CrystEngComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/crystengcomm

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

A perspective on the Growth-only Zone, the Secondary Nucleation Threshold and Crystal Size Distribution in Solution Crystallisation

Terence L. Threlfall*[a] and Simon J. Coles[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

The metastable zone between the solubility curve and the crystallisation curve can be divided into two regions, separated by the secondary nucleation threshold. At a supersaturation less than that of the secondary nucleation threshold seeds will grow but not proliferate. At greater supersaturations, they will proliferate. Although this concept has been known in connection with the commercial crystallisation of sucrose and of lactose for almost a century it is virtually unrecognised beyond the sugar and dairy industries. The consequences are explored of the existence of such a boundary on the crystallisation process and the influence on crystal size distribution. It is recognised that the secondary nucleation threshold is the 'seeded metastable limit', but there are important consequences of avoiding the latter term. Mullin's 'Latent Period' is shown to be closely related to the 'growth only' zone within the metastable zone.

Introduction

Within the metastable zone of solution crystallisation lies an area within which only crystal growth can occur unaccompanied by crystal proliferation. In one sense this is extremely well known and has been well used for over a century. Yet it is a concept that is barely mentioned, recognised or acknowledged in the mainstream crystallisation literature. This is partly due to a lack of consensus on nomenclature describing the concept, but generally down to a lack of recognition or awareness of it. This perspective article aims, for the first time and often drawing on examples hidden unrecognised in the literature, to provide a standard set of terms and fully characterise this phenomenon.

We shall refer to the boundary separating the growth-only zone and the crystal proliferation zone as the 'Secondary Nucleation Threshold'. As will be seen in the historical summary immediately following, there have been a large number of terms used to describe what appears to be the same dividing line, a sure indication that the universality of the phenomenon has remained unrecognised.

During the preparation of a paper on inert seeds[1], the concept of a 'forced crystallisation line' was encountered in connection with the crystallisation of lactose[2-4]. Lactose crystals as small as 30 micrometres can feel gritty in the mouth, so a particle size of less than this is usually the goal of the crystallisation. This is achieved by forcing secondary nucleation by seeding solutions lying in the metastable zone at supersaturations greater than the forced crystallisation line with massive amounts of finely milled lactose: hence the designation. In academic papers on the crystallisation of sugars, what appears to be the same crystallisation line is called the 'secondary nucleation threshold' (SNT). The concept is also known in connection with the commercial crystallisation of sucrose although, surprisingly, it does not appear to have acquired any name in that context[5]. The secondary nucleation threshold is used in this case for the opposite purpose, the production of larger crystals, because the goal of sugar crystallisation is usually to produce crystals of uniform size for 'granulated sugar', or 'caster sugar' or 'preserving sugar'. There has been occasional mention of the existence of such a boundary within the metastable zone in connection with crystallisation beyond the context of the dairy and sugar industries. However, it neither seems to have been thoroughly discussed, nor the consequences of its existence explored. It was noted by Ting and McCabe[6] in 1934 that some spontaneous nucleation of magnesium sulphate heptahydrate occurred first at a temperature intermediate between the solubility curve and the point of massive secondary nucleation marking the limit of the metastable zone. These initial seeds however failed to proliferate, as has been discussed in detail in the previous paper[1], both for magnesium sulphate and for many other compounds. Ting and McCabe's data was converted into a line within the metastable zone by Soehnel[7], but not further explored. Tung[8] and Tung, Paul, Midler and McCauley[9] draw concentration/temperature diagrams with such a line displayed, but fail to comment on the consequences. Sangwal[10] assumes that proliferation occurs only near the metastable limit and thereby suggests the existence of a proliferation boundary. Kraus and Nyvlt[11] mention the concept of a region within which supersaturation is entirely consumed by existing crystals, implying a growth-only region. In connection with preferential seeding for chiral separation, Hongo, Yamada and Chibata[12] (and references therein) divide the metastable region into two.

Strickland–Constable[13] refers to a region in which only crystal growth can occur, but fails to elaborate on the implications of such a region.

Flood and co-workers[14] have identified a 'secondary nucleation threshold' in the crystallisation of carbohydrates and have extended the concept to the crystallisation of methionine[15]. The earliest use of the term secondary nucleation threshold appears to have been by White. White, Marziano and co-workers[16] have used the concept of a secondary nucleation threshold in the crystallisation of ibuprofen. It has also been investigated in connection with lysozyme crystallisation[17]. Khaddour and Rocha[18] have investigated 'secondary nucleation within the metastable zone', which is presumably the same phenomenon, but have not considered it in relation to product formation. The terms 'surface secondary nucleation', 'activated secondary nucleation' and 'surface-activated secondary nucleation' have been used by Puel, Veesler and co-workers[19-21] to describe what appears to be the onset of the same boundary. Agrawal and Paterson[22] show a concentration-temperature diagram for lactose with numerous lines between the solubility and metastable line variously named as 'forced crystallisation line'[2], 'critical supersaturation for growth'[23], 'secondary nucleation threshold'[24], 'forced secondary nucleation'[23] and 'nuclei first detection curve'[25]. It seems probable that these all represent the same thing, but measured under different conditions. It will be noted that there may be at least two phenomena being discussed under one umbrella here: failure of added seeds to proliferate and failure of spontaneously generated seeds to proliferate until the supersaturation increases beyond a certain point. A further issue concerns the careful experiments of Miers and Isaacs[26], who observed primary nucleation at or near the solubility curve and proliferation only at the metastable zone limit. It is not even certain that the primary product crystals are responsible for the subsequent nucleation, and therefore whether it can properly be described as secondary nucleation according to the accepted definition of the term. In describing the division as the 'dynamic metastable limit', further confusion about the nature of the boundary appears to have been generated[25]. It is important to realise that the secondary nucleation threshold marking the limit of this area is not the 'dynamic metastable limit'. The latter is the limit of spontaneous crystallisation, whilst the secondary nucleation threshold is the onset of crystal proliferation under seeding.

