Check for updates

Mesoscale clusters of organic solutes in solution and their role in crystal nucleation

Michael Svärd [iD]

It is becoming evident that primary nucleation of crystals of organic molecules from solution is often anything but 'classical' in its complexity. It is also becoming increasingly clear that mesoscopic clusters of molecules are at least partly involved in the phenomenon. On-going advances are rapidly enabling these clusters to be investigated with ever more accuracy. However, despite numerous recent studies, fundamental knowledge of the properties, thermodynamics and kinetics of these clusters, and of their role in nucleation, is still limited.

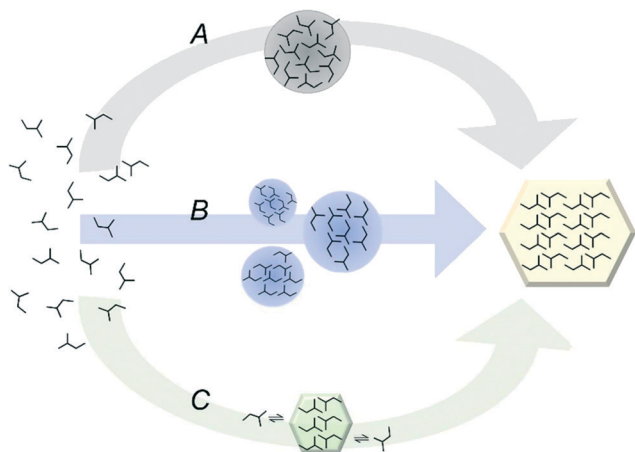## Crystal nucleation and mesoscale clusters

Nucleation is the birth process of one phase in another, driven by thermodynamics. Zooming in, *primary crystal nucleation* specifically denotes the formation of particles of a solid, long-range ordered phase from a less ordered phase, which could be *e.g.* a melt or a solution. It is a key step in many types of crystallization processes, natural as well as industrial. It has long been the focus of experimental as well as theoretical studies, perhaps most intensely for organic molecules typical of active pharmaceutical ingredients. The reason is most likely found in the particular role of crystallization processes in the pharmaceutical industry, where it is often the key production step rather than just an intermediate purification step. Nucleation at least partly governs properties of the resulting crystals such as size, shape, crystal structure, and physicochemical and mechanical properties derived therefrom. These include solubility and dissolution rate which affect the bioavailability, physical stability which determines storage conditions and shelf life, and powder properties affecting the downstream formulation and manufacturing process.[1] Pharmaceutical manufacturing is subject to regulations that mandate strict control of the crystallization process,[2] particularly concerning the robust production of the desired polymorphic form or solvate.[3] However, crystallization of organic molecular compounds is of importance in many other branches of fine- and specialty chemical industry, *e.g.* in the food industry,[4] and for production of pigments and dyes.[5]

Despite a long history of extensive research, the present level of understanding of the thermodynamics, kinetics and mechanisms responsible for primary nucleation is limited. Empirically, it is a stochastic process, exhibiting a high sensitivity to apparently minor changes in process conditions.[6] Apart from temperature and supersaturation, and the presence of surfaces which can template heterogeneous nucleation, it can also be affected by the volume and geometry of the crystallizing vessel, agitation conditions, shear flow,[7] mechanical shock, and the application of ultrasound, electromagnetic radiation and electrical or magnetic fields. The tendency for a solute to nucleate can also depend on the time and temperature of the previous pre-treatment history of the solution[8] as well as on whether the solution has been filtered. Primary nucleation is characterized by notoriously low predictability, and typically, each crystallization process has to be designed empirically from scratch.

Crystal nucleation is inherently difficult to observe, because it occurs as a singular event occupying a very short time and a small volume, in a system defined by time and volume scales that are, in relative terms, enormous. Until recently these limitations have been prohibitive, and in the absence of actual *in situ* techniques to observe nucleation of crystals in solution, nucleation models have been based chiefly on indirect observations and inferences from later, observable stages of the crystallization process, *e.g.* by measurement of the induction time, and by comparison of crystal properties such as structure and size. For a long time, the dominant descriptive theory of the mechanisms of solute aggregation into crystal nuclei has been the *classical theory of nucleation* (CNT),[9,10] originally designed for condensation of vapour droplets, schematically indicated by path C in Fig. 1. The CNT assumes that a nucleus is formed by a gradual process of attachment and detachment of monomers (*e.g.*

Department of Chemical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. E-mail: micsva@kth.se

Fig. 1 Schematic representation of proposed pathways for primary nucleation from solution: A – nucleation *via* nano- or microdroplets of an amorphous or liquid-like phase; B – nucleation *via* mesoscale pre-nucleation clusters; C – classical nucleation *via* ordered sub-critical clusters.

molecules, ions) to a cluster having a crystalline structure but a thermodynamically unstable size.

Recent advances in analytical techniques based on light scattering, and the development of new *in situ* implementations, are gradually – but at a rapid pace – enabling aggregation and ordering phenomena of molecules in solution to be studied directly.[11,12] Meanwhile, with the gradual increase in computational power and the development of new methods,[13] accurate simulation of nucleation from solutions is fast becoming possible. The growing experimental evidence for the existence of mesoscale aggregates in undersaturated as well as supersaturated solutions,[14–22] with sizes and structures exhibiting complex dependencies on solution conditions – and apparently governed by slow growth- and transformation kinetics – is driving the development of a family of new, *non-classical* nucleation theories, indicated by paths A and B in Fig. 1. One thing in common for these theories is the emphasis on the role of these mesoscale clusters in early aggregation and ordering.[23] In the following, the history of experimental investigations of mesoscale clusters from early indirect approaches to the current state of the art, and the connection to crystal nucleation, is summarized and discussed.

Indirect experimental indications of mesoscale clusters

Gravity and diffusion

There have been various attempts to study clusters indirectly through changes in properties of the solution. As early as 1969, Mullin and Leci performed a famous experiment of aqueous citric acid solutions in columns.[24] At supersaturated conditions, concentration gradients were developed in the column over the course of hours or days, but no concentration gradients were observed at saturated or

undersaturated conditions. The explanation proposed was that solute-rich clusters, of higher density than the bulk solution, would gradually develop and settle in the supersaturated solutions. Larson and Garside expanded this line of investigation to more systems, and using densitometry inferred the presence of aggregated entities whose size would typically increase with increasing supersaturation.[25] However, analysis of the data suggests that the dominating size of entities giving rise to concentration gradients for many systems ranges from 2–3 molecules, corresponding to dimers or other oligomers, up to ~100 molecules.[26] Larger, mesoscale aggregates were either not present, or not present in significant concentrations, until close to the point of nucleation.

Another property of a solution that can be affected by the presence of clusters is its diffusivity. Cussler[27] predicted that the diffusion of clusters will be largely different from that of individual molecules, as traditionally described by Fick's law. Largely on the basis of this analysis, Gouy interferometry has been used to investigate the time-dependence of diffusion coefficients in supersaturated solutions. It was found for a large number of compounds, including glycine and urea, that the diffusion coefficient decreases rapidly towards zero with increasing concentration, and this was attributed to the formation and growth of clusters.[28,29] Moreover, the diffusivity was shown to decrease with increasing solution age. For glycine in aqueous solutions at different levels of supersaturation, the timescale for cluster sizes and diffusion coefficients to reach stable values ranges between 10–100 hours.[28]

Shear flow

Virtually all crystallizing systems are in motion, whether applied mechanically or through convective processes. Although the underlying mechanisms are not well understood, the influence of shear flow on nucleation has been the object of various studies. Early experiments on inorganic molecules from aqueous solution showed that with increasing agitation rate nucleation can be promoted, but at higher agitation rates it can also be obstructed.[30] Comparatively few studies have focused on smaller organic molecules. Using Taylor–Couette flow cells operated under moderate rotation rates, allowing experiments to be conducted under well-controlled shear flows calculatable using the Navier–Stokes equation, nucleation of *e.g.* glycine,[31] butyl paraben,[7] and various pharmaceutical compounds[32,33] has been investigated. In all these studies, the nucleation induction time is found to be inversely related to the fluid shear rate.

