chosen force field.

MD simulations of olanzapine dimerization.

The dimerization equilibrium in both solvents was probed by umbrella sampling MD (USMD) [63-65] simulations at 300 K and 1 bar to compute the potential of mean force $F(r)$ parameterized by the monomer-monomer center-of-mass separation distance r. For the ethanol-water mixture, two OZPN monomers were solvated with 2041 H_2O and 630 EtOH molecules to achieve the same composition (50 vol. %) used in the crystallization experiments. For the pure $CHCl_3$ solvent, the two OZPN monomers were solvated with 1140 molecules. The number of solvent molecules was chosen to ensure that the linear dimensions of the equilibrated simulations cells were ca. 5.0 nm, thereby allowing the disassociated state (large r region of $F(r)$) to be sampled.

Independent USMD simulations were performed in overlapping windows along r to systematically sample the full range of the order parameter space relevant for studying OZPN dimerization. Sampling of each window was performed with a harmonic restraint potential $U(r) = \frac{1}{2}k\,(r - r^*)^2$ applied using PLUMED v. 2.4.3 [66], where k is the spring constant and r^* specifies the window's center. The spring constant k and r^* spacing were chosen to ensure sufficient statistical overlap between simulations performed in adjacent windows. A total of 25 and 37 evenly-spaced windows on the interval $[0.36, 2.03]$ nm were used for the calculations performed in ethanol-water and chloroform solvents, respectively, with k typical values ranging between 2000-5000 kJ/nm^2. Each simulation was equilibrated for 10 ns, followed by a production period of 20 ns. Histogram data collected during the production periods of the simulations were combined using BayesWHAM [67], a Bayes reformulation of weighted histogram analysis method [68], to obtain an estimate of the PMF $F^{WHAM}(r)$. The final estimate of the PMF was obtain by applying the standard Jacobian correction factor [69] associated with the order parameter r via $F(r) = F^{WHAM}(r) + 2kTln(r)$.

The dimerization equilibrium constant, K_D, for OZPN in each solvent was calculated from $F(r)$ using [70]

$$K_D = \frac{4\pi R^3 P_b}{3v^{\emptyset} P_u} = \frac{4\pi R^3 \int_0^{r_b} r^2 \exp[-\beta F(r)]dr}{3v^{\emptyset} \int_{r_b}^{R} r^2 \exp[-\beta F(r)]dr} \quad (9)$$

where P_b and P_u are the probabilities of the bound (dimerized) and unbound (disassociated) states, respectively, $v^{\emptyset} = (c^{\emptyset} N_{Av})^{-1} = 1.66$ nm^3 is the standard volume of a single molecule, $c^{\emptyset} =$ 1 mol/L is the standard concentration, N_{Av} is Avogadro's number, k_B is Boltzmann constant, T is temperature, $\beta = (k_B T)^{-1}$, $R = 1.9$ is the maximum value of r considered in the integration, and r_b is the position of diving surface along the order parameter r that delineates the bound and unbound states. We chose r_b to be near the barrier observed in $F(r)$ in each solvent; uncertainties in K_d were estimated based on the sensitivity of the computed value to the position of the diving surface.

Enthalpies of dimerization were calculated from the differences in enthalpies between the bound and unbound states $\Delta\langle H\rangle_{dimer} = \langle H\rangle_b - \langle H\rangle_u$, where the brackets $\langle \ldots \rangle$ denote ensemble averages, $\langle H\rangle = \langle E\rangle + P\langle V\rangle$, $\langle E\rangle$ is the total energy, $\langle V\rangle$ is the system volume, and P is the imposed pressure. The quantity $\langle H\rangle_b$ was evaluated from data from umbrella sampling windows near the contact minimum in $F(r)$, whereas $\langle H\rangle_u$ was evaluated from windows at large r where the PMF plateaus. Standard reweighting procedures [64,67] were used to remove the effects of the umbrella restraint potentials on the ensemble averages and obtain unbiased estimates of the enthalpies.

