# Effects of Solution Complexation on Crystallization Processes

by

Carlos A. Pons-Siepermann

Bachelor of Chemical Engineering

University of Michigan – Ann Arbor, 2013

M.S. Chemical Engineering Practice

Massachusetts Institute of Technology, 2016

Submitted to the Department of Chemical Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

~~May 2018~~ [June 2018]

## Signature redacted

Signature of Author: _____

Department of Chemical Engineering

May 4, 2018

## Signature redacted

Certified by: _____

Allan S. Myerson

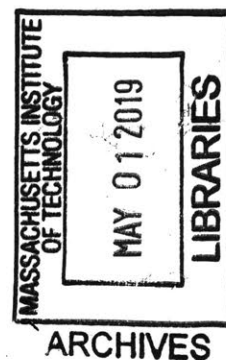Professor of the Practice of Chemical Engineering

Thesis Supervisor

## Signature redacted

Accepted by: _____

Patrick S. Doyle

Robert T. Haslam Professor of Chemical Engineering

Chairman, Committee for Graduate Students

# Effects of Solution Complexation on Crystallization Processes

By

Carlos A. Pons Siepermann

Submitted to the Chemical Engineering Department on May 4, 2018, in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

## ABSTRACT

Crystallization is a separation technique widely used in chemical processes to produce high-purity solid products. The impact of solution chemistry on the kinetics and thermodynamics of crystallization processes is neither well understood nor properly characterized. Therefore, there exists a need for research to develop chemistry that can exploit the effect of impurities, additives and foreign molecules on the chemistry within crystallizing solutions. The use of rational chemical interactions has the potential of enhancing the controllability of crystallization unit operations, providing a new process handle for chemical engineers with which they can create new crystal forms, enhance product purity, improve yields, or inhibit the formation of undesirable crystals.

This thesis focuses on the use of small-molecule chemical additives that exhibit selective intermolecular interactions with crystallizing solutes or their impurities. Within the work reported, there were two major areas of study: purification and nucleation control. Additive-driven solution complexation with impurities was demonstrated to be a powerful tool for enhancing the purity of crystal product, without penalizing the process yield. The technique was implemented for the separation of structural isomers, and tested for the purification of a large pharmaceutical compound with challenging chemical features. The results herein helped elucidate the capabilities of complex-assisted crystallization, and also outline the thermodynamic and chemical limitations of the technique. The second half of the work explored the impact on nucleation rates of dilute impurities that interact with the supersaturated crystallizing solute. For the first time, impurity-driven nucleation inhibition was systematically and quantitatively proven, using high-throughput induction measurements. The experimental results were used to discern the thermodynamic and kinetic impact of the inhibitor, and to elucidate a potential underlying mechanism for the observed behavior. Data demonstrated that even a weakly-interacting dilute additive can lead to massive nucleation rate depression through a kinetic pathway, most-likely due to the disruption of the ordering of the solute molecules within high-concentration clusters during nucleation.

Thesis Supervisor: Allan S. Myerson

Title: Professor of the Practice

## ACKNOWLEDGEMENTS

In a similar train of thought, just as financial resources are necessary, the wisdom to implement them productively and the knowledge to generate true value from what we have are paramount when trying to create meaningful work. Without them, we would go nowhere. At least by myself, I certainly wouldn't have. Which is why I cannot ever thank my research advisor, **Allan S. Myerson**, enough. From the very first time we met, he showed me kindness and patience, the willingness to teach me and let me to go out there (to the lab) and try things for myself. Speaking to Allan was always a breath of fresh air, since I don't think I have ever met a more direct and no-nonsense person in my life. I always joked with my lab peers about the fact that my meetings with him never went past 10 minutes, but the reason for that was always that he would help me get to the core problem of things immediately and devise a plan on how to address it on the spot. When I did need to meet with him for long periods of time, it was because I was stuck, or in trouble. In those instances, he was willing to sit down with me and help me move forward. I will always remember the end of my second year of grad-school. I had just finished the outline for a manuscript and was almost ready to submit; I thought I was going to have my first paper and was excited. Then, as I was finishing up the last data sets, I realized that there was a substantial mistake in how I had designed my control groups, and it invalidated a massive portion of my work, setting me back almost a year. I remember writing to Allan to request a meeting and discuss the issue, feeling crushed and afraid, embarrassed by my carelessness. When he sat down with me, he looked at me and said: "It's OK, you are a student, you are supposed to make mistakes..." Allan could have been upset at me for wasting his time, for my errors, but he supported me. Without that, I don't think I would have been able to go forward with my PhD work. It has been a privilege to work for someone as experienced and knowledgeable in their field. Sometimes I hear other people talk about their relationships and dynamics with their advisors, and I feel guilty for how lucky I was.

From the previous paragraph it should be quite apparent that I deeply admire Prof. Myerson. However, a research group is more than just its advisor. I remember when I was only a prospective student, I actually toured the Myerson lab and was in awe with the quality of equipment and people that were showing me around. It is a little strange, but the first thing I noticed is how clean and organized everything was, followed by the depth of resources available. Once I joined, I was greeted and treated as a colleague and friend by the other members, and I was always able to ask for help when I needed it. I became a proud member of the **Myerson group**, and being a part of it has been an honor and a wonderful experience. I need to particularly acknowledge **Dr. Cameron Weber**, the post-doctorate who tutored me and counseled me during my first semester. Cameron

was infinitely patient with me, and was incredibly knowledgeable in both chemistry and crystallization. His guidance and the groundwork he helped me develop were crucial for my development as a researcher. Another member that I would like to particularly acknowledge is my undergraduate researcher, **Sam Huang**. Sam worked with me for a little over a year, and in that short span, he and I were able to publish two peer reviewed articles together. Sam is one of the brightest and most industrious young men I have had the privilege of working with. His hard work was vital for the advancement of my thesis work, and his friendship was an equally important gift that I will forever be grateful for. Finally, I need to make a special note to my officemate and class peer, **Leia Dwyer**. Since our very first day in the group, Leia and I have sat on adjacent desks in the same office. It takes someone special to be able to withstand four years at such close proximity to me and my hubris. Leia was more than a peer, she was my advisor and confidant. She was always willing to discuss science and research and proof-read my slides and manuscripts, but more importantly she was there to discuss life, sports, vacations, and the future. I could have not asked for a better office companion and friend.

I would also like to thank and acknowledge my **thesis committee members**, Professors **Bernhardt Trout, Yuriy Roman,** and **Tim Jamison**. They have lent me their time and intellect, and guided me during these years, keeping my work both interesting and focused. I am particularly grateful for how flexible my committee members where when it came to scheduling meetings and thesis-related presentations. I count myself within a select and incredibly lucky group of people that never had to truly fret or suffer over the difficulty of scheduling my thesis work. More importantly however, the recommendations and guidance provided by the committee allowed me to stay on schedule and facilitated the development of my PhD work into a coherent thesis.

One of the most rewarding experiences I have had, not only during my PhD, but throughout my life, was the **Chemical Engineering Practice School Program**. During this program, I felt like I grew and earned invaluable experience, which has helped me ever since. In particular, I want to recognize the amazing work from my station directors **Robert Hanlon** and **Robert Fisher**. The patience and guidance from these men was instrumental for me. I cannot overstate how much the practice school has given me. During my job search it felt almost like cheating, since I could answer every situational interview question with an anecdote from my stations. Similarly, I cannot properly acknowledge the practice school without mentioning **Corning Inc.** and **General Mills**, the companies that sponsored us and designed the projects that my colleagues and I worked on.

It is no secret that I am a geek and a nerd. Luckily, this has never been a problem. In fact it has been a blessing during my time in graduate school. Mostly thanks to **my friends**, and the individuals that have helped make my time in MIT beautiful. Two distinct groups of people come to mind: my soccer, football, and kickball teammates, and my board games and Dungeons and Dragons group. I was blessed enough to meet people so fun that they overlapped between those two categories. In particular, I would like to thank **Zsigmond Varga, James Kaczmarek, Harry Watson, Alkiviadis "Orpheus" Chatzivasileiou, Andrew Fiore,** and **Daniel Consoli**. These fine gentlemen have given me some of the greatest adventures and most fun experiences of my life, and thanks to them, my time in Cambridge has been exciting and colorful.

**Family** is a complicated thing, messy and beautiful, just like research. I am very lucky to have my unique brand of craziness with my family, and it is thanks to their teachings, values, and support that I am the person I am. My **parents** have given me more than I deserve, in terms of resources and encouragement, and I hope that one day I can make them proud and sufficiently repay the sacrifices they have made for me, even though I do not know if that is possible. **My siblings** have

given me counsel, friendship, and even taken me in when I needed them. You have kept me sane and grounded in this world, and given me so much inspiration from your own work and accomplishments, even when I was stubborn and angry and tried to quarrel with you. Finally, I need to thank my **wife's family**, which has taken me in without question or hesitation in this foreign land that I now call my home. They have made me feel welcome and loved, when distance and circumstance prevented me from seeing my blood relatives. All of these people have given me love and strength, defining me as a person, and allowing me to move forward. For that, I am eternally grateful.

Last, but not least, I must recognize the most important person in my life: my wife **Maria Pons**. It is an understatement to say that since you came to my life, you have never stopped being there for me. In reality, had it not been for your love and support, I honestly cannot fathom how I would have been able to pull through during the last seven years. Even when you could not be in Cambridge during the first year of the PhD program, our daily calls in the evening were always the highlight of my days and my source of strength and inspiration. Once you finally moved to Massachusetts, you completely changed my life for the better. You have given me happiness that I did not even believe possible. Every day with you is a treasure and an adventure that I would not change for anything in this world, even when sometimes all we do is hang out in our living room and watch TV. I feel like I am the luckiest person in the world, having you to help me proofread my work and listen to my presentations, seeing your eager face as you hear me blabber about my experiments. You are my rock, my happiness, my love, and my life. Thank you.

# Table of Contents

# LIST OF FIGURES

12

# LIST OF TABLES

# Chapter 1 : Introduction

Crystallization is a key separation method for the pharmaceutical, chemical, mining, and food industries[1,2]. The technique allows for the generation of high-purity, ordered solid phases of a chemical entity, which is extremely desirable during product refinement and for the synthesis of specialty solid materials[3]. Traditionally, crystallization processes focus on the control of the supersaturation of the target molecules for a given solvent system, which is most commonly achieved through temperature variations, evaporation, and antisolvent addition[2]. Controlling supersaturation is effective for the generation of crystalline products; however, as a unique process handle, supersaturation lacks versatility and control, being susceptible to the incorporation of impurities to the product lattice, heavily constrained by the solvent system in use, and having limited ability to provide control over the product polymorphism and morphology.

One of the avenues of research for improving crystallization processes focuses on manipulating the chemical environment and the solution composition of the crystallizing system. The manipulation of the interactions occurring in the liquid phase prior to crystallization opens a versatile field of chemistry, in which it is possible to change the purity and structure of the product beyond the limitations of using supersaturation as the unique process parameter. The presence of a carefully designed additive can substantially alter the chemical environment of a solution. As a result, the performance of a crystallization procedure can be altered while still attaining to the constraints of the process design, such as temperature curves and solvent composition.

It is common for crystallization processes to have dissolved impurities and side-products, which can interact in complex and unpredictable ways with the crystallization target, changing solubility and supersaturation[4–20], crystal growth rates[21–25], product purity[1–5,26], product morphology[2,5,21] and crystal form[1,7,27]. The mechanisms for many of the aforementioned impurity-driven disruptions have been extensively reported, and in some cases are well modelled and understood[4–8,21–27]. However, the exact mechanistic effect of impurities and additives on crystal nucleation remains a difficult challenge, both for academia and industry[4,22]. Similarly, the utilization of tailor-made additives as proactive tools to improve crystallization processes, rather than to disrupt them, is a young and unexplored field that requires development and characterization.

This thesis explores the implementation of controlled complexation, studying the mechanisms through which it can alter crystal nucleation rates and enhance purification in crystallization. Focus has been given to applied experimental work, guided by theoretical and computational modelling for the design of successful additives. The work shown has expanded the existing knowledge on the behavior of impurities in solution on crystallization, and provided avenues for controlling yield and purity without requiring alteration in operational conditions and procedures. Additionally, the findings herein have provided the first systematic analysis and demonstration of impurity-driven nucleation inhibition, and provided a theoretical basis for its understanding.

## 1.1 Overview of crystal nucleation, growth, and polymorphism

Crystallization is the process upon which a compound in the liquid or gas phase transforms into its solid state[2]. The formation of a crystal requires two major mechanisms: nucleation, followed by growth. The driver for crystallization is a thermodynamic quantity known as supersaturation, which is related to the excess concentration of the crystallizing species in the non-solid phase with respect to its saturation concentration, leading to a high chemical potential[2,28–31]. The excess in chemical potential leads to the instability and metastability of the non-solid phase. Although nucleation and growth are driven by supersaturation, they proceed independently form one another,

17

and possess unique kinetic parameters that vary according to the experimental conditions under which the crystallization occurs[2,5,32,33]. The combined effect of growth and nucleation, coupled with the innate chemistry of the solute, define the polymorphism, purity, size distribution, morphology, and physicochemical properties of the product crystal. As a result, proper control of these two steps is crucial for engineers in the design of industrial crystallization processes.

The understanding of nucleation itself is limited, given that the most prevalent theory to model it (classical nucleation theory) is ineffective and inaccurate, having little-to-no predictive capabilities for the kinetics of true crystallization processes[1–3,34,35,]. The classical nucleation theory (CNT) model was proposed by Gibbs[1,2,36]. This model assumes that in a supersaturated solution, nuclei of various sizes form within the solution. These nuclei are assumed to have identical morphology and properties to the final crystal. Small nuclei are too unstable and dissolve back to the supersaturated solution. Eventually, a large enough nucleus of a "critical size" is formed, and proceeds to grow, one molecule at a time. CNT proposes an extremely idealized notion of how crystallization develops. As a result, CNT lacks any predictive capability for applied crystallization processes. Similarly, several phenomena related to crystallization cannot be rationalized by the CNT model[37,38], such as the effect of impurities and heterosurfaces on nucleation rates and the unique nature of protein crystallization. Due to the limitations of the CNT model, and thanks to observations made during the crystallization of proteins, an alternative model was proposed, called the two-step mechanism of crystallization[1,3,35,38,39]. The two-step model stipulates that within a supersaturated solution, unstable liquid clusters form intermittently with high concentration of the solute. Within these clusters, the solute molecules come close to one another and eventually experience an ordering step, in which they arrange into the morphology of a crystal. The two-step model is substantially more versatile than CNT and has the ability to better rationalize true crystallization phenomena. However, the two-step model is a relatively modern theory that is still under debate and has not been fully characterized from a fundamental and theoretical perspective, even though it has already been embraced by most of the academic crystallization community[1].

Nucleation can be split into two types, primary and secondary. The first refers to the formation of a new crystal from a homogeneous solution, the second is the formation of new crystals from the breakage of already existing seeds and nuclei. Primary nucleation can proceed homogeneously (from a clear, undisturbed liquid), or heterogeneously. The later mechanism takes place due to the presence of surfaces, dust, and particles that lower the energy barriers required for nucleation, hence facilitating crystallization. In real-life applied crystallization processes, nucleation almost exclusively onsets from heterosurfaces, given that homogeneous nucleation is extremely slow and unreliable, and that it is virtually impossible to remove the presence of surfaces and particulates during a crystallization experiment. In industrial processes, secondary nucleation is often the preferred mechanism for crystallization, and it is controlled using what are known as "seeds". Seeds are a controlled quantity of the desired crystal that are added to the crystallization reactor, with the objective of promoting secondary nucleation and crystal growth. The presence of seeds minimizes the formation of undesirable small nuclei, known as fines, which are difficult for post-processing and product handling. Additionally, the presence of seeds allows for better control of the product morphology and polymorphism.[1,2,40]

Crystal growth is the process that takes place once a nucleus or a seed exists within the parent solution, while the liquid phase remains supersaturated. Growth refers to the attachment of additional solute molecules to a parent crystal, increasing its size. The kinetics of growth can be quite complex, as they are heavily dependent on the chemistry of solute, the surface chemistry of

the crystal form that is nucleated, and the presence of impurities and additives that can poison or hinder the attachment frequency[32,33,40,41]. Even within the same crystal, different faces can have varying linear growth rates, leading to anisotropic behavior, the formation of preferred faces, and non-spherical crystal morphologies. Nucleation and growth occur simultaneously in solution, as long as the liquid phase remains supersaturated. This implies that old nuclei grow bigger, as new crystals are formed. Because of this behavior, the final equilibrated system displays a large amount of crystals with a wide range of sizes, known as the crystal size distribution (CSD). Control of the primary ranges of CSD at the end of crystallization experiments is one of the most crucial parameters that process engineers must tackle when designing crystallization operations[2,3,42].

Crystals are ordered solid structures of the molecules of a given material. The arrangement, order, and orientation in which those molecules are packed determines what is known as the crystal form. When a given material can have multiple crystal forms, thanks to the ability of its molecule to pack in different orientations or conformations, that material is defined as having "polymorphs"[2]. Polymorphism is a crucial phenomenon in crystallization processes and a major consideration for industry and crystal engineers. Different polymorphs of the same material can have different densities, solubilities, stabilities, surface chemistry, and even reactivity. As a result, it is imperative for experiment and process design to understand and control the specific desired polymorph of a given product. It is important to make a distinction between crystal form and crystal morphology. Polymorphism refers to the molecular packing of the crystallizing solute, while morphology is the physical shape of a crystal (plate-like, needle, prism, etc.)[2,43]. Different polymorphs are not guaranteed to have different morphologies, given that crystal shape depends on which functional groups are exposed in each of the crystal faces of a given polymorph, and how these surfaces interact with their solvent environment, and any impurities present in solution. Because of this, a polymorph can have multiple morphologies depending on the conditions in which it is crystallized. In fact, the use of tailor-made additives to control crystal morphology is a well-known and documented technique, extensively characterized and reported both in literature and industry[2,21,44].

