LETTER

Predicting crystal growth via a unified kinetic three–dimensional partition model

Michael W. Anderson[1], James T. Gebbie-Rayet[1]†, Adam R. Hill[1], Nani Farida[1], Martin P. Attfield[1], Pablo Cubillas[1]†, Vladislav A. Blatov[2,3], Davide M. Proserpio[2,4], Duncan Akporiaye[5], Bjørnar Arstad[5] & Julian D. Gale[6]

Understanding and predicting crystal growth is fundamental to the control of functionality in modern materials. Despite investigations for more than one hundred years[1–5], it is only recently that the molecular intricacies of these processes have been revealed by scanning probe microscopy[6–8]. To organize and understand this large amount of new information, new rules for crystal growth need to be developed and tested. However, because of the complexity and variety of different crystal systems, attempts to understand crystal growth in detail have so far relied on developing models that are usually applicable to only one system[9–11]. Such models cannot be used to achieve the wide scope of understanding that is required to create a unified model across crystal types and crystal structures. Here we describe a general approach to understanding and, in theory, predicting the growth of a wide range of crystal types, including the incorporation of defect structures, by simultaneous molecular-scale simulation of crystal habit and surface topology using a unified kinetic three-dimensional partition model. This entails dividing the structure into 'natural tiles' or Voronoi polyhedra that are metastable and, consequently, temporally persistent. As such, these units are then suitable for re-construction of the crystal via a Monte Carlo algorithm. We demonstrate our approach by predicting the crystal growth of a diverse set of crystal types, including zeolites, metal–organic frameworks, calcite, urea and L-cystine.

By understanding crystal growth at the molecular scale it is possible to control crystal habit, crystal size, the elimination or incorporation of defects and the development of intergrowth structures. Because crystals are used, for example, in technologies including pharmaceuticals and gas storage, as separation materials, in optoelectronic devices and as heterogeneous catalysts, such understanding is vital. We can illustrate many of the problems that must be addressed in crystal growth by considering zeolites[12] as an example, which represent a very complex, yet important, crystal type that forms the backbone of the heterogeneous catalysis industry. Zeolites are nanoporous materials for which the framework of the material is constructed from a strong, covalently bonded network of Si–O and Al–O bonds. The pores of the material are filled with water and cations that balance the negative charge on the framework. Crystals of zeolites grow from aqueous solutions at temperatures up to about 230 °C and it is well known from nuclear magnetic resonance spectroscopy that the solution phase exhibits very complex speciation[13–16]. This is a seemingly intractable problem in terms of defining a simple set of rules that govern the hundreds of different zeolite structures, let alone the thousands of related crystal structures such as metal–organic frameworks (MOFs)[17–19]. However, the course of a crystallization is relatively predictable and, therefore, there must be a relatively small number of rules that govern the most important aspects of the crystal growth, with subsidiary rules governing deviations.

The starting point in our simplification comes from a general Monte Carlo simulation applied to the growth of fats[20]. In this work it was shown that the principal determinant of crystal growth was the local internal energy at the crystal surface in relation to the chemical potential of the phase from which the crystal grows. This is a very important simplification because it allows the growth medium—solution, melt, gas and so on—to be considered to have only a growth potential, and so the speciation does not need to be considered in detail. Although this growth potential will be a result of the speciation, this can be treated as a subsidiary effect to be considered subsequently[13–16]. In the case of multicomponent crystals—such as MOFs or co-crystals for which species in solution, for example, linkers and metal centres, cannot interchange—a driving force for each component needs to be considered, unless the stoichiometries of the two phases are matched. For zeolites, or any system in which the nutrient is interconverting, a single driving force can be considered equivalent to a single-component system.

The crystal structure then needs to be broken down into 'units of growth', a process that is normally referred to as 'coarse-graining the problem'. To deconstruct the problem, we require a distinction between 'unit of growth' and 'growth unit'. To identify a growth unit, we need to know the growth mechanism, however, a unit of growth is just a suitable division of the structure in terms of metastability. If the material of interest is a molecular crystal, then a unit of growth would be a single molecule, because this represents a strongly bonded entity that remains intact during crystallization, forming relatively weak bonding with neighbours to yield the crystal. Such a unit of growth is probably, in many cases, the actual growth unit for the crystal, assuming that the unit does not dimerize in solution. However, for a zeolite, which is a fully connected three-dimensional network of covalent bonds, a single-molecule unit of growth is not viable. For our analysis, the unit of growth is any structural element that represents a metastable surface structure with small enough dimensions to describe all of the intricacies of the crystal formation. As a metastable entity it will be persistent in time at the crystal surface during growth and can therefore be considered to determine the overall rate of crystal growth. For the simulation of the full three-dimensional growth of a crystal, for example via the development of a kinetic Monte Carlo model, only the rate-determining steps are required.

We now return to the problem of nanoporous zeolites composed of condensed tetrahedral silicate units forming cage-like structures. These cages are strongly related to metastable surface entities because the cage wrapping permits maximum condensation of the cage[11]. Consider a cage within the bulk of the zeolite with all tetrahedral silicon sites fully condensed (termed Q^4 units). When the same cage is located at the surface of the zeolite crystal, most of the silicon sites will lose only one bond of condensation, making this structural configuration a minimum in energy for a surface moiety. Of the more than 200 zeolite structures, around one quarter consist of Q^4 units that will lose only

[1]Centre for Nanoporous Materials, School of Chemistry, The University of Manchester, Oxford Road, Manchester M13 9PL, UK. [2]Samara Center for Theoretical Materials Science (SCTMS), Samara University, Academician Pavlov Street 1, Samara 443011, Russia. [3]School of Materials Science and Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China. [4]Università degli Studi di Milano, Dipartimento di Chimica, Via Camillo Golgi 19, 20133 Milano, Italy. [5]SINTEF Materials and Chemistry, PO Box 124, Blindern, 0314 Oslo, Norway. [6]Curtin Institute for Computation, Department of Chemistry, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia. †Present addresses: Scientific Computing Department, STFC Daresbury Laboratory, Warrington WA4 4AD, UK (J.T.G.-R.); Earth Sciences Department, Durham University, Lower Mountjoy, South Road, Durham DH1 3LE, UK (P.C.).
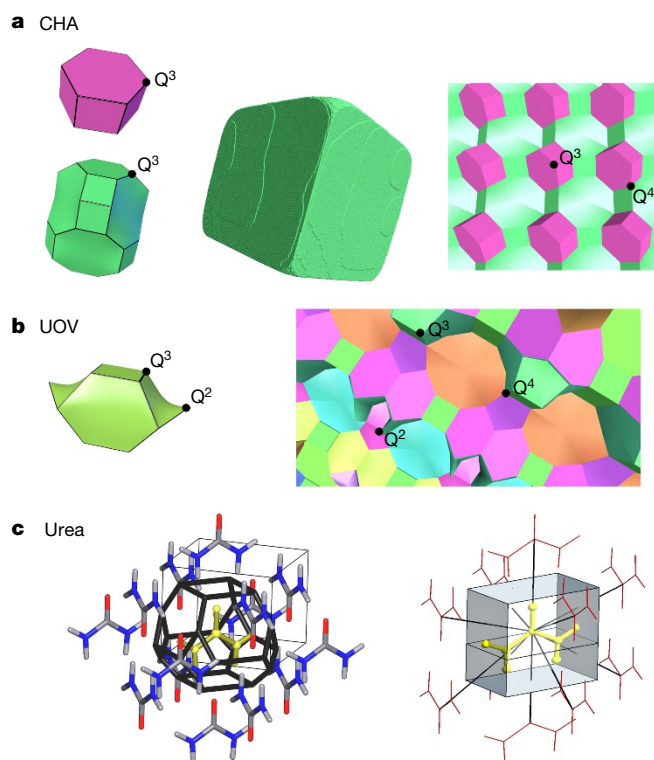
a CHA

Q^3

Q^3

Q^3

Q^4

b UOV

Q^3

Q^2

Q^3

Q^4

Q^2

c Urea

Figure 1 | Demonstration of the tiling and Voronoi-polyhedra partitioning methods. a, The chabazite structure (CHA), which is composed of two cages (tiles), the double 6-ring (pink) and chabazite cage (green). Both of these tiles consist entirely of Q^3 tetrahedra and are considered as closed cages or tiles. As the crystal grows (middle), the cages condense and the Q^3 vertices are converted into Q^4 vertices (right), thereby stabilizing the cage relative to the solution phase. **b**, The complex UOV structure, which consists of 16 tiles (see Extended Data Fig. 3). One of these tiles is shown, one that consists of both Q^3 and Q^2 vertices and is therefore considered as an open cage or tile (see Methods for definition of open versus closed cages and tiles). Condensation again results in stabilization of these tiles relative to the solution phase. **c**, A tile representation of the molecular crystal urea. The left panel shows the initial Voronoi construction with 14 urea neighbours surrounding the central molecule. Four of these interactions are very weak and can be neglected, leaving the ten interactions represented by the black lines in the right panel. Each interaction passes through the face of the ten-sided tile. Grey, red and blue represent carbon, oxygen and nitrogen atoms, respectively, the yellow represents the central urea molecule, and the black edges illustrate the Voronoi polyhedron created by these molecular interactions.

one condensation at the surface ($Q^4 \rightarrow Q^3$). The other three quarters have Q^4 units that may lose two condensations ($Q^4 \rightarrow Q^2$); however, these structures will still be at an energy minimum. Therefore, the cages become a suitable unit of growth even though they are not the growth unit. These cages are three-dimensional space-filling tiles that can be computed in a relatively straightforward manner using algorithms such as those implemented in ToposPro[21,22] (Fig. 1). Consequently, this establishes a simplified route to coarse-graining the zeolite problem into energetically minimized, metastable, rate-determining steps that, when balanced against a potential energy driving force from the growth medium, permits generation of a general kinetic Monte Carlo algorithm for zeolites. We use the so-called 'natural tiles' that describe the system of minimal cages in a unique and unambiguous way according to a strict algorithm[21]. In this regard it is very different to the diverse systems of secondary building units often chosen by structural chemists according to their personal view on how to divide up structures. The units of growth are space-filling and, although the crystal is nanoporous, they are considered to fill all space during growth (the voids within cages are filled with water and cations). This property is the same for the growth of any crystal, whether nanoporous or not.

Our approach is extendable to any cage-like structure, regardless of the bonding type, so MOFs with extensive coordination bonds are immediately treatable.

To extend our approach to other crystals we use the Voronoi partitioning procedure to fill the space with polyhedral units, in a manner opposite to that used for the tiling method: with Voronoi polyhedra, the objects (atoms or molecules) occupy the centres, whereas with tiles they occupy vertices. Molecular crystals, such as aspirin, urea and water, can be categorized as a three-dimensional Voronoi partition, in which the molecule sits at the centre of a Voronoi polyhedron and the faces of the polyhedron represent the interactions with neighbouring molecules (Fig. 1). Similarly, for ionic crystals, such as calcite and zinc oxide, the ions sit at the centres of Voronoi polyhedra with faces representing the interactions between cations and anions. In these last two examples, the network of interactions can be considered without needing to introduce three-dimensional partitioning; however, it is useful to realize that all crystal systems can be treated in the same manner. To summarize, we assume the units of growth to be polyhedral (tiles or Voronoi polyhedra depending on nature of the crystal). The Voronoi partition can also be used for structures that have no tiling, for example, for polycatenated networks.

Having partitioned the crystal space, the problem then is to establish the energies of all of the three-dimensional polyhedral units in any configuration and the degree of condensation/attachment at the surface of a growing crystal relative to the bulk phase. For complex crystal systems there could be thousands of possible types of surface site, although, in principle, only a fraction of these will be topologically viable during crystal growth. By interfacing our kinetic Monte Carlo code with the three-dimensional partitioning approach of ToposPro[22], we can compute all of the possible connectivities for any partitioning pattern and, consequently, any crystal structure. Then, to a first approximation, the energies of the polyhedral units are directly related to the degree of condensation or attachment (see Extended Data Fig. 1 for the LTA zeolite system). Secondary energetic effects can be computed at a much higher level of simulation to determine subsidiary effects, but most structural features are determined purely by connectivity. Common defects, such as screw dislocations, can be incorporated by displacing three-dimensional polyhedral units to equivalent sites along the screw core, resulting in perfect crystal re-connection. Growth modifiers can also be simulated by poisoning units of growth accordingly (that is, by reducing their probability of growth). This approach permits both growth and dissolution at individual surface sites, depending on whether the chemical potential of the growth medium is above or below the energy of that surface site. In this manner, by changing the driving force systematically within the simulation, the equilibrium morphology is found when the rates of growth and dissolution are balanced. Examples for the LTA and FAU zeolite structures are shown in Extended Data Figs 1 and 2, respectively, illustrating how both the habit of the crystal and the much more sensitive surface topology can be matched with experiment across all crystal faces.

This approach enables straightforward computation of crystals no matter what degree of complexity exists in the structure; for example, the UOV zeolite structure (Fig. 2 and Extended Data Fig. 3), which has a very large unit cell and is constructed from 16 tiles in a mixture of open and closed environments, is readily treated in an efficient manner. For such a system, even using the same energy penalty for every tile vertex gives both a crystal habit and surface topology very similar to that observed experimentally. This computation yields the terrace structure, which also includes the nature of the surface termination that, for nanoporous materials, is the gateway to the internal porosity. This approach also demonstrates how framework crystals such as NES (Extended Data Fig. 4) have great difficulty circumventing large cages that will necessarily represent large energy barriers. The resulting crystals are very thin plates and any modification to this morphology would require careful attention to the stabilizing of the large cage through templating.

