Predicting crystal growth via a unified kinetic 3-D partition model

Michael W. Anderson[1], James T. Gebbie[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 the course of crystal growth is fundamental to the control of functionality in modern materials. Despite investigations for over one hundred years[1-5] it is only recently that the molecular intricacies of these processes have been revealed by scanning probe microscopies[6-8]. In order to bring some order and understanding to this vast amount of new information requires new rules to be developed and tested. To date, because of the complexity and variety of different crystal systems, this has relied on developing models that are usually constrained to one system only[9-11]. Such work is painstakingly slow and will not be able to achieve the wide scope of understanding in order to create a unified model across crystal types and crystal structures. Here we describe a new approach to understand and, in theory, predict the growth of crystals, including the incorporation of defect structures, by simultaneous molecular-scale simulation of crystal habit and surface topology using a unified kinetic 3-D partition model. We exemplify 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 we have the possibility to control crystal habit, crystal size, the elimination or incorporation of defects and the development of intergrowth structures. As crystals are used in technologies from pharmaceuticals to gas storage and separation materials, from optoelectronic devices to heterogeneous catalysts, such understanding is vital. If we take an example of a very complex and yet very important crystal type, that of zeolites[12] which form the backbone of the heterogeneous catalysis industry, then many of the problems that must be addressed in crystal growth can be illustrated. Zeolites are nanoporous materials were 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 $230^{\circ}C$ and it is well known from NMR 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 even the hundreds of different zeolite structures, let alone the thousands of related crystal structures such as metal organic frameworks (MOFs)[17-19]. However, we know that the course of a

[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, Ac. Pavlov St. 1, Samara 443011, Russia.
[3]School of Materials Science and Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, People's Republic of China.
[4]Università degli Studi di Milano, Dipartimento di Chimica, Via C. 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 6845, Western Australia.
[†] Current address: Department of Chemistry, Durham University, Lower Mountjoy, South Road, Durham, DH1 3LE, UK.

crystallisation 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 through the work of Boerrigter et al. on general Monte Carlo simulation applied to the growth of fats[20]. In this work they showed that the principal determinant 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 *etc.* – to be considered only to have a growth potential rather than considering the speciation in detail. Of course this growth potential will be a result of the speciation, however, that can be treated as a subsidiary effect considered subsequently[13-16]. In the case of multicomponent crystals, such as MOFs or co-crystals, where species in solution, e.g. linkers and metal centres, cannot interchange then a driving force for each component needs to be considered – unless the stoichiometries of the two phases are matched. For zeolites, or any system where the nutrient is interconverting, a single driving force may 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. In order to deconstruct the problem, we require a distinction between *unit of growth* and *growth unit*. The latter assumes that we know the growth mechanism and know what growth units are actually attaching to the crystal surface. If the material of interest is a molecular crystal, then a *unit of growth* would naturally be a single molecule as this represents a strongly bonded entity that remains intact during crystallisation, forming relatively weak bonding with neighbours to yield the crystal. In fact, such a *unit of growth* is probably, in many cases, the actual *growth unit* for the crystal assuming that it does not dimerise in solution. However, for a zeolite, which is a fully connected three-dimensional network of covalent bonds, such an approach is not viable. But for our analysis the *unit of growth* is just a structural element that represents a metastable surface structure with small enough dimensions to describe all 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 in the crystal growth process. 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.

Returning then to the problem of nanoporous zeolites composed of condensed tetrahedral silicate units forming cage-like structures. We know from previous work that these cages are strongly related to metastable surface entities because the cage wrapping permits maximum condensation of the cage[11]. Imagine a cage within the bulk of the zeolite with all tetrahedral silicon sites fully condensed (termed Q^4 units). Then imagine the same cage at the surface of the crystal where most of the silicon sites will only lose one bond of condensation and is, therefore, a minimum in energy for a surface structure. In fact, of the 200+ zeolite structures, around one quarter of these structures, consist of Q^4 units that will only suffer loss of one condensation at the surface ($Q^4 \rightarrow Q^3$). The other three quarters have Q^4 units that may suffer the loss of two condensations ($Q^4 \rightarrow Q^2$), however, these 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 3-dimensional space-filling tiles that can be computed in a relatively straightforward manner using algorithms such as those implemented in ToposPro[21,22] (*Figure 1*). Consequently, this establishes a simplified route to coarse-grain the zeolite problem into energetically minimised, 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. The choice of *natural* tiles[21], which are used in our study, is unique and does not permit an alternative. In this aspect it is very different to the choices made by structural chemists when dividing up the structures, such as the secondary building units. The *units of growth* are space-filling and, although the crystal is nanoporous, it is considered as filling all space during growth (the voids within cages are filled with water and cations). This will be the same for the growth of any crystal whether nanoporous or not. This approach is naturally extendable to any cage-like structure, regardless of the bonding type, such that MOFs with extensive coordination bonds are immediately treatable.

To extend this approach to other crystals we use the Voronoi partitioning procedure, which is the dual compared to the tiling method, to fill the space with polyhedral units. In particular, in the Voronoi-polyhedra the objects (atoms or molecules) occupy the centres, while in tiles they occupy vertices. Thus, molecular crystals, such as aspirin, urea or water, can be categorised as a 3-D-Voronoi partition, where the molecule sits at the centre of a Voronoi-polyhedron and the faces of the polyhedron represent the interactions with neighbouring molecules (*Figure 1*). Similarly, for ionic crystals, such as calcite or zinc oxide, the ions sit at the centre of Voronoi polyhedra with faces representing the interactions between cations and anions. In these last two cases the network of interactions can be considered without the need to introduce the concept of 3-D partitioning, however, it is useful to realise 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 the structures that have no tiling, e.g. for polycatenated networks.

The problem then is to establish the energies of all these 3-D polyhedral units in any configuration and degree of condensation/attachment at the surface of a growing crystal relative to the solution phase. For complex crystal systems there could be thousands of possible surface site types although, in principle, only a fraction of these will be topologically viable during crystal growth. By interfacing our kinetic Monte Carlo code with the 3-D partitioning approach of ToposPro[22] we can compute all 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/attachment (see *Figure E1* for the LTA zeolite system). Secondary energetic effects can be computed at a much higher level of simulation in order to determine subsidiary effects, but most structural features are determined purely by connectivity. Common defects, such as screw dislocations, can be incorporated by displacement of 3-D polyhedral units to equivalent sites along the screw core resulting in perfect crystal re-connection. Growth modifiers can be simulated by poisoning *units of growth* accordingly. 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 very important LTA and FAU structures are shown in Figures *E1* and *E2* respectively and illustrate how both the habit of the crystal and the much more vulnerable surface topology can be matched with experiment across all crystal faces.

This approach allows the straightforward computation of crystals no matter what degree of complexity exists in the structure. For example, the UOV structure (Figures *2* and *E3*) with a very large unit cell and 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 that also includes the nature of the surface termination which, for nanoporous materials, is the gateway to the internal porosity. This approach also demonstrates how framework crystals such as NES (*Figure E4*) 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 stabilising of this large cage through templating.

The MFI zeolite framework type, which is one of the most important industrial catalysts known as ZSM-5, reveals not only the surface 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 plethora of internal silanols that is well-established for ZSM-5. The interesting discovery is that, because the growth mechanism on different faces of the crystal is necessarily different, the silanols are confined to zones of different density within the crystal (Figures *2* and *E5*). This mimics almost exactly the optical microscopy images that show identical zoning and has been a source of debate for many years[23]. Similarly, in ETS-10 which displays rod growth (*Figure E6*) 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 just as observed experimentally by Raman microscopy[24]. Our kinetic 3-D partitioning model shows that a straightforward growth mechanism can explain these optical phenomena without the need to resort to 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 CHA and LTA (Figures *E7* and *E1* 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 system (*Figure E8*), 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 (*Figure E9, left*). Indeed, MOFs are as readily treated using this approach either as cage (partitioned) structures as in HKUST-1 or as molecular crystals as shown for MOF-5 (*Figure E9, right*). In the latter case it is necessary also to consider the solvent as an important element in the crystal growth since without this it is impossible to replicate the observed crystal habit and surface topology. This is a good example of a multicomponent crystal and demonstrates the power of this approach to this important general class of materials.

Both molecular crystals and ionic crystals (Figures *3* and *4* respectively) are amenable to this treatment and for calcite the values of crystallisation energies are in broad agreement with those calculated using a combination of interatomic potentials and a continuum solvent model, see methods section. For the L-cystine system it has been shown[29,30] that growth on the <001> face proceeds predominantly via screw dislocations. When this is augmented with the 6_1 screw axis of the crystal structure and highly anisotropic rates of crystal growth a complex pin-wheel surface topology is generated. Our simulations are able to faithfully reproduce all these growth features based upon four interaction energies and reveal the importance of the interplay between different growth modes in the complete crystallisation mechanism (Figures *4* and *E10*).

Finally, addition of growth modifiers is also readily achieved (*Figure 4*) where targeting of specific growth sites can be examined in relation to their effect on the growth topology giving similar results to those observed experimentally[25]. The power of this approach is in the general applicability across crystal systems and gives a window of understanding that can then be readily explored through higher-level calculations on each individual system.

1. Ostwald, W. Studien über die bildung und umwandlung fester körper. *Z. Phys. Chem.* **22**, 289-330 (1897).
2. M. Volmer, M. Kinetics of phase formation (Kinetik der phasenbildung), *Theodor Steinkopf, Dresden and Leipzig*, (1939).
3. Stranski, I.N. Zur Theorie des kristallwachstums. *Z. Phys. Chem.* **136**, 259-278 (1928).
4. Kossel, W. Zur energetik von oberflächenvorgängen nachrichten der gesselschaft der wissenschaften Göttingen. *Mathematisch-Physikalische Klasse* **Band 135** (1927).
5. Burton, W.K., Cabrera, N. & Frank, F.C. The growth of crystals and the equilibrium structure of their surfaces. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **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 *Surface Science* **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. Amer. 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. DeYoreo, J.J. *et al.* Rethinking classical crystal growth models through molecular scale insights: consequences of kink-limited kinetics. *Cryst. Growth & Design* **9**, 5135-5144 (2009).
11. Brent, R. *et al.* Unstitching the nanoscopic mystery of zeolite crystal formation. *J. Amer. 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.   Ferey 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 Cryst.* **A63**, 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 and Design* **14**, 3576-3586 (2014).

23.   Roeffaers, M.B.J. *et al.* Morphology of large MFI crystals unravelled by fluorescence microscopy. *J. Amer. Che. 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 $TiO3^{2-}$ quantum wires in an ETS-10 crystal revealed by laser-scanning confocalpolarized micro-Raman spectroscopy. *Angew. Chem. Int. Ed.* **50**, 8697-8701 (2011).

25.   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).

26.   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. *Miner. Mag.* **62A**, 1194-1195 (1998).

27.   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).

28.   Walker, A.M., Slater, B., Gale, J.D. & Wright, K. Predicting the structure of screw dislocations in nanoporous materials. *Nature Materials* **3**, 715-720 (2004).

29.   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).

30.   Rimer, J.D. *et al.* Crystal growth inhibitors for the prevention of L-cystine kidney stones through molecular design. *Science* **330**, 337-341 (2010).

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).

Author Contributions

MWA conceived ideas and wrote *CrystalGrower* growth code and performed simulations, **JTG** wrote early version growth and visualisation code and performed simulations, **ARH** wrote *CrystalGrower* visualisation code and performed simulations, **NF** and **PC** recorded AFM images, **MPA** coordinated MOF work, **VAB** modified TOPOSPro code to interface with *CrystalGrower*, **DMP** developed ideas to integrate tiling methodology, **DA** and **BA** funded **ARH** and contributed to crystal growth mechanism discussions, , **JDG** performed computed energetics on calcite and L-cystine systems.

MAIN FIGURE LEGENDS

Figure 1 | Demonstration of the tiling and Voronoi-polyhedra partitioning methods. Top sequence shows the chabazite structure (CHA) which is composed of two cages (tiles), the double six ring (pink) and chabazite cage (green). Both these tiles consist entirely of Q^3 tetrahedra and are considered as closed cages. As the crystal grows, shown in the middle, the cages condense and Q^3 is converted into Q^4, right image, thereby stabilising the cage relative to the solution phase. The middle sequence shows the complex UOV structure that consists of 16 tiles (see *Figure E3*). One of these tiles is shown that consists of both Q^3 and Q^2 vertices and is therefore considered as an open tile. Condensation again results in stabilisation of these tiles relative to the solution phase. The lower image shows a tile representation for the molecular crystal urea. On the left is 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 on the right represented by the black lines. Each interaction passes through the face of the ten-sided tile.

Figure 2 | Collection of images presenting the results of simulations run on a variety of framework types. 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 *Figure E1* and Supplementary Movie 1). 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 respectively. Interrogation of the internal structure of the crystal reveals an hour-glass structure similar to that observed experimentally by optical microscopy. This is a result of incomplete tiles having silanol groups known to be present in the ZSM-5 structure (see *Figure E5* and Supplementary Movie 4). ETS-10 is an octahedral/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 seen here the rod growth nature of the crystal is immediately apparent and leads to the incorporation of defects (see *Figure E6* and Supplementary Movie 5). UOV is one of the most complex zeolite structures with a very large unit cell and 16 tiles, both open and closed. The present methodology is able to efficiently grow such a complex structure with both surface topography and habit matching that observed experimentally. The surface structure is also gleaned from the calculations as well as the nature of partially constructed layers at intermediated metastable steps (see *Figure E3* and Supplementary Movie 3). MOFs can be modelled by two differing methods. First, treating them as multicomponent molecular crystals with metal clusters and organic linkers treated as separate molecules (as in MOF-5 shown above). Second using the same treatment as zeolite frameworks (as seen with HKUST-1 in *Figure E9, left* and Supplementary Movie 6). Again the crystal habit and surface topography match that observed experimentally with different crystallisation conditions (further examples are shown for MOF-5 in *Figure E9*, right).

Figure 3 | Simulations of an ionic crystal (calcite) and a molecular crystal (urea). Simulations of calcite and urea, demonstrating the universality of the present approach to

different crystal classes. All simulations are shown under equilibrium conditions. For calcite the reaction energy for the conversion of solubilised ions to the crystal per coordination to the crystal is set at 5, 10 and 15 kcal/mol. Calculation shows that the value lies between 10 and 15 kcal/mol and the crystal habit and surface topography of these two crystal simulations match experiment closely. At 5 kcal/mol the terrace edges are much more rounded than observed experimentally. The major difference between 10 and 15 kcal/mol is the terrace density that can also be used as a distinguishing factor. For urea three different reaction energies are used depending primarily upon the strength of interaction in the urea crystal (discussed in methods section). The large {110} faces are flat and dominated by terraces elongated in the c-direction. The smaller pseudo-{111} faces are rough and generated in large part by dissolution when the supersaturation is close to equilibrium.

Figure 4 | Images demonstrating the results of incorporating screw dislocations in growing LTA, L-cystine and calcite crystals. Screw dislocations may be computed using the present 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. Upper left images show simulations of LTA structure with and without a screw dislocation running along [100] through the crystal. Immediately apparent is a lengthening of the crystal along the [100] direction owing to the greater ease for growth at the spiral growth front. This demonstrates how to determine the relative growth rates of layer-by-layer growth versus spiral growth. Upper right image shows 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 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 Movie 7. The lower sequence shows screw dislocations in calcite. Left image 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 image 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 image shows the effect of selective poisoning at two-coordinate sites along terrace edges that has the effect to produce rounding of terrace features.

