Fundamentals of Crystallisation

Richard Sear

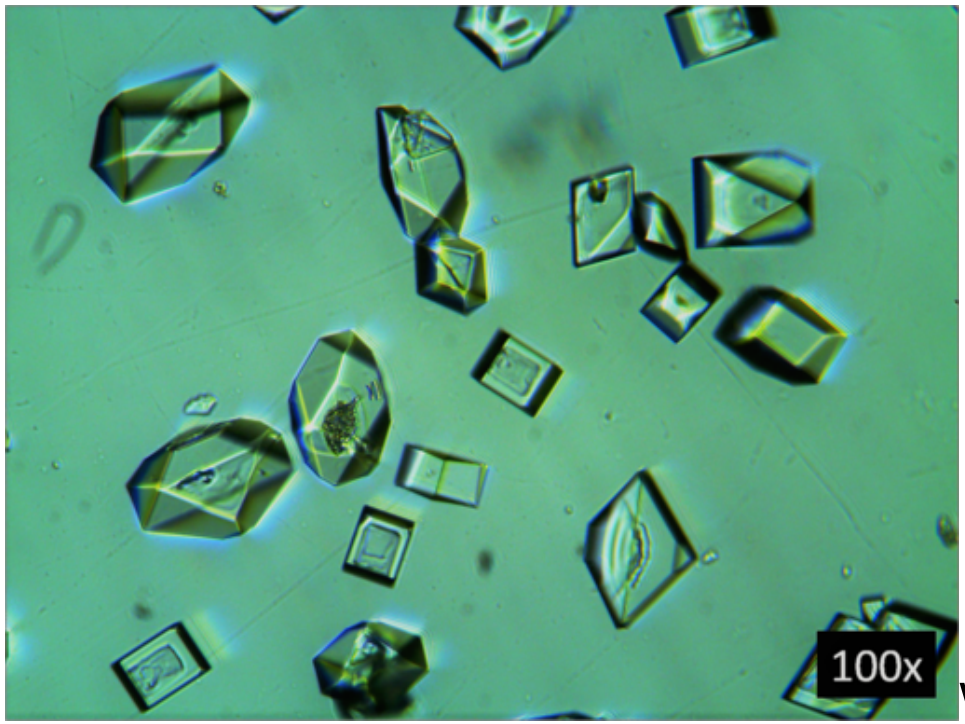
University of Surrey

RAMP (MSCA ITN) School, Ireland May 2018

Outline

Introduction: Crystals

Crystals: substances where the molecules are arranged in regular repeating lattice

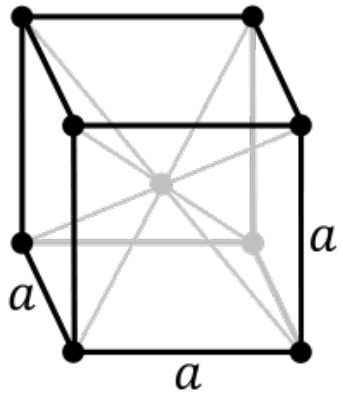


Wikimedia

We want to control crystallisation.

The defining feature of crystals is their crystal lattice

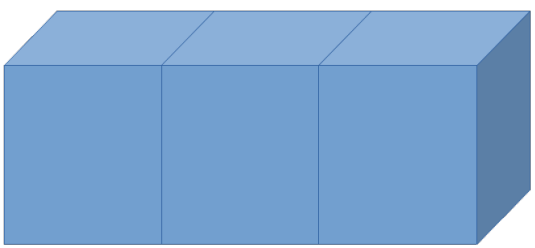
In, eg BCC, each molecule has 8 neighbours arranged at corners of cube. This is a statement about local ordering.



BCC

Wikimedia

This arrangement repeats periodically, for a macroscopic single crystal, over billions of molecules. This is a statement about long-ranged translational ordering, and hence about Bragg peaks in XRD.



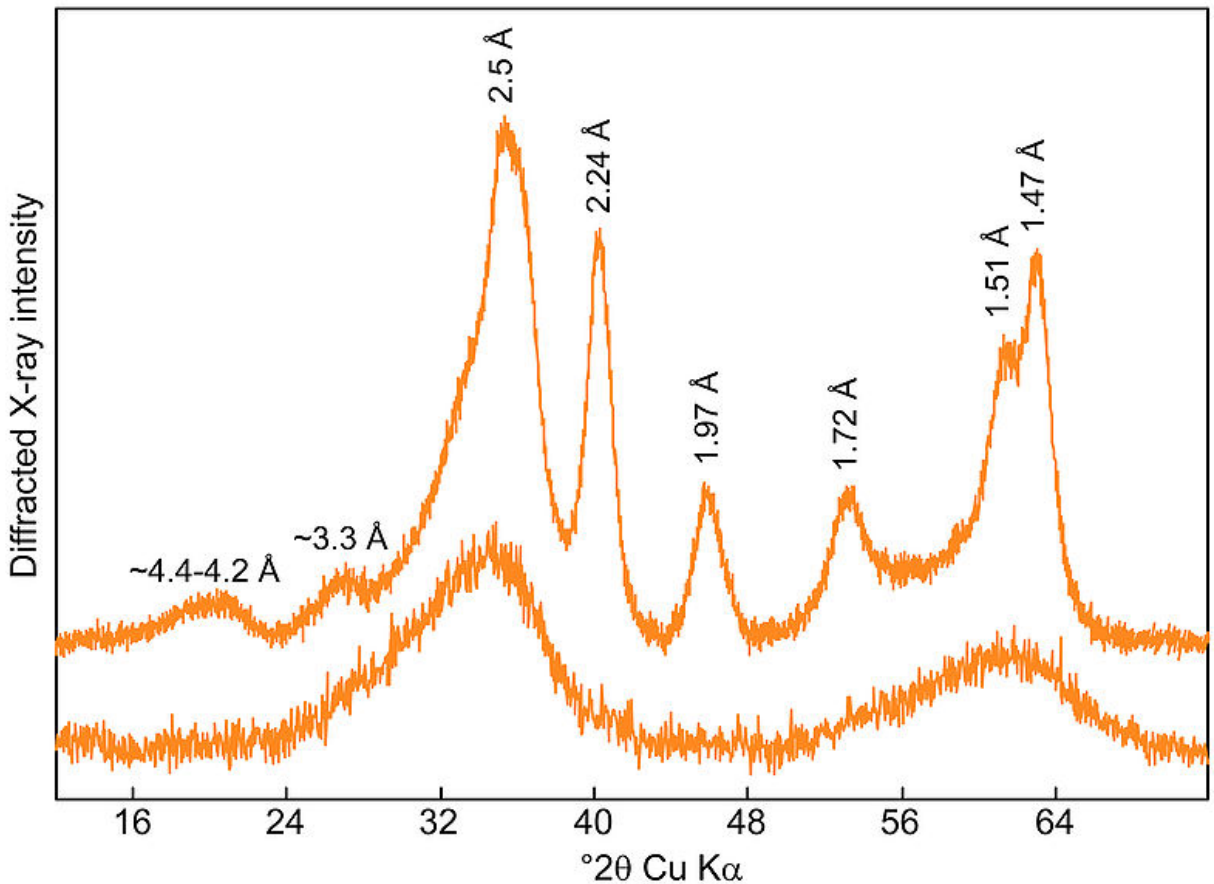
Some substances form huge crystals

Gyspsum (calcium sulphate), Naica, Chihuahua, Mexico (Wikimedia)

Note human for scale, and that these took many thousands of years to grow.

Some substances seem never to form crystals more than nms across

XRD patterns of two forms of ferrihydrite: bottom is "two-line" ferrihydrite, top is "six-line" ferrihydrite.



Drits *et al.* Clay Min. 1993

"two-line" ferrihydrite (bottom) is amorphous

"six-line" ferrihydrite (top) is kind of a crystal

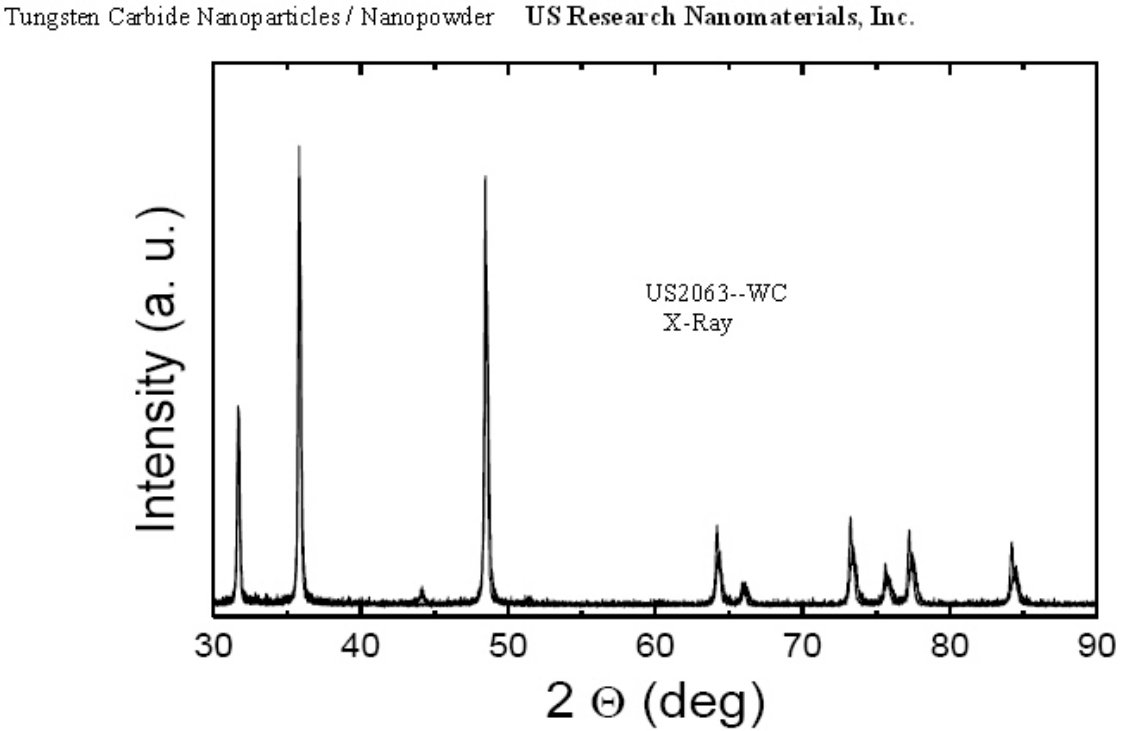
Some substances can be easily grown as large crystals, some have never been grown as large crystals

For poorly understood reasons, the facility with which things crystallise varies enormously.