MD simulations of adsorption of OZPN monomers and dimers at the solution-crystal interface.

To model adsorption, we assumed that the (002) surface of the 2OZPN·EtOH·2H_2O crystal described above was exposed to solvent. For this, we increased the size of simulation cell to 7.6 nm in the direction parallel to the crystallographic *z*-direction to create an infinite slab with empty vacuum on either side. All other cell parameters left were set fixed to the average values computed from simulation (Supplementary Table 3). A single OZPN dimer or monomer was inserted in the

vacuum space above the crystal, and the remaining empty space was solvated with the ethanol-water mixture.

Enthalpies of adsorption for the OZPN dimer and monomer were estimated using [71] $\Delta\langle H\rangle_{ads.} = \langle H\rangle_{surf} - \langle H\rangle_{bulk}$, where the brackets $\langle \dots \rangle$ denote ensemble averages, and $\langle H\rangle_{surf}$ and $\langle H\rangle_{bulk}$ are the enthalpies computed for states in which the species is adsorbed at the surface and in the bulk liquid region far away from the surface, respectively. The enthalpy in each state was computed via $\langle H\rangle = \langle E\rangle + \sigma_{zz}\langle V\rangle$, where σ_{zz} is the magnitude of the imposed stress normal to the surface. Sampling of the two states was achieved by using PLUMED v. 2.4.3 [66] to apply restraint potentials along the order parameter z, define as the distance between the specie's center of mass and the plane containing the nitrogen atoms of the piperazine rings, which are exposed on the surface of the crystal. To sample the adsorbed state, a quadratic wall $U^{res}(z) = k(z - z^*)^2$ for $z > z^*$; 0 otherwise, with force constant $k = 1000\ \mathrm{kJ\ nm^{-2}}$ was placed at $z^* = 1.2$ nm above the surface. To sample the desorbed state, lower and upper quadratic walls $k = 1000\ \mathrm{kJ\ nm^{-2}}$ with were applied at $z^* = 2.5$ and $z^* = 3.0$ nm to keep the species in the bulk liquid region, where interactions with the wall crystal surface are negligible.

The systems were equilibrated at 300 K for 10 ns at constant volume, followed by a 500 ns simulation in which a semi-isotropic Parrinello-Rahman barostat [59] was applied to allow the z-dimension of the simulation cell to fluctuate in order to impose an isostress condition of σ_{zz} =1 bar; all other dimensions of the cell remained fixed. Averages for each state were computed from statistics collected from 10 independent simulations, discarding data from the first 200 ns of each MD trajectory to ensure equilibration. Ensemble statistics were reweighted [65,67] to obtain unbiased estimates of $\Delta\langle H\rangle_{ads.}$

The appreciable uncertainties in the computed ΔH^o_{ads} values reflect the slow structural relaxations near the surface of the crystal, which frustrate sampling of OZPN monomer and dimer conformations in the adsorbed state. These slow relaxations arise from solvent structuring at the

crystal interface and barriers to reorientation of the monomer and dimer due to transient hydrogen bonding with the surface (Extended Data Figure 7 b – g). As a result, long simulations were needed to sample configurations of the OZP monomer and dimer near the surface and converge estimates of ΔH^{o}_{ads}.

The correlation between the step velocity and total solute concentration in the presence of dimers.