The MFI zeolite framework type, also known as ZSM-5—one of the most important industrial catalysts—reveals not only the surface
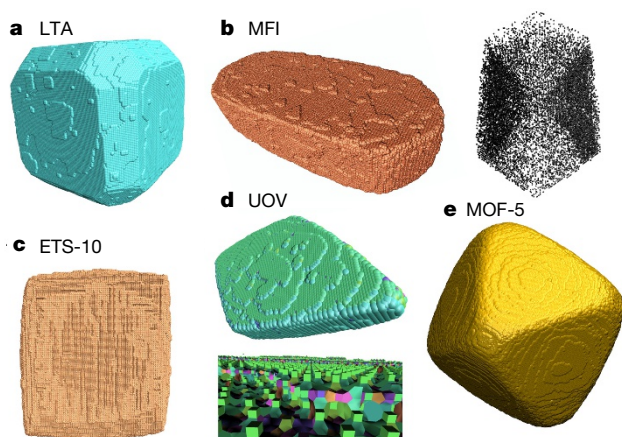
Figure 2 | Results of simulations run on various framework types.
a, LTA (zeolite A) consists of three closed cages (tiles), and the
experimental morphology can be achieved by adjusting the energy of
these tiles independently relative to solution (see Extended Data Fig. 1 and
Supplementary Video 1). **b**, MFI (also known as ZSM-5 or silicalite) is a
complex structure consisting of 10, all open, tiles. The morphology and
topology can be simulated very well using different energy penalties for
large and small tiles. Interrogation of the internal structure of the crystal
reveals an hourglass structure similar to that observed experimentally
by optical microscopy. This structure is due to crystallization with
incomplete condensation of tiles resulting in the silanol groups that are
known to be present in the ZSM-5 structure (see Extended Data Fig. 5 and
Supplementary Video 4). **c**, ETS-10 is a mixed-coordination octahedral
and tetrahedral nanoporous framework structure that consists of titanate
rods that are stacked layer by layer in an orthogonal arrangement. Viewed
down the [001] axis, as is the case here, the rod-based crystal growth
mechanism is immediately apparent; such a growth mechanism leads to
the incorporation of defects (see Extended Data Fig. 6 and Supplementary
Video 5). **d**, UOV is one of the most complex zeolite structures, with a
very large unit cell and 16, both open and closed, tiles. Our methodology
is able to efficiently grow such a complex structure with both surface
topography and habit matching experimental observations. The surface
structure (top) is determined from the calculations, as is the nature of
partially constructed layers (bottom) at intermediate metastable steps
(see Extended Data Fig. 3 and Supplementary Video 3). **e**, MOFs can
be modelled using two differing methods: first, by treating them as
multicomponent molecular crystals, with metal clusters and organic
linkers treated as separate molecules (as in MOF-5); and second, by using
the same treatment as zeolite frameworks (as seen with HKUST-1 in
Extended Data Fig. 9a–d and Supplementary Video 6). Again the crystal
habit and surface topography match those observed experimentally with
different crystallization conditions (further examples are shown for
MOF-5 in Extended Data Fig. 9e–i).

structure but also the internal structure of the crystals. Within the bulk
of the crystal, tiles remain incomplete—in other words, they possess
dangling silanol groups—consistent with the many internal silanols
that are well-established to be present in ZSM-5. The interesting
discovery is that, because the growth mechanisms on different faces of
the crystal are necessarily different, the silanols are confined to zones
of different density within the crystal (Fig. 2 and Extended Data Fig. 5).
This finding mimics almost exactly the optical birefringence, showing
zoning identical to that observed experimentally (Extended Data Fig. 5)
and has been a source of debate for many years[23]. Similarly, in ETS-10,
which displays rod growth (Extended Data Fig. 6) rather than layer
growth, the incompleteness of the rods results in internal defects that
congregate in a zone from the (001) facets to the centre of the crystal,
as observed experimentally by Raman microscopy[24]. Our kinetic three-
dimensional partitioning model shows that a straightforward growth
mechanism can explain these optical phenomena without the need for
complex arguments related to twinning of the crystals. Common defect
structures, such as screw dislocations, are able not only to replicate the
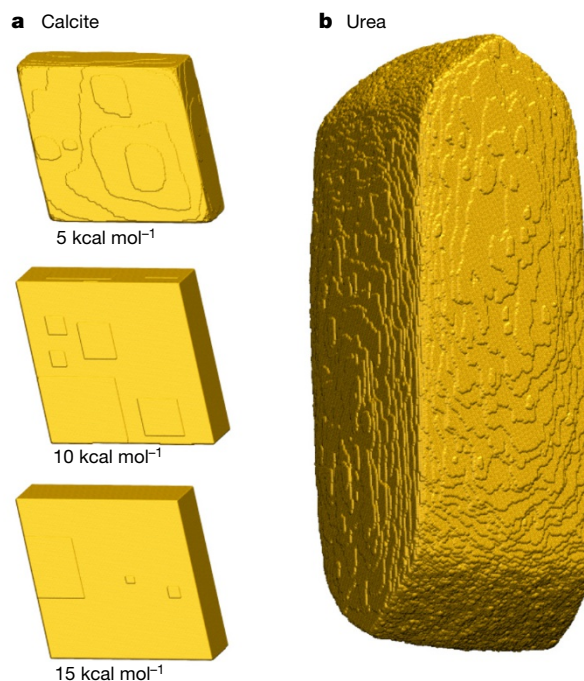spiral topology, such as in LTA and CHA (Extended Data Figs 1 and 7,

**Figure 3 | Simulations of an ionic crystal (calcite) and a molecular crystal
(urea). a**, **b**, Simulations of calcite (**a**) and urea (**b**), demonstrating the
universality of our approach to different crystal classes. All simulations are
shown under equilibrium conditions. **a**, For calcite, the reaction energy for
the conversion of solubilized ions to the crystal per coordination to the crystal
is set at $5\,kcal\,mol^{-1}$, $10\,kcal\,mol^{-1}$ or $15\,kcal\,mol^{-1}$. Calculations show
that the value lies between $10\,kcal\,mol^{-1}$ and $15\,kcal\,mol^{-1}$, and the crystal
habit and surface topography of the two corresponding simulations match
experiment closely. At $5\,kcal\,mol^{-1}$, the terrace edges are much more rounded
than observed experimentally. The main difference between the simulations
for $10\,kcal\,mol^{-1}$ and $15\,kcal\,mol^{-1}$ is the terrace density, which can also be
used as a distinguishing factor. **b**, For urea, three different reaction energies
are used, depending primarily on the strength of interaction in the urea
crystal (discussed in Methods). The large {110} faces are flat and dominated
by terraces elongated in the *c* direction of the crystal (vertical in the image).
The smaller pseudo-{111} faces are rough and generated in a large part by
dissolution when the supersaturation is close to equilibrium.

respectively), but also indicate the relative growth rates of the screw in
relation to the layer growth. Complex interleaving of screw formation
owing to fast and slow growth directions, such as seen in the AEI
zeolite system (Extended Data Fig. 8), can be faithfully reproduced.
Also, the direction of the screw core can be interrogated according to
the multiplicity of the spiral growth emanating at the crystal surface,
such as in the metal–organic HKUST-1 (Extended Data Fig. 9a–d).
Indeed, MOFs are as readily treated as zeolites using this approach,
either as cage (partitioned) structures as in HKUST-1 or as molecular
crystals as shown for MOF-5 (Extended Data Fig. 9e–i). In the latter
case, it is necessary also to consider the solvent as an important element
in the crystal growth, because without it the observed crystal habit and
surface topology cannot be replicated. MOF-5 is a good example of a
multicomponent crystal, demonstrating the power of our approach to
this important general class of materials.

Molecular crystals and ionic crystals (Figs 3 and 4, respectively) are
both amenable to this treatment and, for calcite, the crystallization
energies are in broad agreement with those calculated using a combi-
nation of interatomic potentials and a continuum solvent model (see
Methods). For the L-cystine system, it has been shown[25,26] that growth
on the ⟨001⟩ face proceeds predominantly via screw dislocations. When
this growth mechanism is augmented with the 6_1 screw axis of the
crystal structure and with highly anisotropic rates of crystal growth,
a complex pin-wheel surface topology is generated. Our simulations,
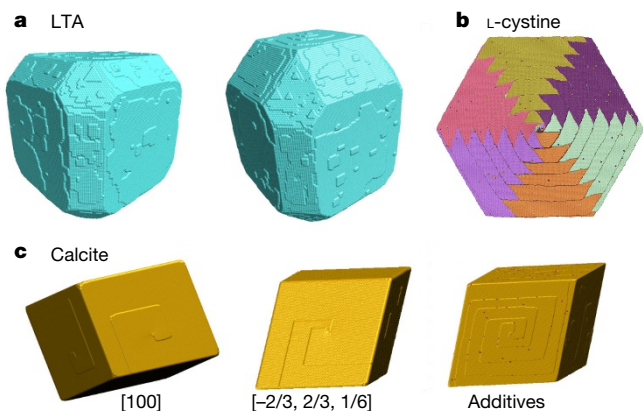which are based on four independent interaction energies, faithfully

Figure 4 | Results of incorporating screw dislocations in growing LTA, L-cystine and calcite crystals. Screw dislocations may be computed using our methodology for any crystal system along any crystal direction. This method operates according only to topology and does not account for the energy of the crystal at the screw core. Nonetheless, it allows all possible topologically permitted structures to be tested for growth morphology; energy considerations can be determined separately. **a**, Simulations of LTA structure with (right) and without (left) a screw dislocation running along [100] through the crystal, vertically. A lengthening of the crystal along the [100] direction is immediately apparent, owing to the greater ease for growth at the spiral growth front. This demonstrates how the relative growth rates of layer-by-layer growth versus spiral growth can be determined. **b**, Pin-wheel crystal growth formation in the L-cystine system caused by the 6_1 screw axis, with hexagonal terraces consisting of six individual L-cystine layers forming a step bunch circumscribed by the slow growth directions. Progression of the step bunches and of single steps is the result of a complex interplay between attachment at single step edges, step bunches and surface sites that can be seen in Supplementary Video 7. **c**, Screw dislocations in calcite. The left panel shows a single screw dislocation with screw core along [100]. Such a dislocation emanates on two adjacent {104} crystal faces. Calcite is also known to exhibit double screw dislocations and the middle panel shows a double Burgers vector screw along [−2/3, 2/3, 1/6], which has the smallest displacement possible for such a double screw. The right panel shows the effect of selective 'poisoning' (via the addition of additives) at two-coordinate sites along terrace edges (red dots), which produces rounding of terrace features.

reproduce all of these growth features and reveal the importance of the interplay between different growth modes in the complete crystallization mechanism (Fig. 4 and Extended Data Fig. 10).

Finally, the addition of growth modifiers is also readily achieved (Fig. 4), and was used to examine the targeting of specific growth sites in relation to the effect of these sites on the growth topology, yielding similar results to those observed experimentally[27]. The power of this approach is in the general applicability across crystal systems, and it provides a window of understanding that can be explored through higher-level calculations on each individual system.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