Thermodynamically, polymorphism can be monotropic or enantiotropic, depending on whether the relative stability of the polymorphs of a compound is a function of temperature or not. When the particular stable form of a molecule changes at a given temperature, then the system is enantiotropic. Conversely, when the stable form is the same regardless of temperature, the system is monotropic[2,22,43,45–50]. In terms of the actual geometric and physicochemical arrangement of the molecules, there are two types of polymorphism: packing and conformational[2,43,47,50]. Conformational polymorphism occurs when the molecule can have different stable structural conformers. The possible conformers are stabilized differently as the crystallizing solvent and temperature are changed. As a result, the change in the relative geometry of the molecule can lead to different crystal forms, and changes in the lattice interactions. A molecule can still exhibit multiple polymorphs, even when it is more structurally simple and possesses only one single stable conformer. Multiplicity in crystal forms arises in this case from different packing arrangements of the same molecule. This is what is known as packing polymorphism. This polymorphism stems from how the molecules are oriented with respect to one another in the unit cell, rather than how the molecule itself is structured.

## 1.2 Impurities in crystallization

The presence of undesired molecules in a product stream is a ubiquitous problem in the chemical industry. The need for chemical separations, product processing, and impurity removal is one of the main reasons process designers and chemical engineers are vital players for chemical process

design. Crystallization unit operations are in no way exempt from impurities; in fact, one of the major benefits of crystallization is that it serves as a superb purification technique, given that the solid products it generates tend to reject a substantial portion of the chemicals present within the original mixture[2]. The presence of impurities within the parent solution can have a major impact on the way a crystallization step proceeds, potentially affecting all stages of the process: nucleation, growth, morphology, polymorphism, and product composition[2,4,5,7,27,43,51,52].

Different types of impurities exist within the context of crystallization, depending on how the undesired molecule attaches and remains within the final crystal. One of the most common types is surface impurities, which deposit on the solid-liquid interface between the solvent and the crystals, and are either attached through a chemical interaction with the crystal surface, or become deposited and adhered as the solvent is evaporated. This type of impurity is easily preventable, as long as an appropriate washing step is performed to remove the remaining parent solution (also known as "mother liquor") once the crystallization step has finished. A similar type of impurity arises from solvent-entrainment, and usually occurs for systems with large or porous crystals, where parts of the liquid phase can become trapped in the solid product and are not easily washed away. This type of impurity can be avoided by carefully controlling the change in supersaturation as a function of time during the crystallization, to ensure that the crystals do not grow too quickly. Slower-growing crystals are less likely to entrain solvent, and tend to have smoother morphologies, preventing impurity encapsulation. The third and arguably most complex impurity incorporation mechanism occurs for lattice impurities. This type of impurity places itself within the crystal lattice, either within lattice imperfections and defects or within the crystal unit cell, if the impurity is structurally similar to the crystallizing molecule. Lattice impurities are by far the most tenacious and difficult to remove. To address lattice impurities, it is necessary to either change the crystallization conditions (nucleation and growth rates, solvent, pre-processing to reduce impurity concentration in feed stream, etc.) or perform multiple consecutive crystallizations with the same material to progressively diminish the number of impurities remaining. These solutions can prove to be expensive or difficult to carry out, as they can compromise the product yield (when successive crystallizations are applied) or require the re-design of an entire unit operation, which in turn might not even be possible due to the physical and chemical constraints of the system.[2]

As impurities adhere to a crystal's surface, they generate a chemical and physical "blockage" which can substantially impact the way that crystal develops. It is relatively common for the presence of a specific impurity to completely change which faces of a crystal grow, given that depending on how the molecules interact, it is possible for the impurity to severely poison one of the surfaces and prevent it from growing[2]. This phenomenon can actually be used proactively when crystals are not desired, which is the case in some industrial and medical applications. The use of "tailor-made" additives to manipulate crystal growth and morphology is well-known and a fertile area of research and development within the crystallization process development community[2,21,23,53]. The impact of impurities on crystal polymorphism and nucleation is much less understood[38,51,52] and is an area of research that needs to be expanded upon, given that there is evidence of such effects taking place. However, to the knowledge of the author, the use of tailor-made additives to manipulate polymorphism and nucleation rates has not been properly characterized and explored.

## 1.3 Solution complexation and cocrystallization

The existence of strong intermolecular interactions between the nucleophilic and electrophilic functional groups of different molecules is a well-known chemical phenomenon that has massive implications on the behavior of multi-component chemical systems. Often called hydrogen bonding, the existence of this of electrostatic effect is a fundamental component of nature and chemistry, governing and modifying the behavior of molecules and dictating the physical properties of most materials.[54-56]

When two or more molecules share strong hydrogen bonds in solution, they form what is known as a solution complex, and the individual chemical species are called co-formers of one another. The chemistry behind hydrogen bond formation is extensive and varies wildly depending on the specific chemical structure of the species involved[54]. Correlating the structure and properties of molecules is a key aspect of the work done to select successful interacting pairs. Additionally, the presence of solvent shielding and interactions can severely complicate and alter the strength of a complex, since the solvent helps define the chemical environment and the entropy of the thermodynamic entity that is a complex[54-58]. Generally speaking, when two molecules come together, they do so because it is energetically favorable, and their functional groups are stabilized as they share partial charges and electrons. However, as the molecules become bound to one another, they lose part of their ability to vibrate freely within their chemical environment, and are no longer able to interact as well with the solvent molecules around them. As such, for a complex to be thermodynamically favorable it must not only be energetically beneficial, but the entropic penalties associated with it must also not counteract the enthalpic gains.

Complexes display unique behaviors in solution and can impact the solubility and equilibrium of the species, potentially affecting all physicochemical events that transpire in the solution in which they are present, crystallization being one of such potential events. Within the context of applied crystallization, co-formers have recently been a topic of extensive discussion, as they are crucial in predicting and facilitating the formation of cocrystals; cocrystallization is one of the most prolific areas of research for crystal engineers and scientists. A cocrystals is a crystal that contains more than one molecular species within its unit cell and repeat unit. A great amount of debate exists on the exact definition of the term "cocrystal", given that it is unclear whether it encompasses crystal hydrates and solvates of molecules, or if it specifically refers to crystals of non-solvent molecules. Interest in cocrystals arises from the fact that the presence of multiple molecules allows tuning the solid properties of the desired product, by changing its co-former. For example, an active pharmaceutical ingredient (API) that has poor solubility and bioavailability by itself might have favorable properties if it is manufactured as a cocrystal, alongside a co-former that facilitates its aqueous solubility, or stabilizes it, preventing its decomposition, or limiting its reactivity when exposed to the atmosphere. [59-64]

Solution complexation is useful for more than just cocrystal design. Complexes in solution can be used to control and manipulate the solubility of molecules by generating speciation equilibrium between the free molecules and the bound complex[13-15,18,19]. Additionally, molecular interactions between an additive and the crystallizing solutes are indicative of the potential of that additive to act as a growth inhibitor[21,23,32] or as an impurity[2,65,66]. Finally, as will be discussed herein, complexation is a powerful tool that can be used to tailor-design additives to prevent lattice-impurities for incorporating into a growing crystal. Additionally, complexation can be used to hinder the formation of nucleation clusters, slowing down the onset of nucleation and changing the kinetics of crystallization.

## 1.4 Thesis outline

This thesis and its sub-projects are divided into two clearly defined categories: purification and nucleation inhibition. These two sub-sections represent different uses for solution complexation within the context of applied crystallization and help elucidate the importance of understanding solution chemistry when designing crystallization procedures. Each of the core chapters in this report provides a detailed review of a well-defined project that was completed as part of the body of work of this thesis. Each project has either been published in a peer-reviewed journal, or officially finalized without publication, but written as an independent complete entity. Chapters 2 and 3 focus on the utilization of complex-assisted crystallization as a technique for enhancing the purification capabilities of crystallization without compromising yield. Chapter 4, 5 and 6 study the use of solution complexation to controllably regulate nucleation rates of small-molecules and to manipulate their polymorphism.

Chapter 2 provides a study of the first-ever successful implementation of complex-assisted crystallization for the separation of structural isomers. The work described in this chapter expands on past discoveries and thesis work from members of the Myerson group[65–68]. However, previous work using solution complexation as a purification tool focused on systems with a very limited set of functional chemistry. This chapter explores for the first time how to establish selectivity of interaction when both the crystallization target and the impurity shares the same functional groups, differing only in the relative spatial arrangement of these groups within the molecular backbone. The experimental analysis focuses on the comparison of final purity and yield on batch crystallizations in the presence or absence of candidate complexation additives. The additive selection is outlined in detail, comprising a combination of experimental and computational tools. The purification data has also been analyzed and correlated to obtain insight on the chemical characteristics that define a successful complexing agent. The intent of this analysis is generating fundamental understanding that can facilitate future molecular selection in future complex-assisted crystallization projects.

Chapter 3 explores the purification of an industrially-relevant active pharmaceutical ingredient (API) using complex-assisted crystallization. The molecule studied is a proprietary drug belonging to Novartis International AG. The purification of this API was recommended by Novartis as a test for the limitations and applicability of complex-assisted crystallization as a viable separations technique for pharmaceutical crystallization processes. The selected API has a much larger and more complex structure than any chemical previously purified using solution-complexation. Additionally, the structural differences between the impurity and the API are subtle and represent a new challenge within the context of additive selection and interaction selectivity. The work discussed in this chapter is procedurally similar to that of chapter 2, requiring batch experiments and computational analysis for complex selection. However the conditions of the batch experiments are unique due to the stringent purity and yields required for the commercial drug that is being separated. The crystallization data is a key part for delineating the achievable thermodynamic limits for complex-assisted crystallization.

Chapter 4 is the first section of the thesis focused on the second topic of work: nucleation inhibition. The project therein studies the stochastic phenomenon of nucleation using high-throughput experimentation to measure the probability distributions of induction time. Benzoic acid was the chemical chosen for the measurements, based on its simple structure and favorable solubility and polymorphic behavior. This chapter investigates the effect of varying concentrations of a strongly-interacting complexing agent on the nucleation rate of benzoic acid. The experimental work

focuses on de-convoluting the kinetic and thermodynamic drivers for changes in the measured induction times. Particular care is placed in supersaturation monitoring and control to ensure that the driving force for crystallization is properly accounted for and normalized, even in the presence of the additive inhibitor. This work is, to the knowledge of the author, the first to systematically demonstrate the effect of a small-molecule complexing agent on the nucleation rate of a solute using high-throughput induction time data. The impact of impurities on crystallization kinetics is a crucial question commonly addressed by crystallization and purification engineers. However, past studies and analysis have mostly focused on studying the effect of impurities on growth rates rather than nucleation. This is in great part due to the tough hardware requirements for completing a proper nucleation rate measurement. The results reported in this chapter demonstrate fascinating trends that pose provocative questions about the impact of small concentrations of impurities and additives on crystal nucleation, laying the ground-work for future projects in the Myerson group, including the ones reported in chapters 5 and 6.

Chapter 5 further explores the effect of dilute impurities and complexing agents on nucleation rates. The project reported in this chapter considers an alternative set of molecules to the ones discussed in chapter 4. The new chemistry focuses on demonstrating that the findings of chapter 4 are generally applicable and to determine if a much weaker complexation interaction can also lead to significant nucleation inhibition. The objective of the experimental work within this project is finding the nucleation rate kinetics of the chosen solute (3-nitrophenol) in the presence and absence of a complexing agent. The resulting kinetic parameters are used to rationalize the underlying inhibition mechanism driving the experimental observations. The proposed mechanism considers and compares the two prevalent theoretical models for nucleation: classical nucleation theory and the two-step nucleation model. The findings provides new key insights for the theoretical understanding of the effect of impurities on crystal nucleation and for the possibility of using solution complexation as a process handle to enhance control over crystallization procedures.

Chapter 6 summarizes the last project of this thesis work. The premise of this project is that complexation can be used to manipulate the conformational arrangement of a crystallizing solute in solution, which may in turn affect the product polymorphism. The effect of impurities on product polymorphism is a known phenomenon[27,69]; however, it has not been systematically and experimentally rationalized within the context of crystal nucleation. This chapter explores using a model solute that exhibits conformational polymorphism to demonstrate that altering the conformation arrangement in solution using impurities can affect the final crystal polymorph. The experiments are focused on high-throughput nucleation measurements to analyze the statistical distribution of possible polymorphs as a function of the additive concentration. Experiments are still being completed by the time of writing this report, but are expected to be completed within the next couple of months.

Chapter 7 provides concluding remarks and a round-up summary of the major discoveries and accomplishments elaborated within the body of this thesis. Finally, the chapter provides recommendations for future work and application of solution complexation.

# Chapter 2 : Purification of nitrophenols using complex-assisted crystallization

The contents of this chapter were adapted from work published in a peer-reviewed journal article[70]:
Pons-Siepermann, C.; Huang, S.; S. Myerson, A. Purification of Nitrophenols Using Complex-Assisted Crystallization. *Cryst. Eng. Comm.* **2016**, 18 (39), 7487–7493.

## 2.1 Introduction

Crystallization is the preferred method for the production of high-purity solid phases[1]. However, in some cases, impurities which are structurally similar can incorporate into the crystalline lattice in amounts as high as 1-3%. This can be a particular problem in the crystallization of active pharmaceutical ingredients[2].

Recent work performed by Weber[66] and Hsi[65,67] focused on a new method for the enhancement of crystallization performance through the reduction of impurity concentrations in crystal products without decreasing yield. Their efforts explored the purification of small-molecule organics by removing a single impurity that was structurally similar to the purification target. The strategy employed by Weber and Hsi to remove impurities consisted on adding a complexing agent to the crystallizing solution that would interact with the impurity molecule. The impurity/complexing agent interacting pairs were shown to be too bulky to incorporate into the crystals of the purification target, producing high-purity solids in a single crystallization step.

The work done in complex-assisted crystallization has shown success for relatively simple systems. Results have been obtained for the purification of benzamide, cinnamamide[66,65], and amoxicillin[67], by the removal of their respective impurities: benzoic acid, cinnamic acid, and 4-hydroxyphenylglicine. In all of these case-studies, the impurities have a common carboxylic acid functionality that is not present on the purification target. This common structural moiety observed for the target/impurity pairs studied was instrumental to the success of the experiments: these works exploited the acidity and excellent hydrogen bond donating capacities of the carboxylic acid groups[55,54] of the impurities as the target for their complexation motifs.

Despite the success of the preliminary efforts on complex-assisted crystallization, the scope and limitations of this technique are not well-known and need to be characterized. There is currently no known example of successful implementation of this technique on systems where the impurity molecules do not contain carboxylic acid functionalities. The present investigation focused on expanding the known applicability of complex-assisted crystallization by studying the purification of 4-Nitrophenol (4NP) and the removal of its isomer 3-Nitrophenol (3NP) as a lattice-incorporating impurity. This system was chosen because it departs from the framework employed in previous research on complex-assisted crystallization, attempting to expand the known limits of the technique. The chosen molecules are structural isomers, none of which contain a carboxylic acid functional group. Given that the target and its impurities contain the same functional groups, the complexation had to be implemented targeting the differences in relative proximity of the functional groups present within the molecules.

This chapter reports the experimental results obtained for the purification of 4NP in water, using 3-aminobenzoic acid (3ABA) as a complexing agent for 3NP. The structures of 4NP and 3NP are shown in Figure 2-1. The effect on purity and yield caused by varying the amount of complexing agent employed were also explored. The crystal form of the products was assessed using X-Ray powder diffraction (PXRD) to establish the effect of the complexing agent on the product crystallinity. Additionally, a separation procedure using nanofiltration membranes was used to

demonstrate the formation of a 3NP-3ABA complex in solution, by monitoring changes in 3NP rejection coefficients in the presence of the complexing agent. The intent of this project is to expand the known applicable types of systems that can be purified through complex-assisted crystallization and further delineate the scope of the technique.
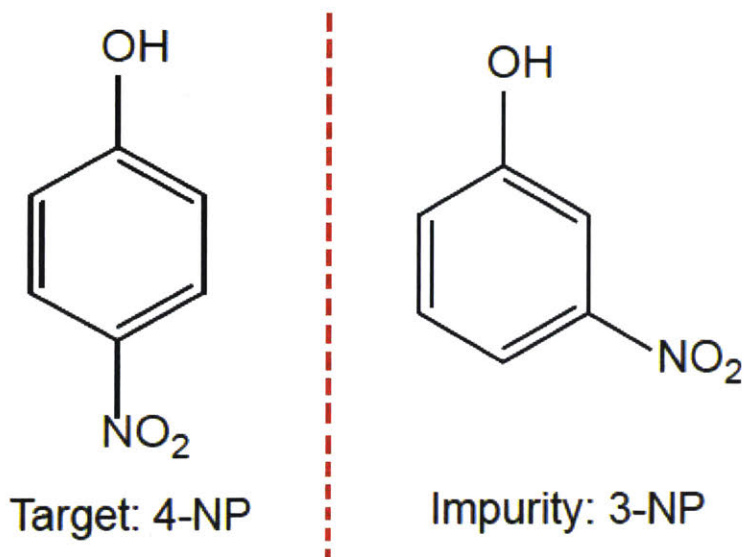


**Figure 2-1:** Molecular structure of 4-nitrophenol (left), and 3-nitrophenol (right). These molecules are respectively the purification target and impurity.

## 2.2 Experimental

### Materials

All compounds were used as received. 4-nitrophenol (4NP), 3-nitrophenol (3NP), 3-aminobenzoic acid (3ABA), succinamic acid, guanidineacetic acid, γ-Aminobutyric acid, taurine, β-alanine, water, methanol, and trifluoroacetic acid (TFA) were all purchased from Sigma Aldrich. The nanofiltration membrane, 24-PBI-DBX was obtained from the Livingston group from the Imperial College of London.