Some crystals may stop growing due to composition fluctuations, lattice defects forming and causing strain, impurities being incorporated into the lattice, ...
Better news: some molecular interactions are known to result in rapid crystallisation and the formation of large crystals, eg, spherical molecules (atoms) do this
make proteins interact as spheres using polymers ...

X-Ray Diffraction (XRD) of crystals

Powder XRD, $S(q)$, pattern of tungsten carbide



US Research Nanomaterials Inc.

Bragg peaks due to long-range translational order of crystal lattice
Here will consider powder and similar XRD patterns. Full single-crystal pattern is 3D object but I will not consider that here.

Disorder in crystals

An infinite crystal with perfect ordering of the molecules with give an XRD pattern with perfectly sharp (delta function) peaks, from which a perfect structure can be obtained.

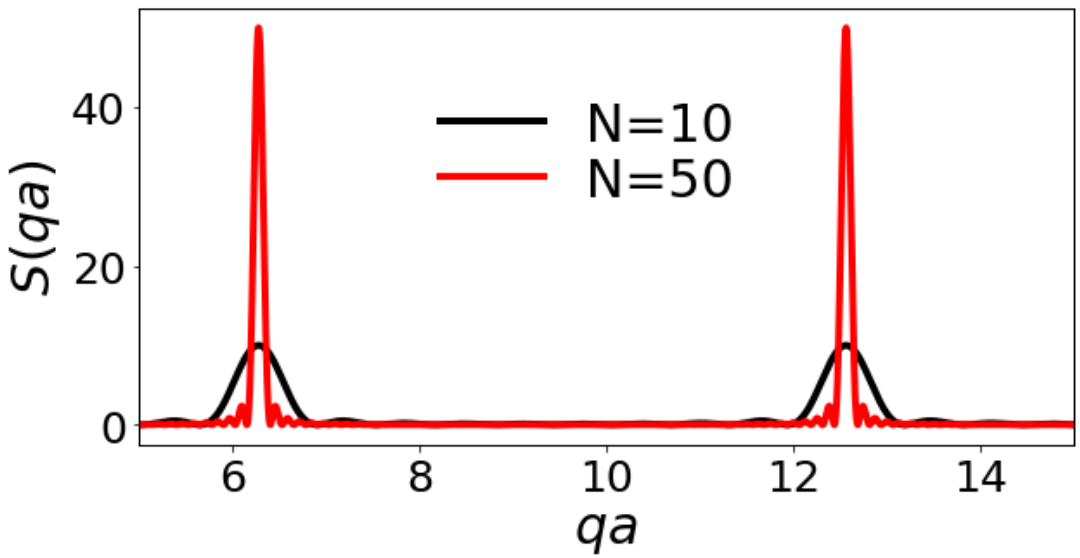
In the real world, no crystals are infinitely big and in no crystal are the molecules perfectly ordered $\rightarrow$ XRD peaks are broadened, and have lowered intensity, and this limits the resolution of the protein structure that can be obtained.

Crystals are finite

To the right, black curve is for 10 perfect crystal planes, red curve is for 10 perfect crystal planes.

Note that for perfect ordering, the more crystal planes in the lattice. the narrower the peaks.

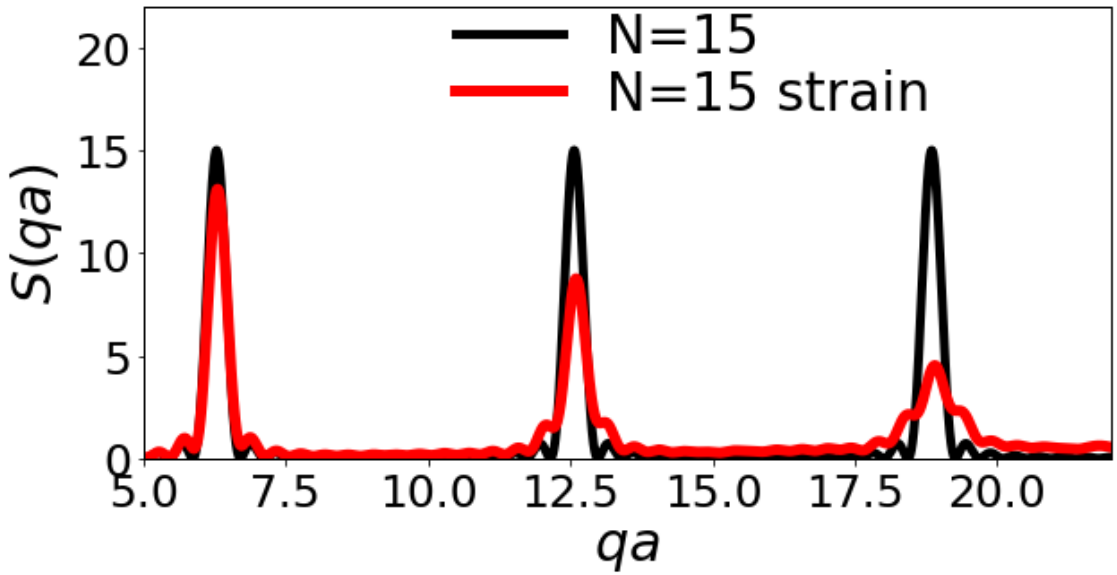
This was appreociated by Paul Scherrer, who realised that peak width (FWHM) varies as 1/(number of lattice planes) — for perfect lattice.



Structure factors for perfectly ordered stacks of $N = 10$ and 50 planes. Peak width $\sim 1/N$. a = lattice constant.

Crystals are disordered

In real crystals there is disorder: the molecules are not perfectly in step, the periodicity of the arrangement of the molecules is not perfect.



Structure factors for a perfectly ordered stack of $N = 15$ planes, and for a stack of $N = 15$ planes with strain from surface + disorder. a = lattice constant.

Disordered crystals

Note that if successive planes slip out of perfect periodicity by some amount δ (=length), this broadens peak widths by an amount

$$\text{broadening} \sim q_{PEAK}\delta$$

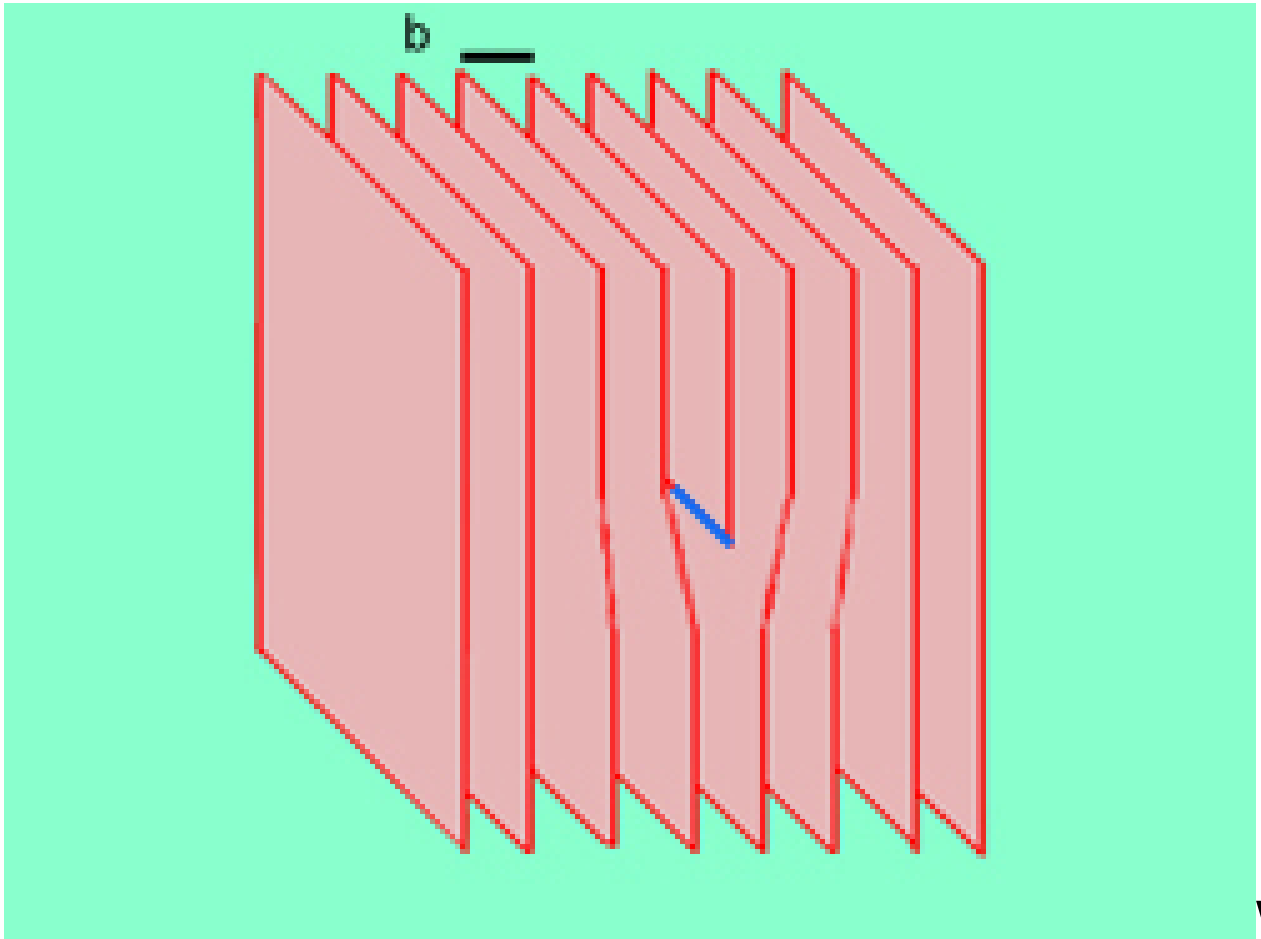
i.e., broadens peaks at high q much more than small q peaks — different from pure finite-size effects that broaden all peaks equally.

