Full Paper

Israel Journal of Chemistry

doi.org/10.1002/ijch.202100081

How to Identify the Crystal Growth Unit

Lakshmanji Verma[+],[a] Monika Warzecha[+],[b] Rajshree Chakrabarti[+],[a] Viktor G. Hadjiev,[c] Jeremy C. Palmer,[a] and Peter G. Vekilov*[a, d]

Abstract: The structure and composition of the crystal growth unit are of huge fundamental and practical consequence. We propose a method to identify the solute species that incorporates into the growth site on crystal surfaces, the kinks, which rests on the kinetics of the elementary reaction at the kinks. We use as model crystals olanzapine, an antipsychotic medication, and etioporphyrin I, a field-effect transistor. We combine time-resolved *in situ* atomic force microscopy with Raman and absorption spectroscopies, complemented by density functional theory and all-atom molecular dynamics modeling of the solutions. We show that the structure of the growth unit cannot be deduced neither from the solute oligomers nor from the crystal structure. Chemical kinetics analyses reveal that if the dominant solute species is the one that incorporates into the crystal growth sites, then the kinetics of layer growth complies with a monomolecular rate law. By contrast, if the crystal growth unit assembles from two units of the dominant solute form, a bimolecular rate law ensues. Solutions of both olanzapine and etioporphyrin I are dominated by solute monomers, which exist in equilibrium with a minority of dimers. Whereas numerous olanzapine crystal structures incorporate dimer motifs, etioporphyrin I crystals organize as stacks of monomers. Olanzapine crystal grow by incorporation of dimers. One of the studied face of etioporphyrin I grows by incorporation of the majority monomers, whereas the other one selects the minority dimers as a growth unit. The results highlight the power of the crystallization kinetics analyses to identify the growth unit and illuminate one of the most challenging issues of crystal growth.

Introduction

The interactions between the solute molecules in solution may lead to self-assembly into dimers, trimers, or higher oligomers.[1] The assembled solute oligomers may, in some cases, mirror the structural units of the crystal.[2] Examples include self-associates of isonicotinamide,[3] sulfamerizine,[4] and tetrolic acid,[2c,d,5] which appear identical to respective crystal structural elements. The correlation between structural motifs in the solution and in the crystal was construed to affirm that the common structural elements represent the solute species that associates to the crystal growth sites, the kinks. The identity of the unit by which crystals grow is of singular fundamental significance. The composition and structure of the species that incorporates into the kinks is also of substantial practical importance. It constitutes the central assumption of molecular and coarse-grained models that aim to predict crystal growth rates and crystal habits.[6] It also presents a target for modification by solvents and additives that are applied to select alternative crystal polymorphs and forms.[5,7] The elementary acts of incorporation into kinks, however, have only been directly visualized for relatively few crystals, mostly of proteins with large molecules.[8] Given these constrains to directly identify the growth unit, the promotion of the solute oligomers to the rank of the incorporating species has remained largely unchallenged.

The complexity of deducing the growth unit from crystal structural data is illustrated by the extensively studied crystallization of glycine. Analyses of diffusion coefficients in aqueous glycine solutions implied the presence of hydrogen-bonded cyclic dimers, a structural unit of the α-glycine crystal lattice, which were then invoked as building blocks for crystal growth.[9] Measurement of the solution freezing point depression, however, revealed that in aqueous environments glycine mostly existed as monomers.[10] Molecular dynamic (MD) simulations informed that only 15% of glycine molecules were held within any type of dimers and that open chain dimers, bound by single hydrogen bonds, were more stable than the cyclic dimers and dominated the small dimer population.[11] These findings belie the conjecture that the preference for α-glycine over other polymorphic forms is dictated by the abundance of cyclic dimers in aqueous

[a] *L. Verma,[+] R. Chakrabarti,[+] J. C. Palmer, P. G. Vekilov*
Department of Chemical and Biomolecular Engineering,
University of Houston,
4726 Calhoun Rd., Houston, TX 77204-4004, USA.
[b] *M. Warzecha[+]*
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.
[c] *V. G. Hadjiev*
Texas Center for Superconductivity,
University of Houston,
3369 Cullen Blvd., Suite 202, Houston, Texas 77004-50024, USA.
[d] *P. G. Vekilov*
Department of Chemistry,
University of Houston,
3585 Cullen Blvd., Houston, TX 77204-5003, USA
E-mail: vekilov@uh.edu
[+] *These three authors contributed equally to this paper.*

solutions, which then assemble into the crystal and impose its structure.[11]

The thesis that solute oligomers define the growth units and correlate with the crystal structural blocks is further refuted by the mismatch between solute state and the crystal structure found with benzoic acid,[12] mandelic acid,[2c] and inosine dihydrate.[7c] Solutions of these compounds do not exhibit the dimers that are present in the crystal structures, certifying the disconnection between the solution species and crystal structural elements.

A missing element in the analyses of crystallization by incorporation of complex solution species is direct identification of the growth unit, whether it be monomer, dimer, or higher oligomer. The challenges arise from the combination of nanoscopic length and time scales that characterize incorporation into the growth sites, the kinks, on the crystal surfaces. Whereas current state-of-the-art *in situ* electron microscopy and scanning probe techniques can readily detect the presence of single molecules as small 1 nm in the crystal lattice,[13] the limited temporal resolution of the observations prevents them to discriminate between the incorporation of a dimer and the fast sequential addition of two monomers.

Here we adopt the rate of growth of layers on the crystal surfaces as an indicator of the correlation between the dominant solution species and the crystal building block. We explore the correlation of the growth unit, identified from the kinetics of incorporation into kinks, with both the state of the solute and the crystal symmetry on the examples of two molecular crystals, olanzapine (OZPN) and etioporphyrin I (EtpI). We show that in both cases the dominant solute species may not match the respective crystal structural units. The growth kinetics data identify the growth species for OZPN and one of the EtpI faces as a molecular dimer, which, however, is a minority component in solutions of both compounds. Another of the EtpI faces grows by incorporation of the majority monomers. Whereas dimers motifs abound in OZPN crystal structures, EtpI crystals are built of monomers. The results demonstrate that the crystal growth units may diverge from both the dominant solute species and the crystal structural units and highlight the power of the proposed method to identify the growth unit.