### Complexing Agent Selection

The screening for complexing agents was performed using a coarse computational assay based on energy minimizations, and by direct verification in crystallization experiments. The computational aspect of the work in this chapter is substantially simplified from the original efforts performed by Weber *et al.*[66], implementing forcefield energy minimization in the vacuum of the complex pairs, rather than DFT simulations and solvation models. The purpose of the computational work was not to identify the exact structure of the best complexing agent, but to elucidate the molecular motifs that were likely to yield a stable complex structure and could be ultimately verified in the experimental assays. To perform the computational verification, the molecules were modelled in a vacuum in Material Studio, and allowed to interact with each other using the COMPASS27 forcefield. Further detail on the calculations is provided in the appendix. The selected complexing agents were each tested in crystallization experiments and their performance was compared to the purification results obtained in a control crystallization (no complex added). The best-performing complexing agent was also studied further, characterizing the effect of varying the amount used in each experiment on the product purity and crystal structure.

## Crystallization Experiments

The crystallization experiments were performed in triplicate for each batch. In each standard batch, 4NP (1.125 g, 8.08 mmole) was dissolved with 3NP (100 mg, 0.71 mmole) and varying amounts of complexing agent in DI water (12.5 ml) at 85 °C, with the temperature controlled with a recirculating water bath. The solution was cooled to 5 °C at a rate of 0.44 °C min$^{-1}$ under vigorous stirring. The slurry was stirred at 5 °C for three hours and then vacuum filtered, washed with 15 ml of ice water, dried overnight at 25°C and 25 inHg of vacuum, and weighed. The drying temperature was selected to be close to room temperature to avoid sublimation of the nitrophenols. Sublimation can lead to changes in the solids composition, and was observed at higher temperatures, and verified with thermo-gravimetric analysis (TGA). TGA data is shown in the appendix. The purity of the collected solid was determined using high performance liquid chromatography (HPLC). The mother liquor composition was also monitored with HPLC to determine the product yield and monitor the material mass balances.

The amount of 4NP used experimentally was selected based on the solubility of the solute to obtain a yield of 90% (by mass). The amount of 3NP was selected to be as high as possible to maximize impurity incorporation without incurring co-precipitation of the impurity at the lower experimental temperatures. Co-precipitation was monitored by preparing batches of saturated solution of 4NP and varying amounts of 3NP. The solids were dissolved at room temperature and the solutions were allowed to equilibrate overnight at 5 °C with vigorous stirring. The selected concentration of 8 mg/ml of 3NP was found to be as close to saturation as possible while avoiding co-precipitation. Similar experiments were also carried out with the other nitrophenol isomer, 2-Nitrophenol (2NP). However, the solubility of 2NP is substantially lower in water than that of 3NP and 4NP[71], and it was found experimentally that no significant incorporation of 2NP into the 4NP lattice could be observed without co-precipitation. This was observed with impurity concentrations as low as 0.8 mg/ml in the batch setup. For this reason, no further work was pursued for the removal of 2NP.

Water was selected as the experimental solvent given that 4NP shows a steep solubility curve in it[71]. It was observed that most other less polar solvents dissolved the nitrophenols at very high concentrations, making it difficult to obtain high (>80 wt%) product yields. Furthermore, it was observed that aprotic solvents with reasonable solubility curves (notoriously toluene) did not yield any measurable impurity incorporation, even at high initial impurity concentrations. This difference in behaviour is likely explained by the formation of different product polymorphs, as the change in the lattice structure may lead to different impurity incorporation mechanisms[48]. It was indeed observed that 4NP crystallized from water yielded the thermodynamically unfavourable beta (β) form, while 4NP crystallized from toluene yielded the thermodynamically stable alpha (α) form. These crystallographic results were determined using PXRD and where consistent with reports on 4NP polymorphism[72].

## Membrane Separation

Membrane experiments were carried out in a custom setup shown in Figure 2-2, replicating previous studies[73]. The system is designed as a cross-flow filter: the feed and retentate respectively enter and exit the filter chamber parallel to the membrane surface, whilst the permeate flows perpendicularly across the membrane. The retentate chamber is large (150 ml), allowing for the inclusion of a magnetic stir bar. The feed is pumped at a high pressure (420-450 psig) using a Thermo Scientific Dionex AXP HPLC pump. The system temperature was controlled at 50 °C using a heating plate, which also allowed for a controlled stirring rate of 400 rpm. The retentate

and permeate outlets both have Swagelok® pressure valves, which allowed pressurization of the filter chamber to the desired operating condition. The permeate valve was kept with no resistance to allow for maximum flow across the membrane. The selected operating flowrate was controlled with the pump and was set to 2.5 ml/min, with the actual flowrate measured to be 2.65 ml/min.
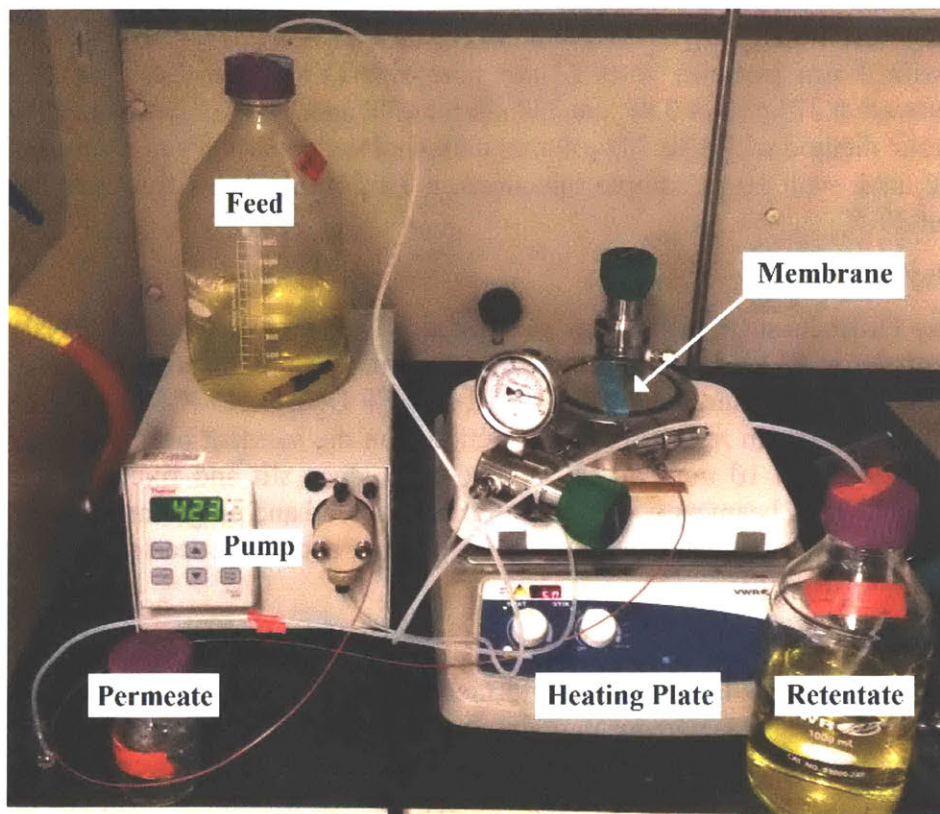


**Figure 2-2:** Picture of the membrane setup for the complexation experiments. The feed solution is supplied to the membrane using an HPLC pump. The membrane casing has valves to regulate the pressure in the retentate and permeate outlets. Temperature was controlled using a heating plate.

The performance metric monitored for the membrane experiments was the rejection factor (RR), as defined by Ferguson et al[73]. The RR is a measure of the fraction of solute that is unable to pass through the membrane. This metric is calculated using equation **2-1**, where $C_p$ is the concentration of solute (3NP) in the permeate stream, and $C_f$ is the concentration of the solute in the feed stream. The concentrations of solute were collected in aliquots and monitored using HPLC. The reported values of RR were determined once the system reached a steady state, which was observed to occur after approximately 6 hours of continuous operation.

$$RR = 1 - \frac{C_p}{C_F} \qquad\qquad 2\text{-}1$$

The membrane used, 24-PBI-DPX, is a polybenzimidazole (PBI) polymer cross-linked with α,α'-dibromo-p-xyelen (DPX) and was prepared following the procedure outlined by Kim et al[74]. The "24" indicates the weight percent of polymer in the membrane dope composition used during synthesis[73–75]. This membrane was selected based on recommendations from the supplier for having the lowest molecular weight cutoff (MWCO). The MWCO is defined as the molecular

weight at which 90% of the solute is retained by the membrane. The MWCO of 24-PBI-DPX is 250-300 g/mole.

## High-Performance Liquid Chromatography

For nitrophenol and 3ABA detection, the HPLC instrument (Agilent 1260) was equipped with a UV-diode array detector (DAD). The column used was a YMC-Pack ODS-A 150 mm X 4.6 mm i.d. packed with 3 μm particles with 12 nm pore size (YMC America Inc.). The detection wavelength was set at 275 nm for 3NP and 230 nm for 4NP and 3ABA. The samples were analyzed using an isocratic method with a 30:70 (volume) methanol/water mobile phase containing 0.3 %vol. trifluoroacetic acid, with 10 μl sample injections, a 1 ml/min solvent flow rate, and a column temperature of 30 °C.

## Powder X-Ray Diffraction

Powder X-Ray Diffraction (PPXRD) was performed on all samples using a PANalytical X'Pert PRO diffractometer at 45 kV with an anode current of 40 mA. The instrument has a PW3050/60 standard resolution goniometer and a PW3373/10 Cu LFF DK241245 X-Ray tube. Samples were placed on a spinner stage in reflection mode. Settings on the incident beam path included soller slit 0.04 rad, mask fixed 10 mm, programmable divergence slit and fixed 1° anti-scatter slit. Settings on the diffracted beam path include: soller slit 0.04 rad and programmable anti-scatter slit. The scan was set as a continuous scan: $2\theta$ angle between 4° and 40°, a scan speed of 0.021989 °/s, and a step size of 0.0167113°.

## Thermogravimetric Analysis

Thermogravimetric analysis (TGA) was performed on a Q500 instrument from TA instruments connected with a nitrogen gas cylinder to maintain a flow rate of 25 mL min$^{-1}$ to keep the sample chamber under an inert gas environment. Between 5 and 10 mg of sample were loaded on platinum sample pans from TA instruments. The samples were allowed to equilibrate at 30 °C, then heated at 10 °C min$^{-1}$ to 100 °C, held isothermal at 100 °C for 10 minutes, and finally heated to 300 °C at 10 °C min$^{-1}$.

## 2.3 Results and Discussion

Complexing agent selection began with the proposal of a complexation motif that exploited the different geometric arrangements of the functional groups between the target (4NP) and impurity (3NP). All nitrophenols contain a weak hydrogen bond acceptor (nitro group) and a hydrogen bond donor (phenol group)[54]. As such, complexation selectivity can be achieved by choosing complexing agents that contain reciprocal hydrogen bond acceptor and donor groups at a relative distance from each other designed to match the desired interaction partner. By constraining the chain length and flexibility of the backbone of the complexing agents, it was possible to promote preferential hydrogen bonding with the impurity molecules[55]. Figure 2-3 illustrates the proposed motif.

The purification of 4NP was considered with multiple complexing agents that followed the motif shown in Figure 2-3. The chemicals considered are listed in Table 2-1.
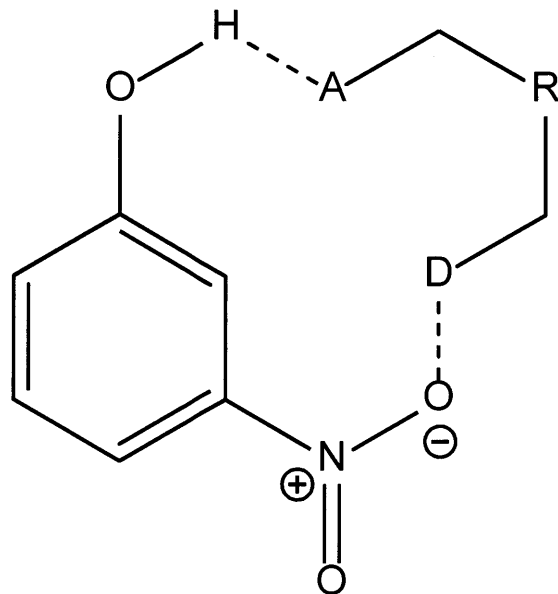
**Figure 2-3:** Complexation motif proposed for 3NP. The complexing agents were selected to have a complementary hydrogen bond acceptor (A) and hydrogen bond donor (D) to the functional groups of the impurity molecules. The exact structure of the backbone (R) is variable.

**Table 2-1:** Relevant information about the selected complexing agents. The interaction energy was calculated as explained in the experimental section. The second column indicates whether each complexing agent forms a zwitterion at the experimental pH.

| Complexing Agent | Zwitterion | MW (g/mol) | $\Delta E$ (kcal/mol) |
|---|---|---|---|
| β-Alanine | Yes | 89.09 | -25.4 |
| Guanidineacetic Acid | Yes | 117.11 | -24.6 |
| γ-Aminobutyric Acid | Yes | 103.12 | -23.5 |
| Taurine | Yes | 125.15 | -22.5 |
| 3-Aminobenzoic acid | No | 138.14 | -15.7 |
| Succinamic Acid | No | 118.1 | -14.0 |

All selected molecules have structures that follow the motif shown in Figure 2-3. The results in Table 2-1 indicate that the selected candidate molecules could all theoretically form energetically favourable interactions with 3NP. These calculations were performed as a coarse test to determine potential complexing agent structures. However, these energy minimizations are purely enthalpy-based and neglect important thermodynamic effects such as entropic penalties incurred during complexation, and solvent interactions with the dissolved nitrophenols and complexes. As a result, experimental verification was needed to conclusively determine the effectiveness of a given complexing agent. Each complexing agent shown in Table 2-1 was tested following the procedures outlined in the experimental section and compared to the performance of a control experiment (done without complexing agent). The results obtained are summarized in Table 2-2, a complete set of all experimental data are shown in the appendix.

**Table 2-2:** Experimental results for the purification of 4NP using multiple complexing agents. The 3NP is shown as a mass fraction of the final product. The yield is measured as the mass fraction of 4NP recovered. All experiments were performed with a 1:1 molar ratio of 3NP:complexing agent. Zwitterionic complexing agents are labeled with a "z".

| Experiment | 3NP (wt%) | 4NP Yield (wt%) |
|---|---|---|
| Control | 2.81 ± 0.31 | 90.6 ± 0.8 |
| β-Alanine$^z$ | 2.42 ± 0.09 | 90.9 ± 0.4 |
| Guanidineacetic Acid$^z$ | 1.58 ± 0.28 | 91.0 ± 0.5 |
| γ-Aminobutyric Acid$^z$ | 1.83 ± 0.57 | 90.7 ± 1.1 |
| Taurine$^z$ | 2.00 ± 0.07 | 91.3 ± 0.2 |
| 3-Aminobenzoic acid | 0.47 ± 0.07 | 89.2 ± 0.4 |
| Succinamic Acid | 1.08 ± 0.40 | 90.6 ± 0.8 |

The data in Table 2-2 shows that all complexing agents induced a measurable improvement in purity with respect to the control with minimal effect on yield. The experimental results indicate that there is no real correlation between the estimated enthalpy of interaction and the actual purity improvement. As explained previously, this results is expected given the coarse nature of the energy calculations and the omission of entropic considerations in the calculations.

Three relevant trends become apparent from the data:

a) The best performing complexing agents, 3ABA and succinamic acid, were the not zwitterionic molecules (see Table 2-1).

b) With the exception of taurine, higher molecular weight complexing agents were more effective (Figure 2-4).

c) The best performing complexing agent (3ABA) has the most rigid backbone structure, containing an aromatic six-membered ring (as shown in Figure 2-5).

The inferior performance of zwitterions can likely be attributed to their higher hydrophilicity compared to the non-zwitterionic complexing agents. Despite theoretically having stronger interactions with 3NP (as shown in Table 2-1), zwitterions are more strongly shielded by water than non-ionic species, making them less effective in the aqueous conditions employed[58].

The data shows a strong correlation between molecular weight and purification, with only taurine behaving as an outlier. This outlier is most likely due to the substantially higher acidity of taurine compared to all other selected complexing agents. As a stronger acid, taurine is more easily dissociated in the water and less likely to have weaker hydrogen bond interactions with 3NP.

The highest 4NP purify was achieved using 3ABA as a complexing agent. The structure of 3ABA is shown in Figure 2-5. The rigid backbone and meta distribution of the functional groups in the ring constrain the interactions to match the motif proposed in Figure 2-3. The aryl backbone of 3ABA is the most rigid of any of the complexing agents considered, making it have the least degrees of freedom of rotation lost upon complexation (minimizing entropic penalties). In addition, 3ABA is the bulkiest of the molecules, which makes its resulting complex larger and less likely to incorporate in the 4NP crystals.
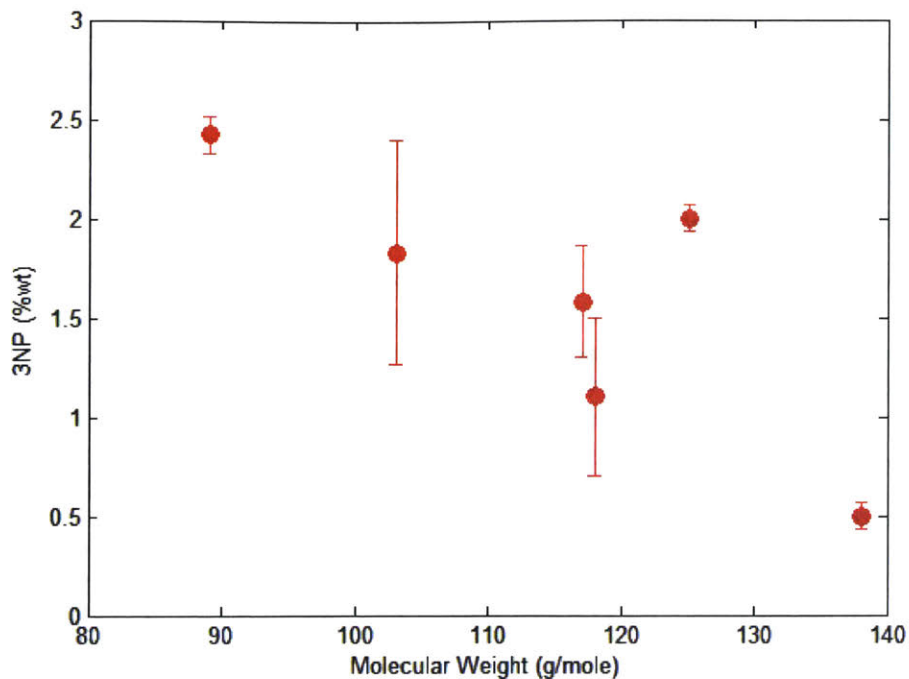
**Figure 2-4:** Correlation between the molecular weight of each complexing agent and the amount of 3NP remaining in the 4NP product after each corresponding crystallization experiment. Taurine is the only outlier of the decreasing trend, corresponding to the data point at 125 g/mol.
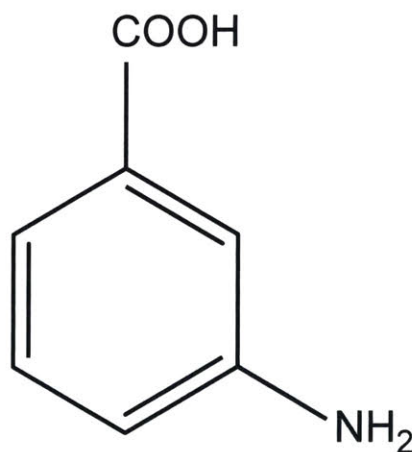


**Figure 2-5:** Structure of 3ABA at the expected experimental protonation state. Both functional groups can act as hydrogen bond acceptors and donors. However the strongest donor site corresponds to the carboxylic acid, and the strongest acceptor site is at the nitrogen's lone pair electron site[54].