1. Ostwald, W. Studien über die Bildung und Umwandlung fester Körper. *Z. Phys. Chem.* **22,** 289–330 (1897).
2. Volmer, M. *Kinetics of Phase Formation (Kinetik der Phasenbildung)* http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA800534 (Theodor Steinkopf, 1939).
3. Stranski, I. N. Zur Theorie des Kristallwachstums. *Z. Phys. Chem.* **136,** 259–278 (1928).
4. Kossel, W. Zur Theorie des Kristallwachstums. *Nachrichten der Gesselschaft der Wissenschaften Göttingen, Mathematisch-Physikalische Klasse* **1927,** 135–143 (1927).
5. Burton, W. K., Cabrera, N. & Frank, F. C. The growth of crystals and the equilibrium structure of their surfaces. *Phil. Trans. R. Soc. Lond. A* **243,** 299–358 (1951).
6. Hillner, P. E., Gratz, A. J., Manne, S. & Hansma, P. K. Atomic-scale imaging of calcite growth in real-time. *Geology* **20,** 359–362 (1992).
7. Land, T. A., DeYoreo, J. J. & Lee, J. D. An in-situ AFM investigation of canavalin crystallization kinetics. *Surf. Sci.* **384,** 136–155 (1997).
8. Agger, J. R., Pervaiz, N., Cheetham, A. K. & Anderson, M. W. Crystallization in zeolite A studied by atomic force microscopy. *J. Am. Chem. Soc.* **120,** 10754–10759 (1998).
9. Cuppen, H. M., van Veenendaal, E., van Suchtelen, J., van Enckevort, W.J.P. & Vlieg, E. A Monte Carlo study of dislocation growth and etching of crystals. *J. Cryst. Growth* **219,** 165–175 (2000).
10. De Yoreo, J. J. *et al.* Rethinking classical crystal growth models through molecular scale insights: consequences of kink-limited kinetics. *Cryst. Growth Des.* **9,** 5135–5144 (2009).
11. Brent, R. *et al.* Unstitching the nanoscopic mystery of zeolite crystal formation. *J. Am. Chem. Soc.* **132,** 13858–13868 (2010).
12. Davis, M. E. Ordered porous materials for emerging applications. *Nature* **417,** 813–821 (2002).
13. Kinrade, S. D. & Swaddle, T. W. ^{29}Si NMR studies of aqueous silicate solutions 2: transverse ^{29}Si relaxation and the kinetics and mechanism of silicate polymerisation. *Inorg. Chem.* **27,** 4259–4264 (1988).
14. Knight, C. T. G. & Harris, R. K. Silicon-29 nuclear magnetic resonance studies of aqueous silicate solutions 8: spin-lattice relaxation times and mechanisms. *Magn. Reson. Chem.* **24,** 872–874 (1986).
15. Harris, R. K. & Kimber, B. J. Si NMR as a tool for studying silicones. *Appl. Spectrosc. Rev.* **10,** 117–137 (1975).
16. Petry, D. P. *et al.* Connectivity analysis of the clear sol precursor of silicalite: are nanoparticles aggregated oligomers or silica particles? *J. Phys. Chem. C* **113,** 20827–20836 (2009).
17. Yaghi, O. M. *et al.* Reticular synthesis and the design of new materials. *Nature* **423,** 705–714 (2003).
18. Férey, G. Hybrid porous solids: past, present, future. *Chem. Soc. Rev.* **37,** 191–214 (2008).
19. Kitagawa, S., Kitaura, R. & Noro, S. Functional porous coordination polymers. *Angew. Chem. Int. Ed.* **43,** 2334–2375 (2004).
20. Boerrigter, S. X. M. *et al.* MONTY: Monte Carlo crystal growth on any crystal structure in any crystallographic orientation; application to fats. *J. Phys. Chem. A* **108,** 5894–5902 (2004).
21. Blatov, V. A., Delgado-Friedrichs, O., O'Keeffe, M. & Proserpio, D. M. Three-periodic nets and tilings: natural tilings for nets. *Acta Crystallogr. A* **63,** 418–425 (2007).
22. Blatov, V. A., Shevchenko, A. P. & Proserpio, D. M. Applied topological analysis of crystal structures with the program package ToposPro. *Cryst. Growth Des.* **14,** 3576–3586 (2014).
23. Roeffaers, M. B. J. *et al.* Morphology of large ZSM-5 crystals unraveled by fluorescence microscopy. *J. Am. Chem. Soc.* **130,** 5763–5772 (2008).
24. Jeong, N. C., Lim, H., Cheong, H. & Yoon, K. B. Distribution pattern of length, length uniformity, and density of TiO$_3^{2-}$ quantum wires in an ETS-10 crystal revealed by laser-scanning confocal polarized micro-Raman spectroscopy. *Angew. Chem. Int. Ed.* **50,** 8697–8701 (2011).
25. Shtukenberg, A. G. *et al.* Dislocation-actuated growth and inhibition of hexagonal L-cystine crystallization at the molecular level. *Cryst. Growth Des.* **15,** 921–934 (2015).
26. Rimer, J. D. *et al.* Crystal growth inhibitors for the prevention of L-cystine kidney stones through molecular design. *Science* **330,** 337–341 (2010).
27. Teng, H. H., Dove, P. M. & DeYoreo, J. J. Reversed calcite morphologies induced by microscopic growth kinetics: insight into biomineralisation. *Geochim. Cosmochim. Acta* **63,** 2507–2512 (1999).

METHODS

Growth model and adaptation to nanoporous materials through three-dimensional tiling. The growth model is adapted from ref. 20: if units of growth can be identified, then the probability for growth relative to the probability of dissolution depends only on the stabilization gained by transferring that unit of growth from solution to the crystal. As the crystal grows, there will be many sites s at which the unit of growth may attach and each of these sites will have its own stabilization energy ΔU_s. The solution is considered to act only as a driving force with a chemical potential $\Delta \mu$, without regard for speciation in solution. At equilibrium, that driving force will be such that the rate of growth and the rate of dissolution are equal. There are three principal approximations[20]: (i) that the processes of growth and dissolution are thermodynamically reversible; (ii) that upon attachment of a unit of growth to the crystal the energy lost through desolvation is proportional to the energy gained through crystal attachment; and (iii) that the entropy change of the growth unit is the same regardless of site type. The probabilities for growth (P_s^{growth}) and dissolution ($P_s^{\text{dissolution}}$) are then

$$P_s^{\text{growth}} = \exp\left[-0.5\left(\frac{\Delta U_s}{kT}\right) + 0.5\left(\frac{\Delta \mu}{kT}\right)\right]$$

$$P_s^{\text{dissolution}} = \exp\left[0.5\left(\frac{\Delta U_s}{kT}\right) - 0.5\left(\frac{\Delta \mu}{kT}\right)\right]$$

where the factor of 0.5 indicates no preference for growth versus dissolution.

In the above approach, the probabilities of growth and dissolution are given by the thermodynamics of the process based on units of growth. However, these quantities are also appropriate as proxies for rate constants in kinetic Monte Carlo simulations. The justification for this comes from the Bell–Evans–Polanyi principle, as widely used in the field of catalysis. For a series of closely related chemical processes that have similar transition states, the relative rates are approximately determined by a factor that depends only on the difference in thermodynamics for the corresponding reactions. Provided that a material is relatively homogeneous, the different pathways for addition of a unit of growth will be rather similar and likely to have common features to their transition states (for example, desolvation of the solution species and surface attachment point). Therefore, we argue that the Bell–Evans–Polanyi principle should be equally valid for crystallization processes as for catalysis. Given this, the probabilities of growth versus dissolution will determine the relative kinetics of crystal growth to within a time constant.

The two principal problems when considering framework crystals such as zeolites or MOFs is (i) to identify the units of growth and (ii) to determine in a simple manner the stabilization energies associated with these units of growth. The units of growth must be distinguished from the growth units. The latter requires full knowledge of the mechanism of growth and the speciation within the solution. The former is a suitable coarse-graining of the problem into rate-determining steps that, for nanoporous materials, are the metastable closed-cage structures. This coarse-graining must be sensible in terms of metastability and small enough to capture the essential growth features, but also large enough to ensure efficiency of calculation. This final provision can be illustrated by reference to MOFs, for which the crystal could be grown as a molecular crystal based on metal clusters, organic linkers and solvent, but as closed cages that are known to be metastable.

Closed cages are readily determined using the approach of three-dimensional tiling achieved by the ToposPro code[22], which computes natural tiles[21]. Of the 229 zeolite frameworks studied, 54 (less than one quarter) of all of the structure codes are constructed entirely from closed cages or tiles made up of Q^3 tetrahedrally coordinated atoms. The remaining 175 structure codes are constructed from tiles that include some Q^2 tetrahedrally coordinated atoms (so-called open cages or tiles). Of these 175 frameworks, 120 are composed entirely of open tiles and 55 include a mixture of open and closed tiles. Nonetheless, these tile types meet the coarse-graining criteria in terms of both metastability and size. Any tiling other than the natural tiling will be a summation of natural tiles and so, in a crystal growth model, will decrease the resolution of the coarse-graining unnecessarily. These larger tiles will themselves be created as the crystal grows. The natural tiles also automatically select tiles that are all Q^3 initially, followed by tiles with both Q^3 and Q^2 (no Q^1 or Q^0), and therefore represent the lowest-energy metastable states.

The relative ΔU_s values associated with these tiles is then necessary to compute probabilities for growth versus probabilities for dissolution. Absolute energies are not required. It is known from experiments[28,29] that the condensation free energy for the reaction

$$\text{Al(OH)}_4^- + \text{H}_4\text{SiO}_4^0(\text{aq}) \rightleftharpoons [(\text{OH})_3\text{Al}-\text{O}-\text{Si(OH)}_3]^- + \text{H}_2\text{O}$$

is $-4.2\,\text{kcal mol}^{-1}$ and so we can expect each condensation of a unit of growth to result in a change of about this size in relative ΔU_s value. Therefore, to a first approximation, the ΔU_s values for a particular tile will form a uniform ladder of

energies with spacing of a few kcal mol^{-1} (because the entropy is approximated to be the same for all sites in the crystal, $\Delta S_s = 0$ and $\Delta U_s \equiv \Delta G_s$, where ΔS_s and ΔG_s are the entropy and free energy changes from surface site to bulk solid respectively). The ladder will range from the tile in its native form consisting of Q^3 and Q^2 units for tetrahedral framework structures to the fully condensed all-Q^4 form. Different chemical condensations will necessarily change the energy spacing as will the stabilization of different cage types through hydration. As a consequence, each cage type can be given a different, independent, energy spacing. The same philosophy can be used for any framework type, whether tetrahedral or not, with the energies depending only on the degree of condensation of the tile. So, for example, octahedral frameworks, combinations of octahedral and tetrahedral frameworks, and any other combination are readily treated. An example of the ladders of ΔU_s values for the simple LTA system is given in Extended Data Fig. 1. For complex systems such as UOV, which has 16 tile types (Extended Data Fig. 3), the number of site types is very large (1,843); however, only a fraction of these are actually ever populated, in this case 168. All of the other tiles that are required for the calculations on systems presented here are given in Extended Data Figures 1–10.

The crystals are seeded by growing the first few units at high supersaturation and then dropping the supersaturation to that of the growing medium. The supersaturation level is easily determined because the simulation can be run to equilibrium (equal growth and dissolution rates) and then the supersaturation set relative to that equilibrium driving chemical potential. The seed is inserted manually because we are growing only one crystal. In the future, competitive growth of crystals, such as Ostwald ripening, could be considered through the inclusion of multiple nuclei.

The goal of this work is to find a generic algorithm that is able to readily describe all of the most important aspects of 'classical' crystal growth, across any crystal system. The trigger for this endeavour is that, with the advent of scanning probe techniques, in particular atomic force microscopy (AFM), the number of experimental parameters available for simulation is much larger than when the only parameter was the crystal habit. The question then arises of whether a model can be derived in which there are fewer parameters than there are experimental observations (an overdetermined model). For example, with zeolite A, the experimental observables are crystal morphology or aspect ratio, terrace morphology on three different facets under different supersaturation conditions, terrace density and terrace heights. This corresponds to about twelve observable parameters with which we are fitting three energy parameters. As a consequence, by performing hundreds of simulations we are able to pinpoint with quite a high degree of accuracy the free energies associated with the rate-determining steps in this process, for this system.

Now consider the UOV system. There are a similar number of experimental observables as for zeolite A; however, the potential number of energy parameters for the model is 16, making the problem underdetermined. An assumption therefore needs to be made. In this case, the assumption we make is that the free energy for condensation of the different cages per Si–O–Si condensation is the same, thereby reducing the number of parameters from 16 to 1. Despite this assumption, we obtain a remarkably good fit with experiments. We need to be aware of the assumption that we have made and predicate further conclusions with that in mind; however, because the mechanism of crystallization of UOV was completely unknown without these simulations, the gain in knowledge is substantial despite the assumption made. It should be noted that our approach cannot be used blindly and requires very careful consideration of any system that is studied. Nevertheless, it provides a route to quickly study the basics of crystal growth for any system, in which the number of fitting parameters will be similar to or less than the number of experimental variables. Peculiarities for a given system would need to be layered on top of this approach.

Different polymorphs can be readily treated because they just represent different crystalline networks and therefore a different three-dimensional partitioning of space. For example, the different zeolite structures in their purely siliceous forms are all polymorphs of silica. Although we have not addressed the possibility of predicting competition between polymorphs here, our approach could be extended: if the different structures are put on a relative scale, then the probability for nucleation of different polymorphs could be tested. This extension would not be straightforward because effecting nucleation via Monte Carlo techniques is technically demanding, owing to the propensity of the system to be perpetually dissolving.

Defects and stacking faults are very important and we believe our approach can be extended to include them. However, at this stage our focus was on keeping the algorithms generic and efficient. Defects and intergrowths are by nature system-specific; consequently, so far we have considered only the most common defect, the screw dislocation. Whether it will be more beneficial to add capability for handling other defects to our algorithms in a bespoke manner for each system

or to develop protocols for introducing some more common intergrowths and defects is yet to be determined.

Our technique can be used for co-crystals with the modification that the driving force for each component would need to be considered separately, although these driving forces would be linked by the stoichiometry of the growing crystal. In principle, any type of interaction can be considered because we are just fitting experimental data. If those interactions are to be computed *ab initio* for subsequent comparison, then different types of interaction may pose different levels of complexity.

In the case of the zeolite A system, the energy parameters were refined by performing more than one thousand calculations that help to pinpoint the solution (see Extended Data Fig. 1g). In this instance, the crystal features are sensitive to the energy parameters (as might be expected), which, consequently, enables the refinement to be precise. Precision will vary from system to system depending on the number of experimental variables versus the number of fitting parameters. Similarly, the rigour of the comparison will depend on what experimental data are available that can be compared. For example, measurement of supersaturation for a molecular crystal might be straightforward, but for a zeolite growing from a gel is extremely difficult. Nonetheless, the supersaturation can become a variable that can be addressed by comparing surface nucleation densities. Also, the rigour of the comparison required depends on the question being asked. For example, if the question is, "how can the aspect ratio of a unidimensional nanoporous system be changed so that the pore length is short, to improve diffusion to a catalytically active centre?" then what is important is, "which parameters influence the aspect ratio of the crystal?" For zeolite L, improving diffusion of a guest species to a catalytically active centre means stabilizing the large cages, which can be achieved through specific templating. To make such a determination, the specific free energies of condensation need to be known to an accuracy of only 0.5 kcal mol^{-1} or less. But it is crucial to know the general mechanism of crystal growth that our simulation provides. Most materials chemists consider the trends to be the most important thing in terms of influencing outcome. If the experimental dataset is large and the number of fitting parameters is small, as is the case for zeolite A, then the free energies of condensation can be determined to an accuracy of about 0.2 kcal mol^{-1}.