Several possible mechanisms have been proposed as explanations for the effect of shear flow on nucleation. Schmidt and Schmidt modelled clusters in turbulent eddies, and suggested that clusters will concentrate at the periphery of eddies under the influence of centrifugal forces, promoting aggregation.[34] Mullin and Raven proposed that

the dependence of nucleation on agitation could be due to a combined effect of diffusion and attrition on the growth of clusters in solution; cluster growth is promoted by agitation-enhanced diffusion at moderate agitation rates, while at higher agitation rates attrition of the clusters leads to obstructed nucleation.[30]

There have also been reports of observations from experiments under less controlled shear conditions with a possible bearing on the relationship between shear and nucleation. The means of stirring is well known to affect the outcome of crystallization. Although it appears not to have been the focus of systematic studies, the use of magnetic stirrer bars is generally considered to promote nucleation.[7] There are also reports of cases where the means of agitation can affect the nucleating polymorph.[32,35] One intriguing example is glycine, which can crystallize in three polymorphs. Whereas the stable γ-polymorph can be crystallized at high supersaturation under quiescent conditions (non-stirred solution), addition of stirring leads to predominant nucleation of the metastable α polymorph.[36] However, it is less certain that such effects are directly related to shear, as there could also be differences in attrition mechanisms and overall different hydrodynamics involved.

History of solution, filtration and nucleation

There are some peculiarities of nucleation that still have not been satisfactorily explained, most notably the so-called *history of solution* effects.[8,37] It has long been known that the propensity for nucleation can depend on the thermal history of the system; decreasing with increasing time and temperature of superheating during pre-treatment. The phenomenon, initially described for crystallization from melts,[38] was later systematically studied for crystallization from solution.[39–43] Nakai[8] proposed a cluster distribution model to explain the kinetics behind the thermal history effects.

More recent work on the nucleation of typical APIs (active pharmaceutical ingredients) and related molecules from organic solutions has shown a systematic dependence of the nucleation propensity on the time and temperature of solution pre-treatment.[37,44] The phenomenon appears to also have a structural dimension, and the time scale of the kinetic processes behind the observed effects is of the order of hours or even days. It has been proposed that, after dissolution, solute-rich aggregates could be present which, to a certain degree and for a certain time interval, could retain some structure remnants from the dissolved phase. With time, clusters in solution will rearrange in size and structure to approach an asymptotic steady-state distribution. A practical conclusion of this hypothesis is that, in order to study nucleation, the history of the solution first has to be standardized, by keeping it at a set temperature for a set length of time.

Besides thermal pre-treatment, several studies have shown that the propensity for nucleation can be adversely affected by filtration of the solution.[41,45–47] Javid et al.[45] investigated the effect of nanofiltration through 0.2 μm filters on laser-induced nucleation of glycine in aqueous solutions at various supersaturations. Filtration resulted in a reduced nucleation tendency overall, and was also shown to affect the polymorphic outcome. Hirschler et al.[46] investigated crystallization of lysozyme from solutions under various combinations of filtration and ageing, and found that filtration reduced the tendency to crystallize. However, in some cases the effects of filtration were absent if followed by a subsequent period of ageing, suggesting a time-dependent aggregation process with slow kinetics is affecting the crystallization.

Direct experimental investigation of mesoscale clusters

There are a number of techniques where the state of the art allows mesoscale clusters to be detected and studied directly in solution. These techniques come with their individual pros and cons, imposing different possibilities as well as limitation on what information they can provide, with regard to e.g. cluster size and size distribution, shape, structure and concentration. Some techniques provide various types of averaged information while others are able to resolve information for each individual cluster or particle. For this reason, many recent studies feature combinations of more than one technique. Such an approach is promising as the use of complementary techniques can provide a clearer, more complete picture, but also because e.g. size information from two different techniques can be compared, improving the confidence in the results.

Most of the direct analytical techniques for detecting and studying clusters in solution can be groups into two main groups: techniques that rely on the scattering of a focused beam of light, and techniques based on electron irradiation. To the former belong static and dynamic light scattering (SLS, DLS), nanoparticle tracking (NTA; also known as Brownian microscopy), and small angle X-ray and neutron scattering (SAXS, SANS). The latter include variants of transmission electron microscopy (TEM). Apart from these main techniques, some studies have also reported detection and analysis of organic mesoscale clusters using electrospray ionization techniques (ESI),[22,48,49] analytical ultracentrifugation (AUC)[22,48] and atomic force microscopy (AFM).[22,49,50]

DLS has possibly been the most frequently used technique for detection of clusters of organic solutes. It is relatively inexpensive and readily available in many labs. Moreover, it is rapid, and has a wide size range. Among its drawbacks is the fact that it is an averaging technique, poorly suited for investigating details of size distributions for polydisperse samples.[51] SAXS, an important part of the toolbox for studying proteins and other macromolecules, has occasionally been tried as a complementary technique to detect, measure the size of, and probe the structure of

mesoscale clusters. It can be combined with wide-angle X-ray diffraction, to evaluate the occurrence of regions of long-range order within the scattering entities.[52] Although the principle of SAXS is straightforward, the technique suffers from some serious practical problems. Because of similar electron densities of the clusters and the surrounding solution, the scattering contrast is weak. The resulting part of the signal emanating from scattering from clusters has very low signal-to-background and signal-to-noise ratios. This necessitates a highly brilliant light source, such as a synchrotron, and finely tuned, high-quality optics.

In comparison, NTA is emerging as a promising alternative. Like DLS, it is based on evaluating scattering of laser light in terms of diffusion coefficients which are translated into cluster sizes. The principal difference is that it allows individual scatterers to be tracked, Fig. 2, enabling size distributions to be measured and number concentrations to be estimated directly. As the technique is still immature, measurements are comparatively time-consuming, and not many models are available, which has restricted its use with regard to e.g. solvents and measurement cells.

Direct *in situ* application of imaging techniques based on TEM have long been impracticable. By cooling the sample to cryogenic temperatures (Cryo-TEM), a still image of the solution structure can be obtained. This technique has been used to study non-classical nucleation of organic molecules[53] and to visualize pre-nucleation clusters of $CaCO_3$.[54,55] Liquid cell TEM is a promising, recently developed technique that combines scanning TEM with a microfluidic cell, which offers high resolution in combination with the possibility to directly capture dynamic events in solution.[56] The equipment is not generally accessible and its operation requires some training. However, it has recently been successfully applied to observe non-classical nucleation and growth of flufenamic acid in an organic solvent.[57] The authors hypothesize that ethanol molecules, undergoing radiolysis from the incident electron beam, promote nucleation.