The complexation of 3NP using 3ABA was further studied, given 3ABA's superior performance compared to the other complexing agents studied. Experiments were carried out to determine: the propensity of 3ABA of incorporating into the 4NP crystals, the effect of varying the amount of complexing agent used on product purity, and the effect of including 3ABA in the experimental solution on the 4NP polymorphism. Additionally, membrane experiments were carried out to demonstrate that the proposed complexation mechanism is taking place and causing the observed improvements in purity.

The amount of complexing agent employed was varied across multiple batch experiments. The variation of the amount of 3ABA was used to optimize the purification process. This strategy is consistent with previous complex-assisted crystallization work[67]. The incorporation of 3ABA into the 4NP product was also monitored for each of the experiments using HPLC. Incorporation of 3ABA is a prevalent concern, given that using a complexing agent would be counterproductive if it behaved as second impurity. The results for the purification of 4NP using 3ABA are summarized in Table 2-3.

**Table 2-3:** Results for the purification of 4NP with 3ABA. The 3NP is shown as a mass fraction of the final product. The yield is measured as the mass fraction of 4NP recovered. 3ABA is reported as a concentration (in ppm) in the solid products. Samples are labeled based on the concentration of 3ABA used in their preparation, the 7.9 mg/ml sample corresponds to a 1:1 molar ratio with 3NP and matches the value reported in **Table 2-1**.

| Sample | 3NP (wt%) | 4NP Yield (wt%) | 3ABA (ppm) |
|---|---|---|---|
| No complex | 2.81 ± 0.31 | 90.6 ± 0.8 | N/A |
| 3ABA 3.9 mg/ml | 2.01 ± 0.07 | 90.0 ± 0.1 | 100 ± 10 |
| 3ABA 5.9 mg/ml | 0.53 ± 0.16 | 90.2 ± 0.9 | 90 ± 5 |
| 3ABA 7.9 mg/ml | 0.47 ± 0.07 | 89.2 ± 0.4 | 220 ± 70 |
| 3ABA 9.9 mg/ml | 0.30 ± 0.04 | 90.6 ± 0.8 | 550 ± 200 |
| Second Crystallization | 0.19 ± 0.02 | 82.3 ± 0.3 | N/A |

The complexation results are presented in comparison to the same control shown in Table 2-2. In addition, a second control is provided as a reference, labeled "second crystallization". This sample is obtained from two consecutive non-complexed crystallization steps. The products obtained in the second crystallization experiment were obtained by re-dissolving and crystallizing the products from the first control. The amount of water used in the consecutive crystallization experiment was chosen to match the solid concentration equivalent to the first control experiment, leading to an immediate expected yield of 90% by mass, which represents approximately a 81% overall yield between the two experiments. Using consecutive controls has been a common reference in past complex-assisted crystallization work[66], given that the second step provides an excellent reference on how much purity can be improved whilst sacrificing yield. The goal of complex-assisted crystallization is to generate high-purity products with negligible yield penalties.

The experimental results shown in Table 2-3 demonstrate a significant improvement in purity upon the addition of 3ABA. The results show a range within the expected purity of the samples between 2.01 wt% 3NP (3.9 mg/ml 3ABA) to 0.3 wt% 3NP (9.9 mg/ml 3ABA). All results show consistent improvement with respect to the control (2.81 wt% 3NP). The observed average sample purity follows an increasing trend with 3ABA concentration. All samples had a detectable amount of 3ABA incorporated in the product, reaching a minimum of approximately 90 ppm for the experiments in which the least 3ABA was used, and increasing as more complexing agent was added. Higher concentrations of 3ABA were not tested because the increased incorporation of complexing agent detracted from the quality of the results. The observed incorporation matches the order of magnitude of purity requirements for organics in pharmaceutical products[76], having 500 ppm or less 3ABA, with the exception of the 9.9 mg/ml sample. The samples crystallized with complexing agent show on average an 8 point percent increase in yield with respect to consecutive crystallizations, and retain yields within 1 point percent to the single crystallization control. The purity increases with increasing 3ABA concentration, but slows down as lower 3NP concentrations are approached, without fully matching the consecutive crystallization purity. Regardless, the

samples prepared with 3ABA achieved up to an 89% increase in purity with respect to the original control, almost matching in a single step the two-step crystallization.

The solid products were analyzed with PXRD and compared in order to determine that the addition of the complexing agents did not affect the expected 4NP polymorph obtained during crystallization. A polymorphic change could indicate that the effect of the change in solution chemistry on the crystal structure, rather than the formation of a complex, is what leads to the improvement in product purity. This is particularly important when considering that no impurity incorporation of 3NP was observed in experiments carried out in solvents where the product 4NP formed in its more stable α form. The PXRD curves obtained are shown in **Figure 2-1**Figure 2-6. The complete set of non-overlaid PXRDs is shown in the appendix.

The data from Figure 2-6 shows that the crystal form of 4NP obtained from the crystallization experiments was the β polymorph, even when a high concentration of complexing agent was used. The lack of a polymorphic change further supports that the cause for the improvement in product purity is the complex formation between 3ABA and 3NP, rather than alternative potential mechanism based on lattice re-arrangement.

Nanofiltration experiments were carried out to demonstrate the formation of 3ABA-3NP complexes in solution as hypothesized in earlier sections. The membrane selected is designed to reject solutes based on their molecular size: the membrane is not permeable to chemicals with bulky structures, but allows small (<250-300 g/mole) structures to pass through. 3NP has a molecular weight of 139.1 g/mole, which means it should easily permeate through the membrane and yield very low RR values. On the other hand, a 3ABA-3NP complex would have a theoretical molecular weight of 277.4 g/mole, which should induce a measurably larger RR.

Experiments were performed using the setup described in the experimental section. A comparison was performed, determining the rejection coefficient of 3NP for a system with and without (control) 3ABA. The solutions chosen had a high concentration of 3NP (7 mg/ml) and the complexation experiment contained 7 mg/ml of 3ABA. The experimental results are summarized in Table 2-4.

The results show a 612.5% average increase of the rejection coefficient upon adding the complexing agent. This significant effect is a strong indication of the formation of a bulky complex that is less likely to pass through the membrane than the original solute, validating the proposed interaction between 3ABA and 3NP. The steady state progression of sample membrane experiments are shown in the appendix.

**Table 2-4:** Rejection coefficients of 3NP through 24-PBI-DBX with and without 3ABA.

| Sample | RR (%) |
|---------|---------|
| Control | 8 ± 1 |
| 3ABA | 57 ± 3 |
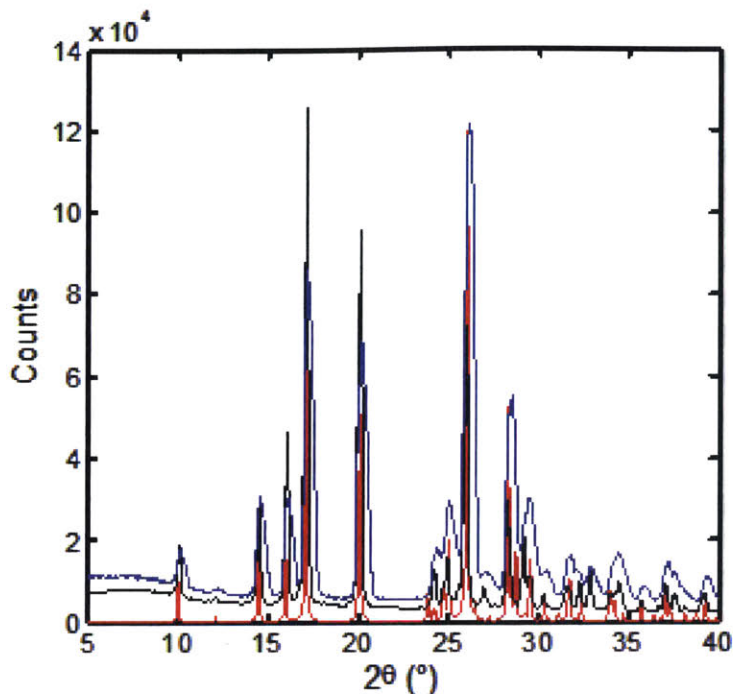
**Figure 2-6:** PPXRD curves of 4NP solid samples after filtration and drying. From bottom to top, the samples are: template of the β form from the Cambridge Structural Database (red)[72,77], control with no complexing agent (black), 4NP from batch prepared with 7.6 mg/ml of 3ABA (blue).

## 2.4 Conclusion

The use of complex-assisted crystallization was successfully implemented for the first time in the separation of structural isomers. The analysis was carried out for the purification of nitrophenols. The removal of 3NP using different complexing agent candidates was studied and shown to have a measurable impact on the product purity with little to no yield penalties. The best performing complexing agent was found to be 3ABA, which had the largest MW and most rigid structure of the molecules tested. The best experiments with 3ABA reduced the impurity inclusion by up to 89% with respect to the single crystallization control, performed without a complexing agent. It was shown using PXRD that the presence of 3ABA did not affect the crystal structure of the product and that the complexing agent would incorporate into the product in quantities of 550 ppm or less, depending on the amount used. The complexation mechanism was validated by monitoring changes in rejection coefficients of the impurity through nanofiltration membranes with and without complexing agent. For this particular case-study with nitrophenols, the observed improvement in purity was not sufficient to match the purity of two consecutive crystallization steps. These results indicate that obtaining complexation selectivity between isomers was more challenging than previous systems[66], where the impurities and the targets had different functional groups. However, this investigation proves that complex-assisted crystallization can induce substantial gains in purity for systems where the purification targets and impurities do not contain different functional groups. This work expands the applicability of complex assisted crystallization beyond the case studies explored in previous research. We believe that current and past data strongly encourages further characterization of the breadth and limitations of complex-assisted crystallization, as the technique shows great potential as a method for improving the effectiveness of crystallization as a purification strategy for structurally similar small-molecule compounds.

## 2.5 Appendix

### Computational Analysis

The energy of interaction between a complexing agent and 3NP was determined as the difference between the reported energy of each molecule placed in proximity to each other (interacting) and separately (non-interacting). Molecules that yielded a significant (<-5 kcal/mole) prediction of beneficial interaction during the formation of the complex were deemed appropriate for experimental verification. The calculated energies are listed below in Table 2-5.

**Table 2-5:** Calculated computational energies of interaction between selected complexing agents and 3NP. All values are reported in kcal/mol.

| Complexing Agent | $E_{separate}$ | $E_{interacting}$ | $\Delta E$ |
|---|---|---|---|
| β-Alanine[z] | -115.5 | -140.9 | -25.4 |
| Guanidineacetic Acid[z] | -200.1 | -224.8 | -24.6 |
| γ-Aminobutyric Acid[z] | -94.2 | -117.7 | -23.5 |
| Taurine[z] | 29.33 | 6.8 | -22.5 |
| 3-Aminobenzoic acid | -16.6 | -44.9 | -15.7 |
| Succinamic Acid | -72.0 | -86.0 | -14.0 |

### Individual Powder X-Ray Diffraction

The PXRD images shown in overlap during the main chapter are shown separately for detail in this section. It must be noted that the count intensity for the 3ABA sample PXRD has lower values than those shown in Figure 2-6 of the chapter because the whole PXRD curve was shifted upward for clarity purposes of that image, the relative intensity and position of the peaks was not modified.
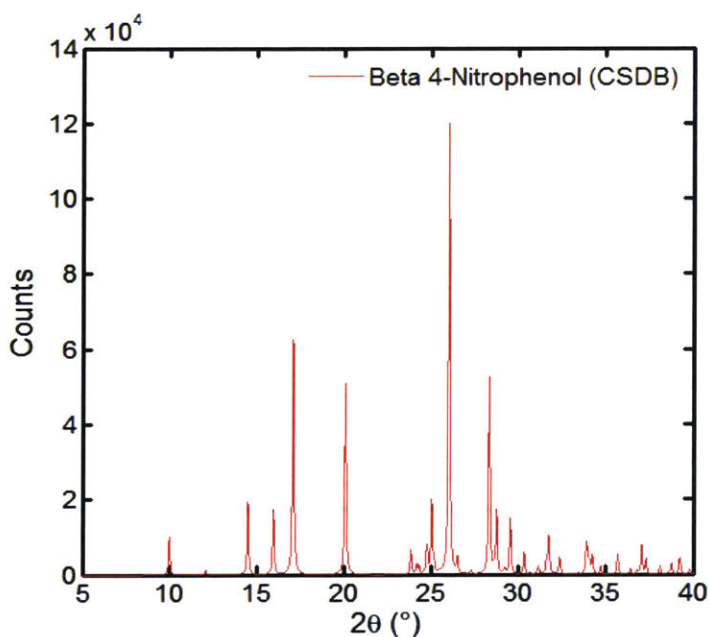


**Figure 2-7:** Theoretical PXRD for the beta polymorph of 4NP, obtained from the Cambridge Structural Database.
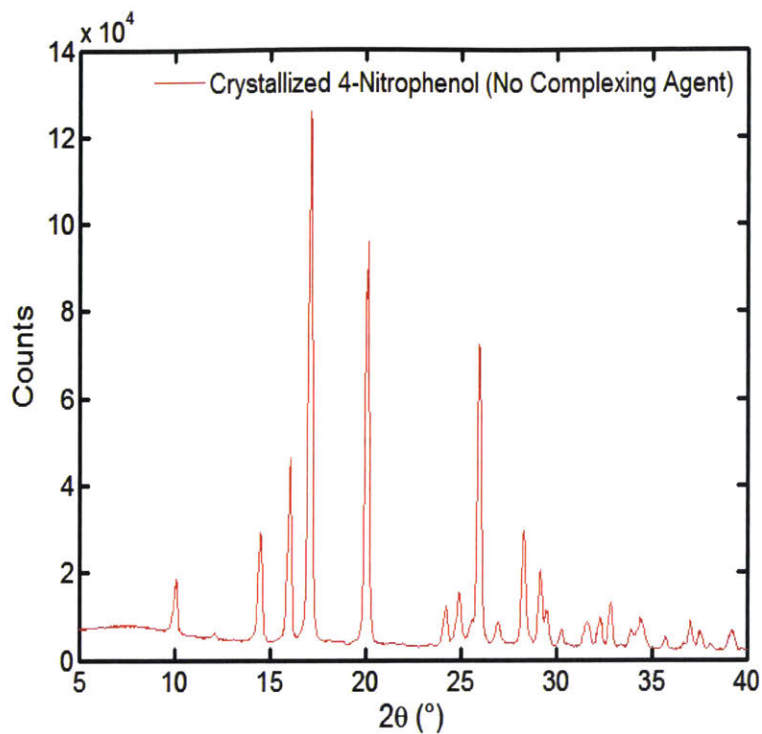
**Figure 2-8:** Experimentally obtained PXRD of 4NP, taken from a sample crystallized in water with no complexing agent. The graph matches with that of the beta form in Figure 2-7.
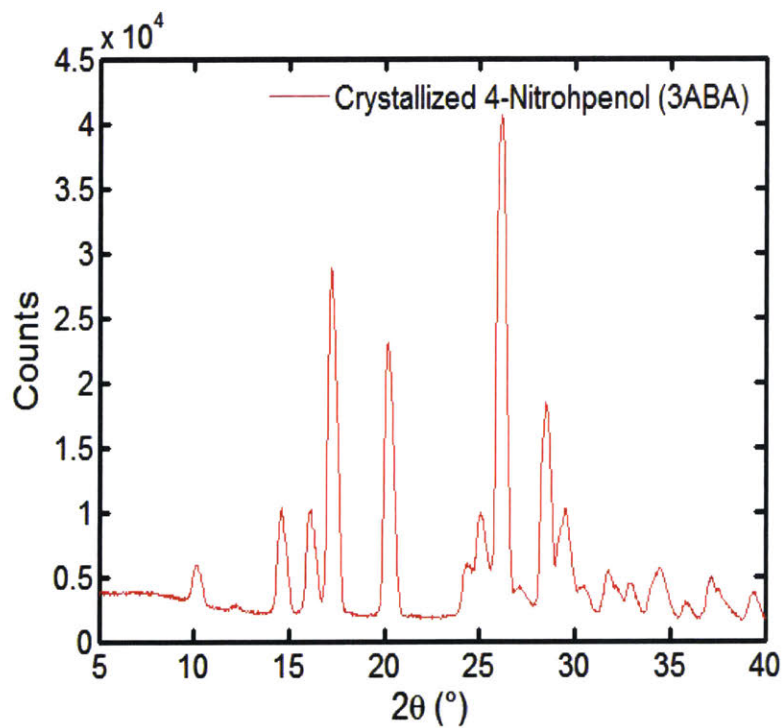