To further test our methodology we looked at crystal growth of the L-cystine system, which has been studied extensively by AFM[25,26] because of its involvement in the pathogenesis of cystine kidney stones. In previous work, detailed images were recorded of the development of complex screw dislocations emanating from the [001] facet, and a Hartman–Perdok analysis is able to describe many of the features observed. Using our approach we are similarly able to simulate all of the main features observed in the L-cystine system (see Extended Data Fig. 10), including the rounded and fractal features that a Hartman–Perdok analysis will not expose. By running multiple simulations in a similar manner to those performed for the zeolite A system, we are able to refine the free energies of crystallization of all of the principal interactions in L-cystine. For example, the two slow growth directions (labelled A+ and A− in refs 25 and 26) must differ by about 1–2 kcal mol^{-1} to explain the step bunches with a total height of one unit cell in the *c* direction (Extended Data Fig. 10b). If the difference between these growth energies is removed or reduced, the height of the step bunches is only half a unit cell in the *c* direction, owing to the symmetry that is repeated every three L-cystine layers. However, what is most interesting from this work is the separation between the step bunches, which changes markedly with either supersaturation or with changes in the four binding energies. This step-bunch separation is discussed at length in ref. 25, in which lattice strain is used to account for deviations between the separation expected from a Hartman–Perdok analysis and that observed in experiments. It is clear from our simulations that the rate of advancement of the slow steps that define the characteristic pin-wheel pattern do not result from binding at this step edge followed by kink growth. Their advancement is precipitated by nucleation and terrace growth on the side of the step bunches (⟨100⟩ facets) as well as surface nucleation. Because of the strong anisotropy of the binding energy in these crystals, nucleation at a ⟨100⟩ surface site is energetically more favourable than binding at a higher-coordination edge site of an individual slow step. Consequently, the rate of growth of the ⟨100⟩ facets and the slow steps is governed by the same process. This effect can be seen in supplementary video 1 of ref. 30, wherein the slow steps are seen to advance principally as a result of advancement of the step bunches; we reproduce this effect in Supplementary Video 7. Step advancement in the slow growth direction is further enhanced by birth-and-spread nucleation on the ⟨001⟩ facet. For this complicated system, being able to visualize the crystal growth in three dimensions helped to disentangle these competing processes. The outcome of a reversal of the handedness of the screw dislocation is shown in Extended Data Fig. 10c. Because the overall advancement of terraces is governed by the crystallography of the material rather than by the handedness of the screw

dislocation, the pattern of growth is remarkably similar, independently of the handedness of the screw dislocation. Indeed, far from the screw core it would be impossible to distinguish the handedness of the screw core. By contrast, near the screw core the terrace structure is necessarily different and is the distinguishing factor between opposite handedness. We simulated both L-cystine and D-cystine (Extended Data Fig. 10d) and, as expected, the presentations of the surface features have obvious opposite handedness. The free energies of binding that we determine directly by simulation of experiment can be compared with the value calculated directly from the experimental solubility of L-cystine[31].

***CrystalGrower* growth code.** The *CrystalGrower* growth code is written in Fortran 95. Input polyhedral units (tiles or Voronoi polyhedra) and their connectivities (partitioning) are provided via an interface with ToposPro[22]. ToposPro builds natural tiling or Voronoi partition in accordance with strict and unambiguous algorithms[21,22]. As a result, all of the information for a given polyhedral unit in terms of the initial vertex condensation is provided for *CrystalGrower*. Polyhedral units neighbouring through faces are identified because these are considered as possible growth sites. Polyhedral units neighbouring through edges or vertices are considered too unstable to act as sensible rate-determining units of growth. Further information that can be subsequently used to determine the consequences of condensation of polyhedral units on vertex condensation is also provided. All information is provided in *P*1 symmetry in the primitive unit cell. The low symmetry is necessary because crystal symmetry is broken at the crystal surface; the primitive cell is used to ensure the most efficient calculation. Equating the crystal network to a network of polyhedral units also provides an effective route for enumeration of screw dislocations. These are computed by dividing the network of polyhedral units by a plane up to the screw core, translating all of the polyhedral units on one side of the plane along the screw direction to an equivalent position, and then reconnecting the network. This procedure results in a new, fully connected network with no dangling bonds. Adjacency matrices for polyhedral units are computed for the screw and for the perfect structure such that look-up tables can be generated to maintain efficient crystal-growth computation even in the presence of such dislocations. The algorithm is completely generic and permits screw dislocations to be generated for any structure in any direction. The energetics of the screw dislocations are not determined and only some dislocations will be energetically feasible. *CrystalGrower* treats the problem simply as a network, but the stability of screw dislocation structures could be tested separately using energy-minimization methods[30].

The key to the program lies in efficient identification of all possible site types for growth; this is done once on initiation for both the perfect structure and any screw dislocations. These site types are not re-computed at each growth or dissolution iteration when an efficient algorithm permits identification of site-type changes as the crystal is modified. Probabilities for growth and dissolution are computed at each iteration as the number of changes of each site type and the driving potential $\Delta\mu$. The value of $\Delta\mu$ can be varied according to a number of protocols; however, most importantly it can be allowed to proceed asymptotically towards the equilibrium value $\Delta\mu_e$ by lowering $\Delta\mu$ for growth and raising $\Delta\mu$ for dissolution. When $\Delta\mu_e$ is found, the numbers of unknowns to be determined is reduced.

The code allows the energy ladders of ΔU_s values to be determined independently for each tile (Extended Data Fig. 1), and different weightings can be applied to different Q^n vertices. The code further permits the poisoning of sites to simulate the addition of growth modifiers. Also, effects of chemical ordering on ΔU_s values within the lattice (for example, Zn/P ordering in framework zinc phosphates or Si/Al ordering in high alumina zeolites) can be computed.

To treat molecular crystals or ionic crystals, site types are determined according to the number of Voronoi polyhedra rather than the number of tile vertices of type Q^n. The philosophy is that the molecule or ion sits within a Voronoi polyhedron built for its centre of mass and that the faces of the polyhedron represent the interactions with neighbours. Each face (or neighbour) can be given a weighting to represent its contribution to the ΔU_s values. All other computation is identical to that for framework crystals, and screw dislocations and growth modifiers may be added accordingly.

***CrystalGrower* visualization code.** The *CrystalGrower* visualization code is written primarily in C++ using OpenGL libraries, and calls to the Windows API. This code was designed to analyse the results of the kinetic Monte Carlo calculations performed by the *CrystalGrower* growth code and allow users to observe the morphology and surface structure of the grown crystal, while also manipulating the crystal in real time. Two information files are required to construct and display the polyhedral units that compose the crystal framework, both of which are generated during a simulation run by *CrystalGrower*.

Two distinct steps are taken to draw the crystal structure. The first, using the partition information file output by *CrystalGrower*, builds a model of each polyhedral unit using a small number of parameters for each polyhedral unit (such

as the number of vertices, the number of faces and whether the faces are flat or non-flat). This model will be a sphere in the case of molecular crystal. Once the parameters for each unit are calculated, the entire unit is stored as a memory object that allows quick, easy access from the graphics card. A series of spheres are also generated per unit for non-molecular crystals, with radii defined by the root-mean-square of the (x, y, z) coordinates of the polyhedral unit, multiplied by a scaling factor to allow changes to be made to the size of the sphere while visualizing the crystal structure as spheres. The second step uses another output file generated by *CrystalGrower*, which provides the visualization package with the (x, y, z) coordinates for each unit cell grown in a run, along with information relating to the type of unit to draw. Coupling these coordinates with those in the partition file, the polyhedral unit is drawn at the correct position to be visualized. After completing the drawing of all of the units for one particular type, the entire array of units for this type is then stored as a new memory object. Storing the results of this second step as another memory object is important because it greatly improves the speed of the program, allowing real-time manipulation of a large number of objects (about 300,000) on a graphics cards with greater than 2 GB of dedicated graphics memory.

Calcite calculations. The energetics for individual growth or dissolution steps were computed using a combination of interatomic potentials and a continuum solvation model. Here the COSMO solvation model[32] was used, with atomic radii fitted such that the experimental hydration free energies of calcium and carbonate were reproduced. A modified force field was then developed by refitting the calcium–carbonate interaction of an earlier model[33], such that the energy difference between calcite and the ions in aqueous solution was consistent with the experimental solubility. Starting from the optimized bulk structure of calcite commensurate with this model, a rhombohedral nanoparticle of calcite was cleaved with dimensions of $16 \times 16 \times 4$ atomic rows oriented such that the long edges run parallel to either the acute or the obtuse step-edge directions. The top layer of this particle was then reduced in size by two molecular layers along each edge to create an island on the surface. Using this structure as a starting point, with the bottom two layers of the slab held fixed during optimization to reproduce the effect of the calcite bulk, various mechanistic pathways for the growth or dissolution of the island were explored. This includes the removal of each of the four distinct ions from both the acute and obtuse step edges, as well as the corresponding additions of ions to each step.

Results from these calculations show that the site types for Ca^{2+} and CO_3^{2-} fall broadly into groups that depend on the connectivity of the ion within the lattice. The difference in reaction energy between the ion in solution and the ion in the crystal is 10–15 kcal mol^{-1} per ligand for both Ca^{2+} and CO_3^{2-}. To provide more accurate results for future studies it will be important to use explicit solvent molecules.
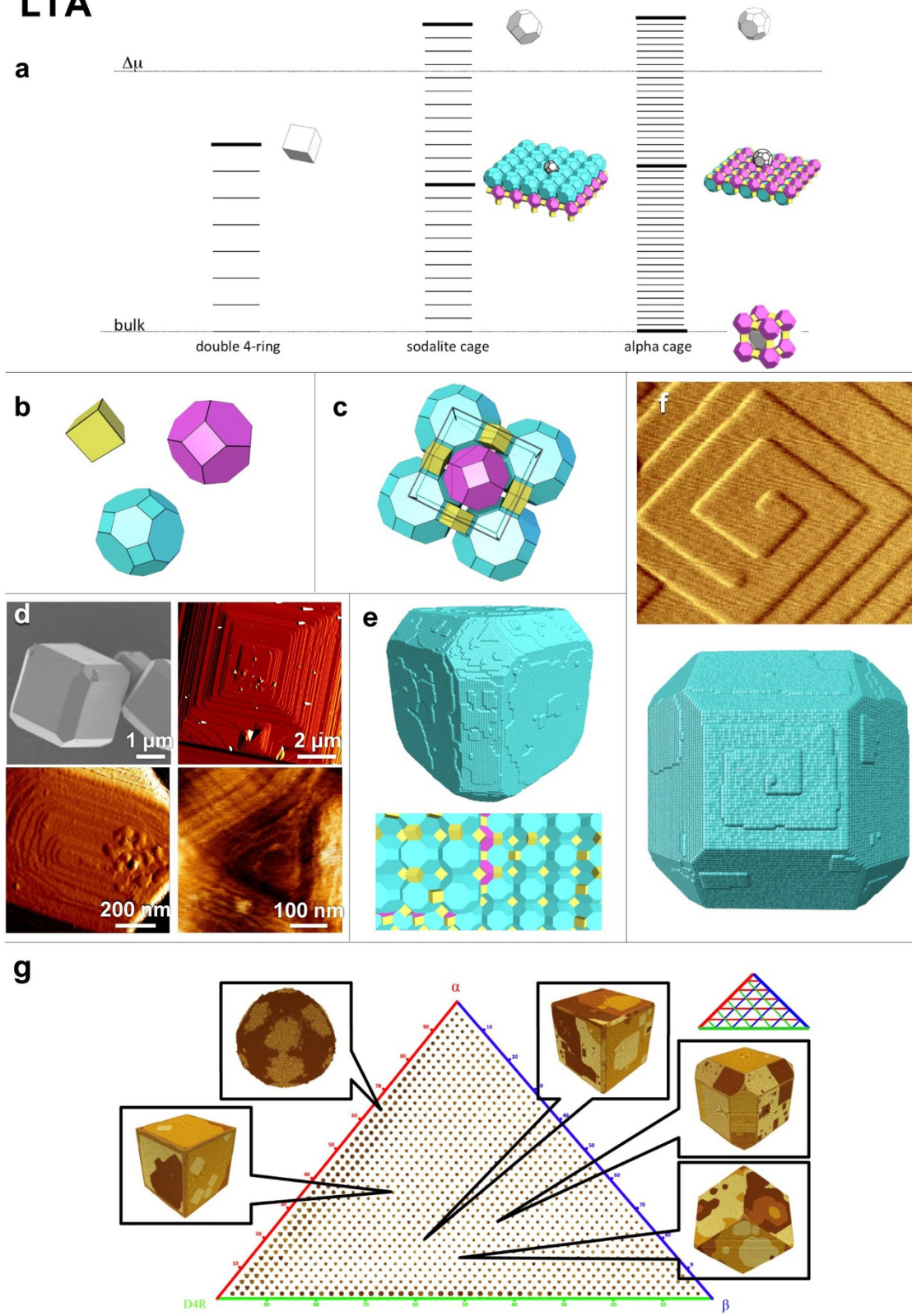
Other features that are revealed using *CrystalGrower*. Although the approach is principally designed to interrogate and simulate crystal surface structure and crystal habit, it is also able to probe internal defects that arise as a result of the growth mechanism. This is particularly relevant in the two structures MFI and ETS-10, both of which are known to exhibit birefringence when viewed under an optical or Raman microscope, respectively. It transpires from our initial studies using *CrystalGrower* that structures with open cages have a tendency to crystallize in a manner that leaves some tiles incomplete, resulting in a high number of internal silanol groups. MFI and ETS-10 (Extended Data Figs 5 and 6) both show internal defect density patterns that almost exactly mimic the birefringence observed experimentally. This is a much more appealing pathway to explaining such effects than that invoking complex twinning mechanisms. We will use this technique in the future to explore a range of framework structures that are known to exhibit such effects.