Results and Discussion

OZPN: Growth by Incorporation of Dimers that Capture a Minor Fraction of the Solute

Olanzapine (OZPN) (Figure 1a) is an antipsychotic drug,[14] which precipitates in more than 60 individual crystal forms.[15] In all of them, but one,[16] OZPN molecules are arranged in a centrosymmetric dimer SC_0, comprised of two conformational enantiomers (Figure 1c).[15] The dominance of dimeric structures has incited the conjecture that the dimers preform in the solution where they capture the majority of the solute.[6c,17]

To test whether OZPN crystal grow by incorporation of dimers, we monitor the growth of $2OZPN{\cdot}EtOH{\cdot}2H_2O$ crystals from a 1/1 (v/v) ethanol/water mixture using time-resolved *in situ* AFM.[17] OZPN crystal expose large {002} faces parallel to the substrate (Figure 1b) to observation by atomic force microscopy (AFM). Time-resolved *in situ* AFM observations reveal that the {002} faces grow by incorporation of solute into steps produced by screw dislocations (Figure 2b).[17] The velocity, v, of evolving steps was determined from the slope of step displacement from a reference point measured as a function of time as OZPN steps grew at steady rates over extended periods. Assuming that steps grow by incorporation of solute monomers[19] implies monomolecular reaction and gives 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. Here 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.[20] Linear $v(C)$ correlations have been observed for numerous solution grown crystals.[8a,c, 21] Unexpectedly, OZPN displays a superlinear

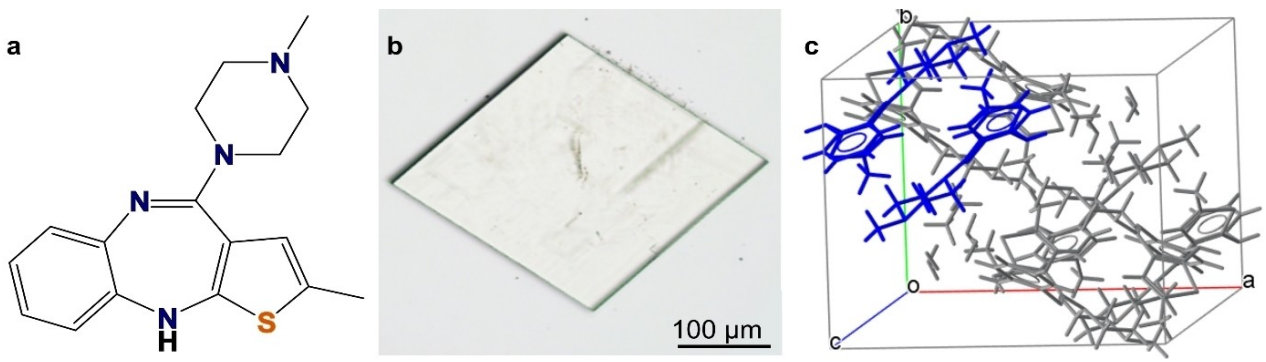

Figure 1. OZPN and its crystals. a. The OZPN molecule. **b**. Optical micrograph of an OZPN crystal; the (002) face faces upwards. **c**. The crystal structure of the dihydrate ethanoate mixed solvate 2OZPN·EtOH·2H2O in space group $P2_1/c$ (Cambridge Structural Database REFCODE WEXQEW [18]). One centrosymmetric OZPN SC_0 dimer is highlighted in blue.

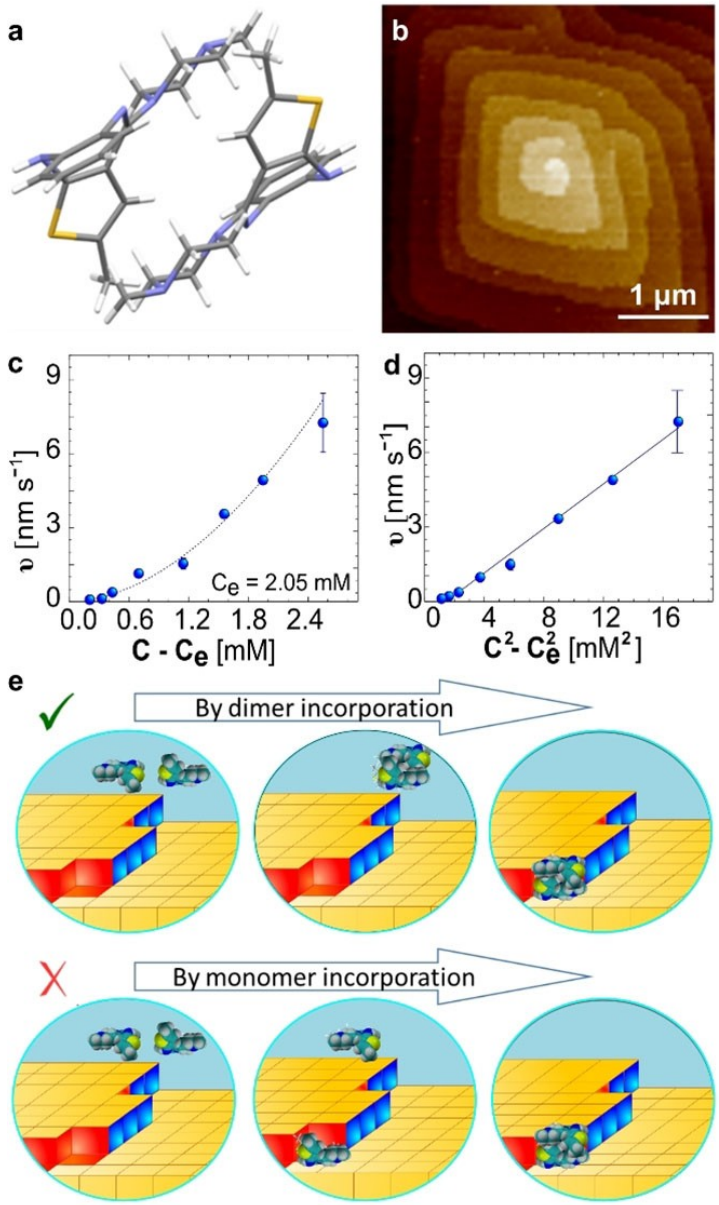

Figure 2. The growth of OZPN crystals. a. The structure of the SC_0 OZPN dimer. **b**. *In situ* AFM image of the surface of a (002) face of an OZPN crystal at $C = 3.87$ mM. New crystal layers are generated by a screw dislocation. **c**. The velocity v of steps in the [110] direction as a function of OZPN concentration C in 1/1 (v/v) $EtOH/H_2O$. **d**. The linear correlation between v and $(C^2 - C_e^2)$ has $R = 0.95$. Dotted lines in **c** and **d** 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 **c** and **d** indicate the standard deviation of v determined as the slope of the displacement-time correlations. Data in c and d are from ref.[13e] **e**. Schematic of two alternative growth mechanisms. Upper path: two OZPN monomers form a dimer in the solution, which incorporates into the crystal as a whole. Lower path: the two monomers incorporate sequentially, forming a SC_0 dimer in the crystal.

$v(C)$ dependence, which extends to concentrations more than twice the solubility C_e (Figure 2c, d).