METHODS

***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 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 as 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 utilised to determine the consequences of condensation of polyhedral units on vertex condensation is also provided. All information is provided in *p1* symmetry in the primitive unit cell. The low symmetry is necessary as crystal symmetry is broken at the crystal surface and the primitive cell is used in order to ensure the most efficient calculation. By equating the crystal network to a network of polyhedral units also provides an effective route for enumeration of screw dislocations. These are computed through cutting the network of polyhedral units by a plane up to the screw core followed by translation of all polyhedral units on one side of the plane along the screw direction to an equivalent position followed by reconnection. This results in a new, fully connected, network with no dangling bonds. Adjacency matrices for polyhedral units are computed for the screw as well as the perfect structure such that look-up tables can be generated in order 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 will be energetically feasible. *CrystalGrower* treats the problem simply as a network but the stability of screw dislocation structures could be tested by energy minimisation methods separately[28].

The key to the program lies in efficient identification of all possible site types for growth and this is done once upon initiation for both the perfect structure and any screw dislocations. These site types are not re-computed at each growth/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 each site type changes as well as 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 this reduces the unknowns to be determined.

The code allows the energy ladders of ΔU_s values to be determined independently for each tile (*Figure E1*) and different weightings can be applied to different Q^n values. The code further permits the poisoning of sites in order to simulate the addition of growth modifiers. Also, effects of chemical ordering on ΔU_s values within the lattice (e.g. Zn/P ordering in framework zinc phosphates or Si/Al ordering in high alumina zeolites) can be computed.

In order to treat molecular crystals or ionic crystals site types are determined according to Voronoi polyhedron neighbour count rather than tile vertex Q^n count. The philosophy is that the molecule or ion sits within a Voronoi polyhedron built for its centre of mass and 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.

Growth model and adaptation to nanoporous materials through 3D-tiling. The growth model is adapted from that described by Boerrigter et al.[20] whereby if *units of growth* can be identified then the probability for growth relative to the probability of dissolution depends only upon the stabilisation 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 stabilisation energy, ΔU_s. The solution is only considered to act 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 is equal. There are three principal approximations in the previous work:[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; (iii) the entropy change of the growth unit is the same regardless of site type. The probabilities for growth, P_s^{growth}, and dissolution, $P_s^{dissolution}$, are then given by equations 1 and 2, where the factor 0.5 indicates no preference for growth versus dissolution.

$$P_s^{growth} = \exp\left(-0.5\left(\frac{\Delta U_s}{kT}\right) + 0.5\left(\frac{\Delta \mu}{kT}\right)\right) \qquad \text{equation 1}$$
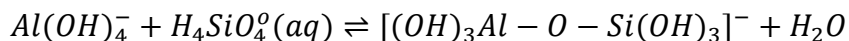
$$P_s^{dissolution} = \exp\left(0.5\left(\frac{\Delta U_s}{kT}\right) - 0.5\left(\frac{\Delta \mu}{kT}\right)\right) \qquad \text{equation 2}$$

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. The justification for this comes from the well-known Bell-Evans-Polanyi principle, as widely employed 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 just depends on the difference in thermodynamics for the corresponding reactions. Providing 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 (e.g. desolvation of the solution species and surface attachment point). Therefore we argue that the Bell-Evans-Polanyi principle should be equally as valid for crystallisation processes as for catalysis. Given this, the probabilities of growth vs 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 to (i) identify the units of growth and (ii) determine in a simple manner the stabilisation 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, *unit of growth*, is just 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 both sensible in terms of metastability, small enough to capture the essential growth features but large enough in order to ensure efficiency of calculation. This final provision can be illustrated by reference to MOFs where the crystal could be grown as a molecular crystal based upon metal clusters, organic linkers and solvent but it can also be grown as closed cages that are known to be metastable.

Closed cages are readily determined using the approach of 3D-tiling achieved by the ToposPro code[22] which computes natural tiles[21]. For the **299** zeolite frameworks studied, this accounts for **54** (less than one quarter) of all the structure codes that are constructed entirely from closed cages / tiles made up of Q^3 tetrahedral atoms. The remaining **175** structure codes are constructed from tiles that include some Q^2 tetrahedral atoms (so-called "open cages/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 both criteria in terms of both metastability and size. Any tiling, other than the natural tiling, will just be a summation of natural tiles and therefore in a crystal growth model will just 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 in order to compute probabilities for growth versus probabilities for dissolution. Absolute energies are not required. It is known from experiment[26,27] that the condensation free energy for the reaction;

$$Al(OH)_4^- + H_4SiO_4^o(aq) \rightleftharpoons [(OH)_3Al - O - Si(OH)_3]^- + H_2O$$

is -4.2 kcal mol^{-1} and so we can expect each condensation of a *unit of growth* to result in this order of magnitude change in relative ΔU_s value. So to a first approximation the ΔU_s values for a particular tile will form a uniform ladder of energies with spacing on the order of a few kcal mol^{-1}. 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 stabilisation 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 just depending upon the degree of condensation of the tile. So, for example, octahedral frameworks, octahedral/tetrahedral or any other combination are readily treated. An example of the ladders of ΔU_s values for the simple LTA system is given in *Figure E1*. For complex systems such as UOV with 16 tile types (*Figure E3*), the number of site types is very

large, 1843, although only a fraction of these are actually ever populated, in this case 168. All other tiles required for the calculations on systems in this paper are given in Figures *E1-E10*.

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 as we are just growing one crystal. In the future one could imagine looking at competitive growth of crystals, such as Ostwald ripening, through the inclusion of multiple nuclei.

The goal of this work is to find a generic algorithm that is able to describe all the most important aspects of "classical" crystal growth, readily, across any crystal system. The trigger for this endeavour is that, with the advent of scanning probe techniques – in particular 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, can a model be derived whereby the number of parameters in the model is less than the experimental observations – i.e. the problem is over-determined. For example, with zeolite A the experimental observables are crystal morphology/aspect ratio, terrace morphology on three different facets under different supersaturation conditions, terrace density, terrace heights. This is on the order of twelve observable parameters to which we are fitting three energy parameters. As a consequence, by performing hundreds of simulations we are able to pinpoint, for this system, with quite a high degree of accuracy the free energies associated with the rate-determining steps in this process. Then consider the UOV system discussed following. There is a similar number of experimental observables, however, the potential number of energy parameters for the model is 16 making the problem underdetermined. So an assumption 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 one. Despite this assumption we obtain a remarkably good fit with experiment. We clearly need to be aware of the assumption we have made and predicate further conclusions with that in mind. However, as the knowledge on the mechanism of crystallisation of UOV was completely non-existent 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 gives a route to study the basics of crystal growth for any system, quickly, whereby the number of fitting parameters will be on the same order of magnitude or less than the number of experimental variables. Peculiarities for a given system would then need to be layered on top of this approach.

Different polymorphs can be readily treated as they just represent different crystalline networks and therefore a different 3D partitioning of space. For example, the different zeolite structures in their purely siliceous forms are all polymorphs of silica. In terms of predicting competition between polymorphs this is not something that we are currently addressing. Nevertheless, one could envision an extension of our approach if the different structures are put on a relative scale then the probability for nucleation of different

polymorphs could be tested. This 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 clearly of enormous importance and we believe our approach can be extended to include this. However, at this stage we were concentrating on keeping the algorithms both generic and efficient. Defects and intergrowths, by their very nature are system specific and consequently so far we have only considered the omnipresent screw dislocation. The question is whether to add this in the future in a bespoke manner for each system or develop protocols for introduction of some more common intergrowths and defects.

The technique can be used for co-crystals with the modification that the driving force for each component would need to be considered separately, although those driving forces would be linked by the stoichiometry of the growing crystal. In principle any type of interaction can be considered as 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 have been refined by performing over one thousand calculations that help to pinpoint the solution (see *Figure E1, g*). What can be seen in this instance is that the crystal features are quite sensitive to these energy parameters (as might be expected). This, consequently, allows the refinement to be quite precise. Precision will vary from system to system depending upon number of experimental variables versus the number of fitting parameters. Similarly, the rigor of the comparison will again depend of what experimental data is available that can be compared. For example, measurement of supersaturation for a molecular crystal may 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 rigor of the comparison required depends upon 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 in order to improve diffusion to a catalytically active centre" – then what is important is "which parameters influence the crystal aspect ratio". For zeolite L that means stabilisation of the large cages, which can be achieved through specific templating. In order to make that determination does not require the specific free energies of condensation to be known to an accuracy of greater than 0.5 kcal/mol. But it is crucial to know the general crystal growth mechanism that our simulation provides. For most materials' chemists the trends are the most important thing in order to influence outcome. If the experimental data set is large and the number of fitting parameters 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.

In order to further test our methodology we have looked at the ʟ-cystine crystal growth system which has been studied extensively by AFM[29,30] owing to its involvement in the pathogenesis of cystine kidney stones. Previous work was able to record detailed images 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. Similarly using our approach we are able to readily simulate all the main features observed in this system (see *Figure E10*) – including in our case the rounded and fractal features that a

Hartman-Perdok analysis will not expose. By doing multiple simulations in a similar manner to than performed for the zeolite A system we are able to refine the free energies of crystallisation of all the principal interactions in this structure. For example, the two slow growth directions (labelled A+ and A- in reference 29) must be on the order of 1–2 kcal/mol different to each other in order to explain the step bunches with height one unit cell in the *c*-direction (*Figure E10, b*). If the difference between these growth energies is removed or reduced the height of the step bunches is only a half unit cell in the *c*-direction owing to the symmetry that is repeated every three L-cystine layers. Most interesting, however, from this work is the separation between the step bunches which changes markedly either with supersaturation or changes in the four binding energies. Reference 29 discusses this step bunch separation at length and invokes lattice strain in order to account for deviations between the separation expected from a Hartman-Perdok analysis and experiment. 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 surface <100> site is energetically more favourable than binding at a higher coordination edge site of an individual slow step. Consequently, the rate of growth of both the <100> facets and the slow steps is governed by the same process. In fact, this effect can be seen in Supplementary Movie 1 to reference 30 whereby the slow steps are seen to advance principally as a result of advancement of the step bunches and we reproduce this in our Supplementary Movie 7. Step advancement in the slow growth direction is further enhanced by birth-and-spread nucleation on the <001> facet. For this somewhat complicated system, being able to visualise the crystal growth in three dimensions helped to disentangle these competing processes. The outcome of a reversal of the handedness of screw dislocation is also simulated in *Figure E10, c*. Because the overall advancement of terraces is governed by the crystallography of the material rather than the handedness of the screw dislocation the pattern of growth is remarkably similar, independent 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. However, near the screw core the terrace structure is necessarily quite different and is the distinguishing factor between opposite handedness. Both L-cystine and D-cystine have been simulated (*Figure E10, d*) 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].

Calcite calculations Calculation of 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 optimised bulk structure of calcite commensurate with this model, a rhombohedral

nanoparticle of calcite was cleaved with dimensions of 16 x 16 x 4 atomic rows oriented such the long edges run parallel to either the acute or 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 optimisation in order 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 4 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 depending upon the connectivity of the ion within the lattice. The difference in reaction energy between the ion in solution and ion in the crystal is between 10 and 15 kcal/mol per ligand for both Ca^{2+} and CO_3^{2-}. In order 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 these crystal systems are known to exhibit optical 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 crystallise in a manner that leaves some tiles incomplete and as a consequence this will result in a high number of internal silanol groups. Both MFI and ETS-10 (*Figures E5* and *E6*) show internal defect density patterns that mimic almost exactly the birefringence observed experimentally. This is a much more appealing pathway rather than the necessity to invoke complex twinning mechanisms to explain such effects. We will use this technique in future to explore a wide range of framework structures that are known to exhibit such effects.

***CrystalGrower* visualisation code** The *CrystalGrower* visualisation code is primarily written 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 *CrystalGrower* growth code and allow users to observe the morphology and surface structure of the grown crystal, whilst also manipulating the crystal in real-time. Two information files are required to construct and display the polyhedral units that compose the crystal framework, which are both 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 based on a small number of parameters for each polyhedral unit (number of vertices, number of faces, flat / non-flat faces *etc*). This will simply be a sphere in the case of molecular crystal. Once these 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 the radius defined by taking the root mean square of the polyhedral unit xyz coordinates, multiplied by a scaling factor to allow changes to be made to the sphere size while visualising the crystal structure as spheres. The second step uses another output file generated by *CrystalGrower*, which provides the visualisation package with the xyz coordinates for each unit cell grown in a run, along with information relating to the type of the unit to draw. Coupling these coordinates with those in the partition file, the polyhedral unit is drawn at the correct position to be visualised. After completing the drawing of all 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 as it greatly improves the speed of the program, allowing real-time manipulation of a large number of objects (~300,000) on graphics cards with greater than 2 GB of dedicated graphics memory.

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 were used.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from MWA and ARH upon reasonable request.

EXTENDED DATA FIGURE LEGENDS

Extended Data Figure E1 | Experimental and simulated data for the LTA framework with energy level diagram. The top segment (denoted with **a**) shows a schematic representation of energy ladders for the LTA structure. Double 4-rings are shown in yellow, sodalite (β) cages are shown in magenta and the α cages are shown in cyan. Each level on the energy ladders moving from bulk to the top, bold line 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 where 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.

The LTA structure consists of three tiles (**b**) connected in a cubic lattice (**c**). Typical crystals exhibit three facets each with unique terrace shape (**d**) all of which can be reproduced along with the crystal habit in a simulation that treats the crystallisation energy of each tile separately (**e**, top), a prediction of the surface termination 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**). A video of the screw dislocation growing through the simulation can be seen in Supplementary Movie 1. The simulated crystals shown in **e** and **f** are approximately 0.15 μm x 0.15 μm x 0.15 μm and 0.20 μm x 0.15 μm x 0.15 μm in size, respectively.

A ternary plot is shown in **g** 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 destabilisation for each cage is at the corner for 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.

Extended Data Figure E2 | Experimental and simulated data for the FAU framework. The FAU structure consists of three tiles (**a**) connected in a face centred cubic lattice (**b**). Simulation dropping from high supersaturation (50 kcal/mol) to low supersaturation (1 kcal/mol) at the halfway point (**c**) matches both 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. Simulations with the same parameters as **c** with varying zoom distances (**f**) show the surface structure predicted for this simulation. Supplementary Movie 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 x 0.20 μm x 0.20 μm.

Extended Data Figure E3 | Experimental and simulated data for the UOV framework. The UOV structure represents one of the most complex framework systems

with a particularly large unit cell with c-dimension 3.87 nm and consists of 16 different tile types (**a**), connected in an orthorhombic lattice (**b**). The germanosilicate crystallites that adopt this framework structure, although intergrown (**c**), show very well defined shape bounded by six faces. 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. Supplementary Movie 3 shows an example of how a UOV crystal grows at both high and low supersaturation values. The simulated crystal shown in all these examples is approximately 0.050 μm x 0.75 μm x 1.40 μm in size.

Extended Data Figure E4 | Experimental and simulated data for the NES framework. The NES framework can usually be expressed with four different tile types. The largest tile – t-nes can be split into three smaller tiles (**a** - green, cyan and blue) to give a total of six tiles possessing similar sizes (**a**) that combine to form an orthorhombic lattice (**b**). These three tiles can no longer be referred to as natural tiles, however they were still generated using the ToposPro software. These tiles play the role of units of growth and are required as CrystalGrower utilises connections through faces, making frameworks that share only corners and edges between small tiles, such as NES, difficult to grow over reasonable timescales. 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**). Lowering the energy penalty for the largest cage changes the morphology of the crystal dramatically, thickening the crystal considerably (**e, 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 (**e, 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 x 0.40 μm x 0.15 μm, whereas the example in **f** is estimated to be 0.10 μm x 0.20 μm x 0.15 μm in size, respectively.

