

Microfluidics and Modeling of Nucleation Rates in Cocrystal Systems

Submitted in partial fulfilment of the requirements for the

degree of

Doctor of Philosophy in

Department of Chemical Engineering

Scott N Pedu

B.S. Chemical Engineering, Rochester Institute of Technology

Carnegie Mellon University

Pittsburgh, PA

January 2022

©Scott N. Pedu, 2021

All Rights Reserved

Acknowledgements

I would like to thank my advisor Shelley Anna, this work could not have been completed without your guidance and support. You have helped me grow as independent scientist and professional.

Thank you to my PhD Thesis Committee – Professors Stephen Garoff, John Kitchin, and Ryan Sullivan, for their support and knowledge over the past few years.

Thank you to my colleagues in the Anna Lab – Charles, Tom, Yanxin, Dylan, Marty, Chijing, Yathu, and Corine. Each of you helped me in your own way, and together we made it through many interesting and challenging years.

Thank you to our Pandemic Safety Officers – Julie Tilton, your tireless efforts ensured we all were kept safe during the ongoing COVID-19 pandemic.

Thank you to my family, without your love and support I could not have even begun this journey.

This thesis research was supported by Department of Chemical Engineering funds and the James C. Meade Fellowship in Chemical Engineering.

Abstract

Understanding the nucleation process is key to controlling the formation of crystalline materials. Previous studies have demonstrated the potential for static arrays of nanoliter-sized drops to act as miniature reactors that probe the nucleation rate for single component crystals. However, there is little work investigating the nucleation of co-crystal systems. This thesis expands upon drop-based microfluidic strategies to develop a platform that allows for the continuous generation of drops and observe nucleation of crystals within them. The continuous production of drops allows for better scalability, by increasing the number of observable drops without increasing the device size. This experimental approach will eventually allow the nucleation rate to be determined as a function of experimental conditions, including the ratio of co-formers, the coformer concentrations, and the temperature. In addition to the experimental platform, this thesis also contributes to a better understanding of how co-crystal nucleation rates depend on experimental conditions by developing a theoretical model for cocrystal nucleation that is built upon the foundational assumptions of Classical Nucleation Theory. Expanding our understanding of co-crystal nucleation facilitates the construction of time dependent phase diagrams showing how the crystal mixture composition evolves as a result of the nucleation rates. Future work may include using the microfluidic platform to identify new co-crystallizing systems or to determine the polymorphs of previously unexplored crystals.

Table of Contents

Chapter 1 : Introduction	1
References	6
Chapter 2 Crystallizing Lysozyme in Continuous Drop Trains	7
Abstract	7
Introduction	7
Materials and Methods	19
<i>Static Array Temperature Step Nucleation</i>	19
<i>Static Array Dehydration Induced Nucleation</i>	23
<i>Continuous Droplet Train Nucleation</i>	24
Classical Nucleation Theory	30
Results	35
<i>Crystallization in Static Arrays</i>	35
Nucleation induced by dehydration in static arrays	44
Continuous Droplet Trains	51
Discussion	63
Conclusions.....	70
References	73
Chapter 3 : Theory of cocrystal nucleation from solution	78
Abstract	78
Introduction	78
Background	83
Model Details.....	84
Classical Cocrystal Nucleation Theory Model.....	86
Initial Complex Cocrystal Nucleation Theory model	92
Results and Discussion	94
Dimensional Analysis	111
Conclusions.....	120
References	122
Chapter 4 : Kinetic pathways to co-crystal nucleation	126
Abstract	126

Introduction	126
Model details	132
Results and Discussion	135
First Nucleating Species Diagrams	135
Modeling the role of concentration change.....	144
Conclusions.....	157
References	159
Chapter 5 : Summary	161
Conclusions.....	161
Future work.....	162

List of Tables

Table 2-1 Experimental Conditions and Solution Compositions	29
Table 2-2 Static array nucleation model parameters. Taken from Akella et al. [21].	41
Table 2-3 Drop volume distribution	52
Table 2-4 Time Resolution Parameters.....	58
Table 2-5 Extracted kinetic fitting parameters from lysozyme nucleation in continuous drop trains.....	63
Table 2-6 Tradeoffs between static and continuous droplet devices for examining crystal nucleation from solution.	70
Table 3-1 Example co-crystal systems, identification techniques, and interaction mechanisms.	81
Table 3-2 Relative rates at different volumes.....	119
Table 4-1 Coformer physical properties	133
Table 4-2 Solvent physical properties	134
Table 4-3 Nucleation and complex formation kinetic parameters	134

List of Figures

Figure 2-1 Droplet observation within microfluidic device. Optically transparent PDMS allows for the direct visual observation of droplets as they flow through the serpentine residence chamber. Constant infusion rates from the syringe pumps driving the flow allow for a direct mapping of droplet position to residence time within the device. 25

Figure 2-2 Schematic diagram of the experimental setup for a continuous crystallization device. Four syringe pumps infuse the three aqueous streams and the oil stream into the microfluidic device. The microfluidic device is placed on top of a transparent moveable stage between the illuminator and sCMOS camera attached to the computer used for data collection. The transparent cooling jacket rests on top of the device and is connected to the recirculating chiller. The crystallized drops and oil are collected in a waste beaker after exiting the device. ... 26

Figure 2-3 Classical Nucleation Theory. Solutes driven by Brownian motion occasionally collide with enough energy to overcome repulsive intermolecular forces. At a critical radius the energy required to create the new interface between the solid and liquid phases is balanced by the reduction in free energy due to adding volume to the solid crystal. Exceeding this radius is referred to as a nucleation event because all subsequent additions to the nucleus reduce the free energy of the system and therefore occur spontaneously. 30

Figure 2-4 Free Energy Change versus Nucleus Radius. The change in free energy of the system, ΔG , is plotted against the radius of the growing nucleus. When the radius exceeds the critical radius r^* the slope of ΔG becomes negative and the process of adding more molecules to the nucleus becomes spontaneous. When the nucleus exceeds this critical radius, we call it a nucleation event. 31

Figure 2-5 Droplet trap in a static array microfluidic design. Schematic diagram of a single microfluidic droplet trap along with relevant device dimensions. The channel depth is fabricated to be $100\mu\text{m}$ 35

Figure 2-6 Drop Generation in static arrays. Left: The aqueous solution is infused into the device, filling all the traps and bypass channels. Right: Oil is infused to displace the solution from the bypass channels, creating drops in the traps. 36

Figure 2-7 Dehydration of microfluidic droplets. Droplets in a static array microfluidic device are exposed to a low relative humidity through a thin layer of PDMS. Water diffuses out of the drops and evaporates into the surrounding air over time. As the droplets dehydrate the supersaturation increases until nucleation occurs. 37

Figure 2-8 Crystals formed in a temperature step static array experiment. The image shows a set of droplets trapped in a microfluidic static array device after temperature induced nucleation has occurred. The resulting crystals are visible within the droplets. 38

Figure 2-9 Crystals formed in a static array as a result of droplet dehydration. The image shows the entire observed grid of static droplets after partial crystallization has occurred due to dehydration of the droplets. The droplet color changes dramatically when nucleation occurs. The drops become significantly more supersaturated, resulting in rapid and complete solidification of the droplet contents. 39

Figure 2-10 Probability that crystallization has not yet occurred versus time for temperature step crystallization. Solid lines indicate the model predictions for the fraction of uncrystallized drops; diamond and star symbols denote measured data points for the lower and higher concentrations, respectively. 40

Figure 2-11 Volume variations in static arrays. Plotted is model predictions for both concentrations with the mean volume as before with two additional model predictions with the a volume one standard deviation above and below the mean. 43

Figure 2-12 Crystal Detection in Static Droplet Arrays. A grid of 182 of the total 720 drops is observed over time. As water slowly leaves the drops by diffusing out of droplets and through the PDMS, the solution concentration within droplets rises until crystals form. The probability function is determined by counting the fraction of drops without observable crystals at each sampling time. 44

Figure 2-13 Droplet volume changes versus time. The left y-axis shows the drop volume as a function of time and the right y-axis shows the corresponding supersaturation within the droplet. As volume decreases due to evaporation of water, the concentration within the droplets increases, raising the supersaturation. 45

Figure 2-14 Probability function versus time for dehydration-induced crystallization. Solid lines indicate model predictions for the fraction of uncrystallized drops; diamond and star symbols denote measured data for the higher and lower concentrations, respectively. 47

Figure 2-15 Dehydration variability in static arrays. Plotted is model predictions for the 30 mg/mL concentration trial with the mean dehydration rate as before with two additional model predictions with the a dehydration rate one standard deviation above and below the mean. 49

Figure 2-16 Continuous drop train microfluidic device design. The microfluidic device consists of two major sections. A) the T-junction drop generation nozzle combines the three aqueous streams as they meet the oil stream where the combined stream is pinched off into individual drops. B) The residence track retains the drops while exposing them to a fixed temperature below the saturation temperature to induce nucleation. The constant cross-sectional area and constant flow rates imply that droplet position and residence time are directly proportional to one another as seen in equation (14)..... 51

Figure 2-17 Drop Volume distributions. Each graph shows the number of drops within a particular volume bin for each of the three solution compositions tested. Notable qualitative differences can be seen between each population. Additionally long tails to the distribution show the presence of many drops well above the expected mean volume 53

Figure 2-18 Crystal detection scheme. A) First the entire frame is subdivided to isolate an individual channel. B) then large background elements are identified and removed. C) Droplet endcaps are identified via edge detection, isolated and paired up to determine each drop position. D) small objects are detected to locate crystals that have formed. E) If the detected small elements are within the drop boundaries the drop is labeled as having nucleated, as indicated by the green box. Otherwise, the drop is counted as part of the total number of droplets. 55

Figure 2-19 Observed probability function versus time for 30 mg/mL Lys 2 w/v% NaCl. The observed probability function is plotted versus time for a solution containing 30 mg/mL lysozyme and 2 w/v% NaCl. The black line represents a least squares linear fit to the log of the average probability from all the runs versus time, assuming the data follows a single exponential decay and the drop volume equals the average volume. The corresponding nucleation parameters were obtained from the fitted slope, and the blue and red lines were then constructed using volumes equal to one standard deviation above and below the average volume recorded over all three runs..... 59

Figure 2-20 Observed probability function versus time for 30 mg/mL Lys 3w/v% NaCl. The observed probability function is plotted versus time for a solution containing 30 mg/mL lysozyme and 3 w/v% NaCl. The black, blue and red lines were constructed as described for Figure 2-17. 60

Figure 2-21 Observed probability function versus time for 20 mg/mL Lys 2w/v% NaCl. The observed probability is plotted versus time for a solution containing 20 mg/mL lysozyme and 2 w/v% NaCl. The black, blue and red lines were constructed the same way as described for Figure 2-17. 61

Figure 2-22 Combined Continuous crystallization results. The graph combines the recorded data points from Figure 2-19, and Figure 2-20, and Figure 2-21, as well as the best fit lines for those figures..... 62

Figure 3-1 Cartoon depiction of two cocrystal nucleation pathways, one with and one without the formation of initial coformer complexes. The mechanism by which dissolved materials arrange themselves into cocrystals is not widely agreed upon. Two possible routes from solution to cocrystal are shown above. The top pathway shows CNT-like cocrystal nucleation assuming single molecule additions to the precritical nucleus that eventually lead to nucleation and growth. The bottom pathway shows initial complex formation assuming cofomers must form complexes in solution that are then added to the precritical nucleus as a single unit. 85

Figure 3-2 (A) Concentrations and (B) probability functions versus time for Classical Cocrystal Nucleation Theory. For CCNT, the solution concentrations are assumed to be constant with time and the resulting probability distributions are simple exponentials like those found in CNT. The total probability of forming a cocrystal is the product of the probabilities of each independent crystallization pathway. 88

Figure 3-3 Fast solution complexation. (Top) Concentrations for pure components and coformer complexes and (Bottom) Cumulative probability functions versus time for pure component and co-crystals in the case of fast solution complexation, with rate constants $k_f=1$ and $k_r=1$, and all other parameter values held fixed. Rapid formation of coformer complexes causes all solution concentrations to reach their equilibrium values before appreciable nucleation has occurred. Note that the resulting cumulative probability function closely resembles that obtained from the classical cocrystal nucleation model. 99

Figure 3-4 Intermediate coformer complexation. (Top) Concentration and (Bottom) probability function for each component versus time. All parameter values are held constant at values used in earlier cases, except the forward and backward rate constants which are $k_f = 0.1$ $k_r = 0.1$. Substantial nucleation has occurred before the solution concentrations reach their equilibrium values. As the coformer complex concentration increases the cocrystal nucleation rate overtakes the single component nucleation rate, and their probability distributions cross at approximately 35s. After the solution concentrations have reached equilibrium, the probability functions approach single decaying exponential functions. 102

Figure 3-5 Slow coformer complexation. (Top) Concentration and (Bottom) probability versus time for individual cofomers and their complexes. Concentrations do not reach equilibrium before the majority of volumes would contain a crystal. Simple exponential behavior is not observed for the probability function, and cocrystals are never the dominant contributor to the net probability. 104

Figure 3-6 Net probability for crystal formation versus time for fast, intermediate and slow complexation rates. This figure compares the net probability P_0 net for the formation of any crystal (either pure component or cocrystals) corresponding to each of the three relative complexation rates considered in **Figure 3-3** through **Figure 3-5**. 106

Figure 3-7 Fast Cocrystal Nucleation. (Top) Concentration and (Bottom) probability function versus time for a cocrystal nucleation rate ten times faster than that considered in Figures 3-4 through 3-6. Although the slowest coformer complexation rate was used here, the faster cocrystal nucleation rate allows cocrystals to become the dominant species forming after approximately 20s. 107

Figure 3-8 Fast nucleation of component 1. (Top) Concentrations and (Bottom) probabilities versus time for fast nucleation of one of the pure components and intermediate complexation rates. The fast rate of nucleation of Component 1 allows it to continue to dominate the process even after the solution has reached concentration equilibrium. 110

Figure 3-9 Pre-exponential factor ratio versus equilibrium constant. Equilibrium constant is plotted on the x-axis and the ratio of the pre-exponential fitting parameters for cocrystals and component 1 is plotted on the y-axis. The lines indicate the critical ratio of pre-exponential factors that gives $N_r = 1$ for a given k_{eq} and ratio of exponential fitting parameters. A crossover point exists if the system parameters fall into the space above the corresponding critical line. 113

Figure 3-10 Crossover $P_{0\ net}$ versus forward complex formation rate. The $P_{0\ net}$ value at the crossover time is plotted as a function of the formation rate of coformer complexes. Note that the case of $k_{eq}=1.34$ corresponds to $N_{r1} = 1.01$, so it is very close to not having a crossover time and the relatively low probability values led to some instability in the numerical integration. 114

Figure 3-11 Effect of volume on cocrystal probability function. The two plots above use identical parameters except that the top graph considers a solution volume of 1 mL and the bottom graph considers a solution volume of 100 mL. Note that the timescale is dramatically shorter for the 100 mL case due to the increased nucleation rates in the larger volume. This also causes the crossover point to occur at a significantly lower $P_{0\ net}$ meaning that the single component crystals are the dominant pathway for a more significant portion of the process using larger volumes. 118

Figure 4-1 Cocrystal ternary phase diagram. Each axis of the triangle represents the concentration of one of the components of the ternary mixture. Each edge of the triangle represents one of three possible binary systems for which the third component has a concentration of zero. Within the triangle, all three components are present in some proportion. Each numbered region represents a different crystal mixture: 1. Single component crystals of coformer 1; 2. Cocrystals containing both coformers; 3. Single component crystals of coformer 2; 4. This region is undersaturated and no crystals are found here; 5,6,7. Each region contains a mixture of crystals corresponding to the overlap of regions 1,2, and 3. 128

Figure 4-2 Probability function for fixed initial coformer concentrations with threshold value. The probability of not finding a crystal of a specific type is represented on the y-axis, as a function of time on the x-axis. The three lines labeled P_{01} , P_{02} and P_{0Co} are the independent probabilities for each of the coformers and the cocrystal. P_{net} is the overall probability of not finding any of the three crystal types. The horizontal line at 0.5 depicts an arbitrary threshold at which there is a 50% probability of finding a crystal. Parameters used to calculate the presented curves are given in the text, and the initial coformer concentrations are $C_{i1}=10$ mg/ml and $C_{i2}=8$ mg/ml. 130

Figure 4-3 Translating probability distributions to transient ternary phase diagrams. Each location on the ternary phase diagram corresponds to a set of initial solution concentrations. At each set of initial conditions, the probability distribution for nucleation evolves over time as shown in the top graph. As time progresses the individual species probability curves will fall below a defined threshold value. $P_0 = 0.5$ is the threshold value indicated by the pink line in the top graph. When the threshold is reached for a given species, the corresponding point on the phase diagram is marked for that species at that time. 136

Figure 4-4 Cocrystal kinetic ternary phase diagrams progressing through time. Each of the four diagrams represent the same system at different points in time, progressing from 1 to 4. The time values corresponding to each panel are: $t_1 = 25$ s, $t_2 = 40$ s, $t_3=55$ s, $t_4= 70$ s and the same parameters are used as given earlier in the chapter. Each colored line represents the boundary at which the probability for the corresponding crystal species is equal to the threshold value. As time progresses these boundaries move from regions of lower concentration to higher concentration. 138

Figure 4-5 Kinetic history of crystal formation for similar initial compositions. Two initial compositions that are close together on the overall phase diagram are considered, denoted by the black dots. The movement of the threshold boundaries for each crystal type relative to the two black dots is depicted at the same four time points considered in Figure 4-4. The probability threshold boundaries reach the two points at different times. At these conditions the left dot is intersected by the single component crystal boundary first, and then later by the cocrystal boundary. The reverse is true for the right dot. While both of these points ultimately end up in the mixed crystal region of the equilibrium phase diagram, their kinetic history difference will likely lead to differences in their final crystal compositions. 140

Figure 4-6 Phase Diagram depicting first nucleating species. Ternary phase diagram in which each location corresponds to a set of initial concentrations of solute and coformers. The shaded regions indicate which crystallizing species is first to reach the threshold probability value at that concentration. The solution containing crystals will continue to evolve from this point until equilibrium is reached. 142

Figure 4-7 Saturation concentration boundaries. With the concentrations of coformer 1 on the y-axis and coformer 2 on the x-axis, each point within the graph represents a unique initial solution concentration pair. The dashed red and green lines represent the saturation concentrations for the respective components. The curved blue dashed line is the saturation concentration for the cocrystal. At the intersection of the cocrystal and single component saturation concentrations, dashed blue lines are drawn with slope corresponding to the stoichiometry to define the region in which equilibrium results in only cocrystals. From an initial concentration marked by the black point, several paths to equilibrium by nucleation are shown by the colored arrows. 145

Figure 4-8 Nucleation Probability Surface. Each x-y pair represents an initial solution concentration. The color of the point at each position represents the relative likelihood of possible nucleation pathways for formation of component 1 crystals (red), component 2 crystals (green), and cocrystals (blue). The relative intensity indicates the magnitude of the net probability at that time. The probability at each point evolves with time based on the ICCNT cocrystal nucleation theory model described in Chapter 3. 147

Figure 4-9 Path to equilibrium by nucleation. The concentrations of the two coformers are plotted along the x and y axes. Each line represents a population of volumes within a solution that start at a high concentration and then reduce the concentration by nucleating solid crystals. From each initial condition the population approaches equilibrium by a different path depending on the nucleation probabilities defined for the instantaneous concentration, depicted in **Figure 4-8**. The color of the line represents the relative proportion of component 1 crystals (red), component 2 crystals (green), and cocrystals (blue). 151

Figure 4-10 Comparing results of nucleation kinetic simulations to the equilibrium phase diagram. Phase diagram plotted with coformer concentrations represented along the x- and y-axes, assuming saturation concentrations of $C_{\text{sat}1}=0.9$ mg/ml and $C_{\text{sat}2}= 0.3$ mg/ml, and cocrystal saturation defined by equation (23) with $K_{\text{sp}}=0.05$. The dashed red and green lines represent the saturation concentration for each of the two coformers. The curved blue line represents the saturation concentration for the cocrystal. The diagonal dashed blue lines bound the region that at equilibrium is predicted to contain only cocrystals. All other regions bounded by saturation concentration lines are labeled according to the crystal mixture that is expected at equilibrium. Around the right and top edges of the plot the colored dashes indicate the results of the kinetic simulations. The position of each dash represents its initial concentrations and its color represents the proportion of each crystal species formed. Red represents component 1 crystals, green represents component 2 crystals and blue represents cocrystals. 153

Figure 4-11 Convergence to continuum behavior. For an initial set of concentrations, the population of crystals was simulated as it evolves to equilibrium via nucleation. 20 trial simulations were conducted, and then compared against the average of all runs. The root-mean-squared error between the average path and each individual path was calculated for a given number of volume elements and plotted above. Each line represents one of the cofomers. Between 400 and 1000 volumes, the deviations from the mean reached a minimum. 154

Figure 4-12 Convergence with decreasing time step. The RMS error is plotted on the y-axis versus the time step along the x-axis, decreasing from left to right. As the time step becomes smaller, the difference between the numerical and analytical solution to P_{net} decreases, reaching a minimum at approximately 0.01s..... 155

Figure 4-13 Convergence of numerically computed P_{net} . Plot of probability on the y axis versus time on the x axis, for several simulations at identical conditions but with different timesteps used in the numerical integration steps. As the time step gets smaller, the difference between the numerical and analytical solution for P_{net} gets smaller as well. 156

Chapter 1 : Introduction

Crystallization plays a critical role in diverse industries such as pharmaceuticals and food sciences [1]. Pharmaceutical companies invest heavily in the research and development of new Active Pharmaceutical Ingredients (APIs) [2,3]. Merck spends in excess of \$4 billion per quarter on R&D costs [4], and Forbes predicts total prescription drug sales will surpass \$1 trillion in the next four years [5]. More than 90% of APIs involve a crystallization step as part of their production [6]. Purification, manipulation of dissolution rates, and control of macroscopic properties for easier handling are common uses of crystallization [6]. Drugs that require a crystallization step as part of their synthesis include familiar over-the-counter products like aspirin and ibuprofen, and a wide range of others like Januvia, a Type 2 diabetes medication and one of the top selling Merck products [7]. Despite the evident importance of crystallization processes in industry, control of such processes is still largely empirical due to the very large parameter space that governs nucleation and growth of specific crystal forms with desired properties. The problem is even larger when one considers more complex crystallization processes such as cocrystallization of two or more compounds. Empirical determination of effective API formulations and processes is remarkably resource intensive, consuming materials, money, and labor. Improved screening tools and approaches that consume less material and allow for greater generalization of the fundamental physico-chemical mechanisms underlying crystallization would have a significant impact on the industry by potentially reducing development costs and identifying improved formulations.

Manufacturers of crystalline products typically require specific targets for their physical properties and compositions [17]. One key property is the crystal habit, or the

macroscopic shape of the crystal, which can range from nearly spherical to long, thin needle-like structures. It is common for materials to be able to form several different crystal habits, and it is usually desirable to design a process that selectively produces only one. In addition to the shape of the crystal, size polydispersity is often of concern when producing crystals. More monodisperse crystals allow for easier material handling and have more consistent surface-area-to-volume ratios. Finally, the specific polymorph of the crystal, or the arrangement of molecules in the crystal lattice, can affect shelf life and bioavailability, or the proportion of the substance which is absorbed by the body, for APIs. Control over the physical properties of a crystal requires an understanding of the crystal formation process. Nucleation is the critical first step in that process.

The primary thermodynamic property that controls the nucleation rate is the supersaturation of the solute. Each solvent can dissolve a fixed concentration of a solute depending on experimental conditions. To form a supersaturated solution, a change to the system must reduce the solubility of a solute. Reduced solubility favors phase separation, but kinetic barriers that prevent instantaneous agglomeration of solutes hinder phase separation. The competition between the chemical potential driving the dissolved solute to a solid crystal phase and kinetic barriers gives rise to a characteristic of the crystallization process known as the “metastable zone,” defined as the region in which the system is supersaturated, but crystals do not readily nucleate. Experimental conditions such as solvent composition and cooling rate control the width of this region. There are several proposed empirical models to calculate the metastable zone width [8-11]. These models define the metastable zone width as the difference between the temperature at which the solution first becomes supersaturated and that at

which the first crystal nucleates. The temperature difference is a function of cooling rate, with faster cooling rates leading to wider metastable zone widths.

Supersaturation is a function of many more parameters than just temperature. Ionic strength, pH, and other components in the system can all affect the solubility of a solute. There are previous explorations of this phase space by groups like Zukoski et al. [2], who utilized microliter-sized drops in chambers with controlled water vapor pressures. They defined the “supersolubility boundary” by slowly removing water from drops of lysozyme solution until the first crystals formed. Varying the concentrations of lysozyme and NaCl salt in the initial drops led to different final concentrations, mapping out the supersolubility boundary. Supersaturation depends on many physical properties, creating a very large associated parameter space for nucleation conditions.

The large parameter space explains why it is still common to experimentally determine the nucleation rate for proteins and other complex molecules, as there is not a predictive model for determining nucleation rates [12]. However, control over nucleation is a key to controlling the properties of crystallized products [13]. Determining crystallization conditions empirically on a per material basis consumes R&D resources and poses a serious impediment to the development of new products that require a crystallization step in their manufacturing. For many applications like APIs, the material cost of macroscale crystallization studies is prohibitive due the large volumes involved. Drop based microfluidic strategies offer an alternative platform that can be used to quantify nucleation rates and associated kinetic constants over a wide range of conditions, while using nanoliter volumes of material per drop. One goal of this thesis is

to develop improved drop based microfluidic approaches to studying crystal nucleation rates and investigating their behavior for a wide range of experimental conditions.

In some cases, there are no crystallization conditions that lead to a desired product. In these cases, more complex forms of crystallization may be employed. For example, by forming an API into a salt, the bioavailability of a previously insoluble compound can increase [14]. In other cases, forming a cocrystal with one or more additional components can lead to desired properties [15]. Cocrystallization increases the complexity of the system considerably compared to single component crystallization by adding additional dimensions to the already large parameter space. This can make it difficult to effectively screen for possible cocrystallizing conditions for APIs [16]. Methods like the Kofler technique, in which drops of potential cocrystal forming solutions are placed under a microscope slide on a heated stage and observed until crystals form, require sample volumes of the order of hundreds of milliliters [15]. Such volumes are unfeasible for many expensive APIs. The lack of practical techniques to measure the nucleation rates of cocrystal systems has resulted in a general lack of study of this phenomenon. A drop based microfluidic platform offers the ability to systematically explore the cocrystallization phase space while minimizing the consumption of potentially expensive components by reducing the required volumes by at least 1000-fold.

This thesis presents both an experimental approach to measure the homogenous nucleation rate of crystallization from solution using drop-based microfluidic devices, as well as a theoretical model for the nucleation rate of a cocrystal system. Chapter 2 presents an overview of the state-of-the-art of nucleation modeling and a survey of

macroscopic experimental techniques currently used to measure nucleation rates. A comparison is presented of two different drop microfluidic approaches: a static array of a fixed number of drops and a continuously flowing train of drops. Results for a model crystallizing system are examined and the advantages and pitfalls of each method are outlined. Chapter 3 introduces the state-of-the-art of cocrystal research, highlighting the current knowledge gap for cocrystal nucleation modeling and experimentation. A novel kinetic model for cocrystal nucleation is developed by considering the rates of formation of cocrystal complexes as a part of the crystallization process. The relevant timescales for each part of the process are examined, revealing several controlling dimensionless parameters. The salient features of the cocrystallization model are investigated as a function of these dimensionless timescale ratios, and potential experimental approaches to validation of the model are outlined. The model and relevant fundamental timescales also suggest approaches for separating and measuring the relevant timescales using the small volumes available in drop based microfluidics. Chapter 4 demonstrates an application of the novel kinetic model for cocrystallization. The model is used to explore the pathways to cocrystal formation and to subsequently determine phase diagrams for equilibrium and metastable cocrystal systems. Chapter 5 provides a summary of the contributions of this thesis and considers potential future work.

References

- [1] Jon F. Edd, Katherine J. Humphry, Daniel Irimia, David A. Weitz, Mehmet Toner. *Lab Chip*, 2009, 9, 1859–1865
- [2] M. R. Thorson, S. Goyal, B. R. Schudel, C. F. Zukoski, G. G. Zhang, Y. Gong and P. J. Kenis, *Lab Chip*, 2011, 11, 3829–3837.
- [3] S. L. Morissette, O. Almarsson, M. L. Peterson, J. F. Remenar, M. J. Read, A. V. Lemmo, S. Ellis, M. J. Cima and C. R. Gardner, *Adv. Drug Delivery Rev.*, 2004, 56, 275–300.
- [4] <https://investors.merck.com/news/press-release-details/2018/Merck-Announces-Fourth-Quarter-and-Full-Year-2017-Financial-Results/default.aspx>
- [5] <https://www.forbes.com/sites/johnlamattina/2018/06/12/pharma-rd-investments-moderating-but-still-high/#2000ea0e6bc2>
- [6] Narayan Variankaval and Aaron S. Cote. *AIChE Journal* July 2008 Vol. 54, No. 7.
- [7] https://www.ema.europa.eu/documents/scientific-discussion/januvia-epar-scientific-discussion_en.pdf
- [8] L. Xiong et al. *Journal of Molecular Liquids* 269 (2018) 805–815
- [9] J. Nývlt, R. Rychlý, J. Gottfried, J. Wurzelová, Metastable zone-width of some aqueous solutions, *J. Cryst. Growth* 6 (2) (1970) 151–162, [https://doi.org/10.1016/0022-0248\(70\)90034-5](https://doi.org/10.1016/0022-0248(70)90034-5).
- [10] N. Kubota, An interpretation of the metastable zone width concerning primary nucleation in anti-solvent crystallization, *J. Cryst. Growth* 310 (2008) 4647–4651.
- [11] K. Sangwal, A novel self-consistent Nývlt-like equation for metastable zone width determined by the polythermal method.pdf, *Cryst. Res. Technol.* 44 (3) (2009) 231–247.
- [12] Peter G. Vekilov, 2010, 2, 2346–2357
- [13] David I. Trauffer, Anna K. Maassel, and Ryan C. Snyder, *Cryst. Growth Des.* 2016, 16, 1917–1922
- [14] David J. Berry, Jonathan W. Steed, *Advanced Drug Delivery Reviews*, 117, 2017, 3–24
- [15] Daniel P. McNamara, Scott L. Childs, Jennifer Giordano, Anthony Iarriccio, James Cassidy, Manjunath S. Shet, Richard Mannion, Ed O'Donnell, Aeri Park. *Pharmaceutical Research*, Vol. 23, No. 8, August 2006
- [16] Mary K. Stanton, Annette Bak. *Crystal Growth & Design*, Vol. 8, No. 10, 2008
- [17] Code of Federal Regulations, Title 21, Volume 7, Sec. 610.13

Chapter 2 Crystallizing Lysozyme in Continuous Drop Trains

Abstract

Microfluidics have provided the next step in the miniaturization of homogenous nucleation research. Over time the goal of handling ever smaller liquid samples while increasing statistical accuracy has driven innovation. This work demonstrates the utility of microfluidics to create static arrays of nanoliter sized drops that can be used as independent reaction chambers. These drops can be used to explore nucleation kinetics by controlling both the temperature and the composition of the crystallizing solution. However, static devices have limitations on the number of observable drops, which limits statistical accuracy. To mitigate these limitations, we design and implement a continuous flow droplet-based microfluidic device that is capable of generating a train of drops in which the number of independent drops is no longer coupled with the physical size of the device. Nucleation rate experiments performed in both static array devices and continuous drop train devices allow us to compare the advantages and shortcomings of each design, allowing determination of the best use cases for each type of device.

Introduction

The widespread and common use of crystalline products in diverse industries such as food science, chemical processing, and pharmaceuticals clearly indicates the vital importance of understanding the crystallization process [1-3]. There are several steps in the process of making a crystal that each contribute to the properties of the resulting product. Broadly, the two main processes that drive crystallization are nucleation and growth [4]. Nucleation describes the process of phase transition from solute dissolved in liquid phase solution to a solid phase particle suspended in that

solution [5]. Nucleation is a kinetic process whose rate is controlled by several properties of the solution including temperature, ionic strength, and the presence, chemistry, and amount of other additives [6]. Growth is the process by which material continues to add to existing crystals. The growth stage is controlled by parameters such as the diffusion coefficient, local concentration gradients, and the geometry of the crystal growth planes. While nucleation and growth occur sequentially for an individual crystal, the stochastic nature of the nucleation process means that nuclei can appear over timescales governed by kinetic rates associated with nucleation and the resulting probability distribution. If the fundamental timescales for nucleation and growth are comparable, then the two processes can appear to be mixed as many crystals form within a given solvent volume. Developing models that qualitatively and quantitatively capture the crystallization process, and experimental methods that enable isolation and quantification of each of the underlying fundamental mechanisms will allow for more effective design of industrial scale processes [7].

In this chapter we seek to improve the accuracy of nucleation rate measurements by comparing several droplet-based microfluidic strategies. As new approaches have been developed for measuring nucleation rates, the volume of liquid required has decreased while the number of independent events that can be observed has increased. This has been accomplished in part by using droplets in stabilized emulsions as tiny “reactors” for examining individual nucleation events. However, static droplet methods quickly reach practical limits in both the volume and the number of events that can be readily observed. In this thesis, we demonstrate a continuous droplet train microfluidic device that can be used to observe an arbitrarily large number of

independent nucleation events within a finite-sized microfluidic device. Using this device, we evaluate the strengths and shortcomings associated with using continuous drop generation microfluidics and compare with results from static array microfluidics.

Control over nucleation rates is essential to producing consistent, high quality crystalline products for a variety of reasons. For example, nucleation rate strongly affects crystal size polydispersity. Consider a supersaturated solution in which crystals begin to nucleate. When the first nucleation event occurs, the new nucleus removes solute molecules from solution, decreasing the overall concentration of dissolved solute. If the nucleation rate is slow, few crystals will form initially and begin to grow, allowing the solution to remain supersaturated for a longer time. If this time is long enough to allow further nucleation to take place, the nuclei that form later will grow into smaller crystals than the nuclei that form earlier, due to the reduction in available material in solution [39]. When producing crystalline pharmaceuticals it is often desirable to produce a uniform crystal size, as the crystal surface-area-to-volume ratio can affect the dissolution rate in the body once the drug has reached its treatment target [40]. If nucleation rates are not controlled, the crystals produced can vary significantly in size and require multi-step downstream processing approaches, such as grinding, sieving, and filtration to select for the desired size, which dramatically reduces yield [41].

The nucleation rate plays a role in other features of the final product as well. For example, when characterizing new crystalline materials by methods such as x-ray diffraction (XRD), large single crystals are required [73]. To reliably produce large crystals from solution, control of the nucleation rate is required to produce very few but very large crystals. If nucleation rates are too rapid and many crystals form, nearby

crystals may merge, forming undesired polycrystalline materials. To prevent secondary nucleation and produce only single crystals suitable for XRD, nucleation operations are performed near the saturation concentration as described by Shanthi *et al.* [42].

Additional considerations beyond crystal size polydispersity include determining and controlling which polymorph, or configuration of the crystal lattice, is formed in the nucleation process. Even relatively simple materials like amino acids can have multiple polymorphs, and the rate of nucleation affects which polymorph emerges. For example, the amino acid glycine can form three different polymorphs, known as alpha, beta and gamma, and each can be formed from solution depending on the nucleation conditions. He *et al.* showed that gamma-glycine polymorphs could be produced only at slow nucleation rates, whereas fast rates produced the alpha or beta polymorphs [43].

Polymorph is an important characteristic of the final crystal product since the macroscopic geometry of each polymorph, known as the crystal habit, is different and can affect how the material must be processed. If the expected product exhibits a spherical crystal habit that flows easily, the unexpected production of a needle-like habit could cause jams or clogs in a system. Beyond mechanical handling considerations, some polymorphs are also more chemically stable than others [44]. Matsuda *et al.* provide an example of polymorph stability by comparing the photodegradation times of three carbamazepine polymorphs [45]. Their work showed that the form II polymorph degraded much more rapidly when exposed to light compared to forms I or III. These examples demonstrate that nucleation plays an important role in the properties of crystalline materials, and therefore understanding and controlling nucleation rate is

useful in determining the most important features and characteristics of the final crystalline product.

Crystallization is often used in industry to accomplish both separation and purification of materials [8]. The most common large-scale crystallization operation, used in the creation of APIs, is cooling crystallization in 1000L semi-batch tanks [9]. Large scale operations can only be conducted successfully once crystallization conditions have been identified in smaller scale empirical studies, which don't always translate to large scale tanks in a straight-forward way due to the large parameter space and interplay of mechanisms like diffusion and kinetics. A common smaller scale, but still macroscopic approach to measuring nucleation rates utilizes small, stirred-tank reactors with volumes of several milliliters [10]. As crystals nucleate in solution, the turbidity rapidly increases, which provides a means of measuring nucleation rate. Typically, micro-batch approaches to crystallization screening induce crystallization by reducing the temperature below the saturation temperature. Micro-batch reactors can induce nucleation through other means as well, including injecting gases into supersaturated glycine solutions and determining the nucleation rate by the change in absorption over time [11]. One challenge in interpreting the results from macroscopic techniques is that the effects of growth and nucleation are often mixed and difficult to de-convolve. Careful handling of the data and planning of multiple experimental trials are needed to isolate the nucleation rates [69].

There are several bench scale methods for measuring nucleation rates of materials crystallizing from solution, with the most common detection method still being turbidity measurements, where a sample is prepared and then rapidly cooled to

increase the supersaturation. Then, optical probes are used to measure the increase in turbidity of the solution, indicating the presence of a new solid phase [14]. The rate at which turbidity changes is used to determine the rate of nucleation within the system. Nucleation is a stochastic process driven by Brownian motion, so it is necessary to capture a large number of nucleation events in order to gather sufficient data to analyze the nucleation statistics [15]. The requirement for large numbers of events has led to the development of parallel nucleation measurement techniques like the Technobis Crystallization Systems “Crystal16”. As the name implies this device can carry out up to 16 independent crystallization experiments at once. With increased parallel experimentation, one concern is that material consumption increases. The Crystal16 system claims to achieve 16 parallel screens using less than 1 mL volumes and less than 100 mg solute, while producing solubility curves in less than four hours [16]. The system works by measuring the transmissivity of 16 individual mL sized vials held at specified fixed temperatures.

Another common method of characterizing crystallization is Differential Scanning Calorimetry (DSC) [17]. DSC uses much smaller volumes of approximately 10 μ L and measures the change in energy required to change the sample temperature as the material undergoes a phase change [46]. Despite the much smaller sample volumes, DSC is not easy to parallelize due to the need to thermally isolate the sample, making it unsuitable for high throughput screening. In general, methods like DSC and the Crystal16 system also rely on expensive proprietary equipment, creating a tradeoff between the cost of purchasing and maintaining the equipment versus the cost of materials saved by the reduced volumes.

Ismagilov *et al.* review three other traditional macroscale techniques for creating protein crystals known as vapor diffusion, dialysis, and free interface diffusion [12]. Vapor diffusion is a crystallization technique in which a small drop containing the crystallizing solution equilibrates with a large reservoir containing similar buffers and precipitants at higher concentration. Dialysis separates the solute from a precipitant by a semi-permeable membrane and allows the precipitant to slowly diffuse into the solution until crystallization occurs. Free interface diffusion is a crystallization technique in which pure solutions and precipitants gradually mix by diffusion under the influence of a concentration gradient across the free interface between the two pure solutions. While effective for investigating properties like the metastable zone width, these three methods all require large solution volumes of a few mL compared to microfluidic methods, which use only a few μL [72]. Additionally, the direct measurement of nucleation statistics is impossible at larger scales, as nucleated crystals remove free solute and reduce the bulk solution concentrations such that the nucleation events are no longer independent [68].

New methodologies have been developed to reduce the solution volume consumed, while simultaneously increasing the number of independent trials. In 2002, Galkin *et al.* developed a method that uses a plate with an array of 400 $7\mu\text{L}$ wells drilled into the surface [13]. Each well acts as an independent reaction chamber for collecting information about nucleation kinetic rate constants. This method has the ability to isolate small liquid volumes and directly observe crystal formation. However, this method lacks flexibility since the geometry is hard-wired and each well must be filled independently.

The development of microfluidic approaches to the characterization of physico-chemical systems has created opportunities to improve upon the earlier crystallization methods. Droplet-based methods are particularly valuable due to the ability to isolate a tiny sample volume within each droplet, to protect the reactor contents using the droplet interface, and to easily manipulate these volumes. There are numerous methods available for generating droplets [52-55]. In general, microfluidic devices with characteristic dimensions between 10 and 100 micrometers can produce drops with volumes from nanoliters to picoliters. Drops can be produced in large numbers at rates ranging from 10s to 1000s of Hz with as small as 1% to 3% polydispersity depending on conditions [56]. In addition, drops are easily manipulated on chip, so they can be collected and stored, observed using various light-based approaches, as well as merged and mixed with other drops [57-58]. Precise control over large numbers of independent volumes of liquid is ideal for the study of stochastic phenomena like nucleation.

Droplet based microfluidic techniques have been used for the study of crystal nucleation kinetics due to the ability to generate large numbers of monodispersed drops that act as independent reaction chambers [47] and the dramatic reduction in material consumption. Individual microfluidic drops are between 10^6 and 10^9 times smaller than the traditional micro-batch characterization process, resulting in substantial reductions in the material cost involved with determining nucleation rates for expensive novel materials. Work by Akella *et al.* determined that at least 500 droplets are required for statistical accuracy, therefore a droplet based microfluidic device using 1 nL sized drops would consume approximately 10^3 less material than a 1 mL micro-batch

experiment [21]. Beyond reducing material consumption, the individual microfluidic drops are isolated, which leads to additional benefits. In larger volumes, nucleation can occur at any location within the solution. Once a crystal has formed, the nucleus consumes material and reduces the bulk concentration in the immediate region, creating concentration gradients within the solution and inducing additional transport-related effects [70]. These effects are very challenging to deconvolute from the already complex mix of nucleation and growth. By isolating the solution within picoliter-scale fluid drops, transport timescales are significantly faster due to the small length scales, and nucleation kinetics therefore become the controlling mechanism for crystallization. In addition, interaction between individual stochastic nucleation events is prevented, creating the opportunity to easily monitor many independent events. Additionally, since microfluidic devices are typically fabricated in optically transparent material, newly formed crystals can be directly observed in a microscope, eliminating the need for any proxy measurement for the presence of nucleation [71].