**Figure 2-9:** Experimentally obtained PXRD of 4NP, taken from a sample crystallized in water with 3ABA as the complexing agent. The graph matches with that of the beta form in Figure 2-7.
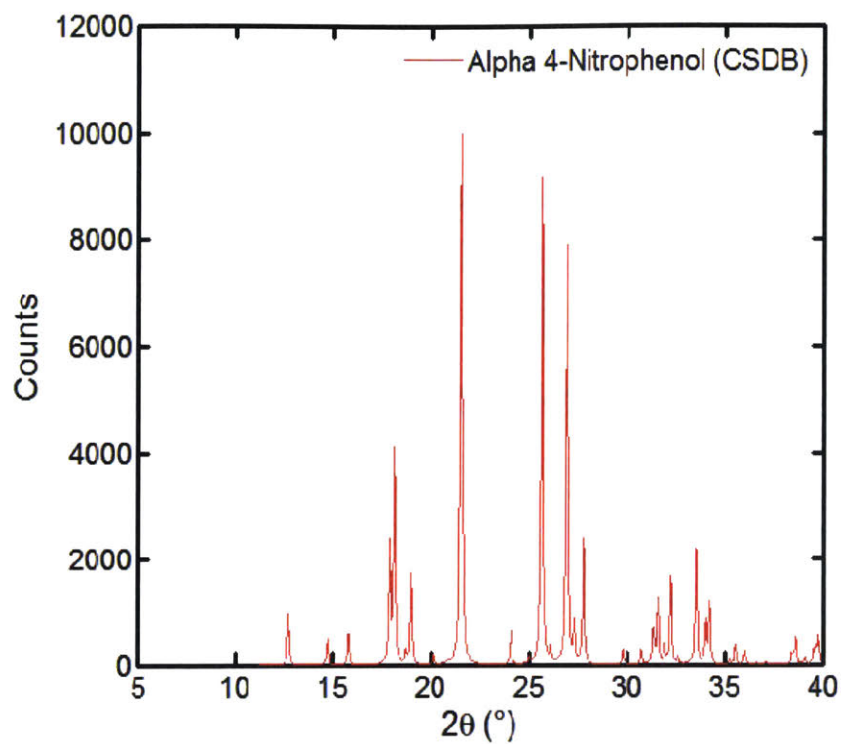
**Figure 2-10:** Theoretical PXRD for the alpha polymorph of 4NP, obtained from the Cambridge Structural Database.
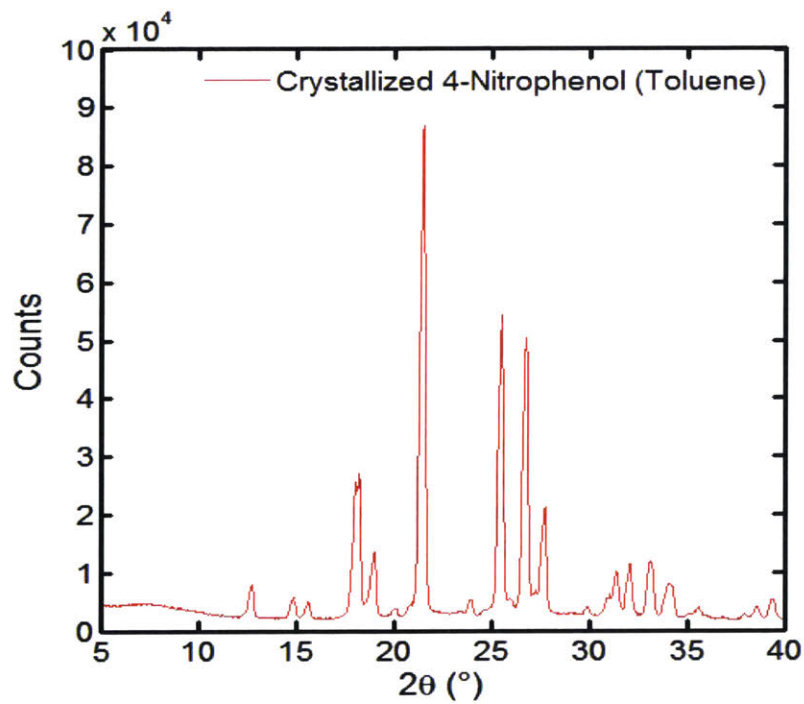


**Figure 2-11:** Experimentally obtained PXRD of 4NP, taken from a sample crystallized in toluene with no complexing agent. The graph matches with that of the alpha form in Figure 2-10.

## Crystallization Data

The individual results for the crystallization experiments are shown below. The data shown corresponds to each of the batches performed in triplicate following the procedure outlined in the main text of the chapter.

**Table 2-6:** Experimental data for the purity of 4NP after crystallization in water. The results are presented in terms of percent weight of 3NP remaining in the solid product.

| Experiment | Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 | Exp. 5 | Exp. 6 | Exp. 7 | Exp. 8 | Exp. 9 | Exp. 10 | Exp. 11 | Exp. 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Control | 3.04 | 3.07 | 3.03 | 2.64 | 2.88 | 2.84 | 3.19 | 2.49 | 2.92 | 2.33 | 2.73 | 2.63 |

| Experiment | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 | Sample 6 |
|---|---|---|---|---|---|---|
| Second Crystallization | 0.21 | 0.19 | 0.17 | N/A | N/A | N/A |
| β-Alanine | 2.32 | 2.47 | 2.49 | N/A | N/A | N/A |
| Guanidine Acetic Acid | 1.40 | 1.47 | 1.44 | 2 | N/A | N/A |
| γ-Aminobutyric Acid | 1.18 | 2.19 | 2.13 | N/A | N/A | N/A |
| Taurine | 2.07 | 1.93 | 1.99 | N/A | N/A | N/A |
| 3-Aminobenzoic Acid | 0.52 | 0.53 | 0.51 | 0.41 | 0.37 | N/A |
| Succinamic Acid | 1.46 | 0.86 | 1.33 | 1.35 | 0.4 | 1.05 |

| Experiment | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 | Sample 6 |
|---|---|---|---|---|---|---|
| 3ABA 3.9 mg/ml | 1.93 | 2.03 | 2.07 | N/A | N/A | N/A |
| 3ABA 5.9 mg/ml | 0.68 | 0.61 | 0.24 | 0.52 | 0.49 | 0.62 |
| 3ABA 7.9 mg/ml | 0.52 | 0.53 | 0.51 | 0.41 | 0.37 | N/A |
| 3ABA 9.9 mg/ml | 0.28 | 0.34 | 0.28 | N/A | N/A | N/A |

**Table 2-7:** Experimental data for the yield of 4NP after crystallization in water. The results are presented in terms of percent weight of 4NP recovered with respect to the original amount in the batch.

| Experiment | Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 | Exp. 5 | Exp. 6 | Exp. 7 | Exp. 8 | Exp. 9 | Exp. 10 | Exp. 11 | Exp. 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Control | 92.0 | 91.6 | 91.8 | 90.0 | 90.3 | 89.8 | 90.2 | 91.2 | 90.8 | 90.1 | 90.0 | 89.9 |

| Experiment | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 | Sample 6 |
|---|---|---|---|---|---|---|
| Second Crystallization | 82.6 | 82.1 | 82.3 | N/A | N/A | N/A |
| β-Alanine | 90.5 | 90.8 | 91.3 | N/A | N/A | N/A |
| Guanidine Acetic Acid | 91.4 | 90.9 | 90.4 | 91.4 | N/A | N/A |
| γ-Aminobutyric Acid | 91.4 | 91.4 | 89.4 | N/A | N/A | N/A |
| Taurine | 91.4 | 91.1 | 91.3 | N/A | N/A | N/A |
| 3-Aminobenzoic Acid | 89.3 | 89.0 | 88.8 | 89.8 | 89.1 | N/A |
| Succinamic Acid | 91.5 | 91.5 | 91.0 | 90.0 | 90.1 | 89.7 |

| Experiment | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 | Sample 6 |
|---|---|---|---|---|---|---|
| 3ABA 3.9 mg/ml | 90.1 | 90.1 | 89.9 | N/A | N/A | N/A |
| 3ABA 5.9 mg/ml | 91.2 | 89.4 | 91.5 | 89.6 | 90.3 | 89.4 |
| 3ABA 7.9 mg/ml | 89.3 | 89.0 | 88.8 | 89.8 | 89.1 | N/A |
| 3ABA 9.9 mg/ml | 89.9 | 91.5 | 90.4 | N/A | N/A | N/A |

## Membrane Experiments

The curves shown in Figure 2-12 and Figure 2-13 show the change in concentration over time of the retentate and permeate in the membrane experiments. The values reported in the chapter are taken as the average of the rejection coefficient of the last three data points reported.
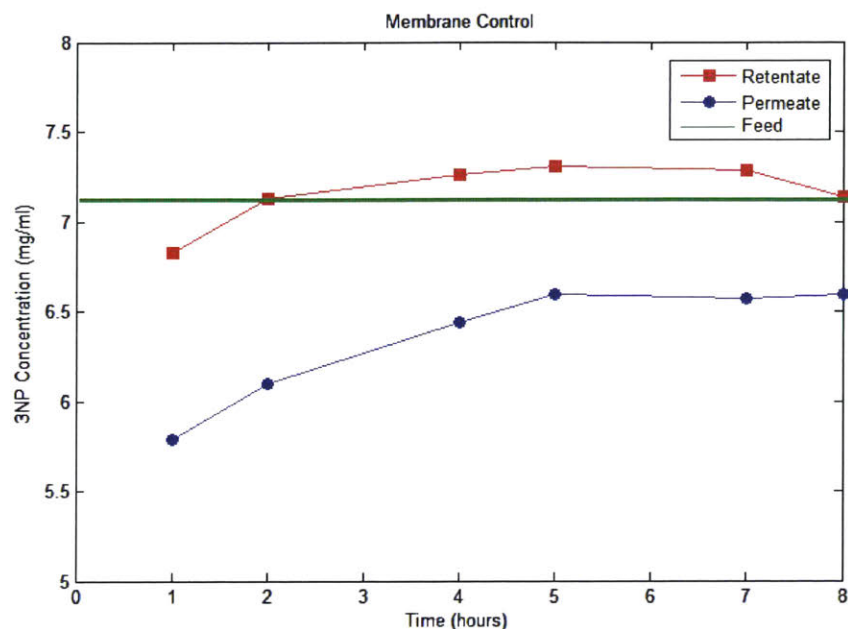


**Figure 2-12:** Evolution of the retentate and permeate concentrations over time in the control membrane experiment (no complexing agent used).
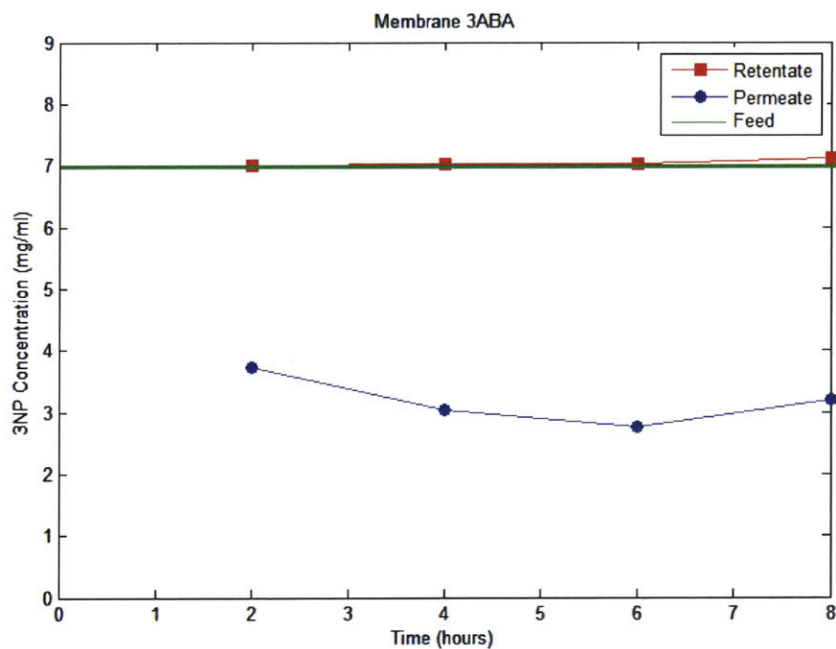


**Figure 2-13:** Evolution of the retentate and permeate concentrations over time in the 3ABA+3NP membrane experiment.

## TGA Testing of Nitrophenols

TGA was used to optimize the drying conditions of the samples in the crystallization experiments. TGA was performed in order to compare the evaporation and sublimation profiles of the nitrophenols. This was done because it was observed that upon long exposure at medium temperatures (greater than 40 at °C) in a vacuum oven (25 inHg of vacuum) the composition of the crystallized solids would change, with a decrease in the relative amount of 3NP. The TGA curves are shown in Figure 2-14. The data in Figure 2-14 shows that 3NP is more easily sublimated than 4NP. This result is expected given that 3NP has a lower boiling point than 4NP and it is expected to have a higher vapour pressure. This is evidenced by the fact that a substantially larger mass fraction of 3NP is lost upon reaching 100 °C. Subsequently, 3NP was completely removed at almost 25 °C lower temperature than 4NP. The heating profiles used for both compounds were identical and follow the program described in the experimental section. As a consequence of these results, the crystallization samples of the purification experiments were dried using a vacuum oven held at 25 °C for 12 hours. Under these conditions no changes in relative nitrophenol concentration were observed.
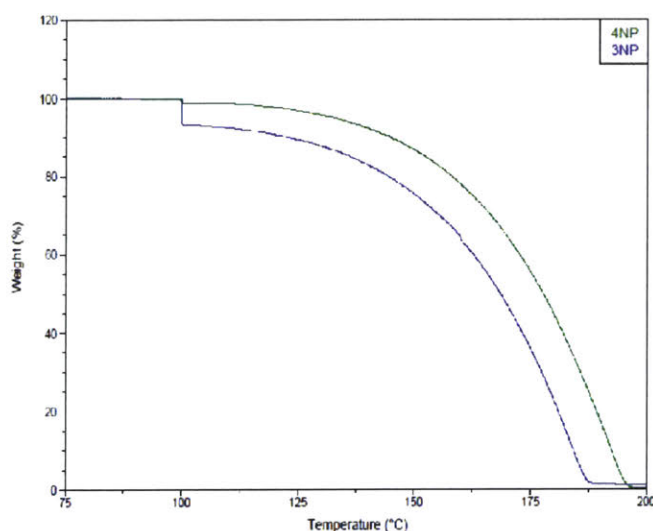


**Figure 2-14:** TGA curve of 3NP and 4NP. The curves are overlapped to show the differences in mass loss between the two compounds.

# Chapter 3 : Purification of a complex and industrially relevant API

The contents of this chapter represent unpublished original work. The findings of this project were deemed to be inadequate for publication in a peer-reviewed journal; however, the work shown herein provides key insights on the implementation and limitations of complex-assisted crystallization as a viable technique for pharmaceutical crystallization.

## 3.1 Introduction

Pharmaceutical molecules are becoming increasingly complex, even for small-molecule drugs[78,79]. This indicates that the ability to separate structurally complex molecules using complex-assisted crystallization must be studied and demonstrated in order to elucidate the viability and usefulness of the technique within the context of applied pharmaceutical processes. The results shown in chapter 2, as well as the findings of previous research[65–68,70,80] have demonstrated that the use of complexation in solution can lead to massive purity improvements without compromising yields on crystallization model systems. Despite the progress achieved, the work performed has been limited to relatively small impurity molecules (MW < 200 g/mol) [65–68,70,80] with simple chemistry and functional groups that can be easily targeted for complexation and intermolecular interactions[54].

Past work has demonstrated that small, structurally similar compounds with a differing functional group can be easily separated using complex-assisted crystallization[65,66]. The research performed for the purification of nitrophenols[70] elucidated the applicability of complex-assisted crystallization for the separation of structural isomers and allowed expanding the scope of the technique. Finally, work performed by *Hsi et al.* demonstrated that simple API small-molecules can be purified using complexation. However, the successful implementation of complex-assisted crystallization on a modern API has not been demonstrated.

Novartis International AG has been the primary contributor and funder of the research summarized within this thesis, under the supervision of the Novartis-MIT Center for Continuous Manufacturing (NMCCM). Based on recommendations from the NMCCM, a drug recently approved by the Food and Drug Administration (FDA) and European Medicines Agency (EMA) was selected as a candidate for purification using complex-assisted crystallization. The selected molecule will be referred to as "Compound C" throughout this chapter. This API is significantly larger than any compound previously studied with complex-assisted crystallization (MW> 300 g/mol). Furthermore, the structural difference between the purification target and the impurity is completely different to any system previously studied using this technique: the impurity is lacking a functional group present within the API. Figure 3-1 shows an abbreviated version of the chemical structures of the two molecules, the non-specific structure is provided to preserve confidentiality of the structure, as requested by Novartis.