In our earlier perspective article[1], the discussion about the occurrence of inert seeds generally concentrated on intrinsically inert seeds, because the majority of the literature appeared to be concerned with these. Seeds inert in nature as a result of their position within the metastable zone of the concentration-temperature diagram were discussed, but only briefly. It has since been realised that the latter occurrence of inert seeds is in fact far more common, so the present paper is an attempt to redress the balance. Everyone who has grown crystals of alum, copper sulphate or other salts[27] at school has in fact observed non-proliferating crystals, yet, as already stated, it is barely mentioned in the research crystallisation literature. Every time a seeded crystallisation is conducted, the presence of an SNT is confirmed, but the necessary consequence of such an observation, namely that at supersaturations below those delineated by the SNT a non-proliferation zone must exist is rarely realised.

We felt that it was worthwhile to repeat the crystallisation of selected compounds in order to confirm and understand the generality of the phenomena being discussed, despite the overwhelming evidence from the lactose and sucrose industries of the reality and applicability of the concept. The extensive experimental results from this work are not included here, but will be reported in a separate paper. The only results presented here are some preliminary manual determinations of metastable zone widths and secondary nucleation thresholds shown in Table 1. The function of entries in this table is to facilitate the discussion by showing that the concepts under discussion are not confined to the crystallisation of sugars but are applicable to organic substances in general. The results presented in Table 1 were determined initially by manual experiments with detection by eye and refined and confirmed by turbidimetric measurements with an HEL CrystalEyes system.

Discussion

Secondary Nucleation Threshold

That there is a variation in the tendency for seed proliferation across the metastable zone appears to be widely accepted[28]. What is less recognised is that there may be a discrete boundary separating proliferating and non-proliferating regions. That a growth-only region without seed proliferation exists is supported by the school crystal growth experiments mentioned in the introduction. A wide range of supersaturations, often uncontrolled, results in crystal growth without extraneous seeding. This must have been observed on many millions of occasions. If any crystal proliferation is observed the simple advice is to try diluting the solutions. So at least for such substances the growth-only region must have a substantial width. But there are also accounts of the existence of a line separating proliferation and non-proliferation of crystals, some of which have been referenced above. The usual concentration/temperature diagram of solution crystallisation can therefore be elaborated, Figure 1. Solutions to the right of the boundary as drawn allow existing crystals to grow, but no proliferation occurs as a result of the presence of those crystals. Solutions to the left allow crystal growth and proliferation but do not spontaneously nucleate until the metastable zone boundary is reached. The amorphous and oiling out lines shown by Tung at al. have not been included in Figure 1, because they have been discussed in detail[9] so will not be further considered here.

Figure 1. The generalised solution crystallisation diagram, showing the thermodynamic solubility curve, the dead zone limit, the secondary nucleation threshold and the metastable limit. §

§A is the undersaturated region. B, C, and D are usually grouped together as the metastable region, but are here divided into B, the dead zone, C the growth only region, and D the crystal proliferation region. For unseeded crystallisation the dead zone limit will mark the primary heterogeneous crystallisation, and C will be the Latent Period. E is the region of spontaneous nucleation.

The Dead Zone

A further boundary has been drawn in Figure 1 within the metastable zone near the solubility curve. This is the limit of the dead zone within which even a seeded solution is incapable of either crystal growth or crystal proliferation. Such a crystallisation gap was first described by Mullin[29], and has also been noted by Flood and Wantha[30] and by Heffels, de Jong and Sinke[31] and discussed by the present authors[32] but one of the problems which this engenders deserves elaborating. If there is such a zone, showing an hysteresis, how is it possible to derive an exact thermodynamic solubility value, since an hysteresis and a thermodynamic equilibrium are incompatible for the same event? The explanation is probably that any collection of crystals will display a range of stabilities. The smallest crystals will be less stable as will crystals of extreme habit. Crystals with greater numbers of defects are likely to possess higher Gibbs' energies. Even within a perfect crystal, whatever is to be understood by that term, the edges and particularly the corners will be less stable. So a collection of crystals will start with a range of stabilities, but as the sizes and the imperfections are gradually ironed out, the measured solubility will approach the thermodynamic value. This sequence is probably a further reason why the *exact* measurement of solubility is so difficult and time-consuming[33]. The observations of Malivuk *et al.*[34] on the dissolution and growth of crystals at concentrations respectively greater and less than the thermodynamic solubility value support what is being described here. The inevitable conclusion from this discussion is that the dead-zone will be narrowed by the presence of greater numbers of seed crystals, and may eventually be reduced effectively to zero with time and seed area. However, this cannot happen when the limit is determined from isothermal crystallisation studies[32].

	Metastable Zone Width	Crystal Proliferation Zone
Ascorbic acid	16° (lit. 36-17°)[61]	4°
Glutamic acid[62]	47-37°	
D,L-Aspartic acid	40°	12°
L-Aspartic acid	45°	15°
D,L-Malic acid	19°	3°
D-Malic acid	11°	
Famotidine Polymorph A[63]	44°, 16°	
Famotidine Polymorph B[63]	3°	
Traditional sugars	~50°	
Xylose	36°	18°
Sorbitol	10°	
Mannitol	17°	9°
Isoascorbic acid	21°	6°
Raffinose	30°	10°
Ibuprofen[16]	5°	4°

Table 1. A selection of compounds displaying wide metastable zone widths §

§The ideal compound for these investigations would possess a wide metastable zone width but less than 40°C, and adequate solubility in an aqueous solution of moderate viscosity. The 'crystal proliferation zone' is the temperature interval between the secondary nucleation threshold and the metastable zone limit. There are two entries with narrow metastable zone widths, Famotidine Polymorph B, for comparison with Polymorph A, and Ibuprofen to show how close to the solubility curve the secondary nucleation threshold can lie.