Extended Data Figure E5 | Experimental and simulated data for the MFI framework. The MFI structure consists of ten tile types (**a**) connected in an orthorhombic lattice (**b**). A consistent feature in MFI crystals is the optical hourglass effect seen in crystals of silicalite shown in (**c**). 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 incorporated as a result of the crystal growth mechanism being different on different faces of the crystal. Our simulations (**d**) reflect well the crystal habit and surface topology and allow investigation of the surface termination. The internal structure (**e**) shows that the density of silanol groups mirrors very closely the sectoring of the crystal, as only tiles with incomplete coordination are shown. Such a change in crystal chemistry would, almost certainly, be associated with a change in crystal refractive index and hence the observed optical effect. Supplementary Movie 4 demonstrates growth of the hourglass feature as the simulation progresses as well as showing how different tiles express this feature with differing degrees. The simulated

crystal shown in **d** is approximately 0.10 µm x 0.05 µm x 0.40 µm in size, whilst the crystal shown in **e** is approximately 0.10 µm x 0.05 µm x 0.20 µm.

Extended Data Figure E6 | Experimental and simulated data for the titanosilicate material ETS-10. ETS-10 is a nanoporous titanosilicate with a structure consisting of five tiles (**a**) connected in a monoclinic lattice (**b**). The structure has a very similar symmetry to zeolite beta, however, the titanium is incorporated into titania rods surrounded by silica (**c,d**) – HREM courtesy of O. Terasaki. These rods run alternatively in orthogonal [110] and [-110] directions and it is found that these rods are the dominant components for the crystal growth (**f**). Further, because the rods do not always connect and heal inside the crystal structure there is a high concentration of internal silianol groups in a sector from the (001) faces to the centre of the crystal (**e**). Such sectoring in ETS-10 has been observed in Raman microscopy[24]. A movie demonstrating how the rods in the ETS-10 framework grow in alternating orthogonal directions can be viewed in Supplementary Movie 5. The simulated crystals shown in **d** and **e** are approximately 0.25 µm x 0.25 µm x 0.05 µm, and 0.10 µm x 0.10 µm x 0.10 µm in size, respectively.

Extended Data Figure E7 | Experimental and simulated data for the CHA framework. The CHA structure consists of two tiles (**a**) connected in a monoclinic lattice (**b**). Typical silicoaluminophosphate (SAPO) 34 crystals show distorted cube morphology when viewed under a SEM (**c**) exhibiting <100> faces with isotropic growth of terraces. A plethora of screw dislocations can be observed in this system (**e**) and this is reproduced for a dislocation with a screw core running along [100] (**d**). Simulations predict that the surface terminates at the double six rings (**f**). The simulated crystal shown in **d** has an approximate size of 1.50 µm x 0.10 µm x 0.10 µm due to elongation caused by the screw dislocation.

Extended Data Figure E8 | Experimental and simulated data for the AEI framework. The AEI structure consists of two tiles (**a**) connected in an orthorhombic lattice (**b**). The structure is related to CHA in that it can be built by connecting double six rings. However, because of the alternating orientation of these units along the c-direction this results in a switching of the fast and slow growth rates along the a-direction and b-direction (shown on AFM of aluminophosphate (AlPO) 18 in **c**). The result is a complex interleaved spiral growth that is faithfully simulated according to this space partitioning methodology (**d-f**). The simulated crystal shown in d – f is approximately 1.70 µm x 0.05 µm x 0.05 µm in size due to elongation caused by the screw dislocation.

Extended Data Figure E9 | Experimental and simulated data for two separate metal-organic frameworks, HKUST-1 (left) and MOF-5 (right). The left partition (**a-e**) shows HKUST-1, a metal-organic framework 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 results in excellent simulation 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 as such a screw core would result in multiple spirals (**d**). Simulations also allow investigation of the surface termination for this framework (**e**). A video of the screw dislocation growing and

migrating across the surface of the crystal can be viewed in supplementary movie 6. The approximate size of the crystal shown in **d** and **e** is 0.35 μm x 0.35 μm x 0.35 μm.

The right partition (**f-j**) shows MOF-5 [$Zn_4O(bdc)_3$, bdc=1,4-benzenedicarboxylate] which is a permanently porous metal-organic framework with a simple cubic arrangement of metal centres (red) and linkers (grey), along with a solvent molecule occupying the centre of each cube (purple) (**f, left**). Treating the structure as a multi-component 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 and solvent (**f, right**). Square terracing on the (100) face (**g**) can be in one of two orientations, dependent on the synthesis conditions. The crystal morphology can also be changed to exhibit (111) faces as an octahedral crystal, again dependent on synthesis conditions (**h, left**), as opposed to only (100) faces (**h, right**) in a cubic crystal. (111) faces exhibited by the crystal are much more isotropic (**i**) compared to the (100) faces (**f**) and do not adopt different orientations demonstrated by the (100) faces. The simple cubic network alone will not permit the development of (111) facets and is immediately an indication of the importance of the solvent N,N'-diethylformamide (DEF) in the preparation. By varying driving forces for binding of the linker, metal centre and solvent in the final crystal all crystal habits and surface topologies may be generated (**g – i**). Double spirals through growth at screw dislocations on the (100) face can also be simulated (**j**). The estimated sizes of the crystals shown in **g** are 0.76 μm x 0.75 μm x 0.75 μm and 0.70 μm x 0.80 μm x 0.80 μm, **h** are 0.50 μm x 0.50 μm x 0.50 μm and 0.75 μm x 0.75 μm x 0.75 μm, **i** is 0.50 μm x 0.50 μm x 0.50 μm, and **j** is 0.75 μm x 0.75 μm x 0.75 μm, respectively.

Extended Data Figure E10 | Collection of simulated data for L-cystine along with a single simulation of D-cystine. The panel labelled **a** presents a suite of calculations of the crystal growth of L-cystine at supersaturation 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 crystallisation: (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, known as A+ and A- in references 29 and 30, with one hydrogen bond; (iv) in the c-direction with two hydrogen bonds. Interaction type (ii) and (iii) are maintained at half that of (iv) with a minimal difference of (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 in kcal mol^{-1}. When all these interactions are too weak isotropic sticky growth is observed and when they are too high the growth is too geometrical. The optimum balance is highlighted in **a** where the sum of the free energies of crystallisation is 7.0 kcal mol^{-1}, compared to 4.3 kcal mol^{-1} derived directly from the solubility of L-cystine at pH 7[31]. Similar to that reported in reference 29, the step bunching is closer than observed experimentally as lattice strain at the screw core will weaken interactions and prevent the core region, which is advancing by attachment to the slow step alone, forming a closed hexagonal step bunch as quickly as predicted. A set of simulations demonstrating the effect of small differences in the energy of slow growth directions termed A+ and A- in references 29 and 30 is shown in **b**, noted 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 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. Reversal of the handedness of the screw core is not immediately apparent (**c**) as 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. L-cystine and D-cystine will present identical surface structures with opposite handedness, as expected, shown in panel **d**.

SUPPLEMENTARY MOVIE LEGENDS

LTA – Growth via Screw Dislocation – Supplementary Movie 1

In this movie, an LTA crystal is grown with a screw dislocation running in the [100] direction through the crystal. Similar to the FAU simulation, the crystal is first grown under high supersaturation conditions (90 kcal/mol) for 5 million iterations, followed by a drop in supersaturation to slightly above equilibrium (1.5 kcal/mol) to allow a more detailed study of the growing process. During the slow growth, the propagation of the spiral from the screw dislocation is clear. The crystal is then rotated to the opposing face, demonstrating the elongation of the morphology in the [100] as expected as the screw dislocation creates a low energy surface for the growth units to attach to. After arriving at the opposing face, the slow growth process is reversed then replayed to again demonstrate the spiral-growth of the crystal.

FAU – Layer Growth – Supplementary Movie 2

A FAU crystal is grown under high supersaturation conditions (50 kcal/mol) for 5 million iterations, showing rapid growth and the adoption of an octahedral crystal morphology. The second 5 million iterations occur after the supersaturation is lowered to slightly above equilibrium (1 kcal/mol), allowing the growth process to continue slowly. At this point the triangular shape of the terraces can be observed, growing in the opposing direction to the facet where the growth is occurring. The final portion of the movie shows the surface structure viewed from the top, followed by the side of a grown terrace.

UOV – Layer Growth – Supplementary Movie 3

Growth of a UOV crystal is demonstrated using similar conditions to the previous movies: 5 million iterations at high supersaturation (70 kcal/mol) followed by 5 million iterations at a supersaturation value slightly above equilibrium (1 kcal/mol). The diamond-like morphology of the UOV crystal can clearly be seen during panning and rotation. Rounded / isotropic layer growth can be observed during the growth process, along with the fusion of terraces once the growth process is slowed. Close-up shots are then shown of the surface structure on the (100) face, followed by the side-wall of the crystal: the (013) face.

MFI – Hourglass Growth and Tile Type Relation – Supplementary Movie 4

To begin, the t-mfi-1 tiles are shown during the growth process of an MFI crystal to allow a view into the internal structure of the crystal as it grows, showing the formation of the hourglass structure mentioned previously. Following this growth, a series of frames are shown where the tile type shown is cycled on the (001), (100) and (001) faces respectively, demonstrating how different tile types contribute to this internal feature in varying degrees.

ETS-10 – Orthogonal Rod Growth – Supplementary Movie 5

This movie demonstrates the growth mechanism of ETS-10 *via* titanate rods in alternating directions, orthogonal to each other. Each alternating layer is represented by

different colours to draw attention to the orthogonal directions. The conditions for this growth process are changed slightly compared to the previous movies. The crystal is first grown at high supersaturation (140 kcal/mol) to ensure the growth process begins, however no frames are recorded. The supersaturation is then lowered to 1 kcal/mol, and frames are taken every 200,000 iterations for the first 20 frames, followed by 40 frames recorded at every 10,000 iterations. During these slowed down frames, a closer view of the surface is shown, with arrows overlaid following the direction of rod growth that can be observed. A final 20 frames with a spacing of 20000 iterations are shown, along with close-ups of the surface structures adopted by the vertical and horizontal rods that express the structure.

HKUST-1 – Growth via Screw Dislocation – Supplementary Movie 6

This supplementary movie shows the growth of an HKUST-1 crystal under similar conditions to the FAU, LTA and UOV crystals simulated previously. The first 5 million iterations are run at high supersaturation conditions (100 kcal/mol) followed by a drop to a lower supersaturation (1 kcal/mol). Slowing down the growth clearly demonstrates how the screw dislocation running along the [110] direction completely alters the final morphology of the crystal. The growth is then reversed, and the screw dislocation is followed to the point where it migrates back onto the (111) that it originally appeared from. Rotating the crystal then allows the viewer to observe the elongation of the crystal shape caused by the spiral growth on both sides. Rotating the crystal to view the opposite side of the screw dislocation, the location of the core is again followed whilst re-growing the crystal at low supersaturation (1 kcal/mol) ending with the final growth frame, and a panned out view of the entire crystal.