We eliminated four probable scenarios of apparent growth acceleration at high supersaturation:[13e] mesoscopic OZPN-rich clusters[17,22] which provide additional OZPN molecules to the steps; inaccurate solubility, increasing kink density at higher supersaturation[8c,19, 23] and step pinning by impurities.[21a]

We propose that the superlinear $v(C)$ is a consequence of crystal growth by incorporation of dimers that exist in equilibrium with OZPN monomers in the growth solution (Fig.2e). Elevated OZPN concentrations shift the dimerization equilibrium towards dimers and nonlinearly enhance the dimer

concentration. Towards an analytical relation between the total solute concentration and the step velocity, which accounts for partial or total dimerization of the solute, we consider the equilibrium between two monomers M and a dimer D

$$2M \rightleftarrows D \text{ , for which } K_D = C_D/C_M^2, \tag{1}$$

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

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

We obtain for C_M and C_D

$$C_M = \frac{\sqrt{1+8K_DC}-1}{4K_D}, \tag{3}$$

and

$$C_D = K_DC_M^2 = \frac{1}{16K_D}\left[\sqrt{1+8K_DC}-1\right]^2. \tag{4}$$

The step velocity v scales with the difference between opposing fluxes: the flux of solute molecules into the kink site j_+ and that of the solute molecules from the kink site j_-

$$v = \Omega(j_+ - j_-) \tag{5}$$

where Ω is the molecular volume in the crystal. In turn, the flux j_+[19,24]

$$j_+ = \frac{D_i}{\Lambda_i}C_i \tag{6}$$

where D_i is the diffusion coefficient, Λ_i, resistance to enter the step, and i designates either monomer or dimer. At equilibrium, $v = 0$, and

$$j_- = j_+ = \frac{D_i}{\Lambda_i}C_{ie},$$

where C_{ie} is the concentration of the respective species at equilibrium. Because j_- is independent of C_i, we obtain

$$v = \frac{D_i\Omega_i}{\Lambda_i}(C_i - C_{ie}). \tag{7}$$

We derive expressions for the step velocity v as a function of total OZPN concentration C for four scenarios that involve dimers present in the growth solution:

(i) Monomers dominate in the solution and growth occurs by the attachment of monomers. The latter assumption transforms Eq. (7) into

$$v = \frac{D_M\Omega_M}{\Lambda_M}(C_M - C_{Me}) \tag{8}$$

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 $K_D \ll 1/C_M$, $K_DC_M \ll 1$ and $8K_DC_M < 1$. With this, Eq. (3) transforms to

$$C_M = \frac{1}{8K_D}8K_DC = C \tag{9}$$

A linear relationship is expected between the step velocity, v, and the total OZPN concentration

$$v = \frac{D_M\Omega_M}{\Lambda_M}(C - C_e). \tag{10}$$

(ii) Monomers dominate in the solution, although growth occurs by attachment dimers. In this case, $C_D \ll C_M$, $K_D \ll 1/C_M$, and $8K_DC_M < 1$. With this, Eq. (4) becomes

$$C_D = \frac{1}{16K_D}\frac{64K_D^2C^2}{4} = K_DC^2 \tag{11}$$

We obtain for j_+, j_-, and v

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

and

$$v = \frac{D_D\Omega_D}{\Lambda_D}K_D\left(C^2 - C_e^2\right) \tag{12}$$

Eq. (12) implies that the correlation between the step velocity and the total solute concentration is quadratic.

(iii) If dimers dominate in solution but the growth occurs by attachment of monomers, $C_D \gg C_M$ 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. (7),

$$v = \frac{D_M\Omega_M}{\Lambda_M}\sqrt{\frac{1}{2K_D}}\left(\sqrt{C} - \sqrt{C_e}\right) \tag{13}$$

Eq. (13) corresponds to a sublinear dependence between the step velocity and the total solute concentration.

(iv) If dimers dominate in the solution and associate to the steps, $C_D \gg C_M$ and $K_D \gg 1/C_M$, $8K_DC > 8K_DC_M > 1$ and $\sqrt{1+8K_DC} > 1$. We obtain

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

And

$$v = \frac{D_D \Omega_D}{2\Lambda_D}(C - C_e). \quad (14)$$

The presented kinetic scheme reveals that if the dominant solute species, whether it be monomer of dimer, is the one that incorporates in the kinks, the kinetics of layer growth will be linear, Eqs. (10) and (14). If the growth units derive from the decay of larger assemblies, sublinear rate law, as in Eq. (13), ensues. If the 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 $v = \beta_D K_D \Omega_D (C^2 - C_e^2)$, Eq. (12), where $\beta_D = D_D/\Lambda_D$, the subscript D denotes dimer and K_D is the dimerization equilibrium constant. The $v(C)$ data for OZPN are consistent with this latter functional relation (Figure 2c, d).

To assess the formation of OZPN dimers in solution we employ Raman spectroscopy. We compare OZPN spectra at a low concentration, 0.005 M, to spectra at relatively high concentrations, up to 0.043 M, and to spectra of the corresponding solid form, 2OZPN·EtOH·2H_2O, in which OZPN is arranged as SC_0 dimers (Figure 3a). To assign the origins of the observed peaks, we model spectra for an OZPN monomer and an OZPN SC_0 dimer using density functional theory (DFT) (Figure 3a). The model spectra for the OZPN monomer show multiple Raman peaks between 1200 and 1500 cm^{-1} (Figure 3a), whereas the model spectra for the dimer show strong peaks around 1000 cm^{-1} and in the range 1500–1600 cm^{-1}.[13e]

Raman spectra of OZPN water ethanoate solvate crystals show prominent dimer peaks comparable to spectra obtained for solution samples at high concentrations (Figure 3a). The low-concentration solution spectra reveal strong monomer peaks and lack of dimer peaks. Raman spectra at increasing intermediate concentrations in both solvents display gradual contraction of the monomer peaks and an increase in intensity of the dimer peaks (Figure 3a). The concentration dependence of the monomer and dimer Raman peak intensities enable estimation of the OZPN dimerization constant K_D. The Raman intensity directly relates to the concentration of the species generating the signal

$$I = JC \quad (15)$$

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

The band at 1517 cm^{-1} relates with the dimer and it preserves its shape at all OZPN concentrations (Figure 3a). The intensity of this peak at the highest tested C was selected as an internal intensity standard I_{st}. We test several values of K_D and calculate the corresponding C_D for each of the