Zone Boundary Sharpness

Each of the boundaries considered here, namely the metastable zone limit, the secondary nucleation threshold and the dead zone boundary are seen to be broad demarcations rather than sharp lines. In view of what is discussed next under 'Mechanistic Considerations', the SNT is not a sharp line and acquires its breadth because it moves according to the crystallisation parameters and so covers a range of values. Instead it must be an intrinsically broad boundary, although also subject to further broadening under the influence of crystallisation parameters. The reason for this intrinsic breadth is probably that crystals of different sizes appear to have different propensities for growth[1,31,35] so that any collection of seeds whether spontaneously generated (primary nucleation) or added, will have a range of sensitivities to growth and proliferation. The metastable zone limit and the secondary nucleation threshold are both dynamic boundaries dependent on kinetics. Therefore they could be determined either by polythermal or isothermal methods. There have been recent attempts to relate the metastable zone limit temperature to the induction times[36]. Conducting both these types of measurement would give a more complete picture of the crystallisation landscape. However, with the exception of the experiments described in reference 16, all investigations have been by polythermal methods. Solutions to the left of the metastable boundary, as is well known, can spontaneously nucleate so will not be considered further in this paper.

It will be proposed below that the secondary nucleation threshold is the representation of a change of crystal growth mechanism with increasing supersaturation, from ordered

crystal growth to disordered crystal growth.

Mechanistic Considerations

It is not obvious why seed proliferation should occur. Generally, despite the contrary examples collected in the earlier paper[1], the addition of one crystal to a sufficiently supersaturated solution will cause rapid proliferation, often at considerable distance from the seed. Under conditions of stirring this is normally considered to be due to collision with the impeller, although the number of new crystals often appears greatly to exceed the number of particles that might be generated by the impact. However it also occurs in quiescent solutions and with tethered crystal seeds, as was confirmed by experiments to be reported elsewhere on magnesium sulphate heptahydrate. It is only necessary to think of the total solidification brought about by a single pulse on a melt of sodium acetate, or the effect of adding a single seed crystal to a highly supersaturated solution of sodium thiosulfate to realize how rapidly crystallisation can spread in quiescent liquids. Contact nucleation has been much investigated, but before contact nucleation can occur in unseeded situations or in the presence of tethered seeds, a non-contact nucleation must take place to allow seeds to disperse throughout the solution. Even under seeded conditions, non-contact nucleation must be occurring, although it will be increasingly swamped during the course of the crystallisation by the contact nucleation.

The key to the problem of propensity for non-contact proliferation may lie in the statement by Randolph and Larson[35] that 'only growing crystals can proliferate'. We shall go further and suggest that only a rapidly growing crystal can proliferate. A seed crystal that is not growing, or one that is growing only by orderly addition of molecules at steps on the surface to allow layers to spread (birth and spread or spiral growth mechanisms[37]) has no means of throwing off particles to disseminate and provide new growth centres throughout the solution. At least, the classical descriptions of such processes fail to suggest any boundary layer-surface growth interaction which could generate such particles. By contrast, rapid growth at higher supersaturations, the so-called rough growth, might allow this, although not if considered solely as an 'adhesive' process[38]. The clustering in solution derived from molecular dynamics simulation[39] suggests a possible mechanism for such crystallisation transfer.

Throughout the literature many suggestions have been made that the process of growth must be more complex than that of the simple transfer of molecules from the solute phase to the crystal surface. So a transfer layer adjacent to the crystal surface is implied or proposed. Miers[40] in 1903 showed by refractive index measurement that the solution adjacent to the crystal face was more concentrated than that of the bulk, thereby indicating a boundary layer. Berg's interferometric measurements[41] have never been satisfactorily explained, which implies that the layer adjacent to the crystal must be complex in nature. A layer next to the surface of a growing crystal in which incipient crystallisation is taking place has been referred to as the Powers fluidized layer[42,43]. It has been suggested that this layer could either be scraped off[44-46] or removed by fluid shear[47-51] and so be a source of crystal proliferation[22]. Bilgram and Steininger[52] have suggested that their light scattering experiments during the crystal growth process can only be explained by the existence of a boundary layer. Qian and Botsaris[53] have proposed one of the most detailed accounts of the process of crystal growth involving the attraction and aggregation of colloidal, sub-nucleating particles to the surface. The strongest evidence for feedback between crystal surface and the immediately adjoining layer of solution derives from the work of Flood[54]. Transferring crystals from higher to lower supersaturation immediately results in a reduction of the growth rate to zero. This indicates that either the surface itself or the adjacent layer to it, or both, have a memory imprint of the previous growth conditions. What is missing from many of these accounts is measurement over a range of supersaturation and therefore of the association of particular growth mechanisms with the phenomena discussed.