L-Cystine – Growth via Screw Dislocation – Supplementary Movie 7

This supplementary movie shows the growth L-cystine at a supersaturation of 0.6 kcal mol^{-1}. The optimal binding free energies where: (i) strong binding 3.5 kcal mol^{-1}; (ii) and (iii) weak binding 0.78 kcal mol^{-1} and 0.98 kcal mol^{-1}; (iv) binding in the c-direction 1.75 kcal mol^{-1}. Advancement of the step bunches is through birth-and-spread growth on the <100> side faces, which is encouraged through the strong binding direction. This birth-and-spread growth precipitates growth at the single slow growth edges. Growth is further complicated by a small amount of birth-and-spread growth on the <001> face.
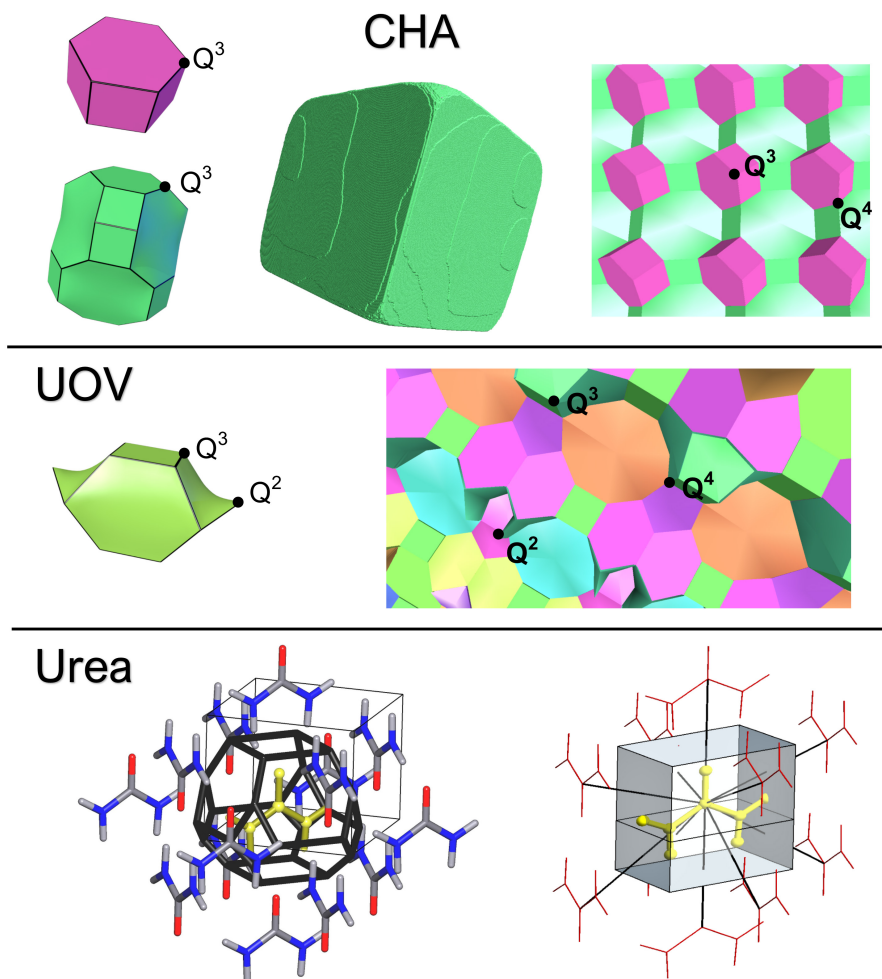
Figure 1 Top sequence shows the chabazite structure (CHA) which is composed of two cages (tiles), the double six ring (pink) and chabazite cage (green). Both these tiles consist entirely of Q^3 tetrahedra and are considered as closed cages. As the crystal grows, shown in the middle, the cages condense and Q^3 is converted into Q^4, right image, thereby stabilising the cage relative to the solution phase. The middle sequence shows the complex UOV structure that consists of 16 tiles (see supplementary material). One of these tiles is shown that consists of both Q^3 and Q^2 vertices and is therefore considered as an open tile. Condensation again results in stabilisation of these tiles relative to the solution phase. The lower image shows a tile representation for the molecular crystal urea. On the left is 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 on the right represented by the black lines. Each interaction passes through the face of the ten-sided tile.
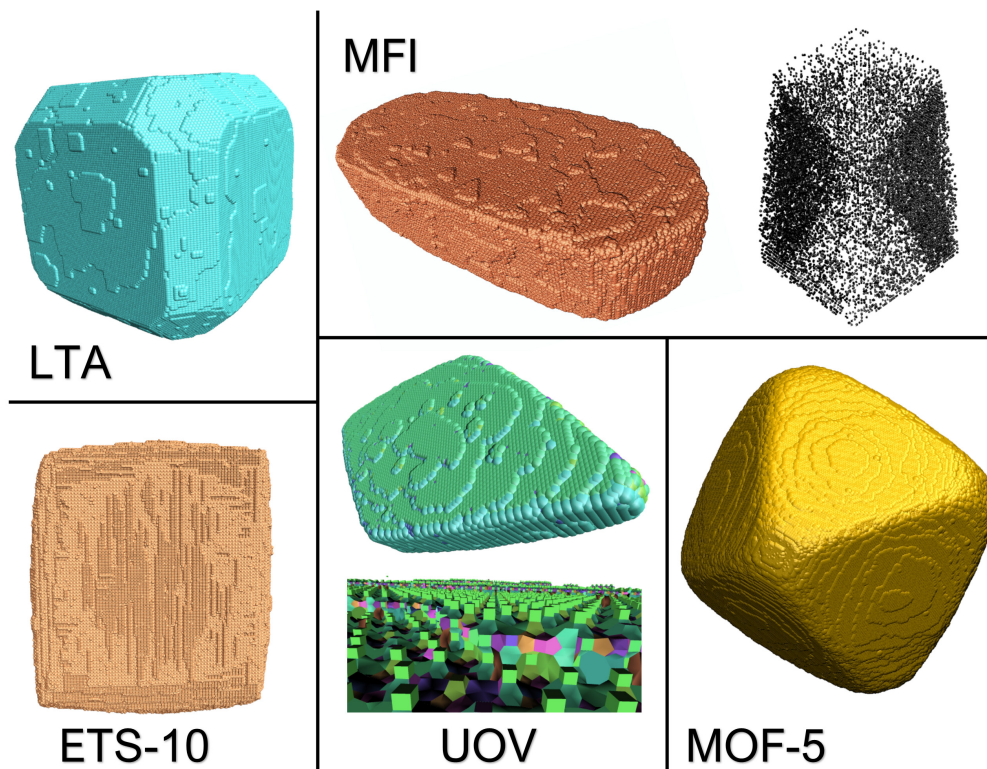
Figure 2 Simulations of a variety of nanoporous crystals exhibiting a wide selection of structural diversity. LTA, zeolite A, consists of three closed cages (tiles) and the experimental morphology can be achieved by adjusting the energy of these tiles independently relative to solution (see figure S3 and movie S2). 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 respectively. Interrogation of the internal structure of the crystal reveals an hour-glass structure similar to that observed experimentally by optical microscopy. This is a result of incomplete tiles having silanol groups known to be present in the ZSM-5 structure (see figure S6 and movie S5). ETS-10 is an octahedral/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 seen here the rod growth nature of the crystal is immediately apparent and leads to the incorporation of defects (see figure S7 and movie S4). UOV is one of the most complex zeolite structures with a very large unit cell and 16 tiles, both open and closed. The present methodology is able to efficiently grow such a complex structure with both surface topography and habit matching that observed experimentally. The surface structure is also gleaned from the calculations as well as the nature of partially constructed layers at intermediated metastable steps (see figure S4 and movie S3). MOFs can be modelled by two differing methods. First, treating them as multicomponent molecular crystals with metal clusters and organic linkers treated as separate molecules (as in MOF-5 shown above). Second using the same treatment as zeolite frameworks (as seen with HKUST-1 in figure S10 and movie S6). Again the crystal habit and surface topography match that observed experimentally with different crystallisation conditions (further examples are shown for MOF-5 in figure S11).
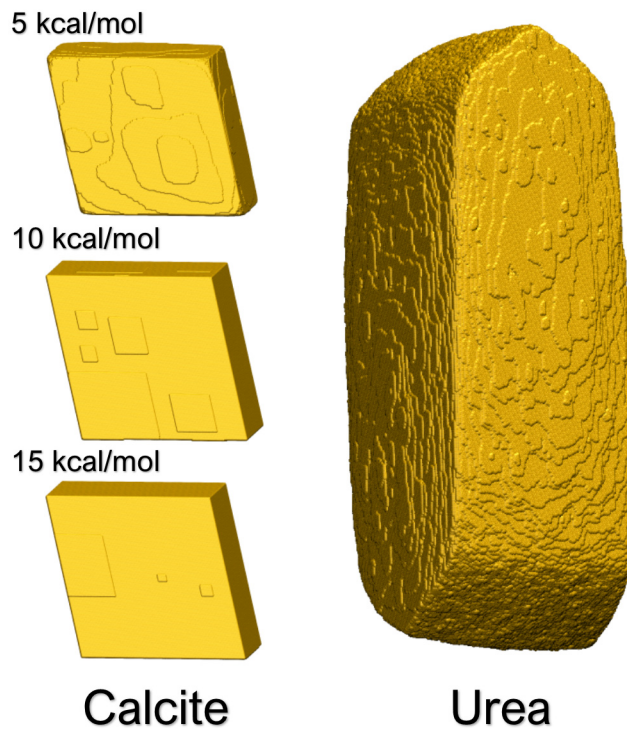
Figure 3 Simulations of the ionic crystal, calcite, and the molecular crystal, urea, demonstrating the universality of the present approach to different crystal classes. All simulations are shown under equilibrium conditions. For calcite the reaction energy for the conversion of solubilised ions to the crystal per coordination to the crystal is set at 5, 10 and 15 kcal/mol. Calculation shows that the value lies between 10 and 15 kcal/mol and the crystal habit and surface topography of these two crystal simulations match experiment closely. At 5 kcal/mol the terrace edges are much more rounded than observed experimentally. The major difference between 10 and 15 kcal/mol is the terrace density that can also be used as a distinguishing factor. For urea three different reaction energies are used depending primarily upon the strength of interaction in the urea crystal (discussed in supplementary material). The large {110} faces are flat and dominated by terraces elongated in the c-direction. The smaller pseudo-{111} faces are rough and generated in large part by dissolution when the supersaturation is close to equilibrium.
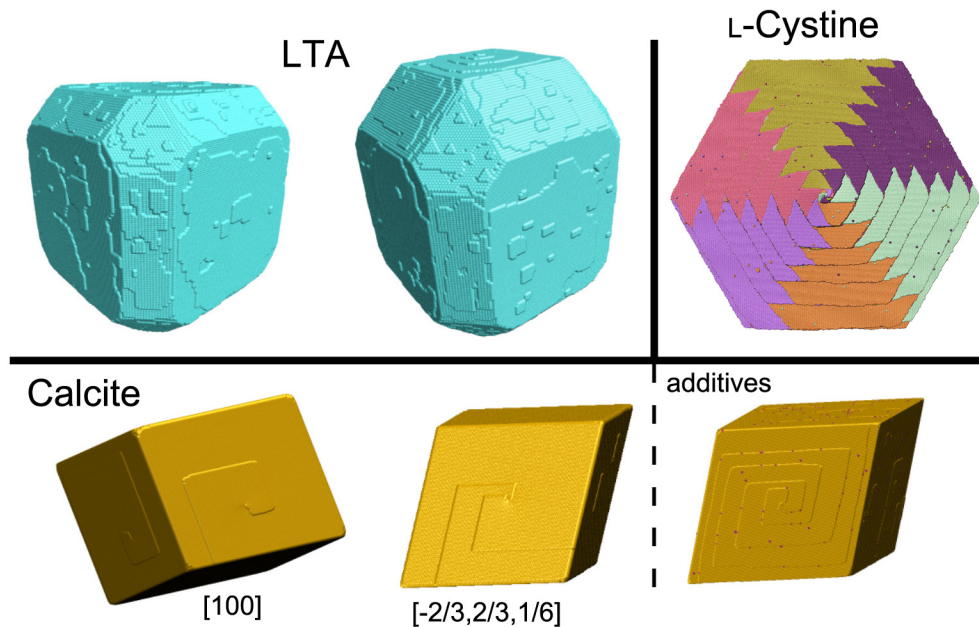
Figure 4 Screw dislocations may be computed using the present methodology for any crystal system along any crystal direction. This method operates according only to topology and does not account for the energy of the crystal at the screw core. Nonetheless, it allows all possible topologically permitted structures to be tested for growth morphology – energy considerations can be determined separately. Upper left images show simulations of LTA structure with and without a screw dislocation running along [100] through the crystal. Immediately apparent is a lengthening of the crystal along the [100] direction owing to the greater ease for growth at the spiral growth front. This demonstrates how to determine the relative growth rates of layer-by-layer growth versus spiral growth. Upper right image shows 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 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 Movie 7 of the supplementary material. The lower sequence shows screw dislocations in calcite. Left image 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 image 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 image shows the effect of selective poisoning at two-coordinate sites along terrace edges that has the effect to produce rounding of terrace features.

METHODS

***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 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 as 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 utilised to determine the consequences of condensation of polyhedral units on vertex condensation is also provided. All information is provided in P1 symmetry in the primitive unit cell. The low symmetry is necessary as crystal symmetry is broken at the crystal surface and the primitive cell is used in order to ensure the most efficient calculation. By equating the crystal network to a network of polyhedral units also provides an effective route for enumeration of screw dislocations. These are computed through cutting the network of polyhedral units by a plane up to the screw core followed by translation of all polyhedral units on one side of the plane along the screw direction to an equivalent position followed by reconnection. This results in a new, fully connected, network with no dangling bonds. Adjacency matrices for polyhedral units are computed for the screw as well as the perfect structure such that look-up tables can be generated in order 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 will be energetically feasible. *CrystalGrower* treats the problem simply as a network but the stability of screw dislocation structures could be tested by energy minimisation methods separately[28].

The key to the program lies in efficient identification of all possible site types for growth and this is done once upon initiation for both the perfect structure and any screw dislocations. These site types are not re-computed at each growth/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 each site type changes as well as 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 this reduces the unknowns to be determined.

The code allows the energy ladders of ΔU_s values to be determined independently for each tile (*Figure* S1) and different weightings can be applied to different Q^n values. The code further permits the poisoning of sites in order to simulate the addition of growth modifiers. Also, effects of chemical ordering on ΔU_s values within the lattice (e.g. Zn/P ordering in framework zinc phosphates or Si/Al ordering in high alumina zeolites) can be computed.

In order to treat molecular crystals or ionic crystals site types are determined according to Voronoi polyhedron neighbour count rather than tile vertex Q^n count. The philosophy is that the molecule or ion sits within a Voronoi polyhedron built for its centre of mass and 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.

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 were used.

SUPPLEMENTARY MATERIAL

Growth model and adaptation to nanoporous materials through 3D-tiling. The growth model is adapted from that described by Boerrigter et al.[20] whereby if *units of growth* can be identified then the probability for growth relative to the probability of dissolution depends only upon the stabilisation 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 stabilisation energy, ΔU_s. The solution is only considered to act 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 is equal. There are three principal approximations in the previous work:[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; (iii) the entropy change of the growth unit is the same regardless of site type. The probabilities for growth, P_s^{growth}, and dissolution, $P_s^{dissolution}$, are then given by equations 1 and 2, where the factor 0.5 indicates no preference for growth versus dissolution.

$$P_s^{growth} = \exp\left(-0.5\left(\frac{\Delta U_s}{kT}\right) + 0.5\left(\frac{\Delta \mu}{kT}\right)\right) \qquad \text{equation 1}$$

$$P_s^{dissolution} = \exp\left(0.5\left(\frac{\Delta U_s}{kT}\right) - 0.5\left(\frac{\Delta \mu}{kT}\right)\right) \qquad \text{equation 2}$$

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. The justification for this comes from the well-known Bell-Evans-Polanyi principle, as widely employed 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 just depends on the difference in thermodynamics for the corresponding reactions. Providing 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 (e.g. desolvation of the solution species and surface attachment point). Therefore we argue that the Bell-Evans-Polanyi principle should be equally as valid for crystallisation processes as for catalysis. Given this, the probabilities of growth vs 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 to (i) identify the units of growth and (ii) determine in a simple manner the stabilisation 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, *unit of growth*, is just 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 both

sensible in terms of metastability, small enough to capture the essential growth features but large enough in order to ensure efficiency of calculation. This final provision can be illustrated by reference to MOFs where the crystal could be grown as a molecular crystal
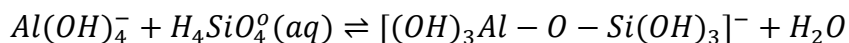
Framework composition	Number of structures
Total	229
All closed tiles	54
All open tiles	120
Mixture of open and closed tiles	55

Table S1 Statistics on the number of zeolite frameworks composed of only closed tiles, only open tiles and a combination of the two.

based upon metal clusters, organic linkers and solvent but it can also be grown as closed cages that are known to be metastable.

Closed cages are readily determined using the approach of 3D-tiling achieved by the ToposPro code[22] which computes natural tiles[21]. For zeolites this accounts for about one quarter of all the structure codes that are constructed entirely from closed cages made up of Q^3 tetrahedral atoms. The other three quarters are constructed from some tiles that include some Q^2 tetrahedral atoms. Nonetheless, these meet both criteria in terms of both metastability and size. Any tiling, other than the natural tiling, will just be a summation of natural tiles and therefore in a crystal growth model will just 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 in order to compute probabilities for growth versus probabilities for dissolution. Absolute energies are not required. It is known from experiment[26,27] that the condensation free energy for the reaction;

$$Al(OH)_4^- + H_4SiO_4^o(aq) \rightleftharpoons [(OH)_3Al - O - Si(OH)_3]^- + H_2O$$

is -4.2 kcal mol[-1] and so we can expect each condensation of a *unit of growth* to result in this order of magnitude change in relative ΔU_s value. So to a first approximation the ΔU_s values for a particular tile will form a uniform ladder of energies with spacing on the order of a few kcal mol[-1]. 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 stabilisation 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 just depending

upon the degree of condensation of the tile. So, for example, octahedral frameworks, octahedral/tetrahedral or any other combination are readily treated. An example of the ladders of ΔU_s values for the simple LTA system are given in *Figure* S1. For complex systems such as UOV with 16 tile types (*Figure* S4), the number of site types is very large, 1843, although only a fraction of these are actually ever populated, in this case 168. All other tiles required for the calculations on systems in this paper are given in *Figure* S2-S11.

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 as we are just growing one crystal. In the future one could imagine looking at competitive growth of crystals, such as Ostwald ripening, through the inclusion of multiple nuclei.

The goal of this work is to find a generic algorithm that is able to describe all the most important aspects of "classical" crystal growth, readily, across any crystal system. The trigger for this endeavour is that, with the advent of scanning probe techniques – in particular 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, can a model be derived whereby the number of parameters in the model is less than the experimental observations – i.e. the problem is over-determined. For example, with zeolite A the experimental observables are crystal morphology/aspect ratio, terrace morphology on three different facets under different supersaturation conditions, terrace density, terrace heights. This is on the order of twelve observable parameters to which we are fitting three energy parameters. As a consequence, by performing hundreds of simulations we are able to pinpoint, for this system, with quite a high degree of accuracy the free energies associated with the rate-determining steps in this process. Then consider the UOV system discussed following. There is a similar number of experimental observables, however, the potential number of energy parameters for the model is 16 making the problem underdetermined. So an assumption 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 one. Despite this assumption we obtain a remarkably good fit with experiment. We clearly need to be aware of the assumption we have made and predicate further conclusions with that in mind. However, as the knowledge on the mechanism of crystallisation of UOV was completely non-existent 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 gives a route to study the basics of crystal growth for any system, quickly, whereby the number of fitting parameters will be on the same order of magnitude or less than the number of experimental variables. Peculiarities for a given system would then need to be layered on top of this approach.

Different polymorphs can be readily treated as they just represent different crystalline networks and therefore a different 3D partitioning of space. For example, the different

zeolite structures in their purely siliceous forms are all polymorphs of silica. In terms of predicting competition between polymorphs this is not something that we are currently addressing. Nevertheless, one could envision an extension of our approach if the different structures are put on a relative scale then the probability for nucleation of different polymorphs could be tested. This 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 clearly of enormous importance and we believe our approach can be extended to include this. However, at this stage we were concentrating on keeping the algorithms both generic and efficient. Defects and intergrowths, by their very nature are system specific and consequently so far we have only considered the omnipresent screw dislocation. The question is whether to add this in the future in a bespoke manner for each system or develop protocols for introduction of some more common intergrowths and defects.