Microfluidic approaches offer several methods for creating isolated drops, including both continuous drop production as well as creation of static drop arrays. One common method of generating an array of stationary drops is known as the 'store-and-create' method. This approach uses an interconnected series of capillary traps to capture small volumes of a dispersed phase fluid as it flows through the device. Initially, a fluid fills a series of wells interconnected by small exit channels and larger bypass channels. A second immiscible fluid is introduced such that it displaces the first fluid. The invading fluid flows into both the wells and the bypass channels, but it cannot pass through the small exit channel, since that would require a higher driving pressure to

overcome the excess capillary pressure created by the more highly curved interface. The original fluid is introduced again to displace the invading fluid. As the original fluid re-fills the channels, the invading fluid cannot escape the wells due to the excess capillary pressure, and small volumes break off and become trapped in the wells. Boukellal *et al.* demonstrated this technique, noting the potential to use the trapped drops as independent reaction vessels [18]. A related technique, dubbed 'create-and-store', utilizes a similar geometry but generates drops upstream before flowing them through the device, capturing them within the capillary traps [48]. Both techniques generate or capture a fixed number of drops prescribed by the fabricated geometry. Static array microfluidics that store drops in individual traps can isolate hundreds of drops simultaneously, but the density of capillary traps per unit area is limited by the necessary surrounding geometry such as the bypass channel. Limited ability to pack large numbers of drops together in a region creates challenges for visualization of the crystallization events with the limited field of view available in microscope and camera systems, and the need for sufficient magnification to resolve individual picoliter sized drops. To further increase the drop density, microfluidic devices were developed to collect and hold surfactant stabilized emulsions by generating thousands of drops and storing them all in a single reservoir, with coalescence prevented by a carefully chosen surfactant [59-60]. However, studies by Abedi *et al.* have shown that despite the surfactant-coated drop interfaces, drops can still influence one another during nucleation experiments [49]. They demonstrated that nucleation events frequently occurred in neighboring clusters of drops in a stabilized emulsion, suggesting that the state of a neighbor drop can influence when nucleation occurs. Inter-drop interactions

cause the observed nucleation rates to differ from the values that are expected if events are truly independent.

Since the close packing of stabilized emulsions represents the upper limit of drops per area density, other approaches are needed to increase the number of independent observable drops. Continuous production of drops and continuous flow through a detection region is one approach to increasing the number of independent events observed. Stan and Whitesides implemented this technique to study the freezing behavior of supercooled water drops generated by a flow focusing nozzle [51]. The drops passed through a linear channel over thermoelectric coolers which carefully controlled the temperature until the drops froze. The 80 μ m-diameter drops were generated at a rate of 75 Hz and frozen within less than 400 μ s of being generated. Zhang *et al.* demonstrated the formation of metal nanoparticles in a continuous drop generation device that fed into millimeter sized tubing, greatly increasing the residence time for nucleation time, but at a larger scale than is typical for microfluidics [64]. Work by Teychené *et al.* showed the ability to measure nucleation rates for organic materials in glass microfluidic devices [63]. Their experiments showed clear improvements in residence time and material consumption but struggled to adapt the method to soft lithography due to solvent compatibility issues. These examples and others show a clear progression toward flexible microfluidic platforms for the study of nucleation from solution.

There are numerous methods of producing drops continuously, including co-flow, crossflow, and flow focusing geometries [50-51]. Continuous drop generation methods have the immiscible continuous and dispersed phase liquids flowing simultaneously and

continuously into the device, meeting at a junction with a specific shape designed to promote drop breakup. The continuous pressure driven flow of the two immiscible liquids results in generation of a continuous ‘train’ of drops with control over drop sizes and low polydispersity [19]. Drop trains allow for an unlimited number of drops to pass through a microfluidic device and be observed, offering a promising approach to further improve upon nucleation kinetic measurements.

In this chapter, we investigate two droplet based microfluidic approaches to the measurement of crystal nucleation kinetics. The ‘store-and-create’ method is developed to measure nucleation using either control of temperature or control of concentration to induce supersaturated conditions. The results obtained from static arrays are compared with a continuous droplet train method. In each case, the design of the geometry and the operation of the device are examined, and image analysis is used to determine the fraction of nucleation events over time, which is related to the nucleation rate function. Experimental factors influencing the measured nucleation events, such as droplet polydispersity and temperature uniformity, are investigated for each method. A model crystallizing system composed of lysozyme, an enzyme found in chicken egg whites that crystallizes from aqueous salt solution, is used to compare with prior reported results and to assess the advantages and disadvantages of each approach.

The next section describes the materials and methods used in these experiments. Then, a brief description of nucleation kinetic modeling is introduced using classical nucleation theory as the foundation. The connection between nucleation kinetic modeling and droplet-based observations of nucleation events is described. Next, the design, operation, and analysis of static drop arrays is presented along with

nucleation results for the lysozyme system using this approach. The design, operation, and analysis of continuous drop trains is presented along with nucleation results for the lysozyme system using this approach. Finally, the two methods and their results, advantages, and pitfalls are discussed.

Materials and Methods

Static Array Temperature Step Nucleation

Microfluidic devices were fabricated in polydimethylsiloxane (PDMS) using standard soft lithography techniques for the purpose of creating and trapping drops of crystallizing solution; the design of the devices will be discussed later in this chapter and can be seen in **Figure 2-5** [61-62]. The molded PDMS containing the microchannel design is plasma bonded to another slab of PDMS, ensuring that all microchannel walls have the same wettability. Once bonded, the devices are soaked in 100 cSt silicone oil for at least 3 days prior to use, to ensure that the permeable PDMS is fully saturated with oil.

Crystallization studies are conducted using lysozyme as the crystallizing molecule. Lysozyme was selected because its crystallization behavior has been examined previously in static drop arrays. Nucleation rate constants were reported by Akella *et al.*, providing a basis for comparison with the experiments presented here [21]. Lysozyme crystallization is influenced by the presence of salt and other additives, and we selected a mixture composition for our experiments that allowed for comparison with previous studies. Lysozyme derived from chicken egg white was obtained from Sigma-Aldrich (cat. no. L6876) and used as received. Lysozyme concentrations vary from 20 to 30 mg/mL, and were selected to match supersaturation values reported previously [20-

21]. In addition to lysozyme, the solutions contain 12.5 w/v% 10kD polyethylene glycol (PEG, obtained from Sigma-Aldrich, cat. no. 309028), 2 to 5 w/v% NaCl depending on the specific trial, an equilibrium concentration of Brij 93 surfactant (Sigma-Aldrich, cat. no. 388866), and a 0.1M sodium acetate (NaAc) buffer to maintain a constant pH of 4.8. These additives serve to alter the solubility of lysozyme through several different pathways. The pH of the solution affects the surface charge of the protein, and the addition of salt acts to screen this charge and reduce electrostatic repulsion between protein molecules. The addition of PEG aids in the aggregation of lysozyme through the depletion attraction effect [67].

When using drops, it is important to remember that the surface area to volume ratio is significantly higher than in larger scale experiments. Since proteins like lysozyme can also adsorb to oil-water interfaces, a surfactant is used to inhibit adsorption of the protein [65]. The surfactant must meet a few key requirements to be useful for microfluidic protein crystallization applications. First, the surfactant must be biocompatible meaning it does not denature or otherwise damage the protein, and must prevent the adsorption of the protein at the oil-water interface. Adsorbed proteins can denature, and reduce the bulk concentration by their removal from solution. Block-copolymer surfactants that include a hydrophilic PEG group have been shown to create biocompatible interfaces [66]. Additionally, the surfactant must be compatible with the surface of the microfluidic device to avoid surface wetting and permit the generation of drops. Several PEG-based surfactants were tested with wettability and biocompatibility considerations in mind, and it was found that Brij 93 produced stable drops in static array devices. Independent experiments using the pendant drop method showed that

drops created in solutions of 0.02M Brij 93 or greater exhibited identical surface tension with or without lysozyme in solution. This suggests that Brij 93 dominates the interface and prevents lysozyme from absorbing, and for that reason it was selected to be the surfactant used for all static array experiments considered here. Brij 93 dissolves very slowly into water and a high concentration was desired to ensure that its adsorption would dominate the interface. Therefore, when preparing solutions, excess Brij 93 was added and the solution was left to sit for at least a day to allow for equilibrium dissolution of the surfactant. It was noted that in the containers where the water-surfactant mixture was prepared a thin film formed at the air-water interface when excess Brij 93 was added. All further solutions were made using liquid collected by inserting a syringe substantially below this film, and no notable effects on nucleation were recorded in the presence of excess Brij 93.

The solution compositions selected are used as the immiscible dispersed phase liquid when generating droplets for the microfluidic experiments. All solutions used in the experiments considered here are supersaturated with respect to lysozyme to promote crystallization, so they were formulated in two parts to prevent crystallization prior to droplet formation. Both solution parts contained Sodium Acetate(NaAc), Brij 93 surfactant, and PEG. One portion contained lysozyme and the other contained NaCl, which acts as a precipitant to promote crystallization. Each microfluidic geometry is designed to maintain separation of the two solution parts for as long as possible before generating droplets. For the static array experiments, the two solutions are mixed immediately prior to introducing the dispersed phase fluid into the device.

The two solution parts were created in the following manner: First, NaAc is added to DI water to create 60 *mL* of 0.1 *M* solution. HCl is added to adjust the pH to 4.8. 1 *mL* of Brij 93 surfactant is added and thoroughly mixed with the solution. Brij dissolves slowly into water, so the solution is allowed to rest for a day to ensure complete dissolution. Excess Brij forms a film at the surface so the saturated solution is withdrawn using a syringe and transferred to a second container to remove this film. Next, 12.5 *w/v%* PEG is added. To fully dissolve the polymer, a magnetic stir bar is used to mix the solution for at least 6 hours. The final mixture is used as the stock solution that forms the basis for both the lysozyme and precipitant solutions. The stock solution is divided in half. To one half, 10 *mg/mL* of NaCl is added. To the other half, twice the desired final concentration of lysozyme is added to create the protein solution. When mixed in a 1:1 ratio by volume, the salt and lysozyme are diluted to the desired concentrations, while the PEG concentration remains constant. The continuous phase fluid is 100 cSt silicone oil.

The oil with 1 *wt%* Brij 93 surfactant is first injected into the channels at a flow rate of 12 $\mu\text{L}/\text{min}$ for 15 minutes. This initial introduction of oil displaces the surfactant free oil in which the devices were soaked prior to use, and it later helps prevent the aqueous solutions from wetting the channel walls. The precipitant and protein solutions are mixed manually at a 1:1 ratio, immediately loaded into a syringe, and then injected into the device at a flow rate of 1 $\mu\text{L}/\text{min}$ for 45 minutes, partially filling all the capillary traps and surrounding channels. At this point the bypass channels contain enough solution to fill the remaining traps when the oil is injected in the next step. Finally, the silicone oil is injected into the device again at 12 $\mu\text{L}/\text{min}$ for 10 minutes, displacing

excess aqueous solution from the bypass channels and leaving behind only the newly formed drops in each of the trap wells. At this point, a filled device is ready to be used to perform a crystal nucleation experiment.

To control the temperature, a recirculating water bath is first equilibrated at the desired temperature. A bath temperature setting of 4 °C results in droplet temperatures of 5 °C. The cooling jacket was constructed from an acrylic block with a reservoir machined into the middle, through which the cooling fluid is continuously pumped. The reservoir is sealed off by a glass slide upon which the microfluidic device is supported. The loaded device is placed into the cooling apparatus with the thin base layer closest to the cooling liquid. To prevent air bubbles in the cooling water from obstructing the view, the entire assembly is then inverted. Images are recorded, noting any time delay from placing the device to the start of recording. The device is monitored until all droplets have formed a visible crystal. Images are captured at a rate of one frame every two minutes by a microscope camera installed on a light table. (Richards Corporation Image Interpretation Systems HFO-4 with a Bausch and Lomb microscope fitted with a CCD camera - Panasonic WV-CD22).

Static Array Dehydration Induced Nucleation

Two different static array experiments are conducted. The first experiment type mirrors standard nucleation experiments in macroscale solution, in which the temperature of the solution is decreased rapidly to increase the supersaturation to initiate crystal nucleation. The second type of experiment maintains a fixed temperature, but gradually increases the solution concentration above supersaturation via slow

dehydration of the droplets. These two approaches represent two methods of obtaining supersaturated solution and result in different observed nucleation rate curves.

For static array dehydration experiments, the device is inverted such that the thin base layer faces upward, and is placed into the humidity control chamber, which passes dry air over the device to maintain a 5% relative humidity environment in the environment just outside the device. The concentration gradient of water across the surfaces of the device causes water to diffuse out of the droplets and into the environment. The droplet volumes decrease steadily with time. As soon as the dry air stream is activated, images are captured at a rate of one frame every two minutes by the same optical setup used for the temperature step experiments. Data is recorded for up to 12 hours, or until all droplets in the device exhibit a visible crystal.

Continuous Droplet Train Nucleation

For the continuous droplet train experiments, the two solutions are injected into the device simultaneously in separate inlet channels that come together along with a central stream containing a buffer solution just prior to entering the droplet generation region. The length of the channel segment in which the three streams co-flow together is 0.2 *mm*, which is significantly shorter than the length needed for full mixing to occur. Therefore, the solution parts enter the generated drops together at the appropriate concentrations, but do not mix at all until they are together within the drops.

For experiments using continuous drop trains, the continuous phase oil used is the fluorinated oil FC-70 (Sigma-Aldrich, cat. no. F9880), containing 0.1 *wt%* PicoSurf fluorosurfactant (Sphere Fluidics, cat. No. C012). In this case, we found that the fluorinated oil and fluorosurfactant combination performed better in preventing droplet

coalescence in the residence channel. Values listed in **Table 2-1** correspond to the final drop compositions resulting from the mixing of the three inlet streams, or from off-chip mixing for static arrays. Supersaturation values shown are calculated based on equation (5) described in the next section using parameter values obtained from Forsythe *et al.* [33]

The PDMS microfluidic devices used in this chapter were fabricated using standard soft lithography techniques and placed on the stage of a Nikon Eclipse Ti-U Inverted Microscope at 2X magnification (Nikon Imaging, Melville, NY) [34]. Images were captured with an sCMOS pco.panda 4.2 camera (Pco-tech; Wilmington, DE). Example images captured from the continuous droplet train device can be seen in **Figure 2-1**. For this work, we automated the image collection process using

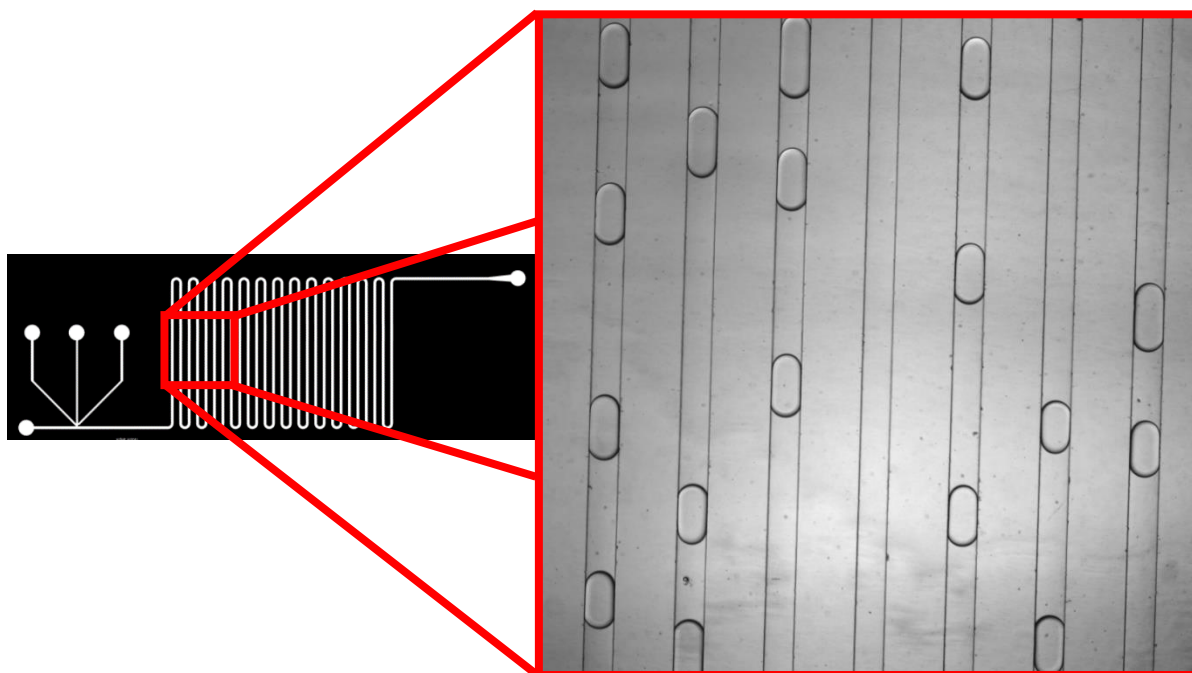


Figure 2-1 Droplet observation within microfluidic device. Optically transparent PDMS allows for the direct visual observation of droplets as they flow through the serpentine residence chamber. Constant infusion rates from the syringe pumps driving the flow allow for a direct mapping of droplet position to residence time within the device.

MicroManager Software. A custom cooling jacket connected to a recirculating chiller (VWR AD15R-40-V11B, Cat. no. 89202-994) was used to maintain the temperature within the device. Unlike with the static array devices, the cooling jacket was placed on top of the microfluidic device to accommodate the working distance of the microscope. This required a minor redesign of the acrylic block, but its basic function remained the same. Three Harvard Apparatus PHD2000 syringe pumps infused the three aqueous streams and the oil stream into the device. A schematic diagram of the continuous droplet train experimental setup is shown in **Figure 2-2**.

The three aqueous streams were connected to the device along with the continuous oil stream such that the aqueous streams meet before entering the T-junction at which droplets are formed. All aqueous streams were driven at $0.075 \mu\text{L}/\text{min}$ and the oil stream was driven at $1.00 \mu\text{L}/\text{min}$. Infusion rates were chosen to balance

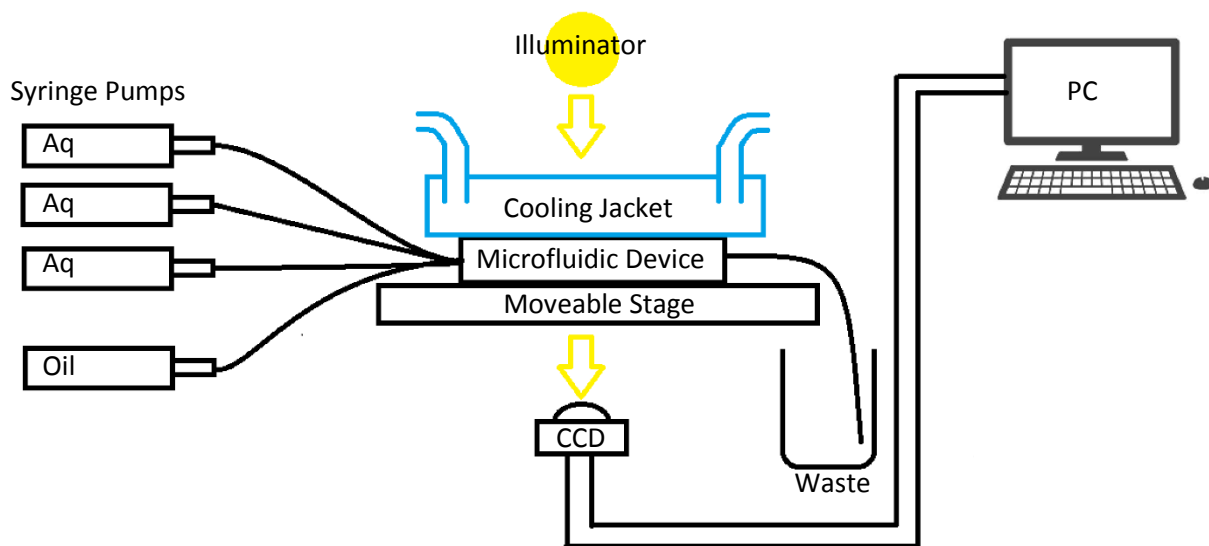


Figure 2-2 Schematic diagram of the experimental setup for a continuous crystallization device. Four syringe pumps infuse the three aqueous streams and the oil stream into the microfluidic device. The microfluidic device is placed on top of a transparent moveable stage between the illuminator and sCMOS camera attached to the computer used for data collection. The transparent cooling jacket rests on top of the device and is connected to the recirculating chiller. The crystallized drops and oil are collected in a waste beaker after exiting the device.

several competing effects related to droplet generation and residence time. The primary and most obvious consideration is the desired residence time within the serpentine channel. A mass balance in the fixed volume of the single-inlet single-outlet residence channel indicates that the volume flow rate of fluid being infused into the channel will equal that of fluid exiting the channel. Dividing by the cross-sectional area of the rectangular channel gives the average velocity of the fluid within the channel. From the known length of the residence channel, a simple relationship can be formed between the total infusion flow rate and the total residence time,

$$\dot{Q}_{net} = \dot{Q}_{protein} + \dot{Q}_{salt} + \dot{Q}_{buffer} + \dot{Q}_{oil}, \quad (1)$$

where \dot{Q}_{net} is the total volumetric flow rate being delivered by the syringe pumps. The terms on the right-hand side of the equation indicate that when multiple inlet streams come together at a junction, the total volumetric flow rate is given by the sum of the volumetric flow rates of the inlet streams of protein solution, salt solution, buffer solution, and the continuous phase oil. The average velocity \bar{V} in the downstream channel is then given by

$$\bar{V} = \frac{\dot{Q}}{A} = \frac{\dot{Q}}{w*d}, \quad (2)$$

where A is the cross-sectional area given by the product of the channel width w and the channel depth d . The residence time t_r can be evaluated using the average velocity and the fixed channel length L ,

$$t_r = \frac{L}{\bar{V}}. \quad (3)$$

It is desired to maximize the residence time, since that is the total amount of time that the drops will be observable during the experiment. One way to do that is to simply run the syringe pumps as slowly as they can operate. However, since the T-junction for forming droplets is in line with the residence chamber, the role of flow conditions on droplet generation must also be accounted for [35]. In a T-junction geometry, the ratio of the flow rates of the oil and aqueous streams has the largest impact on the droplet volumes. This ratio combined with the operational range of the syringe pumps limits how low the volumetric flowrate of a given fluid stream can go. The total flow rate also determines the capillary number at which droplets are formed. The capillary number affects whether the T-junction operates in the squeezing or dripping regimes of droplet formation, which also impacts the resulting size of the droplets. In addition, syringe pumps function by way of discrete motion of a stepper motor to advance the plunger inside a syringe at a constant velocity. At very slow speeds near the operating limit of the pump, the discrete motions can introduce unsteadiness to the flow rate.

The flow rates given above result in an oil:aqueous flow rate ratio of 4:1 and a capillary number of $Ca=109$, which has a predicted droplet volumes of $12.2 \pm 3.8 \text{ nL}$. The compositions of the incoming streams result in droplets containing an initial composition of protein and salt given in **Table 2-1** for each of the types of devices considered here. The total volumetric flow rate results in an average velocity in the temperature controlled serpentine channel of 24.6 mm/min . The droplets fill the cross-sectional area of the channel, and this confinement leads them to travel at a speed comparable to the average velocity. Images were collected every 500 ms at several

locations along the channel. After image collection, a custom MATLAB script processes the images to determine where nucleation events have occurred.

Table 2-1 lists the experimental conditions and solution compositions considered in each of the droplet microfluidic approaches to solution crystallization. The concentrations of protein and salt resulting from mixing of the inlet streams are shown along with the pH, the temperature at which crystallization is induced, and the supersaturation, defined as $\sigma = \ln\left(\frac{C}{C_{sat}}\right)$ where C_{sat} is the saturation concentration of Lysozyme.

Table 2-1 Experimental Conditions and Solution Compositions

Method	Lys (mg/mL)	NaCl w/v%	pH	C _{sat}	Temp °C	σ
Static Array	30	5	4.8	0.45	5	4.2
Static Array	25	5	4.8	0.51	5	3.9
Continuous	30	3	4.8	1.82	8	2.8
Continuous	20	2	4.8	4.04	8	1.6
Continuous	20	3	4.8	1.81	8	2.4

Classical Nucleation Theory

Nucleation is the phase change event in which a new solid crystal phase forms from a species dissolved in a solution. Classical nucleation theory (CNT) is one of the most common models used to describe nucleation events, and it is based on the assumption that solute monomers collide principally due to Brownian kinetics, and that free energy changes involved in forming a nucleus can describe the conditions for their formation. Before nucleation takes place, dissolved members of the crystallizing species driven by Brownian motion occasionally collide with one another as shown in **Figure 2-3**. If this collision is sufficiently energetic, the two particles will form the beginning of what is called “a pre-critical nucleus”. The interface created between the pre-critical nucleus and the surrounding solution has an energy cost that increases with the surface area of the nucleus. There is also a reduction in free energy for the solid crystal state compared with the dissolved state, where the free energy change decreases with increasing volume of the nucleus [21]. The interaction between these two energy

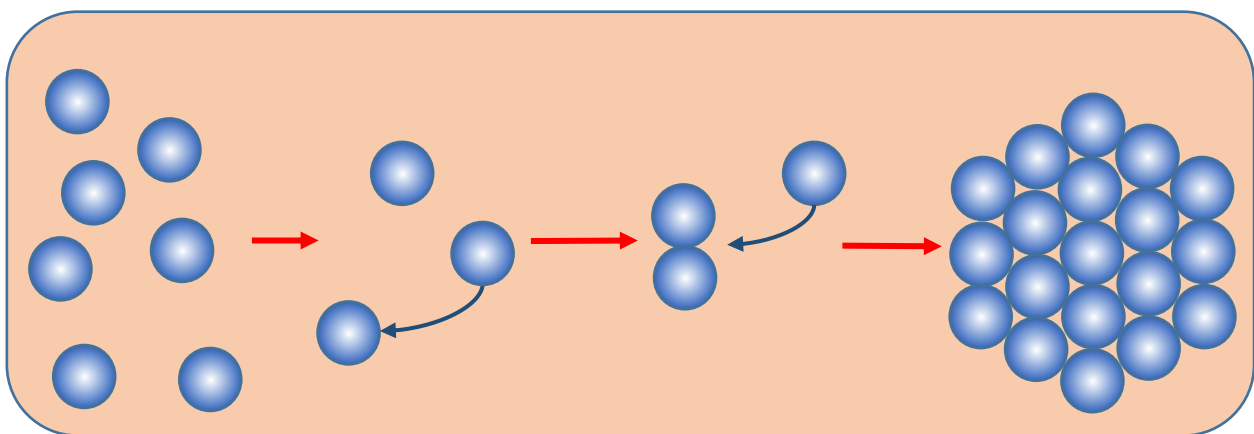


Figure 2-3 Classical Nucleation Theory. Solutes driven by Brownian motion occasionally collide with enough energy to overcome repulsive intermolecular forces. At a critical radius the energy required to create the new interface between the solid and liquid phases is balanced by the reduction in free energy due to adding volume to the solid crystal. Exceeding this radius is referred to as a nucleation event because all subsequent additions to the nucleus reduce the free energy of the system and therefore occur spontaneously.

changes is shown in **Figure 2-4**. When the nucleus grows sufficiently large it passes a critical radius and it becomes energetically favorable for more particles to attach to it. The transition beyond the critical radius is the point in the process that we call nucleation [22]. There is then a period of growth before the crystal becomes observable under microscopy [76]. Classical Nucleation theory does not always provide accurate predictions but is a useful starting point for understanding nucleation [77].

The driving force behind nucleation is supersaturation, or the ratio of the solute concentration to its saturation concentration. Supersaturation is a function of

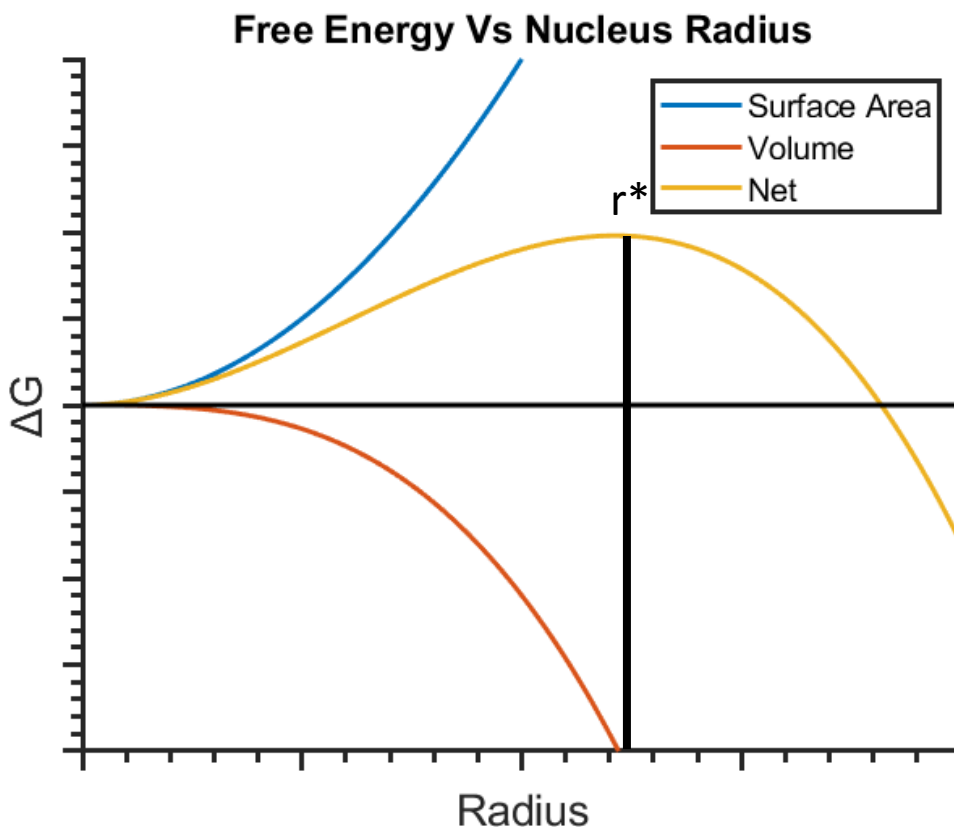


Figure 2-4 Free Energy Change versus Nucleus Radius. The change in free energy of the system, ΔG , is plotted against the radius of the growing nucleus. When the radius exceeds the critical radius r^* the slope of ΔG becomes negative and the process of adding more molecules to the nucleus becomes spontaneous. When the nucleus exceeds this critical radius, we call it a nucleation event.

concentration and temperature, and also can depend on other additives to the solution that influence the intermolecular forces. Salts like NaCl are commonly added due to their effect on electrostatic repulsion [25]. It is common for proteins to acquire a surface charge when dissolved in water, and this surface charge will be equal for all identical proteins in the same solution [26]. Thus the dissolved solutes will experience electrostatic repulsive interactions. The range over which this repulsive force acts is defined by the Debye length, which is affected by the ionic strength of the solution. The addition of salts reduces the Debye length due to the formation of an electrical double layer, where ions in solution act to screen the charge of the solute [27]. The reduction in electrical repulsive forces results in a lowering of the energy barrier to solute collisions. For this reason, it is common for salts to be added to nucleating systems to increase the supersaturation and thereby increase nucleation rates.

Since the system is driven by Brownian motion, the rate of nucleation events within a given volume is most accurately described in terms of a probability function [21] given by

$$\frac{dP_0(t)}{dt} = -JP_0(t) , \quad (4)$$

where $P_0(t)$ is the probability of finding zero crystals in a volume V at time t , and J is the nucleation rate. If the nucleation rate and volume are constant, this differential equation is easily solved, yielding an exponential decay of the probability of finding zero crystals,

$$P_0(t) = e^{-J^Vt} \quad (5)$$

Equation (5) shows that if one can measure the probability as a function of time, then the nucleation rate J can be extracted, and then its value can be correlated to material properties and experimental conditions.

In physical systems for which CNT accurately represents the nucleation process being observed, the value of J can be determined from expressions of the free energy changes that arise from creation of new interface and increased nucleus volume. In terms of physical properties of the nucleating materials, J can be expressed as

$$J = \rho_N Z j * e^{-\frac{\Delta G^*}{k_B T}} \quad (6)$$

Where $\rho_N e^{-\frac{\Delta G^*}{k_B T}}$ represents the number of critical nuclei per unit volume, with ρ_N representing the number density of nucleation sites and $e^{-\frac{\Delta G^*}{k_B T}}$ representing the energy barrier relative to thermal energy. Z is the Zeldovich Factor, which describes the probability that a nucleus at the critical radius will proceed forward into a crystal or dissolve back into solution [23]. The rate at which new particles attach to the nucleus, j , is defined as

$$j = 4\pi\rho DR^* e^{-\frac{\Delta F}{k_B T}}, \quad (7)$$

where ρ is the number density of monomers, D is the diffusion constant of dissolved monomers, R^* is the radius of the critical cluster, and ΔF is the activation energy for the addition of a monomer to an existing cluster. Since the physical parameters represented within J are not always easy to determine, it is common to condense the expression for J into a form that has two fitting parameters, A and B , and to express concentration in dimensionless terms as the supersaturation σ of the solution [24], where

$$\sigma = \ln\left(\frac{C}{C_{sat}}\right) = \frac{\Delta\mu}{k_B T} , \quad (8)$$

$$\frac{B}{\sigma^2} = \frac{\Delta G^*}{k_B T} , \text{ and} \quad (9)$$

$$J = ACe^{-B/\sigma^2} , \quad (10)$$

where C is the concentration of the nucleating species and C_{sat} is the saturation concentration. When the probability function $P_0(t)$ defined in Equation (4) is plotted on a log-linear graph, J represents the slope of the distribution. When concentration, volume, and temperature are constant, J will also be constant and $P_0(t)$ will be linear on a log-linear graph. If one or more parameters is not constant, then the instantaneous slope J will change with time, and the probability distribution will not be linear on a log-linear graph.

While there are macroscopic techniques available to measure nucleation rates, droplet based microfluidic systems allow for the direct observation of the nucleation probability distribution by creating large numbers of isolated drops which each serves as an independent reaction vessel [28]. In droplet based microfluidic systems the value of $P_0(t)$ is directly equivalent to the fraction of drops without crystals in them at a given time, and V is the drop volume. Thus, determining the probability function is as simple as capturing images of a set of drops over time and counting those without crystals divided by the total number of drops.

One method that has previously been used to determine the nucleation probability function involves creating a static array of droplets containing a solution of a crystallizing material [29-32]. A static array uses the microfluidic geometry to trap or

stabilize droplets in a grid within the device allowing the droplets to be observed over time as the nucleation process takes place. This method allows for the observation of hundreds to thousands of drops over the course of an experiment.

Results

Crystallization in Static Arrays

Crystallization was induced in two separate ways in the static array devices. In the first method, supersaturation is increased by reducing the temperature suddenly and

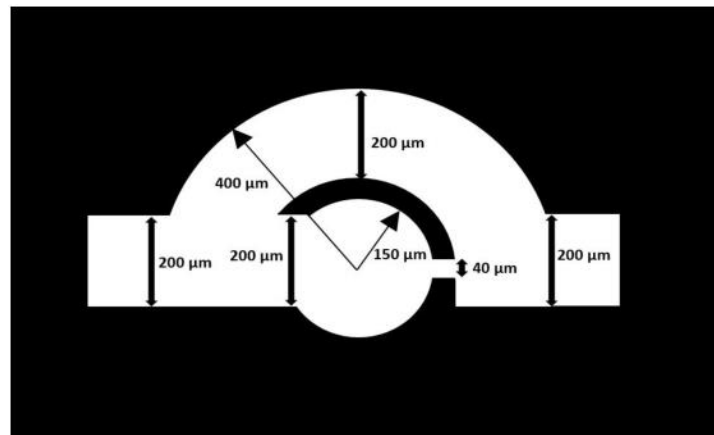


Figure 2-5 Droplet trap in a static array microfluidic design. Schematic diagram of a single microfluidic droplet trap along with relevant device dimensions. The channel depth is fabricated to be 100μm.

maintaining a lower temperature of the device and the droplets confined within it. In the second method, supersaturation is increased directly by allowing the droplets to dehydrate, effectively increasing the solute concentration over time. The same device design was used for both methods of inducing crystallization.

For the static array experiments, a grid of store-and-create microfluidic droplet traps was used to generate the droplets. Relevant dimensions for one trap can be seen in **Figure 2-5**. The entire grid contained 720 identical traps, with 45 identical traps connected in series, and 16 parallel rows. In the droplet trap, the $40\ \mu\text{m}$ exit channel attached to the downstream side of the round trap prevents the solution-filled droplet from leaving the trap as a result of the increased capillary pressure required to deform the droplet interface enough to push the droplet through the smaller channel. Droplet generation in the static array is shown in the images presented in **Figure 2-6**. Initially the supersaturated solution is infused into the device filling all of the capillary traps, and then a later infusion of oil with surfactant clears out the bypass channels, creating isolated droplets in the traps. In the geometry shown here, the average drop volume formed was $9.3 \pm 0.5\ \text{nL}$, with a maximum recorded volume of $11\ \text{nL}$ and a minimum of $7.5\ \text{nL}$. This represents a polydispersity of 5.3% which is similar to previous reports for other static array devices [74].

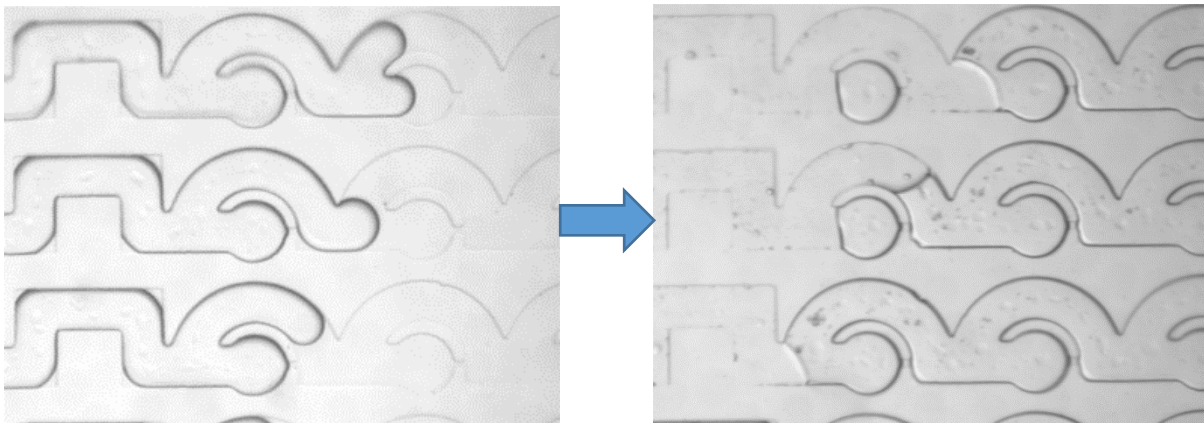


Figure 2-6 Drop Generation in static arrays. Left: The aqueous solution is infused into the device, filling all the traps and bypass channels. Right: Oil is infused to displace the solution from the bypass channels, creating drops in the traps.

Temperature control

A custom-built water jacket connected to an external chiller was used to control the temperature within the static array microfluidic device. Temperature measurements were collected at two locations within the device: at the entrance to the residence channel and at its exit. Temperature was separately measured by thermocouple at various locations within the device, and it was found to be uniform within the 0.1 °C accuracy of the measurement device. These temperature measurements were taken during all trials to ensure consistency between runs.

Humidity Control

PDMS is permeable to many gasses and liquids, including water. This is normally regarded as a detriment to its utility but in this case we use it to advantage. By placing the device in a low humidity environment, a driving force is created for water to leave the drops. The solute remains within the droplets, thereby causing the solution concentration to increase as the droplet volume decreases. Once the droplet contents exceed the saturation concentration, crystallization is induced. In the present

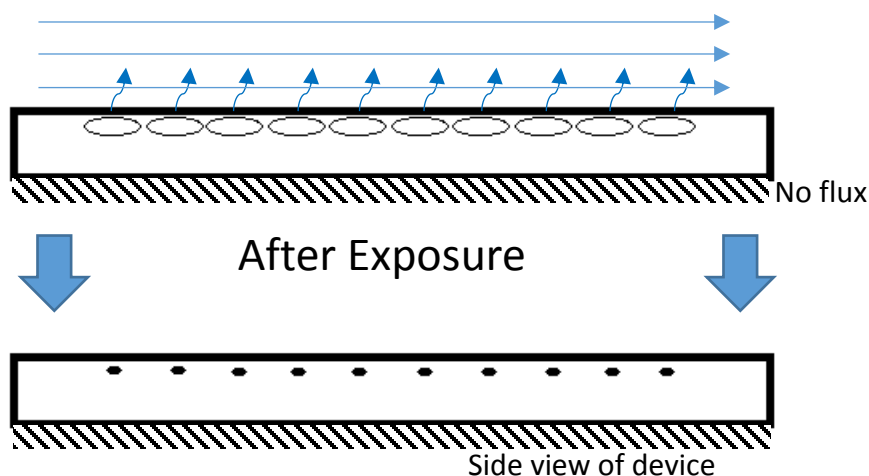


Figure 2-7 Dehydration of microfluidic droplets. Droplets in a static array microfluidic device are exposed to a low relative humidity through a thin layer of PDMS. Water diffuses out of the drops and evaporates into the surrounding air over time. As the droplets dehydrate the supersaturation increases until nucleation occurs.

experiments, the low humidity environment is created by placing the loaded static array device into a small acrylic case. Dry air flows through the case, maintaining a low relative humidity within the case, but external to the microfluidic device. The low relative humidity creates a concentration gradient of water, driving water to diffuse out of the drops, through the surrounding oil phase liquid in which water is sparingly soluble, through the permeable PDMS, and finally evaporating into the surrounding dry air. The dehydration process is illustrated schematically in **Figure 2-7**.

Crystal detection

One advantage of the static array device is that the droplets remain stationary throughout the crystallization process, which can take place on timescales of the order of hours. Once the temperature is decreased, the droplets are monitored by capturing images every 2 minutes. Once nucleation has occurred, crystals rapidly become visible

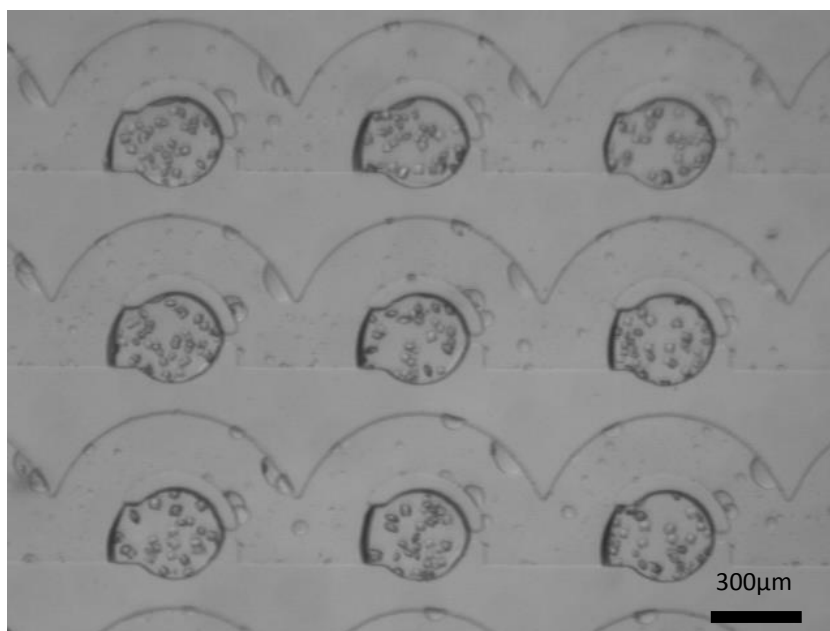


Figure 2-8 Crystals formed in a temperature step static array experiment. The image shows a set of droplets trapped in a microfluidic static array device after temperature induced nucleation has occurred. The resulting crystals are visible within the droplets.

in the magnified view of the microscope, and thus the fraction of droplets containing crystals at each time step can be counted. This fraction can be used to construct the probability function $P_0(t)$. To improve the detection accuracy, a background image is captured before the temperature is decreased and is subtracted from each subsequent captured image. The background-subtracted image is then checked for the presence of new objects at each of the known drop locations. An example image of droplets containing crystals is shown in **Figure 2-8**. The process is monitored until all the drops exhibit crystal formation.