The example of glycine

The majority of direct experimental investigation of mesoscale clusters of organic solutes has targeted rather

small molecules, with few studies focusing on larger and more flexible molecules, such as typical pharmaceuticals. A selection is shown in Fig. 3. In particular, glycine and other simple amino acids in aqueous solution have been the object of a large number of studies using various analytical techniques to probe the degree and type of self-association, at supersaturation prior to nucleation as well as in undersaturated solutions. As previously mentioned, early work using diffusivity and gravity-based analysis found evidence of clustering of entities larger than a dimer in supersaturated solutions.[28] By means of electrospray ionisation mass spectrometry (ESI-MS), clusters of up to 30 molecules could be detected from solutions under a large range of conditions, for glycine as well as other amino acids.[58,59] The analysis indicated that these aggregates could have a preferred size, with a stability exceeding that of both smaller and larger aggregates. Direct *in situ* measurements of glycine clustering in aqueous solutions of various amino acids at undersaturated conditions have been accomplished using AUC,[48] showing that the cluster concentration is typically very low; involving less than 0.2% of the glycine molecules. Studies using small angle neutron and X-ray scattering[18,60] have found evidence of clustering at supersaturation, and analysis of the power law behaviour of X-ray scattering indicates that glycine dimers organise into larger amorphous clusters on nucleation.[18]

Perhaps the most interesting contributions to the clustering of amino acids are the comprehensive studies incorporating more than one technique to study the same system and conditions.[20,49,60–62] The combination of more or less complementary techniques allows a more robust picture of the system to be obtained. Using a combination of NTA, AFM and ESI-MS, Hagmeyer et al.[49] report that spherical clusters of amino acids with a diameters between 100–200 nm spontaneously arise upon dissolution in water, apparently to exist in a dynamic equilibrium with dissolved molecules. Researchers from Jan Sefcik's lab at Strathclyde (Jawor-Baczynska et al.[20,62] and Zimbitas et al.[61]) in a series of studies investigated clustering in aqueous solutions of glycine and DL-alanine using a combination of techniques, DLS, NTA, SAXS, NMR and Cryo-TEM. They detected two distinct populations of aggregates, present in supersaturated
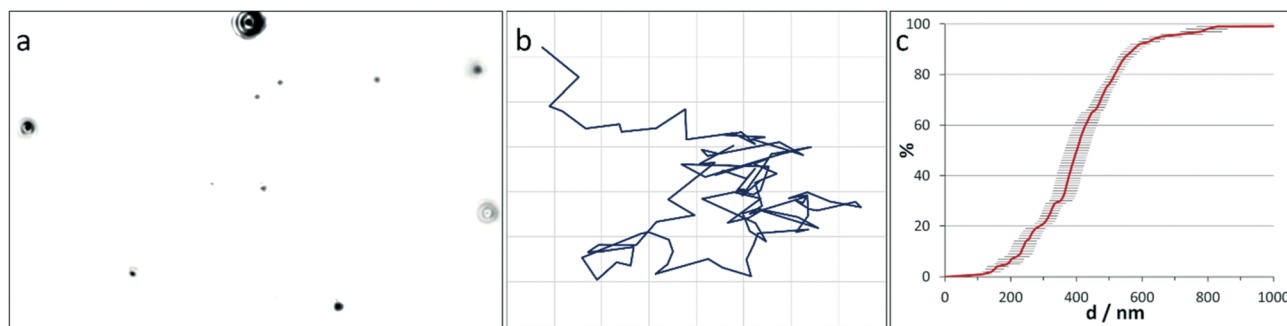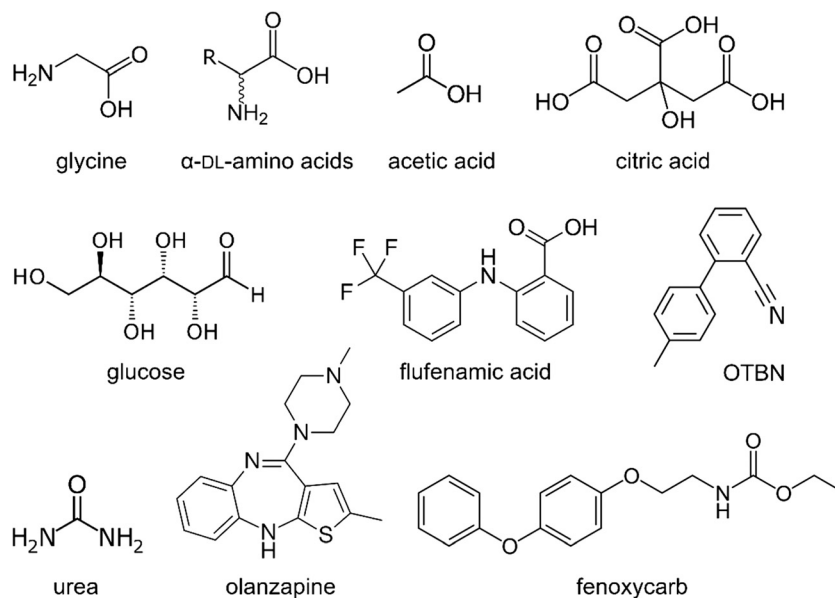

Fig. 2 a) An example of a colour-inverted still frame obtained by NTA, showing scattering species as dark points; b) an example of a trajectory of a scatterer tracked by NTA; c) an example of a cumulative cluster size distribution obtained from averaged NTA runs.

Fig. 3 A selection of organic molecules investigated in clustering studies; glycine[18,20,28,48,60–62] and other DL-amino acids,[22,48,49,58,59,62] acetic and citric acid,[63,64] glucose,[63–65] flufenamic acid,[57] OTBN,[66] urea,[63,64,67] olanzapine[50] and fenoxycarb.[68]

as well as undersaturated solutions. Molecular clusters of around 1 nm in diameter, consistent with hydrated dimers, exist in combination with significant concentrations of liquid-like mesoscale clusters, "nanodroplets". The latter are reported to be distributed in size with mean diameters between 100–300 nm, in detectable amounts in solutions at all concentrations, even well below the solubility limit at 1 mg g^{-1} of water, with larger aggregates appearing with increasing concentration. They report that the clusters do not constitute a separate phase, describing the solution structure as "a thermodynamically stable mesostructured liquid containing solute-rich domains dispersed within bulk solute solution".[62]

Sedlák carried out a large, systematic study of the scattering from mesoscale clusters in aqueous solutions of various small molecules, including citric acid, acetic acid, glucose and urea.[63,64] Using a combination of DLS and SLS on carefully prepared and filtered solutions, he reports the presence of close-to-spherical regions within a broad size range up to several hundred nm. By combining analysis with a centrifugation technique, it could be established that the density of the clusters exceeded that of the bulk solution. The kinetics involved in the cluster development process are reported to be very slow; time scales for the development of the clusters was found to range from minutes to weeks, depending on the system. The growth of clusters with time was generally observed to lead to a broadening of the size distribution.