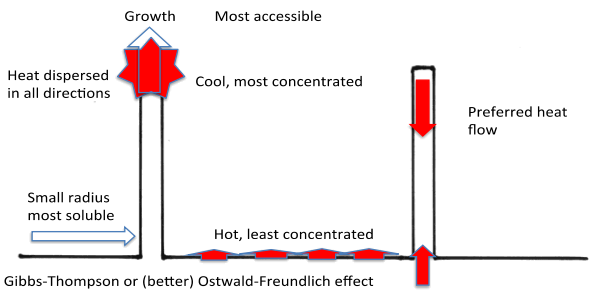
Dendrite Formation

One phenomenon which is clearly associated with higher supersaturations is dendrite formation. The presence of dendrites in particular has been noted especially on magnesium sulfate. In magnesium sulfate solutions proliferation occurred under flow conditions when the supersaturation exceeded 2.3°C [49]. Visible dendrites appear on magnesium sulfate heptahydrate crystals at 8°C supersaturation, but surface imperfections ('veiled growth') appear at 4°C supersaturation. The rough growth features are likely to be very much below the limit of visual detection, so begin to occur nearer 2.3°C supersaturation. The growth and detachment of dendrites on crystal surfaces of ammonium chloride[55,56] has been recorded microphotographically∗, including the 'dendrite coarsening' step[57]. In this the dendrite bases narrow although the tips are still extending. The dendrites eventually snap off as a result. It is the release of dendrites into the solution which offers the possibility of proliferation. The reason for this behaviour would appear to be differential supersaturation at the base and tip of the dendrites, due both to concentration and the temperature effects, as shown in Figure 2. Heat transfer is generally considered only in connection with melt crystallisation. It is nearly always assumed that in solution crystallisation only mass transfer is important. Whilst this is true at a macroscopic level, it is not so at the microscopic domains under consideration. Recently the extent of heat transfer into the liquid phase during crystallisation has been assessed[58]. Heat is generated and concentration lost at the crystal surface during growth. The solution is therefore less saturated in the vicinity of the surface and the diameter of the dendrites small, leading to dissolution dependent on the Gibbs-Thompson[29] or Ostwald-Freundlich[59] effect (smaller particles are more soluble than larger ones). At the tip the concentration is higher and the heat can escape more readily: this is also the reason for the formation of the dendrites in the first instance. The slightest irregularity of the surface or protrusion on the surface gives rise to an immediate acceleration of growth due to improved

∗Astonishingly similar cinephotomicrographs of dendrite growth and detachment in a very different system, the solidification of liquid crystal surfaces, have been published[56]

opportunity of heat loss at the tip, see Figure 2. This is also the mechanism by which needle growth is encouraged during rapid crystallisation. Crystallites generated in a boundary layer might attach also to the crystal and be the source of irregularities for the growth of dendrites. This shows that dendrite growth and a boundary layer are not mutually incompatible concepts.

Figure 2. Factors influencing the thinning of the base of dendrites, leading to their detachment.[§]



[§]The heat released by crystallisation is added to the heat travelling along the dendrite. The solution near the base of the dendrite is less concentrated, and the radius of the dendrite makes it effectively more soluble than bulk crystal.

Insufficient is known about the mechanisms of rough growth to be sure of the detailed processes by which crystal proliferation can be engendered by rapid growth of a seed crystal, but the mechanisms suggested above, which are merely a condensate of the ideas scattered through the literature, seem rational ones. These would seem enough to support the idea that rough growth can support the dispersion of nuclei.

One consequence of the effect of heat of crystallisation on growth characteristics is that the secondary nucleation threshold is likely to be closer to the solubility curve for substances with high solubility and large heats of crystallisation. This is likely to generate higher local surface temperatures, for that will aid dendrite detachment, if indeed that is the defining mechanism.

The Temperature-Concentration Diagram: Analysis of Figure 1

The undersaturated region is labelled A. The area marked B is the dead zone, considered in reference 32 and earlier in this discussion under 'Secondary Nucleation Threshold'. The area C is the one of central interest to this present account. It is the region where crystal growth on existing seeds can occur, but proliferation will not occur. Area D is the proliferation zone in which growth on, and proliferation of, existing seeds will occur, but spontaneous nucleation will not. Area E is where spontaneous crystallisation can occur. The metastable zone encompassing areas B, C and D is generally considered as the area within which spontaneous nucleation cannot occur. There is one type of 'nucleation' which does take place within the metastable zone, which was first pointed out[6] in connection with Miers and Isaacs observations[26]. This is primary heterogeneous nucleation. It is perhaps more suitable that this should not be regarded as nucleation at all, but primary heterogeneous crystallisation. The nuclei are already there, albeit as foreign particles. Primary heterogeneous nucleation is thus essentially crystal growth, although it seems to be characterised by a great curiosity, namely that it can dramatically slow or possibly cease soon after initiation. It has been suggested that the effectiveness of the seeds bringing about primary heterogeneous nucleation is due not to their surface characteristics or chemical nature, but to their radius*[1], which would certainly explain this.[60]

The use of the terms 'seeded' and 'unseeded' metastable limits shows the distinction between the metastable zone limit and the secondary nucleation threshold. However, in referring to 'seeded' metastable limits confusion is generated, because the subsequent proliferation is not spontaneous nucleation, since it is not within the spontaneous nucleation area. We also avoid the use of the term seeded metastable limit because it implies deliberate seeding, whereas as can be seen, the term is equally applicable to the consequences of the primary 'seeding' or accidental seeding scenarios. With narrow metastable widths commonly encountered with inorganic salts and small organic molecules such as constitute the solutes in the vast majority of literature MZW investigations, it will be really difficult to separate the secondary nucleation threshold from the metastable zone limit. So wider MZWs are desirable for clear identification of the secondary nucleation threshold.

The weak relationship between molecular complexity, particularly conformational multiplicity and metastable zone widths,[32] indicates that many of the molecules now appearing from the pharmaceutical industry may have wide metastable zone widths. An example is famotidine polymorph A with a metastable zone width of 44 to 15°C, dependent on solvent[61]. That famotidine Polymorph B has a much narrower MZW (3°C) shows how little is known about the factors governing the MZW. There are a few small molecules with wider MZWs than those commonly reported, such as glutamic acid[62] with a MZW of 48°C or ascorbic acid[63] with a MZW of 40 to 18°C. These both have the features of competing hydrogen bonding groups and conformational choice. These are also the characterising features of the sugars with traditionally wide MZWs, although a further effect is due to the presence of anomers[14] and viscosity. We set out to identify possible examples of wide MZWs on this premise. A tabulation of some of the wider metastable zone widths is presented in Table 1.