Disorder that limits periodicity characteristically suppresses large q peaks

Simple thermal motion about position in perfect lattice only reduces peak height (Debye-Waller factor).

Dislocations strain and disrupt the lattice, broadening peaks

During growth, crystal planes may end or begin, creating large defects in the crystal lattice: dislocations.



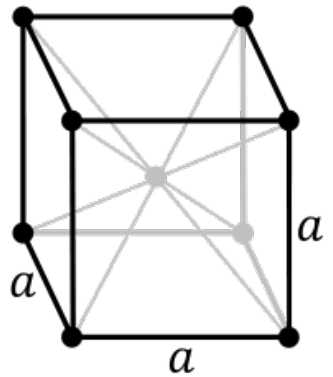
Wikimedia

Dislocations will also reduce the information in an XRD pattern

Many substances can form more than one crystal lattice: Polymorphism

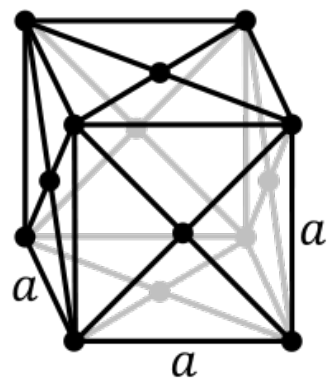
Many substances crystallise into more than one crystal lattice.

We don't understand the competition between polymorphs well enough to make quantitative predictions for how they compete or how such competition may prevent large single crystals of either polymorph from forming



BCC

Wikimedia



FCC

Wikimedia

Summary for: Crystals

1. No crystal is infinite and perfectly ordered
2. Disorder limits the information in XRD patterns
3. Disorder that weakens the correlations in lattice spacing weakens large q peaks

Questions? Thoughts?

Crystallisation: Nucleation then Growth

Crystallisation is the process of going from no crystal, eg just supersaturated solution, to a crystal, or more commonly many crystals.

Crystallisation = Nucleation then Growth??

Crystal Nucleation

Nucleation time t_{NUC} is how long we have to wait for a crystal to first appear: ranges from < 1 s, to $>$ age of universe — i.e., in some cases crystals appear instantly while in other crystallisation never happens.

t_{NUC} often found to be very sensitive to: supersaturation, impurities, surfaces, mechanical shock, ...

Crystal Growth

Growth is a pre-existing crystal becoming larger by adding additional crystal planes

Crystalline silver can grow at up to 100 m/s into supercooled liquid silver. In other systems growth can arrest: growth rate = 0, or it can take seconds to add a single extra crystal plane - contrast this with the μs it takes a protein to diffuse its own diameter

Nucleation and growth not independent, formation of the initial nucleus must involve growth

Crystallisation has been studied for thousands of years

Crystallisation has been studied for thousands of years and in systems from iron to water to plastics to food ....

Wikimedia: Alexander Mosaic (from Pompeii, now in Napoli Arch. Museum)
Alexander the Great & Darius both relying on crystalline materials

know a lot about crystallisation, that appears to apply to most crystallising systems, We can apply that knowledge to membrane proteins.

Crystallisation has been studied for thousands of years

Wikimedia: cirrus clouds

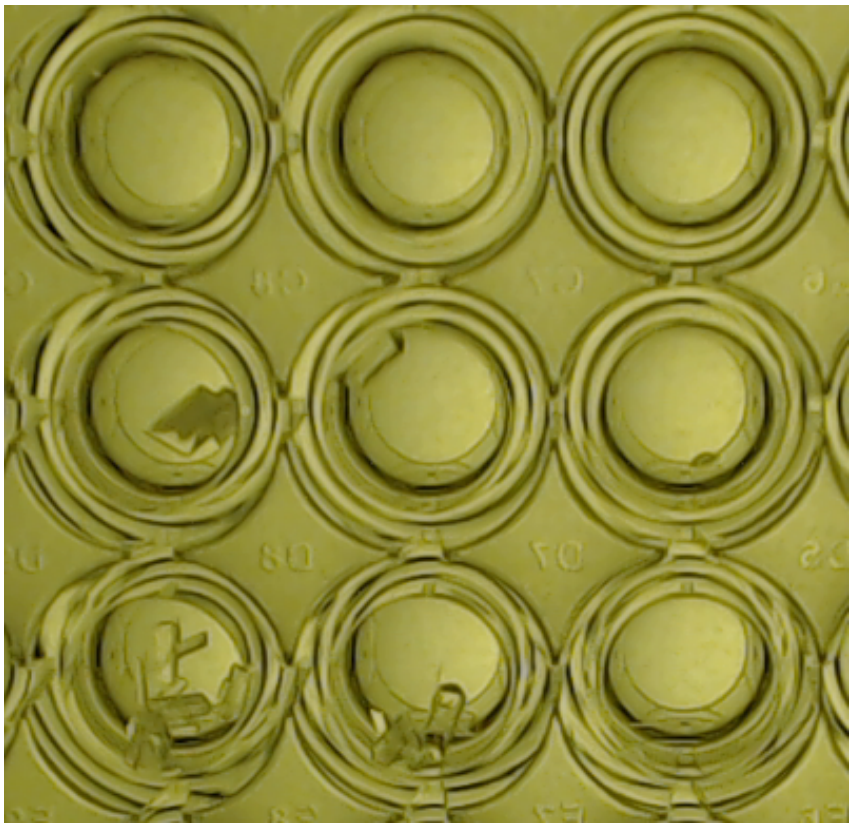
Membrane proteins are hard to work with and expensive, so few quantitative studies, but large literature on water freezing — which is key part of climate models.

Nucleation

1. Supersaturation independent of time
2. Supersaturation varying with time
3. Supersaturation varying with time & space
4. Two-step nucleation, three-step nucleation ...

Nucleation is stochastic

In two apparently identical droplets, a crystal may appear in one after an hour, in the other after 50 hours.



$P(t) =$ fraction of samples where nucleation has not occurred

If nucleation occurs at constant rate:
$P(t) = \exp(-kt)$.

Little et al, J Chem Phys 2017

Nucleation is a stochastic (random) process

Nucleation is known to to be stochastic, i.e., random, nucleation will occur at different times even in identical droplets.

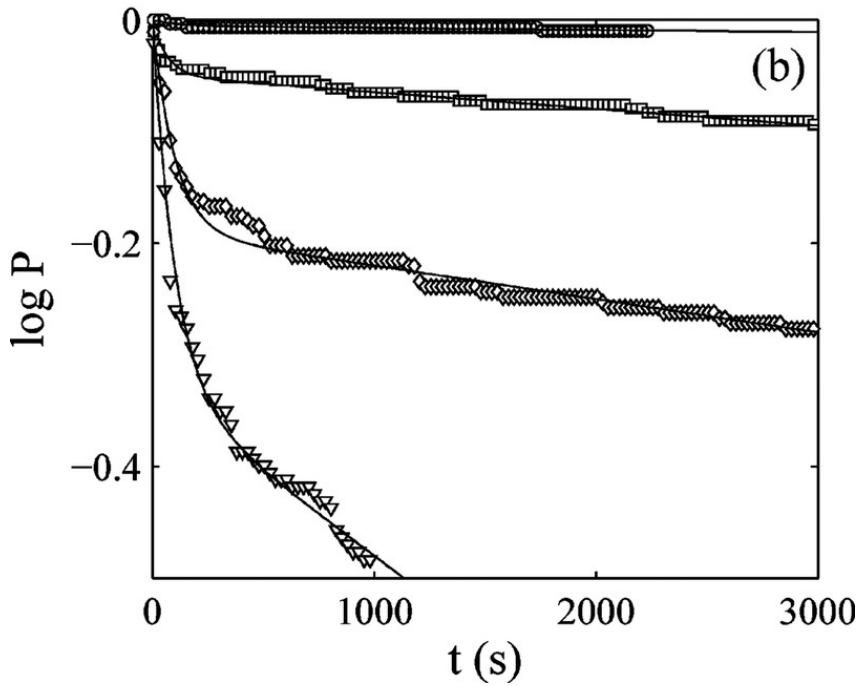
This is like the decay of a radioactive nucleus, whose decay is also random. Radioactive decay occurs at a constant rate, k, so probability nucleus has *not* decayed $P(t)$ is

$$P(t) = \exp(-kt)$$

Nucleation similar except that effective rate k is *not* a constant, so $P(t)$ is not a simple exponential.

Nucleation is stochastic, typically with time-dependent rate

Plot of log of survival function $P(t)$ = fraction of samples where nucleation has not occurred, as function of time t.
Laval *et al.* Langmuir 2009. System = KNO_3 from solution.



Simple model (decay of radioisotope): $P(t) = \exp(-kt)$ clearly wrong. Effective nucleation rate (= slope of this plot) is not a constant.