Detecting crystal formation in a static array with dehydration driven nucleation is carried out in nearly the same manner as for the temperature step experiments. We made slight modifications to the image processing routine to account for the change in droplet size over time. The color contrast between crystallized and uncrystallized drops

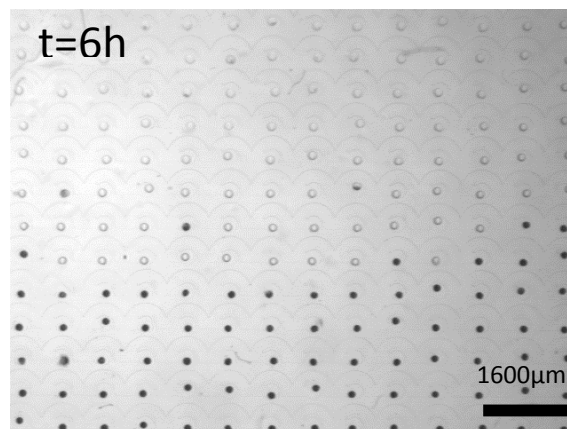


Figure 2-9 Crystals formed in a static array as a result of droplet dehydration. The image shows the entire observed grid of static droplets after partial crystallization has occurred due to dehydration of the droplets. The droplet color changes dramatically when nucleation occurs. The drops become significantly more supersaturated, resulting in rapid and complete solidification of the droplet contents.

shown in **Figure 2-9** facilitated automated crystal detection using a custom MATLAB script.

Observed Probability Distributions

For the static array experiments a total of eight crystallization trials were completed, four imposing a temperature step, and four imposing dehydration of the drops. Two different initial concentrations of lysozyme were used in each type of experiment, 30mg/mL and 25mg/mL, so that two trials for each method and each concentration were completed. **Figure 2-10** below shows the experimental results for the temperature step experiments. The y-axis represents the fraction of drops without observed crystals versus time on the x-axis. Each marker represents a captured image,

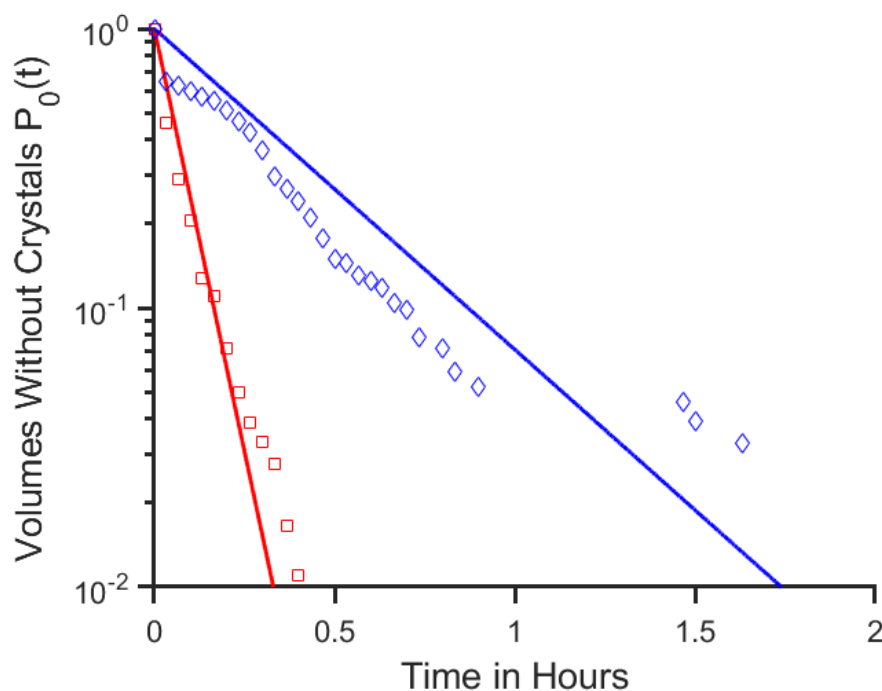


Figure 2-10 Probability that crystallization has not yet occurred versus time for temperature step crystallization. Solid lines indicate the model predictions for the fraction of uncrystallized drops; diamond and star symbols denote measured data points for the lower and higher concentrations, respectively.

with the y-axis magnitude equal to the fraction of drops without crystals at that time. The continuous lines represent model prediction curves generated in MATLAB using the analytical solution for probability distribution obtained using the parameters reported by Akella *et al.*, and assuming all droplet volumes were equal to the mean volume of 9.3 nL [21]. The measured data points are obtained by visually counting the number of drops that do not yet exhibit visually observable crystals in each frame, and then calculating the fraction of uncrystallized drops within the entire array. Experimental results conducted at each concentration show a decrease in the fraction of uncrystallized drops that is roughly linear on a log-linear scale, with the experiment conducted at higher solution concentration exhibiting a steeper slope than the probability function at the lower concentration. The model predictions are computed using equation (5) for the probability function $P_0(t)$ which is expected to decrease exponentially with time, with the higher concentration yielding a shorter time to crystallization. The model parameters are summarized in **Table 2-2**.

Table 2-2 Static array nucleation model parameters. Taken from Akella *et al.* [21].

	30 mg/mL	25 mg/mL
Temperature T	5.0 °C	5.0 °C
Fitting parameter A	$6.2 \times 10^7 \text{ mg}^{-1} \text{ s}^{-1}$	$6.2 \times 10^7 \text{ mg}^{-1} \text{ s}^{-1}$
Fitting parameter B	293.2	293.2
Supersaturation σ	4.25	4.10
Nucleation rate J	$416.3 \text{ ml}^{-1} \text{ s}^{-1}$	$108.5 \text{ ml}^{-1} \text{ s}^{-1}$

At the higher concentration of 30 *mg/mL*, the experimental probability function correlates closely with the model predictions. At the lower concentration of 25 *mg/mL*, the discrepancy between the experimental results and the model predictions is larger, with the model overpredicting the experiment at a given time. One possible contributing factor to this deviation at lower concentrations is that variability in droplet size has more time to impact the crystallization process during slower crystallization processes. Another consideration is the possible presence of contaminants that can act as heterogeneous nucleation sites, allowing for quicker than expected crystallization in some drops. Contamination could help explain the rapid crystallization seen at early times in the 25 *mg/mL* trial. This rapid initial decrease may also have been present in the 30 *mg/mL* trial, but may have been masked by the overall faster rate of nucleation. Images were recorded every 2 minutes during all static array experiments which gives us the maximum uncertainty smaller than the markers in the horizontal axis. Quantifying uncertainty in the fraction of drops crystallized poses significantly more challenging. Primarily the droplets volume polydispersity affects the nucleation rate because as seen in equation (5) volume directly contributes to the value of $P_0(t)$. With a range of drop volumes, in this case 9.3 ± 0.5 *nL*, the larger droplets will have lower values of $P_0(t)$ compared to model predictions and vice versa for droplets smaller than average. Additionally, whether or not a crystal is detected is dependent on a number of factors including the resolution of the image, the position of the crystal within the drop, and the orientation of the crystal. For this reason calculating a numerical value for the uncertainty in the vertical direction proves to be quite difficult. A minimum uncertainty of

a one droplet being misidentified would give a value of $1/N$, which in our case becomes a significant magnitude as the number of crystallized drops approaches N .

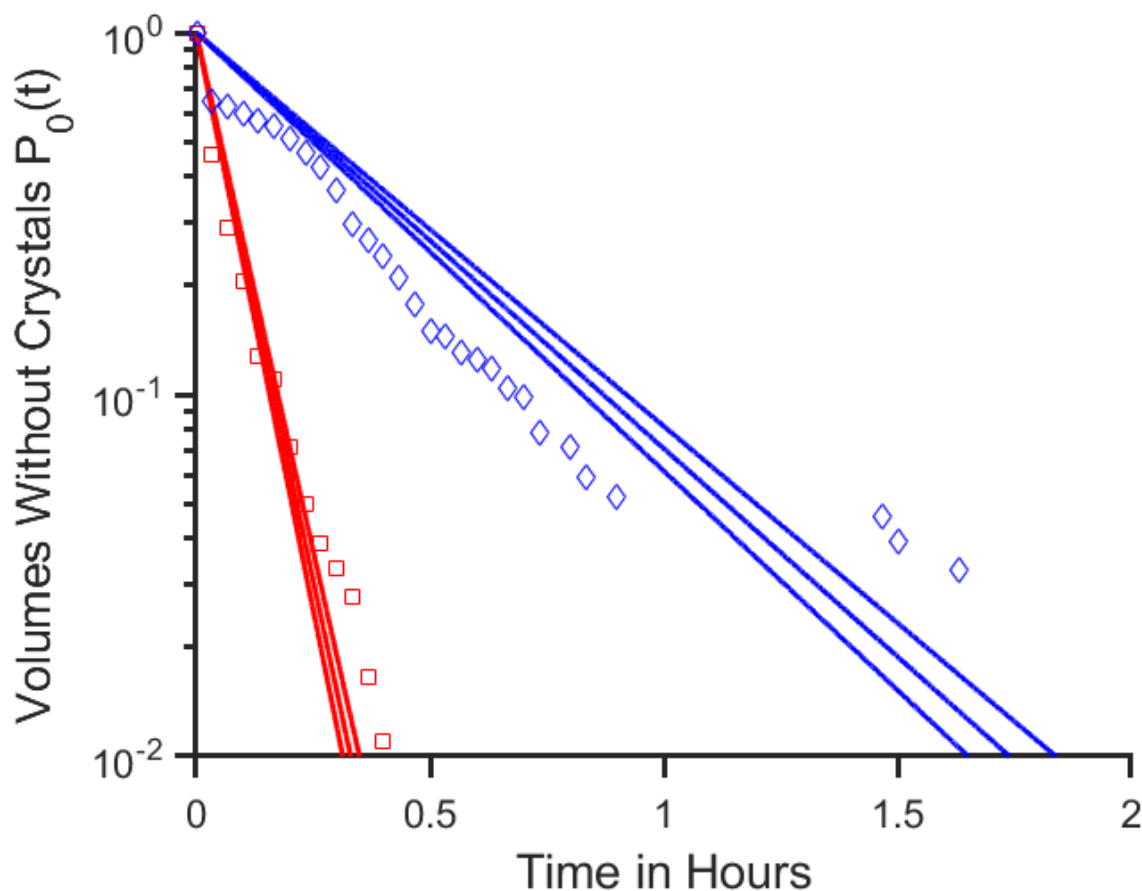


Figure 2-11 Volume variations in static arrays. Plotted is model predictions for both concentrations with the mean volume as before with two additional model predictions with the a volume one standard deviation above and below the mean.

As seen in **Figure 2-11** the volume variability does not fully explain the differences between model predictions and the experimental results. When plotting the model predictions, represented by the solid red and blue lines, one standard deviation above and below the mean volume, the experimental results still lay outside of the expected range, with the lower concentration consistently below the model. These deviations may be due to crystallization that occurred during the loading process.

Nucleation induced by dehydration in static arrays

Supersaturation is also increased by removing water from the droplets over time, decreasing the droplet volumes and increasing the concentration of all components of the solution as shown schematically in **Figure 2-7**, and captured images show this in **Figure 2-12**.

Analysis of dehydration induced crystallization experiments is complicated by the simultaneous variation of volume and concentration. Volume varies nonlinearly with time, as the changing surface area affects the rate of mass transfer out of the drops.

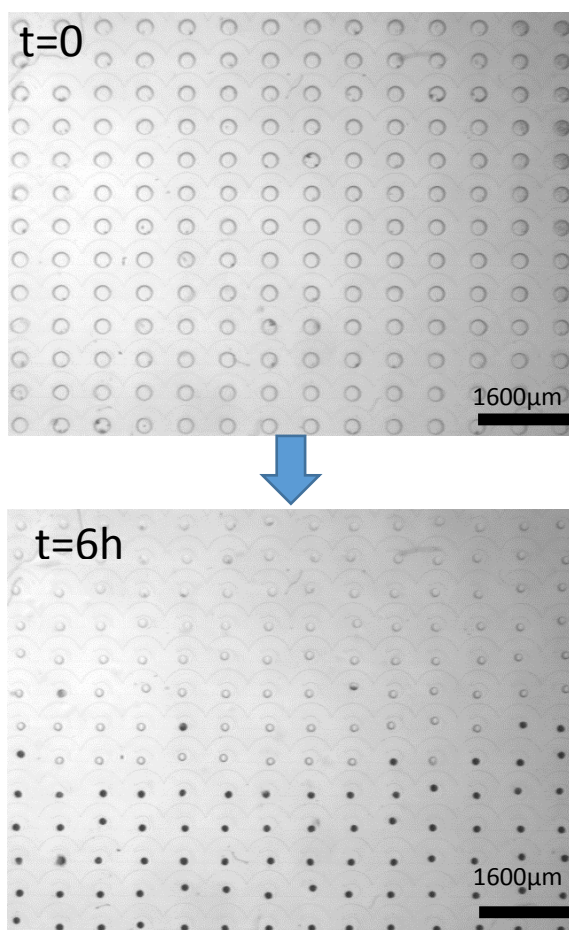


Figure 2-12 Crystal Detection in Static Droplet Arrays. A grid of 182 of the total 720 drops is observed over time. As water slowly leaves the drops by diffusing out of droplets and through the PDMS, the solution concentration within droplets rises until crystals form. The probability function is determined by counting the fraction of drops without observable crystals at each sampling time.

While the governing equation for the probability of crystallization is still expected to remain the same, the differential equation for $P(t)$ must be solved while incorporating the time dependence of volume and concentration, which renders $J(\sigma)$ no longer constant with time. The resulting nonlinear differential equation does not have an analytical solution, however it is easily solved by numerical integration. The resulting $P_0(t)$ curve is not expected to have the same form as is observed in a temperature step driven crystallization trial, since the rate of nucleation is expected to continuously increase as the droplet dehydrates.

In dehydration-induced nucleation the drop volumes and the supersaturation change with time, as water leaves the droplets but the solutes do not. These changes are shown in **Figure 2-13** where the left y-axis shows the predicted volume of the

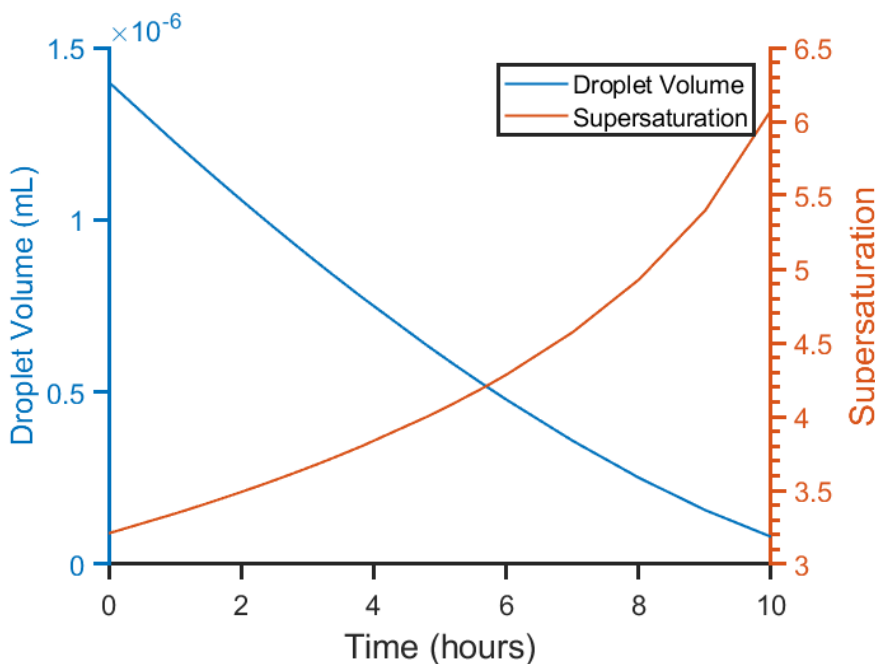


Figure 2-13 Droplet volume changes versus time. The left y-axis shows the drop volume as a function of time and the right y-axis shows the corresponding supersaturation within the droplet. As volume decreases due to evaporation of water, the concentration within the droplets increases, raising the supersaturation.

droplet versus time and the right y-axis shows the corresponding supersaturation at that time. Non-constant volumes and concentrations change equation (4) so that J and V are now functions of time, given by

$$\frac{dP_0(t)}{dt} = -J(t)V(t)P_0(t) \quad (11)$$

$$V(t) = \frac{4}{3}\pi(r_0^2 - 2\alpha * t)^{\frac{3}{2}} \quad (12)$$

$$C = \frac{C_0 * V_0}{V(t)} \quad (13)$$

Where r_0 is the initial radius of the drop, α is a constant volume change rate to be $5.7 \pm 9.7 \times 10^{-14} \text{ m}^2/\text{s}$. The resulting differential equation must be solved numerically in conjunction with mass transport equations governing the volume and concentration changes as water leaves the droplets.

Model predictions for a dehydration-induced crystallization experiment are shown in **Figure 2-14**. Similar to the temperature-induced experiments, the y-axis shows the fraction of drops without crystals versus time on the x-axis. All times earlier than 6 hours are omitted from the graph as no nucleation took place during those times. Each marker represents a recorded image with its y-axis magnitude equal to the fraction of drops without observable crystals at that time. The fraction of drops without crystals decreases non-linearly on a log-linear scale, with the higher solution concentration exhibiting crystals earlier and at a faster formation rate than the measurements at lower concentration.

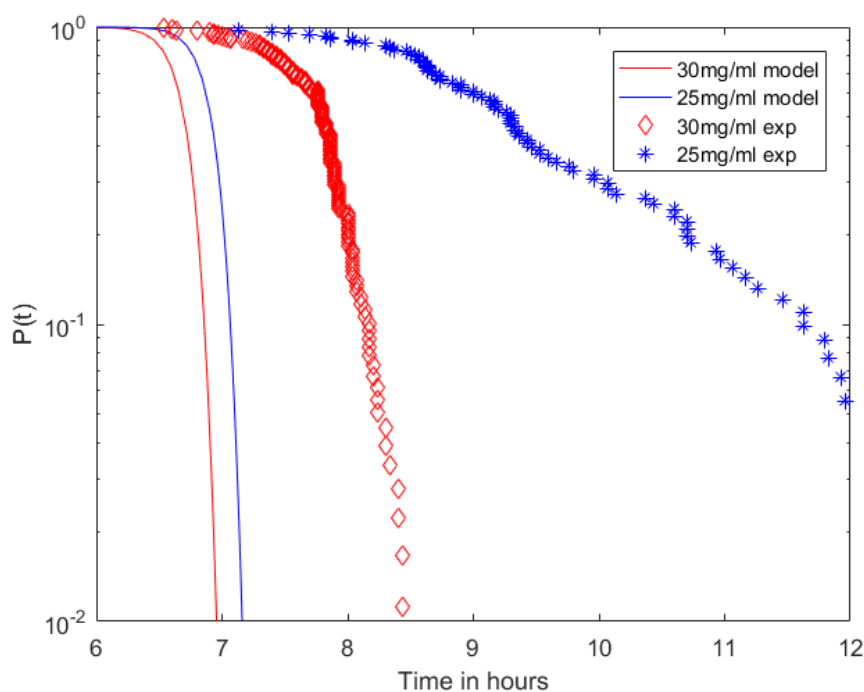


Figure 2-14 Probability function versus time for dehydration-induced crystallization. Solid lines indicate model predictions for the fraction of uncrystallized drops; diamond and star symbols denote measured data for the higher and lower concentrations, respectively.

Model prediction curves are generated in MATLAB using the subroutine ode23s to solve equation (11) for the probability versus time. The volume change for all drops was assumed to be uniform and equal to the average value of α . The model curves also exhibit a nonlinear and faster-than-exponential decrease, with similar qualitative shape to the experimental observations. The model predicts an earlier time to achieve complete crystallization of all the droplets than observed in the experiments. The experiments show that nearly all the droplets have formed a crystal within a range of 8 to 12 hours, with the lower concentration extending to the maximum time and the higher concentration droplets crystallizing within about 8.5 hours. The model by contrast predicts that crystallization is complete within about 7 hours for both concentrations,

although the higher initial protein concentration results in slightly faster overall crystallization time compared with the lower concentration.

The deviation between the model predictions and the experimental results is greater for the dehydration induced nucleation experiments compared with the temperature induced experiments. The predicted time to completion is approximately 1.5 hours shorter than that observed in experiments for the 30mg/mL trial, and the 25mg/mL trial took over 5 hours longer to reach complete crystallization compared with its corresponding model prediction. The discrepancies between the measured data and the model predictions most likely arise from at least two factors: the variability in drop sizes, $9.3 \pm 0.5 \text{ nL}$, and the variability in dehydration rates, $5.7 \pm 9.7 \times 10^{-14} \text{ m}^2/\text{s}$, within the same static array device. All data was collected from drops near the center of the device, ensuring that for each set of 182 droplets observed out of the possible 720, at least 3 rows separated them from the edge of the device. Despite this, the observed droplets that were closer to the edge of the device showed higher dehydration rates compared to droplets closer to the center of the device. This spatial variation creates dehydration gradients within the device, which in turn causes interior drops to dehydrate and then crystallize more slowly than exterior drops. This variability was difficult to control, and help motivate the switch to the continuous flow microfluidic device discussed in the next section, in which each droplet can experience the same history and conditions while crystallizing.

Figure 2-15 shows the dehydration nucleation data from the 30 *mg/ml* trial with the model predictions for the mean dehydration rate as well as model predictions for dehydration rates one standard deviation above and below the mean. This large change

in the dehydration rate combined with the observed systematic behavior of the droplets closer to the edge nucleating before drops more centrally located to the device fully explains the variability in nucleation rates observed.

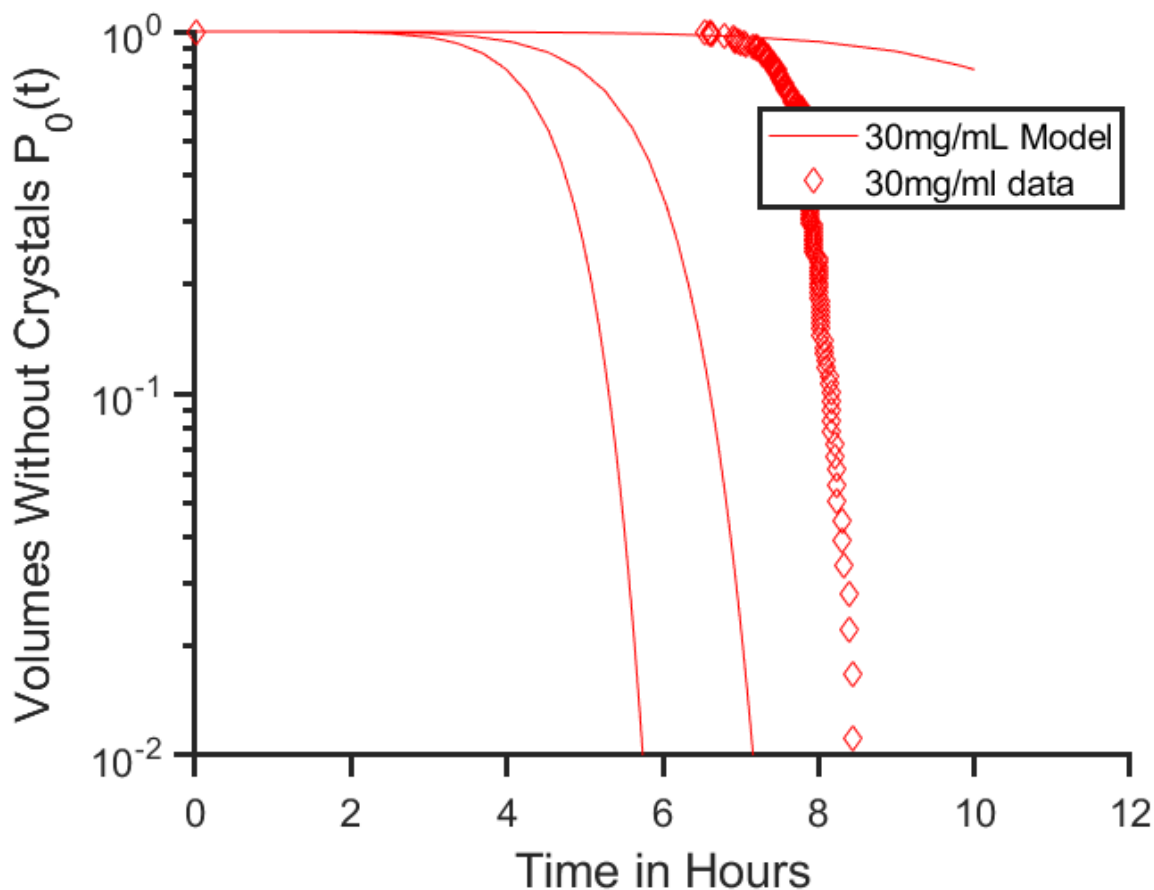


Figure 2-15 Dehydration variability in static arrays. Plotted is model predictions for the 30 mg/mL concentration trial with the mean dehydration rate as before with two additional model predictions with the a dehydration rate one standard deviation above and below the mean.

The results presented here illustrate several advantages and drawbacks of using static arrays of microfluidic droplets to study crystal nucleation from solution. Static arrays can retain droplets for extended periods of time to allow for complete crystallization of all observed droplets. Both fabrication and operation are relatively

simple, allowing these devices to be utilized in a wide range of situations. However, droplet may experience non-uniform experimental conditions arising from their locations within the device, leading to non-uniform dehydration rates. The fixed number N of drops that can be held within the device is the maximum number of independent experiments that can be observed, which implies that the smallest measureable value of P_0 is $\frac{1}{N}$. If this resolution is not sufficient to observe a phenomenon of interest, such as the rate of slowly nucleating species in the same solution as quickly nucleating ones, then, the resolution for P_0 can only be improved by increasing the number of droplets per device. Increasing droplet numbers can be done by either increasing the size of the device or by increasing the density of the droplet traps within the same device footprint. Both of these approaches have limitations. Increasing the device size presents challenges with efficiently filling the droplet traps and with visualizing all of the droplets. Increasing the density of traps presents challenges with fabrication, since too little space between traps can lead to challenges with properly sealing the channels. Thus, neither approach can substantially increase the number of drops. If greater resolution in the measurement of P_0 is desired, then a different device style is required. In the next section, we describe the development of one possible alternative design using a continuous train of discrete droplets. This design permits the observation of a much larger number of independent events, but also requires significant adjustments to the calculation of P_0 and limits the amount of time over which crystallization can be observed.

Continuous Droplet Trains

The design of the microfluidic device used for generating static arrays showed many areas with room for improvement. Specifically we aimed to improve the number of drops that could be observed in a single experiment, eliminate the spatial variation's effect on the conditions the drop experiences, and integrate on chip mixing so that the solution would not begin to crystallize before droplet formation under any circumstances. Solving all of these problems simultaneously in a static array style device proved to be quite challenging, and eventually a new perspective was needed. Inspired by microfluidic devices that generate a continuous train of droplets by utilizing a T-junction, we sought to implement a similar device to investigate nucleation [75]. The continuous flow style device could mix the protein and precipitant solution directly before droplet generation, and since each droplet traces the same path through the residence channel they would all experience the same history of experimental conditions.

For continuous drop train experiments, the design of the microfluidic device can be seen in **Figure 2-16**. This device was fabricated in PDMS with a constant channel

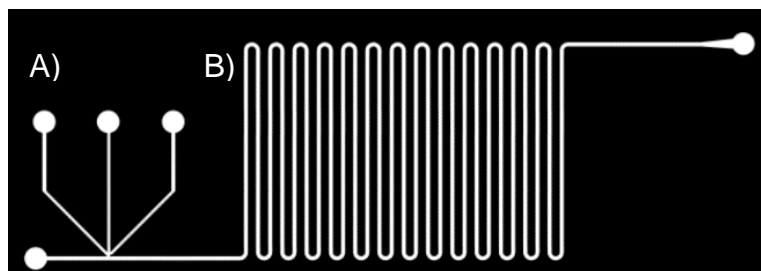


Figure 2-16 Continuous drop train microfluidic device design. The microfluidic device consists of two major sections. A) the T-junction drop generation nozzle combines the three aqueous streams as they meet the oil stream where the combined stream is pinched off into individual drops. B) The residence track retains the drops while exposing them to a fixed temperature below the saturation temperature to induce nucleation. The constant cross-sectional area and constant flow rates imply that droplet position and residence time are directly proportional to one another as seen in equation (14)

depth of 100 μm . All channels had a constant channel width of 200 μm . Aqueous streams were infused from the three vertical inlet channels, while the oil was infused into the horizontal inlet. All four inlet streams meet at a T-junction where droplets are formed and then flow into the residence channel for observation.

Drop size distribution

Droplet size distributions were recorded for each of the three solution concentrations listed in **Table 2-3** were measured. Droplets created for trials of 30 mg/mL lysozyme and 3 w/v\% NaCl had volumes of $10.8 \pm 3.3 \text{ nL}$, droplets created for trials of 20 mg/mL lysozyme and 3 w/v\% NaCl had volumes of $10.9 \pm 2.7 \text{ nL}$, and drops created for trials of 20 mg/mL lysozyme and 2 w/v\% NaCl had volumes of $15.0 \pm 5.6 \text{ nL}$. Although there is variation in these average droplet sizes, all three are within one standard deviation of one another and therefore can be considered nearly equivalent. Histograms of the measured volumes are shown in **Figure 2-17** and summarized in **Table 2-3**. Notable differences in the shapes of the volume distributions for each trial can be observed. The trial with the narrowest drop size distribution has the lowest lysozyme concentration, indicating that the protein itself may play a role in the drop generation process. It is possible that lysozyme could influence the interfacial properties of the drops by competing with the surfactant at higher concentrations [38]. This could also influence the wettability of the aqueous and oil streams at the T-junction.

Table 2-3 Drop volume distribution

Lysozyme (mg/mL)	NaCl (w/v%)	Mean volume (nL)	Volume St.Dev (nL)	PDI
30	3	10.8	3.3	30.5%

20	2	15.0	5.6	37.3%
20	3	10.9	2.7	24.8%

Table 2-3 shows that the polydispersity of the droplet sizes resulting from the continuous droplet train design is in the range of 25% to 37%. While microfluidic devices are capable of producing very monodisperse drops (as low as 1-3% polydispersity), the drop size distribution in the continuous drop train trials had higher polydispersity than expected. A contributing factor to this polydispersity is likely the slow flow rates used

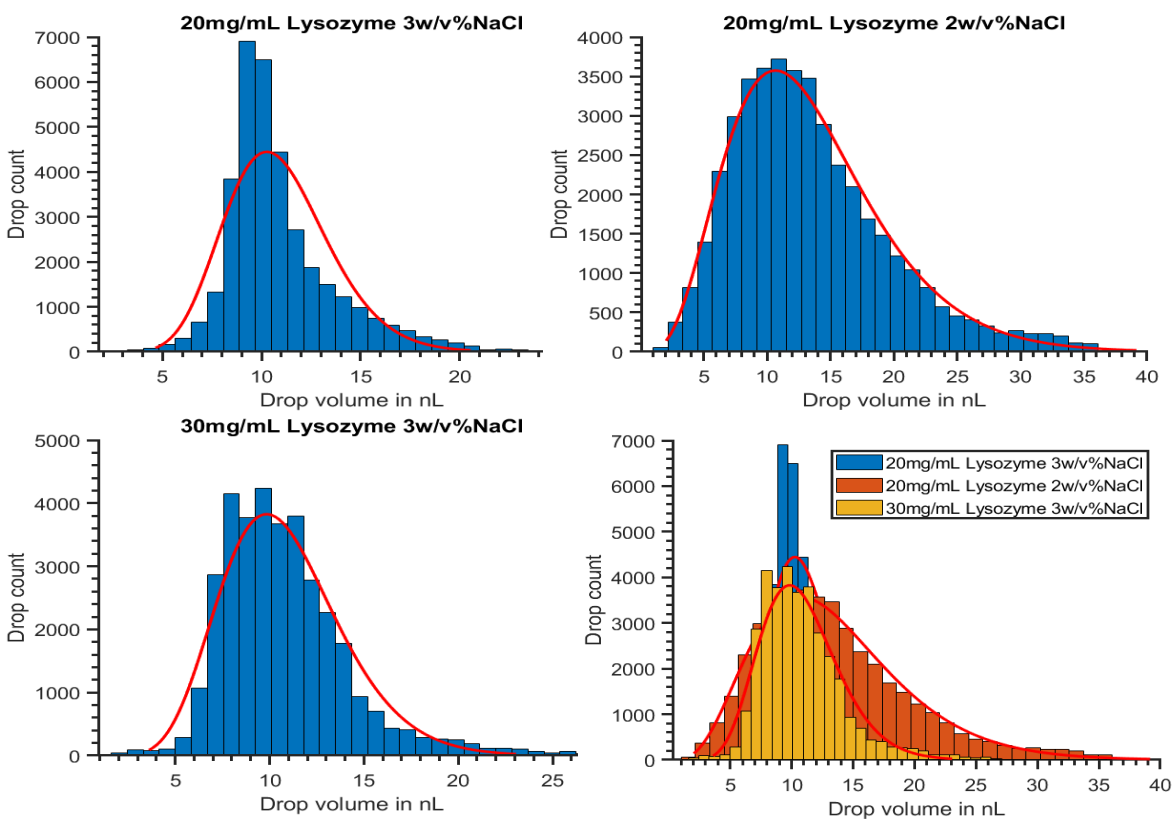


Figure 2-17 Drop Volume distributions. Each graph shows the number of drops within a particular volume bin for each of the three solution compositions tested. Notable qualitative differences can be seen between each population. Additionally long tails to the distribution show the presence of many drops well above the expected mean volume

within the devices. Slow infusion rates are necessary to maximize the amount of time a drop is observable within the residence channel. However, at slower flow rates the ratio of viscous forces to surface tension, known as the capillary number, is closer to unity [36]. At capillary numbers near unity, where viscous forces and capillary forces are comparable, some instability can occur in drop generation and lead to higher polydispersity. Another potential factor leading to run-to-run polydispersity is that PDMS devices are disposable and a new device is used for every experiment. Slight variations in the molding process may affect the drop generation. To avoid this in future, a reusable microfluidic device made from fused silica or other materials could be employed [37]. All three drop size distributions show the presence of a long tail at higher drop volumes, indicating a considerable population of larger than average drops.

Crystal Detection

When observing crystal nucleation events in the continuous droplet train experiments, we employ a different approach to constructing the probability function. Counting the number of drops that pass by a viewing region without crystallizing and the number that enter that same region already crystallized provides the fraction of drops with crystals at that location in the device. Since a given droplet position within the device corresponds to a specific residence time, carrying out these measurements at several locations along the residence channel helps construct the probability function versus time. The residence time is proportional to the steady infusion rates provided by the syringe pumps and the distance downstream of the T-junction where droplets are first formed. The residence time therefore corresponds to the incubation time of the droplet.

Figure 2-18 shows a schematic diagram of the image processing algorithm used for the crystal detection. An inverted microscope and attached high speed camera is used to capture an image of a portion of the device that contains several channel segments. The image is divided into separate regions that each include only one channel. The background is removed, leaving only the drops and any crystals within them. Edge detection is then used to identify the circular end caps of all the visible

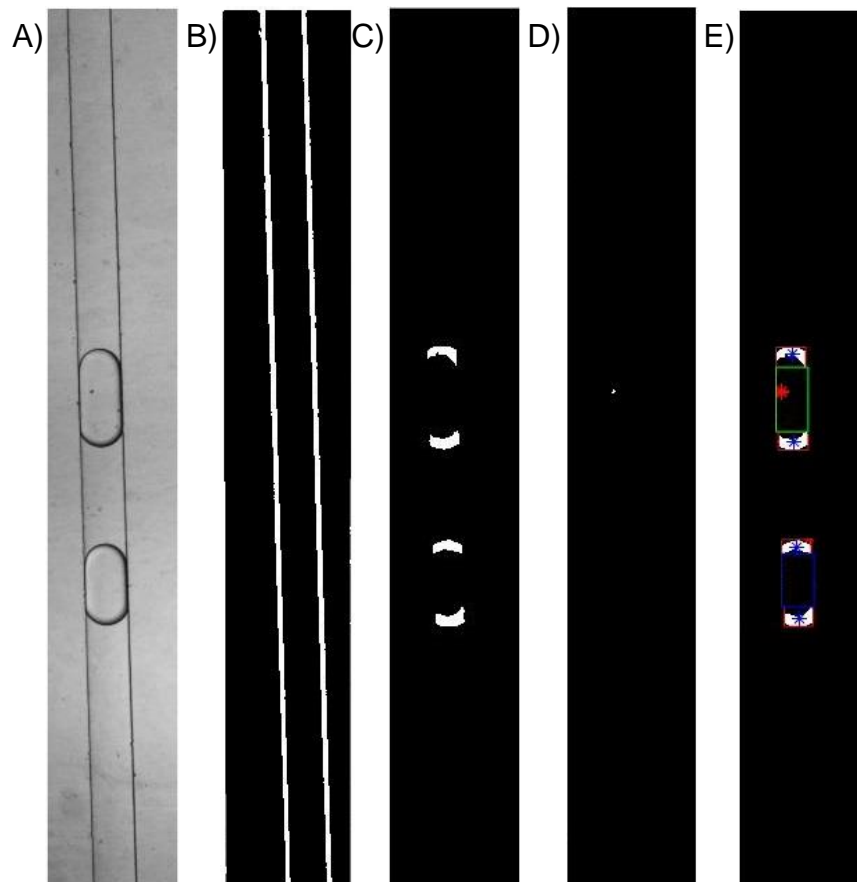


Figure 2-18 Crystal detection scheme. A) First the entire frame is subdivided to isolate an individual channel. B) then large background elements are identified and removed. C) Droplet endcaps are identified via edge detection, isolated and paired up to determine each drop position. D) small objects are detected to locate crystals that have formed. E) If the detected small elements are within the drop boundaries the drop is labeled as having nucleated, as indicated by the green box. Otherwise, the drop is counted as part of the total number of droplets.

drops, and these endcaps are paired up to define the boundary of the entire drop. Within this boundary, if an object is detected, this is counted as a crystal within the drop. If there are crystals present, the previous frame is also checked to determine when the crystal became visible. If there was no visible crystal in the previous frame and there is a crystal in the current frame, the algorithm notes this as a nucleation event occurring at the midpoint between the drop locations in each frame. At pixels upstream of the midpoint, the drop is counted as uncrystallized and at points downstream of the midpoint, the drop is counted as crystallized. If a drop enters and leaves the field of view without developing a new crystal, either because it had already crystallized or because it did not contain a crystal, the drop is simply counted appropriately for all points along the channel. Repeating this process for all recorded images builds up a list of all locations at which a droplet has crystallized and the total number of drops that have passed each point. From this we calculate the fraction of drops that have formed a crystal at each location along the channel and at the corresponding incubation time, which allows us to plot the overall probability function $P_0(t)$. The slope and linearity of the curve can be used to analyze the kinetics of the crystallization process as we have previously discussed for the static arrays.

Practical limits on the time resolution of crystal nucleation events occurring in drops moving through the continuous drop train device are imposed by the drop velocity, V_d , the drop length L_d , the time between frames captured by the observing camera, Δt_f , a conversion factor between pixels and distance, C_{pix} , and the frame rate of the camera f . Between each frame captured by the camera the drop will move a distance proportional to its velocity.

$$d_f = \frac{V_d}{f} \quad (14)$$

A drop must be fully within a frame before it can be accurately measured, as we cannot be certain that a drop does or does not contain a crystal if the whole drop is not in view. The time it takes for a drop to fully enter the viewing region from the first instant at which its leading edge enters the region is given by:

$$\Delta t_0 = \frac{L_d + (d_f - C_{pix})}{V_d} \quad (15)$$

If this time is greater than the time between frames, then it sets the minimum time resolution for the device; otherwise the time between frames determines the minimum time resolution. The minimum spatial resolution is defined by the minimum time resolution multiplied by the drop velocity.

$$\Delta d = \Delta t * V_d \quad (16)$$

For our system the values of the limiting time resolution are summarized in **Table 2-4**. Since $\Delta t_0 > \frac{1}{f}$ the limiting time scale for our measurements is the time it takes for a drop to fully enter the viewing region and not the frame rate of the camera recording the images. This is expected as the flow rates in our system are relatively slow.

Table 2-4 Time Resolution Parameters

L_d	149 μm	d_f	0.2 mm
V_d	24.6 mm/min	Δt_0	0.86s
C_{pix}	1.1 $\mu\text{m/pixel}$	Δd	0.2 mm
$\frac{1}{f}$	500 ms		

Observed Probability Function

For continuous drop train experiments, after analyzing the captured images using the algorithm described earlier and shown in **Figure 2-19**, probability distributions can be constructed by first counting the total number N of drops observed. The numerator of P_0 can be calculated by subtracting from unity a value of $\frac{1}{N}$ for each drop that enters the viewing region already containing crystals, and another $\frac{1}{N}$ for each nucleation event recorded prior to that region. Very few drops transition from containing no crystal to containing a visible crystal within a given viewing region since nucleation is a relatively rare event, and high magnification is needed to observe early nuclei. As a result, the major contribution to the probability distribution arises from the fraction of drops entering

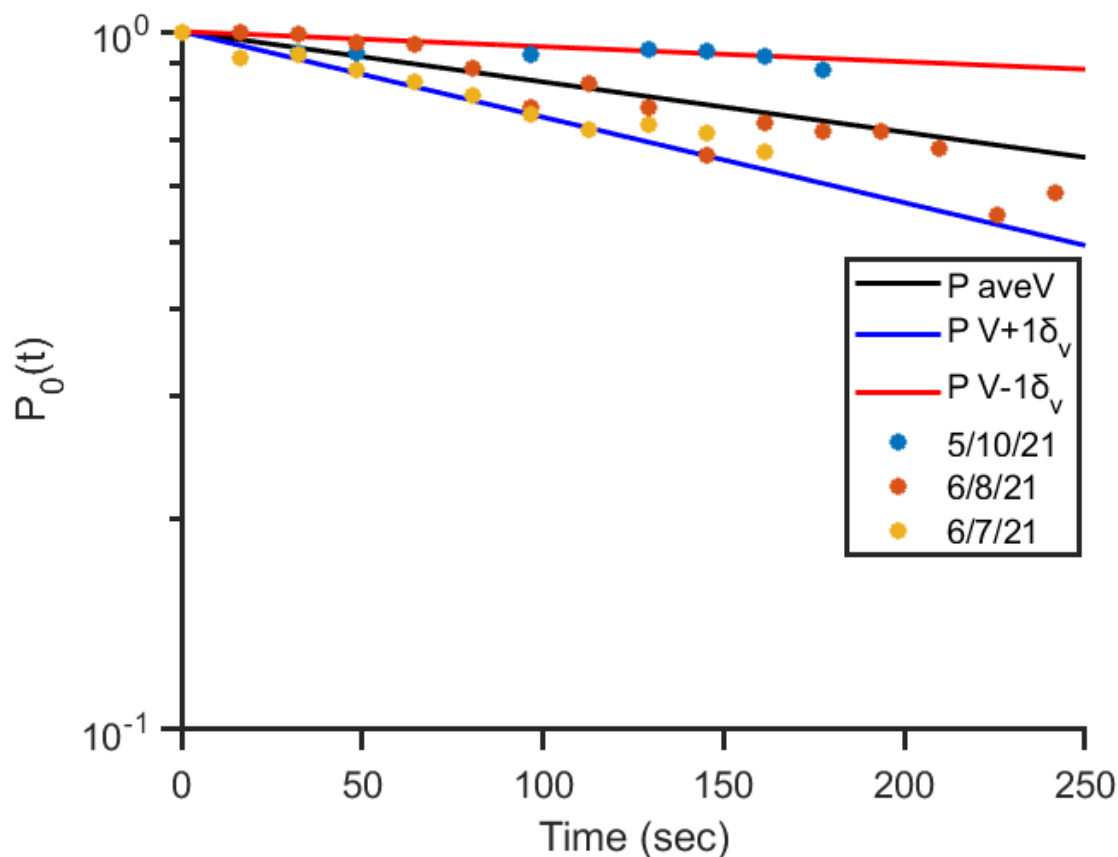


Figure 2-19 Observed probability function versus time for 30 mg/mL Lys 2 w/v% NaCl. The observed probability function is plotted versus time for a solution containing 30 mg/mL lysozyme and 2 w/v% NaCl. The black line represents a least squares linear fit to the log of the average probability from all the runs versus time, assuming the data follows a single exponential decay and the drop volume equals the average volume. The corresponding nucleation parameters were obtained from the fitted slope, and the blue and red lines were then constructed using volumes equal to one standard deviation above and below the average volume recorded over all three runs.

each viewing region, and the probability value is nearly constant within each region.