It can be seen that the secondary nucleation threshold can occur anywhere within the metastable zone. Ibuprofen, which has a narrow MZW has been added to this table because the secondary nucleation threshold in this case lies very near the

*[1] An observation supporting this view was encountered during a crystallisation study on 4-chloropropionanilide[32]. A fragment of Teflon tape was removed from the solution, whereupon the crystallisation point immediately decreased by 3°C. It is difficult to think of any organic material more inert than polyfluoroethylene, so the seeding effect must have been connected with other characteristics of the particle.

solubility curve. The width of the MZW appears to be independent of the position of the secondary nucleation threshold. In all the published cases the course of the secondary nucleation threshold is approximately parallel to that of the metastable zone limit. This has been confirmed by our own observations on a variety of compounds.

To the list of factors contributing to the width of the metastable zone, conveniently summarised in reference 32, a further one was discovered during the course of this work. When the gap between the end of the fibre optic and the mirror was decreased the sensitivity of the probe was likewise reduced and the metastable zone width changed significantly. It had been expected that halving the gap would halve the sensitivity, but the resulting charts suggested that the sensitivity may have decreased by a factor of a hundred-fold. Consequently the metastable width was much enlarged, even at the slowest temperature ramps. The dependence of the metastable zone width on the sensitivity of instrumental detection appears not to have been discussed in detail in the literature. Dependent on whether the detection method is proportional to the surface area of the particles, or numbers of particles, or volume of particles, different relationships would be expected between increase in detector sensitivity and increase in MZW. Indeed, detection may involve mixed relationships, although these are likely to be linear at the initial detect point.

The Latent Period

Although area C has been called the growth area it can be recognised from the previous paper[1] that it is also forms the start of the 'Latent Period' for unseeded crystallisation. This was probably first observed by Ayerst and Phillips[64] but designated originally by Mullin[29], and seen since by many observers, although not always recognised. This phenomenon arises because initial crystal growth on the particles always present in normal solvents brings about the primary crystallisation event ('primary heterogeneous nucleation'). The crystals or particles so formed do not proliferate until the metastable zone limit is reached, or sufficient time has elapsed for a statistically unlikely event to occur[65]. This sequence will therefore be more prominent in aqueous organic solvents, because the organic solvent precipitates out the traces of inorganic salts dissolved in the aqueous medium which then act as additional heterogeneous nuclei. As suggested earlier, it is their shape which is critical to their effectiveness as nuclei[60]. Crystals generated by growth on heterogeneous nuclei are not identical in their behaviour to added seed crystals[1,32]. Hence the 'growth' region and the Latent Period are not identical, but they are clearly subject to the same conditions and must be closely related. It may be objected that it is generally understood and stated in the textbooks primary nucleation requires a higher supersaturation than secondary nucleation[66]. Presumably this is because of the expectation that seeds of the same substance would grow more readily than heterogeneous seeds and also of the observation that the metastable zone limit occurs at higher supersaturations than the SNT. This shows the danger of equating 'nucleation' with 'first easily observable crystallisation' or 'crystal proliferation'. When crystallisation was observed visually the distinction was more easily recognisable. The universal use today of automatic instrumentation allows the first signs of nucleation to escape notice, because it represents such a small change of baseline.

In the case of the crystallisation of a very large number of simple inorganic salts and organic compounds examined, to be reported later, in aqueous or aqueous organic solutions under isothermal or nearly isothermal conditions an initial turbidity was followed, often many hours later, by a sudden crystal proliferation. This effect has already been reported for propionanilide[32]. The most probable explanation for this behaviour is that, as described above, the initial crystallisation observed is due to primary heterogeneous nucleation, catalysed by the appropriate radii of the heterogeneous particles. As soon as crystal growth begins, the radii are enlarged so the growth propensity of the particles is reduced. Since they are in the growth only region, they are also incapable of proliferation.

In another set of experiments conducted by the authors, also not reported here, on Design of Experiment investigations into the seeding of magnesium sulfate solutions, in those cases where only 3 or 4 seeds were added growth of the magnesium sulfate seeds preceding the proliferation could be clearly seen. This observable behaviour would seem to be closely aligned to the Latent Period case, suggesting that the seeded and non-seeded crystallisation phenomena are essentially the same. The question then arises as to why the metastable zone limit and the SNT are not coincidental. This must be a feature of the vast crystal size difference of heterogeneous nuclei and added seeds. It is extremely difficult to grind seeds below a few micrometres[67]. By contrast the particles in well-filtered solutions are likely to be in the nanometre region. A final comment concerns the range of the Latent Period. A certain degree of supersaturation is clearly essential for the initial nucleation/growth, so a line has been drawn in Figure 3 representing this. The crystal proliferation point must by definition be the heterogeneous metastable zone limit.

Figure 3. The generalised solution diagram showing the effect of choice of temperature on the attempt to grow crystals and remove fines by annealing across the secondary nucleation threshold. Mullin's latent period is also included.

The course of the crystallisation

The single nucleation event has been advocated and shown in some cases to determine the outcome of crystallisation[65]. In practical seeded situations, however, there are going to be many simultaneous nucleation events and hundreds of consecutive nucleation events. The result of this is that it is impossible to tell from the mere course of the crystallisation as to the details of the seed proliferation. In each case the trajectory of the crystallisation, e,g. the turbidimetric–value, will appear to slope upwards at an increasing incline, but whether this is due in any particular case to several overlapping primary crystallisation events followed by varying intervals to the secondary nucleation, or to single or simultaneous nucleation events followed by crystal growth and continuously increasing proliferation, is impossible to determine merely by the shape. So careful analysis of the crystallisation curves of the kind carried out by Dugua and Simon will be necessary to establish the true course of the crystallisation events[68]. An explanation of such nucleation events has been suggested by Qian and Botsaris.[53] This is yet another manifestation of the complexity of crystallisation which remains insufficiently explored.