The step velocity relates to the molecular fluxes of association j_+ and dissociation j_- of molecules to a step as

$$v = \Omega(j_+ - j_-) \tag{10}$$

where Ω is the volume occupied in the crystal by the unit entering the step; in the case of monomer and dimer coexistence, $\Omega_D = 2\Omega_M$. Assuming that the kinetics of molecular association to a step are monomolecular [72,73],

$$j_+ = \frac{D_i}{\Lambda_i} C_i \tag{11}$$

where D_i is the diffusion coefficient of the associating species and Λ_i, resistance to enter the step in units of length; the ratio D_i/Λ_i is equal to the step kinetic coefficient β_i introduced by Chernov [73,74]. At equilibrium, $v = 0$, and

$$v = \frac{D_i \Omega_i}{\Lambda_i}(C_i - C_{ie}) \; . \; ,$$

where C_{ie} is the concentration of the associating species at equilibrium. Assuming that j_- is independent of C_i we obtain

$$v = \frac{D_i \Omega_i}{\Lambda_i}(C_i - C_{ie}) \; . \tag{12}$$

Below, we derive expressions for the step velocity v as a function of total solute concentration C for four possible cases when dimers are present in the growth solution:

(i) Monomers dominate in the solution and growth occurs by the association of monomers.

(ii) Monomers dominate in solution, although growth occurs by association dimers.

(iii) Dimers dominate in solution but the growth occurs by monomers.

(iv) Dimers dominate in solution and growth occurs by association of dimers.

(i) Monomers dominate in the solution and growth occurs by incorporation of monomers. In this case, Eq. (12) transforms into

$$v = \frac{D_M \Omega_M}{\Lambda_M}(C_M - C_{Me}) \tag{13}$$

Assuming that the concentration of dimers in the solution is much lower than the concentration of monomers, $C_D \ll C_M$, the dimerization constant $K_D \ll 1/C_M$, $K_D C_M \ll 1$ and $8K_D C_M < 1$. Using the approximations valid for small x

$$(1+x)^{1/2} \cong 1 + \frac{1}{2}x - \frac{1}{8}x^2$$

$$\left[(1+x)^{\frac{1}{2}} - 1\right] \cong \frac{1}{2}\left[x - \left(\frac{x}{2}\right)^2\right] \cong \frac{x}{2}$$

and designating $x = 8K_D C$, Eq. (6) from Methods transforms to

$$C_M = \frac{1}{8K_D} 8K_D C = C \tag{14}$$

A linear relationship is expected between step velocity, v, and the total OZPN concentration (Extended Data Figure 4)

$$v = \frac{D_M \Omega_M}{\Lambda_M}(C - C_e)\,. \tag{15}$$

Numerous such dependencies have been recorded in the literature [75-78].

(ii) Monomers dominate in the solution, but growth occurs by incorporation of dimers. In this case, $C_D \ll C_M$, $K_D \ll 1/C_M$, and $8K_D C_M < 1$. Using the approximations for $x \ll 1$

$$(1+x)^{1/2} \cong 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \cdots$$

$$\left[(1+x)^{\frac{1}{2}} - 1\right]^2 \cong \frac{1}{4}\left[x - \left(\frac{x}{2}\right)^2\right]^2 \cong \frac{x^2}{4}$$

and assuming $x = 8K_DC$, Eq. (7) from Methods transforms to

$$C_D = \frac{1}{16K_D}\frac{64K_D^2C^2}{4} = K_DC^2 \tag{16}$$

This result is intuitively expected for $C_D \ll C_M$ since $C_M = C - 2C_D \cong C$ and $C_D = K_DC_M^2 = K_D(C - 2C_D)^2 \cong K_DC^2$. We get for j_+, j_-, and v

$$j_+ = \frac{D_D}{\Lambda_D}K_DC^2, \qquad j_- = j_e = \frac{D_D}{\Lambda_D}K_DC_e^2\ ,$$

and

$$v = \frac{D_D\Omega_D}{\Lambda_D}K_D(C^2 - C_e^2) \tag{17}$$

Eq. (17) indicates that in the case of growth by dimer incorporation from a solution, in which the dimers are the minor component, the correlation between the step velocity and the total solute concentration is quadratic.