Nucleation is stochastic, typically with time-dependent rate

The effective nucleation rates $h(t)$ is

$$h(t) = -\frac{\mathrm{d} \ln P(t)}{\mathrm{d}t}$$

i.e., slope of $\ln P$ vs t

$P(t) =$ fraction of samples where nucleation has not occurred

effective rate $h(t)$ typically depends on time, eg nucleation rate of KNO_3 starts high, but slows

$h(t)$ time dependent because of:

1. Crystallising droplets are not the same. Nucleation is known to be almost always heterogeneous — occurs on impurity — and these impurities vary from one droplet to another
2. There are time-dependent processes going, eg slow chemical reactions, protein unfolding, ...

Discussed in review, Sear CrystEngComm 2014

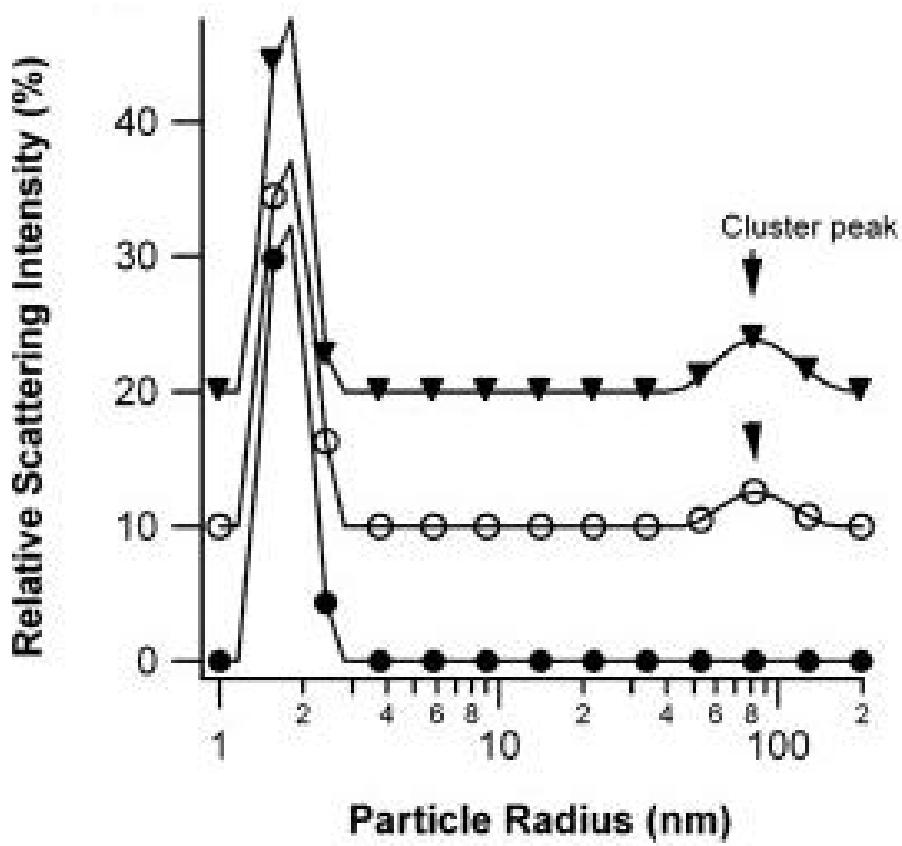
Nucleation occurs with crystal nucleus in contact with a surface or other heterogeneity

Nucleation of crystals is almost always heterogeneous: the crystal nucleus is not just in the bulk solution, it is contact with a surface (eg of dirt), or contains impurities, etc. These impurities have huge effects on nucleation.

For example, after ~ 60 years work on the nucleation of ice in water, we now know that the nucleation of ice is almost always heterogeneous above temperatures $\simeq -35$ C.

Protein aggregation → proteins can generate their own nucleation-active surfaces

Parmar et al, Biophys Chem **129**, 224 (2007):



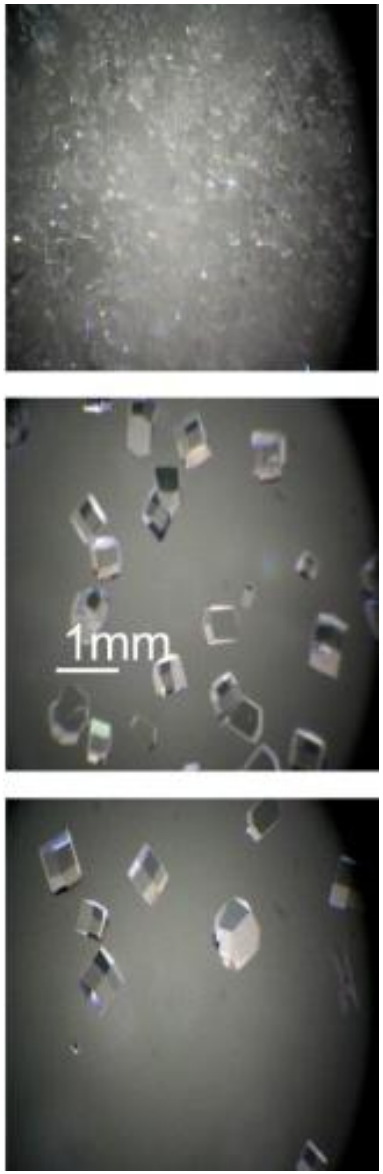
Dynamic light scattering (DLS) measurements of lysozyme solutions show big peak for particles $\simeq 2$ nm across — lysozyme (monomers), plus another peak ~ 80 nm — lysozyme aggregates

Proteins nucleating on protein aggregates & where the protein comes from matters

Parmar et al, Biophys Chem **129**, 224 (2007):
From top:

1. Seikagaku lysozyme filtered with 220 nm filter
2. Seikagaku lysozyme filtered with 20 nm filter — lysozyme aggregates removed
3. Worthington lysozyme filtered with 220 nm filter — source makes a difference!

No protein is 100% pure and this makes a difference to nucleation



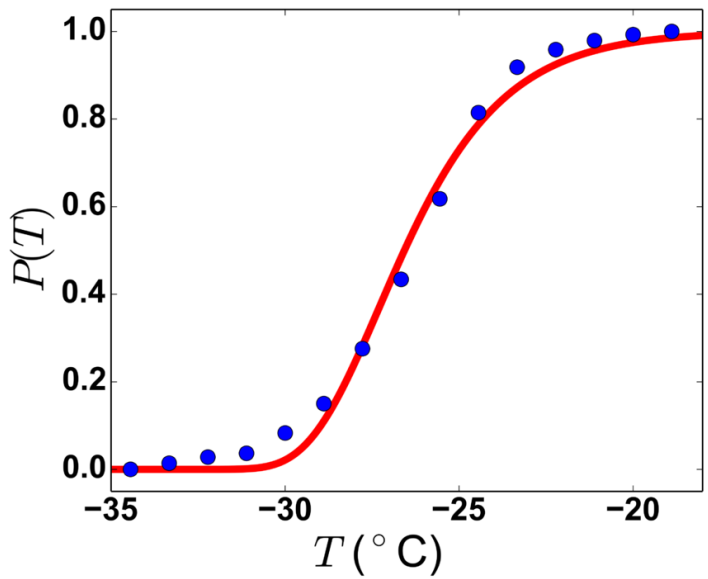
Nucleation when supersaturation is increasing with time

Often nucleation is done not at constant supersaturation (constant temperature, concentrations, etc) but at a supersaturation that increases with time.

Often time-dependent systems are also spatially non-uniform (supersaturation different at different points in the droplet. But here I consider spatially uniform systems.

Nucleation when supersaturation is increasing with time

Fraction of small (34.5 μm) water droplets that have crystallised, as a function of temperature/time — they are being cooled at a constant rate. Data from Dorsch & Hacker, NACA report 1950. Note for droplets this small, growth of ice is fast, $<$ 1 s.



red curve Gumbel function

$$P(T) = \exp\left[-\exp\left(-(T-\mu)/\beta\right)\right]$$

μ and β fit parameters

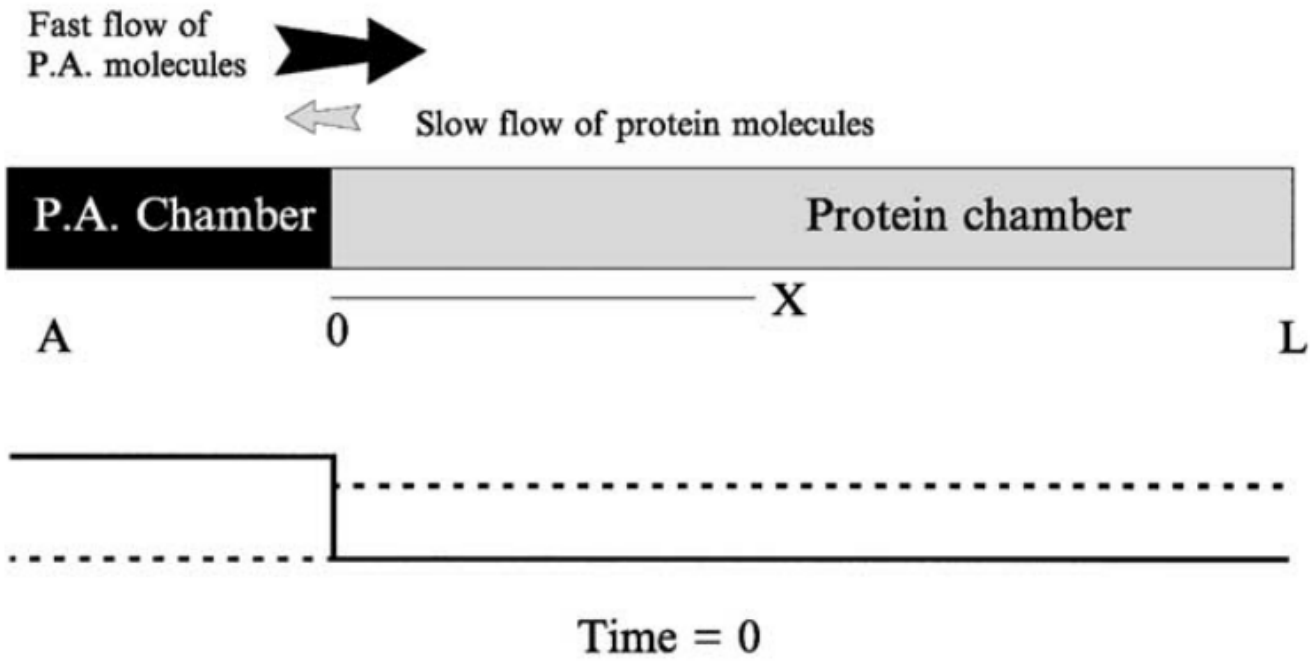
Nucleation when supersaturation is increasing with time

In the case of water droplets studied under conditions of continuously increasing supersaturation, nucleation occurs over a broad range of supersaturations — equivalently a range of times.