Experimental methods. AFM images were acquired using a JPK instruments Nano Wizard II. Images were taken in contact mode. Standard silicon nitride tips (NP Bruker Probes), with a nominal radius of 20 nm and a nominal spring constant of 0.58 N m^{-1}, were used.

Data availability. The data that support the findings of this study are available from M.W.A. and A.R.H. (adam.hill@manchester.ac.uk) on reasonable request.

28. Pokrovski, G. S., Schott, J., Salvi, S., Gout, R. & Kubicki, J. D. Structure and stability of aluminum-silica complexes in neutral to basic solutions. Experimental study and molecular orbital calculations. *Min. Mag.* **62A,** 1194–1195 (1998).
29. Tagirov, B., Schott, J., Harrichoury, J.-C. & Escalier, J. Experimental study of the stability of aluminate-borate complexes in hydrothermal solutions. *Geochim. Cosmochim. Acta* **68,** 1333–1345 (2004).
30. Walker, A. M., Slater, B., Gale, J. D. & Wright, K. Predicting the structure of screw dislocations in nanoporous materials. *Nat. Mater.* **3,** 715–720 (2004).
31. Carta, R. & Tola, G. Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations. *J. Chem. Eng. Data* **41,** 414–417 (1996).
32. Klamt, A. & Schüürmann, G. COSMO — a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* 799–805 (1993).
33. Raiteri, P., Demichelis, R. & Gale, J. D. A thermodynamically consistent force field for molecular dynamics simulations of alkaline-earth carbonates and their aqueous speciation. *J. Phys. Chem. C* **119,** 24447–24458 (2015).

LTA

a

Δμ

bulk

double 4-ring

sodalite cage

alpha cage

b

c

f

d

1 μm

2 μm

200 nm

100 nm

e

g

α

D4R

β

Extended Data Figure 1 | See next page for caption.

Extended Data Figure 1 | Experimental and simulated data for the LTA framework with energy level diagram. a, A schematic of energy ladders for the LTA structure. Double 4-rings (D4Rs) are shown in yellow, sodalite (β) cages are shown in magenta and α cages are shown in cyan. Each level on the energy ladders, moving from bulk to the top, corresponds to the loss of coordination (or gain in hydration) of one Q^4 site to a Q^3 site. In the case of a double 4-ring, the bottom level corresponds to a cage with zero Q^3 sites and eight Q^4 sites, whereas the top level corresponds to a fully solvated double 4-ring with zero Q^4 sites and eight Q^3 sites. The driving force $\Delta\mu$, which is the chemical potential of the solution, can be considered as a single value for a zeolite in which the nutrient is interconverting. However, for a system with species that do not interconvert, such as a co-crystal, more than one driving force would be required—although these driving forces would be interrelated by the stoichiometry of solution species and their relative rates of consumption. **b–f**, The LTA structure consists of three tiles (**b**) connected in a cubic lattice (**c**). Typical crystals exhibit three facets, each with a unique terrace shape (**d**), all of which can be reproduced along with the crystal habit in a simulation that treats the crystallization energy of each tile separately (**e**, top); a prediction of the surface termination on (100) can also be made (**e**, bottom). Screw dislocations along [100] in the LTA structure are ubiquitous and the same set of parameters also reproduces the nature of this defect (**f**). Supplementary Video 1 shows the screw dislocation growing through the simulation. The simulated crystals shown in **e** and **f** are approximately $0.15\,\mu\text{m} \times 0.15\,\mu\text{m} \times 0.15\,\mu\text{m}$ and $0.20\,\mu\text{m} \times 0.15\,\mu\text{m} \times 0.15\,\mu\text{m}$ in size, respectively; $\Delta U_s(\alpha) = 1\,\text{kcal mol}^{-1}$, $\Delta U_s(\beta) = 5.4\,\text{kcal mol}^{-1}$, $\Delta U_s(\text{D4R}) = 8\,\text{kcal mol}^{-1}$. **g**, A ternary plot consisting of 1,176 simulation data points, corresponding to exploring the energy space at 2% intervals. All images are recorded at equilibrium driving force. The highest destabilization for each cage is at the corner at which its name appears, and the corresponding axes are colour coded. The directions of the grid lines for each axis can be seen in the small diagram top right. Examples of interesting morphologies for particular energy combinations are highlighted.

FAU
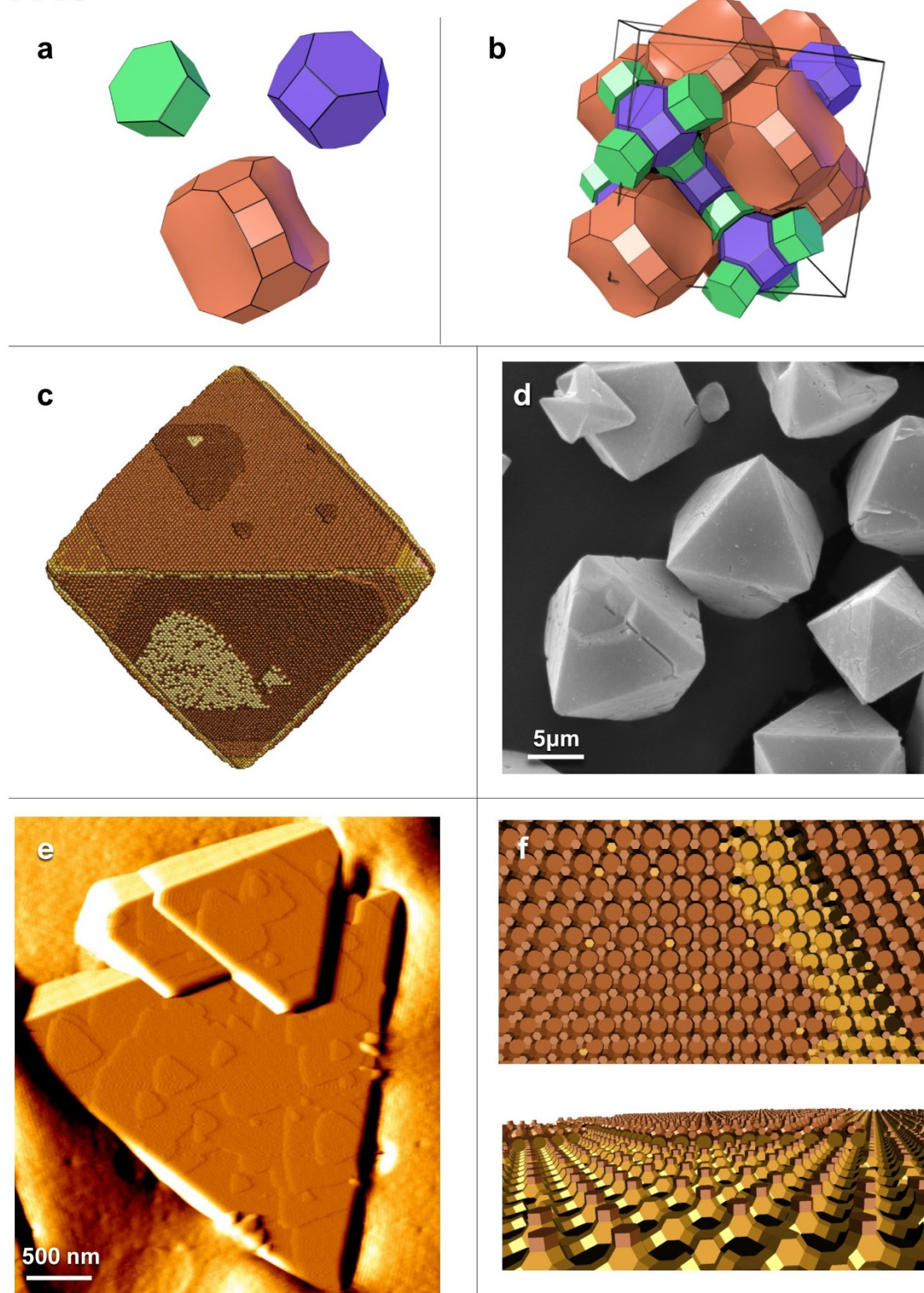
Extended Data Figure 2 | Experimental and simulated data for the FAU framework. a, **b**, The FAU structure consists of three tiles (**a**) connected in a face-centred-cubic lattice (**b**). **c**–**e**, A simulation in which the supersaturation is decreased from high (50 kcal mol^{-1}) to low (1 kcal mol^{-1}) at the halfway point, with $\Delta U_s = 2$ kcal mol^{-1} (**c**), matches scanning electron microscopy (SEM; **d**) and AFM (**e**) data on a zinc phosphate structure, in terms of both the octahedral habit and the triangular terraces on the (111) face with correct orientation. **f**, Simulations with the same parameters as for **c** with varying zoom distances show the surface structure predicted for this simulation. Supplementary Video 2 shows an example of how a FAU crystal grows at both high and low supersaturation. The simulated crystal shown in **c** and **f** has an approximate size of 0.20 μm × 0.20 μm × 0.20 μm.

UOV
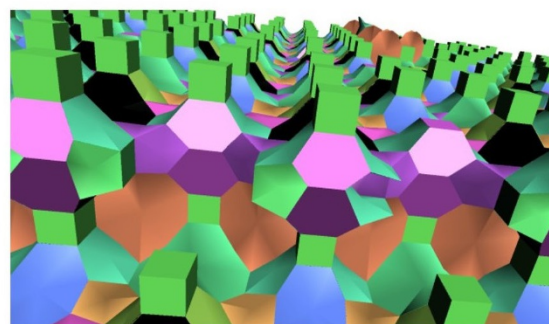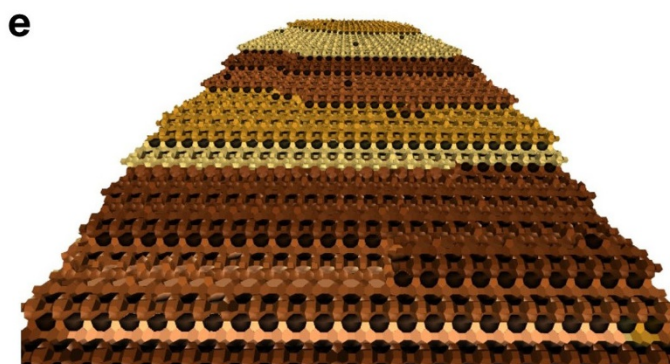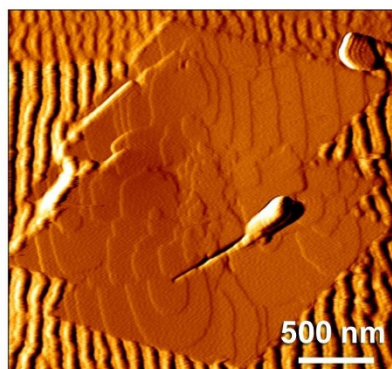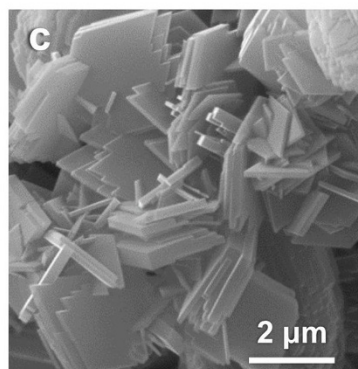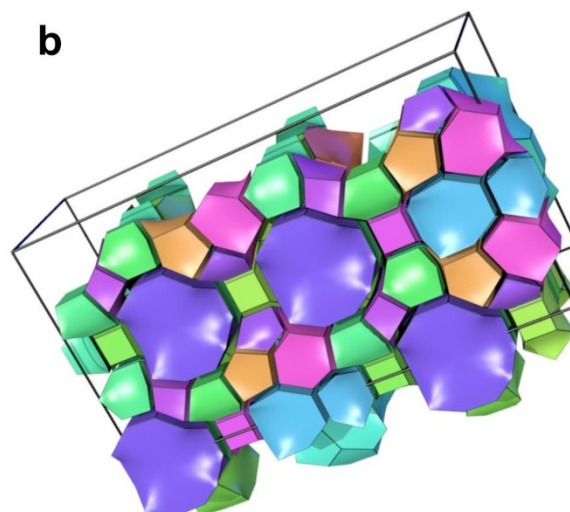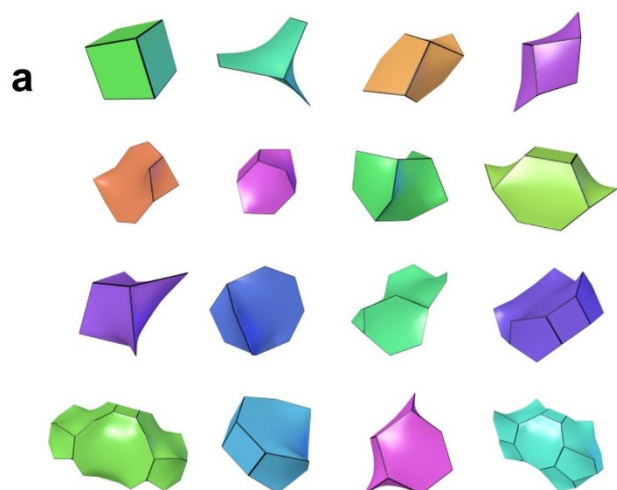
Extended Data Figure 3 | Experimental and simulated data for the UOV framework. a, **b**, The UOV structure represents one of the most complex framework systems, with a particularly large unit cell with dimensions of $1.27\,nm \times 2.20\,nm \times 3.88\,nm$, and consists of 16 different tile types (**a**) connected in an orthorhombic lattice (**b**). **c**, The germanosilicate crystallites that adopt this framework structure, although intergrown, show very well defined shape bounded by six faces. **d**, **e**, Simulation confirms the two large faces to be {100} (**d**) and the four side faces to be {013} (**e**, left). Simulations also reveal the structure at the crystal surface on the side wall (**e**, left) and top face (**e**, right), with differing colouring methods; $\Delta U_s = 2\,kcal\,mol^{-1}$. Supplementary Video 3 shows an example of how a UOV crystal grows at both high and low supersaturation values. The simulated crystal shown in all of these examples is approximately $0.050\,\mu m \times 0.75\,\mu m \times 1.40\,\mu m$ in size.