More complex organic compounds

Although the number of studies is still low, the situation appears similar for more complex organic molecules in organic solvents. Using a combination of SAXS, DLS and NTA, clustering in the system fenoxycarb/isopropanol was recently investigated over a wide range of concentrations, covering the range from dilute, undersaturated to supersaturated solutions. The study revealed that distributions of clusters ranging in size up to a micrometer are established in stagnant solutions, with mean sizes adjusting to conditions of temperature and conditions, but governed by slow kinetics (of the order of days).[68] The direction of the phenomenon as well as the slow kinetics involved agree well with the observed dependence of nucleation on solution history for this system.[44]

For the pharmaceutical intermediate 2-cyano-4′-methylbiphenyl (OTBN), clustering in two organic solvents (methanol and chloroform) was investigated by DLS, NTA and SAXS.[66] Mesoscale clusters were detected in undersaturated solutions above a certain concentration threshold in both solvents, with mean sizes showing only a slight or no dependence on the concentration above this threshold. An intriguing dependence of cluster size and agitation was observed; in solutions subjected to a period of agitation before analysis, mean clusters sizes increased by an order of magnitude. The exact mechanism behind this observation is not understood. Similar observations have been made for the API olanzapine.[50] Using NTA and AFM, mesoscopic clusters of a disordered nature, as verified by the absence of rotation of polarized light, were observed in aqueous-ethanol solutions. Cluster sizes were comparatively small, 35 nm on average, and exhibited no dependence on the solute concentration within the investigated range. However, the concentration of clusters was found to increase both with the bulk solute concentration and time, with a time scale of the order of hours. Moreover, the same dependencies were observed when the solvent composition was varied; the mean size remaining the same, but the concentration of clusters changing with the ethanol content.

Properties of mesoscale clusters

The various indirect observations of mesoscale clusters of organic molecules in solution, as well as the growing direct experimental evidence of their existence, have actualized a large number of fundamental questions as to the nature of these species. Examples of major questions – by no means all the relevant questions that could be posed – could be: i) how big are they, and do they have an equilibrium size or size distribution? ii) What is their composition and structure? iii) At what concentrations are they present in solution? iv) Do they represent a generally occurring phenomenon? v) How do the size and concentration depend on bulk properties, such as temperature, solute concentration and solvent? And vi) what kinetics control the dynamics and establishment of steady state cluster populations? Various studies have tried to answer some of these questions. To some extent, the picture is starting to clear, and some consensus is developing, but in other areas very little is still known. It should be stressed that the majority of studies have focused on a small set of molecules (*cf.* Fig. 3). To what extent observations of mesoscale clusters in the investigated systems can be generalized to other systems is still not clear. This is partly due to selection and publication bias. A broader survey of the generality of the phenomenon, and a more systematic investigation of its manifestation, is sorely needed.

Size. The mean size reported for clusters of different organic compounds varies in the range tens to hundreds of nm. The mean size is however not a very good descriptor of these entities except as a simple response variable. Rather, it has been repeatedly shown that the clusters can occur in rather broad distributions. There is a notable consensus in the literature that mesoscale clusters constitute a distinct population, separate from the smaller *molecular* clusters (*e.g.* dimers and other oligomers).

Structure and composition. Here, the main question is whether the clusters contain regions of long-range order, or whether they are best characterized as amorphous or liquid-like. There are indirect indications that clusters could contain some kind of structure, either reminiscent of a previously dissolved phase or which can at least template nucleation of a particular structure, but this requires further study. A range of spectroscopic and scattering-based techniques have been employed to obtain information about molecular interactions in clusters, and how they relate to different crystal structures. In a novel approach, Ward *et al.*[67] measured the level of second-harmonic scattering from concentrated aqueous urea solutions at under-as well as supersaturated conditions. This type of scattering is produced by certain crystal symmetries (typically non-centrosymmetric), to which the known polymorph of urea belongs. The measurements indicated that the clusters feature at least some degree of order resembling the final crystal structure. As to the composition, several studies suggest that the clusters could contain a significant amount of solvent molecules. Overall, there seems to be agreement that they have a higher solute density than the bulk solution.

A related question is whether the clusters are smooth and uniform in shape, or whether they comprise loose aggregates of smaller entities. For glycine, this has been investigated through the fractal dimension, using information from SAXS measurements.[18] The tentative conclusion is that clusters are initially mass fractals (loose aggregates or liquid-like clusters) which transform *via* surface fractals to smoother non-fractal entities.

Concentration. The number concentration of mesoscale clusters is not typically reported in studies, as it is not straightforward to measure with DLS. However, based on NTA measurements, it seems clear that in the majority of cases, the fraction of solute molecules involved in mesoscale clustering is very small. As an example, it has been reported that about 10^{-7} to 10^{-5} (*i.e.* less than 0.001%) of all olanzapine molecules are part of mesoclusters in aqueous or aqueous-ethanol solutions.[50]

Dependence on solute concentration and temperature. A diverging picture is emerging over how the cluster size depends on the solute concentration. Some studies have found virtually no such dependence, with mean cluster sizes remaining steady over a wide range of concentrations,[50,66] while others indicate that the steady state hydrodynamic diameter increases with solute concentration.[65,68] A third possibility is that the size distribution broadens to include larger clusters with increasing solute concentration. As to the number concentration of clusters, however, it seems well established that it increases with solute concentration. The dependence on temperature seems to have been explored to a lesser extent, and with partly contradictory indications. Given the importance of temperature in connection with solution history effects, it is clear that this needs to be systematically investigated.

Kinetics of clustering processes. Perhaps the most intriguing aspect of mesoscale clustering is the strikingly slow kinetics involved in the establishment or readjustment of steady state cluster populations. Some studies suggest that time constants involved are of the order of hours or days. If this is indeed so, it needs to be clearly established, as it has important implications for crystallization. The propensity of nucleation as well as growth, and the possibility to crystallize different polymorphs, could have a non-negligible dependence on the pretreatment of solutions, which is connected to mesoscale clustering, as has indeed been suggested for some systems.[37,68]

Diversity and evolving theory: beyond classical nucleation

There is no doubt that the classical nucleation theory is an oversimplification, even for cases where it works reasonably well. It is well known that theoretically calculated nucleation rates and kinetic factors based on the classical theory can

deviate significantly from experimental values.[69,70] During the last decade, a family of new, non-classical theories have emerged (Fig. 1). These theories, acknowledging the existence and important role of mesoscale entities for the nucleation process, attempt to address a number of major issues with the CNT:

i) that it does not separate the processes of increasing density and increasing order (assuming fully crystalline clusters and a single thermodynamic barrier);

ii) that it does not allow for sufficient flexibility regarding the growth units of the pre-critical cluster (assuming growth proceeds through attachment of monomers);

iii) that a steady-state distribution of clusters and a steady-state nucleation rate is immediately established, and

iv) that the thermodynamic treatment, at least in its most simple form, assumes that the surface tension of the nucleus is equal to that of a completely flat interface, neglecting its shape or curvature.