(iii) Dimers dominate in the solution, but growth occurs by incorporation of monomers. In this case, $C_D \gg C_M$ and $K_D \gg 1/C_M$, $8K_DC > 8K_DC_M > 1$ and $\sqrt{1 + 8K_DC} > 1$. With this,

$$C_M = \frac{\sqrt{8K_DC}}{4K_D} = \frac{1}{2}\sqrt{\frac{2C}{K_D}}$$

and, from Eq. (13),

$$v = \frac{D_M\Omega_M}{\Lambda_M}\sqrt{\frac{1}{2K_D}}\left(\sqrt{C} - \sqrt{C_e}\right) \tag{18}$$

Eq. (18) corresponds to a sublinear correlation between the step velocity and the solute concentration, illustrated in Extended Data Figure 4.

(iv) Dimers dominate in the solution and associate to the steps. In this case, $C_D \gg C_M$ and $K_D \gg 1/C_M$, $8K_DC > 8K_DC_M > 1$ and $\sqrt{1 + 8K_DC} > 1$. From Eq. (4) in Methods we obtain

$$C_D = \frac{C}{2}$$

and

$$v = \frac{D_D\Omega_D}{2\Lambda_D}(C - C_e)\ . \tag{19}$$

The four expressions derived here are summarized in Supplementary Table 2.

Factors that favor faster growth by dimers, a minority solution component.

Experimental evidence, discussed in the main text, indicates that OZPN building blocks reach the steps from the state of adsorption on the terraces. According to a classical model [74], if transport of solute from the solution bulk toward the crystal surface is not a limiting factor for step motion [79-81]

$$v = \frac{\lambda}{h}\frac{\Omega D}{\Lambda_{ads}}(C - C_e)\left(\frac{\Lambda_s}{\lambda} + \frac{1}{2}\coth\frac{l}{2\lambda}\right)^{-1}, \tag{30}$$

where h is step height, l is the separation between adjacent steps, λ is the characteristic length of surface diffusion, D is the bulk diffusivity, Λ_{ads} is a resistance to adsorb on the terraces from the solution in units length, Λ_s is the resistance to incorporate into kinks from the surface, also measured as length.

The derivation of Eq. (30) relies on the assumption that the solute molecules persist throughout all segments of their path to the steps. The monomer/dimer dynamics may modify the solute distribution near steps and induce a kinetics law that diverges from Eq. (30). Before such law is developed, we will use Eq. (30) for qualitative deductions about the growth mechanism and stop short of direct quantitative comparisons.

The parameters the kinetic law in Eq. (30) are governed by respective transition state enthalpies [74]

$$\lambda = a\exp\frac{\Delta H^{\neq}_{des} - \Delta H^{\neq}_{SD}}{2k_BT} \tag{31}$$

$$D = D_0\exp\left(-\frac{E_{BD}}{k_BT}\right) \tag{32}$$

$$\Lambda_{ads} = a\exp\frac{\Delta H^{\neq}_{ads} - E_{BD}}{k_BT} \tag{33}$$

$$\Lambda_s = a\exp\frac{\Delta H^{\neq}_{kink}}{k_BT} \tag{34}$$

where a is the characteristic surface length, $\Delta H^{\neq}_{ads}$ is the kinetic barrier for adsorption on a terrace, $\Delta H^{\neq}_{des}$ is the kinetic barrier for desorption, $\Delta H^{\neq}_{SD}$ is a kinetic barrier for two-dimensional diffusion

along the surface toward a step, $\Delta H^{\neq}_{kink}$ is a kinetic barrier for attachment of a growth unit to a growth site, and E_{BD} is the effective energy barrier of bulk diffusion.