The measured probability function versus time for each of the three solution

concentrations are shown in **Figure 2-19** through **Figure 2-21**. In each figure, the y-axis represents the fraction of drops without crystals and the x-axis represents time. Each marker represents the fraction of droplets without crystals passing through a particular viewing region within the residence channel, whose position is related to time using the

constant volumetric flowrates and the cross-sectional area of the microchannel at that location. The black solid lines in each figure represent a linear least-squares fit to the measured data on a log-linear scale, and the resulting fitted slope is used to calculate the nucleation rate used in equation (10), assuming the droplet volume is equal to the average volume for the entire droplet population. The fitting parameters A and B given in equation (10) are obtained from this fitted slope. The red and blue solid lines on each plot represent the model predictions using the same kinetic fitting parameters A and B

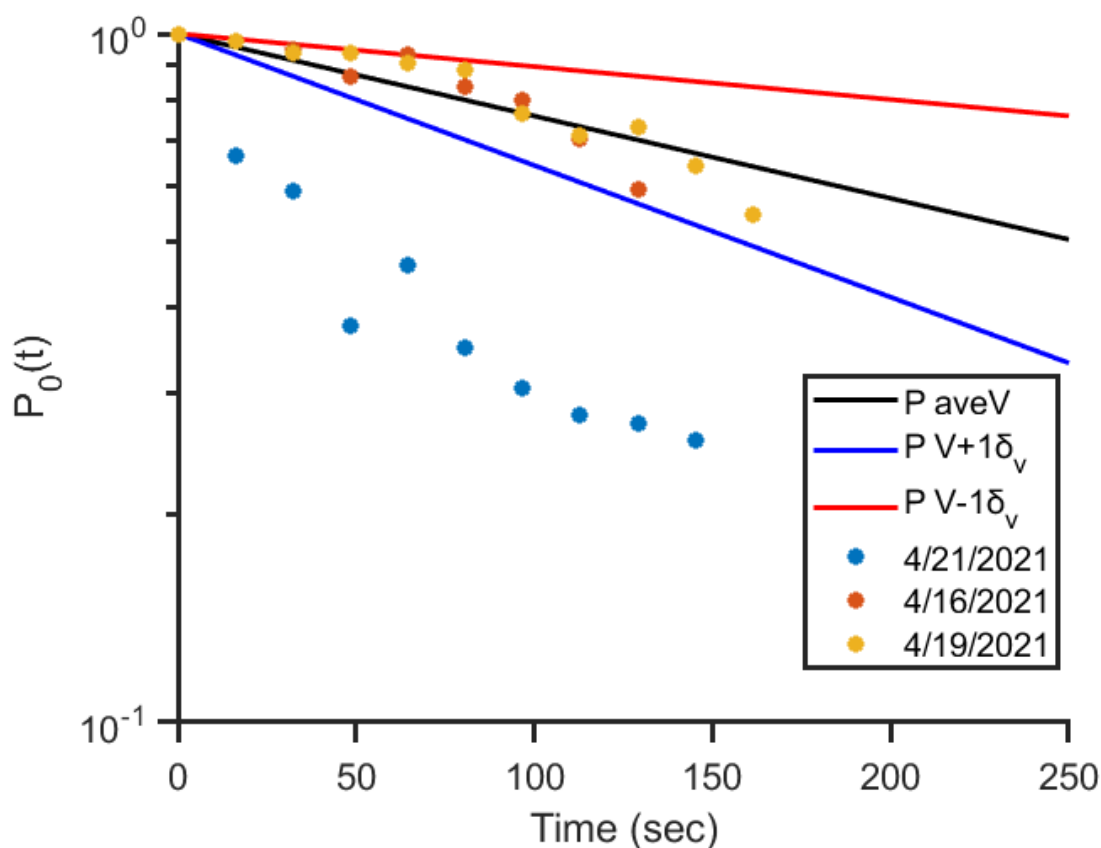


Figure 2-20 Observed probability function versus time for 30 mg/mL Lys 3w/v% NaCl. The observed probability function is plotted versus time for a solution containing 30 mg/mL lysozyme and 3 w/v% NaCl. The black, blue and red lines were constructed as described for **Figure 2-19**.

obtained from the best fit slope, but using a volume equal to one standard deviation higher or lower, respectively, based on the volume distributions shown in **Figure 2-17** and corresponding standard deviation values listed in **Table 2-3**. The slopes of the red and blue lines were then obtained and used to determine the change in nucleation

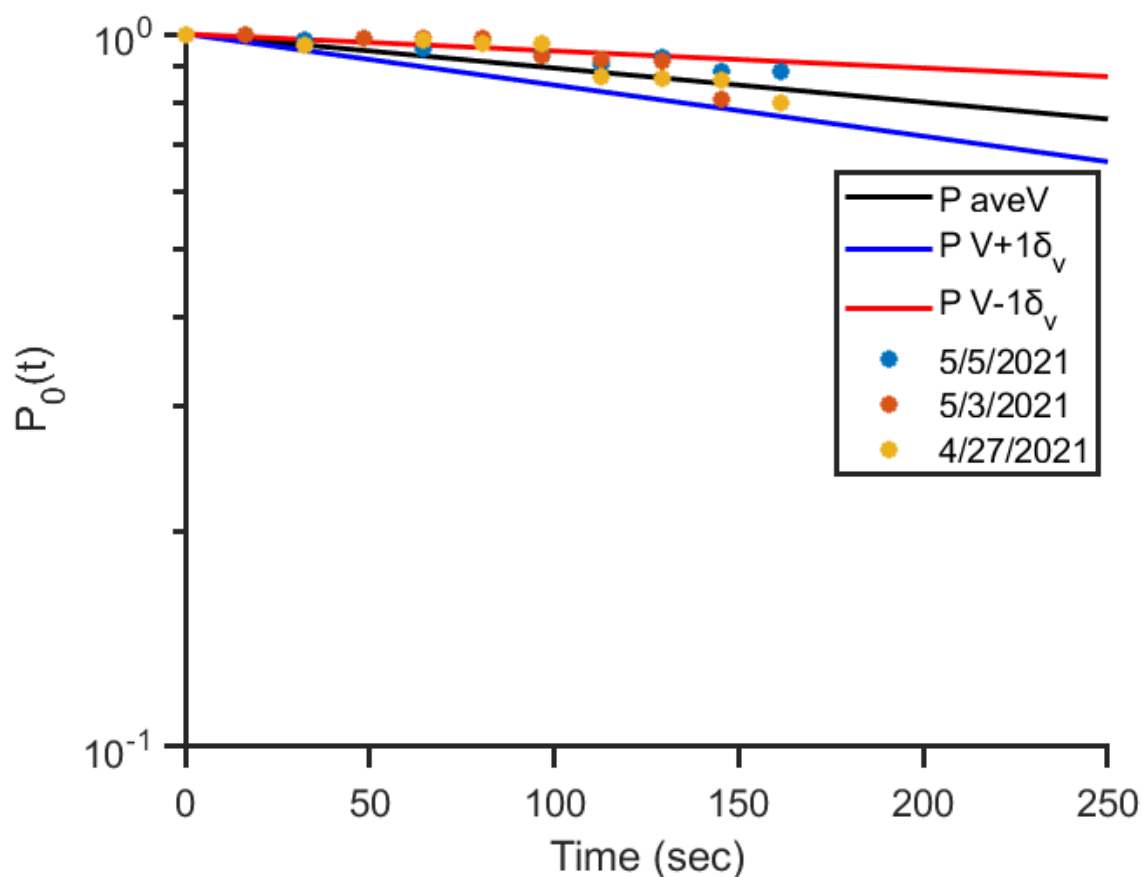


Figure 2-21 Observed probability function versus time for 20 mg/mL Lys 2w/v% NaCl. The observed probability is plotted versus time for a solution containing 20 mg/mL lysozyme and 2 w/v% NaCl. The black, blue and red lines were constructed the same way as described for Figure 2-17.

parameters that would have resulted if the volume were assumed to remain constant at the average value. This analysis provides an estimate of the uncertainty in nucleation fitting parameters that could be obtained from these experiments. All resulting fitting

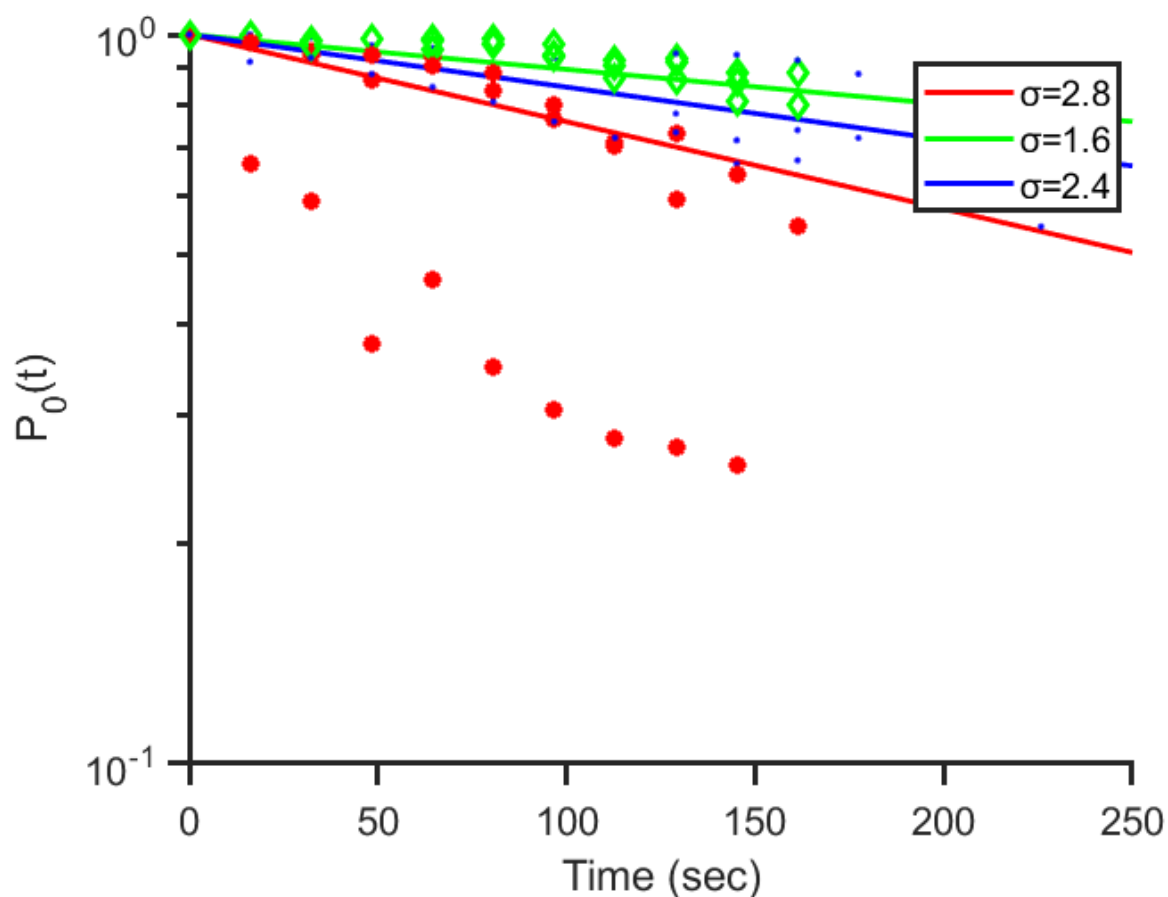


Figure 2-22 Combined Continuous crystallization results. The graph combines the recorded data points from **Figure 2-19**, and **Figure 2-20**, and **Figure 2-21**, as well as the best fit lines for those figures.

parameters are summarized in **Table 2-5**, and we will examine the observed trends next.

Figure 2-22 as well as the parameter values given in **Table 2-5** demonstrate that the higher the supersaturation, the faster the nucleation. However at each supersaturation significant run-to-run variability is observed, particularly at the highest supersaturation as shown in **Figure 2-20**. In this case, data collected on 4/21/2021 deviated significantly from the other trials at that same composition. The reasons for this deviation are not clear, it could be due to contamination allowing for more rapid

heterogeneous nucleation. To show the error associated with the uncertainty in drop size, along with the best fit line, two other probability distributions are modeled with identical kinetic parameters, but volume adjusted up and down by the standard deviation for the population.

Table 2-5 *Extracted kinetic fitting parameters from lysozyme nucleation in continuous drop trains*

σ	A $mg^{-1} s^{-1}$	B	J $ml^{-1} s^{-1}$	+A $mg^{-1} s^{-1}$	+B	+J $ml^{-1} s^{-1}$	- A $mg^{-1} s^{-1}$	-B	-J $ml^{-1} s^{-1}$
2.8	4.5×10^7	290	183.5	2.9×10^7	299	293.4	4.4×10^7	281	72.3
1.6	3.0×10^7	230	100.9	2.1×10^7	235	151.2	3.2×10^7	225	50.3
2.4	9.0×10^7	240	153.7	5.1×10^7	249	261.3	10.9×10^7	235	47.3

Observed nucleation rates are on the same order of magnitude as those from from the static array trials $A=6.2 \times 10^7 \text{ mg}^{-1} \text{ s}^{-1}$ and $B=293.2$. Directly comparing the graphs of the continuous crystallization device to the static array is of little utility because the significantly shorter residence time causes all of the observed times in the continuous device occur within the first two frames of recorded data in the static arrays. However the two graphs do still show agreement, with the continuous data points falling within the range measured in static arrays.

Discussion

Static arrays and continuous drop train microfluidics both show promise for investigating nucleation of crystals from solution. Static arrays offer significantly simpler operation and data processing, as well as the ability to monitor processes within drops

for significantly longer residence times. On the other hand, there are several downsides of static array devices. Generally, since drops are trapped and prevented from moving until the process is complete, the device must be made physically larger to accommodate additional drops representing additional independent experiments. there are significant limitations to the number of droplets that can reasonably be observed within the viewing window for a typical optical microscope at a typical objective magnification. For the droplet volumes considered in the present experiments, a 2X objective magnification only permitted 182 droplets to be observed simultaneously, and a maximum of 720 droplets could be fabricated within the current microfluidic device design. With the limited number of droplets viewable in a single frame, rastering across devices is needed to allow for monitoring every drop in a larger array. Rastering takes time and therefore limits the available time resolution for observations. These limitations can be mitigated somewhat by capturing lower magnification images at the expense of spatial resolution, or by modifying the geometry of the trap array; however the overall number of droplets that can be monitored within a static array device will always be moderate. The limited number of total experiments can be observed in the experimental results, for example those presented in **Figure 2-10**, where an array of 182 drops containing a solution of lysozyme was monitored in a static array microfluidic device corresponding to the geometry shown in **Figure 2-5**. While the observed fraction of drops without crystals agreed well with model predictions at earlier times, at later times the data points become more widely spaced on the vertical axis as lower probabilities are attained.

The increased vertical spacing of the data points is a direct result of the finite number of drops that can be observed at once. Since the lowest resolution observable probability value is $1/N$, on a log scale the data points will appear to be spaced further apart at values of P corresponding to a mostly-crystallized droplet array. Therefore, the ability to accurately measure low nucleation probability values that typically at later times in the process is severely limited within static array microfluidic devices. There are several situations in which this limitation may present a challenge. For example, if a crystal can nucleate two different polymorphs, but one forms at a much higher nucleation rate, it may be difficult to observe the slower nucleation rate if a majority of the drops nucleate the faster crystal first. To get around this you need to increase the number of drops observed, which will come at the cost of either optical resolution or temporal resolution. Additionally increasing the number of drops requires the fabrication of physically larger devices which may introduce other issues related to being able to flow fluids into them uniformly.

The nucleation experiments conducted in static arrays in which droplets dehydrate over time demonstrate the role of drop position within a microfluidic device on the observed nucleation rate. The drops closer to the edge of device are observed to dehydrate more quickly than those in the center, likely due to the edges of the device having more surface area exposed to the surrounding dry air. This leading to uneven rates of concentration change across the device and results in the large observed standard deviation of dehydration rates. Devices and humidity chambers must be designed carefully to avoid this effect. Without mitigating the spatial variation in

dehydration rates, the observed nucleation rates can exhibit significant deviation from model predictions, as shown in **Figure 2-14**.

In contrast to static array devices, continuous flow drop based microfluidic devices instead allow the drops to flow through the device during the experiment, permitting the number of observable drops to be decoupled from the physical size of the device, and as large as desired given the ability to observe droplets for an indefinite length of time. The dramatically increased number of available independent experiments allows for any desired statistical accuracy to be achieved within a finite device footprint. Spatial limitations do still arise, however, since the residence time available within a given device is determined by the drop positions and drop velocities within the device. The typically larger droplet velocities of 10-100mm/min combined with the finite channel length that can occupy a microfluidic footprint leads to very short total residence times within the continuous flow devices. While the static arrays considered here permitted incubation times of the order of tens of hours or longer, the continuous flow devices resulted in observable incubation times of less than ten minutes. Thus, the static array devices are well suited for observing relatively slow processes for which the early stages of nucleation contain the most relevant information. The continuous drop train devices are better suited for relatively fast processes for which greater statistical accuracy of the early time nucleation probability is desired. The two methods may reveal complementary information for the same process.

In static droplet arrays, observing $P(t)$ is relatively straightforward since the determination of the probability simply requires counting the number of drops that contain a visible crystal. In continuous droplet trains, the residence time of a particular

drop is related to its position, so the probability value corresponding to a given incubation time only becomes evident once a large enough number of passing droplets have been observed.

The observed probability distributions exhibit significant run-to-run variability. Based on the analysis presented earlier, it appears that much of the observed variability arises from variations in the drop volumes. The polydispersity of droplet volumes leads directly to variations in the observed nucleation rate as indicated in equation (10), which shows that the nucleation rate is directly proportional to the volume. The large polydispersity of droplets formed in these devices makes it difficult to obtain precise nucleation rates with confidence. However, the overall trends show promise for improving this analysis capability in the future. For example, solutions with higher supersaturation exhibited higher average nucleation rates as expected from Classical Nucleation Theory. We expect that reducing the polydispersity of the droplet volumes will directly improve the confidence in the fitted nucleation parameters.

Comparing the two microfluidic devices examined in this chapter, each presents unique advantages and tradeoffs that must be considered when attempting to measure nucleation kinetics. Static arrays offer clear advantages in terms of total incubation time for the drops, allowing them to be isolated and observed for practically infinite lengths of time. Continuous drop train devices have a rather short finite residence time dictated by the flow rates and geometry of the residence channel. For a slowly nucleating system, static arrays offer superior incubation time and a larger span of the probability distribution, from $P_0(t) = 1$ down to $P_0(t) = 1/N$, where N is the number of drops. For particularly slowly nucleating systems the number of observable drops can be increased

by utilizing externally moving camera stages to raster over a larger device area. Slowly nucleating systems are also well suited for the create-and-store device used in this study, as the solution is supersaturated throughout the loading procedure, and it is preferred to have no crystals form before observation begins. Nucleation rates can be slowed by operating at higher temperatures, lower concentrations, or lower ionic strengths. All of these factors work to reduce the supersaturation of the solution, which reduces the driving force for nucleation. Although a large range of $P_0(t)$ can be observed the data becomes increasingly discrete as the number of drops without crystals decreases.

Continuous drop train microfluidics offer an entirely different set of strengths which can be leveraged in other operating conditions. The finite residence time within the channel imposes a maximum incubation time for observing droplets, so this style of device is better suited for rapidly nucleating systems. Since the drops are generated continuously, as many drops can be produced as are needed to achieve the desired statistical accuracy. The nucleation rate for a material can be increased by operating at lower temperatures, higher concentrations and higher ionic strengths. All of these factors increase the supersaturation of the solution and lead to faster crystal formation. Although there is some variability in nucleation rates possible by tuning temperature, concentration, and ionic strength, these factors also usually have practical constraints for a given crystallizing system. For example, too low a temperature will result in formation of crystals before droplet breakup, which complicates analysis of the observed probability distribution. At higher concentrations, it may be possible for

crystals to form before the droplets can achieve complete mixing, violating the assumption of a well mixed solution that is foundational to Classical Nucleation Theory.

For continuous droplet trains, the range of observable $P_0(t)$ can be extended by retaining the drops within the channel for longer. This can be done by increasing the channel length, increasing the cross-sectional area of the channels, or lowering the flow rates within the channel. Each of these options has practical limits. For example, the pressure drop along microfluidic channels increases with length and can eventually exceed the tensile strength of the device or the allowable back pressure of the syringe pump. Syringe pumps also have minimum operating flow rates and can begin to exhibit pronounced discrete motions at lower flow rates. With these constraints in mind, there are still many applications for which continuous drop train microfluidics are well suited. Primarily quickly nucleating systems, such as solutions of concentrated proteins or super cooled ice nucleating droplets where the limited residence time does not interfere with the observation of interest. These devices provide arbitrarily accurate statistics for shorter range of incubation times.

Further refinement of either design could potentially help emphasize the strengths or mitigate the shortcomings of the devices, but it is clear that the two devices operate in very different regimes of timescales and numbers of independent experiments observable. A brief summary of the tradeoffs between the two styles of microfluidic devices is summarized in **Table 2-6**.

Table 2-6 Tradeoffs between static and continuous droplet devices for examining crystal nucleation from solution.

	<i>Static droplet array</i>	<i>Continuous droplet train</i>
<i>Drops per area</i>	Finite	Infinite
<i>Residence time</i>	Infinite	Finite
<i>Drop age</i>	Uniform	Function of position
<i>Nucleation rate</i>	slow	fast
<i>Statistical accuracy</i>	Geometry dependent	Arbitrarily high
<i>Statistical range</i>	1 to $1/N$	Limited by residence time

Conclusions

Isolating small volumes of supersaturated solutions in microfluidic devices provides the ability to directly observe the nucleation rates for the system. Microfluidics have an intrinsic advantage over benchtop scale experiments in terms of material consumed due to the small length scales involved. The unique ability to isolate nanoliter scale droplets of water is hugely beneficial for the study of nucleating systems. In this study we compared two microfluidic approaches to study the nucleation of lysozyme. First we utilized a static array microfluidic device and applied a temperature step to raise the supersaturation and induce nucleation within the droplets. This method provided good agreement with values reported in the literature but helped to highlight some of the issues related to static arrays. The finite number of drops prescribed by the device geometry limits the resolution in the probability distribution as the system approaches $(N-1)/N$ drops crystallized.

Additionally in another set of experiments with the same devices, we allowed the droplets to dehydrate by exposing the device to a low relative humidity at room temperature, causing the supersaturation to increase as the solutes concentrated in the shrinking droplets. This is a unique method of inducing crystallization within droplets permitted by the long residence time and water permeability of PDMS. However analysis of this data was greatly limited due to the droplets located closer to the edge of the device experiencing different dehydration rates compared to the drops located more towards the center of the device. Future experiments of this kind should seek to prevent the edges of the device from being exposed to the low relative humidity so spatial variations within the device can be eliminated.

Next we use drop based continuous flow microfluidic device to also study the nucleation of lysozyme in a different microfluidic device geometry. Continuous flow drop based microfluidics provide a powerful and flexible platform to determine the probability distribution for crystallization. The number of independent nucleation events that can be observed is decouple from the physical size of the microfluidic device, unlike in static array microfluidics. The ability of this design to scale up the number of drops without necessitating are physically larger device dramatically improves the possible statistical accuracy of continuous flow devices compared to static array style microfluidic devices. This platform can readily be expanded to measure many complex systems including co-crystallization. Should this platform be adapted to cocrystal systems, additional instrumentation may be required to identify the crystal species that nucleates. The transparent nature of PDMS allows for the potential addition of inline spectroscopy techniques for crystal identification provided that the peaks of interest are not similar to

those of PDMS. Experimental results highlight the importance of volume distribution and temperature control in reducing the uncertainty of measured nucleation events.

Refinement of the drop generation nozzle and temperature control would be a first priority for future work on this device.

The limited residence time within the device provides some limitations on the type of nucleating systems that can be observed. For example, slowly nucleating systems may leave the observable residence channel before significant nucleation occurs. This can be slightly mitigated by careful experiment planning but ultimately can never be completely avoided.

For our system, droplet volume variations played a significant role in the uncertainty of the measurements taken from the device. Refinement of the drop generation to reduce the polydispersity is required to reduce this uncertainty. Changes in drop generation to reduce polydispersity will aid in increasing the accuracy of measurements. Changes may include switching from T-junction drop generation to another method such as co-flow or flow focusing or modifying the T-junction to better accommodate the slow flow rates. Additionally changes to the channel geometry such as widening the residence channels would allow for the syringe pumps to operate at higher flow rates, which could potentially stabilize drop generation.

Comparing the nucleation results between the two devices shows that the measured nucleation fitting parameters are on the same order of magnitude as each other and as values seen in the literature for similar experiments. While the two measured probability distributions agree with each other, it is difficult to directly compare them due to the massive difference in time scales.

References

- [1] Jon F. Edd, Katherine J. Humphry, Daniel Irimia, David A. Weitzbc, Mehmet Toner. Lab Chip, 2009, 9, 1859–1865
- [2] M. R. Thorson, S. Goyal, B. R. Schudel, C. F. Zukoski, G. G. Zhang, Y. Gong and P. J. Kenis, Lab Chip, 2011, 11, 3829–3837.
- [3] S. L. Morissette, O. Almarsson, M. L. Peterson, J. F. Remenar, M. J. Read, A. V. Lemmo, S. Ellis, M. J. Cima and C. R. Gardner, Adv. Drug Delivery Rev., 2004, 56, 275–300. [1] J. Am. Chem. Soc., Vol. 122, No. 1, 2000
- [4] Leubner, IH. “Particle Nucleation and Growth Models.” Current opinion in colloid & interface science 5.1-2 (2000): 151–159. Web.
- [5] Karnal, A. K. “Nucleation-Trap Crystallizer for Growth of Crystals from Solutions.” Journal of crystal growth 297.1 (2006): 152–156. Web.
- [6] RATH, BB. “KINETICS OF NUCLEATION AND GROWTH-PROCESSES.” Materials science & engineering. 32.3 (1995): 101–106. Web.
- [7] Verdurand, Elodie. “Secondary Nucleation and Growth of Organic Crystals in Industrial Crystallization.” Journal of crystal growth 275.1-2 (2005): E1363–E1367. Web.
- [8] Narayan Variankaval and Aaron S. Cote. AIChE Journal July 2008 Vol. 54, No. 7.
- [9] [https://chem.libretexts.org/Textbook_Maps/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_\(Physical_and_Theoretical_Chemistry\)/Equilibria/Solubility/Types_of_Saturation](https://chem.libretexts.org/Textbook_Maps/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Equilibria/Solubility/Types_of_Saturation)
- [10] Yan Xiao, Jingkang Wang, Xin Huang, Huanhuan Shi, Yanan Zhou, Shuyi Zong, Hongxun Hao, Ying Bao, Qiuxiang Yin. Cryst. Growth Des. 2018, 18, 540–551
- [11] Myerson A. S., Cryst. Growth Des. 2017, 17, 3287–3294
- [12] Liang Li and Rustem F. Ismagilov. Annu. Rev. Biophys. 2010. 39:139–58.
- [13] Peter G. Vekilov, Oleg Galkin. Physicochem. Eng. Aspects 215 (2003) 125-130
- [14] Shiao, Lie-Ding. “Determination of the Nucleation and Growth Kinetics for Aqueous L-Glycine Solutions from the Turbidity Induction Time Data.” Crystals. 8.11 (2018): n. pag. Web.
- [15] Izmailov, AF. “A Statistical Understanding of Nucleation.” Journal of crystal growth 196.2-4 (1999): 234–242. Web.
- [16] <https://www.crystallizationsystems.com/crystal16>
- [17] Desnos, Hugo. “Ice Induction in DSC Experiments with Snomax.” Thermochemica acta 667 (2018): 193–206. Web.
- [18] Hakim Boukellal, Seila Selimovic, Yanwei Jia, Galder Cristobal, Seth Fraden. Lab Chip, 2009, 9, 331–338

- [19] Shelley L. Anna, Nathalie Bontoux, and Howard A. Stone. Appl. Phys. Lett. 82, 364 (2003)
- [20] Galkin, O. "Direct Determination of the Nucleation Rates of Protein Crystals." The journal of physical chemistry. 103.49 (1999): 10965–10971. Web.
- [21] Akella, Sathish V. "Emulsion-Based Technique To Measure Protein Crystal Nucleation Rates of Lysozyme." Crystal growth & design. 14.9 (2014): 4487–4509. Web.
- [22] Galkin, O. "Nucleation of Protein Crystals: Critical Nuclei, Phase Behavior, and Control Pathways." Journal of crystal growth 232.1-4 (2001): 63–76. Web.
- [23] Baidakov, Vladimir G. "Crystal Nucleation Rate Isotherms in Lennard-Jones Liquids." The Journal of chemical physics. 132.23 (2010): n. pag. Web.
- [24] Karthika, S. "A Review of Classical and Nonclassical Nucleation Theories." Crystal growth & design. 16.11 (2016): 6663–6681. Web.
- [25] Goegelein, Christoph. "A Simple Patchy Colloid Model for the Phase Behavior of Lysozyme Dispersions." The Journal of chemical physics. 129.8 (2008): n. pag. Web.
- [26] PARK, JM. "EFFECTS OF PROTEIN CHARGE HETEROGENEITY IN PROTEIN-POLYELECTROLYTE COMPLEXATION." Macromolecules. 25.1 (1992): 290–295. Web.
- [27] Stern, Eric. "Importance of the Debye Screening Length on Nanowire Field Effect Transistor Sensors." Nano letters 7.11 (2007): 3405–3409. Web.
- [28] Zheng, B. "Using Nanoliter Plugs in Microfluidics to Facilitate and Understand Protein Crystallization." Current opinion in structural biology. 15.5 (2005): 548–555. Web.
- [29] Dewan, Alim. "Growth Kinetics of Microalgae in Microfluidic Static Drop Arrays." Biotechnology and bioengineering. 109.12 (2012): 2987–2996. Web.
- [30] Sun, Meng. "Microfluidic Static Drop Arrays with Tuneable Gradients in Material Composition." Lab on a chip 11.23 (2011): 3949–3952. Web.
- [31] Jin, Si Hyung. "A Programmable Microfluidic Static Drop Array for Drop Generation, Transportation, Fusion, Storage, and Retrieval." Lab on a chip 15.18 (2015): 3677–3686. Web.
- [32] Jeong, Heon-Ho. "Microfluidic Static Drop Array for Analyzing Microbial Communication on a Population Gradient." Lab on a chip 15.3 (2015): 889–899. Web.
- [33] Forsythe, E. L.; Judge, R. A.; Pusey, M. L. J. Chem. Eng. Data 1999, 44, 637–640.
- [34] McDonald, JC. "Fabrication of microfluidic systems in poly(dimethylsiloxane)" Electrophoresis Vol.21 Issue 1. Page 27-40. (2000)

- [35] Garstecki, P. "Formation of Drops and Bubbles in a Microfluidic T-Junction - Scaling and Mechanism of Break-up (vol 6, Pg 437, 2006)." Lab on a chip 6.5 (2006): 693–693. Print.
- [36] Christopher, G. F. "Microfluidic Methods for Generating Continuous Drop Streams." Journal of Physics D: Applied physics 40.19 (2007): R319–R336. Web.
- [37] Bings, NH. "Microfluidic Devices Connected to Fused-Silica Capillaries with Minimal Dead Volume." Analytical chemistry. 71.15 (1999): 3292–3296. Web.
- [38] Wu, Dan. "Interaction Between Proteins and Cationic Gemini Surfactant." Biomacromolecules. 8.2 (2007): 708–712. Web.
- [39] Nanev, Christo N. "Relationship Between Number and Sizes of Crystals Growing in Batch Crystallization: Nuclei Number Density, Nucleation Kinetics and Crystal Polydispersity." Journal of crystal growth 546 (2020): n. pag. Web.
- [40] Ji Yeon Jun, "Preparation of size-controlled bovine serum albumin (BSA) nanoparticles by a modified desolvation method" Food Chemistry, Volume 127, Issue 4, (2011) Pages 1892-1898,
- [41] Nanev, Christo N. "Steering a Crystallization Process to Reduce Crystal Polydispersity; Case Study of Insulin Crystallization." Journal of crystal growth 480 (2017): 164–169. Web.
- [42] D. Shanthi, P. Selvarajan, R. Jothi Mani, "Nucleation kinetics, growth and hardness parameters of L-alanine alaninium picrate (LAAP) single crystals", Optik, Volume 125, Issue 11, (2014) Pages 2531-2537
- [43] He, Guangwen. "Direct Growth of γ -Glycine from Neutral Aqueous Solutions by Slow, Evaporation-Driven Crystallization." Crystal growth & design. 6.8 (2006): 1746–1749. Web.
- [44] Dharmendra Singhal, William Curatolo, "Drug polymorphism and dosage form design: a practical perspective". Advanced Drug Delivery Reviews, Volume 56, Issue 3, (2004), Pages 335-347
- [45] Matsuda Y, Akazawa R, Teraoka R, Otsuka M. Pharmaceutical evaluation of carbamazepine modifications: comparative study for photostability of carbamazepine polymorphs by using Fourier-transformed reflection-absorption infrared spectroscopy and colorimetric measurement. J Pharm Pharmacol. 1994 Mar;46(3):162-7. doi: 10.1111/j.2042-7158.1994.tb03770.x. PMID: 8027920.
- [46] http://www.tainstruments.com/pdf/literature/TN071_REV-B-DSC_Sample_Pan_Guide.pdf
- [47] Bithi, Swastika S. "Behavior of a Train of Drops in a Fluidic Network with Hydrodynamic Traps." Biomicrofluidics. 4.4 (2010): n. pag. Web.
- [48] Boukellal, Hakim. "Simple, Robust Storage of Drops and Fluids in a Microfluidic Device." Lab on a chip 9.2 (2009): 331–338. Web.

- [49] Abedi, Samira. "Collective Nucleation Dynamics in Two-Dimensional Emulsions with Hexagonal Packing." *Physical review*. 101.3 (2020): n. pag. Web.
- [50] Toldy, Arpad I. "Spherical Crystallization of Glycine from Monodisperse Microfluidic Emulsions." *Crystal growth & design*. 12.8 (2012): 3977–3982. Web.
- [51] Stan, Claudiu A. "A Microfluidic Apparatus for the Study of Ice Nucleation in Supercooled Water Drops." *Lab on a chip* 9.16 (2009): 2293–2305. Web.
- [52] Whitesides, George M. "The Origins and the Future of Microfluidics." *Nature*. 442.7101 (2006): 368–373. Web.
- [53] Squires, Todd M. "Microfluidics: Fluid Physics at the Nanoliter Scale." *Reviews of modern physics* 77.3 (2005): 977–1026. Web.
- [54] Macosko, Evan Z. "Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Drops." *Cell*. 161.5 (2015): 1202–1214. Web.
- [55] Stone, HA. "Engineering Flows in Small Devices: Microfluidics Toward a Lab-on-a-Chip." *Annual review of fluid mechanics* 36.1 (2004): 381–411. Web.
- [56] Teh, Shia-Yen. "Drop Microfluidics." *Lab on a chip* 8.2 (2008): 198–220. Web.
- [57] Sung Kwon Cho, Hyejin Moon and Chang-Jin Kim, "Creating, transporting, cutting, and merging liquid drops by electrowetting-based actuation for digital microfluidic circuits," in *Journal of Microelectromechanical Systems*, vol. 12, no. 1, pp. 70-80, Feb. 2003, doi: 10.1109/JMEMS.2002.807467.
- [58] Haeberle, Stefan. "Microfluidic Platforms for Lab-on-a-Chip Applications." *Lab on a chip* 7.9 (2007): 1094–1110. Web.
- [59] Baret, Jean-Christophe. "Surfactants in Drop-Based Microfluidics." *Lab on a chip* 12.3 (2012): 422–433. Web.
- [60] Dangla, R. "Drop Microfluidics Driven by Gradients of Confinement." *Proceedings of the National Academy of Sciences of the United States of America*. 110.3 (2013): 853–858. Web.
- [61] Xia, Younan. "SOFT LITHOGRAPHY." *Annual review of materials science*. 28.1 (1998): 153–184. Web.
- [62] Qin, Dong. "Soft Lithography for Micro- and Nanoscale Patterning." *Nature protocols*. 5.3 (2010): 491–502. Web.
- [63] Teychené, Sébastien. "Microfluidic Device for the Crystallization of Organic Molecules in Organic Solvents." *Crystal growth & design*. 11.11 (2011): 4810–4818. Web.
- [64] Zhang, Lei. "Continuous and Scalable Production of Well-Controlled Noble-Metal Nanocrystals in Milliliter-Sized Drop Reactors." *Nano letters* 14.11 (2014): 6626–6631. Web.

- [65] Arooj, Mahreen. "Adsorption and Unfolding of Lysozyme at a Polarized Aqueous–Organic Liquid Interface." *The journal of physical chemistry*. 120.12 (2016): 3100–3112. Web.
- [66] Clausell-Tormos, Jenifer. "Drop-Based Microfluidic Platforms for the Encapsulation and Screening of Mammalian Cells and Multicellular Organisms." *Chemistry & biology*. 15.5 (2008): 427–437. Web.
- [67] Kuhl, Tonya. "Direct Measurement of Polyethylene Glycol Induced Depletion Attraction Between Lipid Bilayers." *Langmuir : the ACS journal of surfaces and colloids*. 12.12 (1996): 3003–3014. Web.
- [68] Simões, Ilka Sumiyoshi and Gioielli, Luiz Antonio. "Crystal morphology of binary and ternary mixtures of hydrogenated fats and soybean oil." *Brazilian Archives of Biology and Technology* [online]. 2000, v. 43, n. 2, pp. 241-248.
- [69] Andrzej Jeziorny, Parameters characterizing the kinetics of the non-isothermal crystallization of poly(ethylene terephthalate) determined by d.s.c., *Polymer*, Volume 19, Issue 10, 1978, Pages 1142-1144, ISSN 0032-3861, [https://doi.org/10.1016/0032-3861\(78\)90060-5](https://doi.org/10.1016/0032-3861(78)90060-5).
- [70] Gusak, A M. "Kinetics of Nucleation in the Concentration Gradient." *Journal of Physics. Condensed matter* 13.12 (2001): 2767–2787. Web.
- [71] Laval, Philippe. "A Microfluidic Device for Investigating Crystal Nucleation Kinetics." *Journal of crystal growth* 303.2 (2007): 622–628. Web.
- [72] Halldorsson, Skarphedinn. "Advantages and Challenges of Microfluidic Cell Culture in Polydimethylsiloxane Devices." *Biosensors & bioelectronics*. 63 (2015): 218–231. Web.
- [73] https://serc.carleton.edu/research_education/geochemsheets/techniques/SXD.html
- [74] Thomas Brubaker, Michael Polen, Perry Cheng, Vinay Ekambaram, Josh Somers, Shelley L. Anna & Ryan C. Sullivan (2020) Development and characterization of a "store and create" microfluidic device to determine the heterogeneous freezing properties of ice nucleating particles, *Aerosol Science and Technology*, 54:1, 79-93, DOI: 10.1080/02786826.2019.1679349
- [75] Deqing Huang, Kang Wang, Yaolei Wang, Hejia Sun, Xingyuan Liang, Tao Meng, Precise control for the size of droplet in T-junction microfluidic based on iterative learning method, *Journal of the Franklin Institute*, Volume 357, Issue 9, 2020, Pages 5302-5316
- [76] *Chem. Rev.* 2014, 114, 15, 7610–7630 Publication Date: July 8, 2014
<https://doi.org/10.1021/cr400544s>
- [77] Karthika, S. "A Review of Classical and Nonclassical Nucleation Theories." *Crystal growth & design*. 16.11 (2016): 6663–6681. Web.

Chapter 3 : Theory of cocrystal nucleation from solution

Abstract

Homogenous nucleation of cocrystals from solution is a relatively unexplored area of research, with no widely accepted model to predict nucleation kinetic rates. Cocrystal research currently focuses on high throughput screening of co-former pairs and equilibrium phase behavior [1-5]. This chapter aims to fill this knowledge gap by proposing an extension of Classical Nucleation Theory (CNT) considering two cases: one in which coformers add directly to the crystal nucleus and one in which the formation of coformer complexes is considered. Numerical integration allows for the determination of underlying physical behavior of the cocrystallizing system. Qualitative observation of the cumulative probability distributions allow for insight into the role of the competing rates found in this process. Further analysis of the characteristic timescales present in the process allows for the definition of dimensionless parameters that elucidate the challenges of scaling up homogenous cocrystal nucleation processes. The insights provided by this new proposed model for cocrystal nucleation will be valuable in designing systems that can reliably and efficiently produce new cocrystal products.

Introduction

Cocrystals contain two or more constituents as part of their lattice structure [6]. It is often desirable to form a cocrystal when there are some beneficial properties of the single component crystal, but other properties are not well suited for the desired application [7-9]. In this chapter we address the knowledge gap surrounding the modeling of cocrystal nucleation rates. Currently there is no widely accepted model for predicting the nucleation rates in cocrystal systems. Here we will propose an extension of Classical Nucleation Theory in which we include reaction kinetics describing the

formation of cocrystallizing complexes and explore the new controlling parameters for the process. The insights gained from exploring this model allow for a more complete understanding of the cocrystal nucleation process and the implications it has for scaling up operations. Cocrystals can be differentiated from a solid mixture of the two coformers by techniques such as raman spectroscopy and x-ray diffraction [52]

One application of cocrystals is in the manufacturing of Active Pharmaceutical Ingredients (APIs). In some cases, an API may have the desired therapeutic function but poor bioavailability in its obtainable single component crystal form [10].