The two-step theory

Observations of crystallisation of macromolecules and colloids, for which nucleation can be orders of magnitude slower than for smaller molecules, has resulted in the proposal of a *two-step theory* of nucleation.[69,71,72] In the first step, local fluctuations in concentration result in the formation of clusters. The clusters have an amorphous, or liquid-like structure, an elevated solute concentration, and can range in size from 10^1 to 10^3 nm.[23] Nucleation of the crystal phase can then take place within the clusters. The lifetime, composition, internal and interfacial structure of the clusters or dense liquid droplets is a subject of debate.[69]

The two-step theory separates nucleation into two order parameters, local concentration and local structure, where the CNT assumes simultaneous fluctuations in both parameters. In free energy terms, shown schematically in Fig. 4, the two-step nucleation pathway features not one minimum as in the CNT, but two; the first corresponds to the formation of the dense, amorphous clusters and the second the nucleation of a periodically ordered solid phase inside the clusters. Two cases can be distinguished. If the dense liquid is thermodynamically stable with respect to the bulk solution, the phenomenon corresponds to a liquid–liquid phase separation (LLPS, or *oiling out*).[15,73–75] If the macroscopically stable state of the solution is as one liquid phase, the amorphous clusters are metastable.[76]

Evidence for two-step-type nucleation is not limited to macromolecules and colloids, however. For example, fluorescence spectroscopy has been used to show that the compound 4,4′-di-*tert*-butyldibenzoylmethanatoboron difluoride nucleates in a two-step process from dense, amorphous clusters.[77] Alison *et al.*,[52] using a combination of SAXS and WAXD, showed that the compound 2,6-dibromo-4-nitroaniline nucleated in two steps, initially as a non-diffractive (amorphous) phase that subsequently became diffractive (crystalline). Based on molecular dynamics
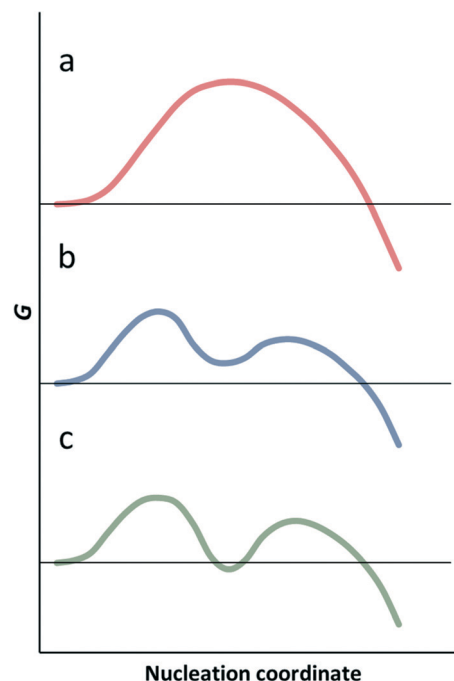

Fig. 4 Gibbs energy profiles for classical nucleation (a), and two-step nucleation with metastable clusters (b) and stable clusters (c).

simulations, two-step nucleation has been predicted for several systems, including glycine during antisolvent crystallization,[78] and acetic acid from tetrachloromethane during evaporative crystallization.[79]

Nucleation from pre-nucleation clusters

An alternative, but in many ways similar and at least partly overlapping theory, was developed for nucleation of calcium carbonate in biomineralization processes. According to this theory, nucleation will proceed from aggregation and ordering processes involving *pre-nucleation clusters* (PNCs). A significant body of work has been contributed by the Cölfen group from Konstanz, Germany.[14,21,80] Although the lion's share of studies has concerned inorganic systems, the mechanism has also been proposed for organic systems such as amino acids. Unlike the case for proteins, the PNCs are postulated to be thermodynamically stable entities, lacking a defined phase boundary, and to be present in solutions at undersaturated as well as supersaturated conditions in a distribution governed by a dynamic equilibrium with the solution. They are proposed to be amorphous or nano-crystalline, often containing significant amounts of solvent molecules, and with variations in size and structure occurring on timescales of hundreds of picoseconds. In addition, they are postulated to be able to have structural motifs resembling the structure of a bulk crystal polymorph.[80] Cluster formation has been proposed to be driven by the entropy gain from reduced solvation.[81] Nucleation is thought to occur *via* aggregation, interface elimination and growth, either by classical mechanisms or by

oriented attachment. Crystalline particles may then form within the aggregates, in analogy with the two-step theory.

Interestingly, the average size of mesoscale clusters, as discussed in a previous section, has repeatedly been shown to be either independent of the supersaturation, or even increase as a function of supersaturation. Inasmuch as these mesoscale clusters can be considered pre-nucleation clusters, this appears to contradict the classical theory, which states that the size of a critical cluster should become smaller with increasing supersaturation.[82] However, given the differences between the two theories, and the fact that little is actually known about the mechanism of PNC aggregation and growth, the comparison is not straightforward.

It must be stressed that there is a significant amount of terminological overlap or even confusion in the literature, where terms like 'clusters' or even 'prenucleation clusters' are used to describe stable as well as metastable entities by proponents of both two-step and PNC theories. Moreover, the terms are frequently used to denote a range of different types of aggregates, ranging from smaller aggregates of the order of molecules, such as dimers, through larger mesoscale aggregates up to liquid droplets. This reflects the complex energy landscape of clusters, with many minima corresponding to various entities deviating from the concept of monomers in the CNT. As will be discussed in the following, there is an emerging debate over to what extent either of these nucleation theories are universally applicable, to what extent they in fact overlap, and to what extent, as some recent studies suggest, a more versatile view of nucleation should be recognized.[83]

Emerging diversity

Recently, there is a trend in the field of nucleation research towards recognizing the possibility for more complex and system-specific nucleation mechanisms. This development accompanies the increased capabilities of – and the increased output from – scattering- and electron-based, and spectroscopic, imaging and analysis techniques, as well as computational modelling. For example, liquid cell TEM was recently used to show that nucleation of flufenamic acid from ethanol appears to involve elements of both the PNC and two-step theories.[57]

Innovative experimental studies of inorganic, metal–organic and macromolecular systems indicate that nucleation of the final crystalline particle can occur along surprisingly complex and hierarchical pathways. For example, using *in situ* liquid cell TEM, nanocrystals of palladium have been shown to emerge after a series of steps involving highly hydrated atomic, amorphous clusters of a size around 1 nm, aggregating to form so-called *cluster clouds*, from which a nucleus emerges. The crystallinity of the nucleus is then gradually improved through a rearrangement process involving interaction with the cluster cloud.[84] In another example, where high-res TEM was used to sample quenched solutions at various points along the nucleation coordinate,

is has been shown that gypsum originally nucleates as bassanite nanocrystals at undersaturated conditions, which transform into gypsum by hydration in a multi-step process.[85] A three-step mechanism (liquid–liquid phase separation into dense clusters, condensation into larger aggregates, and crystallization of the aggregates) has been observed for nucleation of the metal–organic framework material ZIF-8.[86]

One particularly well-studied system is calcium carbonate, $CaCO_3$, with three polymorphs (vaterite, aragonite and calcite). Different studies have suggested that this compound nucleates through a two- or a multi-step process from dense, amorphous or liquid-like phases,[87] or mediated by pre-nucleation clusters.[14] MD simulations predict the emergence of a stable population of liquid-like hydrated $CaCO_3$ clusters in solutions where pre-nucleation clusters have been reported.[88] In one well-cited study, liquid-cell TEM studies covering the entire nucleation process have shown competing nucleation mechanisms in play simultaneously. Calcite appears to nucleate only *via* direct association of ions, *i.e.* through a classical mechanism, while *e.g.* vaterite is observed to nucleate from spherical amorphous aggregates.[89] The study, like others, also suggests that there could be added diversity as regards the degree of hydration of clusters and the mechanisms whereby crystal nuclei are formed from amorphous clusters.