To assess how the activation enthalpies of the constituent processes contribute to the total $\Delta H^{\neq}$, we consider two limiting cases of Eq. (30). If $l \gg 2\lambda$, then $\coth(l/2\lambda) \cong 1$, and when $l \ll 2\lambda$, $\coth(l/2\lambda) \cong 2\lambda/l$. Denoting for brevity $(C - C_e)/C_e$ as σ

$$\frac{\sigma}{v} = \frac{h\Lambda_{ads}}{\lambda\Omega C_e D}\left[\frac{\Lambda_s}{\lambda} + \frac{1}{2}\right], \text{ for } l \gg 2\lambda \tag{35}$$

and

$$\frac{\sigma}{v} = \frac{h\Lambda_{ads}\Lambda_s}{\lambda^2\Omega C_e D} + \frac{\Lambda_{ads}}{\Omega C_e D}\frac{h}{l}, \quad \text{for } l \ll 2\lambda. \tag{36}$$

Eq. (35) describes distant steps that do not compete for supply of solute, whereas Eq. (36) depicts the growth of closely spaced steps that deplete each other's supply and, as a result, grow slower. Eqs. (35) and (36) predict sharp transition from a linear to a sublinear correlations between v and $(C - C_e)$ with decreasing l that is particularly conspicuous if the step velocity data are presented in coordinates $[\sigma/v](1/l)$. In these coordinates, the proportionality between v and σ manifests as a constant value of σ/v for large step separations, corresponding to low $1/l$. The competition for supply between adjacent steps that arises at small ls enforces a linear increase of σ/v with increasing $1/l$. Sharp transition between constant and linear $[\sigma/v](1/l)$ correlations have been observed at $l = 2\lambda$ with several solution growth systems [78-82]. The correlation between the step velocity v and the step-step separation l for steps on the (002) OZPN crystal face likewise complies the prediction of Eqs. (35) and (36) (Extended Data Figure 5). This correlation suggests that the transition from non-interacting to competing steps occurs at $l \cong 250$ nm. Based on this critical parameter, the characteristic surface diffusion length $\lambda \cong 125$ nm.

The step velocities comprising the quadratic correlation between v and C (Figure 1e) were measured only with steps separated by distances greater than 250 nm (= 2λ) (Extended data Figure 6), which ensures that the steps do not compete for supply and the step kinetics obeys Eq.

(35). Combining Eqs. (31) – (35) and assuming strong resistance for molecular incorporation into kinks, i.e., $\Lambda_s/\lambda > 1/2$, as with KDP [79,80] and calcite [81], then

$$\left(\frac{D}{\Lambda}\right)_{eff} = \frac{D}{\Lambda_{ads}}\frac{\lambda^2}{h\Lambda_s}$$

and
$$\Delta H^{\neq} \cong \Delta H^{\neq}_{ads} - \Delta H^{\neq}_{des} + \Delta H^{\neq}_{SD} + \Delta H^{\neq}_{kink} . \quad (37)$$

The enthalpy barriers encountered by a molecule reaching a step *via* adsorption on the terraces are schematically depicted in Figure 5e.

Lastly, since

$$\Delta H^{o}_{ads} = \Delta H^{\neq}_{ads} - \Delta H^{\neq}_{des} , \quad (38)$$

then
$$\Delta H^{\neq} \cong \Delta H^{0}_{ads} + \Delta H^{\neq}_{SD} + \Delta H^{\neq}_{kink} . \quad (39)$$

where ΔH^{0}_{ads} is the equilibrium adsorption enthalpy. Thus, the main components of the activation barrier for growth are $\Delta H^{\neq}_{kink}$ and ΔH^{0}_{ads}. Distinct kinetics of growth by monomers or dimers should correlate with dissimilar values of these two parameters.