Bioavailability is the extent to which an API can enter the biosystem of interest and have an active effect. For the human body, bioavailability is usually strongly affected by the water solubility of the drug. Solubility, and thus bioavailability, can be increased in several ways, such as by forming a salt from the pure molecule [45]. Forming a cocrystal with the API as one component is another potential path to improving the desired crystal properties when other options such as forming a salt are unavailable [11]. While cocrystals are not widely used in practice, they are gaining popularity [12]. A notable proof of concept is carbamazepine nicotinamide cocrystals. Carbamazepine is an anticonvulsant used to treat epilepsy and bipolar disorder. However, its pure crystal form has poor bioavailability. Studies have shown that forming a cocrystal with nicotinamide greatly increases the solubility of carbamazepine, which in turn increases its bioavailability [13-14].

Other uses of cocrystals include modifying the macroscopic shape of the crystal for easier handling. If the preferred crystal structure of a material causes it to form long needles, this can cause problems with the flowability of the solid as the high aspect ratio

shapes struggle to roll or slide past one another [44]. Cocrystal polymorphs can take on different shapes compared to their single component counterparts, so if a desired product can be cocrystallized with an inert coformer it is possible to create a more workable final product. When working with APIs specifically, a key step in the manufacturing process is often the creation of a tablet, which also depends on the physical properties of the materials. All these considerations have been assessed simultaneously for the case of the API griseofulvin, an antifungal used to treat infections of the scalp, fingernails and toenails. Forming a cocrystal of griseofulvin with acesulfame, an artificial sweetener, results in a cocrystal with superior dissolution rates and powder flowability compared to the single component crystal, while maintaining tabletability [32]. Producing materials that are more workable while simultaneously more effective clearly demonstrates how cocrystals can offer significant advantages in industrial applications.

While the features of a material that promote or inhibit cocrystallization are not fully understood, there are some common traits among known coformers. In general pairs of materials that form cocrystals need a mechanism by which the system free energy can be reduced because of the association of the two components. A list of common API coformers, techniques used to identify them, and their applications can be found in **Table 3-1**. This list continues to grow as new materials are developed.

Table 3-1 Example co-crystal systems, identification techniques, and interaction mechanisms.

<i>API</i>	<i>Co-former</i>	<i>Interaction method</i>	<i>identification techniques</i>
<i>Caffeine</i>	maleic acid	Hydrogen bonds [35]	Raman [47]
<i>Lysozyme</i>	Benzamidine	Van der Waals [36]	XRD [36]
<i>Carbamazepine</i>	nicotinamide	Pi-stacking [37-38]	Raman [48]
<i>2-aminopyrimidine</i>	Benzoic acid	Halogen bonds [39]	IR spectroscopy [39]

Table 3-1 also indicates the mechanism by which the two components associate with one another in a cocrystal. Studies have shown that polar molecules can find minimum energy configurations by forming hydrogen bonds with their corresponding coformers [40]. Small molecules with aromatic rings in their structures can form cocrystals via an attractive phenomenon known as pi-stacking where the pi-bonds of one molecule align with those of another molecule. While the exact interaction mechanism of a given coformer pair is not always reported in the literature, pi-stacking and hydrogen bonds appear to be the most common mechanisms at play [33-34]. Both interactions are known to form rapidly, on the order of picoseconds for hydrogen bonding [46] and a few hundred femtoseconds for pi-stacking of small molecules like benzene [40]. On the other hand, complexes of larger molecules like proteins or polymers have been reported to form on the order of seconds to minutes [41]. The wide range of kinetic timescales of association leaves open the possibility that complex formation may compete with nucleus formation, which can occur over timescales from seconds to minutes or longer [49-50].

Hydrogen bonds contributing to the formation of cocrystals explains the presence of many weak acids and bases among the list of coformers, as they become polar after deprotonation in a solvent. Similarly, pi-stacking is common among small molecules with aromatic rings. The generality of these associative mechanisms indicates that there is potential to create many new pharmaceutical products if cocrystals can be produced reliably. However, there are several barriers to the development of new cocrystalline APIs. One barrier is that the parameter space for forming cocrystals is enormous, making experimental exploration of this phase space extremely labor intensive and costly. There is also no widely accepted model for cocrystal nucleation. Without added insight from theory, systematically exploring the large phase space requires expensive research and development time to find appropriate co-formers to make new API cocrystals. This lack of fundamental knowledge also limits the ability to design control systems that will enable the industrial scale production of cocrystal-based products. High throughput screening methods for cocrystal pairs have been developed to mitigate some of these limitations [15-18]. Computational approaches have also been employed to calculate important interaction parameters for cocrystals that have been identified [19]. However, an analytical model capturing the salient features of the cocrystal nucleation behavior, including relevant timescales and compositions, would go a long way toward enabling the design of more efficient routes to cocrystalline API development.

Industrially the most common method for producing cocrystals is wet grinding, where a solid mixture of the two coformers is milled with minimal solvent additions until cocrystals form and can be collected [28-30]. This is an energy intensive process that

scales poorly. In some cases, cooling reactor vessels or solvent evaporation have been implemented to prepare cocrystals [31]. These operations are often seeded with pre-existing cocrystals prepared by a smaller scale method. The relative abundance of literature detailing seeded reactors and other non-nucleating techniques like wet grinding indicates that there is a strong practical preference to avoid the nucleation step in producing cocrystals, presumably due to the lack of clear understanding of the conditions that promote cocrystal nucleation. Having a robust theory of cocrystal nucleation would facilitate the direct generation of cocrystals from solution, which would enable development of new materials for which seeds are not available. This chapter addresses this knowledge gap by presenting a cocrystal nucleation model built upon the foundational assumptions of Classical Nucleation Theory, combined with additional elements capturing the rate of formation of coformer complexes that then nucleate cocrystals. Factors specific to cocrystallization are considered, such as the definition of saturation concentration for the cocrystallizing mixture. Crystal growth, which occurs after a stable nucleus is formed, will largely be neglected in this chapter, although it is another important feature of crystal engineering.

Background

For a material to spontaneously transition from being dissolved in solution to forming a stable crystal nucleus, it must take a path that ultimately reduces its free energy. For single component crystals, Classical Nucleation Theory (CNT) explains this pathway. Although CNT does not perfectly describe all crystal nucleation, it still serves as a good basis upon which to build other theories, as it captures much of the underlying physics as it is currently understood [20-22]. Using CNT as a starting point, we can begin to create a model describing the transformation of a cocrystal system from

a collection of independently dissolved co-formers to a final macroscopic crystal. As explained in Chapter 2, CNT describes the nucleation of single component crystals from solution by the repeated addition of molecules to the precritical nucleus. If we assume that cocrystal nucleation operates by similar mechanisms, modifications must be made to account for incorporating the coformer species into the nucleus. We will assume here that the underlying mechanism is still the occasional collision of dissolved solute molecules driven by Brownian motion with sufficient energy to aggregate, and that after surpassing a critical size, growth of the nucleus becomes spontaneous.

Model Details

In this chapter we consider two different mechanisms for the nucleation of cocrystals as illustrated in **Figure 3-1**. The orange and blue dots on the left represent the two coformers dissolved a supersaturated solution. Proceeding to the right, the coformers add to a nucleus by one of two pathways which will be described in detail later, until a cocrystal has formed, as depicted on the right-hand side. Both mechanisms build upon the core assumptions of CNT. Each mechanism introduces a modification to CNT to address the increased complexity associated with cocrystal formation. For both proposed models we assume that the particle motions are driven by Brownian motion and that they occasionally collide with sufficient energy to overcome the energy barrier to aggregation. Consistent with CNT, we assume a quasi-stationary distribution, implying that the molecules added to the pre-critical nuclei do not significantly change the bulk concentration of the solution [51]. Additionally, we consider only homogenous nucleation, and only one nucleation pathway for each single component crystal and cocrystal, meaning that no alternative polymorphs or crystal stoichiometries with their own nucleation rates are considered.

The first route to cocrystal nucleation considers the case in which the coformers do not have strong attractive forces between them, and thus do not associate before forming a cocrystal. This model maintains the underlying assumptions of CNT, and assumes that there are three possible pathways for crystallization in the mixture: the formation of a pure crystal of either component, and the formation of a cocrystal by addition of both types of monomer to the pre-critical nucleus. Additionally, we assume that the nucleation rate for each pathway is unaffected by the presence of the other pathways. This assumption implies that the kinetic fitting parameters A and B found in the definition of the nucleation rate constant J remain the same for each single

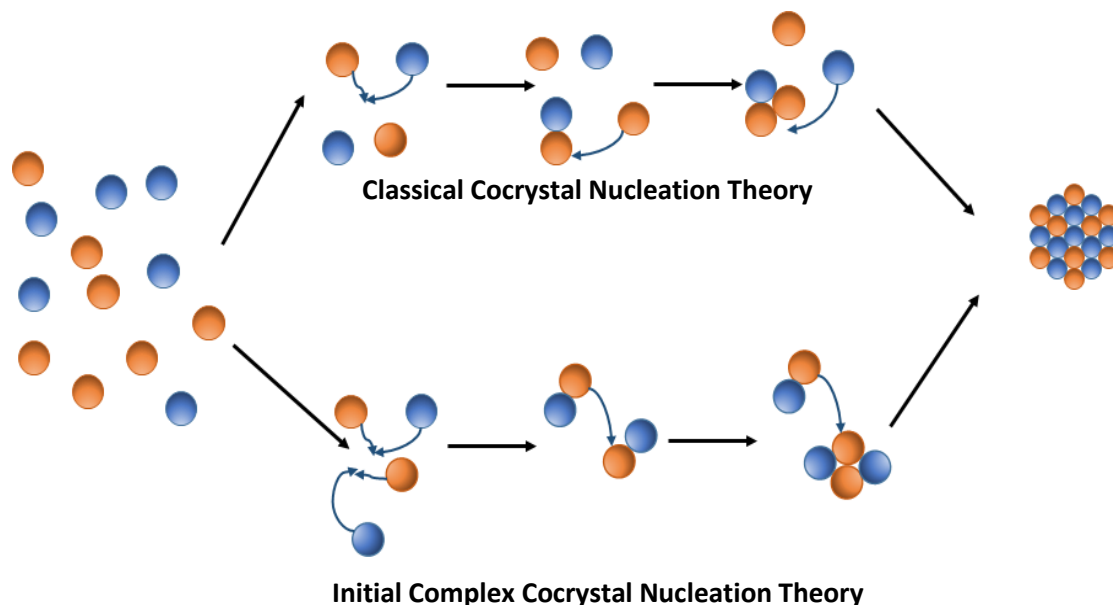


Figure 3-1 Cartoon depiction of two cocrystal nucleation pathways, one with and one without the formation of initial coformer complexes. The mechanism by which dissolved materials arrange themselves into cocrystals is not widely agreed upon. Two possible routes from solution to cocrystal are shown above. The top pathway shows CNT-like cocrystal nucleation assuming single molecule additions to the precritical nucleus that eventually lead to nucleation and growth. The bottom pathway shows initial complex formation assuming coformers must form complexes in solution that are then added to the precritical nucleus as a single unit.

component crystal, even in the presence of the other component. We also assume that the saturation concentration for each component remains the same with or without the presence of the other coformer in solution. The supersaturation with respect to the cocrystal is a function of the concentrations of the two coformers, and we assume that the coformers are added to the nucleus in an alternating manner to maintain the stoichiometry of the cocrystal. Due to the close similarity of this route to CNT, we will refer to this model as the Classical Cocrystal Nucleation Theory (CCNT) model.

The second proposed route to cocrystal formation assumes that coformers must associate with one another and form complexes first, and then add to the nucleus as a unit. This mechanism is illustrated by the bottom pathway in **Figure 3-1**. The key new assumption for this model is that the bulk concentrations of each coformer and the complexes change over time due to the reaction kinetics of the complex formation and the stoichiometry, while still assuming a quasi-stationary distribution with respect to the formation of the pre-critical nuclei. As a result, the coformer complexes act as a third independent species in solution, and we assume that there is a saturation concentration associated with the complexes, separate from the individual monomer species. The initial formation of coformer complexes leads to distinct features that will be discussed later in the chapter. We will refer to this route as the Initial Complex Cocrystal Nucleation Theory (ICCNT) model.

Classical Cocrystal Nucleation Theory Model

A successful model for cocrystal nucleation must be able to capture the macroscale observations of the final cocrystal. One observation that is of key importance is the stoichiometry of the cocrystal. Cocrystals have been observed to have

fixed integer ratios between their co-formers [23-24]. This makes intuitive sense for crystals in which the lattice structure of the cocrystal is different from either of the pure component crystals, as crystallization is often used as a purification technique due to the tendency of the crystals to exclude impurities during formation [25]. If a cocrystal lattice structure arranges the coformers so that each addition of a coformer molecule lowers the energy barrier for the addition of the opposite coformer, then once a lattice has begun to form it will self-reinforce the continuation of that same lattice pattern. From an energy standpoint, formation of a specific stoichiometry implies that the energy barrier to adding a new molecule to the cocrystal nucleus depends on the previous addition. Without the association of the coformers before their addition to the nucleus, the mechanism that controls the cocrystal stoichiometry must come at the addition step itself. If there is a lower energy barrier for one coformer to add to the nucleus after the addition of the other, then the composition of the cocrystal can be maintained throughout the nucleation and growth processes.

Figure 3-2 illustrates how to construct the cocrystal probability distribution subject to these assumptions. The figure is divided into two parts where the top graph labeled A) shows the concentrations of each individual component on the y-axis versus time on the x-axis. The bottom graph labeled B) shows the probability of not finding a crystal on the y-axis versus time, plotted on the same scale as the time axis in A). The initial concentrations of the two coformers, C_1 and C_2 , are prescribed and can take on any arbitrary value. The value of C_{sat} is also graphed, while there is a unique saturation concentration for each coformer, their values are very similar so only a single line is displayed here. The red and green lines in **Figure 3-2A** represent coformer

concentrations used in this simulation for illustrative purposes. The red, green, and blue lines in **Figure 3-2B** represent the probability of not finding a crystal of the corresponding type. By our current assumptions of a quasi-stationary distribution, the

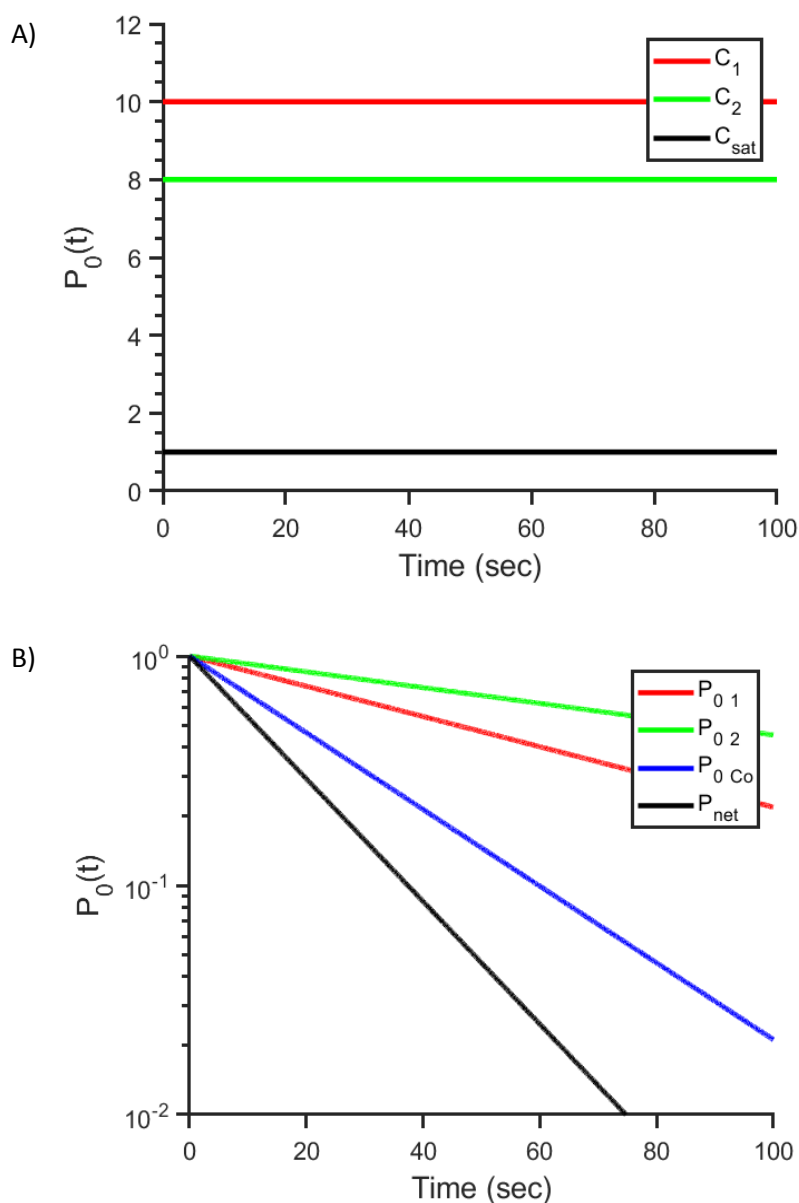


Figure 3-2 (A) Concentrations and (B) probability functions versus time for Classical Cocystal Nucleation Theory. For CCNT, the solution concentrations are assumed to be constant with time and the resulting probability distributions are simple exponentials like those found in CNT. The total probability of forming a cocystal is the product of the probabilities of each independent crystallization pathway.

concentrations do not vary with time. At a constant temperature, the supersaturation for each coformer with respect to its single component crystal is also a constant, given by

$$\sigma_1 = \ln \left(\frac{C_1}{C_{sat1}} \right), \text{ and} \quad (17)$$

$$\sigma_2 = \ln \left(\frac{C_2}{C_{sat2}} \right) \quad (18)$$

where C_{sat1} and C_{sat2} represent the saturation concentrations. The assumed independent nature of the two single component nucleation pathways implies that CNT fitting parameters for each coformer can be experimentally determined in separate experiments using the methods discussed in Chapter 2. The nucleation rate, J_i , of each single component crystal can be modeled by CNT using Equation (10), such that

$$J_i = A_i C e^{-B_i/\sigma_i^2} \quad (19)$$

where A_i and B_i are the nucleation fitting parameters for the i -th coformer, σ_i is given by Equations (17) and (18), and C_i is the concentration of the given component. In this case J_i is a constant determined by the constant concentration of the i -th component, allowing for the analytical determination of the probability distribution for the single component crystals, where $P_{i,0}(t)$ is the probability of finding zero crystals of the i -th crystal species within the nucleation volume V ,

$$\frac{dP_{i,0}(t)}{dt} = -J_i V P_{i,0}(t) \quad (20)$$

$$P_{i,0}(t) = e^{-J_i V t} \quad (21)$$

The probability of not finding a crystal is linear on a log-linear scale as shown in **Figure 3-2B**.

Examining equation (19) leads us into one of the first challenges of modeling cocrystallization: How does one define super-saturation with respect to a cocrystal. The saturation concentration is typically defined as the maximum concentration of a solute that can be dissolved in a solvent. However, in the case of cocrystals, since the solution contains two coformers, the definition of the saturation concentration is not obvious. One way to define supersaturation is to draw an analogy to limiting reagents in a reaction. In this assumption, the saturation concentration will be fixed for a given temperature, and the concentration relevant to cocrystal nucleation is taken to be the lowest saturation concentration among its co-formers,

$$C_{satco} = \text{Min}(C_{sat1}, C_{sat2}) \quad (22)$$

However, these assumptions are not always supported by experimental observations. Several studies have shown that hydrogen bonding between co-formers can lead to the formation of coformer complexes [4]. This is characteristically different from the assumption that coformer coordination occurs at the nucleation step. Cocrystal supersaturation must be redefined under these assumptions. Some previous reports indicate that at equilibrium the saturation concentration of coformer without the presence of complexes is defined by

$$S = \left(\frac{[C_1][C_2]}{K_{sp}} \right)^{\frac{1}{2}} \quad (23)$$

Where $[C_1]$ and $[C_2]$ are the molar concentrations of the co-formers and K_{sp} is a proportionality constant for their solubility [4,27]. In the case of coformer complex formation, the complexes are treated as a new species with their own constant saturation concentration.

Regardless of which definition of saturation concentration is used, since none of the component concentrations change with time, all associated values will remain constant. In this case, there are three ways a crystal can form in the solution: each of the pure components can form single crystals, and cocrystals can form. Here we assume that each of these paths is independent, and follows the behavior given by CNT. This is a relatively straightforward situation for which to find solutions since no parameters change with time, and exact solutions for J and $P_0(t)$ can be found for all three types of nucleation events. Each probability distribution is a simple decaying exponential function as shown in **Figure 3-2B**.

This time independent behavior of the concentrations always results in constant slopes for the probability distribution. The three parallel and independent ways a crystal can form means that the probability of not finding any crystal $P_{0\ net}(t)$ is the product of three separate cumulative probability functions P_{10} , P_{20} , and P_{co0}

$$P_{0\ net}(t) = P_{01} * P_{02} * P_{0co} \quad (24)$$

$P_{0\ net}$ is the quantity that would be directly measured in a droplet microfluidic nucleation experiment such as those described in Chapter 2. Inline crystal species identification capabilities would need to be added to obtain the individual cumulative probability functions. If we assume that P_{01} and P_{02} can be measured independently for the individual coformer species, these functions could then be used along with the measured $P_{0\ net}$ to determine P_{0co} . Fitting the slope of P_{0co} would then allow for the determination of A_{co} and B_{co} . To determine the final mixture composition of crystals, the relative rates of the three nucleation pathways must be examined. In the

example plotted in **Figure 3-2B**, P_{oco} is lower than P_{01} and P_{02} at all times during the trial. Recalling that P_0 is the probability of *not* finding a crystal within the volume, this result implies that we are always more likely to form a cocrystal at these conditions. Therefore, cocrystals will dominate the expected final crystal mixture for this set of conditions. Under this model, since all nucleation rates are constant, the species with the highest rate will always be the most likely to form a crystal first.

Initial Complex Cocrystal Nucleation Theory model

Let us consider an alternative model for cocrystal nucleation from solution. In this case we will assume that coformer complexes must form *before* entering the nucleus. We will model the formation of the complexes using a reaction kinetic model, which will allow for the possibility that a finite time is needed for complex formation. Subsequent to their formation, the coformer complexes add to the pre-critical nucleus as a unit, thereby enforcing the crystal stoichiometry. Key differences in the underlying physics arise for this model, most notably that there are now additional timescales to consider in the process. The rate of coformer complex formation competes with the nucleation rate of pure component crystals in concentration spaces where either or both of the pure components are supersaturated.

To evaluate the competing effects of coformer complexation, pure component nucleation, and cocrystal nucleation, we first prescribe the initial concentrations of the two co-formers and we set the initial concentration of the coformer complexes to zero,

$$C_1(0) = C_{1,0} \quad C_2(0) = C_{2,0} \quad C_{co}(0) = 0$$

As time progresses, the concentration of the coformer complexes will increase as they form. For this chapter, we will assume that the stoichiometry of the pure components in

the complex is 1:1 and thus the complex formation reaction proceeds according to a first order equilibrium reaction given by the reaction equation



where *API* denotes the active pharmaceutical ingredient of interest, *ICF* denotes the inert coformer. The *API* and *ICF* components associate to form the coformer complex denoted *CoC* in solution. k_f and k_r are the forward and backward rate constants for the formation of the coformer complexes. The solution to the first order reaction given in equation (25) results in time-dependent concentrations for each component given by

$$[API](t) = [API_0] * \left(\frac{k_r + k_f e^{-(k_f + k_r)t}}{k_f + k_r} \right), \quad (26)$$

$$[ICF](t) = [ICF_0] - ([API_0] - [API](t)), \text{ and} \quad (27)$$

$$[CoC](t) = [API_0] - [API](t) \quad (28)$$

where $[API]$, $[ICF]$, and $[CoC]$ are the molar concentrations of the two pure components and the coformer complexes respectively. $[API_0]$ and $[ICF_0]$ are the initial molar concentrations of the co-formers.

Each of the three concentrations described in equations (26) through (28) also influence the nucleation of both pure components and coformer complexes, and thus they contribute to the overall probability of not finding a crystal within a particular volume. In each case, the nucleation rate constant can be determined using the CNT expression for J for each component, given by

$$J_1 = A_1 [API] e^{-\frac{B_1}{\sigma_1^2}}, \quad (29)$$

$$J_2 = A_2[ICF]e^{-\frac{B_2}{\sigma_2^2}}, \text{ and} \quad (30)$$

$$J_{co} = A_{co}[CoC]e^{-\frac{B_{co}}{\sigma_{co}^2}} \quad (31)$$

where A_x , B_x and C_x are the CNT fitting parameters and the concentrations corresponding to the three potential nucleation events, resulting in three independent nucleation rates J_x and three independent nucleation probabilities given by

$$\frac{dP_{x0}(t)}{dt} = -J_x V P_{x0}(t)$$

where $P_{x0}(t)$ is the probability of not finding a crystal in volume V at time t .

In this case, since the concentrations change with time, the nucleation rates J_i also change with time, and the cumulative probability function is no longer a simple exponential function. The coupling of the reaction and nucleation equations makes it impossible to evaluate the solution to the differential equation for P_{net} analytically. However, the system of equations can be readily solved by numerical methods. In this case we use the MATLAB ode23 solver to determine J_1 , J_2 , and J_{co} as well as $P_{0,1}$, $P_{0,2}$, and $P_{0,co}$.

Results and Discussion

In practice, droplet-based microfluidic approaches such as those described in Chapter 2 provide the opportunity to directly observe $P_{0,net}$ by counting the fraction of drops without crystals in them as a function of time [26]. Direct observations like this allow for the determination of several important physical parameters describing the cocrystal nucleation process. The functional form of $P_{0,net}$ is determined by the forward and reverse reaction rates k_f and k_r for the formation of coformer complexes as well as

the nucleation parameters A_{co} and B_{co} . Provided that independent measurements for A_1, B_1 and A_2, B_2 are acquired from single component nucleation experiments, then k_f , k_r , A_{co} and B_{co} can be found by fitting the observed $P_{net}(t)$. This assumes that the single component nucleation parameters are independent of concentration and are not influenced by the formation of coformer complexes.

Qualitative differences in the shape of $P_{0\ net}$ also provide insight into the nature of the underlying nucleation process. The two rate processes of complex formation and nucleus formation compete during cocrystal formation, and the relative rates will influence the observed kinetics. To compare the rate processes of complex formation and cocrystal nucleus formation, we assume that the formation of coformer complexes is primarily controlled by the forward reaction rate k_f . This rate determines the time needed for the coformer complexes to reach an equilibrium concentration. The total nucleation rate, on the other hand, removes material from solution by forming the new crystalline solid phase. In concentration spaces where the pure components are initially supersaturated with respect to the single component crystal, it is possible for nucleation to begin before the coformer and complex concentrations have reached equilibrium. Depending on which rate is dominant, the observed shape of $P_{0\ net}$ as well as the composition of the expected final crystal mixture will differ. If the shape of $P_{0\ net}$ resembles a simple exponential with no noticeable curvature, the rate of formation of coformer complexes is assumed to be much faster than the rate of cocrystal nucleation.

At this point it is important to appreciate how large a parameter space governs the process of homogenous cocrystal nucleation from solution. Each component has its own kinetic fitting parameters A, B and supersaturation σ , which all contribute to the

value of J for each species. For the coformer complexes, the addition of k_f and k_r further complicate the process. To get a handle on how each of these parameters affects the process, let us examine them individually. The pre-exponential factor A is directly proportional to the nucleation rate, J , so a linear increase in A will cause a linear increase in the nucleation rate. The exponential rate parameter B relates to the free energy change of the process. A linear increase in B will result in an exponential decrease in J . The rate constants k_f and k_r determine the shape of the transient concentration curves for the process. The forward formation rate of the coformer complexes k_f primarily controls the timescale at which the solution reaches an equilibrium concentration, and the ratio of $\frac{k_f}{k_r}$ determines the magnitude of the equilibrium concentration of coformer complexes. Since k_f and k_r determine the evolution of the component concentrations with time, these two parameters also strongly influence the value of supersaturation as it changes with concentration and thus time according to equation (23). Generally, a linear increase in k_f will lead to a linear increase in the equilibrium concentration of coformer complexes, which in turn results in a logarithmic increase in supersaturation. These changes will be directly correlated with the cocrystal nucleation rate, as faster formation of coformer complexes leads to higher saturations being reached more rapidly. Increases in k_f will also be inversely correlated with the single component crystal nucleation rates, since the formation of complexes reduces the concentrations of the unpaired solutes. For a majority of cocrystal systems, these parameters are not known, hindering the ability to validate any cocrystal nucleation model. Therefore, to demonstrate the potential range of behavior captured in this model, we select a few examples using previously reported

parameter values whenever possible, using estimated values when none are available.

The microfluidic techniques described in Chapter 2 offer the potential to eventually obtain the needed data to validate the model. Computational methods like Molecular Dynamics or Density Functional Theory may also provide insights into the rates of formation of coformer complexes for small molecules, and light scattering could be used to measure the complexation rates for larger molecules like proteins [19].

To illustrate the range of possible behaviors within the model, we will first consider three examples in which we will initialize the simulation with identical values for all parameters except for the forward and reverse reaction rates. We will consider a constant ratio $k_f/k_r = 1$, and we will consider three scenarios: fast, intermediate, and slow solution complexation, in which k_r and k_f take on large, medium, and small values to illustrate different regimes of competition between complex formation and nucleus formation. We consider for illustrative purposes fixed initial solution concentrations of

$$C_{1i}(0) = 10 \frac{gm}{ml} \quad C_{2i}(0) = 8 \frac{mg}{ml} \quad C_{coi}(0) = 0 \frac{mg}{ml}$$

For the single component crystal pathways, the saturation concentration values were selection to correspond to a known cocrystal system reported by Holan *et al.* for agomelatine (component 1) and citric acid (component 2) in methanol at 26°C [1]. For this system, the saturation concentrations are given by

$$C_{1sat} = 0.369 \frac{mg}{ml} \quad C_{2sat} = 0.947 \frac{mg}{ml}$$

Since the component concentrations decrease with time, the value of supersaturation also decreases with time.

$$\sigma_1(t) = \ln\left(\frac{C_1(t)}{C_{sat1}}\right) = \ln\left(\frac{C_1(t)}{0.369 \frac{mg}{ml}}\right) \quad \sigma_2(t) = \ln\left(\frac{C_2(t)}{C_{sat2}}\right) = \ln\left(\frac{C_2(t)}{0.947 \frac{mg}{ml}}\right)$$

The initial supersaturations for the two coformers in this case $\sigma_1 = 3.30$ and $\sigma_2 = 2.13$.

The initial saturation concentration for the cocrystal nucleation pathway is defined by equation (23) where $[C_1] = C_{1i}$ and $[C_2] = C_{2i}$ are the initial concentrations of the two coformers, and K_{sp} is the solubility product reported as the product of the two coformer concentrations when a cocrystal is in phase equilibrium with the solute. Its value is reported as $K_{sp} = 0.0525$ for the agomelatine and citric acid coformers [1]. The instantaneous supersaturation for the cocrystal pathway is given by

$$\sigma_{co}(t) = \ln\left(\frac{C_{co}(t)}{C_{co sat}}\right) \quad (32)$$

$$C_{co sat} = 0.1 \frac{mg}{ml}$$

The fitting parameters A and B for agomelatine and citric acid in methanol are not readily available, so their values were chosen arbitrarily and kept fixed for this example. We will explore the effect of changing these values later in the chapter. The selected values for this illustrative example are

$$A_1 = 4 \cdot 10^3 \text{ } mg^{-1} \text{ } s^{-1} \quad B_1 = 10 \quad A_2 = 2 \cdot 10^3 \text{ } mg^{-1} \text{ } s^{-1} \quad B_2 = 10$$

$$A_{co} = 8 \cdot 10^3 \text{ } mg^{-1} \text{ } s^{-1} \quad B_{co} = 10$$

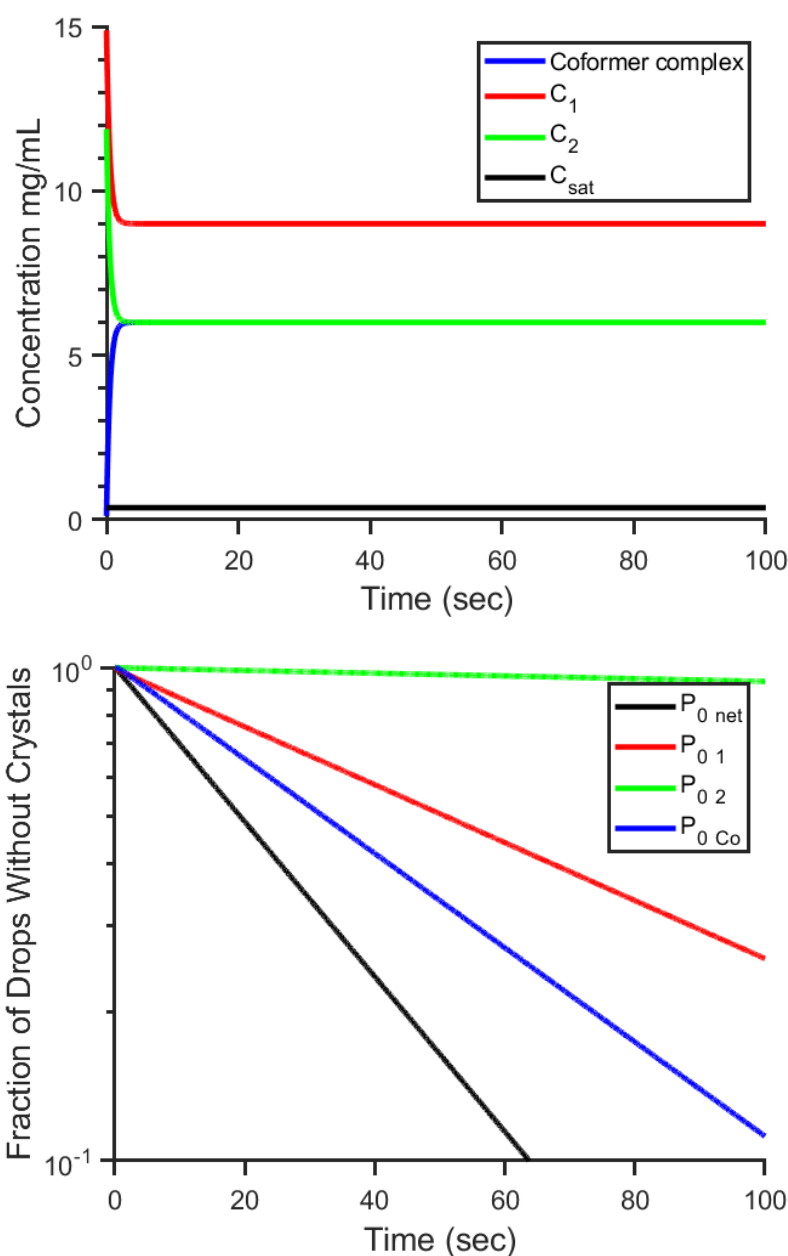


Figure 3-3 Fast solution complexation. (Top) Concentrations for pure components and coformer complexes and (Bottom) Cumulative probability functions versus time for pure component and co-crystals in the case of fast solution complexation, with rate constants $k_f=1$ and $k_r=1$, and all other parameter values held fixed. Rapid formation of coformer complexes causes all solution concentrations to reach their equilibrium values before appreciable nucleation has occurred. Note that the resulting cumulative probability function closely resembles that obtained from the classical cocrystal nucleation model.

The ratio of the rates k_f and k_r is kept constant so that the equilibrium concentration remains fixed for these examples. Note that we happened to select comparable nucleation parameters, and since the equilibrium concentrations are fixed, the net probability distribution will have a similar final slope regardless of which crystal species ends up dominant at the end of the process. A fixed volume of 1 mL was considered for this example.

For the case of fast solution complexation, the values considered here for the complex formation rate constants are

$$k_f = 1 \text{ s}^{-1} \quad k_r = 1 \text{ s}^{-1}.$$

Figure 3-3 shows two graphs. The top graph plots concentration on the y-axis versus time on the x-axis. The red and green lines represent the concentrations of the two coformers and the blue line represents the coformer complexes. The horizontal black line shows the saturation concentration for the materials in solution. The bottom graph shows the fraction of drops without crystals on the y-axis versus time on the x-axis, shown with the same scale as the top graph. The lines represent the probability of not finding a crystal for each of the three components in solution versus time, with red and green representing the two single component crystals and blue representing the cocrystal. The black curve represents the probability of not finding any of the three possible crystals. At these conditions the fast formation of coformer complexes brings the component concentrations to equilibrium rapidly, and nucleation occurs at effectively constant concentrations during most of the process. This reduces the complexity of the model since the time dependence of the nucleation rate J_x decays rapidly and the rates

are nearly independent of time after a short initial transient. This behavior allows us to approximate the behavior of the system as three parallel nucleation pathways that each follow Classical Nucleation Theory. In this case, we expect the final crystals observed to contain a mix of pure-component and cocrystals that is determined by the relative magnitudes of the nucleation rates for each pathway. In other words, if the single component nucleation rate is much faster than the cocrystal nucleation rate, single component crystals should make up most of the final crystal mixture. The lack of dependence of the concentration on time simplifies the analysis for this case. Note that with the rapid establishment of the coformer complexes, and consequent reduction of the single component concentrations, the cumulative cocrystal probability distribution is consistently the lowest of the three independent probabilities. Recalling that these values indicate the probability of *not* finding a crystal, this result indicates that cocrystals are always the most likely to form in this case.

When the rate of coformer complex formation is comparable to the rate of nucleation, we consider this an intermediate complexation rate. To illustrate this regime, we consider values of the reaction rate constants of

$$k_f = 0.1 \text{ s}^{-1} \quad k_r = 0.1 \text{ s}^{-1}$$

Figure 3-4 plots the same quantities as the figure above, with the top graph showing the solution concentrations and the bottom graph showing the probability distribution that results from those concentrations. The slower complex formation rates cause a longer transient for the initial concentration to evolve to the equilibrium concentrations. During this transient period, the concentrations of the pure component molecules steadily decrease, reducing the rates of nucleation for each until the concentrations

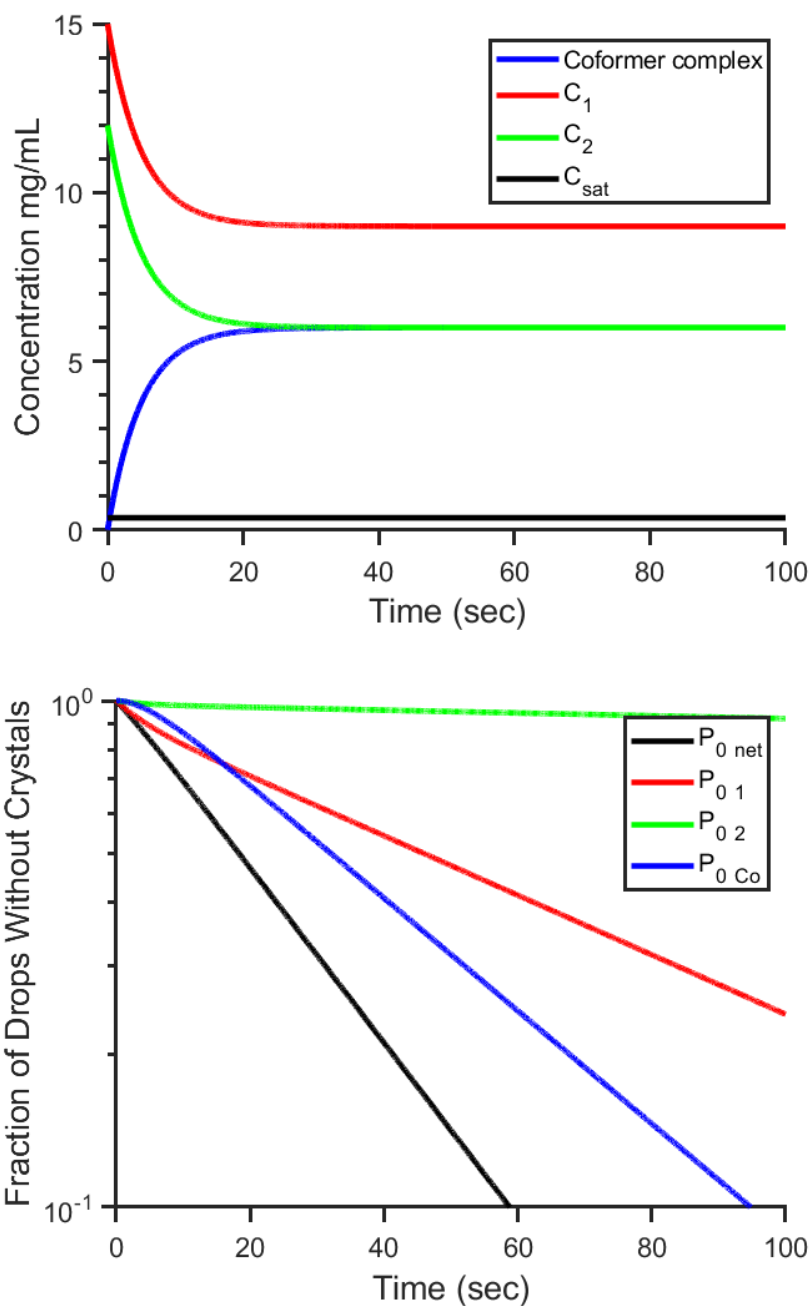


Figure 3-4 Intermediate coformer complexation. (Top) Concentration and (Bottom) probability function for each component versus time. All parameter values are held constant at values used in earlier cases, except the forward and backward rate constants which are $k_f = 0.1$ $k_r = 0.1$. Substantial nucleation has occurred before the solution concentrations reach their equilibrium values. As the coformer complex concentration increases the cocrystal nucleation rate overtakes the single component nucleation rate, and their probability distributions cross at approximately 35s. After the solution concentrations have reached equilibrium, the probability functions approach single decaying exponential functions.

reach a constant value. Simultaneously, the rate of cocrystal nucleation rises during the transient until it also reaches a constant value. This interplay leads to the observed shapes of the individual probability curves. Similar to the fast complexation case, we can compare the relative magnitudes of the probability functions to determine which crystal types are most likely to form at these conditions, with the lowest probability corresponding to the highest likelihood of finding a crystal of that species at that time. In Figure 3-4 the component 1 probability is dominant for approximately the first 18s, at

which point the cocrystal takes over as the dominant crystal pathway. This behavior is more complex than the CCNT model, and indicates that some component 1 crystals are likely to have nucleated before a majority of cocrystals form. Being aware of complex dynamics like this can help in designing more selective nucleation processes. Due to the stochastic nature of the process, one cannot be completely certain that cocrystals will not be present. At these conditions P_1 decreases rapidly from its initial value, indicating that the probability of finding a component 1 crystal is rapidly increasing at early times. However as the solution concentrations approach equilibrium the slope of P_1 begins to decrease until it reaches a new lower constant value. Simultaneously P_{co} has an initial slope of zero since no coformers are initially present. Once coformer complexes begin to form, the slope of the probability function steadily increases until it reaches a constant slope after the solution concentrations have reached equilibrium. The interplay between these two curves is explained through equations (26) through (31) which detail the relationship between the concentrations in solution and the

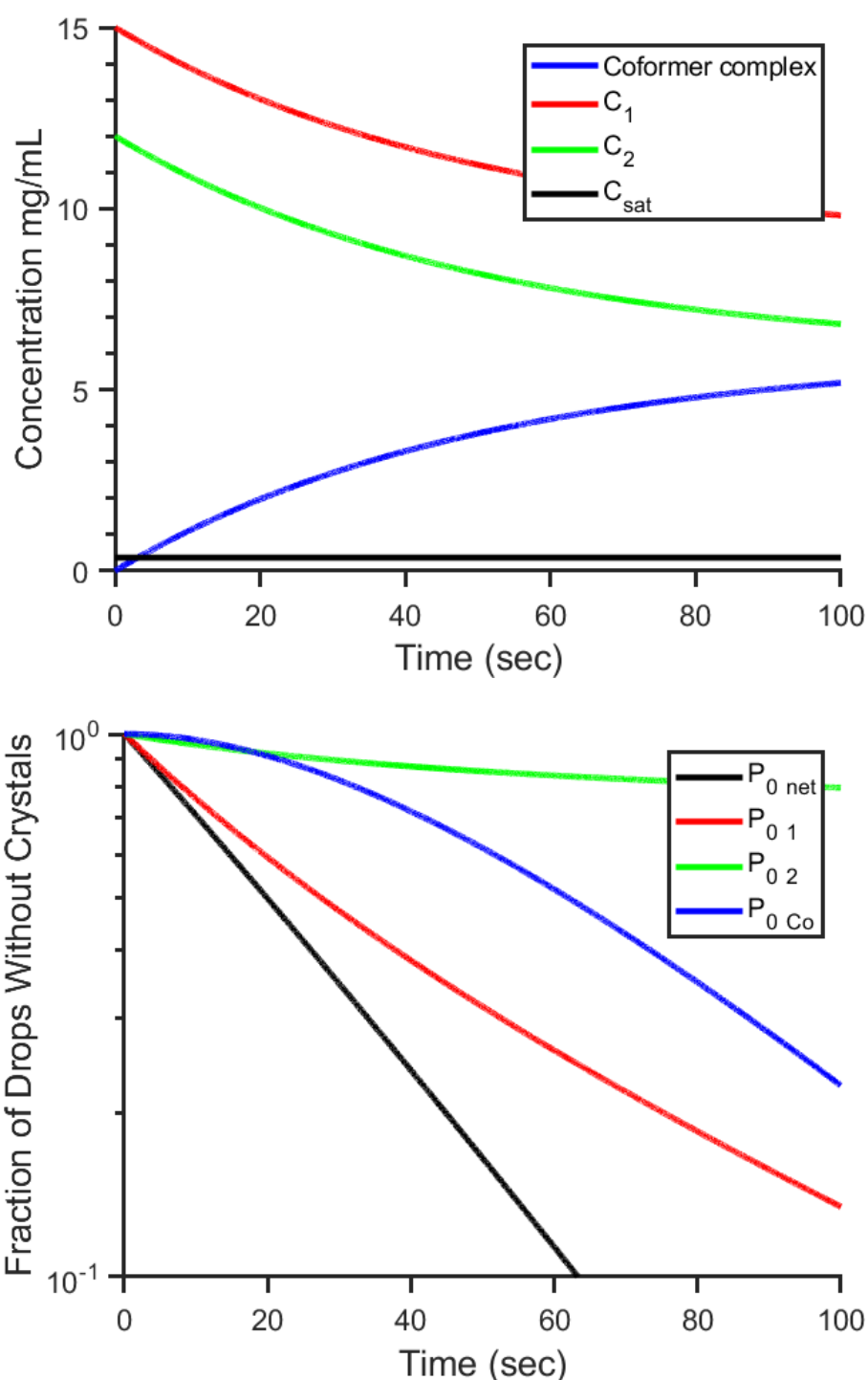


Figure 3-5 Slow coformer complexation. (Top) Concentration and (Bottom) probability versus time for individual cofomers and their complexes. Concentrations do not reach equilibrium before the majority of volumes would contain a crystal. Simple exponential behavior is not observed for the probability function, and cocrystals are never the dominant contributor to the net probability.

resulting nucleation rates. As component 1 is consumed to produce the coformer complexes its supersaturation decreases, lowering the nucleation rate.