The technique can be used for co-crystals with the modification that the driving force for each component would need to be considered separately, although those driving forces would be linked by the stoichiometry of the growing crystal. In principle any type of interaction can be considered as 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 have been refined by performing over one thousand calculations that help to pinpoint the solution (see *Figures* S12-S14). What can be seen in this instance is that the crystal features are quite sensitive to these energy parameters (as might be expected). This, consequently, allows the refinement to be quite precise. Precision will vary from system to system depending upon number of experimental variables versus the number of fitting parameters. Similarly, the rigor of the comparison will again depend of what experimental data is available that can be compared. For example, measurement of supersaturation for a molecular crystal may 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 rigor of the comparison required depends upon 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 in order to improve diffusion to a catalytically active centre" – then what is important is "which parameters influence the crystal aspect ratio". For zeolite L that means stabilisation of the large cages, which can be achieved through specific templating. In order to make that determination does not require the specific free energies of condensation to be known to an accuracy of greater than 0.5 kcal/mol. But it is crucial to know the general crystal growth mechanism that our simulation provides. For most materials' chemists the trends are the most important thing in order to influence outcome. If the experimental data set is large and the number of fitting parameters 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.

In order to further test our methodology we have looked at the L-cystine crystal growth system which has been studied extensively by AFM[29,30] owing to its involvement in the

pathogenesis of cystine kidney stones. Previous work was able to record detailed images 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. Similarly using our approach we are able to readily simulate all the main features observed in this system (see *Figure* S15) – including in our case the rounded and fractal features that a Hartman-Perdok analysis will not expose. By doing multiple simulations in a similar manner to than performed for the zeolite A system we are able to refine the free energies of crystallisation of all the principal interactions in this structure. For example, the two slow growth directions (labelled A+ and A- in reference 29) must be on the order of 1–2 kcal/mol different to each other in order to explain the step bunches with height one unit cell in the *c*-direction (*Figure* S16). If the difference between these growth energies is removed or reduced the height of the step bunches is only a half unit cell in the *c*-direction owing to the symmetry that is repeated every three L-cystine layers. Most interesting, however, from this work is the separation between the step bunches, which changes markedly either with supersaturation or changes in the four binding energies. Reference 29 discusses this step bunch separation at length and invokes lattice strain in order to account for deviations between the separation expected from a Hartman-Perdok analysis and experiment. 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 surface <100> site is energetically more favourable than binding at a higher coordination edge site of an individual slow step. Consequently, the rate of growth of both the <100> facets and the slow steps is governed by the same process. In fact, this effect can be seen in movie number two of the supplementary material to reference 30 whereby the slow steps are seen to advance principally as a result of advancement of the step bunches and we reproduce this in our Movie 7. Step advancement in the slow growth direction is further enhanced by birth-and-spread nucleation on the <001> facet. For this somewhat complicated system, being able to visualise the crystal growth in three dimensions helped to disentangle these competing processes. The outcome of a reversal of the handedness of screw dislocation is also simulated in *Figure* S17. Because the overall advancement of terraces is governed by the crystallography of the material rather than the handedness of the screw dislocation the pattern of growth is remarkably similar, independent 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. However, near the screw core the terrace structure is necessarily quite different and is the distinguishing factor between opposite handedness. Both L-cystine and D-cystine have been simulated (*Figure* S18) 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].

Calcite calculations Calculation of 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 optimised bulk structure of calcite commensurate with this model, a rhombohedral nanoparticle of calcite was cleaved with dimensions of 16 x 16 x 4 atomic rows oriented such the long edges run parallel to either the acute or 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 optimisation in order 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 4 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 depending upon the connectivity of the ion within the lattice. The difference in reaction energy between the ion in solution and ion in the crystal is between 10 and 15 kcal/mol per ligand for both Ca^{2+} and CO_3^{2-}. In order 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 these crystal systems are known to exhibit optical 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 crystallise in a manner that leaves some tiles incomplete and as a consequence this will result in a high number of internal silanol groups. Both MFI and ETS-10 (*Figures* S6 and S7) show internal defect density patterns that mimic almost exactly the birefringence observed experimentally. This is a much more appealing pathway rather than the necessity to invoke complex twinning mechanisms to explain such effects. We will use this technique in future to explore a wide range of framework structures that are known to exhibit such effects.

***CrystalGrower* visualisation code** The *CrystalGrower* visualisation code is primarily written 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 *CrystalGrower* growth code and allow users to observe the morphology and surface structure of the grown crystal, whilst also manipulating the crystal in real-time. Two information files are required to construct and display the polyhedral units that compose the crystal framework, which are both generated during a simulation run by *CrystalGrower*.

Two distinct steps are taken to draw the structure. The first, using the partition information file output by *CrystalGrower*, builds a model of each polyhedral unit based on a small number of parameters for each polyhedral unit (number of vertices, number of faces, flat / non-flat faces *etc*). Once these 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, with the radius defined by taking the root mean square of the polyhedral unit xyz coordinates, multiplied by a scaling factor to allow changes to be made to the sphere size while visualising the crystal structure as spheres. The second step uses another output file generated by *CrystalGrower*, which provides the visualisation package with the xyz coordinates for each unit cell grown in a run, along with information relating to the type of the unit to draw. Coupling these coordinates with those in the partition file, the polyhedral unit is drawn at the correct position to be visualised. After completing the drawing of all 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 as it greatly improves the speed of the program, allowing real-time manipulation of a large number of objects (~300,000) on graphics cards with greater than 2 GB of dedicated graphics memory.

Supplementary figures
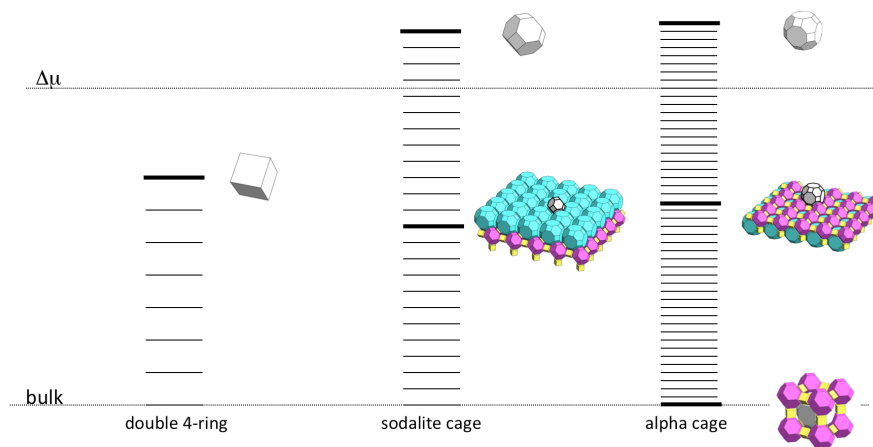
Figure S1 Schematic representation of energy ladders for the LTA structure. Double 4-rings are shown in yellow, sodalite (β) cages are shown in magenta and the α cages are shown in cyan. Each level on the energy ladders moving from bulk to the top, bold line corresponds to the loss of coordination (or 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 where 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.

FAU

Figure S2 The FAU structure consists of three tiles (a) connected in a face centred cubic lattice (b). Simulation dropping from high supersaturation (50 kcal/mol) to low supersaturation (1 kcal/mol) at the halfway point (c) matches both 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. Simulations with the same parameters as c with varying zoom distances (f) show the surface structure predicted for this simulation. Supplementary movie 1 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 x 0.20 μm x 0.20 μm.

LTA
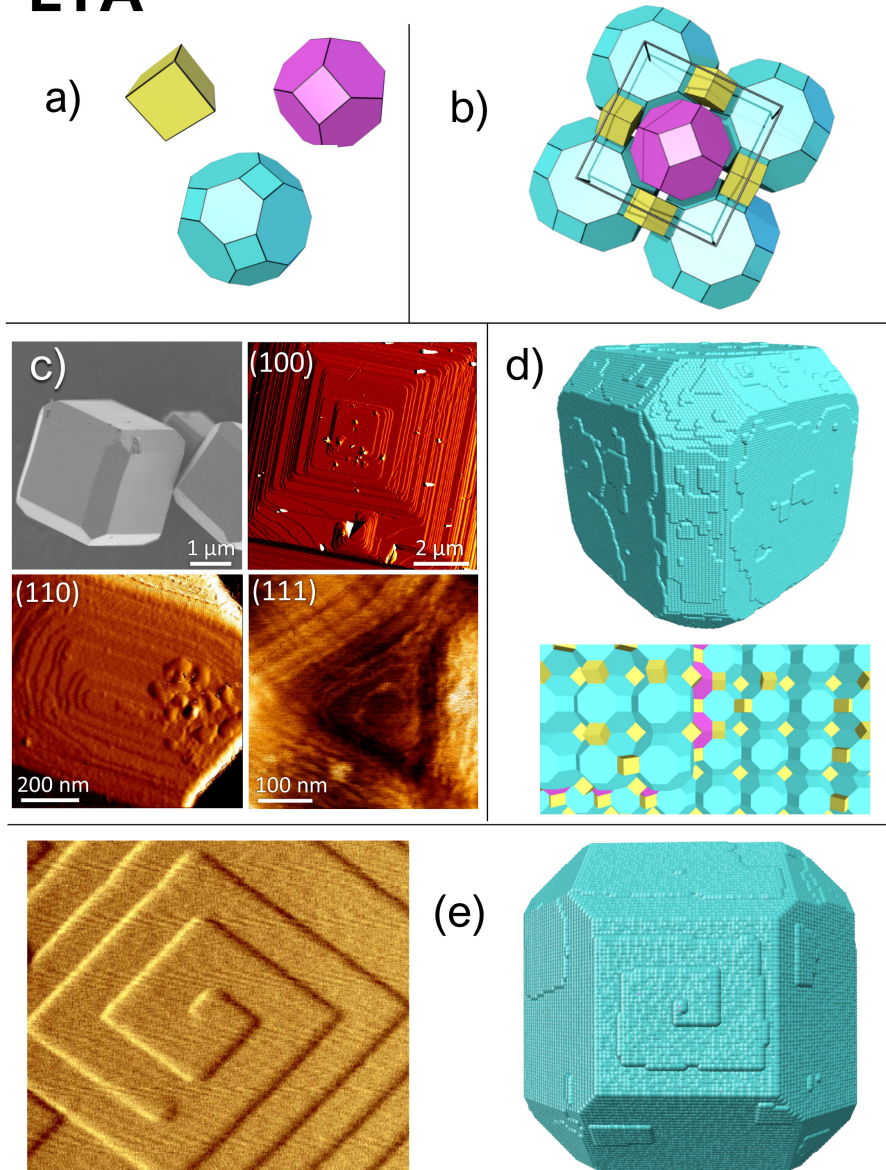
Figure S3 The LTA structure consists of three tiles (a) connected in a cubic lattice (b). Typical crystals exhibit three facets each with unique terrace shape (c) all of which can be reproduced along with the crystal habit in a simulation that treats the crystallisation energy of each tile separately (d, top), a prediction of the surface termination can also be made (d, bottom). Screw dislocations along [100] in the LTA structure are ubiquitous and the same set of parameters also reproduces the nature of this defect (e). A video of the screw dislocation growing through the simulation can be seen in supplementary movie 2. The simulated crystals shown in d and e are approximately 0.15 µm x 0.15 µm x 0.15 µm and 0.20 µm x 0.15 µm x 0.15 µm in size, respectively.

UOV
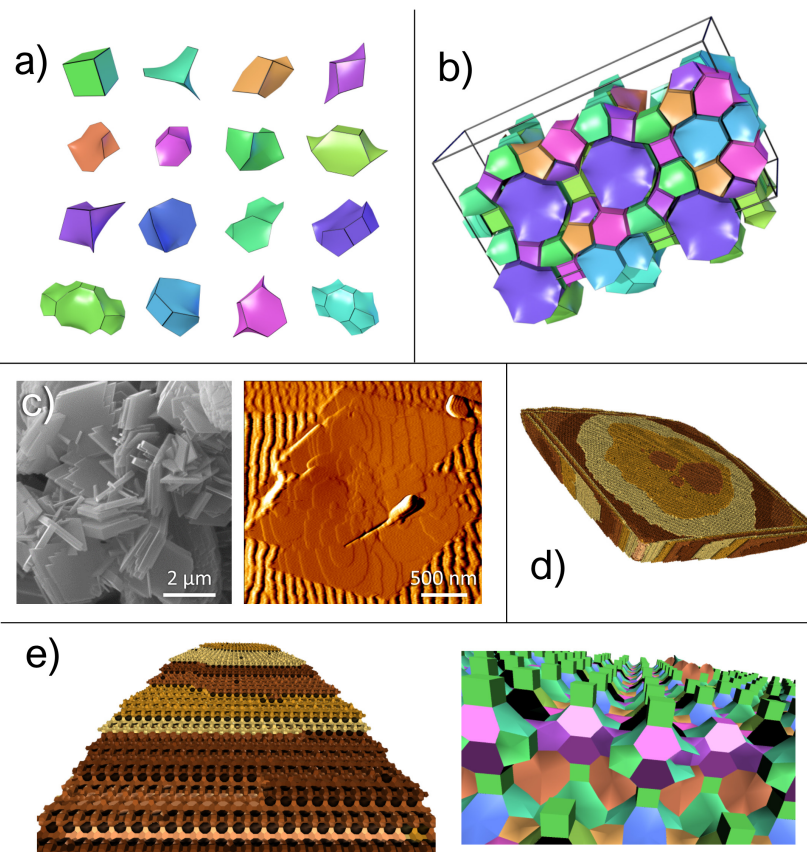
Figure S4 The UOV structure represents one of the most complex framework systems with a particularly large unit cell with c-dimension 3.87nm and consists of 16 different tile types (a), connected in an orthorhombic lattice (b). The crystallites, although intergrown (c), show very well defined shape bounded by six faces. 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). Supplementary movie 3 shows an example of how a UOV crystal grows at both high and low supersaturation values. The simulated crystal shown in all these examples is approximately 0.050 μm x 0.75 μm x 1.40 μm in size.