The main determinant of $\Delta H^{\neq}_{kink}$ are solvent molecules liganded to the solute, which shed upon solute association to the step [77,83-89]. Analysis in the main text suggests that bound waters contribute equally to the kinetic barriers for dimer and monomer incorporation since the strongly bound waters localize on the polar outward surface of the OZPN molecule [42] (Figure 5f). Thus, $\Delta H^{\neq}_{kink}$ for the dimer is likely similar to that of the monomer. Then, the faster kinetics of growth by dimers should be rooted in the strong adsorption of the monomers or dimers on the crystal surface. Strong dimer adsorption enforces lower ΔH^{0}_{ads} for this species, diminishes the effective activation barrier for step growth $\Delta H^{\neq}$, and drives faster growth by dimer incorporation. ΔH^{0}_{ads} of the dimers, evaluated by MD simulations is indeed significantly lower than that of the monomers. The contribution to faster step growth of enhanced dimer adsorption is supplemented by additional dimerization at the surface, as discussed in the main text. The two processes concentrate the dimer at the vicinity of the steps and promote dimer incorporation into the kinks.

References

38 Wawrzycka-Gorczyca, I., Borowski, P., Osypiuk-Tomasik, J., Mazur, L. & Koziol, A. Crystal structure of olanzapine and its solvates. Part 3. Two and three-component solvates with water, ethanol, butan-2-ol and dichloromethane. *Journal of Molecular Structure* **830**, 188-197, (2007).

39 Reutzel-Edens, S. M., Bush, J. K., Magee, P. A., Stephenson, G. A. & Byrn, S. R. Anhydrates and Hydrates of Olanzapine: Crystallization, Solid-State Characterization, and Structural Relationships. *Crystal Growth & Design* **3**, 897-907, (2003).

40 Thakuria, R. & Nangia, A. Olanzapinium Salts, Isostructural Solvates, and Their Physicochemical Properties. *Crystal Growth & Design* **13**, 3672-3680, (2013).

41 Warzecha, M., Safari, M. S., Florence, A. J. & Vekilov, P. G. Mesoscopic Solute-Rich Clusters in Olanzapine Solutions. *Crystal Growth & Design* **17**, 6668-6676, (2017).

42 Warzecha, M. *et al.* Direct Observation of Templated Two-Step Nucleation Mechanism during Olanzapine Hydrate Formation. *Crystal Growth & Design* **17**, 6382-6393, (2017).

43 Klamt, A. & Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions 2*, 799-805, (1993).

44 Qu, H. *et al.* Raman and ATR FTIR spectroscopy in reactive crystallization: Simultaneous monitoring of solute concentration and polymorphic state of the crystals. *Journal of Crystal Growth* **311**, 3466-3475, (2009).

45 Hu, Y., Liang, J. K., Myerson, A. S. & Taylor, L. S. Crystallization Monitoring by Raman Spectroscopy: Simultaneous Measurement of Desupersaturation Profile and Polymorphic Form in Flufenamic Acid Systems. *Industrial & Engineering Chemistry Research* **44**, 1233-1240, (2005).

46 Simone, E., Saleemi, A. N. & Nagy, Z. K. Application of quantitative Raman spectroscopy for the monitoring of polymorphic transformation in crystallization processes using a good calibration practice procedure. *Chemical Engineering Research and Design* **92**, 594-611, (2014).

47 Tanaka, N., Kitano, H. & Ise, N. Raman spectroscopic study of hydrogen bonding in aqueous carboxylic acid solutions. *The Journal of Physical Chemistry* **94**, 6290-6292, (1990).

48 Ayala, A. P. *et al.* Solid state characterization of olanzapine polymorphs using vibrational spectroscopy. *International Journal of Pharmaceutics* **326**, 69-79, (2006).

49 Hunter, C. A., McCabe, J. F. & Spitaleri, A. Solvent effects of the structures of prenucleation aggregates of carbamazepine. *CrystEngComm* **14**, 7115-7117, (2012).

50 Spitaleri, A., Hunter, C. A., McCabe, J. F., Packer, M. J. & Cockroft, S. L. A 1H NMR study of crystal nucleation in solution. *CrystEngComm* **6**, 490-493, (2004).

51 Abraham, M. J., Murtola, Teemu, Schulz, Roland, Páll, Szilárd, Smith, Jeremy C, Hess, Berk, Lindahl, Erik. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25, (2015).