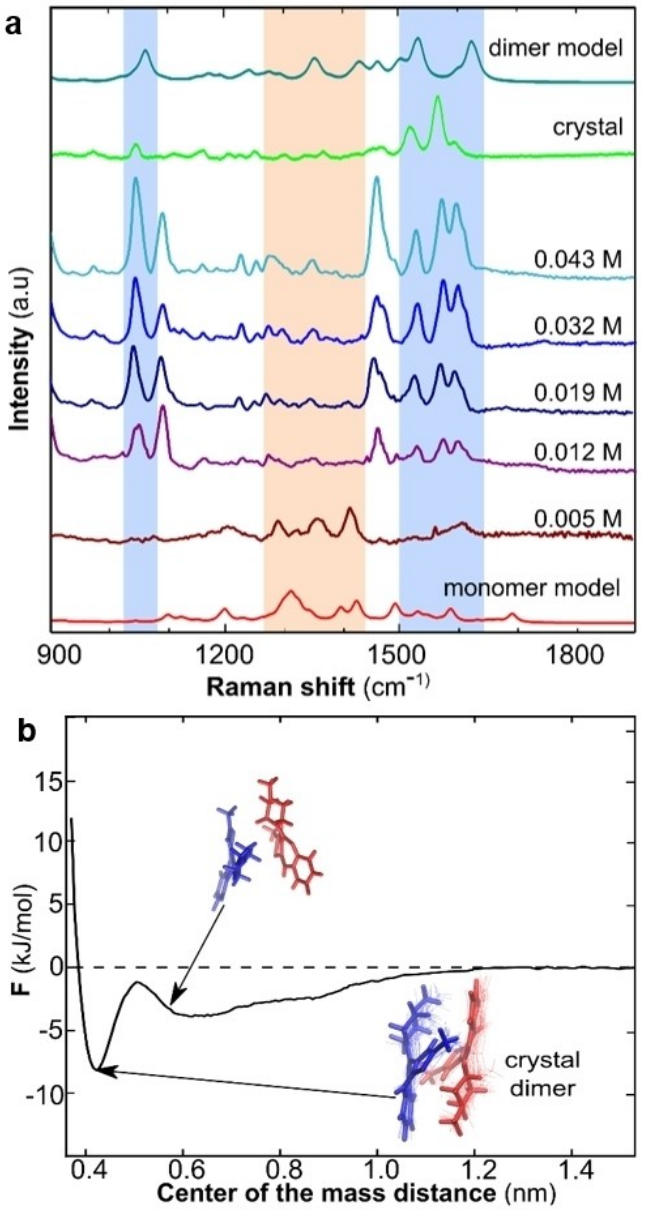

Figure 3. OZPN dimers in solution. a. Raman spectra of OZPN dissolved in 1/1 (v/v) EtOH/H_2O at listed concentrations and of the solid crystalline solvate are compared to spectra for OZPN monomer and dimer calculated using DFT. **b**. The potential of mean force F between two OZPN monomers in 1/1 (v/v) EtOH/H_2O computed using all-atom molecular dynamics. Insets: Lower right, representative snapshots of the configuration of the dimer (the constituent monomers are shown in red and blue, respectively) occupying the deepest F minimum which is nearly identical to the SC_0 dimer found in the crystal structure (one of the ten conformations is highlighted for clarity). Top left: a representative configuration of the two monomers at separations longer than the deep minima. Some of the spectra in a and the mean force potential in b are from ref.[13e]

measured 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 as

$$I_D = \frac{C_D I_{st}}{C_{Dst}} \tag{16}$$

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. The K_D that yielded the lowest RMSD was taken as the best estimate of the dimerization constant. We obtain $K_D = 2.7 \pm 0.1\ M^{-1}$. Mass balance calculations inform that the dimer concentration C_D in the growth solution 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 2c). The superlinear increase of C_D dictates the quadratic $v(C)$ correlation (Figure 2c, d).

All-atom classical molecular dynamics (MD) evaluation of the potentials of mean force F between two OZPN monomers in EtOH/H_2O reveal a relatively deep minimum at center-of-mass separations of 0.43 nm (Figure 3b). The configurations of the two monomers occupying the minimum fluctuate around the structure of the SC_0 dimer observed in OZPN crystal forms (Figure 3b, inset). At larger separations, the SC_0 dimer falls apart, however, F exhibits a shallow secondary minimum that corresponds to loose dimers with variable configuration (Figure 3b). We evaluate the free energy of OZPN dimerization in EtOH/H_2O from the potential value at the deep minimum and obtain –8.1 $kJ\,mol^{-1}$. The dimerization constant K_D was evaluated as the ratio of integrals of F over the closest range minimum and the unbound state and is $2.6 \pm 0.8\ M^{-1}$. 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. Further MD simulations 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.[13e]

The results with OZPN 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. Importantly, observations with OZPN highlight that the solute unit, by which crystals grow, can be identified by correlating solute oligomerization to the solute incorporation rate law. Furthermore, the kinetic, structural, and spectroscopic analyses of OZPN solutions and the growth of OZPN crystals emphasize the disparity between the majority species in the solution and both the crystal structure and the structure of the incorporating solute unit.

EtpI: Distinct Growth Mechanisms of the Anisotropic Crystal Faces

Etioporphyrin I (EtpI, Figure 4a) represents a class of compounds whose crystals (Figure 4b) carry promising optical and electronic properties for use as semiconductors, solar cells, and field-effect transistors.[25] In contrast to OZPN, EtpI has a unique unsolvated triclinic crystal structure built of single molecules stacked in parallel columns (Figure 4c).[25a] The only entry for etioporphyrin I in the Cambridge Structural Database (CSD) is for the non-centrosymmetric P1 crystallographic symmetry group (Figure 4c).

Owing to thier approximately square cross-section, EtpI crystals expose two faces to AFM observation: (010) and (001). The presence of a (101) face in the crystal habit enforces distinct shapes, parallelogram for the (010) face and trapeze for the (001) face. EtpI crystals readily grow in 1-octanol by spreading of layers generated by screw dislocations, observed by *in situ* AFM (Figure 4 a, b) and similar to OZPN.

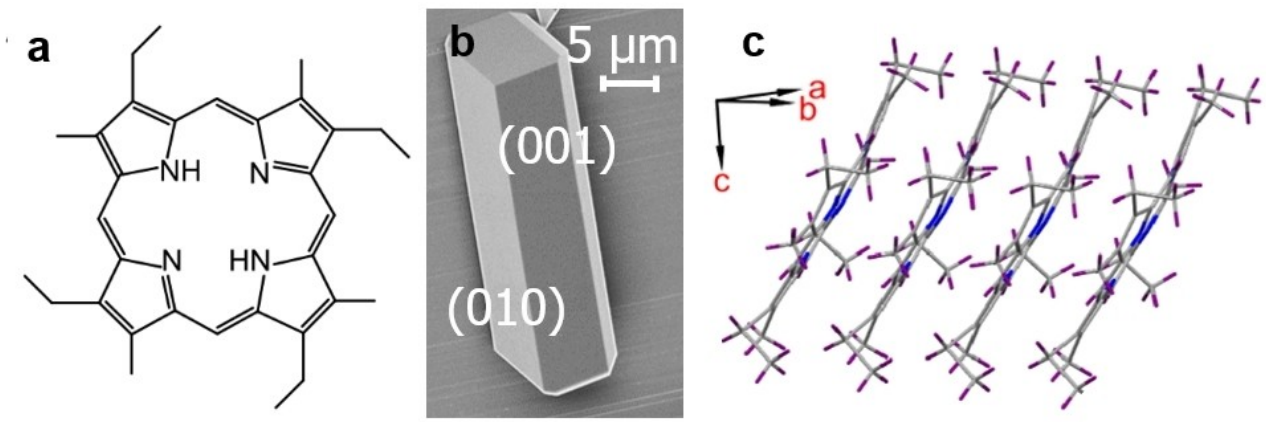

Figure 4. EtpI and its crystals. a, The EtpI molecule. **b**, Scanning electron micrograph of an EtpI crystal. The (010) and (001) faces are labeled. **c**, Molecular arrangement of EtpI in the crystal. Carbon is shown in grey, nitrogen in blue, and hydrogen in purple.

Time resolved AFM measurements of the velocity of steps v reveal that on the (010) face, v scales linearly with the solute concentration C, whereas on the (001) face v is a quadratic function of C (Figure 4c, d).

Eliminating the trivial reasons that may drive a superlinear $v(C)$ dependence (inaccurate solubility value, paucity of kinks, and step inhibition by uncontrolled foreign substances) the two distinct correlations manifest monomolecular kinetics of solute incorporation on the (010) face and bimolecular reaction on the (001) face. A feasible mechanism that guides the divergent kinetics laws on the two crystal faces relies on monomers that dominate the solute speciation. Step growth on the (010) face, which exhibits monomolecular incorporation rate law, selects the monomers as incorporating species since this incorporation mode motivates faster step growth than incorporation of potential oligomers. By contrast, the bimolecular kinetics of step growth on the (010) face manifests two monomers combining into dimers, which remain a minority solute species, but still propel growth by dimers of the distinct steps on the (010) face faster than by incorporation of the majority solute monomers.