This is presumably due to different droplets being different, i.e., one droplet has an impurity that is unusually good at promoting nucleation, and so freezes at a high temperature, while another droplet has only impurities that are bad at helping nucleation, and this droplet freezes at a lower temperature.

Spatially non-uniform systems

Free Interface Diffusion (FID) (also called counterdiffusion) technique, at $t = 0$, protein solution to the right brought into contact with solution of precipitating agent (P.A.) to the left. N.B. Almost all spatially varying systems time dependent, and vice versa.

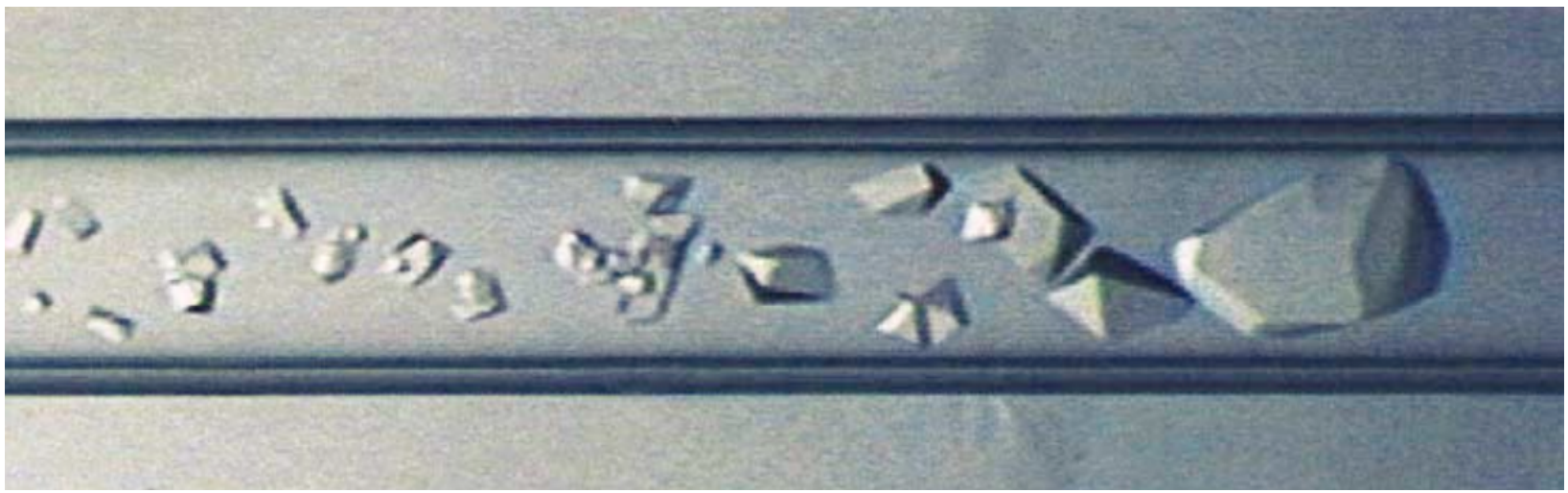


Garcia-Ruiz, Methods Enzymology **368** 130 (2003)

This is a specific technique, but **any** technique will involve diffusive mixing of protein and precipitating molecules at some point in the experiment

Spatially non-uniform systems

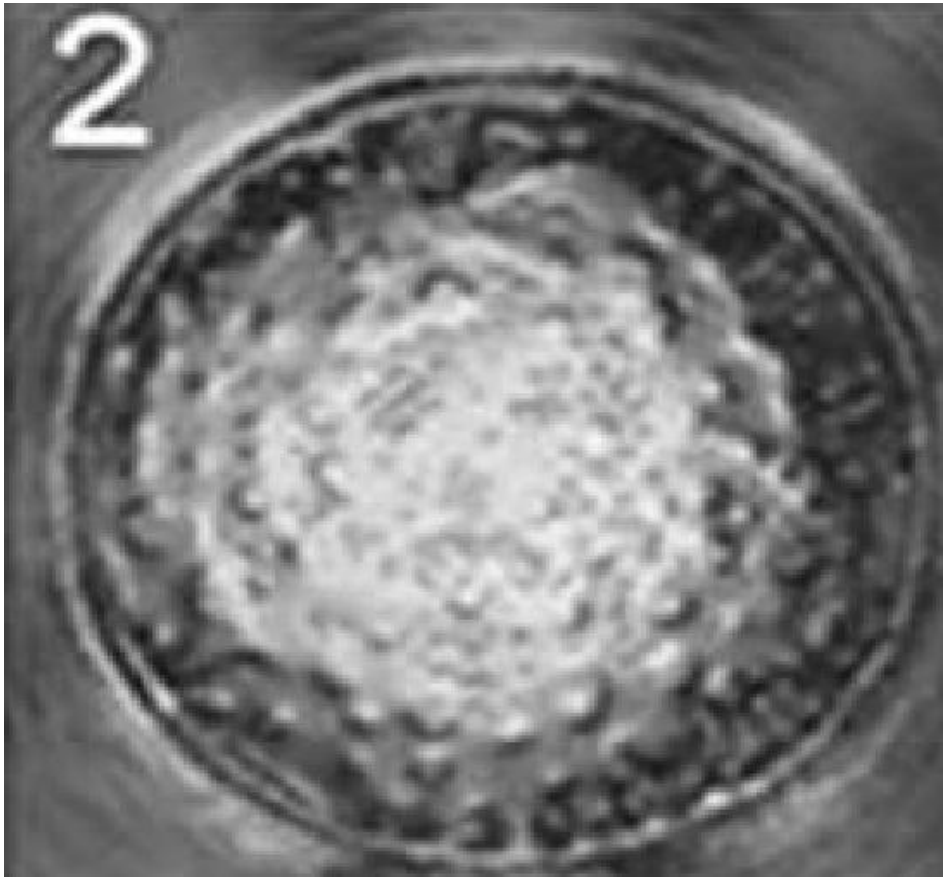
dehydroxiquinase crystals — note crystals are larger to the right, when concentrations of protein and precipitating agent vary in space & time this will give crystals of varying size (and presumably ordering etc)



Garcia-Ruiz, Methods Enzymology **368** 130 (2003)

Show animation for FID

Phase separation, two-step nucleation, three-step nucleation ...



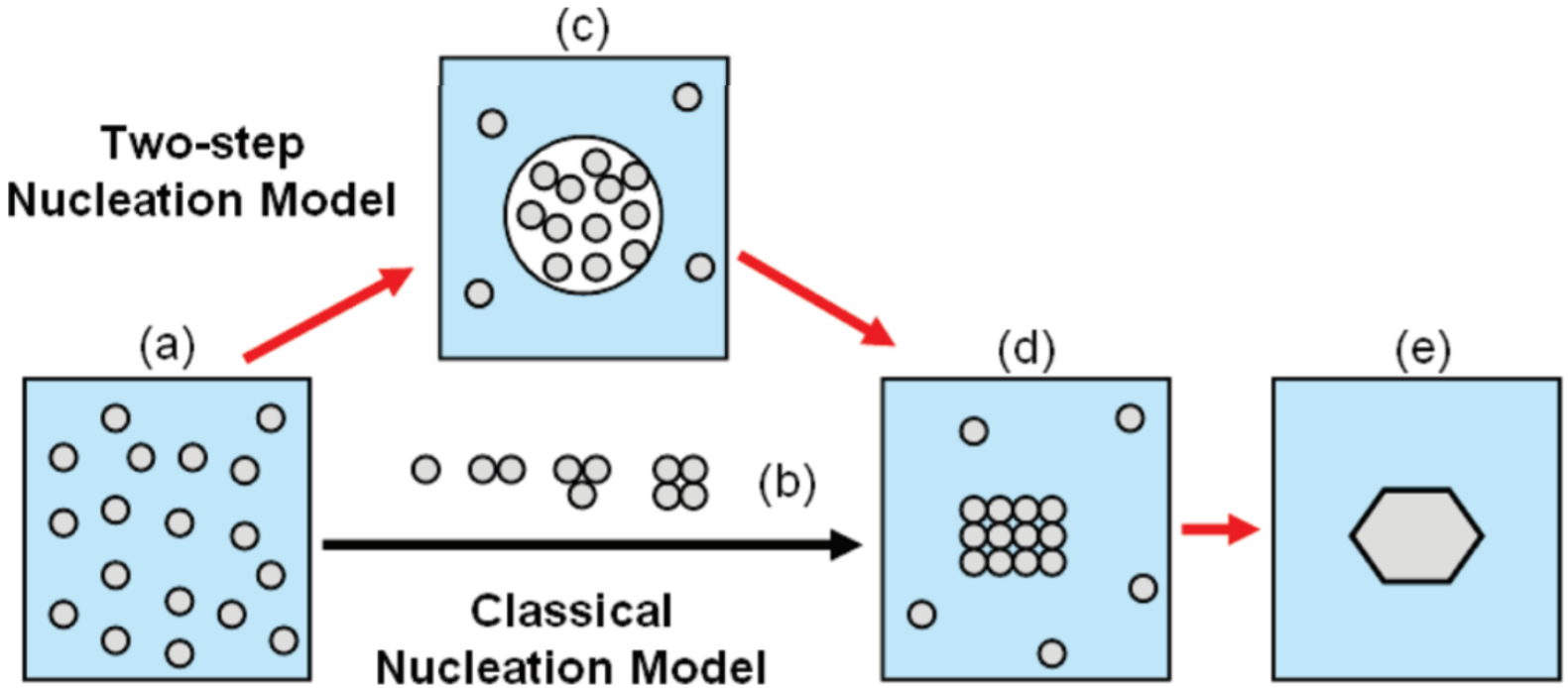
Luft et al., Cryst Growth Design **11**, 651 (2011).