52 Horn, H. W., Swope, William C, Pitera, Jed W, Madura, Jeffry D, Dick, Thomas J, Hura, Greg L, Head-Gordon, Teresa. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics* **120**, 24503, (2004).

53 Dodda, L. S., Cabeza de Vaca, I., Tirado-Rives, J. & Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res* **45**, W331-W336, (2017).

54 Dodda, L. S., Vilseck, J. Z., Tirado-Rives, J. & Jorgensen, W. L. 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *J Phys Chem B* **121**, 3864-3870, (2017).

55 Essmann, U., Perera, Lalith, Berkowitz, Max L. & Darden, T., Lee, Hsing, Pedersen, Lee G. A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577-8593, (1995).

56 Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* **13**, 952-962, (1992).

57 Hess, B., Berendsen, H. J. C., Fraaije, J. G. E. M. & Bekker, H. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463−1472, (1997).

58 Bussi, G., Donadio, Davide, Parrinello, Michele. Canonical sampling through velocity rescaling. *Journal of Chemical Physics* **126**, (2007).

59 Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Strain fluctuations and elastic constants The Journal of Chemical Physics* **52**, 234505, (1981).

60 Guevara-Carrion, G., Vrabec, J. & Hasse, H. Prediction of self-diffusion coefficient and shear viscosity of water and its binary mixtures with methanol and ethanol by molecular simulation. *The Journal of Chemical Physics* **134**, 74508, (2011).

61 Hess, B., Van Der Vegt, Nico F A. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. (2006).

62 Caleman, C., Van Maaren, Paul J, Hong, Minyan, Hub, Jochen S, Costa, Luciano T, and Van Der Spoel, David. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Com-pressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput* **8**, 61-74, (2012).

63 Kästner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 932-942, (2011).

64 Torrie, G. M. & Valleau, J. P. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chemical Physics Letters* **28**, 578-581, (1974).

65 Torrie, G. M. & Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* **23**, 187-199, (1977).

66 Bonomi, M., Branduardi, Davide, Bussi, Giovanni, Camilloni, Carlo, Provasi, Davide, Raiteri, Paolo, Donadio, Davide, Marinelli, Fabrizio, Pietrucci, Fabio, Broglia, Ricardo A, and Parrinello, Michele. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **180**, 1961-1972, (2009).

67 Ferguson, A. L. BayesWHAM: A Bayesian approach for free energy estimation, reweighting, and uncertainty quantification in the weighted histogram analysis method. *Journal of Computational Chemistry* **38**, 1583-1605, (2017).

68 Kumar, S., Bouzida, Djamal, Swendsen, Robert H., Kollman, Peter A. , and Rosenberg, John M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. TheMethod. **13**, 1011-1021, (1992).

69 Fiorin, G., Klein, Michael L, Hénin, Jérôme. Molecular Physics An International Journal at the Interface Between Chemistry and Physics Using collective variables to drive molecular dynamics simulations Using collective variables to drive molecular dynamics simulations. *Molecular Physics* **111**, 3345-3362, (2013).

70 De Jong, D. H., Schafer, L V, De Vries, A H, Marrink, S J, Berendsen, H J, Grubmuller, H. Determining equilibrium constants for dimerization reactions from molecular dynamics simulations. *Journal of Computational Chemistry* **32**, 1919-1928, (2011).

71 Roy, A., Hua, D. P., Ward, J. M. & Post, C. B. Relative Binding Enthalpies from Molecular Dynamics Simulations Using a Direct Method Terms of Use. *J. Chem. Theory Comput* **10**, 2759-2768, (2014).

72 Burton, W. K., Cabrera, N. &. Frank, F.C. The growth of crystals and equilibrium structure of their surfaces. *Phil. Trans. Roy. Soc. London Ser. A* **243**, 299- 360, (1951).