It has even been reported that the predominant nucleation mechanism in a system can shift depending on the conditions. Using AFM, the 2D nucleation of glucose isomerase has been shown to depend on the solute concentration. At low concentrations, a classical mechanism is observed, while at higher concentrations a two-step pathway is followed, as is typically observed for 3D nucleation of this compound.[90] For heterogeneous nucleation of ice, Markov state MD modelling suggests that the relative rates of classical and two-step nucleation can depend on temperature, with a transition point above which classical mechanism is predicted to dominate.[91]

Conclusions and outlook

It is clear that the general trend in nucleation research is moving towards diversification. The classical nucleation theory has in its extended form long been treated as a universal, all-encompassing model. More recent non-classical theories have gradually allowed for a mechanistically more complex process, but also opened up for higher diversity. It is indisputable that some systems are better described by two- or multi-step theories while for many others the classical approach works better. In some cases, competing mechanisms can apparently either coexist or exist in separate occurrence domains. Recent advances in experimental as well as modelling capabilities are continually extending the boundaries of possibilities for studying solution structuring, uncovering a complex landscape of diverse aggregated entities, whose size and structure and growth- and

transformation kinetics exhibit diverse, system-specific dependencies on conditions such as temperature, concentration and bulk supersaturation. This is starting to be mirrored in the diversity of nucleation mechanisms proposed under the umbrella term of *non-classical nucleation*.[92]

As regards mesoscale clusters of organic molecules in solution, there is a great need for further exploration, involving a systematic study of the dependence of different variables, and for a much wider range of systems than have so far been studied. Very little is still known about the steady-state properties of these clusters, such as structure, composition and shape, and how they depend on conditions in the solution, such as concentration, temperature and chemical environment. Even less is understood about the dynamics of these clusters, their lifetime, and the kinetics of dissolution, growth and aggregation. It is becoming increasingly clear that mesoscale clusters can play a decisive role in crystal nucleation. This should make the study of these intriguing entities a top priority for the crystallizing community. The tools for this are already available, and they are rapidly getting more accurate.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

References

1 M. R. Abu Bakar, Z. K. Nagy, A. N. Saleemi and C. D. Rielly, The Impact of Direct Nucleation Control on Crystal Size Distribution in Pharmaceutical Crystallization Processes, *Cryst. Growth Des.*, 2009, **9**, 1378.

2 S. Byrn, R. Pfeiffer, M. Ganey, C. Hoiberg and G. Poochikian, Pharmaceutical solids: a strategic approach to regulatory considerations, *Pharm. Res.*, 1995, **12**, 945.

3 S. R. Chemburkar, J. Bauer, K. Deming, H. Spiwek, K. Patel, J. Morris, R. Henry, S. Spanton, W. Dziki, W. Porter, J. Quick, P. Bauer, J. Donaubauer, B. A. Narayanan, M. Soldani, D. Riley and K. McFarland, Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development, *Org. Process Res. Dev.*, 2000, **4**, 413.

4 S. Sundareswaran and S. Karuppannan, Nucleation control and separation of vanillin polymorphs I and II through the swift cooling crystallization process, *CrystEngComm*, 2021, **23**, 1634.

5 G. Lincke, A review of thirty years of research on quinacridones. X-ray crystallography and crystal engineering, *Dyes Pigm.*, 2000, **44**, 101.

6 D. Kashchiev, *Nucleation: basic theory with applications*, Butterworth-Heinemann, Oxford, 2000.

7 J. Liu and Å. C. Rasmuson, Influence of Agitation and Fluid Shear on Primary Nucleation in Solution, *Cryst. Growth Des.*, 2013, **13**, 4385.

8 T. Nakai, Effects of the thermal history of a solution to nucleation, *J. Chin. Inst. Chem. Eng.*, 1972, **3**, 83.

9 R. Becker and W. Döring, Kinetische Behandlung der Keimbildung in übersättigten Dämpfen, *Ann. Phys.*, 1935, **24**, 719.

10 Y. B. Zeldovich, Theory of new phase formation: cavitation, *Acta Physicochim. URSS*, 1943, **18**, 1.

11 T. Li, A. J. Senesi and B. Lee, Small Angle X-ray Scattering for Nanoparticle Research, *Chem. Rev.*, 2016, **116**, 11128.

12 N. Pienack and W. Bensch, *In Situ* Monitoring of the Formation of Crystalline Solids, *Angew. Chem., Int. Ed.*, 2011, **50**, 2014.

13 F. Giberti, M. Salvalaglio and M. Parrinello, Metadynamics studies of crystal nucleation, *IUCrJ*, 2015, **2**, 256.

14 D. Gebauer, A. Völkel and H. Cölfen, Stable Prenucleation Calcium Carbonate Clusters, *Science*, 2008, **322**, 1819.

15 T. J. Sorensen, Oiling-out and crystallization of vanillin from aqueous solutions, *Chem. Eng. Technol.*, 2014, **37**, 1959.

16 T. J. Sorensen, P. C. Sontum, J. Samseth, G. Thorsen and D. Malthe-Sorenssen, Cluster formation in precrystalline solutions, *Chem. Eng. Technol.*, 2003, **26**, 307.

17 U. Gasser, E. R. Weeks, A. Schofield, P. N. Pusey and D. A. Weitz, Real-space imaging of nucleation and growth in colloidal crystallization, *Science*, 2001, **292**, 258.

18 S. Chattopadhyay, D. Erdemir, J. M. B. Evans, J. Ilavsky, H. Amenitsch, C. U. Segre and A. S. Myerson, SAXS study of the nucleation of glycine crystals from a supersaturated solution, *Cryst. Growth Des.*, 2005, **5**, 523.

19 S. Tanaka, K. Ito, R. Hayakawa and M. Ataka, Size and number density of precrystalline aggregates in lysozyme crystallization process, *J. Chem. Phys.*, 1999, **111**, 10330.

20 A. Jawor-Baczynska, J. Sefcik and B. D. Moore, 250 nm Glycine-Rich Nanodroplets Are Formed on Dissolution of Glycine Crystals But Are Too Small To Provide Productive Nucleation Sites, *Cryst. Growth Des.*, 2013, **13**, 470.

21 D. Gebauer and H. Cölfen, Prenucleation clusters and non-classical nucleation, *Nano Today*, 2011, **6**, 564.

22 Y. Jiang, M. Kellermeier, D. Gebauer, Z. H. Lu, R. Rosenberg, A. Moise, M. Przybylski and H. Colfen, Growth of organic crystals *via* attachment and transformation of nanoscopic precursors, *Nat. Commun.*, 2017, **8**, 6.

23 S. Karthika, T. K. Radhakrishnan and P. Kalaichelvi, A Review of Classical and Nonclassical Nucleation Theories, *Cryst. Growth Des.*, 2016, **16**, 6663.

24 J. W. Mullin and C. L. Leci, Evidence of molecular cluster formation in supersaturated solutions of citric acid, *Philos. Mag.*, 1969, **19**, 1075.

25 M. A. Larson and J. Garside, Solute clustering in supersaturated solutions, *Chem. Eng. Sci.*, 1986, **41**, 1285.

26 R. M. Ginde and A. S. Myerson, Cluster size estimation in binary supersaturated solutions, *J. Cryst. Growth*, 1992, **116**, 41.

27 E. L. Cussler, Cluster diffusion in liquids, *AIChE J.*, 1980, **26**, 43.

28 A. S. Myerson and P. Y. Lo, Diffusion and cluster formation in supersaturated solutions, *J. Cryst. Growth*, 1990, **99**, 1048.

29 R. Mohan, O. Kaytancioglu and A. S. Myerson, Diffusion and cluster formation in supersaturated solutions of ammonium sulfate at 298 K, *J. Cryst. Growth*, 2000, **217**, 393.

30 J. W. Mullin and K. D. Raven, Influence of Mechanical Agitation on the Nucleation of Some Aqueous Salt Solutions, *Nature*, 1962, **195**, 35.