Figure 3-1, shows that Compound C contains an oxy-R group (shown in the right-side of the image) that is missing in the impurity. The central section of both molecules contains the nitrogen-rich imidamide functionality, which experiences an intramolecular interaction with the sulfonyl group to the left, and (in the case of the target API) with the oxy-R group to the right. The presence of such complex functional groups as well as multiple intramolecular hydrogen bonding represents a heightened challenge for complex-assisted crystallization. In addition to the chemical complexity, the purification target for this project, as requested by Novartis, was to reduce the impurity concentration after a single batch crystallization experiment below 500 ppm, from a starting point of 1 %wt. This purification goal far exceeds what has been accomplished in past complex-assisted

crystallization studies[65–68,70,80], which typically explored final impurity concentrations of 0.1-0.5 %wt of the final product. The heightened process challenge brought about by the chemical complexity of Compound C provides an excellent test for complex-assisted crystallization. This project was focused on defining the feasible performance window of the technique and characterizing the limits of purity and chemical complexity that can tackled when attempting to improve crystallization processes with it.
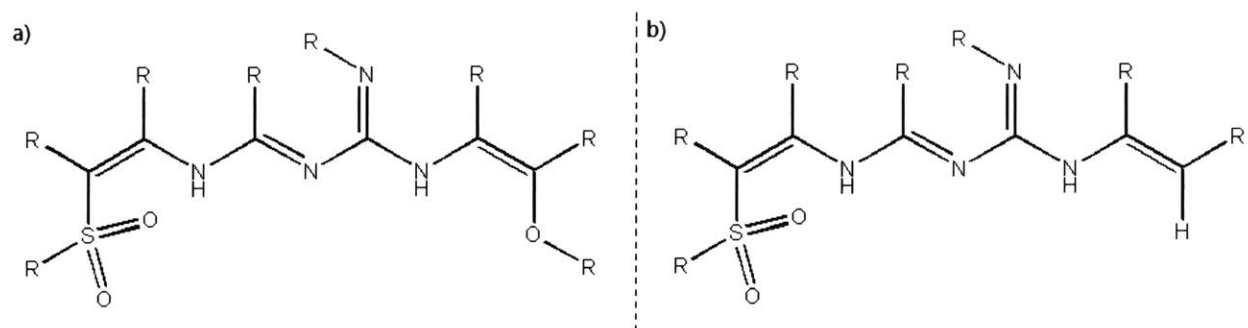


**Figure 3-1:** Chemical structures of the inhibition target, "Compound C" (a), and its impurity (b). The connectivity and functionality of the drug has been simplified as much as possible to keep the confidentiality of the molecule. Sections labeled "***R***" represent any possible functional group.

A complexation motif was proposed to establish selectivity of interaction between an additive and the impurity. The motif focused on exploiting the steric hindrance provided by the extra oxy-R group present within Compound C, and on the presence of a stabilizing intramolecular interaction between the oxy-R group and the imidamide, which is not observed in the impurity. Experiments were carried out using a cooling and antisolvent batch procedure recommended by Novartis, focused on maximizing yield. Batch crystallizations were carried out with and without complexing agents to compare the effect of the presence of the additives on purity and yield of the final product. The selection of complexing agents was done using a mixture of preliminary computational analysis, compounded with direct experimental verification. Membrane separation experiments were done with the feed solution with and without complexing agent to determine the selectivity and effectiveness of the best-performing complexation targets.

## 3.2 Experimental

### Materials

Compound C and its impurity were provided by Novartis International AG and used as received. Tetrahydrofuran (THF), HPLC water, HPLC methanol, and trifluoroacetic acid were purchased from Sigma Aldrich. Ethanol (200 proof) was purchased from VWR International (Edison, NJ). All complexing agents (shown in Figure 3-4) were purchased from either Sigma Aldrich or Alfa Aesar, and used as received.

### High-Performance Liquid Chromatography

For nitrophenol and 3ABA detection, the HPLC instrument (Agilent 1260) was equipped with a UV-diode array detector (DAD). The column used was a YMC-Pack ODS-A 150 mm X 4.6 mm i.d. packed with 3 μm particles with 12 nm pore size (YMC America Inc.). Three detection wavelengths were used for all samples, 275 nm, 240 nm, and 230 nm. For each chemical, the wavelength that provided the highest intensity and sensitivity was used as the calibration signal. The samples were analyzed using an isocratic method with a 55:45 (volume) water/methanol

mobile phase containing 0.3 %vol. trifluoroacetic acid, with 10 μl sample injections, a 1 ml/min solvent flow rate, and a column temperature of 30 °C.

## Membrane Separations

Membrane experiments were using the same setup discussed in Chapter 2. The system is designed as a cross-flow filter: the feed and retentate respectively enter and exit the filter chamber parallel to the membrane surface, whilst the permeate flows perpendicularly across the membrane. The retentate chamber is large (150 ml), allowing for the inclusion of a magnetic stir bar. The feed is pumped at a high pressure (420-450 psig) using a Thermo Scientific Dionex AXP HPLC pump. The system temperature was controlled at 40 °C using a heating plate, which also allowed for a controlled stirring rate of 400 rpm. The retentate and permeate outlets both have Swagelok® pressure valves, which allowed pressurization of the filter chamber to the desired operating condition. The permeate valve was kept with no resistance to allow for maximum flow across the membrane. The selected operating flowrate was controlled with the pump and was set to 2.5 ml/min, with the actual flowrate measured to be 2.65 ml/min.

The performance metric monitored for the membrane experiments was the rejection factor (RR), as defined by Ferguson et al[73]. The RR is a measure of the fraction of solute that is unable to pass through the membrane. This metric is calculated using equation 3-1, where $C_p$ is the concentration of solute (impurity) in the permeate stream, and $C_f$ is the concentration of the solute in the feed stream. The concentrations of solute were collected in aliquots and monitored using HPLC. The reported values of RR were determined once the system reached a steady state, which was observed to occur after approximately 6 hours of continuous operation.

$$RR = 1 - \frac{C_p}{C_F} \qquad \text{3-1}$$

The membrane used, 17-PBI-DPX, is a polybenzimidazole (PBI) polymer cross-linked with α,α'-dibromo-p-xyelen (DPX) and was prepared following the procedure outlined by Kim et al[74]. The "17" indicates the weight percent of polymer in the membrane dope composition used during synthesis[73–75]. This membrane was selected based on recommendations from the supplier for having the lowest molecular weight cut-off (MWCO). The MWCO is defined as the molecular weight at which 90% of the solute is retained by the membrane. The MWCO of 17-PBI-DPX is >600 g/mole.

The membrane experiments were carried out both as an additional verification of the quality and selectivity of the best-performing complexing agents, as well as a tool to potentially increase the effectiveness of complex-assisted crystallization. The use of membrane-separations as a way to enhance continuous crystallizations that relied on complex-assisted crystallization has been successfully demonstrated by Vartak et al[80].

## Complexing Agent Selection

Potential complexing agents were first screened using a computational technique analogous to the one described in Chapter 2, and by other researchers[70,81]. Molecules were built in Avogadro and had their energy minimized using the MMFF94s forcefield[82–84]. Energy minimizations were carried out for Compound C, the impurity, and all potential complexing agents individually. The energy calculations were also performed for each complexing agent interacting with either Compound C or its impurity. Different conformations and starting positions for all molecules were tested each time in order to find the lowest-energy configuration. Once the calculations were

43

performed, the energy of interaction between the complexing agent and the Novartis molecules were estimated using the correlation shown in equation 3-2.

$$\Delta E_{Cmplx} = \Delta E_{int} - \Delta E_{seprt} \qquad \qquad 3\text{-}2$$

In equation 3-2, $\Delta E_{Cmplx}$ is the estimated net energy of interaction between the two molecules, $\Delta E_{int}$ is the minimized calculated energy from minimizing two molecules in proximity with each other, and $\Delta E_{seprt}$ is the energy of the two molecules infinitely far apart from each other (non-interacting). The value of $\Delta E_{Cmplx}$ was determined for a given potential complexing agent with Compound C and its impurity. If a candidate additive had a high $\Delta E_{Cmplx}$, that was stronger for the interaction with the impurity rather than Compound C, then the computational estimate was further refined using Material Studio 6.1 and the COMPASS forcefield. The overall procedure of this calculation was completely analogous to the preliminary one performed with Avogadro. The stronger value of $\Delta E_{Cmplx}$ with the impurity was desired so that the complexing agent would have selectivity towards interacting with it, rather than Compound C.

Once a promising complexing agent candidate was selected, it was directly tested using crystallization experiments in order to determine its effectiveness in improving the purity of the final product. Multiple experiments were carried out with each of the candidates to ensure reproducibility. Additionally, different concentrations of a complexing agent were tested to determine if increasing the amount of additive led to improvements in crystallization performance. Finally, the best performing complexing agents were further tested using membrane experiments (as described above) to determine their selectivity and further optimize their performance.

**Crystallization Experiments**

The procedure for the crystallization of Compound C was a bath with simultaneous cooling and antisolvent addition to induce supersaturation and maximize yields. The solvents, temperatures and times were selected based on recommendations from Novartis. The starting solvent was a 92:8 (v/v) ethanol:THF mixture. The feed solution consisted of 7 %wt of Compound C dissolved in the ethanol:THF solvent. The "crude" of Compound C dissolved in the feed contained 1 %wt of the impurity. This starting recipe was stirred at 250 rpm and heated to 55 °C until complete dissolution was achieved. Complexing agent was also included in this initial solution, the exact amount of which varied across experiments.

Once the starting solution became completely homogeneous, it was slowly cooled to 40 °C. Subsequently, a small injection (5 %wt of the original solid mass) of Compound C seeds were suspended in a 50:50 (v/v) water:ethanol solution and added to the batch. The crystallizer was then allowed to remain on these conditions for 30 min to allow for equilibration and for turbidity (stemming from the seeds) to develop. After the allotted time, water was added as an antisolvent along a period of 90 minutes. The volume of water chosen was such that the final solvent composition was ethanol+THF/Water equal to 60/50 (v/v).

The final suspension of crystals was cooled to 20 °C and allowed to stir and equilibrate overnight. The final crystal products were recovered using vacuum-driven filtration over filter paper. The resulting wet-cake was three-times dried using a 3:1 (v/v) water/ethanol mixture at room temperature. Finally, the crystal products were dried in a vacuum oven at 50 °C overnight. The product was analyzed using HPLC to determine purity, and the mother liquor was tested as well to determine the process overall yield.

44

## 3.3 Results and Discussion

The first set of batch experiments were carried out at control conditions (no complexing agent added) to generate a proper reference baseline of the natural behavior of the crystallization of Compound C. The results for these trials are shown in Table 3-1, the data shows that the control conditions produce outstanding consistency and reproducibility both in purity and yield. The average impurity concentration at the control conditions after a crystallization experiment was measured to be 0.35 ± 0.01 %wt. The yield show similar reliability, averaging 97.3 ± 0.2 %wt. All subsequent crystallization experiments and purification attempts will be compared against these baseline numbers in order to determine whether an additive was successful or not. The objective of this project was reducing the impurity concentration to 0.05% or less (500 ppm) without reducing the yield.

**Table 3-1:** Results for control experiments carried out with Compound C under the batch conditions described in the experimental section. It is clear that excellent reproducibility for both purity and yield were achieved.

| Sample | Final Impurity (%wt) | Yield (%wt) |
|---|---|---|
| Control 1% Exp. 1 | 0.34 | 97.5 |
| Control 1% Exp. 2 | 0.36 | 97.1 |
| Control 1% Exp. 3 | 0.35 | 97.3 |
| | | |
| Average | 0.35 | 97.3 |

In addition to finding the average results for the control set of batch conditions, it was necessary to establish a correlation between the natural yield and impurity concentration in the final product. The reason for this lies on the fact that as yield is raised, so does the amount of impurity that incorporates into the crystal[2]. The relationship between yield and impurity incorporation is a necessary tool for determining the true effectiveness of a complexing agent, given that some additives may affect yield, therefore impacting purity. It is necessary to de-convolute purification effects with yield effects stemming from an additive, which requires understanding of the expected purity at a given final yield. To establish the correlation between yield and purity, the amount of water antisolvent added on a series of control batch experiments was varied. Generally speaking, as the amount of water added is increased, so does the final yield, given that the solubility of Compound C in water is negligible. The results for this set of experiments are shown in Figure 3-2. The data shows an excellent linear correlation of increasing impurity content with increasing yield, as expected[2]. The correlation for the line shown in Figure 3-2 is ppm = 122.56*(yield %) − 8436.6. Using this linear expression, the final purity and yield generated in an experiment with a given complexing agent can be properly analyzed against a normalized baseline. When an additive yields a purity and yield below the line in Figure 3-2, then it has led to an overall improvement of the crystallization process. Conversely, when an additive has a purity and yield that lie above the line, then its overall effect on the crystallization is detrimental, even if the purity was improved with respect to the controls. This is true, because it would imply that such an additive would lead to yield penalties that overcompensate for the purity gains.
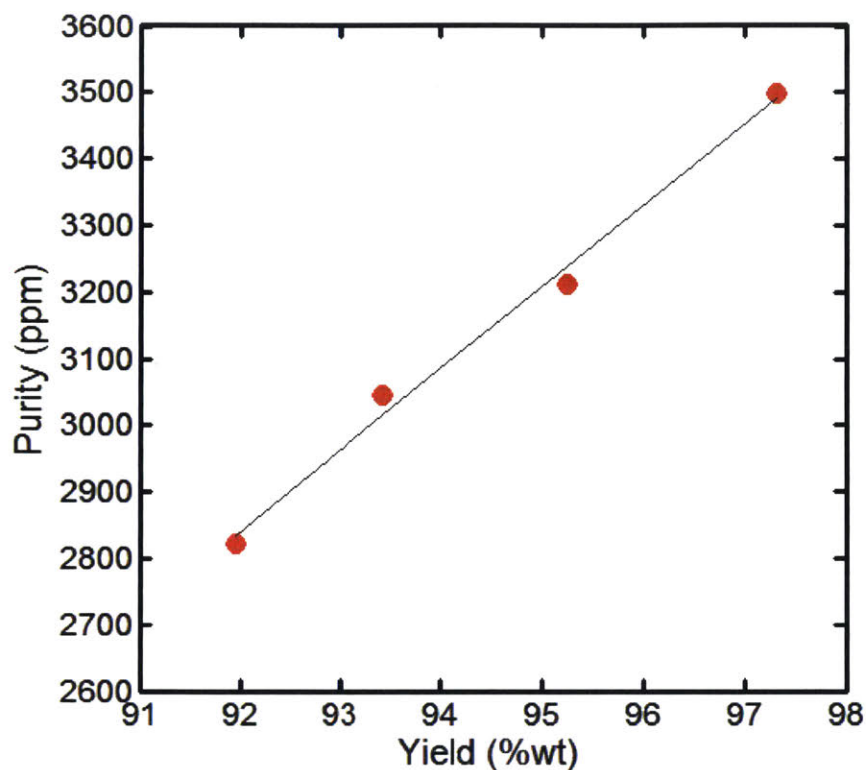
**Figure 3-2:** Correlation between yield of a batch experiment and the final product purity. It is clear that the impurity within the final product increases linearly as a function of yield. The fit has an $R^2$ of 0.997.

Once the control set of conditions and the expected performance of the batch experiments were characterized, it was necessary to select and test complexation candidates that could improve the performance of the crystallization of Compound C, truly enhancing purity. Based on the chemical structures of Compound C and its impurity, it was necessary to device a complexation motif that would promote selectivity of interaction towards the impurity. The major structural elements that helped define the desired motif are illustrated in Figure 3-3. Compound C contains a pair of strong intramolecular hydrogen bonds, one between its sulfonyl group and its imidamide, and another between the oxy-R group and the imidamide. The presence of intramolecular bonding hinders the ability of the molecule to have strong intermolecular interactions, given that the intramolecular bonds are easier to attain, and they stabilize the energy of the chemical with relatively small entropic penalties[54–56]. Additionally, the presence of the oxy-R group generates a steric hindrance for other molecules that could interact with the imidamide core of Compound C. As a result, it was hypothesized that selectivity of interaction towards the impurity could be achieved if these structural limitations of the API were exploited. The proposed model complexing agent would have a series of hydrogen-bond donors and acceptors that would target and complement the nitrogen groups within the imidamide. This bonding portion of the molecule, labeled as "I" in Figure 3-3 would be attached to a large backbone structure, which would hinder the ability of the additive to properly orient itself to interact with Compound C, given its additional steric hindrance from the oxy-R group. A large set of hypothetical structures were explored using the computational strategy outlined in the "Experimental" section of this chapter, a full list of the estimated energies obtained from Material Studio and Avogradro are shown in the Appendix. The full list of complexing agents that were deemed appropriate for experimental verification, as well as their chemical structures, are shown in Figure 3-4.
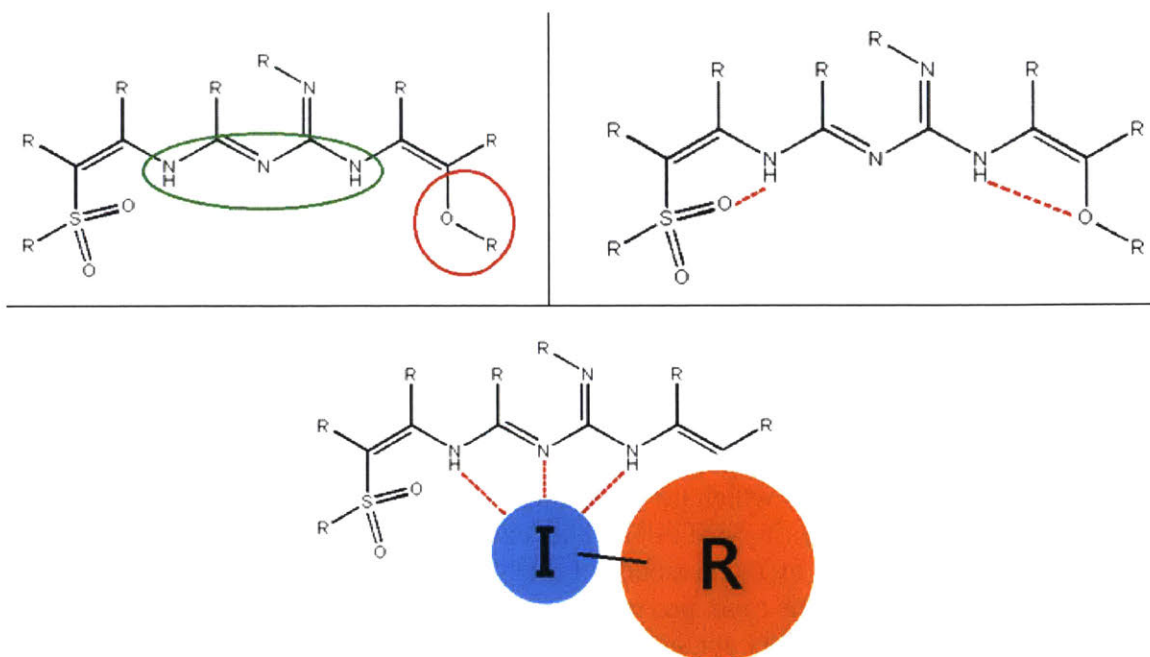
**Figure 3-3:** Chemistry of Compound C and complexation motif proposed to induce selectivity with an additive. Top left: imidamide (green) and oxy-R groups, which are key to the chemical interaction; top right: intramolecular hydrogen bonds present within Compound C, the bond with the oxy-R group is missing in the impurity; bottom: proposed interaction motif, in which the complexing agent will interact with the nitrogen-rich imidamie group, and will simultaneously have a bulky backbone that would be sterically hindered by the presence of the oxy-R group of Compound C.
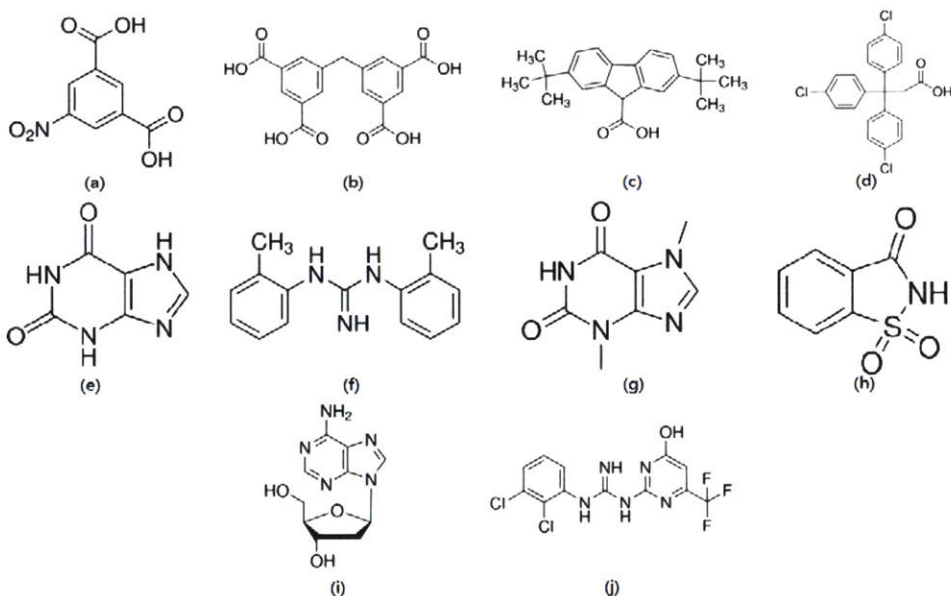