NES

a

b

c

500 nm

500 nm

2 µm

d

e
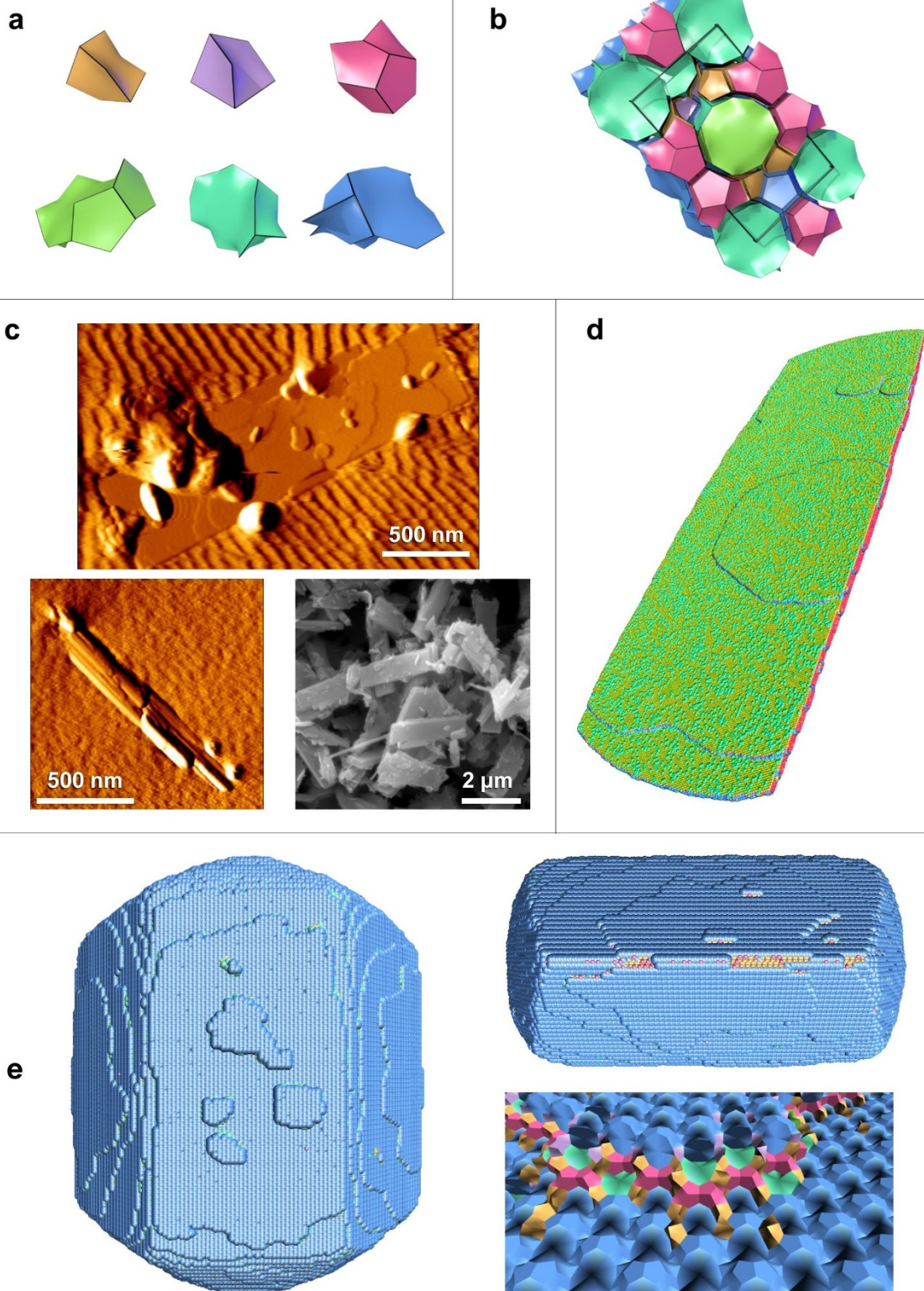
Extended Data Figure 4 | See next page for caption.

Extended Data Figure 4 | Experimental and simulated data for the NES framework. a, b, The NES framework can usually be expressed with four different tile types. The tile with the largest number of vertices, 't-nes', can be split into three smaller tiles (green, cyan and blue) to give a total of six tiles of similar size (**a**) that combine to form an orthorhombic lattice (**b**). These three tiles can no longer be referred to as natural tiles, but they were still generated using the ToposPro software. These tiles play the part of units of growth and are required because *CrystalGrower* utilizes connections through faces, making frameworks that share only corners and edges between small tiles, such as NES, difficult to grow over reasonable timescales. **c, d,** AFM and SEM micrographs show the morphology of aluminosilicate crystals that adopt this framework structure to be long and wafer-like with rounded terraces (**c**), which can be reproduced with simulations using the six smaller tiles chosen (**d**); $\Delta U_s = 2$ kcal mol^{-1}. **e,** Lowering the energy penalty for the largest cage, $\Delta U_s = 1$ kcal mol^{-1} compared to other cages at 3 kcal mol^{-1}, changes the morphology of the crystal markedly, thickening the crystal considerably (left and top right). This is an observable example of a structure being constrained in its propagation in a particular direction owing to the difficulty in growing such large cages. Investigating the surface structure (bottom right) shows that the surface is almost entirely terminated by the largest tile (blue) in the framework, again highlighting this observation. The simulated crystal shown in **d** is approximately 0.05 μm × 0.40 μm × 0.15 μm, whereas the example in **e** is estimated to be 0.10 μm × 0.20 μm × 0.15 μm in size.

MFI
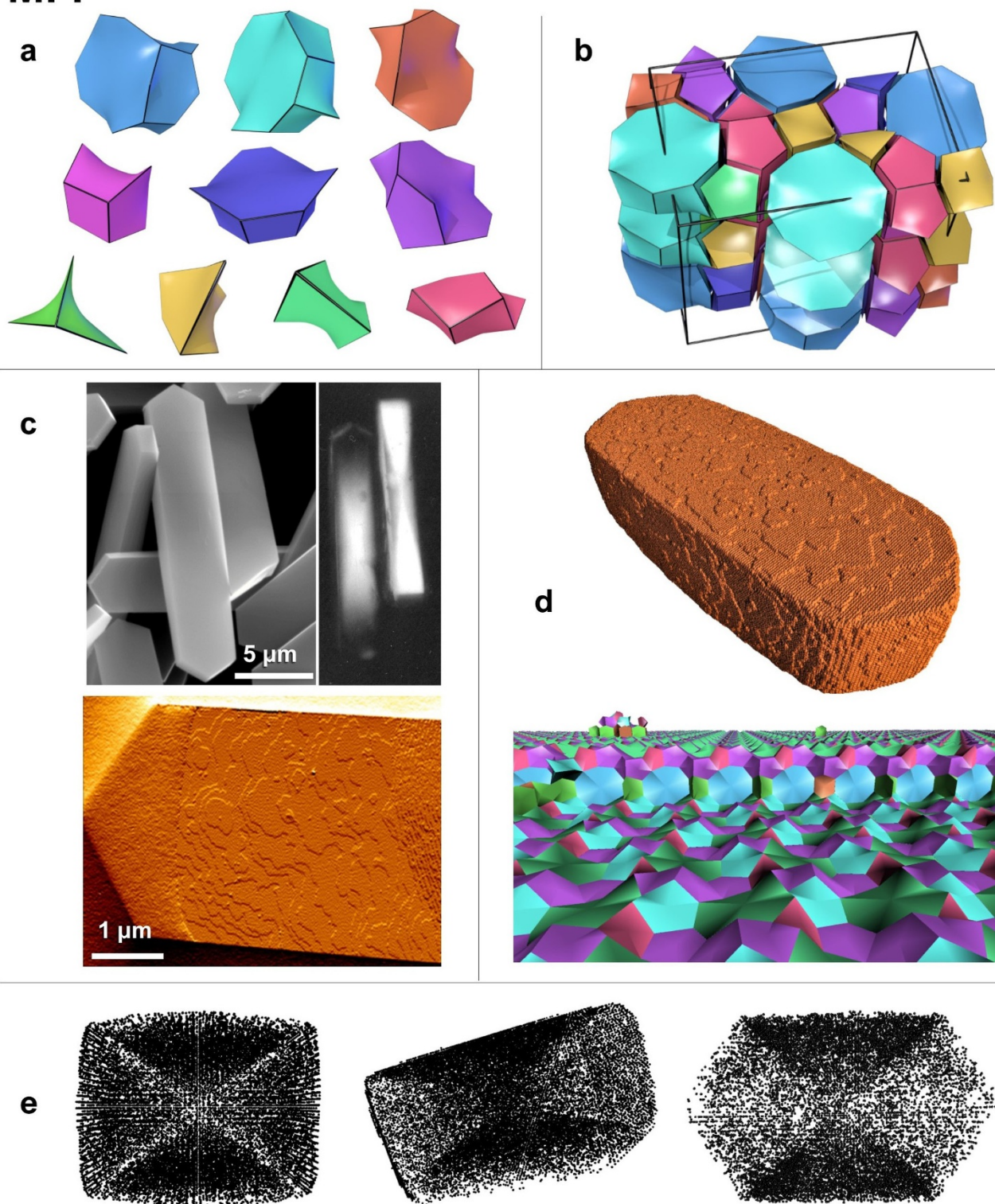
Extended Data Figure 5 | Experimental and simulated data for the MFI framework. a, b, The MFI structure consists of ten tile types (**a**) connected in an orthorhombic lattice (**b**). **c,** A consistent feature in MFI crystals is the optical hourglass effect seen in crystals of silicalite. There has been much conjecture about the origin of this hourglass effect. However, the fact that the optical birefringence, seen in these 300-μm-long crystals, is located in sectors bounded by the crystal faces is a good indication that the origin is a result of the mechanism for crystal growth being different on different faces of the crystal. **d,** Our simulations reflect well the crystal habit and surface topology and allow investigation of the surface termination; $\Delta U_s = 2$ kcal mol^{-1} for small tiles and $\Delta U_s = 4$ kcal mol^{-1} for large tiles.

e, The internal structure shows that the density of silanol groups mirrors very closely the sectoring of the crystal, because only tiles with incomplete coordination are shown. Such a change in crystal chemistry would, almost certainly, be associated with a change in the refractive index of the crystal and hence the observed optical effect. Supplementary Video 4 demonstrates growth of the hourglass feature as the simulation progresses and shows how different tiles express this feature to differing degrees. The simulated crystal shown in **d** is approximately 0.10 μm × 0.05 μm × 0.40 μm in size, whereas the crystal shown in **e** is approximately 0.10 μm × 0.05 μm × 0.20 μm.