73 Chernov, A. A. The spiral growth of crystals. *Sov. Phys. Uspekhi* **4**, 116-148, (1961).

74 Gilmer, G. H., Ghez, R. & Cabrera, N. An analysis of combined volume and surface diffusion processes in crystal growth. *J. Crystal Growth* **8**, 79-93, (1971).

75 Chernov, A. A. & Komatsu, H. Topics in Crystal Growth Kinetics. in *Science and Technology of Crystal Growth* (eds J.P. van der Eerden & O.S.L Bruinsma) 67-80 (Kluwer Academic, 1995).

76 De Yoreo, J. J. & Vekilov, P. G. Principles of crystal nucleation and growth. in *Biomineralization* Vol. 54 *Reviews in Mineralogy & Geochemistry* 57-93 (2003).

77 Vekilov, P. G. What determines the rate of growth of crystals from solution? . *Crystal Growth and Design,* **7**, 2796-2810, (2007).

78 Olafson, K. N., Ketchum, M. A., Rimer, J. D. & Vekilov, P. G. Mechanisms of hematin crystallization and inhibition by the antimalarial drug chloroquine. *Proceedings of the National Academy of Sciences* **112**, 4946-4951, (2015).

79 Vekilov, P. G., Kuznetsov, Y. G. & Chernov, A. A. Interstep interaction in solution growth; (101) ADP face. *J. Crystal Growth* **121**, 643-655, (1992).

80 Vekilov, P. G., Kuznetsov, Y. G. & Chernov., A. A. The effect of temperature on step motion; (101) ADP face. *J. Crystal Growth* **121**, 44-52, (1992).

81 Maruyama, M., Tsukamoto, K., Sazaki, G., Nishimura, Y. & Vekilov, P. G. Chiral and Achiral Mechanisms of Regulation of Calcite Crystallization. *Crystal Growth & Design* **9**, 127-135, (2009).

82 Olafson, K. N., Ketchum, M. A., Rimer, J. D. & Vekilov, P. G. Molecular Mechanisms of Hematin Crystallization from Organic Solvent. *Crystal Growth & Design* **15**, 5535-5542, (2015).

83 Petsev, D. N., Chen, K., Gliko, O. & Vekilov, P. G. Diffusion-limited kinetics of the solution-solid phase transition of molecular substances. *Proc. Natl. Acad. Sci. USA* **100**, 792-796, (2003).

84 Vekilov, P. G. Incorporation at kinks: kink density and activation barriers, . in *Perspectives on Inorganic, Organic and Biological Crystal Growth: From Fundamentals to Applications: AIP Conference Proceedings* Vol. 916 *AIP Conference Series* (eds M. Skowronski, J. J. DeYoreo, & C.A. Wang) 235-267 (AIP, 2007).
85 Elhadj, S., De Yoreo, J. J., Hoyer, J. R. & Dove, P. M. Role of molecular charge and hydrophilicity in regulating the kinetics of crystal growth. *PNAS* **103**, 19237-19242, (2006).
86 Elhadj, S. *et al.* Peptide Controls on Calcite Mineralization: Polyaspartate Chain Length Affects Growth Kinetics and Acts as a Stereochemical Switch on Morphology. *Cryst. Growth Des.* **6**, 197-201, (2006).
87 Dandekar, P., Kuvadia, Z. B. & Doherty, M. F. Engineering Crystal Morphology. *Annual Review of Materials Research* **43**, 359-386, (2013).
88 Teng, H. H. How ions and molecules organize to form crystals. *Elements* **9**, 189-194, (2013).
89 Shtukenberg, A. G., Ward, M. D. & Kahr, B. Crystal growth with macromolecular additives. *Chemical reviews* **117**, 14042-14090, (2017).

Data availability

The datasets generated during and/or analyzed during the current study and the simulation and analysis codes used for these calculations are available from the corresponding authors upon reasonable request.

Code availability

No custom made computer code was used in this work.