**Figure 3-4:** List of complexing agents tested to enhance the crystallization of Compound C. a) 5-nitroisophthalic acid (5NISA) b) 3,3',5,5'-tetracarboxydiphenylmethane (TCBDP) c) 2,7-di-tert-butylfluorene-9-carboxylic acid (27TBCA) d) 3,3,3-tris(4-chlorophenyl)propionic acid (TrisCPA) e) Xanthine (XAN) f) 1,3-di-o-tolylguanidine (DOTG) g) Saccharin (SACC) h) Theobromine (TBR) i) 2-deoxyadenosine monohydrate (2DAMH) j) N-(2,3-dichlorophenyl)-N'-[4-Hydroxy-6-(Trifluoromethyl)-2-Pyrimidinyl]guanidine (NDTFPG)

The structures shown in Figure 3-4 demonstrate that the selected complexing agents all follow the proposed motif of hydrogen bond donor groups, attached to a relatively bulky backbone, such as a benzene ring. There are three main general type of molecules that were selected as potential additives: carboxylic acids (5NISA, TCDBP, 27TBCA, and TrisCPA), xanthines (XAN, SACC, TBR, and 2DAMH), and guanidines (DOTG and NDTFPG). Crystallization experiments were carried out with all the chemicals shown in Figure 3-4, the results of these batches are listed in Table 3-2. For all the experiments listed, the setup followed the exact batch procedure described in the Experimental section and used an amount of water that would be expected to have a yield of 97.3 %wt under control conditions. The second column of Table 3-2, "Complex:Impurity" shows the relative molar quantity of complexing agent used, as a factor of the initial concentration of impurity present within the batch recipe (1 % wt of the Compound C mass used). Some complexing agents were tested at higher ratios to try to force the equilibrium towards the formation of a complex and favor the purification of the product crystals. Two data points worth mentioning are the experiment with TBR, for which no yield could be recovered due to an experimental loss, and the "Extra Water" test with NDTFPG. The extra water experiment was performed to recuperate yield losses incurred by the presence of the NDTFPG complexing agent. This was done because this additive provided the most promising results among the molecules tested in terms of purity gains. Nevertheless NDTFPG consistently caused yield losses. The "Extra Water" test was used to see if yield could be recovered while still maintaining purity gains.

**Table 3-2:** Results for crystallization experiments with Compound C and multiple complexing agents. The initial impurity concentration for all experiments was 1 %wt. See Figure 3-4 for a full list of all the additive names. The "Complex:Impurity" column refers to the molar ratio of complexing agent to impurity within the batch recipe. The "Expected ppm" come from the correlation between purity and yield shown in Figure 3-2. The "Diff. (ppm)" shows the difference between the expected and experimental purity: negative values are underperformers, while positive values are successes.

| Experiment | Complex:Impurity | Impurity (ppm) | Yield (%) | Expected ppm | Diff. (ppm) |
|---|---|---|---|---|---|
| Control | N/A | 3500 | 97.3 | 3388 | -113 |
| 5NISA | 1:1 | 3400 | 95.4 | 3156 | -244 |
| TCBDP | 1:1 | 3100 | 93.6 | 2935 | -165 |
| 27TBCA | 4:1 | 3300 | 96.3 | 3266 | -34 |
| 27TBCA | 10:1 | 3300 | 96.2 | 3254 | -46 |
| TrisCPA | 4:1 | 2800 | 93.1 | 2874 | 74 |
| TrisCPA | 10:1 | 5400 | 96.2 | 3254 | -2146 |
| XAN | 1:1 | 3900 | 95.7 | 3192 | -708 |
| SACC | 1:1 | 3000 | 96.1 | 3241 | 241 |
| TBR | 1:1 | 3400 | N/A | N/A | N/A |
| DOTG | 1:1 | 3300 | 97.1 | 3364 | 64 |
| DOTG | 2:1 | 3400 | 97.0 | 3352 | -48. |
| 2DAMH | 2:1 | 3200 | 96.1 | 3241 | 41 |
| 2DAMH | 10:1 | 3200 | 96.8 | 3327 | 127 |
| 2DAMH | 21:1 | 3000 | 95.5 | 3168 | 168 |
| 2DAMH | 40:1 | 3000 | 96.4 | 3278 | 278 |
| NDTFPG | 1:1 | 3100 | 94.5 | 3045 | -55 |
| NDTFPG | 5:1 | 2600 | 93.3 | 2898 | 298 |
| NDTFPG | 5:1 + Extra Water | 2800 | 96.3 | 3266 | 466 |
| NDTFPG | 10:1 | 2600 | 94.2 | 3009 | 409 |

It is clear from the data shown in Table 3-2 that all complexing agents had some degree of effect on the crystallization experiments. However, the relative significance of such effects was small, with most purities remaining around 3000 ppm (0.30 %wt) of impurity and yields of 95-97 %wt. If we use the correlation from Figure 3-2 between purity and yield, and compare what is the expected purity of each experiment based on the yield observed, it is possible to quantify the degree of "success" of each individual complexation experiment. This analysis is shown in the last two columns of Table 3-2. The experiments highlighted in red indicate samples where the complexing agent represented a detriment for the batch performance, either by decreasing purity or yield (or both). Conversely, the rows highlighted in green correspond to successful experiments, were the complexing agents led to net improvements with respect to the baseline conditions. The best-performing complexing agents were SACC, 2DAMH, and NDTFPG; however, the degree of improvement ("Diff. (ppm)") for these samples was never than 14.3 % of an improvement in purity with respect to the expected value given the measured yield. The most successful experiment was the one performed with 5:1 NDTFPG with an extra amount of water to push the yield upward. This experiment only represented a net gain of approximately 450 ppm, out of the 3250 ppm total of impurity that are anticipated in a control experiment of equivalent yield. For reference, past complex-assisted crystallization research[66,70] achieved improvements of 80-90 % with respect to controls. Additionally, given the required purification target of only 500 ppm of impurity remaining, an improvement of 86 % with respect to the control would needed, at a yield of 97.3 %wt to achieve the purification goal set by Novartis for this project. This means that in order to achieve the target, a complexing agent almost 6 times most effective than the best-performer observed would be required.

The best-performing complexing agents were analyzed using membrane separation experiments analogous to the ones outlined in Chapter 2 and performed in past work within the Myerson group[70,80]. The objective of these experiments was to determine the selectivity and effectiveness of complexation of the additives that led to the most improvement in the batch crystallization experiments. Additionally, it was hypothesized that upon a successful separation using size-based membrane separation, it would be possible to design a continuous-crystallization procedure with membrane recycling, as shown by *Vartak et al.*[80] to maximize the purification and yield achievable. The membrane experiments were performed during a period of 6 hours with regular monitoring of the output concentrations of the permeate and retentate streams in order to determine the proximity to steady state and the effectiveness of separation across the membrane. The overall expectation was that the complexing agent would interact with the impurity, preventing it from going across the membrane. As a result, the permeate stream would have a low concentration of impurity, while the retentate would become more concentrated. Simultaneously, Compound C's concentration across the membrane should remain mostly unchanged, given that its molecular weight should be below the MWCO of the selected membrane, assuming it does not interact with the complexing agents in solution. Finally, the behavior of the complexing agent would be expected to follow an almost identical behavior to the impurity.

The first set of results for membrane experiments reported are illustrated in Figure 3-5. These experiments were carried out with an excess of Compound C to impurity to mimic relative solution conditions representative of the batch crystallizations. Similarly, the complexing agent was placed on a small molar excess to the impurity. The exact molar ratios will remain undisclosed for confidentiality purposes.
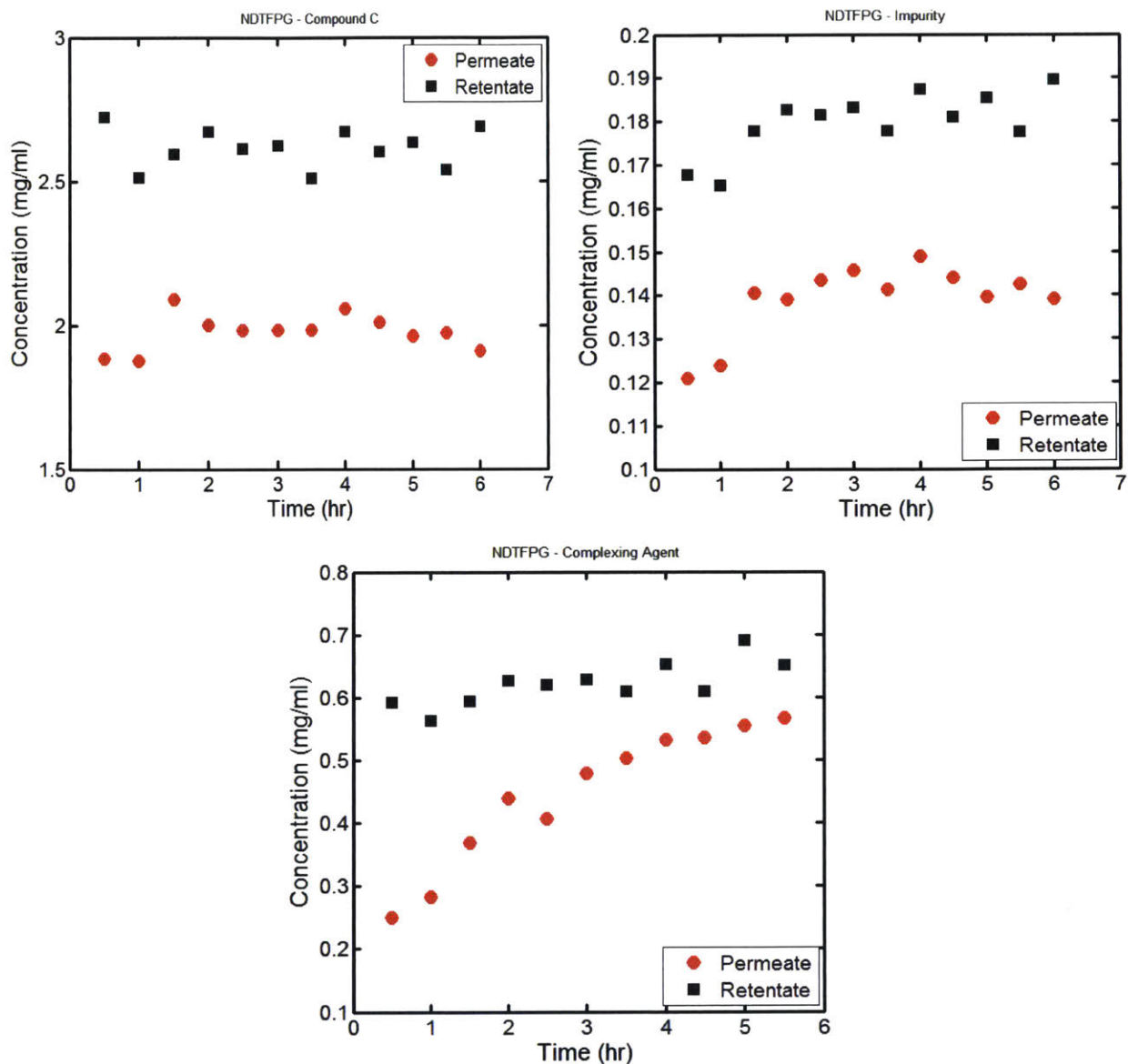
**Figure 3-5:** Membrane experiments performed with NDTFPG. All images show the concentration vs. time behavior of the chemical species for the permeate and retentate streams. Top left: curve for Compound C. Top right: curve for the impurity. Bottom: curve for NDTFPG.

The data shown in Figure 3-5 demonstrates that both Compound C and the impurity had similar behavior across the membrane, with a slightly more concentrated retentate stream. Meanwhile, NDTFPG never truly reached steady state throughout the entire 6 hour run; however, the concentration behavior over time demonstrates that NDTFPG concentration was completely equalizing across the membrane, indicating that it was not being rejected based on its size. The lack of rejection of the complexing agent, coupled with the fact that the RR for both Compound C and the impurity (15.3% and 13.8% respectively) are virtually identical, indicates that there is no strong complexation effect taking place within the solution as it is passed through the membrane system. This result demonstrated that the complex being formed between NDTFPG and either Compound C or its impurity, it is either extremely weak, or completely non-existent. Based on these results, it was clear that no further optimization could be performed with NDTFPG.

50

The second set of membrane experiments were carried out with 2DAMH, the second-best performing complexing agent during the batch crystallization trials. The overall setup was extremely similar to the one applied for NDTFPG, with significant excess of both complexing agent and Compound C. The results from the membrane work with this chemical system are shown in Figure 3-6. The results from this work show a different behavior to the one observed for NDTFPG. For starters, the rejection across the membrane for both Compound C and its impurity were substantially larger than those observed for NDTFPG, indicating that a complex was actually being formed in this instance. This observation was confirmed by the fact that 2DAMH did see a significant difference in concentration between the retentate and permeate streams, indicating that there was a size-based rejection, which was only possible if an extremely large (>600 g/mol) molecular complex was being formed. The RR for 2DAMH was 26 %, while the RR of Compound C was 34 % and for the impurity it was 21 %.
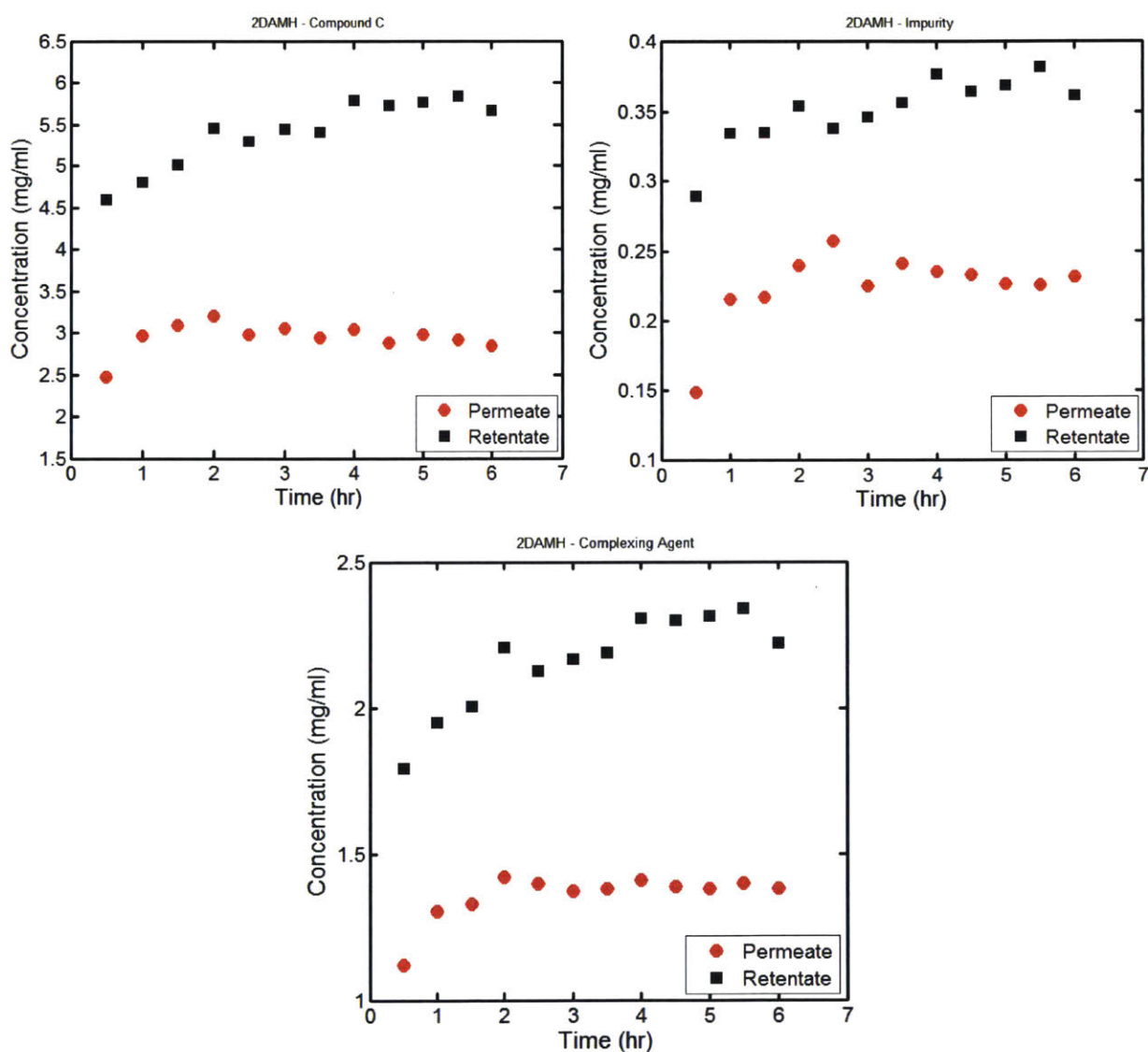


**Figure 3-6:** Membrane experiments performed with 2DAMH. All images show the concentration vs. time behavior of the chemical species for the permeate and retentate streams. Top left: curve for Compound C. Top right: curve for the impurity. Bottom: curve for 2DAMH.