Phase separation/oiling out, i.e., separation into two phases, on higher in protein concentration than the other.

Jen will discuss this, but this can suggest that proteins effectively attract each other — something we expect is required to crystallise from dilute solution

Crystals may, or may not, form in a second stage, from a phase separated

Phase separation, two-step nucleation, three-step nucleation ...



Erdemir et al. Acc Cem Res **42**, 621 (2009)

Note that the system undergoes one transition, to a more concentrated solution, and then and then crystals form in a second step. In some cases crystallisation may involve several steps — defect nucleation, polmymorphs, ...

Phase separation & two-step nucleation, three-step nucleation ...

1. Observing a phase transition in the solution *may* indicate crystallisation conditions are nearby — useful information

2. There are statements in the literature that: 1) a liquid/liquid phase transition helps crystallisation, 2) a liquid/liquid phase transition hinders crystallisation. Both are possible.

3. In multistep crystallisation, there may be one step much slower than the others, if so, the slow rate sets the overall nucleation rate.

Crystal growth

Always need crystal growth to make large crystals, note that can separate out nucleation and growth by using seeds: small crystals made in one experiment that are added to a supersaturated solution in an second experiment, to remove the need for nucleation in the second experiment

Crystal growth: What I will not cover

I am a theoretical physicist, I know less about actually growing crystals in the lab than do Monika, Adrian, Bernadette, Martin, ...

I will however survey some of the data we have about how crystals grow, and some of the models people use to try and understand crystallisation. This will essentially all be in systems other than membrane proteins.

Crystal growth can be horribly complicated :(

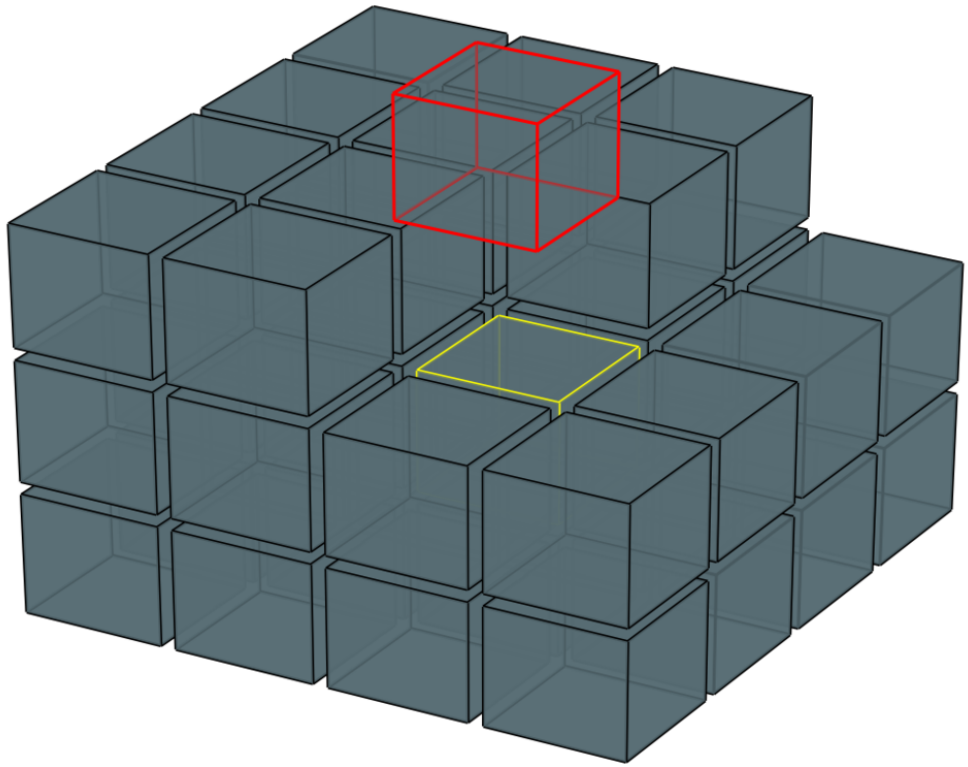
Crystal growth differs from growth of a liquid droplet in that during growth the molecules or ions must fall into the correct lattice positions in order for a well-ordered crystal to grow.
Problems:

1. What determines the growth rate? Is it a process actually at the growing crystal face? Attachment of an extra protein molecule can be an activated (and so) slow process. Is growth limited by transport of protein molecules to to the growing crystal? Or is it a combination of both? If so it will be a complex multi-scale problem.
2. Defects: At least in some cases, it is known that crystals only grow at an appreciable because of a defect, eg spiral dislocation. If so details of this defect critical to growth rates ...
3. Growth can be very sensitive to impurities sticking to growing surface or being incorporated into growing crystal.

Crystal growth: What we know of microscopic mechanisms

Molecules must add in the right place in the lattice, and some places may be much better than others

Expectation is that next molecule will add at corner (kink) in lattice as there it forms most bonds with molecules of the crystal



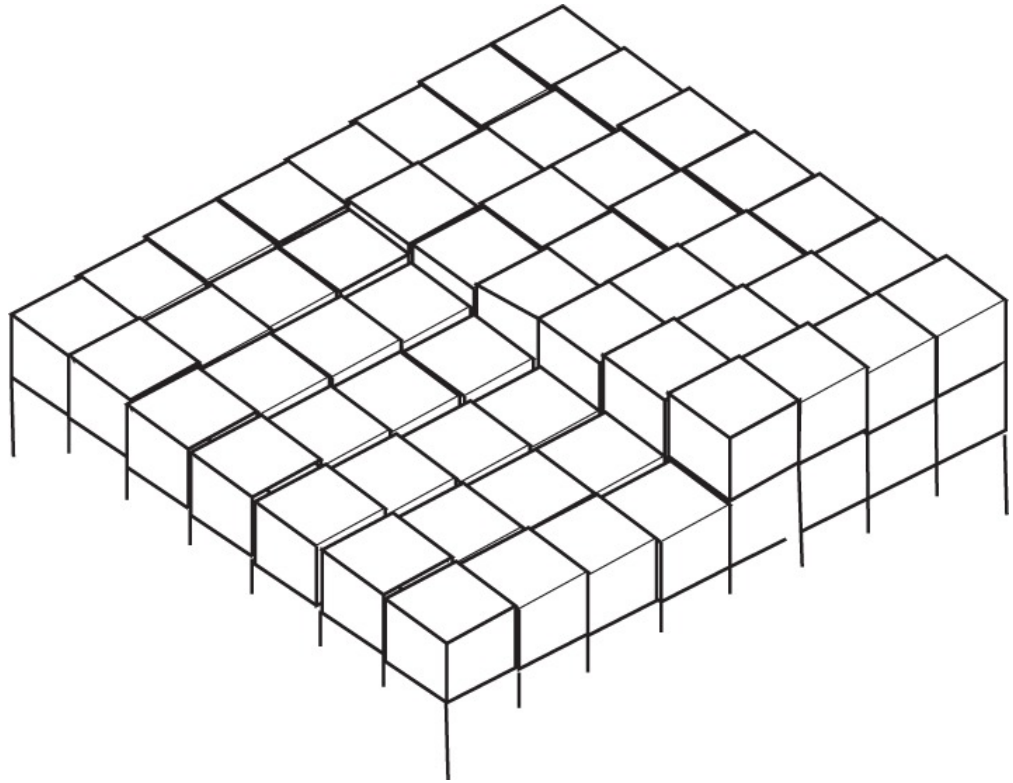
This shows next molecule (red) adding (on top of yellow molecule) in a kink on an incomplete layer of a growing (cubic) crystal.

Molecule adding in corner forms 3 bonds — would only form 1 if it contacts flat surface

Expect molecules preferentially attach in corners, so incomplete lattice planes find it much easier to grow than complete layers — that don't have a corner immediately above another layer.