Large Scale Considerations

There has been no lack of papers on the control of crystallisation conducted at all scales. FDA guidelines have encouraged the use of PAT techniques to monitor and control the progress of crystallisation. A plethora of on-line instrumentation is now available to enable this to be accomplished. However, all this relies on an empirical control of the process. This can involve crossing the solubility curve to induce dissolution of fines or the metastable limit to induce crystallisation[69], or applying 'first-principles' or 'direct design' to the control of crystallisation[70]. In a sense, knowledge of the SNT underpins all of these approaches: it shows that one need not cool as far as the metastable line to induce crystallisation, it is a further first principle and it can guide the area of direct design cooling trajectory. The cooling regimes advocated appear to straddle the SNT line, so complex and accurate computer control[68] for this approach becomes essential.

Undoubtedly the concept of a boundary separating growth and proliferation during solution crystallisation has been encountered, and even commented upon, by process chemists. But it seems never to have been discussed as a coherent concept as a fundamental feature of the crystallisation process. For example Kwokal refers to the need to avoid the 'uncontrolled secondary nucleation zone'[71] and Saleemi *et al*. show automatic direct nucleation control confined to what is here referred to as the growth only region, in order to produce large crystals[72].

Since the concept of the SNT is already used in commercial scale crystallisation, it might be thought that further words on the topic were superfluous. However the narrow applications, the specialized terminology, the less accessible literature and some unique features of carbohydrate crystallisation behaviour mean that the dairy and sugar applications cannot easily be generalized to other areas. Although the experimental work supporting this paper (both presented herein and yet to be reported) is all laboratory based and the discussion has been from the viewpoint of small-scale crystallisation. The thesis being advanced here is that the phenomena are fundamental, universal features of the crystallisation process, so this work clearly has relevance to scaling up. Inert seeds will tend to give a fines content, although this is likely to form but a small proportion of the product mass, which will be largely determined by those seeds which do proliferate. More importantly, the time that a seeded crystallising solution spends in the region where only crystal growth can occur will be an important factor in the ultimate product particle size, dependent both on the increase in size of the seeds during that time, and on the absence of further seeds. The crystals will grow more rapidly nearer to the proliferation boundary, which is desirable for process efficiency, so it is important to know where that boundary lies. It can readily be imagined, and the few examples in Table 1 confirm, that it will be different for every substance, although in the absence of previous recognition of its existence, just how different is indeterminate. The removal of fines and the improvement of crystalline products by temperature cycling of the first-obtained crystal mass is now commonplace, but the recognition of the proliferation boundary now gives an opportunity to control the particle size distribution at an earlier stage. For the solution can be temperature cycled over the proliferation boundary to control the ratio of numbers and size of product particles as an additional weapon in the process chemist's armoury. The main contribution to secondary nucleation is of course collision breeding. But even in this circumstance the position of the cooling solution within the metastable zone will be important, because collision breeding will be minimally effective in spawning vast number of nuclei within the growth only zone. Most controlled crystallisation is carried out with seeding, but the effects of seeding will be restricted in the growth only zone. Seeds will tend to grow rather than proliferate. The consequence of this on product outcome will be that if the solution is seeded on the low supersaturation side of the crystal proliferation region, and the solution remains there for too long, the seeds will grow large before proliferating as the solution cools. This will probably result in a bimodal distribution. If seeding is carried out on the high supersaturation side, then many fines are likely. More advantageous cooling/seeding regimes suggest themselves. A possible solution is to seed and cool to within the proliferation region, then to take the solution back into the growth-only region to allow the particles to grow as the solution is cooled to the required particle size. Clearly, the SNT plays a critical role in the ultimate crystal size distribution (CSD). The ability to carry out a crystallisation in the manner of commercial lactose or sucrose crystallisation on a large scale by cooling and heating to take advantage of the secondary nucleation threshold to produce a desired CSD will depend upon how rapidly the nucleation and crystal growth steps occur. There is a dearth of information in the literature as to the generality of the behaviour of modern pharmaceutical molecules in this regard, although as implied here and previously[21], the complexity of these molecules may encourage wide metastable zones and

slow incorporation into the lattice from solution. If these pre-requisites are not forthcoming, then it may be necessary to cool a sub-batch and mix into the larger remaining warmer solution at the point at which sufficient nuclei have formed. If, despite one's best efforts, it is still not possible to take full advantage of knowledge of the SNT, then one is still better off understanding it than having to work blind.

The lactose crystallisation mentioned earlier confirms the importance of the proliferation region in the commercial crystallisation practice for the production of fine particles. In pharmaceutical and fine chemicals crystallisation, uniform, well-shaped crystals are commonly produced which shows what can be achieved empirically, or with computer control, for generally this must have been achieved without the knowledge of the underlying reasons described here. What is being advocated in this paper is the desirability of measuring both the seeded and the unseeded determination of the crystallisation, in order to have a complete understanding of basis of the solution crystallisation behaviour.

On a smaller scale, we have used temperature cycling to grow larger crystals suitable for single crystal X-ray diffraction measurements. Sometimes this has worked, the cycling producing fewer but larger crystals. Often, however, no increase in crystal size results. Heating the solution dissolves the smaller crystals, but on the cooling cycle as many small crystals reappear as had dissolved. The reason for this is probably that such cycling is between points A and D of Figure 1, whereas cycling between A and C was needed. This again shows the need to ascertain the position of the SNT for good annealing at any scale.