31 C. Forsyth, P. A. Mulheran, C. Forsyth, M. D. Haw, I. S. Burns and J. Sefcik, Influence of Controlled Fluid Shear on Nucleation Rates in Glycine Aqueous Solutions, *Cryst. Growth Des.*, 2014, **15**, 94.

32 J. Liu, M. Svärd and Å. C. Rasmuson, Influence of Agitation and Fluid Shear on Nucleation of m-Hydroxybenzoic Acid Polymorphs, *Cryst. Growth Des.*, 2014, **14**, 5521.

33 P. M. Martin-Soladana, M. Diwan, H. Li, F. L. Nordstrom and A. Sheikh, Investigation of Nucleation under High-Shear Conditions for a Pharmaceutical Compound in an Unseeded System, *Org. Process Res. Dev.*, 2019, **23**, 2627.

34 L. E. V. Shmidt and J. Shmidt, Mechanism of crystallization in agitated solutions, *Chem. Eng. Commun.*, 1985, **36**, 233.

35 K. Sypek, I. S. Burns, A. J. Florence and J. Sefcik, *In situ* monitoring of stirring effects on polymorphic transformations during cooling crystallization of carbamazepine, *Cryst. Growth Des.*, 2012, **12**, 4821.

36 M. J. Vesga, D. McKechnie, P. A. Mulheran, K. Johnston and J. Sefcik, Conundrum of γ glycine nucleation revisited: to stir or not to stir?, *CrystEngComm*, 2019, **21**, 2234.

37 F. L. Nordström, M. Svärd, B. Malmberg and Å. C. Rasmuson, Influence of Solution Thermal and Structural History on the Nucleation of m-Hydroxybenzoic Acid Polymorphs, *Cryst. Growth Des.*, 2012, **12**, 4340.

38 K. Schaum and F. Schoenbeck, Unterkühlung und Krystallisation von Schmelzflüssen polymorpher Stoffe, *Ann. Phys.*, 1902, **313**, 652.

39 M. W. Burke, R. A. Judge and M. L. Pusey, The effect of solution thermal history on chicken egg white lysozyme nucleation, *J. Cryst. Growth*, 2001, **232**, 301.

40 J. Nývlt, Effect of thermal history of aqueous solutions of potassium chloride on the metastable zone width, *Collect. Czech. Chem. Commun.*, 1984, **49**, 559.

41 N. Kubota and Y. Fujisawa, The effects of filtration and thermal history on primary nucleation of potassium bromate from aqueous solution, *Process Technol. Proc.*, 1984, **2**, 259.

42 K. Hussain, G. Thorsen and D. Malthe-Sorenssen, Nucleation and metastability in crystallization of vanillin and ethyl vanillin, *Chem. Eng. Sci.*, 2001, **56**, 2295.

43 J. Nývlt and V. Pekárek, Crystallization studies by thermometric methods. III. The effect of thermal history of solutions on nucleation, *Z. Phys. Chem.*, 1980, **122**, 199.

44 M. Kuhs, J. Zeglinski and Å. C. Rasmuson, Influence of History of Solution in Crystal Nucleation of Fenoxycarb: Kinetics and Mechanisms, *Cryst. Growth Des.*, 2014, **14**, 905.

45 N. Javid, T. Kendall, I. S. Burns and J. Sefcik, Filtration Suppresses Laser-Induced Nucleation of Glycine in Aqueous Solutions, *Cryst. Growth Des.*, 2016, **16**, 4196.

46 J. Hirschler, M. H. Charon and J. C. Fontecilla-Camps, The effects of filtration on protein nucleation in different growth media, *Protein Sci.*, 1995, **4**, 2573.

47 S. V. Gorbachev and A. V. Shlykov, Effect of purification of the solution on the limiting supersaturation, *Zh. Fiz. Khim.*, 1955, **29**, 797.

48 M. Kellermeier, R. Rosenberg, A. Moise, U. Anders, M. Przybylski and H. Colfen, Amino acids form prenucleation clusters: ESI-MS as a fast detection method in comparison to analytical ultracentrifugation, *Faraday Discuss.*, 2012, **159**, 23.

49 D. Hagmeyer, J. Ruesing, T. Fenske, H.-W. Klein, C. Schmuck, W. Schrader, M. E. M. D. Piedade and M. Epple, Direct experimental observation of the aggregation of α-amino acids into 100–200 nm clusters in aqueous solution, *RSC Adv.*, 2012, **2**, 4690.

50 M. Warzecha, M. S. Safari, A. J. Florence and P. G. Vekilov, Mesoscopic Solute-Rich Clusters in Olanzapine Solutions, *Cryst. Growth Des.*, 2017, **17**, 6668.

51 J. J. De Yoreo, A holistic view of nucleation and self-assembly, *MRS Bull.*, 2017, **42**, 525.

52 H. G. Alison, R. J. Davey, J. Garside, M. J. Quayle, G. J. T. Tiddy, D. T. Clarke and G. R. Jones, Using a novel plug flow reactor for the *in situ*, simultaneous, monitoring of SAXS and WAXD during crystallisation from solution, *Phys. Chem. Chem. Phys.*, 2003, **5**, 4998.

53 Y. Tsarfati, S. Rosenne, H. Weissman, L. J. W. Shimon, D. Gur, B. A. Palmer and B. Rybtchinski, Crystallization of Organic Molecules: Nonclassical Mechanism Revealed by Direct Imaging, *ACS Cent. Sci.*, 2018, **4**, 1031.

54 M. Kellermeier, D. Gebauer, E. Melero-García, M. Drechsler, Y. Talmon, L. Kienle, H. Cölfen, J. M. García-Ruiz and W. Kunz, Colloidal Stabilization of Calcium Carbonate Prenucleation Clusters with Silica, *Adv. Funct. Mater.*, 2012, **22**, 4301.

55 M. Pouget Emilie, H. H. Bomans Paul, A. C. M. Goos Jeroen, M. Frederik Peter, G. de With and A. J. M. Sommerdijk Nico, The Initial Stages of Template-Controlled CaCO3 Formation Revealed by Cryo-TEM, *Science*, 2009, **323**, 1455.

56 J. Hermannsdörfer and N. de Jonge, Studying Dynamic Processes of Nano-sized Objects in Liquid using Scanning Transmission Electron Microscopy, *J. Visualized Exp.*, 2017, e54943.

57 J. Cookman, V. Hamilton, S. R. Hall and U. Bangert, Non-classical crystallisation pathway directly observed for a pharmaceutical crystal *via* liquid phase electron microscopy, *Sci. Rep.*, 2020, **10**, 19156.

58 C. K. Meng and J. B. Fenn, Formation of charged clusters during electrospray ionization of organic solute species, *Org. Mass Spectrom.*, 1991, **26**, 542.

59 Z. Takats, S. C. Nanita, R. G. Cooks, G. Schlosser and K. Vekey, Amino Acid Clusters Formed by Sonic Spray Ionization, *Anal. Chem.*, 2003, **75**, 1514.

60 C. E. Hughes, S. Hamad, K. D. M. Harris, C. R. A. Catlow and P. C. Griffiths, A multi-technique approach for probing the evolution of structural properties during crystallization of organic materials from solution, *Faraday Discuss.*, 2007, **136**, 71.

61 G. Zimbitas, A. Jawor-Baczynska, M. J. Vesga, N. Javid, B. D. Moore, J. Parkinson and J. Sefcik, Investigation of molecular and mesoscale clusters in undersaturated glycine aqueous solutions, *Colloids Surf., A*, 2019, **579**, 123633.