Finally, to consider the case in which the rate of complex formation is much slower than the rate of nucleation we assign the following values:

$$k_f = 0.01 \text{ s}^{-1} \quad k_r = 0.01 \text{ s}^{-1}$$

The top graph in **Figure 3-5** plots the solution concentrations as a function of time, while the bottom graph shows the probability functions that result from these concentrations. The concentrations of each component do not reach equilibrium during the timeframe

observed here and instead they continue to change over the entire period, with the two coformers decreasing in concentration and the coformer complexes increasing in concentration. Initially the probability function for each coformer decreases steeply from its initial value, and the rate of decreases slows gradually as time progresses. The cocrystal probability initially exhibits a long initial period in which the probability does not change, and then once coformer complexes begin to form, the rate of formation gradually increases as the concentration of coformer complexes increases with time. None of the probability distributions reach a constant slope by the end of this time span and instead the nucleation rates continue to change throughout the observation time. In this case, throughout the entire observed time, the component 1 probability distribution is always the lowest, meaning you are most likely to find a single component crystal of that species at all times observed here. Coformer complexes have not reached an equilibrium concentration by the time experimental observations are completed and

therefore nucleation rates change steadily throughout the experiment. At these conditions we expect to see very few cocrystals in the final crystal mixture, as single component pathways dominate for a vast majority of the time. For systems that behave this way, homogenous nucleation from solution may result in radically different outcomes from seeded crystallization methods. Even though cocrystals could be favored at equilibrium, if the kinetics don't permit a pathway from solution to cocrystal that can favorably compete with single component nucleation, cocrystals may not be observed in the final product.

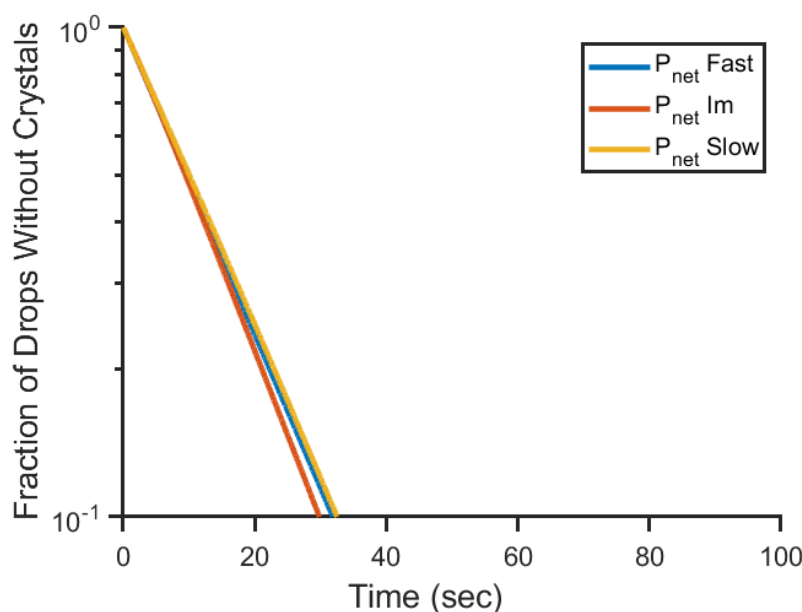


Figure 3-6 Net probability for crystal formation versus time for fast, intermediate and slow complexation rates. This figure compares the net probability $P_{0\ net}$ for the formation of any crystal (either pure component or cocrystals) corresponding to each of the three relative complexation rates considered in **Figure 3-3** through **Figure 3-5**.

While the previous three examples have shown us how the probability distribution for each individual crystal species can be significantly affected by the complex formation rate, comparing the net probability curves reveals a different picture.

The net probability curves for each of the three cases are graphed in **Figure 3-6** with the probability of not finding any crystal plotted on the y-axis versus time on the x-axis.

Each of the three curves shown is the $P_{0\ net}(t)$ taken from the bottom graphs of **Figure**

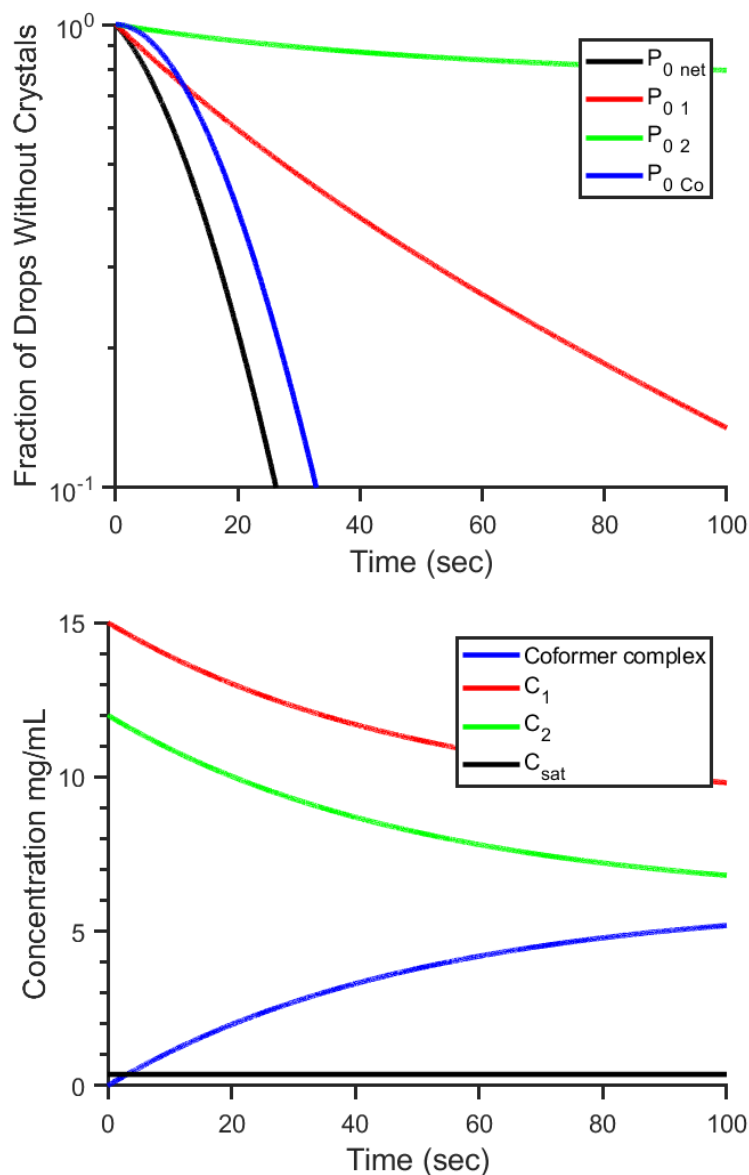


Figure 3-7 Fast Cocystal Nucleation. (Top) Concentration and (Bottom) probability function versus time for a cocystal nucleation rate ten times faster than that considered in Figures 3-4 through 3-6. Although the slowest coformer complexation rate was used here, the faster cocystal nucleation rate allows cocystals to become the dominant species forming after approximately 20s.

3-3 through **Figure 3-5**. Together these net probability curves show that, for this set of nucleation conditions, there is minimal difference in the overall likelihood of finding any type of crystal regardless of the magnitudes of the complex formation rates. Even still, the specific composition of crystal species may vary. This observation arises from the similar magnitudes of the nucleation kinetic fitting parameters selected for all three components. In the previous examples only the rate of formation of complexes controlled which crystal species dominated. Next we will explore another example where the relative magnitudes of the nucleation rates are varied at a constant complex formation rate.

In **Figure 3-7** we again plot the component concentrations versus time on the top graph and the resulting probability functions on the lower graph. In this case, all parameter values from the previous three examples are maintained except that the cocrystal pre-exponential factor is increased by an order of magnitude, thereby increasing the cocrystal nucleation rate. the slowest coformer complex formation rate as shown in **Figure 3-5** is used in this example. the parameter values used are:

$$A_{co} = 8 \cdot 10^4 \text{ mg}^{-1} \text{ s}^{-1} \quad B_{co} = 10$$

In this case, the results show that the probability of not finding cocrystals reaches a lower value than the component 1 crystal probability within 15 s, indicating that it is more likely to find a cocrystal at all times after 15s. Recalling that nucleation rate is linearly correlated with the value of the parameter A , we expect the nucleation rate for the cocrystal to be significantly faster for these parameter settings, so the observed behavior in **Figure 3-7** is consistent with our expectation. Cocrystal formation does not dominate the process initially because the concentration of coformer complexes must

first grow until it exceeds the coformer complex saturation concentration. As the complex concentration rises, the faster nucleation rate allows the cocrystals to take over the process.

Lastly, we compare the case shown in **Figure 3-7** with the case shown in **Figure 3-8** in which the following parameters were used,

$$k_f = 0.1 \text{ s}^{-1} \quad k_r = 0.1 \text{ s}^{-1}$$

$$A_1 = 4 \cdot 10^4 \text{ mg}^{-1} \text{ s}^{-1} \quad B_1 = 10 \quad A_{co} = 8 \cdot 10^3 \text{ mg}^{-1} \text{ s}^{-1} \quad B_{co} = 10$$

Here the cocrystal pre-exponential factor is returned to its original value, and the component 1 pre-exponential factor is increased by an order of magnitude. The rate of complex formation is selected to equal the values used in the intermediate case presented in Figure 3-5. In this case, the component 1 probability distribution dominates the process even after the solution has reached a concentration equilibrium.

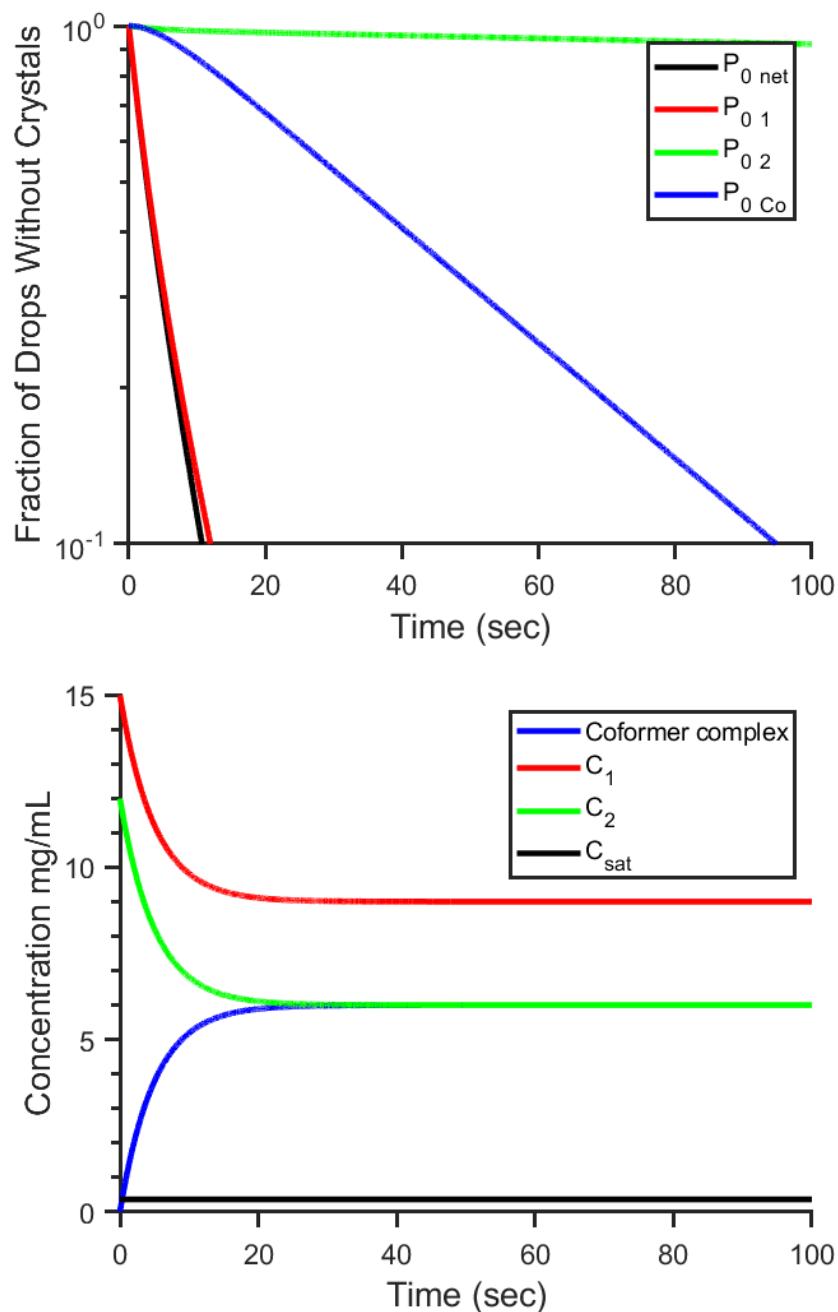


Figure 3-8 Fast nucleation of component 1. (Top) Concentrations and (Bottom) probabilities versus time for fast nucleation of one of the pure components and intermediate complexation rates. The fast rate of nucleation of Component 1 allows it to continue to dominate the process even after the solution has reached concentration equilibrium.

Dimensional Analysis

The previous examples show that the competition between the several different process timescales during cocrystallization can yield very different results, even from the situation that may be favored at equilibrium. The complexity of the cocrystallization process and the multiple competing timescales suggest that it may be useful to identify dimensionless numbers to compare the relative magnitudes of the governing timescales. As we described in the previous section, the lowest value of P among the three independent crystallization pathways determines which species has the greatest probability of nucleating a crystal at that time. Assuming that the concentration of coformer complexes is initially zero, the value of the cocrystal nucleation rate J_{co} will also be zero initially. This means that for cocrystals to become the dominant nucleation pathway, they must develop a nucleation rate greater than both of the single component crystal nucleation rates, and this higher rate must be maintained for long enough to result in $P_{0,co}(t) < P_{0,1}(t)$ and $P_{0,co}(t) < P_{0,2}(t)$. Since nucleation is driven by supersaturation, the rate of cocrystal nucleation is maximized when the solution concentrations have reached equilibrium. Equilibrium also results in the minimum values of J_1 and J_2 . The equilibrium concentrations are determined by the values of k_f and k_r . Therefore we can *a priori* calculate the values of J_{co} , J_1 , and J_2 at the equilibrium concentrations to determine whether cocrystals will be the dominant nucleation pathway. The nucleation ratios can be obtained by dividing equation (31) by equation (29) and equation (31) by equation (30),

$$Nr_1 = J_{co}/J_1 \quad (33)$$

$$Nr_2 = J_{co}/J_2 \quad (34)$$

If both of these ratios are greater than unity at the equilibrium concentrations, then there will be a time at which the slope of P_{co} will become steeper than that of P_1 and P_2 . With a steeper slope, it is possible for $P_{0,co}(t) < P_{0,1}(t)$ and $P_{0,co}(t) < P_{0,2}(t)$ given enough time. This can be seen in **Figure 3-4** at approximately 15s, where the cocrystal probability curve (blue) crosses over the single component curve (red). If the values of Nr_1 or Nr_2 are equal to or less than unity, a crossover point will not occur and there will be no time at which cocrystals are the dominant nucleation pathway. Since the nucleation rates J_{co} , J_1 , and J_2 depend on the nucleation fitting parameters A and B , as well as the concentrations determined by k_{eq} where

$$k_{eq} = k_r/k_f, \quad (35)$$

then we can find analytical solutions to the system of equations that define the values of J_{co} , J_1 , and J_2 at equilibrium. **Figure 3-9** shows the ratio of the pre-exponential factors A_{co} to A_1 on the y-axis versus k_{eq} on the x-axis. Each line represents a different ratio of B_{co} to B_1 . Each line on the plot depicts the critical ratio of nucleation fitting parameters that is needed in order for a crossover point to exist for a given equilibrium constant and corresponding equilibrium concentration. For ratios of the pre-exponential factors that are greater than the critical values, a crossover point exists.

The presence of a crossover point is necessary but not sufficient for designing a system with the goal of producing cocrystals. The time at which the crossover point occurs must also be considered. If the crossover point occurs after a significant amount

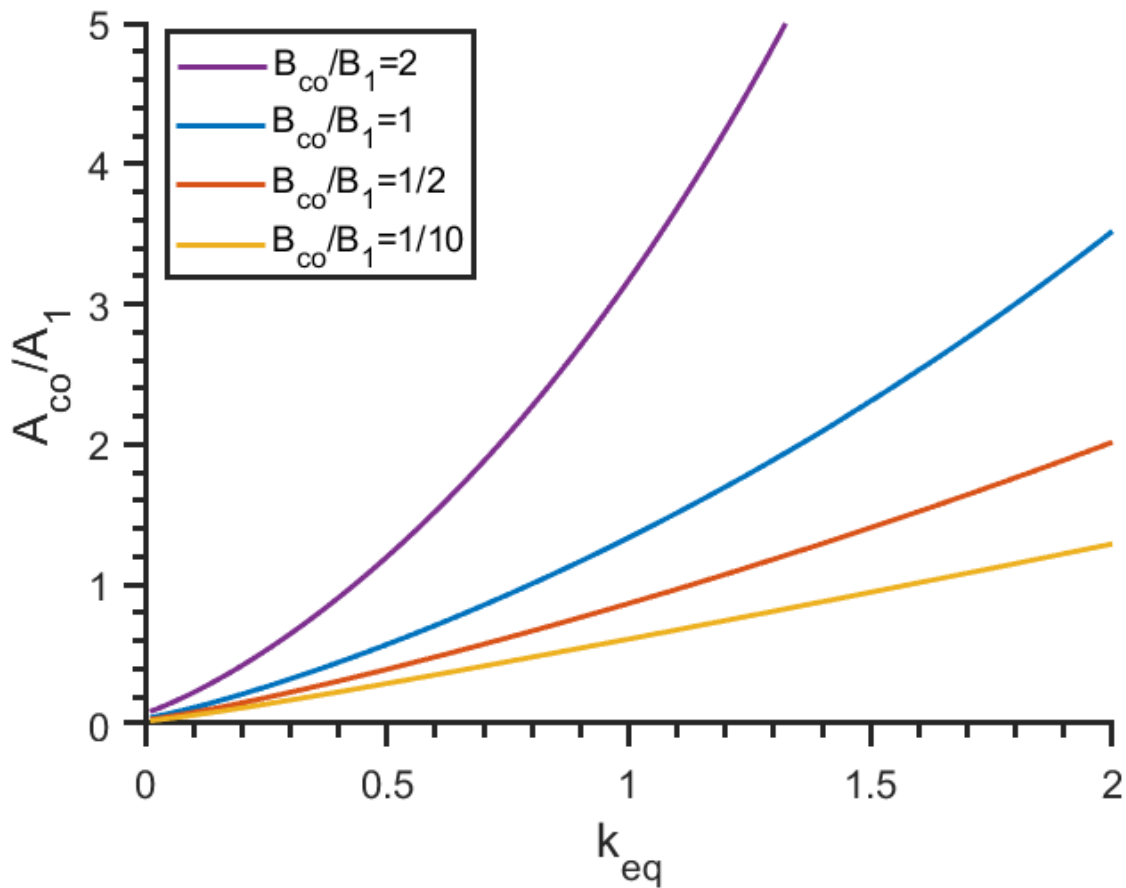


Figure 3-9 Pre-exponential factor ratio versus equilibrium constant. Equilibrium constant is plotted on the x-axis and the ratio of the pre-exponential fitting parameters for cocrystals and component 1 is plotted on the y-axis. The lines indicate the critical ratio of pre-exponential factors that gives $N_r = 1$ for a given k_{eq} and ratio of exponential fitting parameters. A crossover point exists if the system parameters fall into the space above the corresponding critical line.

of time, then it is possible that substantial amounts of a single component crystal will have nucleated before cocrystal formation could become the dominant pathway. For this reason, it is preferable for the crossover point to occur as early in the process as possible. The value of the crossover time is difficult to determine analytically since the concentrations vary with time, but it can be easily obtained by numerically solving the model equations.

Let us explore how several parameters affect the crossover time, starting with the forward reaction rate of complex formation k_f as shown in **Figure 3-10**. The figure shows the value of $P_{0,net}$ at which cocrystals become the dominant nucleation pathway on the y-axis versus the formation rate of cofomers on the x-axis. Here the value of k_{eq} is kept constant for each plotted curve by changing both the forward and backward rate proportionally. A fixed value of the equilibrium constant also fixes the equilibrium concentrations, but the timescale to reach equilibrium is shorter for larger complex

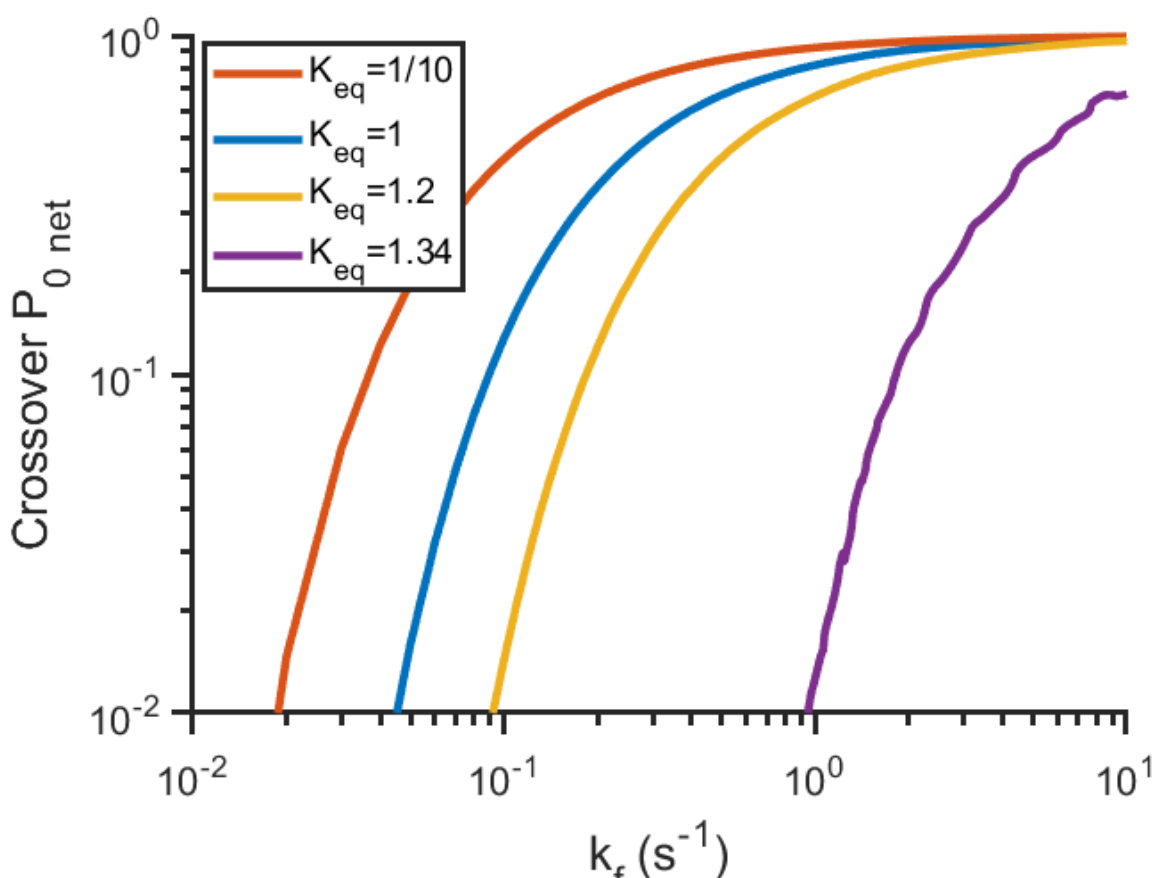


Figure 3-10 Crossover $P_{0,net}$ versus forward complex formation rate. The $P_{0,net}$ value at the crossover time is plotted as a function of the formation rate of cofomer complexes. Note that the case of $k_{eq}=1.34$ corresponds to $Nr_1 = 1.01$, so it is very close to not having a crossover time and the relatively low probability values led to some instability in the numerical integration.

formation rates. We can use the net probability as a dimensionless proxy for time since $P_{0\text{ net}}$ decreases monotonically with time for a given experiment. Therefore, the lines on Figure 3-10 represent the dimensionless crossover time, expressed as $P_{0\text{ net}}$ for each value of k_{eq} . This dimensionless form allows for direct comparison of the trials corresponding to different parameters. Since $P_{0\text{ net}}$ is the probability of *not* finding a crystal of any species, the closer the crossover probability is to unity, the less likely it is for the system to have nucleated a single component crystal. **Figure 3-10** clearly shows that there are diminishing returns for a wide range of k_{eq} , where faster complex formation rates do not translate into significantly higher crossover probabilities. Generally, values of k_f greater than 2 cause the crossover probability values to be greater than 0.9. Thinking in terms of the droplet based experiments described in Chapter 2, this result implies that less than 1 in 10 drops will have crystallized before cocrystals become the dominant nucleation pathway.

To optimize a microfluidic device in light of the crossover of the probability curves, several strategies can be employed. For static array devices, if the crossover point corresponds to a value of $P_{0\text{ net}}$ equivalent to or less than $1/N$ for the array, it can be assumed that single component crystals will be the dominant pathway at the time of crystallization in less than one droplet. For continuous droplet trains, the nearly infinite number of drops implies that sufficient droplets can always be generated to resolve the crossover probability value. However, since continuous droplet train devices have a limited overall incubation time given by equation (14), the crossover time would need to occur within that overall device timeframe in order to be observed. Because the drops are directly observable while nucleation is occurring, any drop that forms crystals in a

section of the residence channel that corresponds to a time earlier than the crossover point could be sorted differently before collecting the drops to perform post-analysis of the crystal mixture.

The above investigation of the cocrystal nucleation model with coformer complex formation reveals a competition between the rate at which the complexes form and the rate at which crystals nucleate. Thus, we can construct a dimensionless parameter from the associated characteristic timescales for each process. The rate at which coformer complexes form is governed by the forward rate constant k_f which has units of s^{-1} for first order reactions. The nucleation process is governed by the rate of nucleus formation per unit volume multiplied by the volume, Jv , which also has units of s^{-1} [42-43]. Dividing these two quantities results in a dimensionless parameter that indicates the likelihood of formation of a coformer complex or a crystal.

$$Pu = \frac{k_f}{Jv} \quad (36)$$

Since there are multiple nucleation pathways in a cocrystal system, there will be an associated dimensionless parameter for each of those pathways.

$$Pu_1 = \frac{k_f}{J_{0,1}v} \quad (37)$$

$$Pu_2 = \frac{k_f}{J_{0,2}v} \quad (38)$$

$$Pu_{co} = \frac{k_f}{J_{eq,co}v} \quad (39)$$

When Pu is greater than unity, the system will produce coformer complexes at a faster rate than it will produce crystals of that kind. Under the assumption that coformer

complexes form, the concentrations within the solution change with time, and so do the nucleation rates. For this reason, it is important to compare the maximum values of J during the process. For the case of single component crystals, the maximum nucleation rate occurs at the initial concentrations, and for the coformer complexes the maximum nucleation rate occurs when the concentration of complexes has reached equilibrium with components 1 and 2. To denote this, J for single components is given the subscript 0 and the species number, and J for cocrystals is given the subscript eq. It is important to recall that the value J depends both on material properties such as the diffusion coefficient and experimental parameters like temperature and concentration, so some degree of control over the rate values can be asserted. As mentioned above, the timescale for complex formation also varies over a wide range depending on the material, so Pu can take on a large range of values.

Interestingly, the dimensionless parameter Pu reveals that the volume of the nucleating solution directly influences the relative rate of crystal formation versus complex formation. As discussed in Chapter 2, nucleation can be carried out in volumes ranging from 1000L to picoliters depending on the experimental platform used. Therefore the choice of experimental platform can shift the value of Pu dramatically above or below unity, meaning that the volume of the reaction vessel can have a profound effect on which process is dominant. This can be seen in **Figure 3-11**, which plots the probability function versus time for two different volumes. In the top graph, the volume is 1 mL, which is the same as that considered in earlier examples. In the bottom

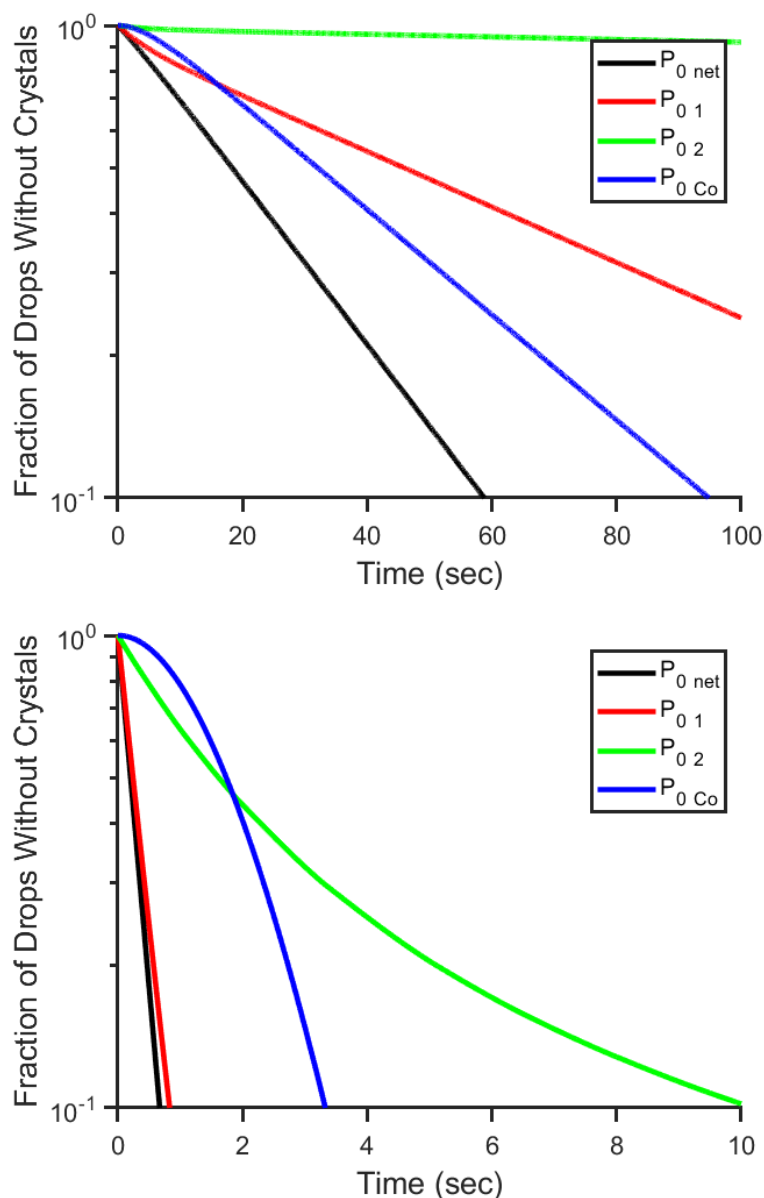


Figure 3-11 Effect of volume on cocrystal probability function. The two plots above use identical parameters except that the top graph considers a solution volume of 1 mL and the bottom graph considers a solution volume of 100 mL. Note that the timescale is dramatically shorter for the 100 mL case due to the increased nucleation rates in the larger volume. This also causes the crossover point to occur at a significantly lower $P_{0 \text{ net}}$ meaning that the single component crystals are the dominant pathway for a more significant portion of the process using larger volumes.

graph, a significantly larger volume of 100mL is considered, with all other parameters values held fixed. When the nucleation volume is 1mL, the cocrystal probability overtakes both of the pure crystal probabilities at a net probability of $P_{0\ net} = 0.8$. However, when the volume is increased to 100mL the cocrystal probability overtakes only the pure component 2 crystal curve at $P_{0\ net} = 0.45$ and it never overtakes the pure component 1 crystal curve. This is reflected in the values of Pu for each condition, as shown in **Table 3-2**. Therefore, changing the solution volume for fixed nucleation parameters can shift the dominant process from complex formation to nucleation as the volume increases. Not only does the qualitative nature of the probability distribution change from being mostly linear at small volumes to nonlinear at higher volumes, the relative magnitudes of each probability function also change. This result clearly demonstrates the importance of using dimensionless parameters – in this case to avoid unexpected behavior during scale up of the cocrystallization process. Therefore, the parameter Pu is an important metric to check when reliable nucleation of cocrystals is desired. If a cocrystal is difficult to produce at larger volumes due to an unfavorable Pu value, it may be possible to produce seed crystals in a microfluidic system, where the significantly lower reaction volumes shift the balance between the competing rates.

Table 3-2 Relative rates at different volumes

	Pu_1	Pu_2	Pu_{co}
1ml	1.9193	12.2149	1.9556
100ml	0.0192	0.1221	0.0196

While the parameter values chosen for this example were selected arbitrarily, they are within reasonable ranges based on the rates observed experimentally in Chapter 2.

Conclusions

In this chapter, we developed two simplified analytical models describing the process of homogenous cocrystallization from solution. There are several potential pathways by which a dilute mixture of co-formers can nucleate cocrystals, including the Classical Cocrystal Nucleation Theory (CCNT) model which assumes no association between coformers before nucleation, and cocrystal stoichiometry is enforced as particles add to the nucleus. We also considered the Initial Complex Cocrystal Nucleation Theory (ICCNT) model in which the coformers form complexes in solution, and the complexes then act as independent nucleating species. In this case the component ratios within the complexes enforce the stoichiometry of the resulting cocrystals. The formation of coformer complexes prior to their crystallization is the mechanism that is most clearly supported by the literature. The predictions made by the CCNT and ICCNT models yield distinct characteristics that can be examined in experimental results. The shape of individual probability distributions are unique between these two models. The CCNT model predicts all simple exponential curves with no time dependence in slope. The ICCNT model predicts that cocrystal probability curves will always have zero slope initially that becomes more negative over time at a rate related to the formation rate of coformer complexes. Independent single component crystallization experiments can provide the nucleation rate parameters for single component crystals, allowing the cocrystal nucleation rate to be isolated and used to determine the rate of formation of coformer complexes. If the rate at which coformer

complexes form is rapid compared to nucleation, it may be difficult to differentiate between the CCNT model and the ICCNT model. This is because the CCNT model has no lag time before cocrystallization occurs, so in the limit of rapid complex formation, these two pathways are indistinguishable.

The existence of competing rates must be considered when scaling up cocrystal nucleation processes. Notably, the nucleation rate for each species scales with the volume of the reaction vessel, but the rate of formation of coformer complexes is independent of volume. Thus, the model presented here can help provide insight into whether single crystals or cocrystals are likely to form and at which rates and which fractions given the conditions of interest. Additionally, the model reveals the possibility that the tiny volumes found in droplet-based microfluidic devices may offer a powerful method for producing seed cocrystals for new cocrystalline materials. Seed crystals can then be used to efficiently scale up industrial processes.

References

- [1] Jan Holaň, "The construction, prediction and measurement of cocrystal ternary phase diagrams as a tool for solvent selection", *European Journal of Pharmaceutical Sciences*, Volume 63, 15 October 2014, Pages 124-131, <https://doi.org/10.1016/j.ejps.2014.06.017>
- [2] Qiushuo Yu, CocrySTALLIZATION of urea and succinic acid in "Nano-Crystallizer", *Chemical Engineering Science*, Volume 229, 16 January 2021, 116082, <https://doi.org/10.1016/j.ces.2020.116082>
- [3] Renato A. Chiarella, Making CocrySTALLS The Utility of Ternary Phase Diagrams, *Crystal Growth & Design*, Vol. 7, No. 7, 2007, <https://doi.org/10.1021/cg070218y>
- [4] Sarah J. Nehm, Phase Solubility Diagrams of CocrySTALLS Are Explained by Solubility Product and Solution Complexation, *Crystal Growth & Design*, Vol. 6, No. 2, 2006, <https://doi.org/10.1021/cg0503346>
- [5] Gislaine Kuminek, CocrySTALLS to facilitate delivery of poorly soluble compounds beyond-rule-of-5, *Advanced Drug Delivery Reviews*, Volume 101, 1 June 2016, Pages 143-166, <https://doi.org/10.1016/j.addr.2016.04.022>
- [6] Paolo P. Mazzeo, Designing a Palette of CocrySTALLS Based on Essential Oil Constituents for Agricultural Applications, *ACS Sustainable Chem. Eng.* 2019, 7, 17929–17940, <https://doi.org/10.1021/acssuschemeng.9b04576>
- [7] Sekhon, B. S., Nutraceutical CocrySTALLS: An overview, *RGUHS J. Pharm. Sci.* 2012, 2, 16 –25. <http://dx.doi.org/10.5530/rjps.2012.2.3>
- [8] Ranjit Thakuria, Pharmaceutical cocrySTALLS and poorly soluble drugs, *International Journal of Pharmaceutics*, Volume 453, Issue 1, 30 August 2013, Pages 101-125. <https://doi.org/10.1016/j.ijpharm.2012.10.043>
- [9] Palash Sanphui, Fast Dissolving Curcumin CocrySTALLS, *Cryst. Growth Des.* 2011, 11, 4135–4145 <https://dx.doi.org/10.1021/cg200704s>
- [10] Min-Sook Jung, Bioavailability of indomethacin-saccharin cocrySTALLS, *Journal of Pharmacy and Pharmacology* 2010; 62: 1560–156. <https://doi.org/10.1111/j.2042-7158.2010.01189.x>
- [11] Shashank P. Patil, Generation of 1:1 Carbamazepine:Nicotinamide cocrySTALLS by spraydrying, *European Journal of Pharmaceutical Sciences* 62 (2014) 251–257 <http://dx.doi.org/10.1016/j.ejps.2014.06.001>
- [12] The Reemergence of CocrySTALLS: The Crystal Clear Writing Is on the Wall Introduction to Virtual Special Issue on Pharmaceutical CocrySTALLS, *Crystal Growth & Design*, Vol. 9, No. 10, 2009, <https://dx.doi.org/10.1021/cg901002y>
- [13] Tolou-Ghamari Z, Zare M, Habibabadi JM, Najafi MR. A quick review of carbamazepine pharmacokinetics in epilepsy from 1953 to 2012. *J Res Med Sci* 2013;18:S81-S5

- [14] Ying, H.R., Zhang, J.N. and Jiang, C.J. (2021) Preparation of Carbamazepine-Nicotinamide Cocrystal. Open Access Library Journal, 8: e7605. <https://doi.org/10.4236/oalib.1107605>
- [15] Sachit Goyal, Microfluidic Approach to Cocrystal Screening of Pharmaceutical Parent Compounds, Cryst. Growth Des. 2012, 12, 6023–6034, <https://dx.doi.org/10.1021/cg3011212>
- [16] Takashi Kojima, High-throughput cocrystal slurry screening by use of in situ Raman microscopy and multi-well plate. International Journal of Pharmaceutics 399 (2010) 52–59. <https://doi.org/10.1016/j.ijpharm.2010.07.055>
- [17] Noriyuki Takata, Cocrystal Screening of Stanolone and Mestanolone Using Slurry Crystallization. Crystal Growth & Design, Vol. 8, No. 8, 2008. <https://doi.org/10.1021/cg800156k>
- [18] Enxian Lu, A rapid thermal method for cocrystal screening. CrystEngComm, 2008, 10, 665–668. <https://doi.org/10.1039/b801713c>
- [19] Christoph Loschen, Cocrystal Ternary Phase Diagrams from Density Functional Theory and Solvation Thermodynamics. Cryst. Growth Des. 2018, 18, 9, 5600–5608. <https://doi.org/10.1021/acs.cgd.8b00923>
- [20] Stefan Auer, Prediction of absolute crystal nucleation rate in hard-sphere colloids. Nature. , 2001, Vol.409(6823), p.1020-1023 <https://doi.org/10.1038/35059035>
- [21] Oleg Galkin, Direct Determination of the Nucleation Rates of Protein Crystals. J. Phys. Chem. B 1999, 103, 10965-10971, <https://doi.org/10.1021/jp992786x>
- [22] Richard P Sear 2007 J. Phys.: Condens. Matter 19 033101. <http://dx.doi.org/10.1088/0953-8984/19/3/033101>
- [23] S. Basavoju, D. Boström, S. Velaga Indomethacin–saccharin cocrystal: design, synthesis and preliminary pharmaceutical characterization. Pharm. Res., 25 (2008), pp. 530-541 <https://doi.org/10.1007/s11095-007-9394-1>
- [24] Gan-Bing Yao, Ternary Phase Diagram for Systems of Succinic Acid + Urea + Water, Glutaric Acid + Urea + Water, and Adipic Acid + Urea + Water at (288.15 and 303.15) K. J. Chem. Eng. Data 2014, 59, 12, 4081–4089. <https://doi.org/10.1021/je500725e>
- [25] Deniz Erdemir. Nucleation of Crystals from Solution: Classical and Two-Step Models. ACCOUNTS OF CHEMICAL RESEARCH 621-629 May 2009 Vol. 42, No. 5. <https://doi.org/10.1021/ar800217x>
- [26] Sathish V. Akella, Emulsion-Based Technique To Measure Protein Crystal Nucleation Rates of Lysozyme. Cryst. Growth Des. 2014, 14, 4487–4509. <https://dx.doi.org/10.1021/cg500562r>
- [27] Rodriguez-Hornedo, Nair. “Reaction Crystallization of Pharmaceutical Molecular Complexes.” Molecular pharmaceutics. 3.3 (2006): 362–367. Web.