Final remarks

The considerations in this paper have avoided the matter of polymorph crystallisation. Clearly each polymorph will have its own solubility, SNT and metastable zone width, so if polymorphism is an issue the crystallisation behaviour of each polymorph will need to be separately investigated. It is however unlikely that the (unseeded) metastable zone limit of the less easily crystallised polymorph will be observable. The benefit of this extra work will be that the propensity for both nucleation and for proliferation of the undesired polymorph will then be better understood. If the SNT of both of the polymorphs lies at similar supersaturation values in a given solvent, then the occurrence of concomitant polymorphs is extremely likely[73].

It has long been known in crystallisation practice that greater control over crystallisation temperatures and cooling rates can improve the crystal product quality dramatically. It is intended that the present analysis provides some understanding of why that may be so, as well as providing a framework and goal within which controlled crystallisation can be better designed. It is surprising that the concepts discussed in this paper particularly that of the SNT which has been used for a century for the understanding of the crystallisation of the world's largest volume organic compound, should have had so little exposure in the general crystallisation literature.

Conclusions

1) The metastable zone can be divided into a 'growth only' region and a crystal proliferation region, separated by the SNT.
2) The division of the metastable zone between crystal growth and crystal proliferation regions is most probably related to a change of crystal growth mechanism.
3) The interval between primary crystal formation ('primary heterogeneous nucleation') and the SNT is the 'Latent Period' described by Mullin and is closely related to the growth area of the metastable zone.
4) Appreciation of the SNT is important for the understanding of crystal size distribution.
5) The increasing complexity of modern pharmaceutical molecules will probably lead to wider metastable zone widths and render the SNT of increasing visibility and importance.

Notes and references

[a] *Chemistry, University of Southampton, Southampton, UK, SO17 1BJ. E-mail: t.threlfall@soton.ac.uk*

1. S.J. Coles and T.L Threlfall, *CrystEngComm*, 2014, **16**, 4355-4364
2. O.F. Hunziker, *Condensed Milk and Milk Powder* 4th Ed., Hunziker, La Grange, Illinois, USA **1926**
3. T.A. Nickerson in *Fundamentals of Dairy Chemistry* Ed. B.H.J. Webb, A.H. Johnson and J.A. Alford, 2nd Ed., AVI publishing, Westport, Conn, USA, 1974
4. S.Y. Wong and R.W. Hartel, *J. Food Sci*., 2014, **79**, R257-R272
5. J.C.P. Chen and C.C. Chang, *Cane Sugar Handbook*, Wiley, N.Y. 1993
6. H. Ting and W.L. McCabe, *Ind. Eng. Chem*., 1934, **26**, 1201-1207
7. O. Soehnel, *Chem. Eng. Res. Des*., 1983, **61**, 186-190
8. H-H. Tung, *Org. Process Res. Dev.*, 2013, **17**, 445-454
9. H-H. Tung, E.L. Paul, M. Midler, J.A. McCauley, *Crystallization of Organic Compounds*, Wiley, New York, 2009
10. K. Sangwal, *CrystEngComm*, 2011, **13**, 489-501
11. J. Kraus and J. Nyvlt, *Zuckerindustrie (Berlin)*, 1994, **119**, 219-222; J. Nyvlt and J. Kraus, *idem*, 407-413
12. C. Hongo, S. Yamata and I. Chibata, *Bull. Soc. Chem. Japan*, 1981, **54**, 1905-1910
13. R.F. Strickland-Constable, *Kinetics and Mechanism of Crystallization*, Academic Press, 1968
14. S.J. Srisa-Nga, A.E. Flood and E.T. White, *Cryst. Growth Des.*, 2006, **6**, 795-801
15. L. Wantha and A.E. Flood, *J. Cryst. Growth*, 2011, **318**, 177-121
16. A. Rashid, E.T. White, T. Howes, L. Liu, J.D. Litster and I. Marziano, www.conference.net.acc/chemeca2011/papers/107.pdf : A. Rashid, E.T. White, T. Howes, J.D. Litster, I. Marziano, *Chem Eng. Res. Des*., 2012, , 158-161
17. S. Tait, *Secondary nucleation characteristics of protein crystallisation*, Ph. D. Thesis, University of Queensland, 2007.
18. I. Khaddour and F. Rocha, *Cryst. Res. Technol*., 2011, **46**, 373-382
19. F. Puel, E. Verdurand, P. Taulelle, C. Bebon, D. Colson, J-P. Klein and S. Veesler, *J. Cryst. Growth*, 2008, **310**, 110-115
20. M. Oullion, F. Puel, G. Fevotte, S. Righini and P. Carvin, *Chem. Eng. Sci*. **2007**, *62*, 820-832, 833-845
21. E. Verdurand, C. Bebon, D. Colson, J-P. Klein, A.F. Blandin and J-M. Bossoutrot, *J. Cryst. Growth*, 2005, **275**, e1363-1367
22. S.G. Agrawal and A.J.H. Paterson, *Chem. Eng. Commun*., 2015, **202**, 698-706
23. Y. Shi, R.W. Hartel and B. Liang, *J. Dairy Sci*., 1989, **72**, 2906-2915
24. B.K. Butler, *Modelling Industrial Lactose Crystallization*, Ph.D. Thesis, University of Queensland, Brisbane, 1998