62 A. Jawor-Baczynska, B. D. Moore, H. S. Lee, A. V. McCormick and J. Sefcik, Population and size distribution of solute-rich mesospecies within mesostructured aqueous amino acid solutions, *Faraday Discuss.*, 2013, **167**, 425.

63 M. Sedlák, Large-Scale Supramolecular Structure in Solutions of Low Molar Mass Compounds and Mixtures of Liquids: I. Light Scattering Characterization, *J. Phys. Chem. B*, 2006, **110**, 4329.

64 M. Sedlák, Large-Scale Supramolecular Structure in Solutions of Low Molar Mass Compounds and Mixtures of Liquids: II. Kinetics of the Formation and Long-Time Stability, *J. Phys. Chem. B*, 2006, **110**, 4339.

65 D. L. Sidebottom, Ultraslow relaxation of hydrogen-bonded dynamic clusters in glass-forming aqueous glucose solutions: A light scattering study, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2007, **76**, 11505.

66 S. Zong, J. Wang, X. Huang, T. Wang, B. Liu, B. Tian, C. Xie and H. Hao, Molecular evolution pathways during nucleation of small organic molecules: solute-rich pre-nucleation species enable control over the nucleation process, *Phys. Chem. Chem. Phys.*, 2020, **22**, 18663.

67 M. R. Ward, S. W. Botchway, A. D. Ward and A. J. Alexander, Second-harmonic scattering in aqueous urea solutions: evidence for solute clusters?, *Faraday Discuss.*, 2013, **167**, 441.

68 M. Svärd, K. R. Devi, D. Khamar, D. Mealey, D. Cheuk, J. Zeglinski and Å. C. Rasmuson, Solute clustering in undersaturated solutions – systematic dependence on time, temperature and concentration, *Phys. Chem. Chem. Phys.*, 2018, **20**, 15550.

69 D. Erdemir, A. Y. Lee and A. S. Myerson, Nucleation of crystals from solution: classical and two-step models, *Acc. Chem. Res.*, 2009, **42**, 621.

70 J. W. P. Schmelzer, A. R. Gokhman and V. M. Fokin, Dynamics of first-order phase transitions in multicomponent systems: a new theoretical approach, *J. Colloid Interface Sci.*, 2004, **272**, 109.

71 P. G. Vekilov, Nucleation, *Cryst. Growth Des.*, 2010, **10**, 5007.

72 W. Pan, A. B. Kolomeisky and P. G. Vekilov, Nucleation of ordered solid phases of proteins *via* a disordered high-density state: phenomenological approach, *J. Chem. Phys.*, 2005, **122**, 174905.

73 M. Svärd, S. Gracin and Å. C. Rasmuson, Oiling out or molten hydrate – liquid-liquid phase separation in the system vanillin-water, *J. Pharm. Sci.*, 2007, **96**, 2390.

74 S. Veesler, L. Lafferrère, E. Garcia and C. Hoff, Phase Transitions in Supersaturated Drug Solution, *Org. Process Res. Dev.*, 2003, **7**, 983.

75 P. E. Bonnett, K. J. Carpenter, S. Dawson and R. J. Davey, Solution crystallisation *via* a submerged liquid-liquid phase boundary: Oiling out, *Chem. Commun.*, 2003, 698.

76 D. N. Petsev, X. Wu, O. Galkin and P. G. Vekilov, Thermodynamic Functions of Concentrated Protein Solutions from Phase Equilibria, *J. Phys. Chem. B*, 2003, **107**, 3921.

77 F. Ito, Y. Suzuki, J.-I. Fujimori, T. Sagawa, M. Hara, T. Seki, R. Yasukuni and M. L. D. L. Chapelle, Direct Visualization of the Two-step Nucleation Model by Fluorescence Color Changes during Evaporative Crystallization from Solution, *Sci. Rep.*, 2016, **6**, 22918.

78 A. Kitayama, K. Kadota, Y. Tozuka, A. Shimosaka, M. Yoshida and Y. Shirakawa, Molecular aspects of glycine clustering and phase separation in an aqueous solution during anti-solvent crystallization, *CrystEngComm*, 2020, **22**, 5182.

79 A. Gavezzotti, Molecular Aggregation of Acetic Acid in a Carbon Tetrachloride Solution: A Molecular Dynamics Study with a View to Crystal Nucleation, *Chem. – Eur. J.*, 1999, **5**, 567.

80 D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergström and H. Cölfen, Pre-nucleation clusters as solute precursors in crystallisation, *Chem. Soc. Rev.*, 2014, **43**, 2348.

81 M. Kellermeier, P. Raiteri, K. Berg John, A. Kempter, D. Gale Julian and D. Gebauer, Entropy Drives Calcium Carbonate Ion Association, *ChemPhysChem*, 2016, **17**, 3535.

82 E. D. Bøjesen and B. B. Iversen, The chemistry of nucleation, *CrystEngComm*, 2016, **18**, 8332.

83 J. De Yoreo, Crystal nucleation: More than one pathway, *Nat. Mater.*, 2013, **12**, 284.

84 B. Jin, Y. Wang, Z. Liu, A. France-Lanord, J. C. Grossman, C. Jin and R. Tang, Revealing the Cluster-Cloud and Its Role in Nanocrystallization, *Adv. Mater.*, 2019, **31**, 1808225.

85 A. E. S. Van Driessche, L. G. Benning, J. D. Rodriguez-Blanco, M. Ossorio, P. Bots and J. M. García-Ruiz, The Role and Implications of Bassanite as a Stable Precursor Phase to Gypsum Precipitation, *Science*, 2012, **336**, 69.

86 X. Liu, W. Chee See, S. Raj, M. Sawczyk, P. Král and U. Mirsaidov, Three-step nucleation of metal–organic framework nanocrystals, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2008880118.

87 E. M. Pouget, P. H. H. Bomans, A. Dey, P. M. Frederik, G. de With and N. A. J. M. Sommerdijk, The Development of Morphology and Structure in Hexagonal Vaterite, *J. Am. Chem. Soc.*, 2010, **132**, 11560.

88 A. F. Wallace, L. O. Hedges, A. Fernandez-Martinez, P. Raiteri, J. D. Gale, G. A. Waychunas, S. Whitelam, J. F. Banfield and J. J. De Yoreo, Microscopic Evidence for Liquid-Liquid Separation in Supersaturated $CaCO_3$ Solutions, *Science*, 2013, **341**, 885.

89 M. H. Nielsen, S. Aloni and J. J. De Yoreo, *In situ* TEM imaging of $CaCO_3$ nucleation reveals coexistence of direct and indirect pathways, *Science*, 2014, **345**, 1158.

90 M. Sleutel, J. Lutsko, A. E. S. Van Driessche, M. A. Durán-Olivencia and D. Maes, Observing classical nucleation theory

at work by monitoring phase transitions with molecular precision, *Nat. Commun.*, 2014, **5**, 5598.

91 C. Li, Z. Liu, E. C. Goonetilleke and X. Huang, Temperature-dependent kinetic pathways of heterogeneous ice nucleation competing between classical and non-classical nucleation, *Nat. Commun.*, 2021, **12**, 4954.

92 B. Jin, Z. Liu and R. Tang, Recent experimental explorations of non-classical nucleation, *CrystEngComm*, 2020, **22**, 4057.