NES
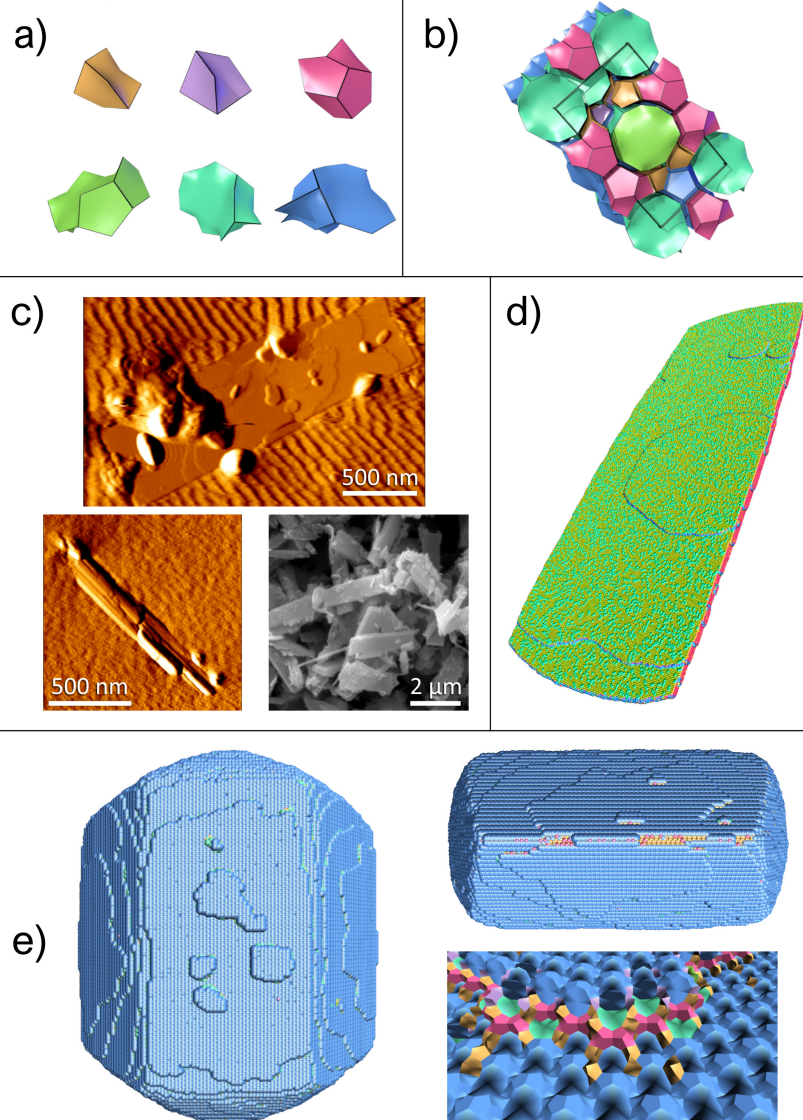
Figure S5 The NES framework can usually be expressed with four different tile types. The largest tile – t-nes can be split into three smaller tiles (a - green, cyan and blue) to give a total of six tiles possessing similar sizes (a) that combine to form an orthorhombic lattice (b). These three tiles can no longer be referred to as natural tiles, however they were still generated using the ToposPro software. These tiles play the role of units of growth and are required as *CrystalGrower* utilises connections through faces, making frameworks that share only corners and edges between small tiles, such as NES, difficult to grow over reasonable timescales. AFM and SEM micrographs show the morphology of these crystals to be long and wafer-like with rounded terraces (c) which can be reproduced with simulations using the six smaller tiles chosen (d). Lowering the energy penalty for the largest cage changes the morphology of the crystal dramatically, thickening the crystal considerably (e, 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 (e 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 x 0.40 μm x 0.15 μm, whereas the example in f) is estimated to be 0.10 μm x 0.20 μm x 0.15 μm in size, respectively.

MFI

Figure S6 The MFI structure consists of ten tile types (a) connected in an orthorhombic lattice (b). A consistent feature in MFI crystals is the optical hourglass effect seen in (c). 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 incorporated as a result of the crystal growth mechanism being different on different faces of the crystal. Our simulations (d) reflect well the crystal habit and surface topology and allows investigation of the surface termination. The internal structure (e) shows that the density of silanol groups mirrors very closely the sectoring of the crystal, as only tiles with incomplete coordination are shown. Such a change in crystal chemistry would, almost certainly, be associated with a change in crystal refractive index and hence the observed optical effect. Supplementary movie 5 demonstrates growth of the hourglass feature as the simulation progresses as well as showing how different tiles express this feature with differing degrees. The simulated crystal shown in d is approximately 0.10 μm x 0.05 μm x 0.40 μm in size, whilst the crystal shown in e is approximately 0.10 μm x 0.05 μm x 0.20 μm.

ETS-10
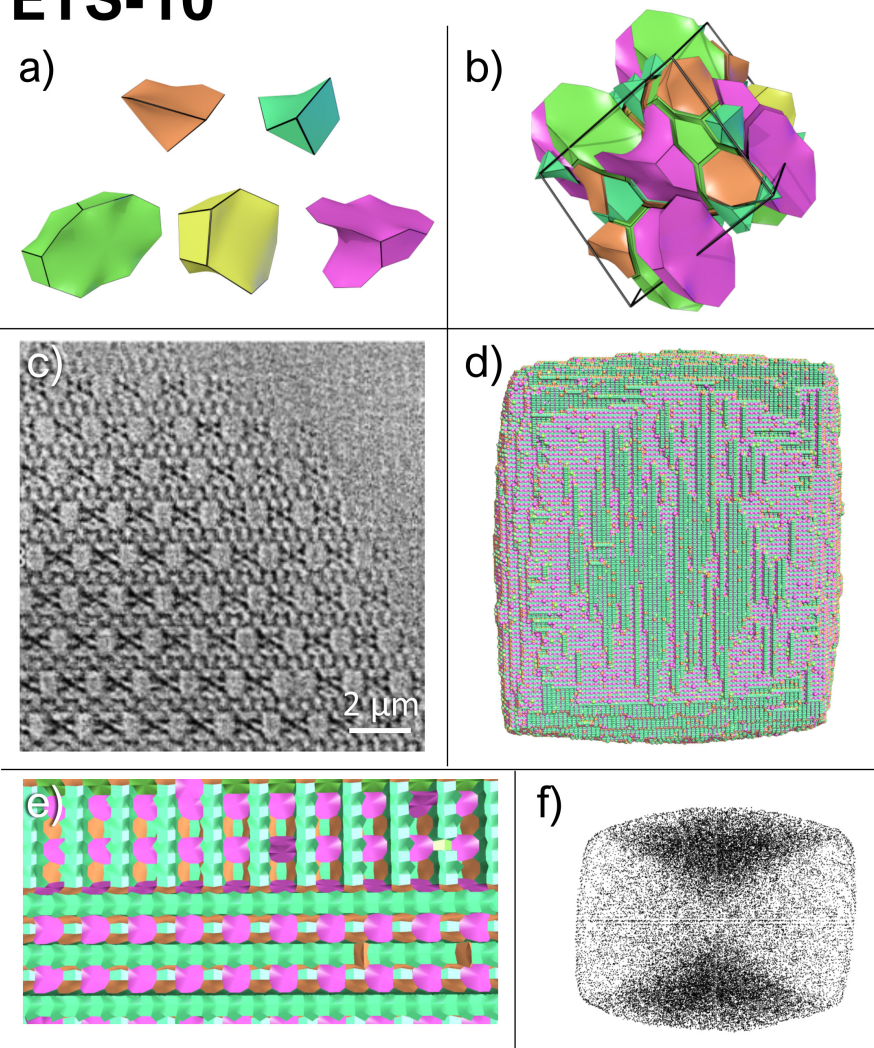
Figure S7 ETS-10 is a nanoporous titanosilicate with a structure consisting of five tiles (a) connected in a monoclinic lattice (b). The structure has a very similar symmetry to zeolite beta, however, the titanium is incorporated into titania rods surrounded by silica (c,d) – HREM courtesy of O. Terasaki. These rods run alternatively in orthogonal [110] and [-110] directions and it is found that these rods are the dominant components for the crystal growth (f). Further, because the rods do not always connect and heal inside the crystal structure there is a high concentration of internal silianol groups in a sector from the (001) faces to the centre of the crystal (e). Such sectoring in ETS-10 has been observed in Raman microscopy[24]. A movie demonstrating how the rods in the ETS-10 framework grow in alternating orthogonal directions can be viewed in supplementary movie 4. The simulated crystals shown in d and e are approximately 0.25 μm x 0.25 μm x 0.05 μm, and 0.10 μm x 0.10 μm x 0.10 μm in size, respectively.

CHA

a)

b)

c)

5 μm

d)

e)

200 nm

f)
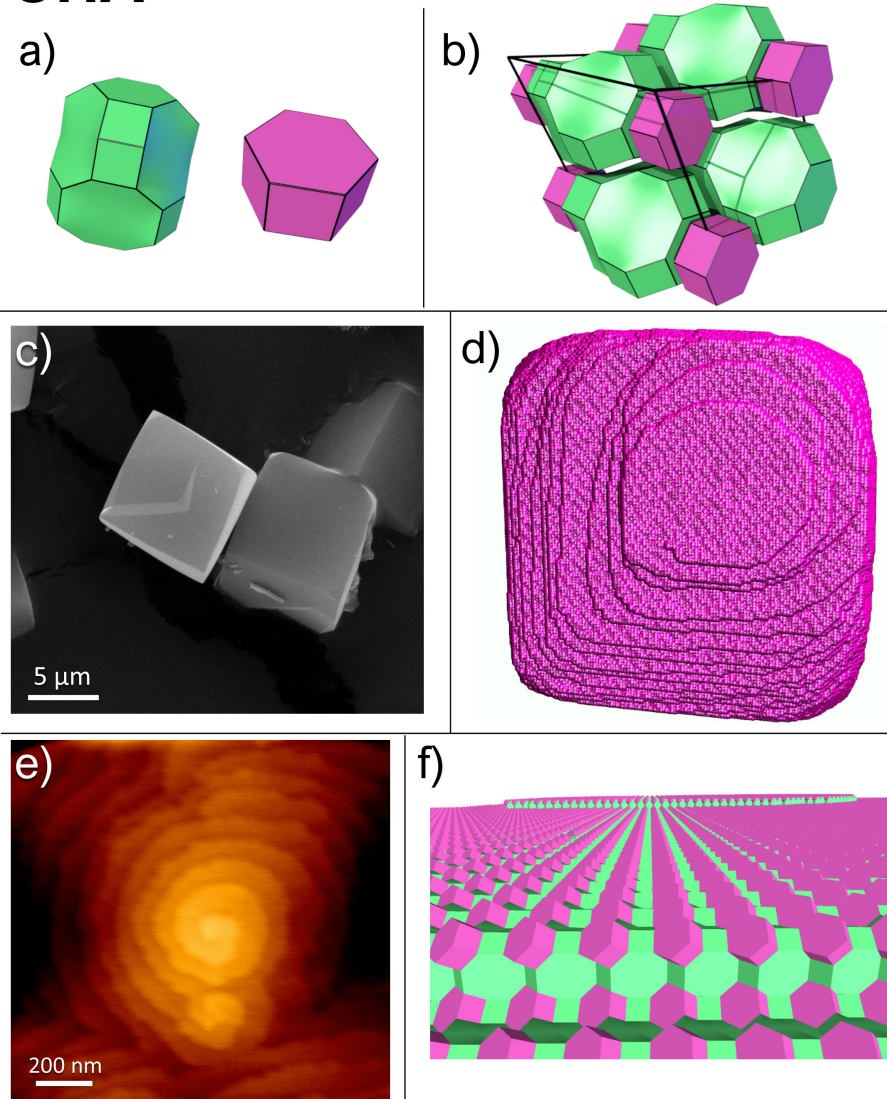
Figure S8 CHA structure consists of two tiles (a) connected in a monoclinic lattice (b). Typical crystals show distorted cube morphology when viewed under a SEM (c) exhibiting <100> faces with isotropic growth of terraces. A plethora of screw dislocations can be observed in this system (e) and this is reproduced for a dislocation with a screw core running along [100] (d). Simulations predict that the surface terminates at the double six rings (f). The simulated crystal shown in d has an approximate size of 1.50 μm x 0.10 μm x 0.10 μm due to elongation caused by the screw dislocation.

AEI
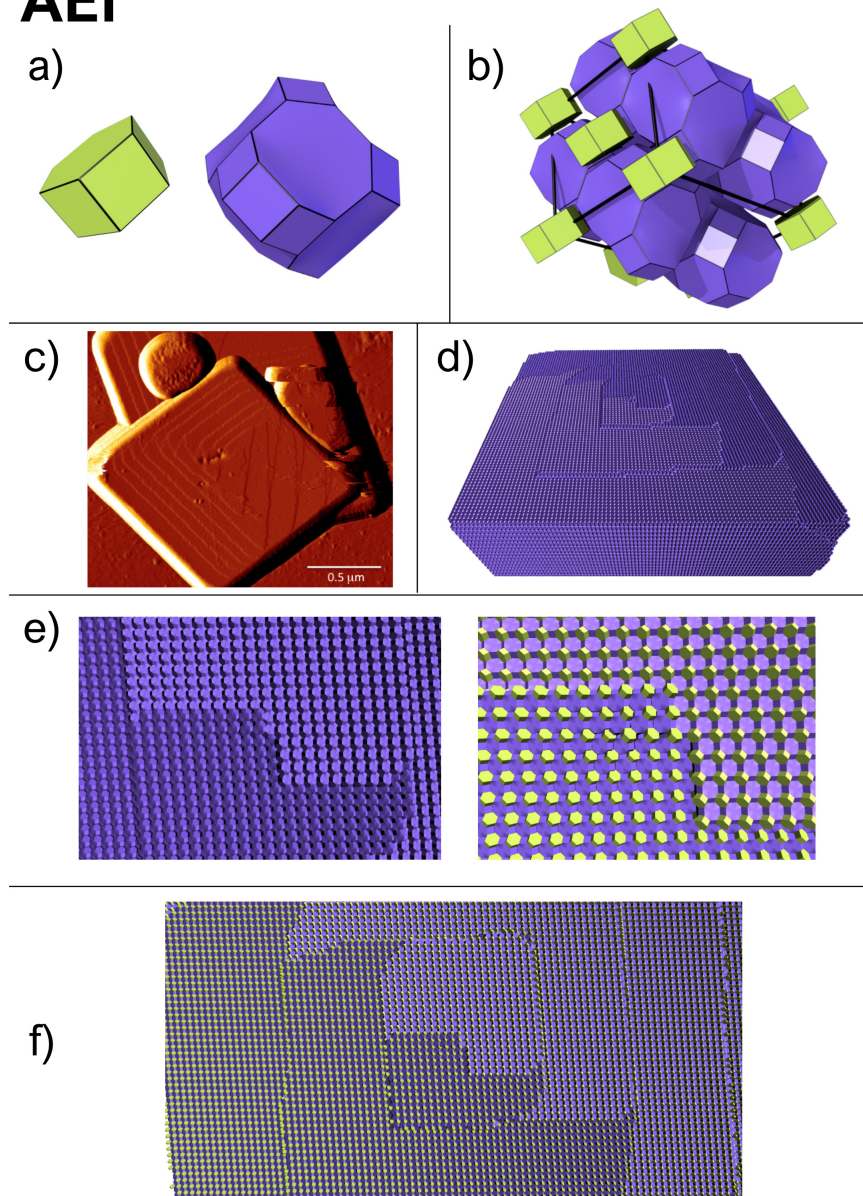
Figure S9 AEI structure consists of two tiles (a) connected in an orthorhombic lattice (b). The structure is related to CHA in that it can be built by connecting double six rings. However, because of the alternating orientation of these units along the c-direction this results in a switching of the fast and slow growth rates along the a-direction and b-direction (c). The result is a complex interleaved spiral growth that is faithfully simulated according to this space partitioning methodology (d–f). The simulated crystal shown in d – f is approximately 1.70 μm x 0.05 μm x 0.05 μm in size due to elongation caused by the screw dislocation.