Characterization of the EtpI oligomerization in octanol solutions by vibrational spectroscopy – the concentration of EtpI in octanol, controlled by the low solubility $C_e =$ 0.17 mM, is below the sensitivity of Raman detection on the background of strong octanol spectra – supports this mechanism. We measured the absorption spectra of EtpI in octanol with a standard UV-Vis spectrophotometer. The recorded absorbance spectra at low and high concentrations were normalized to highlight the change in relative intensity of the absorbance bands (Figure 6a). In the range from 450 to 625 nm the two spectra are similar to published EtpI spectra, where the observed absorbance bands are referred to as Q bands.[26]

We compare the two measured spectra with model spectra computed using a time-dependent density functional theory package (TDDFT) for two EtpI species: a monomer and a dimer, in which the two monomers position as in the crystal lattice (Figure 4c). The monomer model spectrum reveals two purely electronic transitions (0-0) and reproduces only two of the bands in the experimental spectra, at 622 nm (Qx) and 528 nm (Qy) (Figure 6a). The two other relatively broad and structured bands at 566 and 475 nm have evoked conflicting interpretations.[26] The most consensual one is that these are vibronic absorption bands (0-1). The (0-1) bands calculated with full-fledged TDDFT (Figure 6a) confirm the activation of (0-1) absorption in EtpI, albeit with a strength weaker than that of the (0-0) bands. In contrast to the monomer model spectrum, the intensities of most bands in the computed EtpI dimer spectrum are with comparable intensity and at wavelengths similar to the measured Q bands.

The overlapping absorption bands of monomers and dimers hinder the quantification of EtpI dimers using the 450–652 nm wavelength range. Remarkably, the contributions of

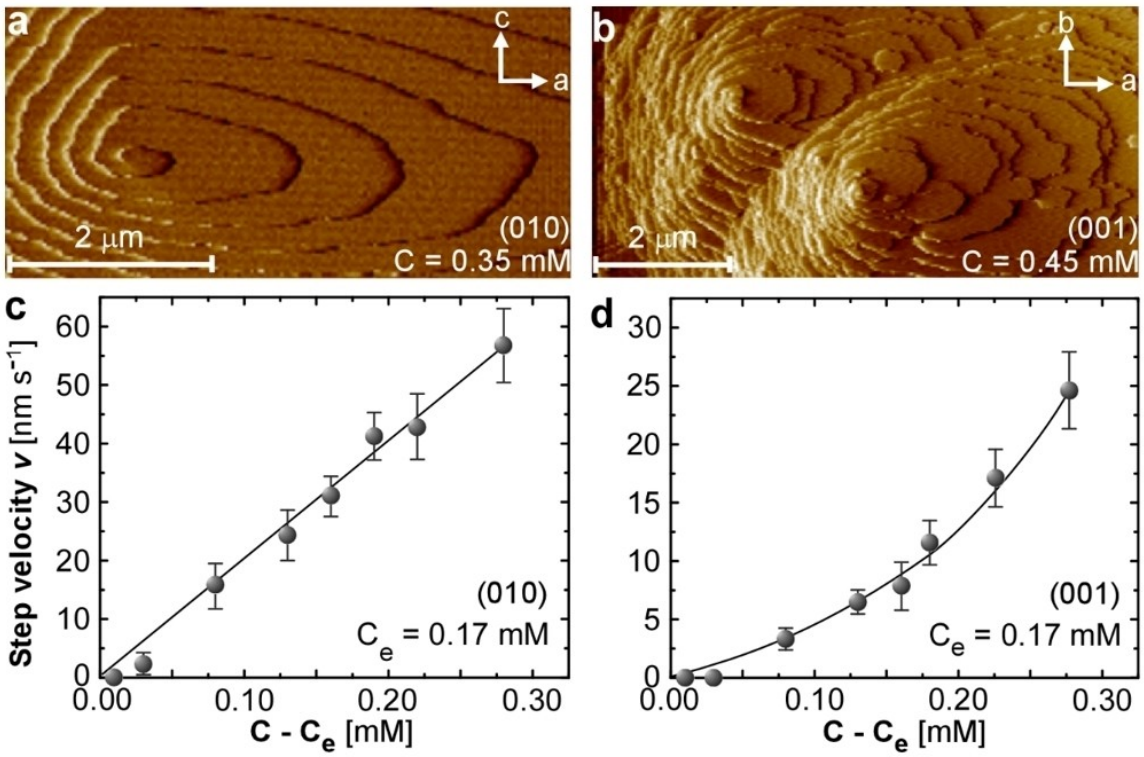

Figure 5. The growth of (010) and (001) faces of EtpI crystals. a, **b**, Generation of new crystal layers by a screw dislocation outcropping on the respective face. **c, d**, The velocity v of steps in the [100] direction as a function of EtpI concentration C in octanol.

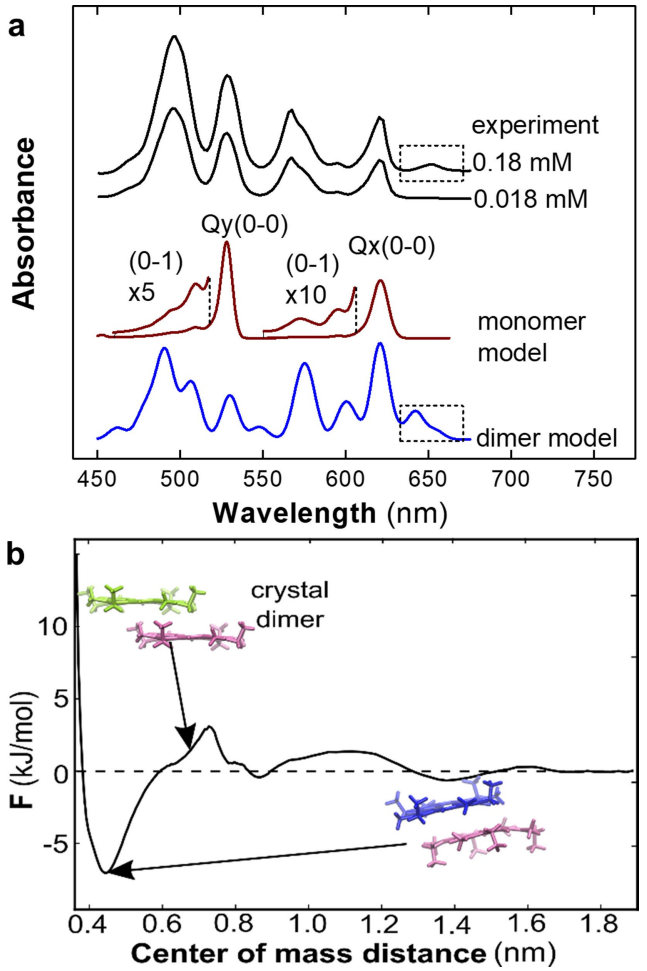

Figure 6. Etioporphyrin dimers in solution. a, Absorbance spectra of EtpI dissolved in octanol, at 0.018 mM and 0.18 mM, and model spectra for EtpI monomer and dimer computed using TDDFT. The monomer model spectrum combines contributions from purely electronic transitions, Qy(0-0) and Qx(0-0), and the vibrational sidebands, (0-1); the latter are displayed exaggerated by 5× and 10× for clarity. **b**, All-atom MD calculation of the potential of mean force F between two EtpI monomers in octanol. Insets: two configurations of the two monomers, top left, as in the crystal lattice, the constituent monomers are shown in purple and green, respectively; and bottom right, shown in violet and pink, as in the deep F minimum, which deviates from the one in the crystal by closer separation, and relative twist and tilts of the monomers.