ETS-10

**a**

b

c

2 nm

d

e

f
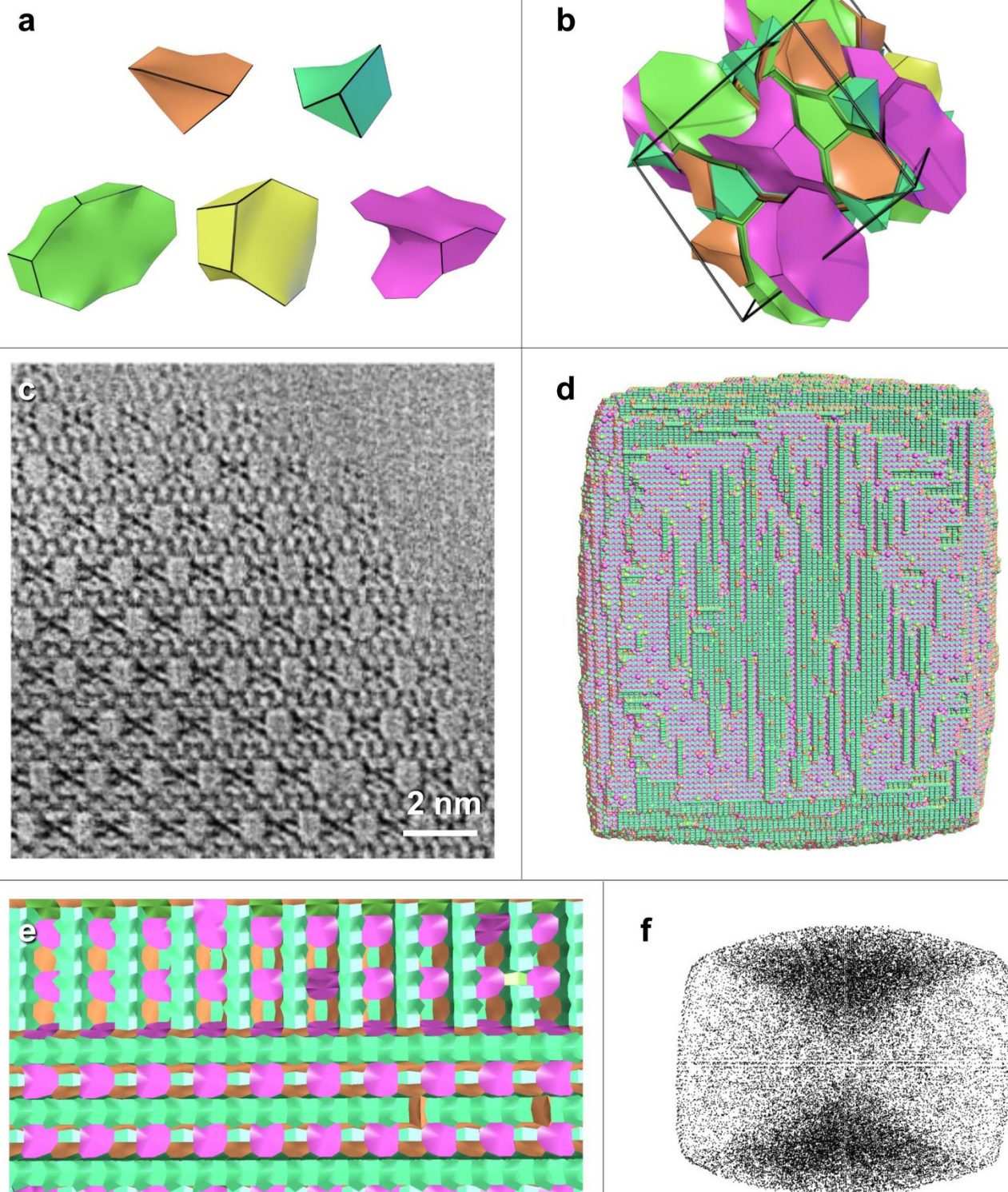
Extended Data Figure 6 | Experimental and simulated data for the titanosilicate material ETS-10. a, b, ETS-10 is a nanoporous titanosilicate with a structure consisting of five tiles (**a**) connected in a monoclinic lattice (**b**). **c,** The structure has a very similar symmetry to zeolite β; however, the titanium is incorporated into titania rods surrounded by silica. High-resolution electron microscopy courtesy of O. Terasaki. **d, e,** These rods run alternately in orthogonal [110] and [Ī10] directions and are found to be the dominant components for the crystal growth. **f,** Because the rods do not always connect and heal inside the crystal structure, there is a high concentration of internal silanol groups in a sector from the (001) faces to the centre of the crystal. Such sectoring in ETS-10 has been observed in Raman microscopy[24]. Supplementary Video 5 demonstrates how the rods in the ETS-10 framework grow in alternating orthogonal directions. The simulated crystals shown in **d** and **f** are approximately $0.25\,\mu m \times 0.25\,\mu m \times 0.05\,\mu m$ and $0.10\,\mu m \times 0.10\,\mu m \times 0.10\,\mu m$ in size, respectively; $\Delta U_s = 2\,kcal\,mol^{-1}$.

CHA
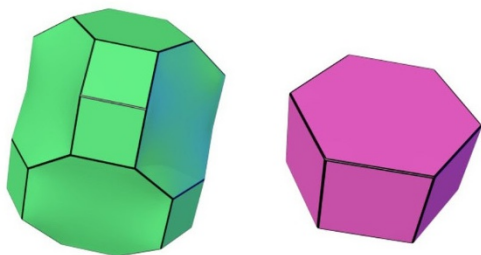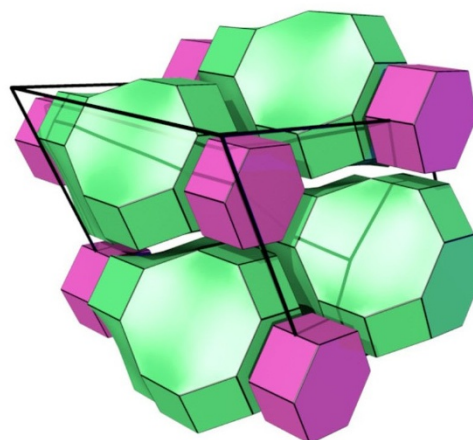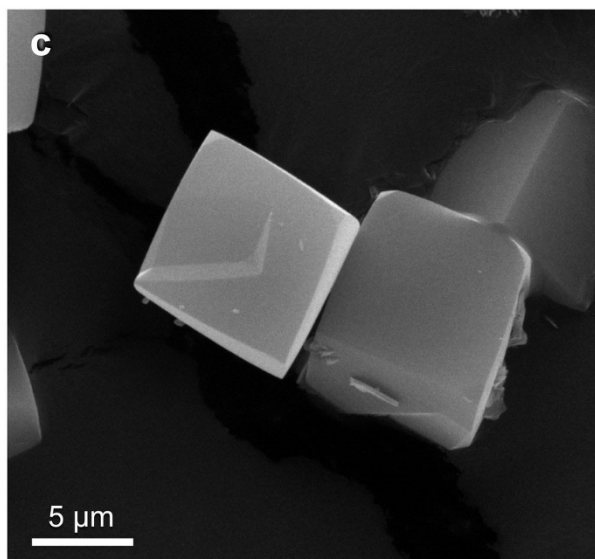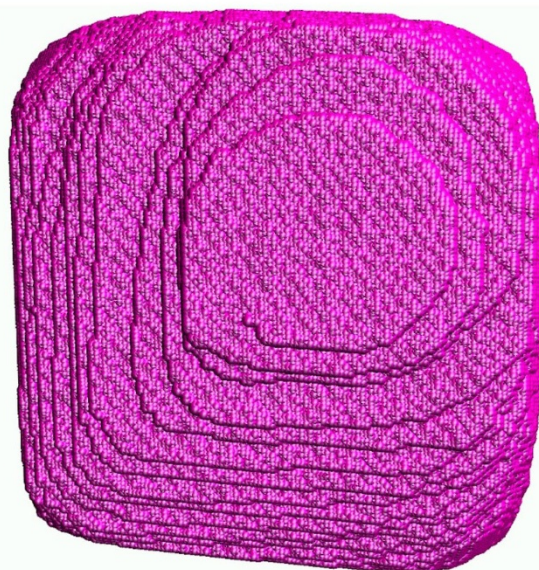
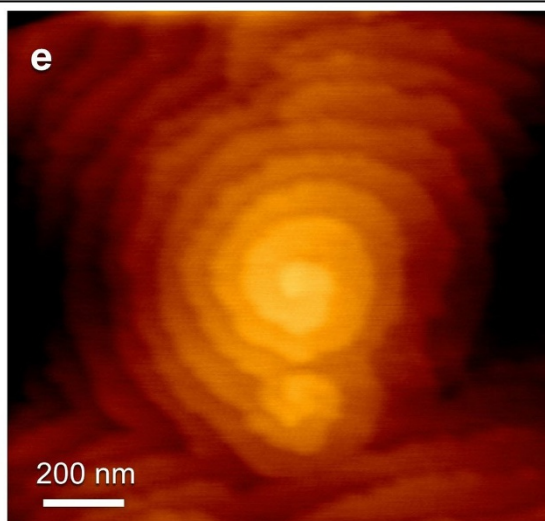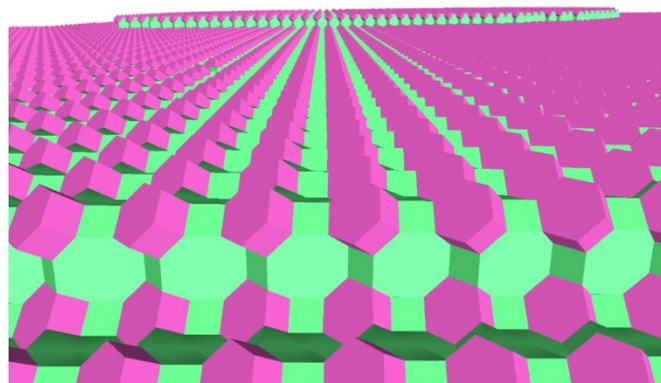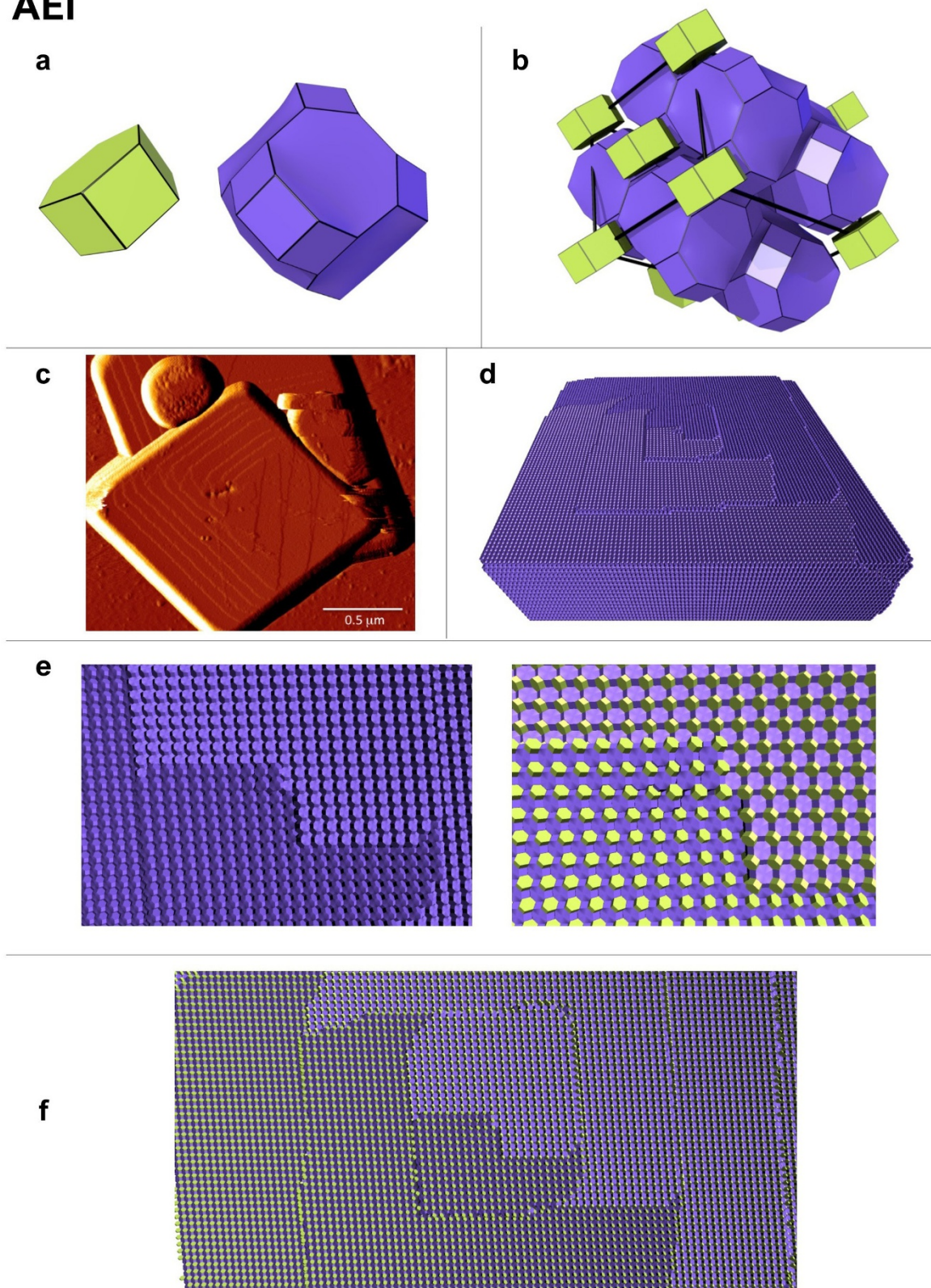**Extended Data Figure 7 | Experimental and simulated data for the CHA framework. a, b,** The CHA structure consists of two tiles (**a**) connected in a monoclinic lattice (**b**). **c,** Typical silicoaluminophosphate (SAPO-34) crystals show distorted cube morphology when viewed using SEM, exhibiting {100} faces with isotropic growth of terraces. **d, e,** Many screw dislocations can be observed in this system (**e**) and one such dislocation with a screw core running along [100] is reproduced by simulations (**d**); $\Delta U_s = 2\,\text{kcal mol}^{-1}$. **f,** Simulations predict that the surface terminates at the double 6-rings. The simulated crystal shown in **d** has an approximate size of $1.50\,\mu\text{m} \times 0.10\,\mu\text{m} \times 0.10\,\mu\text{m}$ due to elongation caused by the screw dislocation.