Defects can dominate growth

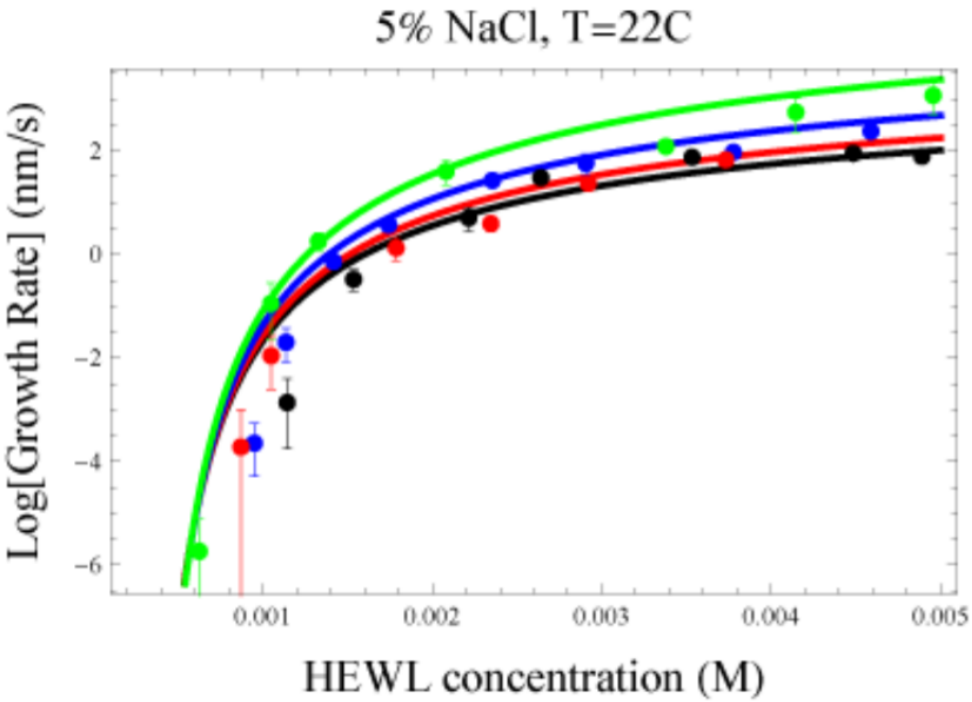
As corners can be so important, defects that continuous growth via corners can accelerate growth...



This shows a defect in crystal lattice called screw dislocation. Crystal grows by molecules adding at advancing edge of dislocation causing spiral growth upwards.

Show calcite screw dislocation growth movie

Few measurements of growth rates of soluble protein crystals



HEWL = Lysozyme

Data from: (18). Gorti et al., Cryst Growth Des **5**, 535 (2005). Curves are fits of unconvincing model of Schmitt et al. JACS **134**, 3934 (2012)

Rate typically strongly dependent on supersaturation, so varies rapidly with concentration. Note rates often slow, nm/s or slower. This is of order one extra layer of protein molecules per second

Compare this to μs timescale for lysozyme to diffuse its own diameter in solution

Crystal growth results on glycine crystallising from solution

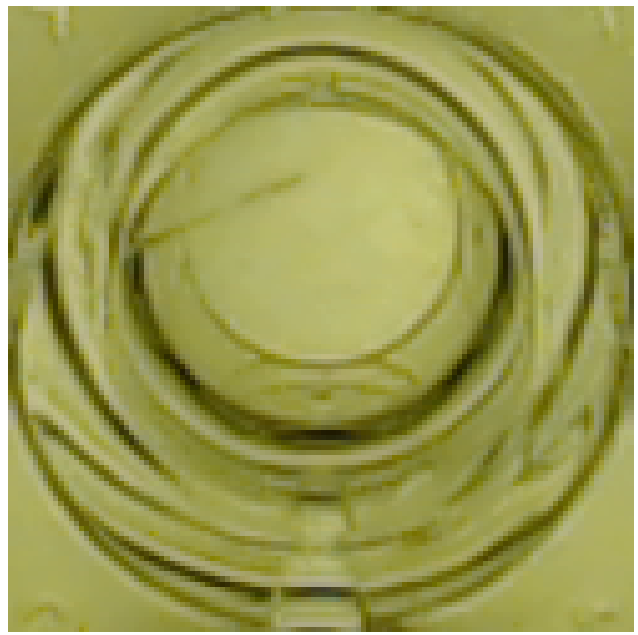
Laurie Little, Joe Keddie & I, J Chem Phys 2017

We cannot identify crystal polymorph from crystal habit (shape)

Both polymorphs have two crystal habits: conventional and needle.

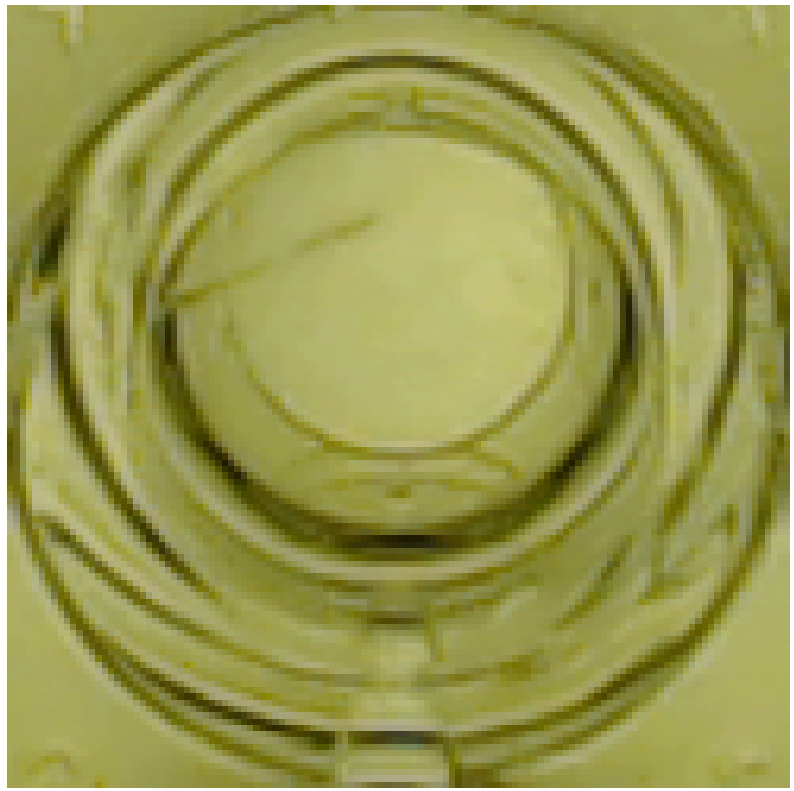
Crystal of α polymorph in conventional habit

Wells are 6.8 mm across



Crystal of α polymorph in needle-like habit. Same Raman spectrum, different shape

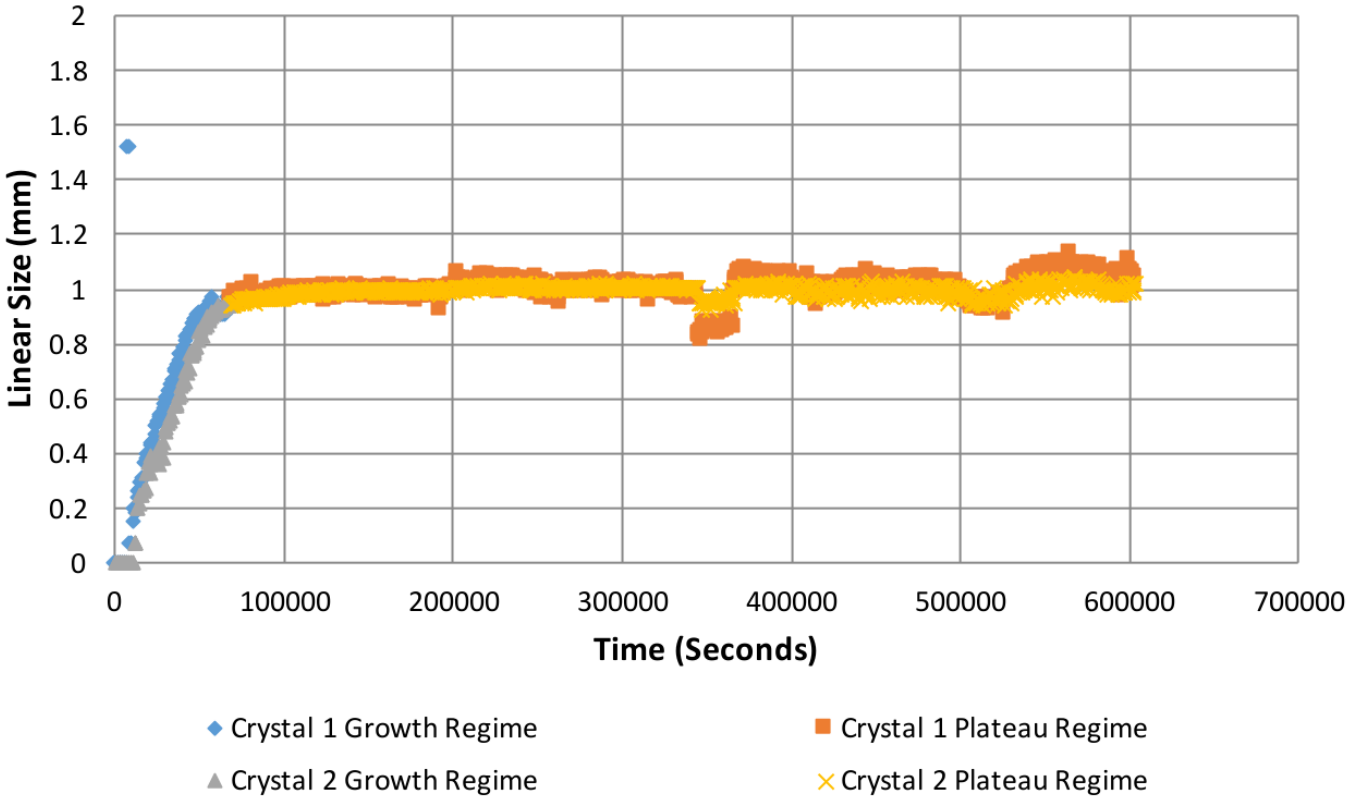
Needle-like crystals form via very rapid growth along long axis



Long axis of needle-like crystals grows much more rapidly: spans well in few tens of mins, conventional habit takes of order an hour to grow.