HKUST-1
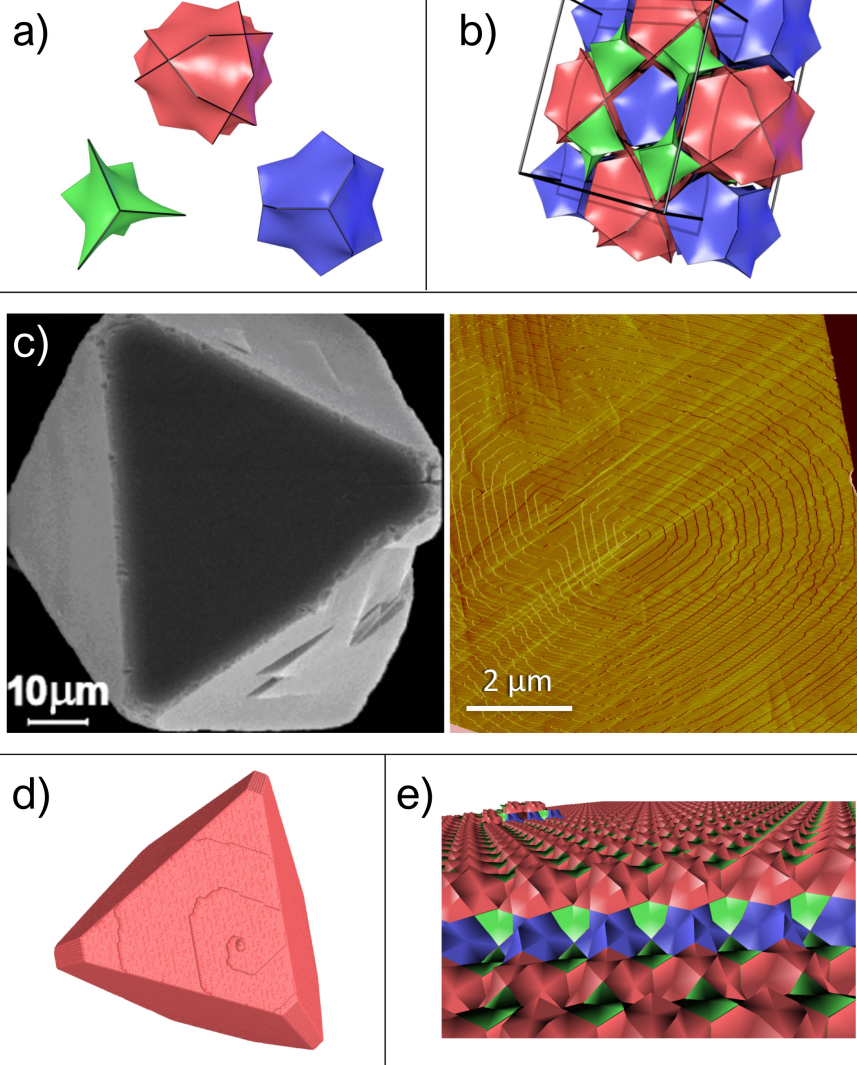
Figure S10 HKUST-1 is a metal-organic framework copper trimesate $Cu_3[(O_2C)_3C_6H_3]_2(H_2O)_3$. By partitioning the structure with nodes at metal clusters and tile edges along linkers the HKUST-1 structure consists of three tiles (a) connected in a face centred cubic lattice (b). Such a partitioning results in excellent simulation 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 as such a screw core would result in multiple spirals (d). Simulations also allow investigation of the surface termination for this framework. A video of the screw dislocation growing and migrating across the surface of the crystal can be viewed in supplementary movie 6. The approximate size of the crystal shown in d and e is 0.35 μm x 0.35 μm x 0.35 μm.
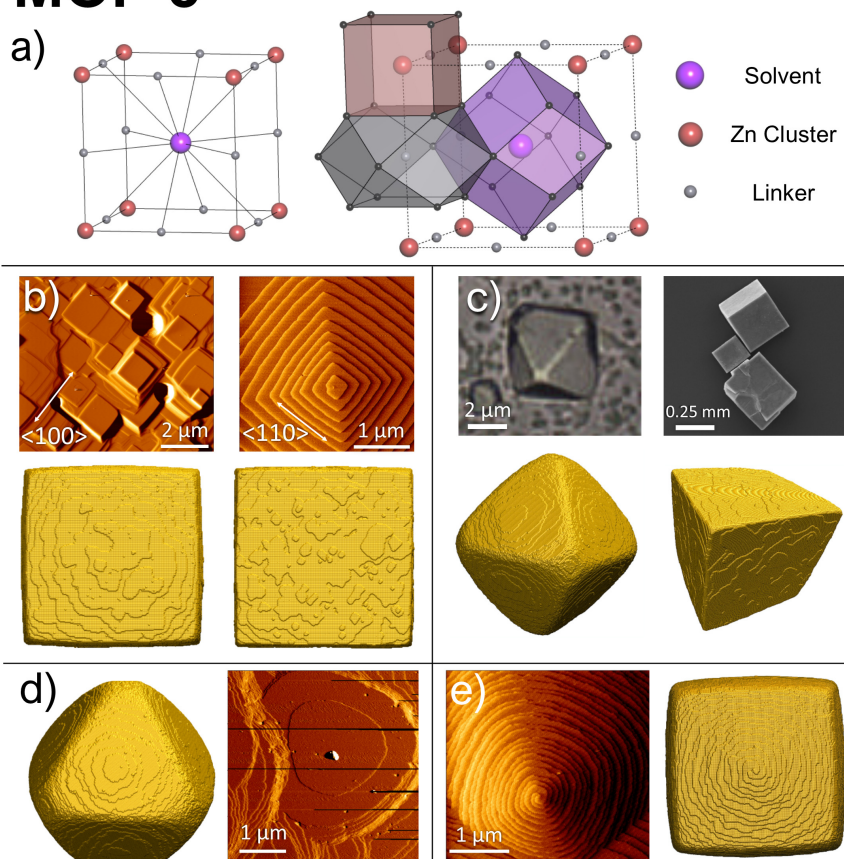
MOF-5

Figure S11 MOF-5 [Zn$_4$O(bdc)$_3$, bdc=1,4-benzenedicarboxylate] is a permanently porous metal-organic framework with a simple cubic arrangement of metal centres (red) and linkers (grey), along with a solvent molecule occupying the centre of each cube (purple) (a, left). Treating the structure as a multi-component 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 and solvent (a, right). Square terracing on the (100) face (b) can be in one of two orientations, dependent on the synthesis conditions. The crystal morphology can also be changed to exhibit (111) faces as an octahedral crystal, again dependent on synthesis conditions (c, left), as opposed to only (100) faces (c, right) in a cubic crystal. (111) faces exhibited by the crystal are much more isotropic (d) compared to the (100) faces (b) and do not adopt different orientations demonstrated by the (100) faces. The simple cubic network alone will not permit the development of (111) facets and is immediately an indication of the importance of the solvent N,N'-diethylformamide (DEF) in the preparation. By varying driving forces for binding of the linker, metal centre and solvent in the final crystal all crystal habits and surface topologies may be generated (b – d). Double spirals through growth at screw dislocations on the (100) face can also be simulated (e). The estimated sizes of the crystals shown in b are 0.76 μm x 0.75 μm x 0.75 μm and 0.70 μm x 0.80 μm x 0.80 μm, c are 0.50 μm x 0.50 μm x 0.50 μm and 0.75 μm x 0.75 μm x 0.75 μm, d is 0.50 μm x 0.50 μm x 0.50 μm, and e is 0.75 μm x 0.75 μm x 0.75 μm, respectively.
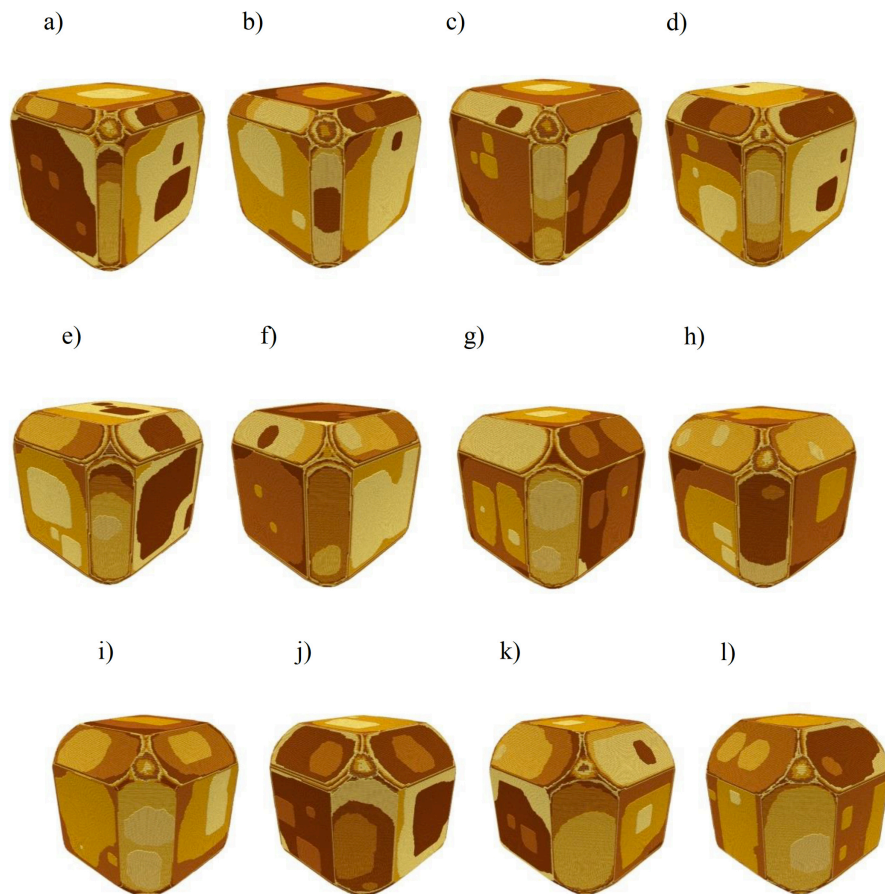
Figure S12 A series of equilibrium images of LTA that demonstrate the sensitivity of the model where in all cases U_{D4R}, = 8 kcal mol^{-1} and U_α = 1 kcal mol^{-1} whilst U_β is varied as follows a) U_β = 1.84 kcal mol^{-1} b) U_β = 1.85 kcal mol^{-1} c) U_β = 1.86 kcal mol^{-1} d) U_β = 1.87 kcal mol^{-1} e) U_β = 1.88 kcal mol^{-1} f) U_β = 1.89 kcal mol^{-1} g) U_β = 1.90 kcal mol^{-1} h) U_β = 1.91 kcal mol^{-1} i) U_β = 1.92 kcal mol^{-1} j) U_β = 1.93 kcal mol^{-1} k) U_β = 1.94 kcal mol^{-1} l) U_β = 1.95 kcal mol^{-1}. Approximate crystal size is 0.5 µm^3..
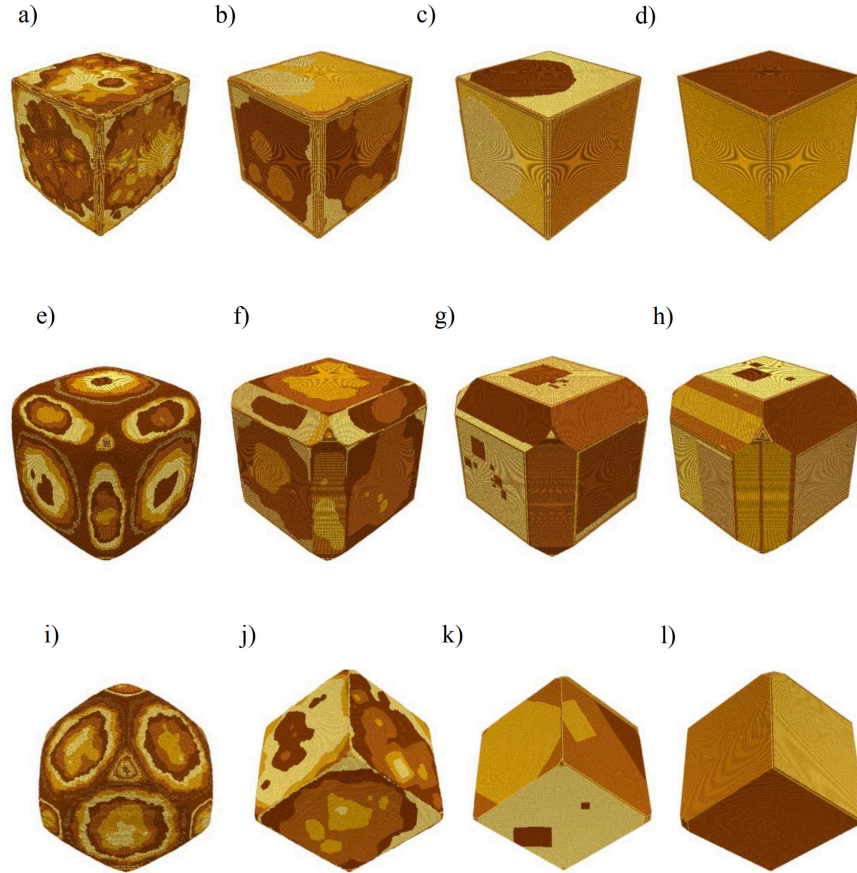
Figure S13 Three series of images of LTA crystals at equilibrium driving force with the same energy ratios a) – d) in the ratio 1:1:1 where a) U_{D4R} = 1 kcal mol^{-1} U_β = 1 kcal mol^{-1} and U_α = 1 kcal mol^{-1} b) U_{D4R} = 2 kcal mol^{-1} U_β = 2 kcal mol^{-1} and U_α = 2 kcal mol^{-1} c) U_{D4R} = 3 kcal mol^{-1} U_β = 3 kcal mol^{-1} and U_α = 3 kcal mol^{-1} d) U_{D4R} = 4 kcal mol^{-1} U_β = 4 kcal mol^{-1} and U_α = 4 kcal mol^{-1} e) – h) in the ratio 5:5:1 where e) U_{D4R} = 2.5 kcal mol^{-1} U_β = 2.5 kcal mol^{-1} and U_α = 0.5 kcal mol^{-1} f) U_{D4R} = 5 kcal mol^{-1} U_β = 5 kcal mol^{-1} and U_α = 1 kcal mol^{-1} g) U_{D4R} = 7.5 kcal mol^{-1} U_β = 7.5 kcal mol^{-1} and U_α = 1.5 kcal mol^{-1} h) U_{D4R} = 10 kcal mol^{-1} U_β = 10 kcal mol^{-1} and U_α = 2 kcal mol^{-1} i) – l) in the ratio 2:2:1 where a) U_{D4R} = 1 kcal mol^{-1} U_β = 1 kcal mol^{-1} and U_α = 0.5 kcal mol^{-1} b) U_{D4R} = 2 kcal mol^{-1} U_β = 2 kcal mol^{-1} and U_α = 1 kcal mol^{-1} c) U_{D4R} = 4 kcal mol^{-1} U_β = 4 kcal mol^{-1} and $\underline{U}_\alpha$ = 2 kcal mol^{-1} d) U_{D4R} = 8 kcal mol^{-1} U_β = 8 kcal mol^{-1} and U_α = 4 kcal mol^{-1}. Approximate crystal size is 0.5 µm^3.
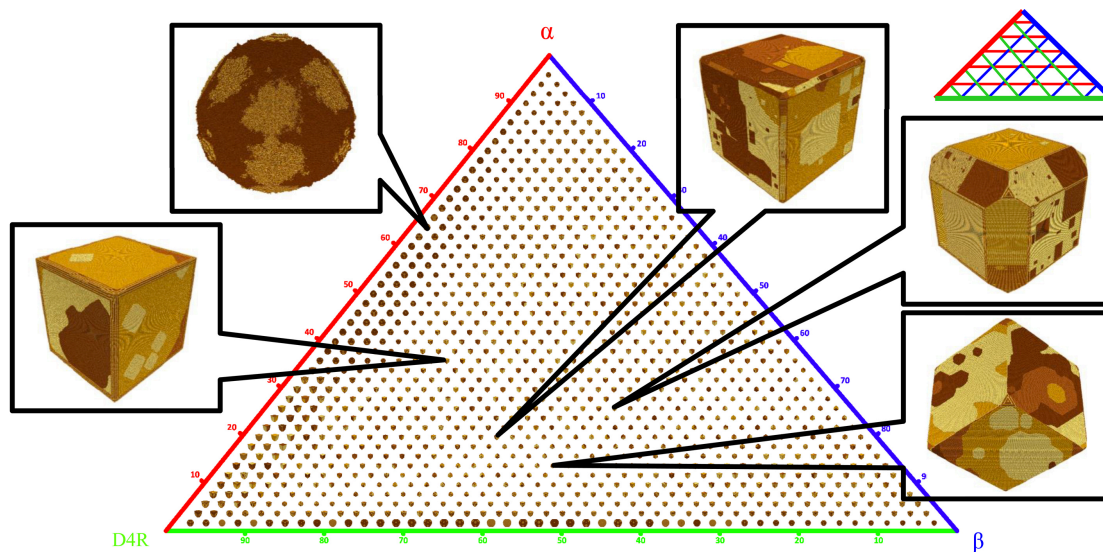
Figure S14 A ternary plot consisting of 1,176 simulation data points of LTA corresponding to exploring the energy space at 2% intervals, all images are at equilibrium driving force. The highest destabilisation for each cage is at the corner for which its name appears and the corresponding axes are colour 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. All crystals are at equilibrium driving force
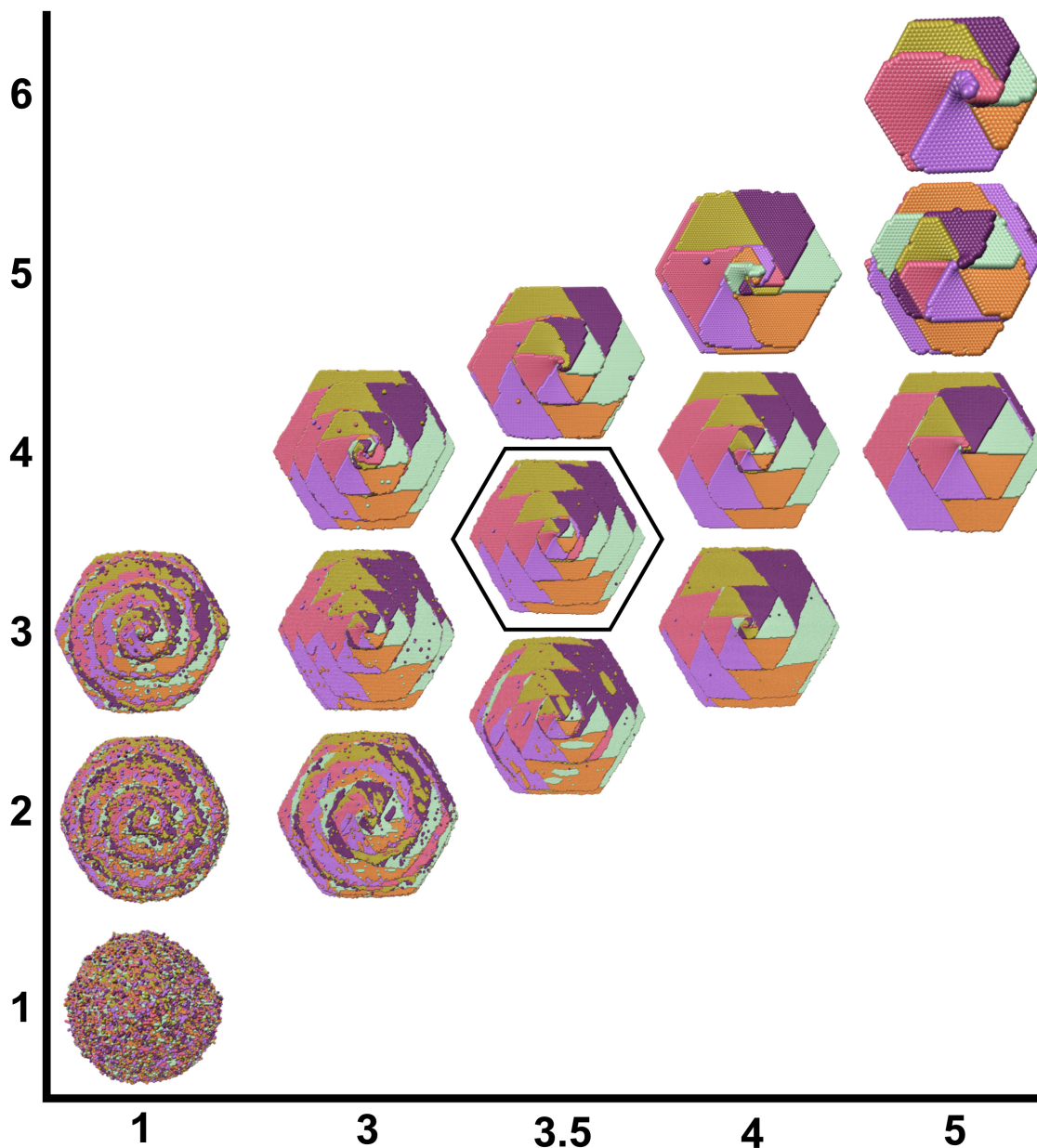
Figure S15 A suite of calculations of the crystal growth of L-cystine at supersaturation 0.6 kcal mol^{-1} viewed down the [001] direction showing characteristic pin-wheel surface topology and hexagonal step bunching. Simulations are based on four characteristic free energies of crystallisation: (i) 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, known as A+ and A- in references 29 and 30, with one hydrogen bond; (iv) in the c-direction with two hydrogen bonds. Interaction type (ii) and (iii) are maintained at half that of (iv) with a minimal difference of (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 in kcal mol^{-1}. When all these interactions are too weak isotropic sticky growth is observed and when they are too high the growth is too geometrical. The optimum balance is highlighted in the figure where the sum of the free energies of crystallisation is 7.0 kcal mol^{-1}, compared to 4.3 kcal mol^{-1} derived directly from the solubility of L-cystine at pH 7[31]. Similar to that reported in reference 29, the step bunching is closer than observed experimentally as lattice strain at the screw core will weaken interactions and prevent the core region, which is advancing by attachment to the slow step alone, forming a closed hexagonal step bunch as quickly as predicted.
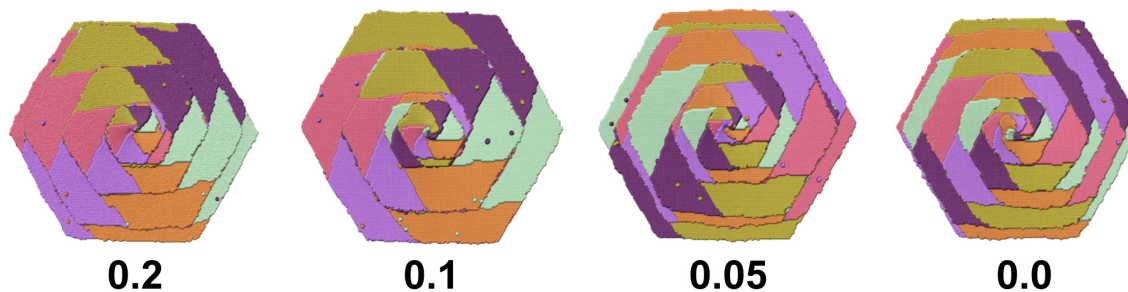
0.2　　　　　**0.1**　　　　　**0.05**　　　　　**0.0**