- [28] Schultheiss et al., (2010). Pterostilbene cocrystals. (US Patent No. 8415507B2). U.S. Patent and Trademark Office.
- [29] Patel et al. (2017). Single-step production method for nano-sized energetic cocrystals by bead milling and products thereof. (US Patent No. 9850181B1). U.S. Patent and Trademark Office.
- [30] Zaworotko, M. (2008). Nutraceutical co-crystal compositions. (EU Patent No. 2167043A2). European Patent Office.
- [31] Bevill et al. (2013). Methods of making cocrystals. (US Patent No. 20130102781A1) U.S. Patent and Trademark Office.
- [32] Chen, Hongbo. "Spherical Cocrystallization—An Enabling Technology for the Development of High Dose Direct Compression Tablets of Poorly Soluble Drugs." *Crystal growth & design*. 19.4 (2019): 2503–2510. Web.
- [33] Zhang, Haobin. "Five Energetic Cocrystals of BTF by Intermolecular Hydrogen Bond and π -Stacking Interactions." *Crystal growth & design*. 13.2 (2013): 679–687. Web.
- [34] Sarma, Bipul. "The Role of π -Stacking in the Composition of Phloroglucinol and Phenazine Cocrystals." *Crystal growth & design*. 8.12 (2008): 4546–4552. Web.
- [35] Leyssens, T. "Solution Cocrystallization, an Effective Tool to Explore the Variety of Cocrystal Systems: Caffeine/dicarboxylic Acid Cocrystals." *CrystEngComm* 16.41 (2014): 9603–9611. Web..
- [36] Yin, X., Scalia, A., Leroy, L., Cuttitta, C. M., Polizzo, G. M., Ericson, D. L., ... Soares, A. S. (2014). Hitting the target: Fragment screening with acoustic in situ co-crystallization of proteins plus fragment libraries on pin-mounted data-collection micromeshes. *Acta Crystallographica Section D: Biological Crystallography*, 70(5), 1177–1189. <https://doi.org/10.1107/S1399004713034603>
- [37] Zhang, Hao. "Preparation and Characterization of Carbamazepine Cocrystal in Polymer Solution." *Pharmaceutics*. 9.4 (2017): n. pag. Web.
- [38] Chieng, N., Hubert, M., Saville, D., Rades, T., & Aaltonen, J. (2009). Formation kinetics and stability of carbamazepine-nicotinamide cocrystals prepared by mechanical activation. *Crystal Growth and Design*, 9(5), 2377–2386. <https://doi.org/10.1021/cg801253f>
- [39] Metrangolo, P.; Meyer, F.; Pilati, T.; Proserpio, D. M.; Resnati, G. Highly Interpenetrated Supramolecular Networks Supported by N...I Halogen Bonding. *Chem. Eur. J.* 2007, 13, 5765–5772
- [40] TACHIKAWA, Hiroto. "Timescale of π -Stacking Formation in a Benzene Trimer Cation Formed by Ionization of the Parent Neutral Trimer: A Direct Ab Initio Molecular Dynamics Study." *ChemistrySelect* 3.4 (2018): 1113–1119. Web.
- [41] Patrickios, Costas S. "Protein Complexation with Acrylic Polyampholytes." *Biotechnology and bioengineering*. 44.9 (1994): 1031–1039. Web.

- [42] Teychené, Sébastien. "Microfluidic Device for the Crystallization of Organic Molecules in Organic Solvents." *Crystal growth & design*. 11.11 (2011): 4810–4818. Web.
- [43] Goh, L.; Chen, K.; Bhamidi, V.; He, G.; Kee, N. C. S.; Kenis, P. J. A.; Zukoski, C. F.; Braatz, R. D. *Cryst. Growth Des.* 2010, 10, 2515–2521.
- [44] Hare, C and Ghadiri, M, The Influence of Aspect Ratio and Roughness on Flowability. 7th International Conference on Micromechanics of Granular Media (Powders and Grains). 2013 | POWDERS AND GRAINS 2013 1542 , pp.887-890
- [45] McNamara, Daniel P. "Use of a Glutaric Acid Cocrystal to Improve Oral Bioavailability of a Low Solubility API." *Pharmaceutical research*. 23.8 (2006): 1888–1897. Web.
- [46] Voloshin, V P. "Hydrogen Bond Lifetime Distributions in Computer-Simulated Water." *Journal of structural chemistry* 50.1 (2009): 78–89. Web.
- [47] M. R. Thorson, S. Goyal, B. R. Schudel, C. F. Zukoski, G. G. Zhang, Y. Gong and P. J. Kenis, *Lab Chip*, 2011, 11, 3829–3837.
- [48] Frederico L.F. Soares, Renato L. Carneiro..*Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 180 (2017) 1–8
- [49] Karthika, S. "Measurement of Nucleation Rate of Ibuprofen in Ionic Liquid Using Induction Time Method." *Journal of crystal growth* 521 (2019): 55–59. Web.
- [50] Nanev, Christo N. "Sigmoid Kinetics of Protein Crystal Nucleation." *Journal of crystal growth* 427 (2015): 48–53. Web.
- [51] Collet, Pierre; Martínez, Servet; Martín, Jaime San (2013). Quasi-Stationary Distributions | SpringerLink. Probability and its Applications
- [52] Wu, Hao-Di. "Preparation of Cocrystal Nanofibres of Cobalt Octaethylporphyrin and Tetracyanoquinodimethane with Good Photoresponse." *Journal of Nanoscience and Nanotechnology* 14.6 (2014): 4097–4100. Web.

Chapter 4 : Kinetic pathways to co-crystal nucleation

Abstract

As a crystallizing system transitions from a supersaturated solution to a suspension of solid crystals at equilibrium with the solution, it must traverse an evolving path through composition phase space. For cocrystal systems, the ability to produce multiple unique solid phases with their own phase equilibria provides significant complications to understanding this process. Utilizing the cocrystal nucleation model described in chapter 3, we explore the influence of nucleation rate on the intermediate crystal populations as the system progresses toward equilibrium. The path taken by a solution through composition space depends on the ratio of nucleation rates of the coformer materials. Certain regions within the phase space provide direct paths to cocrystal nucleation, while others ensure that there will be a mixed population of crystals nucleated.

Introduction

In this chapter we investigate the pathways by which a macroscopic solution of cocrystallizing materials proceeds toward equilibrium. Current literature on cocrystal systems focuses on the equilibrium phase diagram, but there is little information about how the kinetics of cocrystal nucleation may influence the evolution of the crystal population during the transition from supersaturation to phase equilibrium. This chapter uses the insights gained in Chapter 3 to explore the role of kinetics on the final composition of a cocrystal system, compared with the expectations based on equilibrium phase diagrams. Additionally, we will explore the influence of the nucleation

kinetics on the path a solution follows as it evolves from an initial supersaturated state through the formation of single-component crystals and cocrystals. Crystal formation reduces the coformer concentrations in solution and eventually all coformers reach their saturation concentrations. During the evolution from the supersaturated solution to the final equilibrium concentrations, a population of crystals will form, altering the solution composition. These compositional changes in turn influence the nucleation kinetics, further influencing the path a solution takes through composition space.

As Chapter 3 describes, the cocrystallization process is a result of competing nucleation pathways for the formation of single-component crystals and cocrystals. Therefore, depending on the initial solution conditions, it is possible to synthesize an entirely different set of crystal products [1]. Mapping out the transient evolution of the crystal mixture composition is important to understanding and controlling processes that involve cocrystal nucleation from solution [2]. To date, most studies have focused on determining the equilibrium phase behavior for cocrystal systems via methods like wet grinding, isothermal solvent evaporation, and seeded nucleation [3-7]. These methods are useful for determining how the system will behave at long times without considering possible kinetic barriers. Methods like wet grinding are energy intensive, using the input of energy into the system to help overcome kinetic barriers and enable a system to approach a stable equilibrium. Computational chemistry and DFT techniques have also been employed to calculate phase boundaries based on chemical potential differences [8-11]. **Figure 4-1** shows an example equilibrium ternary phase diagram for a cocrystal system. Each edge of the triangle represents the initial mole fraction of one of the three components of the system, with the solvent on the left and the two coformers on the

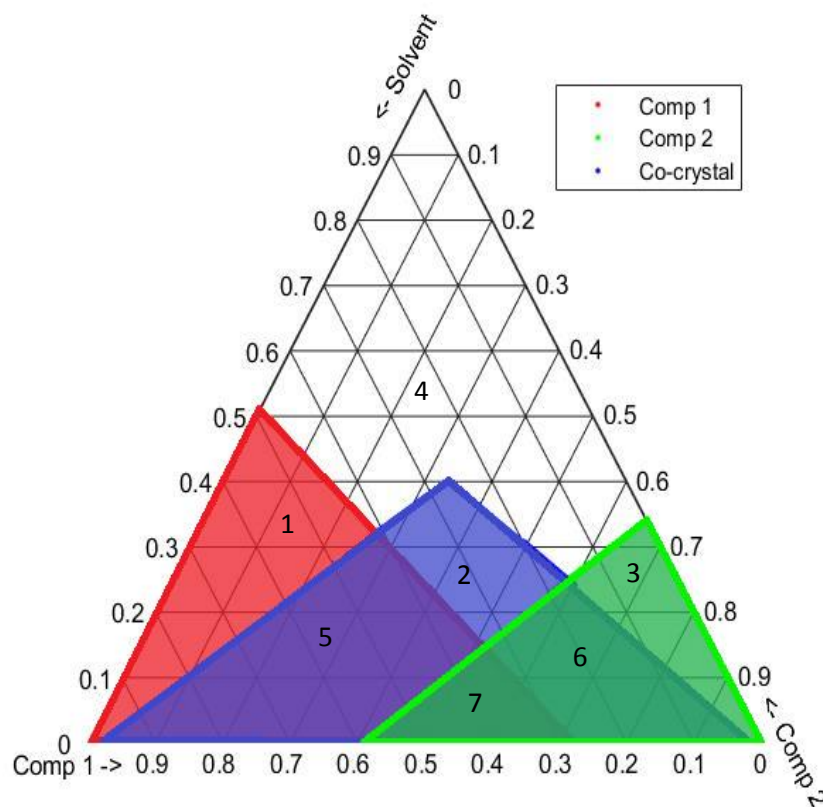


Figure 4-1 Cocystal ternary phase diagram. Each axis of the triangle represents the concentration of one of the components of the ternary mixture. Each edge of the triangle represents one of three possible binary systems for which the third component has a concentration of zero. Within the triangle, all three components are present in some proportion. Each numbered region represents a different crystal mixture: 1. Single component crystals of coformer 1; 2. Cocrystals containing both coformers; 3. Single component crystals of coformer 2; 4. This region is undersaturated and no crystals are found here; 5,6,7. Each region contains a mixture of crystals corresponding to the overlap of regions 1,2, and 3.

bottom and right. The edge of each shaded region of the diagram indicates the solubility limit with respect to each crystalline species in the system. All regions in which the concentration of a given component is above its solubility limit can contain crystals of that component at equilibrium. For heterogeneous crystal growth, in which seed crystals are introduced to the solution, crystals can grow in any region within the respective solubility curve [12]. When using equilibrium to describe the phase diagram of a

cocrystal system in this way, any potential path dependence is neglected. However, without considering the kinetics of nucleation, these diagrams may not fully represent the mixture of products that are accessible by homogenous nucleation from solution.

As described in Chapter 3, kinetics can play an important role in the composition of the crystals that nucleate over time. Here we will show how the kinetics impact the likelihood that a particular crystal species will be the first to nucleating at each set of starting concentration within a phase diagram. We will use the model described in Chapter 3 as a basis to examine this process. Knowing which crystal species nucleates first is relevant to designing a system which is intended to produce cocrystal materials, since once a crystal forms in solution it begins to grow, reducing the supersaturation in the surrounding solution. For cocrystal systems this can mean that the formation of a single component crystal could potentially suppress the formation of cocrystals or vice versa. Several factors impact the relative rates of nucleation for a co-crystal system. From the definition of supersaturation given in equation (23), we can see that for a given initial solution composition it is easily possible to have a difference in supersaturation of the co-crystal versus the single component crystal. Since the driving force for nucleation is supersaturation there is likely to be a difference in nucleation rate for each pathway, regardless of the precise mechanism for nucleation.

For the remainder of this chapter we will examine the role of nucleation kinetics on the transient evolution of the crystal mixture composition, using the Initial Complexation Cocrystal Nucleation Theory (ICCNT) model that assumes the formation of coformer complexes prior to nucleation. **Figure 4-2** shows a plot of cocrystal probability versus time for an assumed ICCNT system as described in Chapter 3. We

assume a two-component system that can form cocrystals or single component crystals, and we use the following parameters as defined in Chapter 3 to describe the nucleation kinetics and kinetics of formation of coformer complexes:

$$A_1 = 4 \cdot 10^3 \text{ mg}^{-1} \text{ s}^{-1} \quad B_1 = 10 \quad A_2 = 2 \cdot 10^3 \text{ mg}^{-1} \text{ s}^{-1} \quad B_2 = 10$$

$$C_{1 \text{ sat}} = 0.369 \frac{\text{mg}}{\text{ml}} \quad C_{2 \text{ sat}} = 0.947 \frac{\text{mg}}{\text{ml}} \quad C_{\text{co sat}} = 0.1 \frac{\text{mg}}{\text{ml}}$$

$$k_f = 0.01 \text{ s}^{-1} \quad k_r = 0.01 \text{ s}^{-1}$$

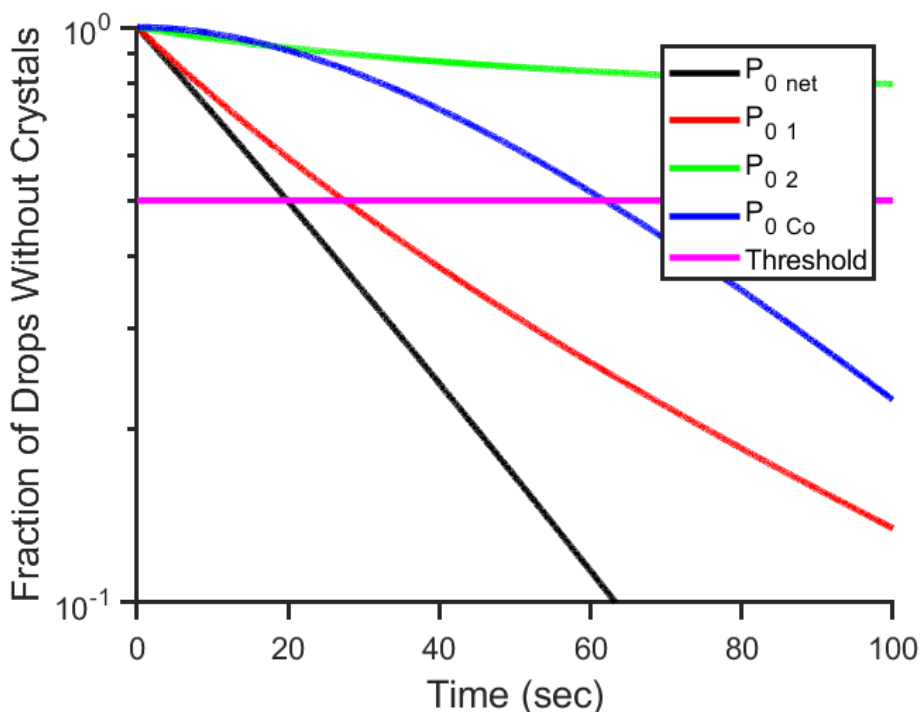


Figure 4-2 Probability function for fixed initial coformer concentrations with threshold value. The probability of not finding a crystal of a specific type is represented on the y-axis, as a function of time on the x-axis. The three lines labeled P_{01} , P_{02} and $P_{0\text{Co}}$ are the independent probabilities for each of the coformers and the cocrystal. P_{net} is the overall probability of not finding any of the three crystal types. The horizontal line at 0.5 depicts an arbitrary threshold at which there is a 50% probability of finding a crystal. Parameters used to calculate the presented curves are given in the text, and the initial coformer concentrations are $C_{i1}=10 \text{ mg/ml}$ and $C_{i2}=8 \text{ mg/ml}$.

These parameters are the same as those considered in Chapter 3 for the slow complexation kinetics and the nucleation fitting parameters are identical to those used in **Figure 3-5**. **Figure 4-2** shows a plot of the probabilities for crystal formation for each of the three crystallizing species in the system: coformers 1 and 2, and the coformer complexes. The figure represents one set of starting concentrations for the coformers, $C_{i1}=10\text{ mg/ml}$ and $C_{i2}=8\text{ mg/ml}$. The red and green lines represent the probability for the single component crystals for components 1 and 2, while the blue line represents the probability for the cocrystals. The solid black line is the net probability of not finding any of the three possible crystal species. The horizontal pink line depicts an arbitrary threshold value of $P_0=0.5$.

Probability curves such as those illustrated in **Figure 4-2** can be used to understand the cocrystallization process in several ways. For example, if we recall the droplet-based microfluidic systems described in Chapter 2, we would expect that at any given time, the fraction of droplets that have not yet formed visible crystals will be equivalent to the probability of not finding a crystal at that time. Although droplet-based systems do not currently permit identification of different crystal types, if in-line identification were possible, we would expect that the fraction of droplets that do not contain a specific crystal type will also be equivalent to its individual probability. Additionally, we can use the relative magnitudes of the probabilities of each crystallizing species to gain insight into the order in which each type of crystal will appear. For example, at or below a P_0 value of 0.5, it is more likely than not that a given crystal type will have formed in a given volume. Choosing $P_0 = 0.5$ as an arbitrary threshold value, we can examine **Figure 4-2** and note that the probability for formation of the single

component crystals of coformer 1 reaches the threshold first among the three species. Furthermore, at the time when $P_{0,1}$ reaches the threshold value, both of the other probability distributions are still relatively close to unity. In a droplet based system we would expect that about 50% of the drops should contain a component 1 crystal, while very few droplets contain cocrystals or single crystals of component 2, despite the system being supersaturated with respect to all three crystal species. Furthermore, if we assume that crystal growth is rapid, we may expect that the predominant formation of one crystal species could lead to the suppressed formation of other species since growth will reduce the overall solution concentrations and thereby reduce the driving forces toward crystallization of any species. This effect could be even more pronounced in systems with larger volumes in which formation of a localized crystal is more likely to affect the solution as a whole.

Model details

To set up a model to determine which crystal species is most likely to nucleate first we first need to define the nucleation parameters for each of the species involved in the cocrystallization process. The governing equations were defined in Chapter 3 in equations (29) through (31). In what follows we will continue to use the nucleation parameters and the coformer complexation parameters given earlier in this chapter. These parameters partially correspond to known values for the coformer system consisting of the API agomelatine and the inert coformer citric acid, as well as additional estimated parameters for those that are not available in previously published reports.

To construct a ternary phase diagram for this system, the concentration of each species is converted to a mole fraction and the volume of the solution is assumed to be the volume of solvent used. The mole fractions are given by

$$M_1 = \frac{C_1 * V}{mw_1} \quad (40)$$

$$M_s = \frac{\rho_s * V}{mw_s} \quad (41)$$

$$x_i = \frac{M_i}{M_1 + M_2 + M_s} \quad (42)$$

where M_1 and M_2 are the molarities of components 1 and 2, M_s is the molarity of the solvent, x_i is the mole fraction of the i th component, mw_1 and mw_2 are the molecular weights of the two coformers, and ρ_s is the density of the solvent. The values used for the example considered here are summarized in **Table 4-1** through **Table 4-3**.

Table 4-1 Coformer physical properties

	Molecular weight (g/mol)	Concentration range (mg/ml)
Component 1 (agomelatine)	243.3	0-15
Component 2 (citric acid)	192.1	0-15

Table 4-2 Solvent physical properties

	Molecular weight (g/mol)	Density (g/ml)
Solvent (Methanol)	32.04	0.791

Table 4-3 Nucleation and complex formation kinetic parameters

	Component 1	Component 2	Cocrystal
<i>A</i>	$4 \cdot 10^3 \text{ mg}^{-1} \text{ s}^{-1}$	$2 \cdot 10^3 \text{ mg}^{-1} \text{ s}^{-1}$	$8 \cdot 10^3 \text{ mg}^{-1} \text{ s}^{-1}$
<i>B</i>	10	10	10
Complex formation			$k_f = 0.01 \text{ s}^{-1}$
			$k_r = 0.01 \text{ s}^{-1}$

For illustrative purposes, the initial volume of the solvent is kept fixed. A minimum and maximum concentration range for each of the two coformers is selected as well as a concentration step size by which we will move linearly between these two extremes as the two coformer concentrations are varied. These bounds determine the area of the phase space that will be explored as well as the concentration resolution that will be probed within that region. The probability functions for each of the three crystallizing species are computed for a total process time of 100 s for each pair of initial concentrations. This generates a large data set describing the probability that each species will have formed a crystal at each time in the process for every initial concentration combination. From this large dataset we will now use construct new composition diagrams that show which species is the most likely to nucleate first at each set of initial concentrations.

Results and Discussion

First Nucleating Species Diagrams

For equilibrium phase behavior it is common to represent the data using a ternary phase diagram with an axis for each of the two co-formers and the solvent [13-16]. This same graphical format will be used to depict the crystal species that is first to nucleate a crystal. Equilibrium ternary phase diagrams depict the boundaries outlining regions in which specific species compositions are found once the system has stopped evolving. Here, we will instead use the kinetic information for the crystallization process to depict aspects of the evolution of the crystal mixture composition with time. As shown in **Figure 4-2**, $P_{0\text{ net}}(t)$ and the individual crystal species probabilities take on unique values for each set of initial concentrations and each time step. This creates a rich but very large dataset depicting evolution of the system, requiring us to consider new ways to parse the information and gain insight into the process. **Figure 4-3** shows one method of examining the time-dependent probability data. Here, the goal is to construct a diagram of the composition space indicating which species is likely to nucleate first at each set of initial concentrations.

The top graph in **Figure 4-3** illustrates the construction of the lower graphs in the same figure panel. The top graph shows the probability functions already presented in **Figure 4-2**, representing one set of initial concentrations, which corresponds to a single point on the lower two ternary phase diagrams, indicated by the red and blue arrows. The edges of these ternary phase diagrams are identical to those presented in **Figure 4-1**. We consider two different time points illustrated by the vertical black dashed lines

on the top plot. The first time point is $t=25s$, which corresponds to the time at which the probability for single crystals of coformer 1 reaches 50%. The second time point is $t=62s$, which is the time at which the probability of cocrystal formation reaches 50%. These time points can be considered “critical” points at which the species that has crossed the probability threshold has become more likely than not to appear in the

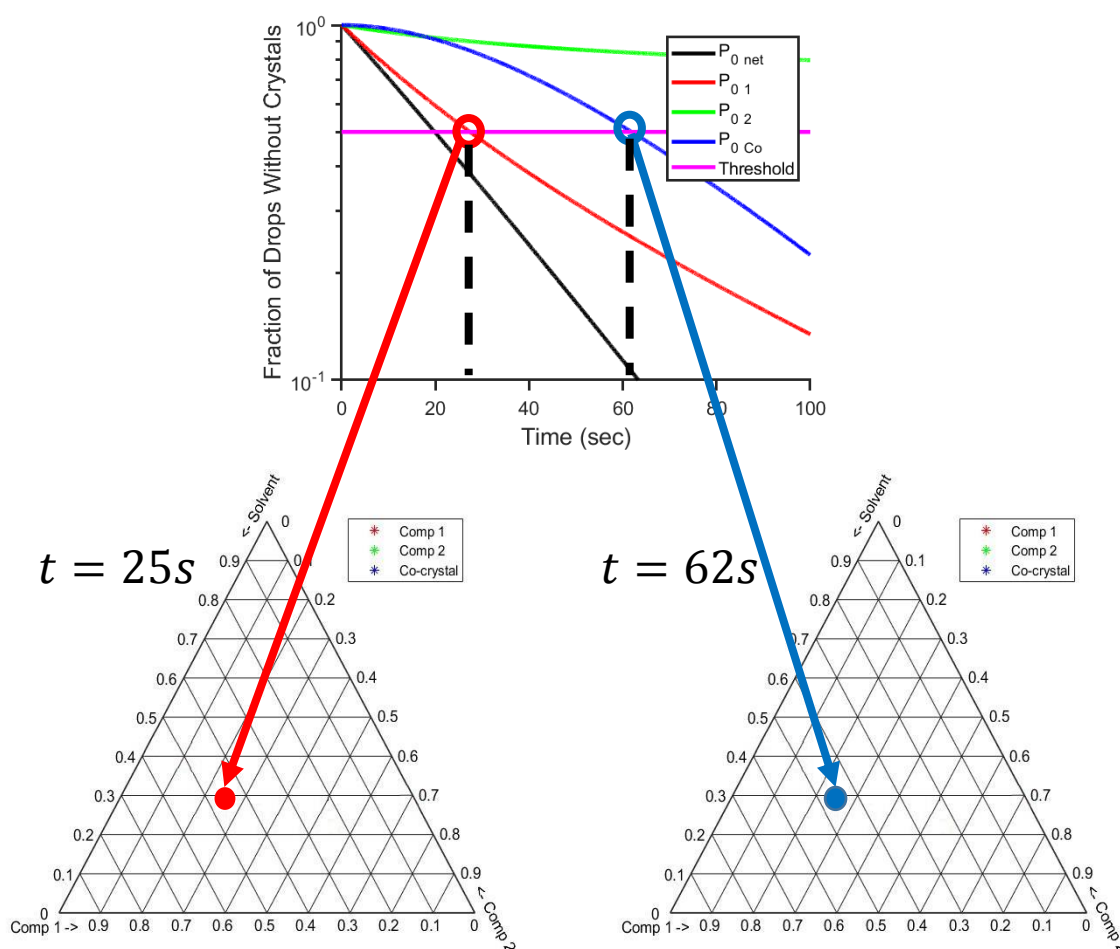


Figure 4-3 Translating probability distributions to transient ternary phase diagrams. Each location on the ternary phase diagram corresponds to a set of initial solution concentrations. At each set of initial conditions, the probability distribution for nucleation evolves over time as shown in the top graph. As time progresses the individual species probability curves will fall below a defined threshold value. $P_0 = 0.5$ is the threshold value indicated by the pink line in the top graph. When the threshold is reached for a given species, the corresponding point on the phase diagram is marked for that species at that time.

mixture. Note that the $P_0 = 0.5$ threshold value is arbitrary; similar diagrams could be created using other values. Using this information, we can create a ternary phase diagram corresponding to each time point. In **Figure 4-3** the lower left diagram indicates the status at time $t = 25$ s, and the lower right diagram indicates the status at $t = 62$ s.

On each time panel, a symbol of a given color indicates that the corresponding crystallizing species has reached the threshold probability for that set of initial concentrations. In **Figure 4-3**, a red symbol is placed on the lower left diagram corresponding to $t = 25$ s to indicate that coformer 1 is 50% likely to have formed a crystal. A blue symbol is placed at the same location (corresponding to the same initial concentrations) at $t = 62$ s to indicate that the cocrystal probability has reached 50%.

The concept illustrated in **Figure 4-3** can be extended to generate a complete ternary diagram depicting the boundaries at which each crystallizing species has reached the probability threshold at a given time. To build up the entire time dependent phase diagram, the probability functions for all species are calculated for the time span of interest, for each set of initial concentrations. Repeating the process illustrated in **Figure 4-3** leads to a collection of compositional points for each crystallizing species that depict the crossover boundaries at each time point. **Figure 4-4** shows four ternary phase diagrams for the same cocrystal system, each representing a different point in time, with the diagram labeled 1 representing the earliest time and 4 representing the latest time. $t=25$ s in the first frame and proceeds linearly to $t=70$ s in the fourth frame. The lines in each diagram are constructed by placing points onto the diagram as described in **Figure 4-3**. Thus, each line represents the concentrations at which the probability of a given type of crystal forming is equal to 50%. These lines border regions

in which all probabilities for that species are below the threshold value at that time, indicating a greater than 50% chance of finding that specific crystal at that time. The red and green lines represent the single component crystals of components 1 and 2, and the blue line represents the cocrystals.

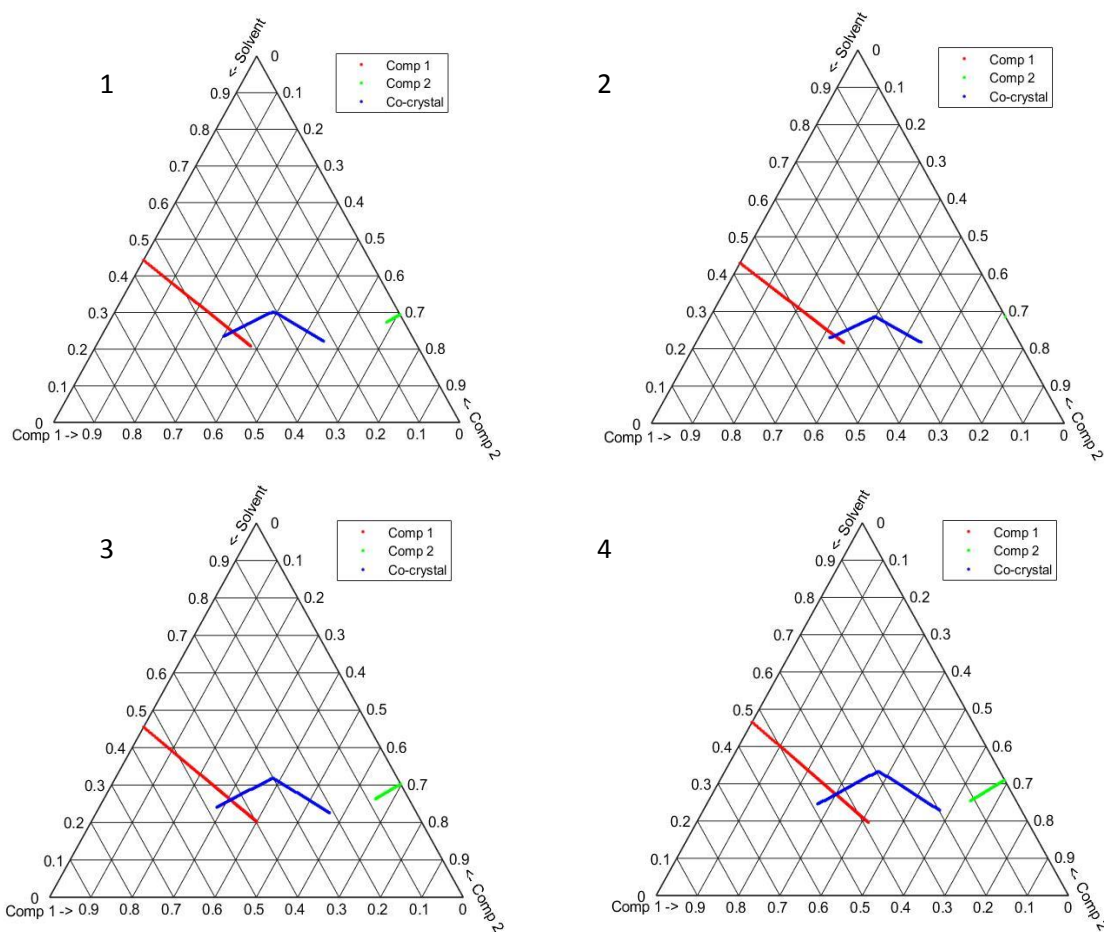


Figure 4-4 Cocrystal kinetic ternary phase diagrams progressing through time. Each of the four diagrams represent the same system at different points in time, progressing from 1 to 4. The time values corresponding to each panel are: $t_1 = 25$ s, $t_2 = 40$ s, $t_3 = 55$ s, $t_4 = 70$ s and the same parameters are used as given earlier in the chapter. Each colored line represents the boundary at which the probability for the corresponding crystal species is equal to the threshold value. As time progresses these boundaries move from regions of lower concentration to higher concentration.

At earlier times, the three boundaries are closest to the bottom edge of the phase diagram, and they shift toward the top apex as time passes. This shows that more concentrated compositions, further from the solvent apex, reach the threshold probability more quickly than less concentrated compositions.

Regions closer to the solvent rich corner tend to have a lower chance of nucleating any crystal, and the regions closer to the other two corners are most likely to form single component crystals. By extension, the cocrystal boundary originates from the point between the two coformer edges that maximizes the specific definition of coformer complex supersaturation prescribed in the model, given by equation (23), relative to the supersaturation of each of the pure components. This point is not always found along the bottom edge of the phase diagram, because in regions highly concentrated in both coformers the single crystal supersaturation will also be high. As time progresses each respective boundary moves from the region with the highest supersaturation for that crystallizing species toward the equilibrium concentration where the supersaturation equals unity. The rate at which the boundaries approach the equilibrium concentration lines is determined by the nucleation rate for that crystal.

The exact way the boundaries change with time provides insight into the system behavior, and illustrates how small changes in the initial concentrations may lead to differences in the final crystal mixture. As shown in **Figure 4-1**, in a typical equilibrium phase diagram there are regions in which both the cocrystal and the single component crystals are expected to exist. This is due to overlap of the regions within the solubility limit for each species. If a solid phase does not yet exist at a given concentration and time, the solution will be supersaturated and eventually a new crystal will nucleate to

reduce the free energy of the system. **Figure 4-5** shows a zoomed in view of one of these transient ternary phase diagrams showing the evolution of the boundaries for the same four time points considered in **Figure 4-4**, but focusing on the two points denoted by the black dots for illustrative purposes. These points represent two initial solution concentrations that are close together in composition, and both are supersaturated with respect to component 1 and the cocrystals.

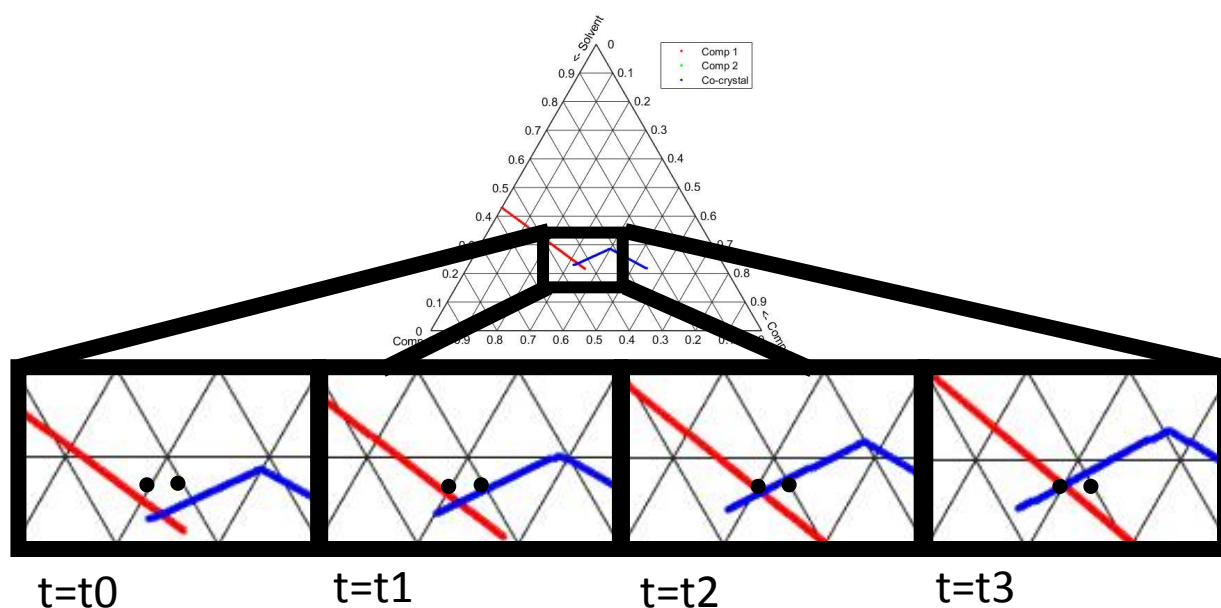


Figure 4-5 Kinetic history of crystal formation for similar initial compositions.

Two initial compositions that are close together on the overall phase diagram are considered, denoted by the black dots. The movement of the threshold boundaries for each crystal type relative to the two black dots is depicted at the same four time points considered in Figure 4-4. The probability threshold boundaries reach the two points at different times. At these conditions the left dot is intersected by the single component crystal boundary first, and then later by the cocrystal boundary. The reverse is true for the right dot. While both of these points ultimately end up in the mixed crystal region of the equilibrium phase diagram, their kinetic history difference will likely lead to differences in their final crystal compositions.

At time $t = t_0$, both points are outside of the regions bounded by the red and blue crystallization probability thresholds. As time passes the probability of forming a crystal increases for all concentrations and the boundaries shift toward the solvent rich region, eventually crossing over the two marked points, with the component 1 boundary reaching the left point and the cocrystal boundary reaching the right point at $t = t_1$. Progressing forward in time, the boundaries continue to shift, and the cocrystal boundary reaches the left point at $t = t_3$.

The difference in arrival times of the two boundaries at the two slightly different initial compositions shows that even compositions very close to one another that are both within the mixed crystal region of the phase diagram, as shown in **Figure 4-1**, can have very different kinetic histories. These differences in kinetic history could lead to significant changes in the final crystal mixture under the assumption that growth is rapid compared to nucleation.

The analysis illustrated in **Figure 4-4** and **Figure 4-5** offers a new way to define the composition boundaries within a ternary phase diagram, by denoting the nucleation pathway that is first to form a crystal at each concentration. An example is shown in **Figure 4-6**. Here we use the same axes as for the **Figure 4-1**, but now the shaded regions have a different interpretation. In **Figure 4-6**, all locations within the red region correspond to compositions at which the component 1 crystal probability reached the threshold value first. Similarly, the green region denotes the region in which the

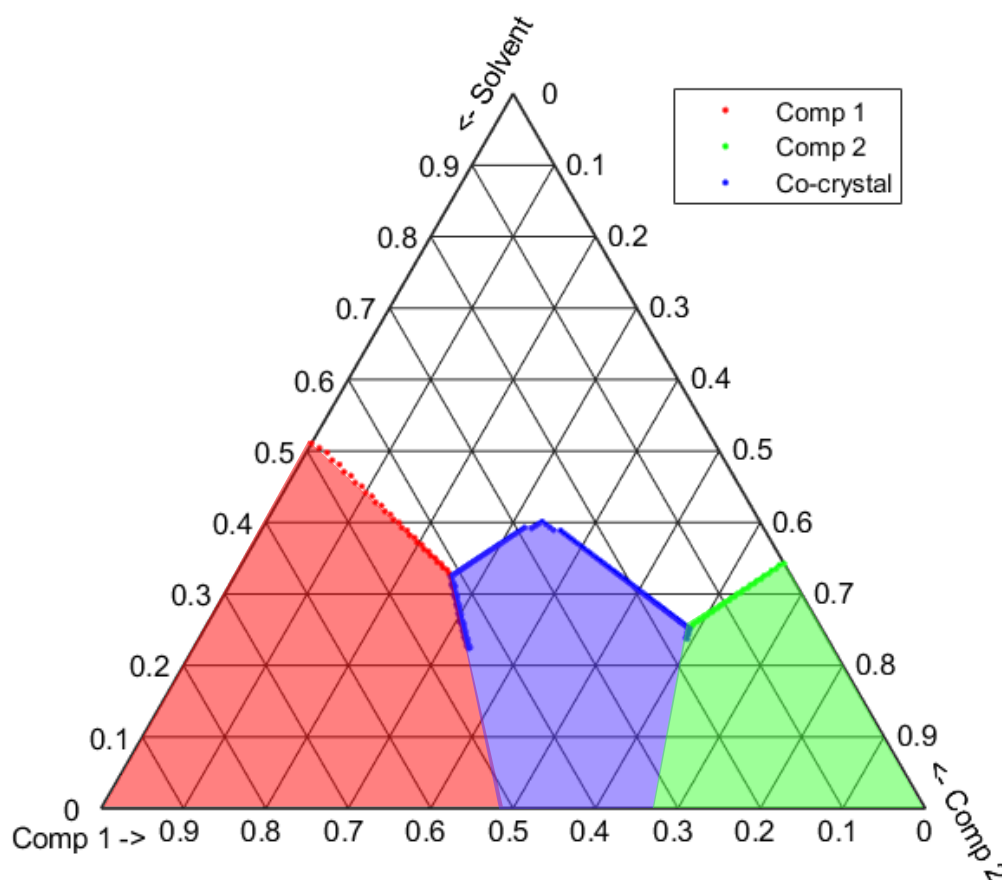


Figure 4-6 Phase Diagram depicting first nucleating species. Ternary phase diagram in which each location corresponds to a set of initial concentrations of solute and cofomers. The shaded regions indicate which crystallizing species is first to reach the threshold probability value at that concentration. The solution containing crystals will continue to evolve from this point until equilibrium is reached.

probability threshold for component 2 was exceeded first, and the blue region denotes the region in which the probability for cocrystal formation exceeded the threshold first. All regions outside of these three areas did not reach the threshold value within the timeframe of the simulation.

Assuming that growth is fast compared to nucleation and at concentrations for which both a single component crystal and a cocrystal are permitted according to thermodynamics, if a single component crystal nucleates first it will rapidly consume material from solution as it grows. This rapid growth will locally reduce the solution concentration for the given component, which also reduces the supersaturation for the corresponding co-crystal and reduces the likelihood of obtaining a cocrystal. It is important to note that the borders between regions in **Figure 4-6** do not imply that it is impossible to obtain mixed crystal phases, rather this figure illustrates the crystal forms that are most likely to be found at a given initial composition after homogenous nucleation from solution. The value of the probability threshold is arbitrary, and nucleation is a stochastic process, so local variations in a real system compared with this phase diagram is to be expected. Within the red region in **Figure 4-6**, which represents the single component crystal of component 1, it is still be expected that cocrystals will also form, since their nucleation probability is nonzero within the solubility limit.