the monomers and dimers to the measured spectra deconvolute owing to the band at 655 nm. This band is seen in the experimental absorbance spectrum at $C_{EtpI} = 0.18$ mM, but vanishes if the EtpI concentration is lowered to 0.018 mM. For EtpI monomers, an absorption band at a wavelength higher than that of the (0-0) transition would correspond to a (1-0) transition from the first vibrational level in the ground state and carry vibrational energy of ca. 87 meV. Vibrations with such high energy, however, cannot be populated at room temperature and render this interpretation not viable. On the other hand, the dimer model produces absorption bands above 625 nm that involve electronic transitions with charge transfer between the two monomers in a dimer. We identify the absorption band at 655 nm in the spectra of EtpI in octanol as a spectral fingerprint of dimers. Its weak intensity at $C_{EtpI} =$ 0.018 mM (Figure 6a) announces a paucity of dimers at this concentration. The strong Q bands at the same C_{EtpI} ascribe the absorbance in the range 450 to 625 nm to the EtpI monomer.

For further insight into the EtpI monomer – dimer dynamics, we carried out all atom MD simulations of the dimerization equilibrium of EtpI dissolved in octanol. Analogously to the MD computations for OZPN in EtOH/H_2O, the potential of mean force F between two EtpI monomers reaches a minimum at short separations (Figure 6b), which signals a EtpI propensity to dimerize. Similarly to OZPN, we evaluate the free energy of EtpI dimerization in octanol from the potential value at the deep minimum. We compute the dimerization constant K_D as the ratio of integrals of F over the closest range minimum and the unbound state and is 4 ± 0.1 M^{-1}. Mass balance calculations reveal that the concertation of dimers in a solution with total EtpI concentration equal to the solubility, 0.17 mM, is ca. 0.0002 mM and it increases to about 0.002 mM at the highest $C_{EtpI} = 0.45$ mM, at which step velocities were measured (Figure 5). The lower dimer content in EtpI solutions than in OZPN solutions is due to the lower total EtpI concentration, dictated by the lower solubility of this compound. The superlinear increase of the dimer concentration with higher C_{EtpI} manifests as quadratic increase of the step velocity on the (001) face, concurrently with the proposed mechanism of growth by dimer incorporation on that face. Notably, the good correspondence of the $v(C_{EtpI})$ correlation to a quadratic rate law minimizes the potential contributions of trimers and higher oligomers – expected to enforce an even steeper $v(C_{EtpI})$ increase – to growth on the (001) EtpI face. The monomer concentration represents the majority of the total C_{EtpI} and increases proportionally to it. This proportionality supports the proposal that the (010) face, on which the step velocity increases linearly with C_{EtpI} grows by incorporation of monomers.

The EtpI dimerization potential diverges from that for OZPN in several key aspects. Most revealingly, the dimer structure found in the crystal lattice does not occupy the global minimum of the potential of mean force, which indicates that dimers that reproduce the arrangement of molecules in the crystal are a minority in the solution. The dimer which occupies the global minimum diverges from the crystallographic configuration in its shorter separation and the relative twist and tilt of the constituent EtpI monomers. In further distinction from OZPN, the global minimum in the EtpI potential of mean force is relatively broad (Figure 6b), which implies that a variety of dimers whose configurations vary by twist and tilt angles and separation exist in a barrier-free equilibrium with each other as they occupy states that differ in free energy by less that the thermal energy, $k_B T \cong 2500$ J mol^{-1} (k_B, Boltzmann constant; T, temperature). Owing to the conformational differences with how EtpI monomers arrange

in the crystal lattice, their incorporation into the kinks must drastically diverge from simple association. We propose that on approach to a kink on the (001) face a dimer reconfigures to attain the crystal structure; this mechanism are subject to further exploration by MD simulations.

The observations and simulations with EtpI reveal complex dynamics of solute oligomerization that impact the kinetics of crystal growth. Monomers dominate the population of EtpI in the solution and the kinetics of layer growth identify them as the solute units that incorporate in the kinks on the (010) face, whereas the minority dimers are selected as the growth unit on the (001) face. The conformation of solute EtpI dimers is not unique, but rather a variety of dimers that differ by separation and tilt and twist angles coexist and continuously interconvert is the solution. The cohabitating dimers structure differently from the arrangement of the EtpI molecules in the crystal lattice. The solution dimers transition to the crystal conformation upon incorporation into kinks on the (001) face.

Summary and Conclusions

Olanzapine and etioporphyrin I, conformally with numerous other crystals, exhibit no correlation between the solute oligomers and identifiable crystal structural blocks. Whereas olanzapine in solution assembles into a unique dimer that captures a minor fraction of the solute, etioporphyrin I engages in exceedingly complex oligomerization dynamics, whereby a variety of dimers coexist in equilibrium but remain a minority of the solute population. Spectroscopy techniques contribute to the understanding of the rich solute behaviors only when integrated with first principles calculations and atomistic simulations.

The crystal growth unit cannot be deduced confidently neither from the solute oligomers, nor from the crystal structure. Olanzapine is an example where the crystal growth unit diverges from the majority solute monomers and matches the crystal structural element. With etioporphyrin I, the monomeric crystal growth unit selected by one of the faces is identical to both the crystal structural element and the majority solute species, whereas the dimeric growth unit on the other face represents just one of the elements of a vast variety of dimers that cohabit in the solution and diverge from the crystal lattice motif.

The spatial and temporal resolutions of the state-of-the-art *in situ* techniques to characterize the molecular processes that comprise crystal growth are limited and attempts to directly image the growth unit may be doomed to failure. Thus, the kinetics of layer growth that manifest the elementary reaction at the kinks, considered in the contexts of solute oligomerization and crystal structure, are an indispensable indicator of the crystal growth unit.

The proposed method to identify the solute species that incorporates into the kinks, an issue of huge fundamental and practical significance, illuminates a pathway through the hurdles along the road to crystallization prediction and control.

Acknowledgments

This work was supported by the National Science Foundation (award no. DMR-2128121), NIH (award no. AI150763), and The Welch Foundation (Grant E-1882). VGH's work was supported by the State of Texas through the Texas Center for Superconductivity at the University of Houston. The authors acknowledge the CMAC National Facility, housed within the University of Strathclyde's Technology and Innovation Centre, and funded with a UKRPIF (UK Research Partnership Investment Fund) capital award, SFC ref H13054, from the Higher Education Funding Council for England (HEFCE) and the EPSRC Future Continuous Manufacturing and Advanced Crystallisation Research Hub (Grant Ref: EP/P006965/1). Computational resources were provided by the Hewlett Packard Enterprise Data Science Institute at the University of Houston and the Texas Advanced Computing Center at the University of Texas at Austin.