Non-Needle-like crystals form via slower growth

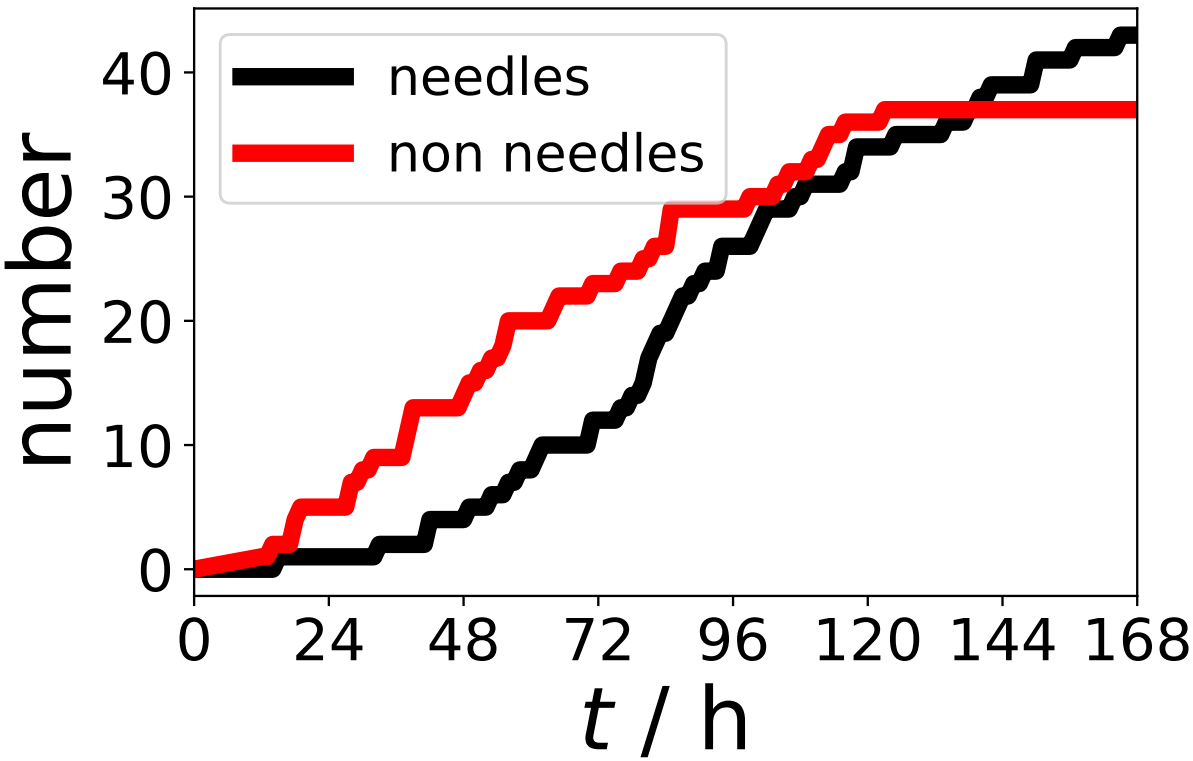
γ polymorph



growth rate of γ non-needle $\sim 1\ \mu\mathrm{m}$/minute — other work, eg Dowling *et al.* 2010, found $\sim 10\ \mu\mathrm{m}$/minute, but that was without NaCl. We find NaCl slows growth.

linear size = square root of area in image. Analysis done by Alex Traher (Surrey)

Early nucleating crystals are different from late nucleating crystals



250mg/ml salt

same trend at all salt concs.

Early nucleating α crystals are rarely needles, late nucleating crystals are mostly needles.

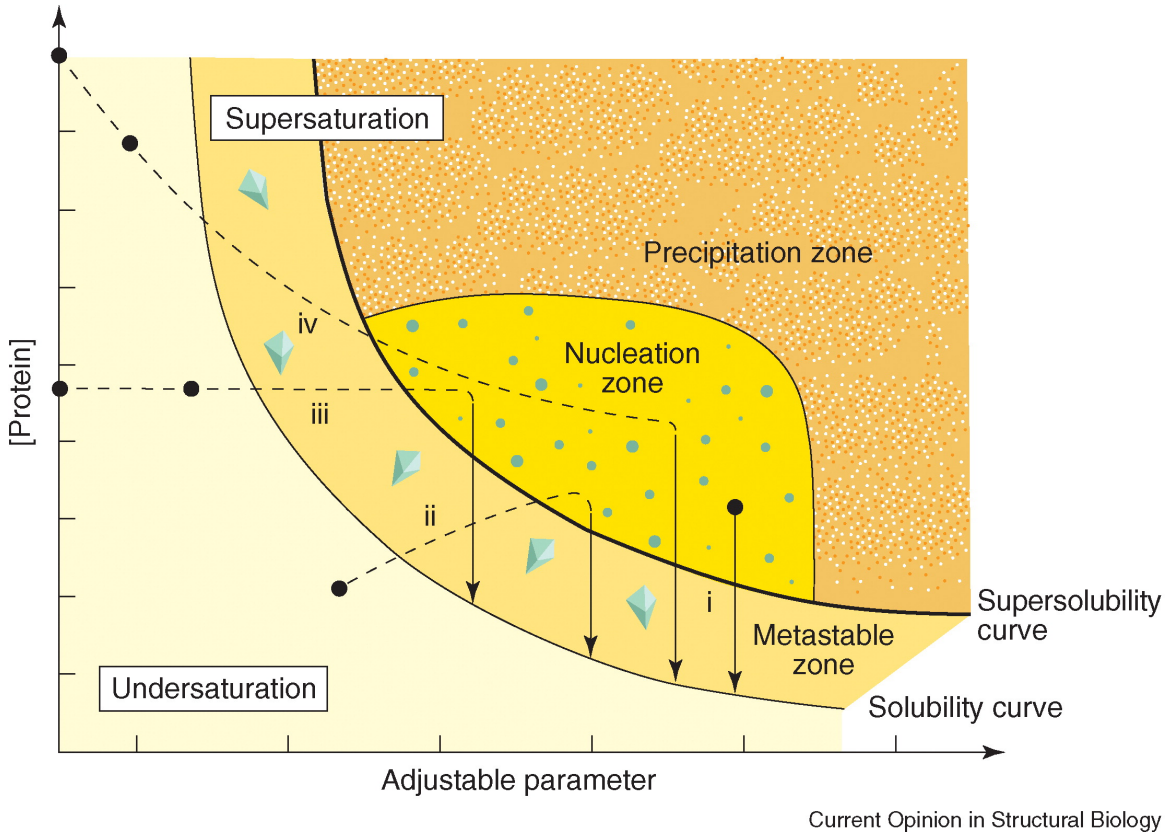
We are surprised by this difference between early and late nucleating crystals.

Summary for mechanisms of growth

1. What determines how fast crystals grow, and how well ordered the resulting crystals are, are very poorly understood
2. Crystal growth presumably involves defects, phenomena over multiple length scales, ... but even in simple systems we do not have either the experimental methods to obtain detailed molecule-scale info on the process, or the computer simulation tools to simulate it

Soluble protein crystallisation a la Naomi Chayen

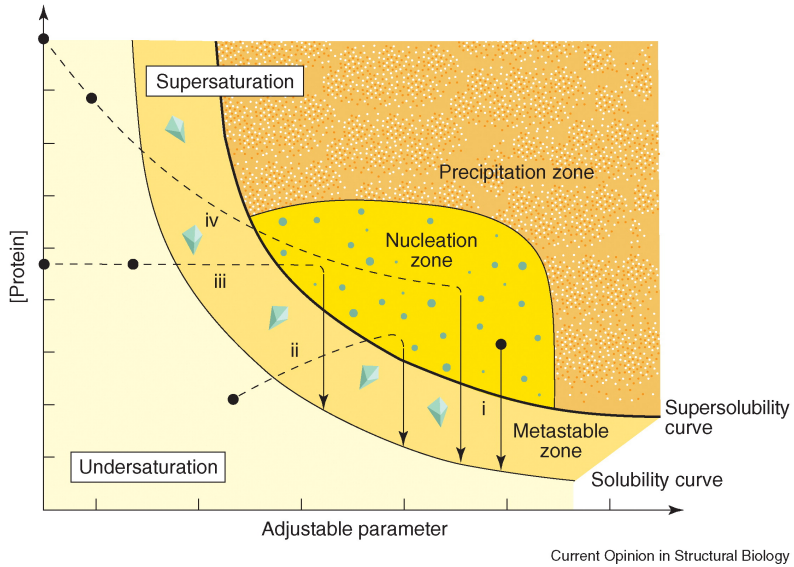
Optimal conditions for crystal nucleation and for crystal growth may be different?



Chayen, Curr Opin Struct Biol **14**, 577 (2004)

Crystallisation of soluble proteins often discussed in terms of "phase diagram" like this. Note many possible x axes as many possible precipitants

Optimal conditions for crystal nucleation and for crystal growth may be different?



Chayen, Curr Opin Struct Biol **14**, 577 (2004)

Empirical evidence is that on increasing precipitant concentration sequence is: soluble $\rightarrow$ growth but no nucleation $\rightarrow$ both nucleation & growth $\rightarrow$ formation of precipitates

This slide & the next is to remind me to link to Jen's part of the course ...

Engineering the effective interaction between proteins

Idea is that soluble protein in just buffer will not crystallise (is undersaturated), so add salts, polymers etc, that in effect modify the effective protein-protein interaction to make this interaction attractive, so proteins stick together to form regular crystal lattice.

This idea implies that the "correct" precipitants will make the protein-protein interaction such that a crystal (as opposed to precipitate) is favoured, favoured in equilibrium sense of low free energy and in sense of favourable crystallisation kinetics.

See Jen's part of course for protein phase diagrams