AEI

a

b

c

0.5 μm

d

e

f

Extended Data Figure 8 | Experimental and simulated data for the AEI framework. a, b, The AEI structure consists of two tiles (**a**) connected in an orthorhombic lattice (**b**). **c,** The structure is related to CHA in that it can be built by connecting double 6-rings. However, because of the alternating orientation of these units along the *c* direction, the fast and slow growth rates switch along the *a* and *b* directions (shown on AFM of aluminophosphate (AlPO-18). **d–f,** The result is complex, interleaved spiral growth that is faithfully simulated using our space partitioning methodology; $\Delta U_s = 2$ kcal mol^{-1}. The simulated crystal shown in **d–f** is approximately $1.70\,\mu m \times 0.05\,\mu m \times 0.05\,\mu m$ in size due to elongation caused by the screw dislocation.
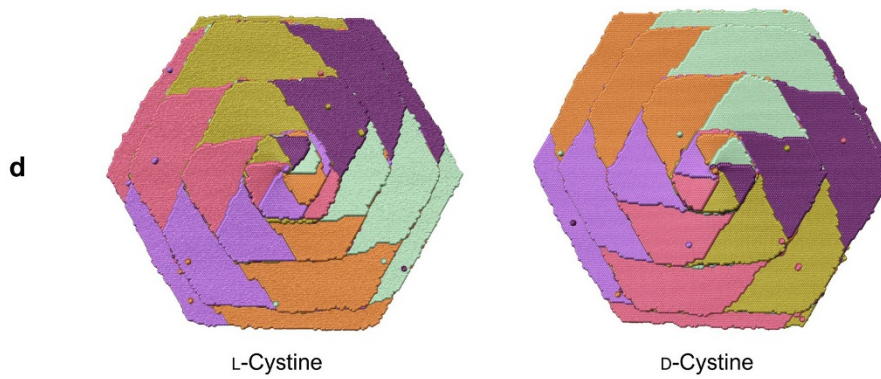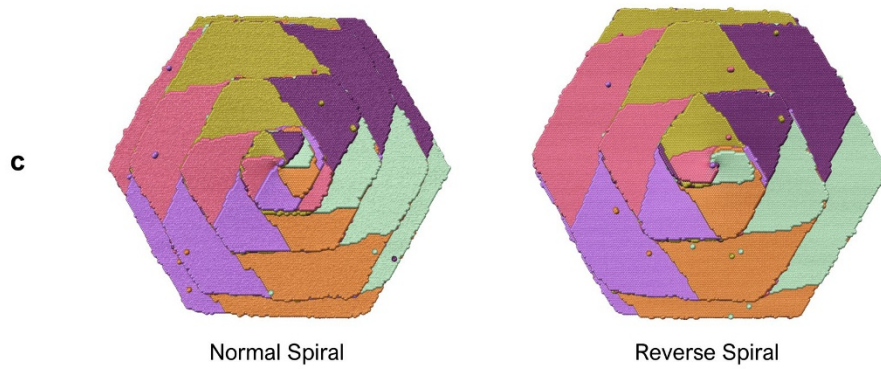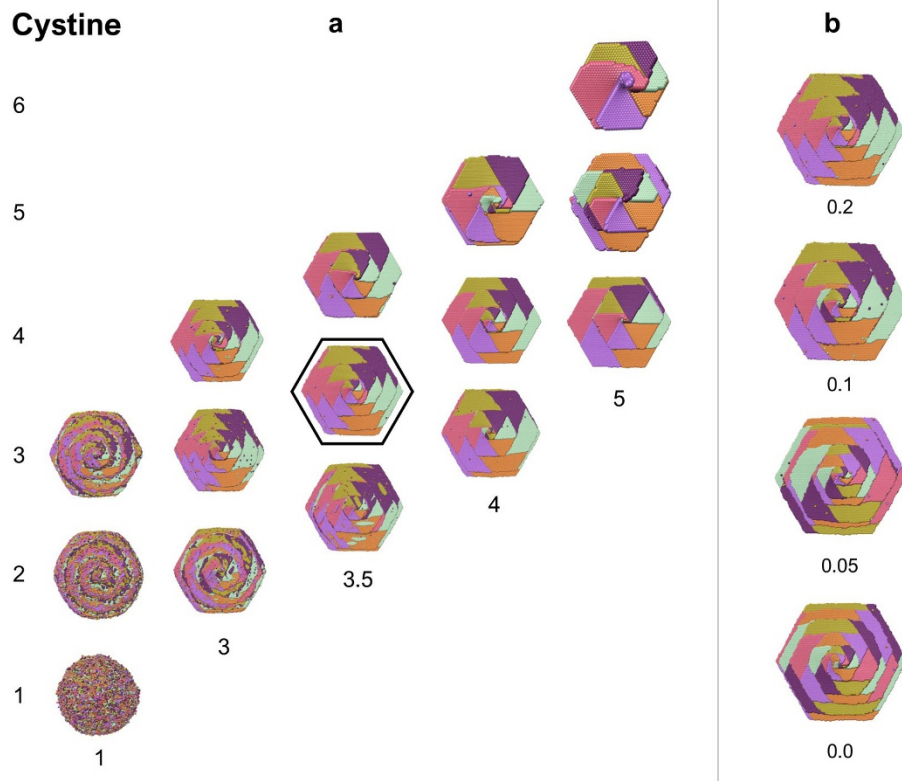
HKUST-1
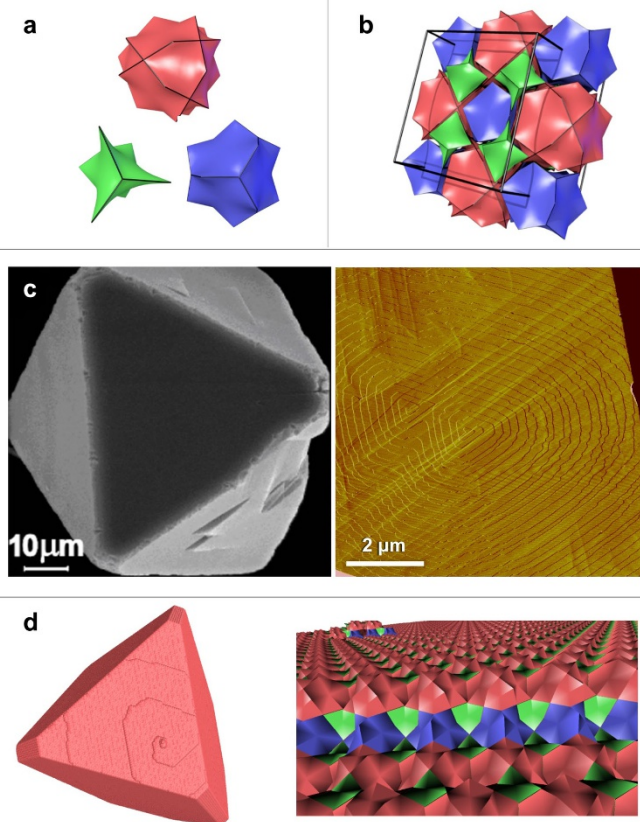
MOF-5

Extended Data Figure 9 | Experimental and simulated data for two metal–organic frameworks, HKUST-1 and MOF-5. a–d, HKUST-1, a porous MOF of copper trimesate, $Cu_3[(O_2C)_3C_6H_3]_2(H_2O)_3$. By partitioning the structure with nodes at metal clusters and tile edges along linkers, the HKUST-1 structure consists of three tiles (**a**) connected in a face-centred cubic lattice (**b**). Such a partitioning produces simulation results (**d**) that are in excellent agreement with experiments in terms of both the crystal habit and the terrace topology of the prevalent screw dislocations (**c**). The simulation predicts that the screw core runs along [110], not perpendicular to the {111} faces because such a screw core would result in multiple spirals (**d**); $\Delta U_s = 2$ kcal mol^{-1} for small tiles and $\Delta U_s = 5$ kcal mol^{-1} for large tiles. Simulations also allow investigation of the surface termination for this framework (**d**). Supplementary Video 6 shows the screw dislocation growing and migrating across the surface of the crystal. The approximate size of the crystal shown in **d** is 0.35 μm × 0.35 μm × 0.35 μm. **e–i**, MOF-5 (Zn$_4$O(bdc)$_3$, where bdc is 1,4-benzenedicarboxylate), a porous MOF with a simple cubic arrangement of metal centres (red) and linkers (grey), along with a solvent molecule occupying the centre of each cube (purple) (**e**, left). Treating the structure as a multicomponent molecular crystal in lieu of a standard
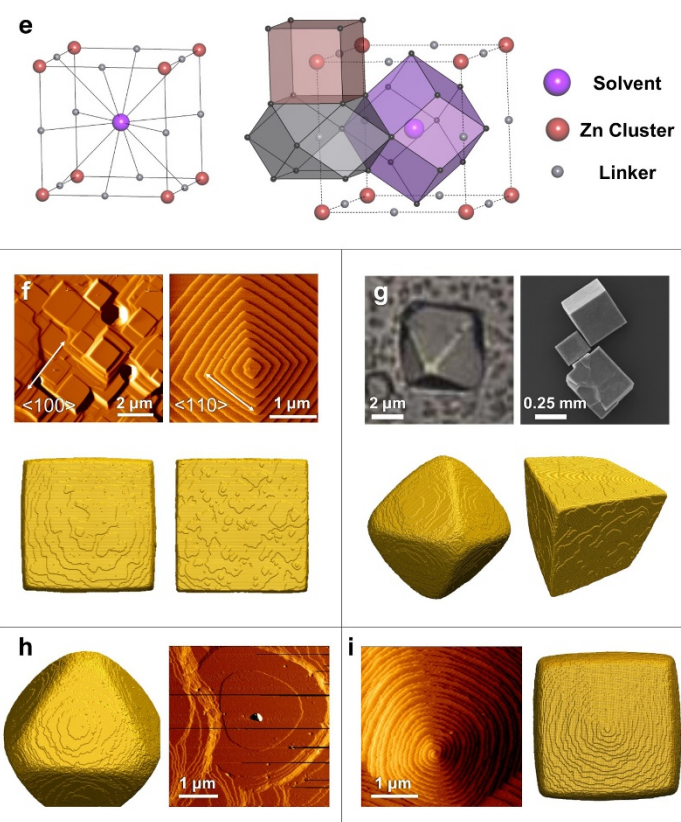
tile-partition as used in other cases, the framework can also be partitioned into Voronoi polyhedra to demonstrate the interactions between each of the linkers, metal centres and solvent molecules (**e**, right). Square terracing on the (100) face (**f**) can be in one of two orientations, depending on the synthesis conditions. The crystal morphology can also be changed to exhibit {111} faces as an octahedral crystal, again depending on synthesis conditions (**g**, left), as opposed to only {100} faces (**g**, right) in a cubic crystal. {111} faces exhibited by the crystal are much more isotropic (**h**) compared to the {100} faces (**f**) and do not adopt the different orientations demonstrated by the {100} faces. The simple cubic network alone will not permit the development of {111} faces and is immediately an indication of the importance of the solvent N,N′-diethylformamide (DEF) in the preparation. By varying the driving forces for binding of the linker, metal centre and solvent in the final crystal, all crystal habits and surface topologies may be generated (**f–h**). Double spirals through growth at screw dislocations on the (100) face can also be simulated (**i**). The estimated sizes of the crystals shown are: **f**, 0.76 μm × 0.75 μm × 0.75 μm and 0.70 μm × 0.80 μm × 0.80 μm; **g**, 0.50 μm × 0.50 μm × 0.50 μm and 0.75 μm × 0.75 μm × 0.75 μm; **h**, 0.50 μm × 0.50 μm × 0.50 μm; and **i**, 0.75 μm × 0.75 μm × 0.75 μm.

Extended Data Figure 10 | See next page for caption.

Extended Data Figure 10 | Collection of simulated data for L-cystine along with a single simulation of D-cystine. a, A suite of calculations of the crystal growth of L-cystine at a supersaturation of 0.6 kcal mol^{-1} viewed down the [001] direction, showing characteristic pin-wheel surface topology and hexagonal step bunching. Simulations are based on four characteristic free energies of crystallization: (i) in the strong binding direction where L-cystine molecules are bound by two hydrogen bonds and S….S contacts; (ii) and (iii) in the weak, slow growth directions, labelled A+ and A− in refs 25 and 26, with one hydrogen bond; and (iv) in the c direction with two hydrogen bonds. Interaction types (ii) and (iii) are maintained at half the strength of (iv), with a minimal difference between (ii) and (iii). These interactions can be combined into 491 growth sites that are computed in the simulation. The grid explores the effect of changing the relative magnitude of interaction (i) (number on the x axis), relative to the sum of the other interactions (number on the y axis); all values given in kcal mol^{-1}. When all of these interactions are too weak, isotropic sticky growth is observed; when they are too high, the growth is too geometrical. The optimum balance is highlighted (black outline), for which the sum of the free energies of crystallization is 7.0 kcal mol^{-1}, compared to the value of 4.3 kcal mol^{-1} derived from the solubility of L-cystine at pH 7

directly[31]. Similar to the results of ref. 25, the step bunching is closer than observed experimentally because lattice strain at the screw core weakens interactions and prevents the core region, which is advancing by attachment to the slow step alone, from forming a closed hexagonal step bunch as quickly as predicted. **b**, A set of simulations demonstrating the effect of small differences in the energy of slow growth directions A+ and A−; energies given in kcal mol^{-1}. If there were no difference in energy, then the pin-wheel structure would have a step bunch height of only half a unit cell, owing to the symmetry that is produced every 180° turn around the screw core. A difference of 0.1 kcal mol^{-1} is sufficient to eradicate this symmetry and make the step bunches one unit cell high. **c**, Reversal of the handedness of the screw core is not immediately apparent because the pin-wheel structure and handedness remain the same whatever the handedness of the screw core as this is driven by the symmetry of the crystal structure at the molecular level. However, differences can be observed in the surface topology close to the screw core, depending on whether the handedness of the crystal structure is the same or the reverse of the handedness of the screw core. **d**, L-cystine and D-cystine present identical surface structures with opposite handedness, as expected.