References

[1] a) M. J. Thomason, C. R. Seabourne, B. M. Sattelle, G. A. Hembury, J. S. Stevens, A. J. Scott, E. F. Aziz, S. L. M. Schroeder, *Faraday Discuss.* **2015**, *179*, 269–289; b) W. Tang, H. Mo, M. Zhang, S. Parkin, J. Gong, J. Wang, T. Li, *J. Phys. Chem. B* **2017**, *121*, 10118–10124; c) V. Cheynier, R. Schneider, J.-M. Salmon, H. Fulcrand, in *Comprehensive Natural Products II* (Eds.: H.-W. Liu, L. Mander), Elsevier, Oxford, **2010**, pp. 1119–1172; d) A. Idrissi, P. Damay, K. Yukichi, P. Jedlovszky, *J. Chem. Phys.* **2008**, *129*, 164512.

[2] a) C. A. Hunter, J. F. McCabe, A. Spitaleri, *CrystEngComm* **2012**, *14*, 7115–7117; b) K. R. Back, R. J. Davey, T. Grecu, C. A. Hunter, L. S. Taylor, *Cryst. Growth Des.* **2012**, *12*, 6110–6117; c) R. Davey, G. Dent, R. Mughal, S. Parveen, *Cryst. Growth Des.* **2006**, *8*, 1788–1796; d) S. Parveen, R. J. Davey, G. Dent, R. G. Pritchard, *Chem. Commun.* **2005**, 1531–1533; e) M. Gardner, A. J. Guerin, C. A. Hunter, U. Michelsen, C. Rotger, *New J. Chem.* **1999**, *23*, 309–316.

[3] S. A. Kulkarni, E. S. McGarrity, H. Meekes, J. H. ter Horst, *Chem. Commun.* **2012**, *48*, 4983–4985.

[4] A. Spitaleri, C. A. Hunter, J. F. McCabe, M. J. Packer, S. L. Cockroft, *CrystEngComm* **2004**, *6*, 490–493.

[5] J. Chen, B. L. Trout, *J. Phys. Chem. B* **2008**, *112*, 7794–7802.

[6] a) P. Dandekar, Z. B. Kuvadia, M. F. Doherty, *Annu. Rev. Mater. Res.* **2013**, *43*, 359–386; b) M. A. Lovette, A. R. Browning, D. W. Griffin, J. P. Sizemore, R. C. Snyder, M. F. Doherty, *Ind. Eng. Chem. Res.* **2008**, *47*, 9812–9833; c) Y. Sun, C. J. Tilbury, S. M. Reutzel-Edens, R. M. Bhardwaj, J. Li, M. F. Doherty, *Cryst. Growth Des.* **2018**, *18*, 905–911.

[7] a) N. Blagden, R. J. Davey, *Cryst. Growth Des.* **2003**, *3*, 873–885; b) I. Weissbuch, V. Y. Torbeev, L. Leiserowitz, M. Lahav, *Angew. Chem. Int. Ed.* **2005**, *44*, 3226–3229; *Angew. Chem.* **2005**, *117*, 3290–3293; c) R. A. Chiarella, A. L. Gillon, R. C. Burton, R. J. Davey, G. Sadiq, A. Auffret, M. Cioffi, C. A. Hunter, *Faraday Discuss.* **2007**, *136*, 179–193; d) N. Varsano, F. Beghi, N. Elad, E. Pereiro, T. Dadosh, I. Pinkas, A. J. Perez-Berna, X. Jin, H. S. Kruth, L. Leiserowitz, L. Addadi, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 7662–7669; e) I. Weissbuch, R.

Popovitz-Biro, M. Lahav, L. Leiserowitz, *Acta Crystallogr. Sect. B* **1995**, *51*, 115–148.

[8] a) S.-T. Yau, D. N. Petsev, B. R. Thomas, P. G. Vekilov, *J. Mol. Biol.* **2000**, *303*, 667–678; b) S.-T. Yau, B. R. Thomas, O. Galkin, O. Gliko, P. G. Vekilov, *Proteins Struct. Funct. Genet.* **2001**, *43*, 343–352; c) S.-T. Yau, B. R. Thomas, P. G. Vekilov, *Phys. Rev. Lett.* **2000**, *85*, 353–356; d) S.-T. Yau, B. R. Thomas, P. G. Vekilov, *J. Cryst. Growth* **2001**, *232*, 188–194; e) M. Sleutel, D. Maes, L. Wyns, R. Willaert, *Cryst. Growth Des.* **2008**, *8*, 4409–4414; f) M. Sleutel, C. Vanhee, C. V. de Weerdt, K. Decanniere, D. Maes, L. W. R. Willaert, *Cryst. Growth Des.* **2008**, *8*, 1173–1180; g) M. Sleutel, R. Willaert, C. Gillespie, C. Evrard, L. Wyns, D. Maes, *Cryst. Growth Des.* **2009**, *9*, 497–504; h) M. Sleutel, J. Lutsko, A. E. S. Van Driessche, M. A. Durán-Olivencia, D. Maes, *Nat. Commun.* **2014**, *5*; i) A. J. Malkin, Y. G. Kuznetsov, R. W. Lucas, A. McPherson, *J. Struct. Biol.* **1999**, *127*, 35–43; j) A. J. Malkin, Y. G. Kuznetsov, A. McPherson, *J. Cryst. Growth* **1999**, *196*, 471–488; k) Y. G. Kuznetsov, A. J. Malkin, R. W. Lucas, M. Plomp, A. McPherson, *J. Gen. Virol.* **2001**, *82*, 2025–2034; l) A. McPherson, A. J. Malkin, Y. G. Kuznetsov, M. Plomp, *Acta Crystallogr. Sect. D* **2001**, *57*, 1053–1060; m) A. J. Malkin, A. McPherson, in *From fluid-solid interfaces to nanostructural engineering. vol. 2. Assembly in hybrid and biological systems*. (Eds.: J. J. De Yoreo, X. Y. Lui), Plenum/Kluwer Academic, New York, **2004**, pp. 201–238; n) A. J. Malkin, R. E. Thorne, *Methods* **2004**, *34*, 273–299; o) D. K. Georgiou, P. G. Vekilov, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1681–1686; p) S.-T. Yau, P. G. Vekilov, *Nature* **2000**, *406*, 494–497; q) S.-T. Yau, P. G. Vekilov, *J. Am. Chem. Soc.* **2001**, *123*, 1080–1089.

[9] D. Gidalevitz, R. Feidenhans′l, S. Matlis, D.-M. Smilgies, M. J. Christensen, L. Leiserowitz, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 955–959.

[10] J. Huang, T. C. Stringfellow, L. Yu, *J. Am. Chem. Soc.* **2008**, *130*, 13973–13980.

[11] a) J. Chen, B. L. Trout, *J. Phys. Chem. B* **2010**, *114*, 13764–13772; b) S. Hamad, C. R. A. Catlow, *CrystEngComm* **2011**, *13*, 4391–4399.