In the event that growth is not rapid compared to nucleation, then the solution composition will gradually change from its initial state, with the component concentrations evolving toward their equilibrium concentrations as crystals begin to nucleate and grow within the solution. The interaction between a changing solution

composition and the nucleation rates is explored in the next section of this chapter using the ICCNT model to represent the nucleation process.

Modeling the role of concentration change

In real systems with finite volumes the first crystal that nucleates will not instantaneously consume all of the solute molecules in a solution. In this section we will explore how allowing for concentration changes, combined with assumed nucleation kinetics based on the ICCNT model, can produce paths through composition space that evolve from supersaturated toward equilibrium. We first examine a case in which growth is assumed to be fast enough to instantly change the solution composition from its supersaturated value to its equilibrium value as soon as a crystal nucleates. This assumption will provide a basis for illustrating how composition within the solution changes when we later explore the more realistic case in which growth is not instantaneous.

Figure 4-7 plots the concentrations of the cofomers on each axis. The horizontal dashed red line represents the saturation concentration of component 1, and the vertical dashed green line represents the saturation concentration of component 2. The curved blue dashed line represents the saturation concentration for the cocrystal according to equation (23). From an initial set of supersaturated concentrations marked by the black point, the system will attempt to minimize its free energy by nucleating a new solid phase until sufficient material has been removed from the solution and the saturation concentration is attained. Starting from the initial concentration marker, the solid arrows indicate the shortest path to the saturation concentration for each species. The diagonal dashed blue lines bound the region in which there is a straight-line path to the cocrystal

equilibrium concentration. The slopes of these lines are determined by the stoichiometry of the cocrystal, which is 1:1 in this example. All initial concentrations outside of this region will result in at least some single component crystals. This is because the solution will still be supersaturated with respect to a single component crystal, even once it attains the cocrystal equilibrium concentration. Eventually these systems will nucleate a single component crystal to further reduce the free energy of the system.

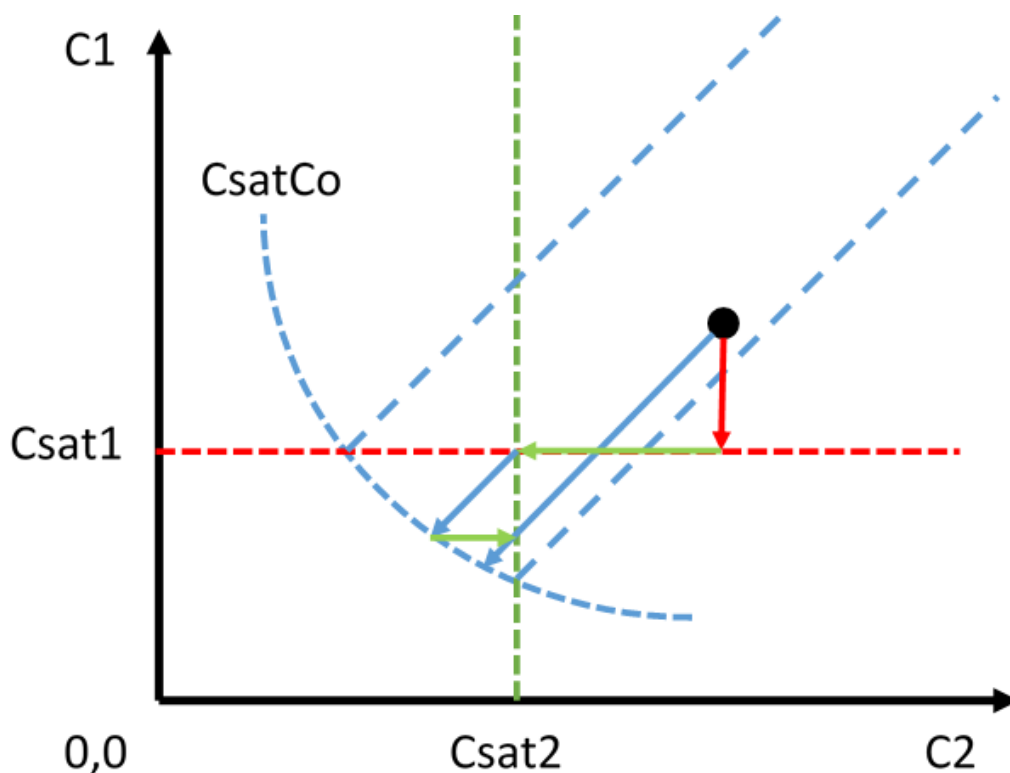


Figure 4-7 Saturation concentration boundaries. With the concentrations of coformer 1 on the y-axis and coformer 2 on the x-axis, each point within the graph represents a unique initial solution concentration pair. The dashed red and green lines represent the saturation concentrations for the respective components. The curved blue dashed line is the saturation concentration for the cocrystal. At the intersection of the cocrystal and single component saturation concentrations, dashed blue lines are drawn with slope corresponding to the stoichiometry to define the region in which equilibrium results in only cocrystals. From an initial concentration marked by the black point, several paths to equilibrium by nucleation are shown by the colored arrows.

Since multicomponent systems that can nucleate cocrystals have a minimum of three saturation concentrations, one for each single component crystal and one for the cocrystal, the path toward phase equilibrium is not obvious. The solid arrows in **Figure 4-7** show two of the many distinct paths that the solution could take from its initial composition to the equilibrium concentration by forming crystals. With many available nucleation pathways, many possible intermediate crystal populations can exist before the solution reaches its final equilibrium state. The compositions of the intermediate crystal populations are determined by the nucleation kinetics of the components in the system. At macroscopic scales it is unlikely that the entire solution will instantaneous change directly from an initial state to phase equilibrium due to transport limitations. This introduces a time dependence for the products of cocrystal nucleating system, where if products are collected before thermodynamic equilibrium is established, a mixture of crystals is almost certain to be observed. Understanding nucleation kinetics in cocrystal systems therefore allows for the possibility to access new or difficult-to-generate products. Starting from any pair of initial concentrations for the coformers, the path through concentration space will be directed by the nucleation probability for that set of concentrations at that time.

To determine which nucleation pathway is the most likely to occur, we can consider probability ratios given by

$$Pr_1 = \frac{P_{01}}{P_{0net}} \quad (43)$$

$$Pr_2 = \frac{P_{02}}{P_{0net}} \quad (44)$$

$$Pr_{co} = \frac{P_{0\ co}}{P_{0\ net}} \quad (45)$$

These three ratios must be determined by simulating the full model described in Chapter 3. At each composition and each moment in time, the three probability ratios will take on unique values that give the relative probability that a particular crystal will nucleate normalized by the total probability. **Figure 4-8** depicts the relative probability

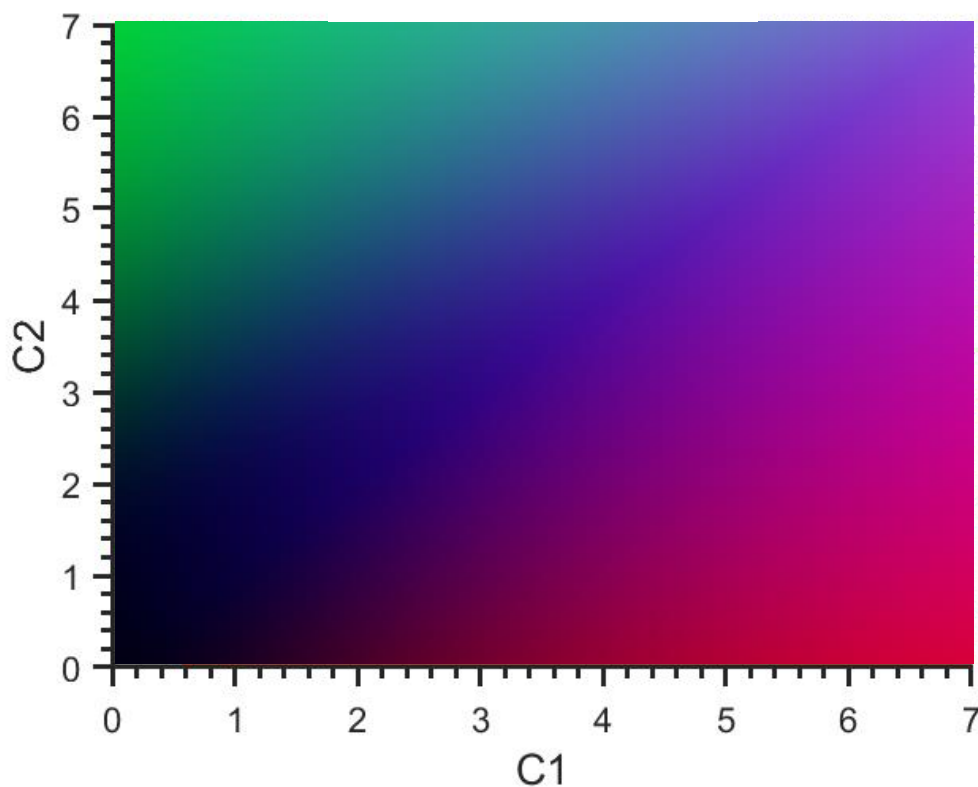


Figure 4-8 Nucleation Probability Surface. Each x-y pair represents an initial solution concentration. The color of the point at each position represents the relative likelihood of possible nucleation pathways for formation of component 1 crystals (red), component 2 crystals (green), and cocrystals (blue). The relative intensity indicates the magnitude of the net probability at that time. The probability at each point evolves with time based on the ICCNT cocrystal nucleation theory model described in Chapter 3.

ratios, where the concentration of one coformer is represented on the y-axis and the other coformer concentration is represented on the x-axis. The plot is a color plot, where the hue of each point is determined by the value of Pr_1 (red), Pr_2 (green), and Pr_{co} (blue), and the intensity is determined by the magnitude of $P_{0,net}$. The color plot can be thought of as a three-dimensional surface where the solution will descend along the gradient of the surface toward phase equilibrium. The shape of the gradient is determined by the Pr values at each instant in time which determine the trajectory along which a solution will move from its initial concentrations toward equilibrium. In the figure, high values of Pr_1 are red, Pr_2 are green, and Pr_{co} is blue. Near the origin, the black region indicates a low probability of forming any crystal at all. Determining the trajectory through the phase space is a complex and computationally intensive problem that would require many computationally expensive simulations to model accurately. To first gain a general understanding of the process let us imagine a simplified case where diffusion is infinitely fast and we neglect growth while still allowing the concentration to change according to the kinetics of complex formation and nucleation.

To account for concentration changes in this manner, we will assume that when a crystal nucleates, it first consumes the material within a critical volume around the nucleus. The dimensions of this volume are determined by physical parameters of the system related to the critical radius of the nucleus r^* , given by

$$r^* = \frac{2\gamma}{\rho_c \Delta\mu} \quad (46)$$

Where γ is the interfacial tension between the crystal and the bulk solution, ρ_c is the number density of the crystal, and $\Delta\mu$ is the chemical potential difference between a

molecule in the solution and a molecule in the crystal. Assuming the critical nucleus can be approximated by a sphere, the number of molecules with in it is defined as:

$$n^* = \rho_c * V_c = \rho_c * \frac{4}{3}\pi r^{*3} = \frac{32\pi}{3\rho_c^2} \left(\frac{\gamma}{\Delta\mu}\right)^3 \quad (47)$$

where V_c is the volume of the critical nucleus. For a critical nucleus to remain intact and not dissolve back into solution, the total number of molecules remaining in solution directly after nucleus formation must be at least equal to the saturation concentration, therefore the total number of molecules n within the critical volume must be greater than or equal to :

$$n = n^* + C_{sat} * V_c \quad (48)$$

The total number of molecules within the critical volume can also be defined by the solution concentration multiplied by the critical volume.

$$C_i * V_c = n \quad (49)$$

Combining these equations allows us to solve for the critical volume corresponding to a given initial concentration, given by

$$V_c = \frac{n^*}{C_i - C_{sat}} \quad (50)$$

Using this critical volume we can begin to investigate how the solution would evolve over time. Consider that when a crystal nucleates, the solution concentration only changes by an amount proportional to n^* , as one critical nucleus worth of material is removed from solution for each nucleation event. This is a simplification and future work could expand upon it by incorporating transport and growth effects. The path a solution

takes from its initial composition toward equilibrium can be seen in **Figure 4-9**. Here, the axes are the same as in the previous figure and each line represents a solution starting at a supersaturated concentration near the top and right edges. As the simulation progressed, the trajectory through the space is determined by the gradient of the surface depicted in **Figure 4-7**, and the solution composition proceeds toward the origin as crystals form and the concentration changes. For this simulation a population of critical volumes all start at an initial supersaturated concentration located near the outer edge of the graph. The system is allowed to progress toward equilibrium by advancing through time and allowing each individual critical volume to randomly form a crystal in accordance with the probability distribution defined by the nucleation model. At each time step it is determined whether or not a volume forms a crystal based on the crystal probability corresponding to that concentration at that time. If a crystal forms, it removes a finite amount of material from solution to reduce the concentration within the volume to a value equal to the saturation concentration. Here the assumption of infinitely fast diffusion implies that the removal of material is distributed across the entire volume, resulting in a change to the solution concentration according to the average of all its constituent volumes. At the next time step, all volumes are examined again for nucleation at the new concentration, and the process repeats until the time span of interest for the simulation is reached.

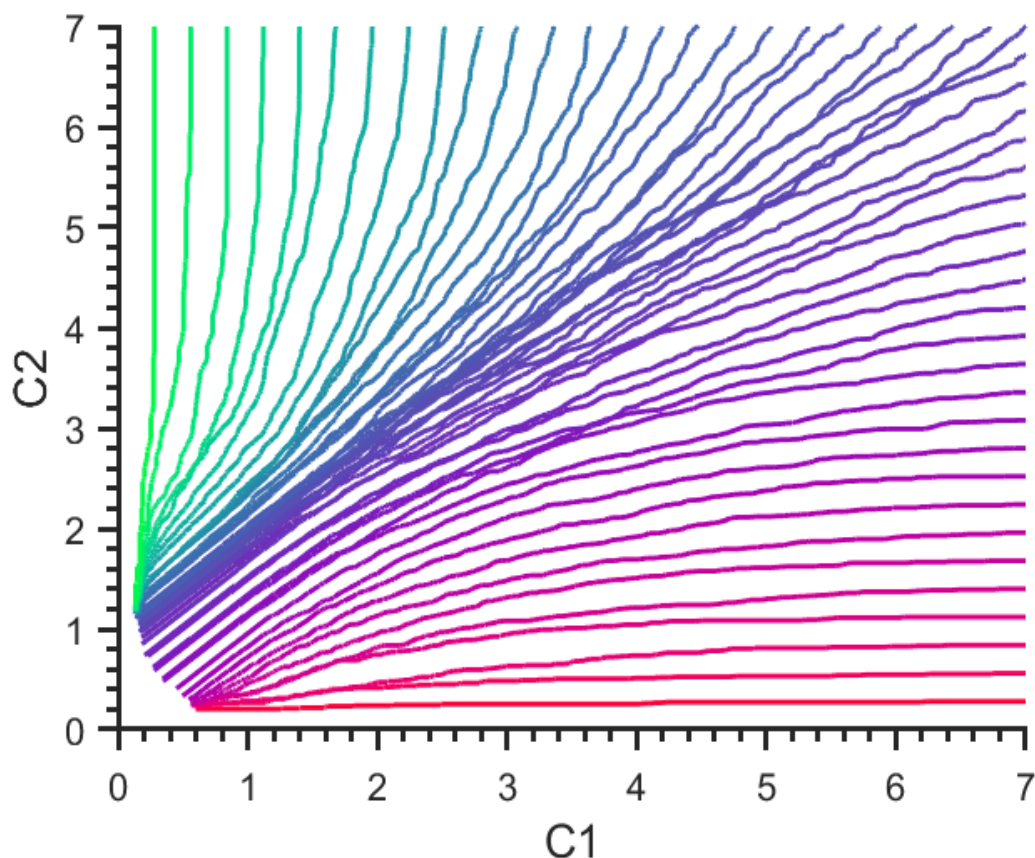


Figure 4-9 Path to equilibrium by nucleation. The concentrations of the two coformers are plotted along the x and y axes. Each line represents a population of volumes within a solution that start at a high concentration and then reduce the concentration by nucleating solid crystals. From each initial condition the population approaches equilibrium by a different path depending on the nucleation probabilities defined for the instantaneous concentration, depicted in **Figure 4-8**. The color of the line represents the relative proportion of component 1 crystals (red), component 2 crystals (green), and cocrystals (blue).

Let us now compare the results from the nucleation kinetic modeling to predictions assuming thermal equilibrium. **Figure 4-10** shows an equilibrium phase diagram overlaid with results from the kinetic simulations. The x and y axes represent the concentrations of the two coformers and the dashed green and red lines show the saturation concentrations for the two single component crystals. The curved blue line shows the saturation concentration for the cocrystal. Each region is labeled with the

crystal composition predicted at thermal equilibrium. Around the edge of the diagram the colored dashes represent the 50 initial concentrations simulated in **Figure 4-9**. The color of each dash represents the proportion of each crystal species that formed during that simulation as a result of the kinetics of the process.

The figure indicates that there are significant differences between the predicted equilibrium crystal compositions and the final compositions resulting from the kinetic rate processes. For example, the sharp transitions depicted between the equilibrium regions of the diagram are much more gradual for the system resulting from the kinetic processes. In particular it should be noted that even within the region that should only contain cocrystals at equilibrium, the kinetic simulations show that at these concentrations single component crystals still nucleated as the system approached equilibrium. This can be seen in the dashes within the region bounded by the two diagonal blue dashed lines. The points simulated using the kinetic model are not purely blue, but instead contain some red and green in their hues. This result suggests that if one is producing cocrystals from solution, and the nucleation process stops before thermal equilibrium is reached, the collected products may differ significantly from the predictions made by equilibrium phase diagrams.

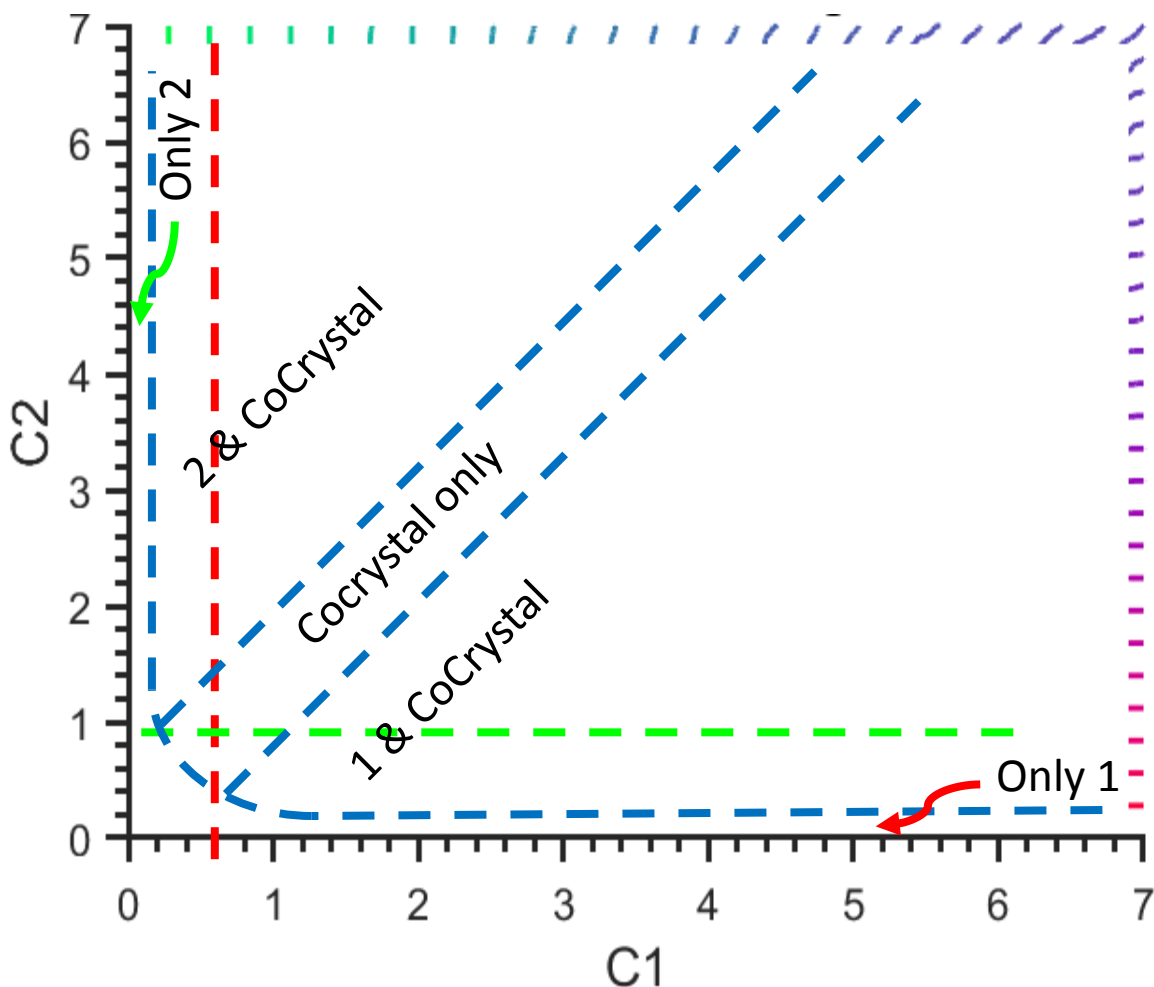


Figure 4-10 Comparing results of nucleation kinetic simulations to the equilibrium phase diagram. Phase diagram plotted with coformer concentrations represented along the x- and y-axes, assuming saturation concentrations of $C_{sat1}=0.9$ mg/ml and $C_{sat2}=0.3$ mg/ml, and cocrystal saturation defined by equation (23) with $K_{sp}=0.05$. The dashed red and green lines represent the saturation concentration for each of the two coformers. The curved blue line represents the saturation concentration for the cocrystal. The diagonal dashed blue lines bound the region that at equilibrium is predicted to contain only cocrystals. All other regions bounded by saturation concentration lines are labeled according to the crystal mixture that is expected at equilibrium. Around the right and top edges of the plot the colored dashes indicate the results of the kinetic simulations. The position of each dash represents its initial concentrations and its color represents the proportion of each crystal species formed. Red represents component 1 crystals, green represents component 2 crystals and blue represents cocrystals.

So far, the simulations presented here have assumed a finite volume of the crystallizing solution. In considering smaller volumes surrounding the critical nucleus, we must also consider the limit at which the system begin to act as a continuum rather than discrete finite elements. At low numbers of discrete volumes the variability inherent in the random process makes it difficult to generate accurate predictions from the model, whereas at larger solution volumes, simulating a large number of discrete

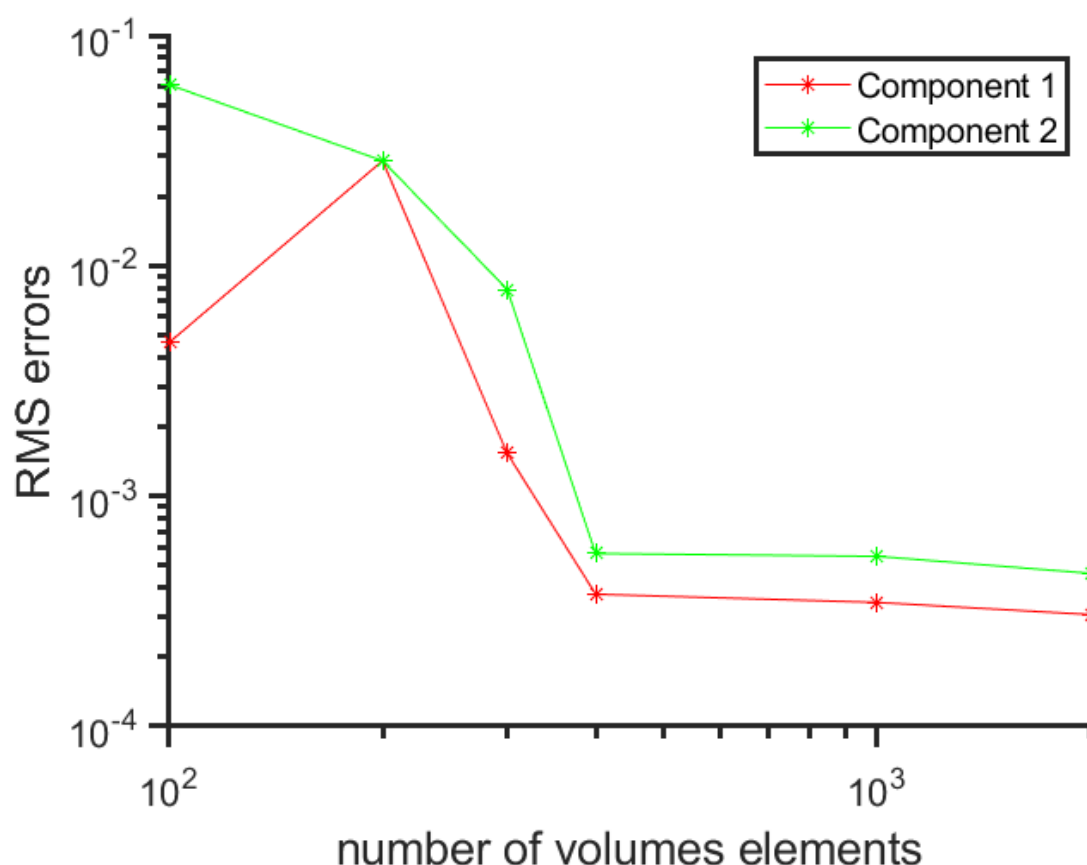


Figure 4-11 Convergence to continuum behavior. For an initial set of concentrations, the population of crystals was simulated as it evolves to equilibrium via nucleation. 20 trial simulations were conducted, and then compared against the average of all runs. The root-mean-squared error between the average path and each individual path was calculated for a given number of volume elements and plotted above. Each line represents one of the cofomers. Between 400 and 1000 volumes, the deviations from the mean reached a minimum.

volumes is computationally expensive. As the number of volume elements increases, the deviation from the mean of all runs decreases. **Figure 4-11** plots the root-mean-squared (RMS) error on the y-axis versus the number of volume elements considered on the x-axis. Since time is a parameter in the previous graphs, we measure the deviation from the expected concentration with respect to each component, noted by the red and green lines on the plot. From this analysis we see that between 400 and 1000 volume elements, the variation in the simulation results approaches a constant

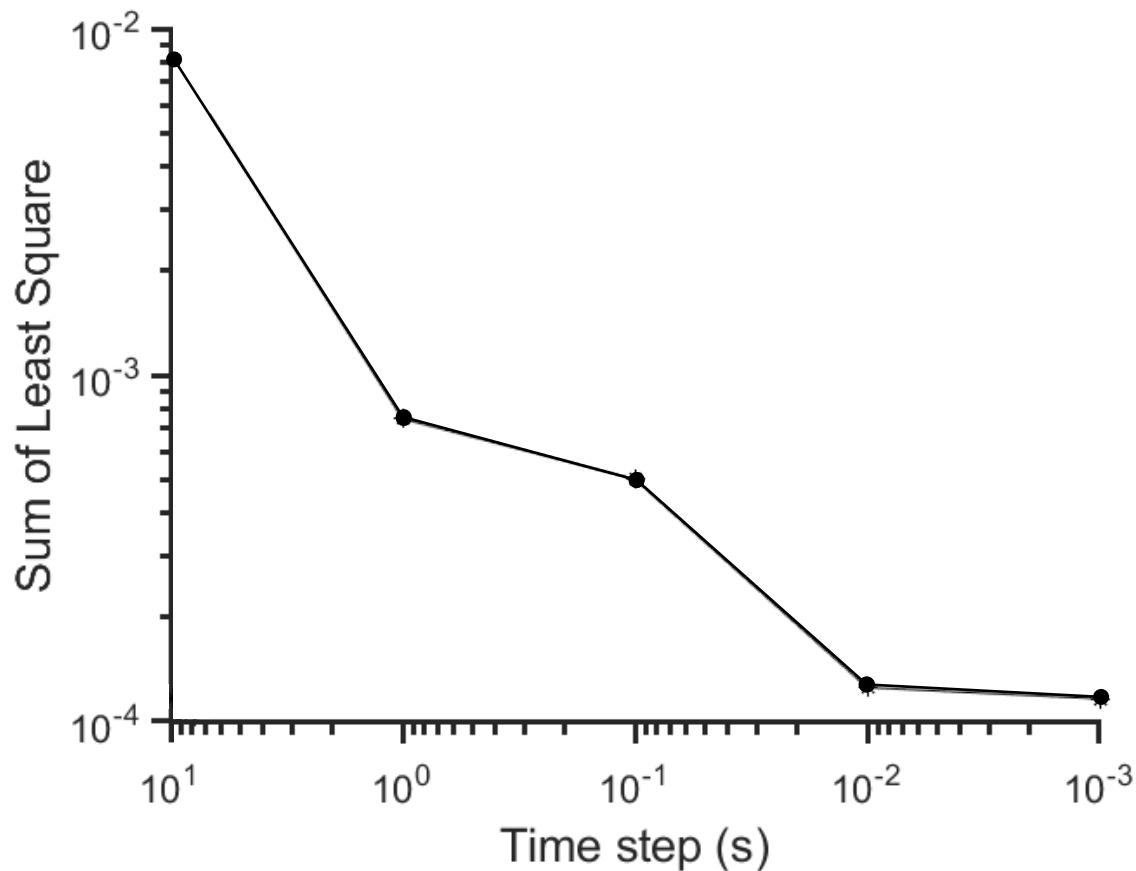


Figure 4-12 Convergence with decreasing time step. The RMS error is plotted on the y-axis versus the time step along the x-axis, decreasing from left to right. As the time step becomes smaller, the difference between the numerical and analytical solution to P_{net} decreases, reaching a minimum at approximately 0.01s.

minimum value. As a result, we use at least 1000 volume elements for all of the simulations considered here.

It is also important to check for numerical stability of the transient simulation, by considering the dependence of the result on the selected time step for the numerical integration steps. To ensure stability, the RMS error between the analytical solution for $P_{0\ net}$ and the numerical solution was calculated at various time step values. As seen in **Figure 4-12**, the difference between the numerical solution and the analytical solution reached an RMS error we considered within tolerance at approximately 0.01s. This

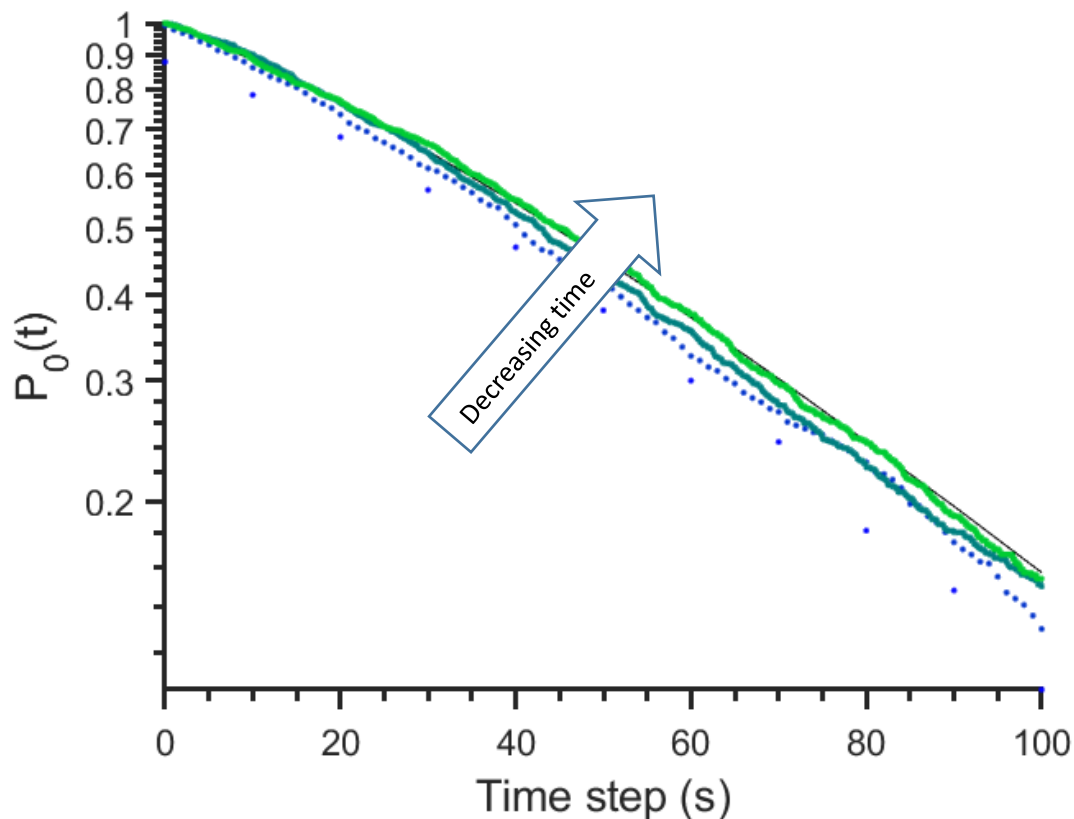


Figure 4-13 Convergence of numerically computed P_{net} . Plot of probability on the y axis versus time on the x axis, for several simulations at identical conditions but with different timesteps used in the numerical integration steps. As the time step gets smaller, the difference between the numerical and analytical solution for P_{net} gets smaller as well.

minimization of the RMS error can also be seen in **Figure 4-13**, which depicts the probability functions produced by the numerical integration corresponding to each time step. The timesteps selected range from 10 s to 0.01 s. As the time step decreases the numerical results converge toward the analytical solution. This result gives us confidence that in cases where analytical solutions are not available, numerical solutions can provide accurate predictions.

Conclusions

The multiple nucleation pathways that are available in co-crystal systems require the investigation of kinetics in addition to thermodynamic equilibrium to accurately predict the outcomes for a given system. Each crystal type that can be generated has a different characteristic timescale for formation at a given initial solution composition. In this chapter we analyzed the kinetics by considering the timescales required to reach equal probabilities of nucleating or not nucleating a given type of crystal, either single crystals of the coformer components or cocrystals. In cases where there is significant lag time between the time at which a single component crystal is likely and the time at which a cocrystal is likely, equilibrium phase diagrams may not provide the full story. We have shown that stepping through time in a kinetic ternary phase diagram can be useful for identifying composition spaces in which the products formed may differ from the expected results at equilibrium. Due to the core assumption that growth is rapid compared to nucleation, once a crystal of a certain kind nucleates from solution it is likely to rapidly consume the remaining dissolved material until the concentration is reduced enough for phase equilibrium to be restored. This consumption of material will reduce the likelihood that a different kind of crystal will nucleate if its rate depends on the supersaturation of the consumed material. At certain conditions this can lead to

entirely different final crystal compositions with only slight variations in the initial solution composition. Therefore, the information provided by nucleation kinetics is invaluable for designing systems that can robustly produce co-crystals at scale.

References

- [1] Ainouz, A. "Modeling and prediction of cocrystal phase diagrams". / International Journal of Pharmaceutics 374 (2009) 82–89
- [2] Rodriguez-Hornedo, Nair. "Reaction Crystallization of Pharmaceutical Molecular Complexes." Molecular pharmaceutics. 3.3 (2006): 362–367. Web.
- [3] Dario Braga. "Reactions Between or Within Molecular Crystals". Angew. Chem. Int. Ed. 2004, 43, 4002 –4011
- [4] Lee, Tu. "Continuous Co-Crystallization As a Separation Technology: The Study of 1:2 Co-Crystals of Phenazine-Vanillin." Crystal growth & design. 12.12 (2012): 5897–5907. Web.
- [5] Yusuke Maeno, "Novel pharmaceutical cocrystal consisting of paracetamol and trimethylglycine, a new promising cocrystal former". International Journal of Pharmaceutics 473 (2014) 179–186
- [6] Shayanfar, Ali. "Physicochemical Characterization of a New Cocrystal of Ketoconazole." Powder technology 262 (2014): 242–248. Web.
- [7] Anderson, Stephen R. "Preparation of an Energetic-Energetic Cocrystal Using Resonant Acoustic Mixing." Propellants, explosives, pyrotechnics 39.5 (2014): 637–640. Web.
- [8] Anderson, Stephen R. "Preparation of an Energetic-Energetic Cocrystal Using Resonant Acoustic Mixing." Propellants, explosives, pyrotechnics 39.5 (2014): 637–640. Web.
- [9] Peng-yuan Chen. "Theoretical study of BTF/TNA cocrystal: Effects of hydrostatic pressure and temperature". Defence Technology 11 (2015) 132-139
- [10] Barbas, Rafael. "Sildenafil-Resorcinol Cocrystal: XRPD Structure and DFT Calculations." Crystals. 10.12 (2020): n. pag. Web.
- [11] Yang, Caiqin. "Experimental and DFT Simulation Study of a Novel Felodipine Cocrystal: Characterization, Dissolving Properties and Thermal Decomposition Kinetics." Journal of pharmaceutical and biomedical analysis. 154 (2018): 198–206. Web.
- [12] Kudo, Shoji. "Production Method of Carbamazepine/saccharin Cocrystal Particles by Using Two Solution Mixing Based on the Ternary Phase Diagram." Journal of crystal growth 392 (2014): 87–91. Web.
- [13] Jia, Qian. "Investigation of the Phase Behavior of a HNIW Center Dot TNT Cocrystal System and Construction of Ternary Phase Diagrams." Crystal growth & design. 19.11 (2019): 6370–6376. Web.
- [14] Jia, Qian. "Investigation of the Solid-Liquid Ternary Phase Diagrams of 2HNIW Center Dot HMX Cocrystal." RSC advances 11.16 (2021): 9542–9549. Web.

[15] Hong, Chao. "A Novel Strategy for Pharmaceutical Cocrystal Generation Without Knowledge of Stoichiometric Ratio: Myricetin Cocrystals and a Ternary Phase Diagram." *Pharmaceutical research*. 32.1 (2015): 47–60. Web.

[16] Croker, Denise M. "Understanding the p-Toluenesulfonamide/Triphenylphosphine Oxide Crystal Chemistry: A New 1:1 Cocrystal and Ternary Phase Diagram." *Crystal growth & design*. 12.2 (2012): 869–875. Web.

Chapter 5 : Summary

Conclusions

In this Thesis I have shown that microfluidics offers a powerful tool for investigating the kinetics of homogenous nucleation of crystals from solution. I have also proposed a realistic model to describe the process of cocrystal nucleation, accounting for nucleation kinetics of single component coformer crystals and cocrystals. Further development and refinement of continuous drop train microfluidic nucleation devices will allow for the investigation of complex crystal forming solutions. The ability to measure nucleation rates accurately is a key tool in understanding the formation of a desired crystalline material. Without control over the nucleation process, resulting crystal populations are unlikely to form in the desired polydispersity and may potentially form undesired or unstable polymorphs. In the case of active pharmaceutical ingredients, this lack of control can lead to the need for additional costly downstream processing to filter or refine the results of the upstream processes. For single component crystal formation, theories like Classical Nucleation Theory can be used to make predictions relevant to the control of these processes, but experimental data is still lacking and is needed to validate these predictions for many systems of interest. When material costs are high, large scale experimentation is not feasible for determining the information required to control the process. These limitations can stifle the development of new materials and APIs. The small length scales inherent to microfluidic systems allow for exploratory research to be performed with a fraction of the material consumption of bench top equipment. Tight control over drop volumes is required to reduce the uncertainty in measurements of probability distributions observed within microfluidic devices.

An additional barrier to the production of cocrystal materials and pharmaceuticals is the lack of a widely accepted model for cocrystal nucleation. Chapter 3 proposes two realistic models for cocrystal nucleation, focusing mainly on the assumption that the formation of coformer complexes precedes nucleation. The presence or absence of an initial lag time for the onset of cocrystal nucleation in experimental studies will help validate this assumption and the corresponding model. Despite the lack of model validation to date, predictions using this model are useful to show the expected effect of nucleation kinetics and complexation kinetics on the final crystal mixture based on the competition of rates within the system. The insights gained from analyzing a single composition in Chapter 3 are expanded to consider the entire phase space and to chart the kinetic history of numerous possible solution compositions in Chapter 4. This allows for a fuller understanding of the process and can inform experimental design considerations. For example, while equilibrium phase diagrams can accurately predict the crystals that are permitted to exist at certain concentrations, they fail to account for the kinetics of nucleation, which will determine the proportion of each type of crystal present at intermediate times. Additionally, if a system does not fully equilibrate and products are collected at an intermediate time in the process, the population of crystals in solution may be entirely different from the predictions made by equilibrium phase diagrams.

Future work

Many avenues for future work exist based on what has been demonstrated in this thesis to date. The results presented in Chapter 2 show that the accuracy of nucleation measurements in droplet-based microfluidic devices can be improved. Reducing the polydispersity of the drop size distribution should be the primary goal for improving the

accuracy of the statistics gathered from both the static arrays and the continuous drop train device. This type of improvement will likely include additional time spent modeling and testing new device designs to balance the considerations of drop size distribution, residence time, and total material consumption. Additionally, while effective at maintaining the target temperatures within the device, the water jacket design is cumbersome for operation and requires long setup times to reach a stable temperature. Redesigning the temperature control system with a thermoelectric cooler would greatly increase the rate at which experiments could be conducted on the platform and would likely provide greater control over the temperature. To expand the system capabilities it may be beneficial to fabricate the microfluidic device in a material other than PDMS. While PDMS is an excellent material for rapid prototyping, due to its low cost and ease of manufacturing, there are some key disadvantages, namely, a polymer matrix of PDMS swells in the presence of many common organic solvents. Swelling occurs due to the solvent entering the interstitial spaces of the matrix, which has the additional effect of changing the contact angle of the solvent on a PDMS surface. Altering the contact angle can severely affect the wettability of the channel walls and therefore the ability to generate drops with the device. Since many known co-crystal systems utilize organic solvents, it is likely that alternative, more solvent stable materials will need to be considered in order to test cocrystal nucleation in microfluidic devices. Common substitutes for PDMS include fused silica and quartz glass, both of which do not swell in the presence of organic solvents, but may need surface treatments to properly adjust the wettability for drop formation. Additionally, hard materials such as these require much longer times and greater expense to fabricate and therefore prototyping designs

is not as rapid as it is for PDMS. Therefore fabrication in alternative materials should only be implemented once the channel design is finalized.

Chapters 3 and 4 deal with the description and prediction of a cocrystal nucleation model which still needs to be experimentally validated. Selecting a compatible cocrystal system and measuring its nucleation rates and the dependence of those rates on conditions including the initial concentrations would provide an excellent starting point for validation of the model predictions. Experiments comparing the final crystal mixture for homogenous nucleation from solution with equilibrium data from literature would be particularly helpful. The concentration range to be considered could be narrowed after determining the nucleation rate constants, assuming experimental results agree well with the model.

Additional computational work could also be conducted to expand upon the work presented in Chapters 3 and 4. Incorporating insights from tools such as molecular dynamics may help quantify parameters used in these models, such as the formation rates of coformer complexes used in the Initial Complexation Cocrystal Nucleation Theory model. Additionally the effect of transport and growth were mostly neglected in the analysis presented in Chapter 4. However, this is a very important aspect of crystal engineering and will result in pronounced effects in real world systems. Additional model refinements to account for the finite rates of diffusion and crystal growth could be incorporated to estimate the local concentration in solution as a function of time. A model including both nucleation and growth would be a useful extension of the models proposed here.