25. S.Y. Wong, R.K. Bund, R.K. Connelly and R.W. Hartel, *Int. Dairy J*., 2011, **21**, 839-847.
26. H.A. Miers, H.A. and F. Isaacs, *J. Chem. Soc. Trans*., 1906, **89**, 413-454
27. A. Holden and P. Morrison, *Crystals and Crystal Growing, MIT Press*, 1982
28. R.J. Davey and J. Garside, *From Molecules to Crystallizers*, Oxford, 2000.
29. J.W. Mullin, *Crystallization* 4th ed., Butterworth-Heinemann, Oxford, 2001
30. A.E. Flood and L. Wantha, *J. Cryst. Growth*, 2013, **373**, 7-12
31. S.K. Heffels, E.J. De Jong and D.J. Sinke, *Zuckerindustrie (Berlin)*, 1987, **112**, 511-518
32. T.L. Threlfall, R.W. De'ath and S.J. Coles, *Org. Process Res. Dev*., 2013, **17**, 578-584
33. T.L. Threlfall, *Analyst*, 1995, **120**, 2435-2460
34. D.A. Malivuk, A.A. Zekic, M.M. Mitrovic and B.M. Misailovic, B.M., *J. Cryst. Growth*, 2013**, 277,** 164-169
35. A.D. Randolph and M.A. Larson, *Theory of Particulate Processes*, 2nd Ed., Academic Press, New York, 1988
36. Yang, H., *CrystEngComm*, 2015, **17**, 577-586
37. M. Ohara and R.C. Reid, *Modeling Crystal Growth Rates from Solution*, Prentice-Hall, 1973
38. I. Sunagawa, *Crystals: Growth, Morphology and Perfection*, Cambridge University Press, 2005
39. Anwar, J. Faraday Disc. 2015, 179. At p.169
40. H.A. Miers, *Phil. Trans. Royal Soc*.,1903, **202**, 459-523
41. W.F. Berg, *Proc. Royal Soc. A*, 1938, **164**, 79-95
42. H.E.C. Powers, *Ind. Chem. Chem. Manuf*., 1963, **39**, 351-355
43. N.A. Clontz and W.L. McCabe, *Chem. Eng. Prog. Symp. Ser.*, 1971, **110 (67),** 6-17
44. G.R. Youngquist and A.R. Randolph, *AIChE J*., 1972, **18**, 421-429
45. D.P. Lal, R.E.A. Mason and R.F. Constable-Strickland, *Trans. Farad. Soc*., 1966, **62**, 455-461
46. N. Kubota and K. Kubota, *J. Cryst. Growth*, 1986, **76**, 69-74
47. T.P. Melia and W.P. Moffitt, *J. Colloid Sci*., 1964, **19**, 433-447; *Ind. Eng. Chem. Fundam*., 1964, **3**, 313-317
48. U. Fasoli and R. Conti, in *Industrial Crystallization*, Ed. J.W. Mullin, Plenum, New York, 1976
49. J. Wang and J. Estrin, *Chem. Eng. Commun*., 1996, **1**, 152-153
50. C.Y. Sung, J. Estrin and G.R. Youngquist, G.R. *AIChE J*., 1973, **19**, 957-962
51. M.M. Reyhani, S. Freij and G.M. Parkinson, *J. Cryst. Growth*, 1999, **198**/**199**, 258-263
52. J.H. Bilgram and R. Steininger, *J. Cryst. Growth*, 1990, **99**, 30-37
53. R-Y. Qian and G.D. Botsaris, *Chem. Eng. Sci*., 1997, **52**, 3429-3440
54. A.E. Flood, *CrystEngComm*, 2010, **12**, 313-323
55. M.O. Kliya, *Krystallografya*, 1956, **1**, 737-742, translation: *Soviet Phys-Cryst*. 1956, **1**, 456-458; D.D. Saratovkin, *Dendritic Crystallization*, Consultants Bureau, New York, 1959
56. J-M. Flesselles, A.J. Simon and A.J. Libchaber, *Adv. Phys.*, 1991, **40**, 1-51
57. E.G. Denk and G.D. Botsaris, *J. Cryst. Growth* 1972, **13**/**14**, 493-499
58. M. Kapembwa, M. Rodriguez-Pascual and A.E. Lewis, *Cryst. Growth Des*., 2014**, 14**, 389-395
59. See Wikipedia articles on the Gibbs-Thompson/Ostwald-Freundlich effect for the complexity of the correct historical assignment of these phenomena.
60. A. Caccuito, S. Auer, S. and D. Frenkel, *Nature*, 2004, **428**, 404-406
61. J. Lu, X-J. Wang, X. Yang and B. Ching, *Cryst. Growth Des*., 2007, **7**, 1590-1598
62. K. Liang, G. White, D. Wilkinson, L.J. Ford, K.J. Roberts and W.M.L. Wood, *Ind. Eng. Chem. Res*., 2004, **43**, 1227-1234
63. B. Wierzbowski, A. Matynia, K. Piotrowski, K and J. Koralewski, *Chem. Eng. Progr*., 2007, **47**, 351-359
64. R.P. Ayerst and M.I. Phillips, *Industrial Crystallisation*, Ed. J.W. Mullin, 1969, 56-68
65. S.S. Kadam, H.J.M. Kramer and J.H. ter Horst, *Cryst. Growth Des*., 2011, **11**, 1271-1277
66. A. Meersmann, C. Heyer and A. Eble, '*Active Nucleation'* in *Crystal Technology Handbook*, Ed. A. Meersmann, 2nd Ed. 2001
67. R. Roberts and R. Rowe, *J. Pharm. Pharmac.*, 2000, **52**, 147-150
68. J. Dugua and B. Simon, *J. Cryst. Growth*, 1978, **44**, 280-286
69. M.R.A. Bakar, Z.K. Nagy, A.N. Saleemi and C.D. Rielly, *Cryst. Growth Des.*, 2009, **9**, 1378-1384
70. M. Fujiwara, Z.K. Nagy, J.W. Chew and R.D. Braatz, *J. Process Control*, 2005, **15**, 493-504
71. A. Kwokal, Course notes, Engineering Crystallography, *Erice International School of Crystallography*, 5-14 June 2015, **48**th course, 237-247
72. A.N. Saleemi, G. Steele, N.I. Pedge, A Freeman and Z.K. Nagy, *Int. J. Pharm*., 2012, **430**, 56-64
73. J. Bernstein, R.J. Davey and J-O. Henck, *Angew. Chem. Int. Ed. Eng*. 1999, **38**, 3440-3461