[12] R. C. Burton, E. S. Ferrari, R. J. Davey, J. L. Finney, D. T. Bowron, *J. Phys. Chem. B* **2010**, *114*, 8807–8816.

[13] a) J. Chen, E. Zhu, J. Liu, S. Zhang, Z. Lin, X. Duan, H. Heinz, Y. Huang, J. J. De Yoreo, *Science* **2018**, *362*, 1135–1139; b) Z. Liu, Z. Zhang, Z. Wang, B. Jin, D. Li, J. Tao, R. Tang, J. J. De Yoreo, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 3397–3404; c) K. N. Olafson, T. Q. Nguyen, J. D. Rimer, P. G. Vekilov, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 7531–7536; d) W. Ma, V. A. Balta, R. West, K. N. Newlin, O. Š. Miljanić, D. J. Sullivan, P. G. Vekilov, J. D. Rimer, *J. Biol. Chem.* **2021**, *296*, 100123; e) M. Warzecha, L. Verma, B. F. Johnston, J. C. Palmer, A. J. Florence, P. G. Vekilov, *Nat. Chem.* **2020**, *12*, 914–920; f) W. Ma, J. F. Lutsko, J. D. Rimer, P. G. Vekilov, *Nature* **2020**, *577*, 497–501.

[14] B. Fulton, K. L. Goa, *Drugs* **1997**, *53*, 281–298.

[15] R. M. Bhardwaj, L. S. Price, S. L. Price, S. M. Reutzel-Edens, G. J. Miller, I. D. H. Oswald, B. F. Johnston, A. J. Florence, *Cryst. Growth Des.* **2013**, *13*, 1602–1617.

[16] S. Askin, J. K. Cockcroft, L. S. Price, A. D. Gonçalves, M. Zhao, D. A. Tocher, G. R. Williams, S. Gaisford, D. Q. M. Craig, *Cryst. Growth Des.* **2019**, *19*, 2751–2757.

[17] M. Warzecha, R. Guo, R. M. Bhardwaj, S. M. Reutzel-Edens, S. L. Price, D. A. Lamprou, A. J. Florence, *Cryst. Growth Des.* **2017**, *17*, 6382–6393.

[18] I. Wawrzycka-Gorczyca, P. Borowski, J. Osypiuk-Tomasik, L. Mazur, A. Koziol, *J. Mol. Struct.* **2007**, *830*, 188–197.

[19] W. K. Burton, Cabrera, N. Frank, *Phil. Trans. R. Soc. A* **1951**, ***243***, 299–360.

[20] a) A. A. Chernov, *Modern Crystallography III, Crystal Growth*,, Springer, Berlin, **1984**; b) G. H. Gilmer, R. Ghez, N. Cabrera, *J. Crystal Growth* **1971**, *8*, 79–93.

[21] a) J. J. De Yoreo, P. G. Vekilov, in *Biomineralization, Vol. 54*, **2003**, pp. 57–93; b) J. J. De Yoreo, T. A. Land, B. Dair, *Phys. Rev. Lett.* **1994**, *73*, 838–841; c) T. A. Land, A. J. Malkin, Y. G. Kuznetsov, A. McPherson, J. J. Deyoreo, *Phys. Rev. Lett.* **1995**, *75*, 2774–2777; d) H. H. Teng, P. M. Dove, C. A. Orme, J. J. De Yoreo, *Science* **1998**, *282*, 724–727; e) T. A. Land, T. L. Martin, S. Potapenko, G. T. Palmore, J. J. De Yoreo, *Nature* **1999**, *399*, 442–445; f) Y. G. Kuznetsov, A. A. Chernov, P. G. Vekilov, I. L. Smol′skii, *Sov. Phys.-Crystallogr.* **1987**, *32*, 584–587; g) P. G. Vekilov, Y. G. Kuznetsov, A. A. Chernov, *J. Cryst. Growth* **1992**, *121*, 643–655; h) P. G. Vekilov, Y. G. Kuznetsov, A. A. Chernov, *J. Cryst. Growth* **1992**, *121*, 44–52; i) P. G. Vekilov, in *Studies and Concepts in Crystal Growth* (Ed.: H. Komatsu), Pergamon, Oxford, **1993**, pp. 25–49; j) P. G. Vekilov, F. Rosenberger, *J. Cryst. Growth* **1996**, *158*, 540–551; k) P. G. Vekilov, F. Rosenberger, *J. Cryst. Growth* **1998**, *186*, 251–261; l) O. Gliko, I. Reviakine, P. G. Vekilov, *Phys. Rev. Lett.* **2003**, *90*, 225503; m) D. N. Petsev, K. Chen, O. Gliko, P. G. Vekilov, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 792–796; n) Y. Qutub, I. Reviakine, C. Maxwell, J. Navarro, E. M. Landau, P. G. Vekilov, *J. Mol. Biol.* **2004**, *343*, 1243–1254; o) O. Gliko, N. Neumaier, W. Pan, I. Haase, M. Fischer, A. Bacher, S. Weinkauf, P. G. Vekilov, *J. Am. Chem. Soc.* **2005**, *127*, 3433–3438.

[22] M. Warzecha, M. S. Safari, A. J. Florence, P. G. Vekilov, *Cryst. Growth Des.* **2017**, *17*, 6668–6676.

[23] M. A. Lovette, M. F. Doherty, *Physical Review E* **2012**, *85*, 021604.

[24] A. A. Chernov, *Sov. Phys. Uspekhi* **1961**, *4*, 116–148.

[25] a) C.-M. Che, H.-F. Xiang, S. S.-Y. Chui, Z.-X. Xu, V. A. L. Roy, J. J. Yan, W.-F. Fu, P. T. Lai, I. D. Williams, *Chem. Asian J.* **2008**, *3*, 1092–1103; b) M. H. Hoang, Y. Kim, M. Kim, K. H. Kim, T. W. Lee, D. N. Nguyen, S.-J. Kim, K. Lee, S. J. Lee, D. H. Choi, *Adv. Mater.* **2012**, *24*, 5363–5367; c) M. H. Hoang, Y. Kim, S.-J. Kim, D. H. Choi, S. J. Lee, *Chem. Eur. J.* **2011**, *17*, 7772–7776.

[26] a) R. Giovannetti, in *Macro To Nano Spectroscopy* (Ed.: J. Uddin), IntechOpen, Rijeka, **2012**; b) T. Bruhn, C. Brückner, *Phys. Chem. Chem. Phys.* **2015**, *17*, 3560–3569; c) K. Kano, H. Minamizono, T. Kitae, S. Negi, *J. Phys. Chem. A* **1997**, *101*, 6118–6124; d) S. S. Dvornikov, K. N. Solov′ev, M. P. Tsvirko, *J. Appl. Spectrosc.* **1983**, *38*, 581–585; e) J. B. PaineIII, D. Dolphin, M. Gouterman, *Can. J. Chem.* **1978**, *56*, 1712–1715.

[27] L. Edwards, D. H. Dolphin, M. Gouterman, A. D. Adler, *J. Mol. Spectrosc.* **1971**, *38*, 16–32.

Manuscript received: July 27, 2021
Revised manuscript received: September 8, 2021
Version of record online: October 27, 2021