Figure S16 Simulations show the effect of small differences in the slow growth directions termed A+ and A- in references 29 and 30, noted on the figure 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 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.
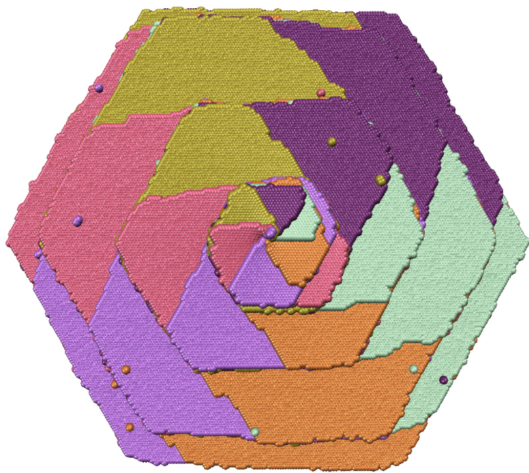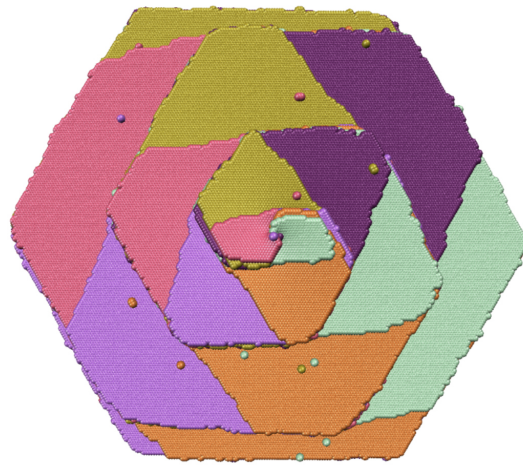
Normal Spiral   **Reverse Spiral**

Figure S17 Reversal of the handedness of the screw core is not immediately apparent as 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 whether the handedness of the crystal structure is the same or the reverse of the handedness of the screw core.
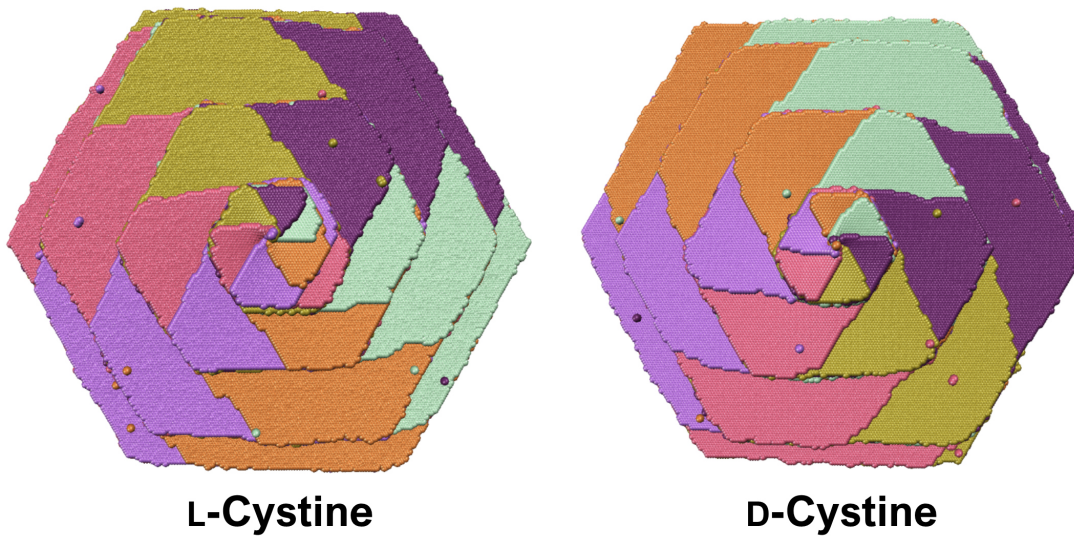
L-Cystine **D-Cystine**

Figure S18 L-cystine and D-cystine will present identical surface structures with opposite handedness, as expected.

Supplementary movies

1   FAU

2   LTA

3   UOV

4   ETS-10

5   MFI

6   HKUST-1

7   L-Cystine

FAU – Layer Growth – Supplementary Movie 1

A FAU crystal is grown under high supersaturation conditions (50 kcal/mol) for 5 million iterations, showing rapid growth and the adoption of an octahedral crystal morphology. The second 5 million iterations occur after the supersaturation is lowered to slightly above equilibrium (1 kcal/mol), allowing the growth process to continue slowly. At this point the triangular shape of the terraces can be observed, growing in the opposing direction to the facet where the growth is occurring. The final portion of the movie shows the surface structure viewed from the top, followed by the side of a grown terrace.

LTA – Growth via Screw Dislocation – Supplementary Movie 2

In this movie, an LTA crystal is grown with a screw dislocation running in the [100] direction through the crystal. Similar to the FAU simulation, the crystal is first grown under high supersaturation conditions (90 kcal/mol) for 5 million iterations, followed by a drop in supersaturation to slightly above equilibrium (1.5 kcal/mol) to allow a more detailed study of the growing process. During the slow growth, the propagation of the spiral from the screw dislocation is clear. The crystal is then rotated to the opposing face, demonstrating the elongation of the morphology in the [100] as expected as the screw dislocation creates a low energy surface for the growth units to attach to. After arriving at the opposing face, the slow growth process is reversed then replayed to again demonstrate the spiral-growth of the crystal.

UOV – Layer Growth – Supplementary Movie 3

Growth of a UOV crystal is demonstrated using similar conditions to the previous movies: 5 million iterations at high supersaturation (70 kcal/mol) followed by 5 million iterations at a supersaturation value slightly above equilibrium (1 kcal/mol). The diamond-like morphology of the UOV crystal can clearly be seen during panning and rotation. Rounded / isotropic layer growth can be observed during the growth process, along with the fusion of terraces once the growth process is slowed. Close-up shots are then shown of the surface structure on the (100) face, followed by the side-wall of the crystal: the (013) face.

ETS-10 – Orthogonal Rod Growth – Supplementary Movie 4

This movie demonstrates the growth mechanism of ETS-10 *via* titanate rods in alternating directions, orthogonal to each other. Each alternating layer is represented by different colours to draw attention to the orthogonal directions. The conditions for this growth process are changed slightly compared to the previous movies. The crystal is first grown at high supersaturation (140 kcal/mol) to ensure the growth process begins, however no frames are recorded. The supersaturation is then lowered to 1 kcal/mol, and frames are taken every 200,000 iterations for the first 20 frames, followed by 40 frames recorded at every 10,000 iterations. During these slowed down frames, a closer view of the surface is shown, with arrows overlaid following the direction of rod growth that can be observed. A final 20 frames with a spacing of

20000 iterations are shown, along with close-ups of the surface structures adopted by the vertical and horizontal rods that express the structure.

MFI – Hourglass Growth and Tile Type Relation – Supplementary Movie 5

To begin, the t-mfi-1 tiles are shown during the growth process of an MFI crystal to allow a view into the internal structure of the crystal as it grows, showing the formation of the hourglass structure mentioned previously. Following this growth, a series of frames are shown where the tile type shown is cycled on the (001), (100) and (001) faces respectively, demonstrating how different tile types contribute to this internal feature in varying degrees.

HKUST-1 – Growth via Screw Dislocation – Supplementary Movie 6

This supplementary movie shows the growth of an HKUST-1 crystal under similar conditions to the FAU, LTA and UOV crystals simulated previously. The first 5 million iterations are run at high supersaturation conditions (100 kcal/mol) followed by a drop to a lower supersaturation (1 kcal/mol). Slowing down the growth clearly demonstrates how the screw dislocation running along the [110] direction completely alters the final morphology of the crystal. The growth is then reversed, and the screw dislocation is followed to the point where it migrates back onto the (111) that it originally appeared from. Rotating the crystal then allows the viewer to observe the elongation of the crystal shape caused by the spiral growth on both sides. Rotating the crystal to view the opposite side of the screw dislocation, the location of the core is again followed whilst re-growing the crystal at low supersaturation (1 kcal/mol) ending with the final growth frame, and a panned out view of the entire crystal.

L-Cystine – Growth via Screw Dislocation – Supplementary Movie 7

This supplementary movie shows the growth L-cystine at a supersaturation of 0.6 kcal mol^{-1}. The optimal binding free energies where: (i) strong binding 3.5 kcal mol^{-1}; (ii) and (iii) weak binding 0.78 kcal mol^{-1} and 0.98 kcal mol^{-1}; (iv) binding in the c-direction 1.75 kcal mol^{-1}. Advancement of the step bunches is through birth-and-spread growth on the <100> side faces, which is encouraged through the strong binding direction. This birth-and-spread growth precipitates growth at the single slow growth edges. Growth is further complicated by a small amount of birth-and-spread growth on the <001> face.