The values observed for RR show an extremely important trend: the rejection for Compound C was greater than for its impurity. This result implies that 2DAMH did not have enough selectivity of interaction towards the impurity molecule, and it was actually interacting with Compound C itself. The insufficient selectivity was an unexpected result for a couple of reasons; first, the batch experiments demonstrated that 2DAMH does have an overall positive effect in the purification of Compund C; second, the computational analysis (see Apendix) had indicated that the energy of interaction between 2DAMH and the impurity should be substantially stronger than that between the additive and Compound C. Based on the data illustrated in Figure 3-6, it became clear that it is not possible to achieve further purification of the target product using a crystallization with a membrane recycle setup, at least with the experimental conditions employed, and the chemicals tested experimentally in this work. For this reason, no further attempts were made to optimize and characterize the membrane experiments.

The fact that the membrane experiments demonstrated that 2DAMH had a stronger interaction with Compound C than with the impurity opens up a relevant point that must be discussed, related to the thermodynamic limitations of the purification of Compound C. This discussion is especially relevant when the fact that the concentrations of the impurity in all of the experimental solutions is 1% of less that of Compound C. The thermodynamic analysis can be done from the perspective of chemical equilibrium and molecular speciation in solution. This kind of analysis has been extensively implemented and is widely accepted in analysis of solution complexation[20,65,85,86]. If the formation of a complex in solution is considered as an equilibrium reaction, then two parallel reactions can take place in solution, one between the complexing agent and the impurity, and another between the complexing agent and the API (Compound C). These expressions are shown in equations **3-3** and **3-4**.

$$I + A \rightarrow IA \qquad\qquad 3\text{-}3$$

$$C + A \rightarrow CA \qquad\qquad 3\text{-}4$$

Where $I$ is the impurity, $A$ is the additive, $C$ is Compound C, $IA$ is the complex between the impurity and complexing agent, and $CA$ is the complex between Compound C and the complexing agnet. The equilibrium from equations **3-3** and **3-4** can be represented mathematically, as shown in equations **3-5** and **3-6** below.

$$K_I = \frac{[IA]}{[I][A]} \qquad\qquad 3\text{-}5$$

$$K_C = \frac{[CA]}{[C][A]} \qquad\qquad 3\text{-}6$$

Where the bracketed terms represent the concentration in solution of each respective compound, and $K_I$ and $K_C$ are the respective thermodynamic equilibrium constants of the reactions shown in 3-3 and 3-4. If equations 3-5 and 3-6 are solved for the concentrations of the complexes in solution, and the resulting expressions divided by each other, expression 3-7 is obtained.

$$\frac{[IA]}{[CA]} = \frac{K_I[I][A]}{K_C[C][A]} = \frac{K_I[I]}{K_C[C]} \qquad\qquad 3\text{-}7$$

The ratio shown in equation 3-7 between the complex with the impurity or Compound C is a direct representation of the selectivity of interaction of a given complexing agent towards the impurity. Throughout this chapter it has been explained that obtaining a high selectivity is key for the

successful implementation of complex-assisted crystallization. However, equation 3-7 shows an extremely important limitation faced in this system: the selectivity is influenced by the relative concentrations of the impurity and Compound C. This correlation means that if there is too much excess of Compound C with respect to the impurity, the selectivity of interaction is reduced, even if the binding constant towards the impurity is greater. This occurs because of Le Chatelier's principle, which is affected by the excess of one of the molecular species (Compound C in this instance), driving its respective reaction towards completion (complex formation). If we consider that for the experimental conditions explored in this project the overall concentration of Compound C is approximately 100 times that of its impurity, equation 3-8 becomes true.

$$\frac{[IA]}{[CA]} \approx 0.01 \frac{K_I}{K_C} \qquad\qquad 3\text{-}8$$

Given the relationship in equation **3-8,** it is clear that to overcome the excess of concentration in compound C it would be necessary for $K_I$ to be substantially larger than $K_C$. In fact, in order to have just as much complex formation with the impurity than with Compound C, it would be required that $K_I \geq 100 K_C$. Considering the definition of the equilibrium constants, as shown in equation 3-9, a binding constant 100 times greater towards the formation of a complex with the impurity would represent a $\Delta G \approx 12 \ kJ/mol$. Based on the predicted $\Delta H$ data obtained from the computational work (results shown in the Appendix), not a single one of the complexing agents found in this project was selective enough towards the impurity. The best performer, according to Material Studio, should have been 2DAMH, with an energy of interaction of approximately 9 kJ/mol. However, this estimate is only for the energy ($\Delta H$), and ignores the detrimental entropic penalties caused by the formation of a complex in solution, which could potentially drive the value of $\Delta G$ lower. As a result, the thermodynamic analysis and the computational data seem to indicate that the limited success of the purification experiments stems from insufficient selectivity towards the binding with the impurity.

$$\ln K = \frac{\Delta G}{RT} \qquad\qquad 3\text{-}9$$

The analysis performed with the equilibrium calculations elucidates some of the most relevant trends observed within this chapter. For starters, the fact that for all the additives tested (as predicted by equation **3-8**) there was as much (or more) complex formed with Compound C than with the impurity explains why almost all crystallization experiments displayed a decrease in the product yield. The formation of a substantial amount of complex with Compound C led to the APIS's solubilization, reducing the supersaturation and driving the yield down. This effect also explains why increasing the amount of complexing agent (as shown in the experiments with excess stoichiometry) improved the effectiveness of the best-performing additives. As the amount of complexing agent was raised, the selectivity limitations were partially overcome, and a significant concentration of the impurity complex could be formed, hence allowing for a true improvement in purification. Regrettably, the purity gains had to always come with increasingly prohibitive yield losses. The presence of such a noticeable yield drawback attached to the purity improvements was not seen in past projects with complex-assisted crystallization[66,67,70] for two reasons. First, in past systems there was not such a small initial amount of impurity driving the excess in equilibrium discussed in the equations of this chapter. In past projects impurity concentrations in initial batch recipes started as high as 6 %wt of the initial solids, making the thermodynamic limitation of equation 3-8 less prohibitive. The second reason is that in past systems, the structure of the

53

molecules involved were substantially simpler, and the chemistry required to obtain selectivity of interaction towards the impurity was less challenging. The large structure and complex functional groups of Compound C and its impurity are a unique design hurdle. Because of it, this project requires specific chemistry and a unique design for complexing agent molecules that was not achievable within the framework of the work laid out herein.

## 3.4 Conclusions

The results summarized in this Chapter provide fascinating insight on the applicability and limitations of complex-assisted crystallization. With the complexing agents tested the best purification that was achieved with respect to controls and after adjusting for yield penalties was of a 14 % reduction of impurity in the final product composition. A large set of complexing agents were studied as potential candidates for purification; regrettably, it was not possible to successfully achieve the ambitious purification goals that were originally outlined by Novartis. The batch experiment results indicated that, at the conditions and concentrations studied, the additives tested tended to have only small effects on the final crystal purity, while consistently incurring on yield penalties. The best-performing additives were found to be 2DAMH and NDTFPG. Membrane experiments were carried out to try to further characterize and optimize the separations with the best-performing additives. However, the membrane trials demonstrated that the complexing agents selected did not provide enough selectivity towards the impurity molecule (2DAMH), or they simply did not interact strongly enough at all with neither Compound C nor its impurity.

Theoretical analysis performed around the equilibrium chemistry of complexation indicated that part of the problem encountered in this project was the extremely low concentrations of impurity with respect to Compound C. This hindrance, coupled with the limited selectivity of interaction that could be achieved with complexing agents, led to the inability to reach the degree of improvement desired on the crystallization procedure. Since only small improvements could be achieved in this project, Novartis requested that complex-assistance crystallization should be pursued no further for the purification of Compound C. The reasons behind this recommendation are the involved chemistry that is required for successful complexation selectivity, as well as the time investment necessary to design and optimize an appropriate additive. The effort required for the successful implementation of this project made no economic sense, especially when an alternative purification method for Compound C exists. It is however the opinion of the author that with sufficient time and synthetic resources, a complexing agent molecule could be custom designed to maximize selectivity to a degree where the desired purification was achievable.

The insights obtained thanks to the work within this chapter are crucial for outlining the useful applicable window of complex-assisted crystallization as a viable technique for crystallization enhancement. It is clear that the case-by-case nature of the technique poses a substantial challenge, especially when facing complex chemical structures, such as the ones discussed in this chapter. Similarly, an important limitation was found, related to the relative concentrations of impurity within the crystallizing solution. In general, it was discovered that the successful implementation of complex-assisted crystallization requires the consideration of a trade-off, where lower initial concentrations of impurity demand higher complexation selectivity towards the impurity. In the particular case of Compound C, the intricate chemistry of the relevant molecules, coupled with the stringent purification requirements led to the inability of reaching the desired purification goals.

## 3.5 Appendix

### Predicted Interaction Energies

As explained in the materials and methods section, computational analysis was used to partially guide the selection of potential complexing agents for the crystallization experiments. Both Avogadro and Material Studio were used to perform this analysis. The data from Material Studio is more thorough given that it was based on a more complete and modern force field for the energy minimization calculations. In the tables below there is a plethora of molecules tested, not all of them were explored experimentally, given that the computational work showed some of them to be completely unfit as additives, lacking any selectivity towards the impurity. In the tables there are three quantities reported. The first is the interaction energy with the impurity, calculated using equation 3-2. This quantity was found subtracting the minimized energy of the non-interacting molecular pair, to the energy of interacting pair. Negative amounts indicate a beneficial interaction between molecules (energy stabilization). An identical calculation was performed between the complexing agent and Compound C ($\Delta E_{API}$). The energy of interaction of each given complexing agents with the impurity and API were then subtracted to one another. The resulting difference is an indication of the selectivity of interaction ($\Delta\Delta E$), or how much stronger was the energy of interaction towards one molecule versus the other. In general, a negative value of $\Delta\Delta E$ indicates a prediction of better selectivity towards the impurity than to the API. As a result, strongly negative numbers are expected to predict better complexing agents.

Avogadro

| Complexing Agent | $\Delta E_{Impurity}$ | $\Delta E_{API}$ | $\Delta\Delta E$ |
|---|---|---|---|
| 5-Nitroisophthalic acid | -51.59 | -43.44 | -8.15 |
| Xanthine | -50.21 | -37.42 | -12.79 |
| Theobromine | -55.21 | -43.72 | -11.49 |
| Saccharin | -47.13 | -33.34 | -13.79 |
| Methyl 2-aminothiophene-3-carboxylate | -33.18 | -33.41 | 0.24 |
| Anthracene-9-carboxylic acid | -87.03 | -81.58 | -5.45 |
| Diphenylacetic Acid | -52.44 | -47.20 | -5.24 |
| DOTG | -48.69 | -30.35 | -18.34 |
| 3,3,3-Tris(4-chlorophenyl)propionic acid | -59.02 | -53.95 | -5.07 |
| 2,7-Di-tert-butylfluorene-9-carboxylic acid | -62.31 | -45.91 | -16.40 |
| Mellitic Acid | -52.07 | -53.25 | 1.18 |
| 9-Fluorenone-1-carboxylic acid | -51.75 | -61.19 | 9.44 |
| 4-Iodobenzoic acid | -50.97 | -41.87 | -9.10 |
| 3-Hydroxy-4-iodobenzoic acid | -52.29 | -41.91 | -10.38 |

Material Studio

| Complexing Agent | $\Delta E_{Impurity}$ | $\Delta E_{API}$ | $\Delta\Delta E$ |
|---|---|---|---|
| 3,3,3-Tris(4-chlorophenyl)propionic acid | -18.56 | -18.40 | -0.16 |
| 27DTBF9CA | -15.66 | -14.86 | -0.80 |
| 3Hydroxy4IodoBA | -12.17 | -15.17 | 3.01 |
| DOTG | -21.86 | -23.73 | 1.86 |
| Saccharin | -13.94 | -17.27 | 3.32 |
| Thymidine | -20.49 | -21.18 | 0.68 |
| 2-Deoxyadenosine Monohydrate | -28.54 | -19.26 | -9.28 |
| 2'-Deoxyguanosine Monohydrate | -18.57 | -18.10 | -0.47 |
| 2'-Deoxygcytidine Hydrocloride | -11.54 | -12.03 | 0.48 |
| Uridine | -17.79 | -17.85 | 0.06 |
| 3,3',5,5'-tetracarboxydiphenylmethane | -18.27 | -27.79 | 9.51 |
| Citrate | -15.03 | -26.06 | 11.03 |
| NDTFPG | -33.48 | -36.90 | 3.42 |

# Chapter 4 : Nucleation inhibition of benzoic acid by means of complexation

The contents of this chapter were adapted from work published in a peer-reviewed journal article[52]:
Pons Siepermann, C. A.; Huang, S.; Myerson, A. S. Nucleation Inhibition of Benzoic Acid through
Solution Complexation. *Cryst. Growth Des.* **2017**, 17 (5), 2646–2653.

## 4.1 Introduction

The prevention of undesired crystal formation in supersaturated systems is a prevalent challenge
faced by the medical, oil, chemical, and water purification industries. Unwanted material
precipitation can lead to product losses, uncontrolled changes in concentration in chemical systems,
and equipment fouling and clogging. Extensive research has been carried out in the search for
efficient mechanisms that can deter the formation of crystalline precipitates, mostly focusing on
the growth inhibition of crystal phases through the implementation of crystal-phase adsorbing
additives[2,21,87]. The mechanisms underlying crystal growth inhibition have been extensively
researched and are well understood[2,41,53]. However, preventing the onset of nucleation (rather than
growth), has proven to be a much greater challenge. Nucleation inhibition has been limited in great
part by lack of fundamental and theoretical understanding of the true underlying mechanisms that
govern nucleation itself, and by the existence of limited tools to properly characterize and quantify
nucleation induction times[1,2].

Recent work[88,89] has shown that high-throughput experiments used to measure induction times for
a large dataset can be used to statistically quantify the induction times for a system with well-
defined conditions at a constant volume and initial supersaturation. These techniques have been
used to successfully account for the stochastic nature of nucleation, and have allowed the
determination of the nucleation rate and induction times for the crystallization of small organic
molecules. In this chapter, we explore the implementation of a complexing agent as a nucleation
inhibitor, quantitatively monitoring the effect of 1,3-di-o-tolylguanidine (DOTG) in the induction
time of benzoic acid (BA) at constant supersaturation.

It is a well-known fact that non-covalent intermolecular bonding between small molecules can
have a significant effect in species solubilities and thermodynamics[12,15]. The traditional
rationalization of this effect is that the formation of the complex establishes a chemical equilibrium
that effectively removes the free species from solution, allowing for a total greater amount of the
interacting molecules to be dissolved at a constant set of experimental conditions. This theoretical
and experimental precedent with solution complexes indicates that by introducing additives that
force strong hydrogen bonds with a target solute, it should be possible to artificially remove them
by "transforming" them into their respective complex, reducing the number of available molecules
that can form a high-concentration cluster and, by extension, an ordered crystal nucleus. It has also
been demonstrated that molecular complexes are less likely to incorporate into growing crystals
of a compound that is highly structurally similar to one of the complexed molecules[65]. By extension,
it is reasonable to postulate that such complexes would be incapable of contributing to the ordering
and nucleation of crystals of one of the constituent molecules.

Within the context of applied crystallization, the use of tailor-designed complexing agents to
establish solution interactions has been explored recently as a tool to improve the purity of
crystallizing systems[65,66,70]. Research has shown that complexing agents can be used to selectively
remove a lattice-incorporating impurity from a crystallizing product without reducing yield, due
to the intermolecular interactions between said impurity and the additive. The prevailing theory to

account for this behavior is that the solution complexes are too bulky to incorporate into the growing crystal lattices, maintaining the desired impurity dissolved. For the specific goal of inhibition, the utilization of complexation additives has also been reported as a successful strategy for the suppression of the crystallization of selected polymorphs of a target molecule in solution[3]. Additionally, recent work with precipitation polymeric inhibitors (PPIs)[90,91] has been successfully demonstrated to lead to kinetic inhibition of precipitation of commercial active pharmaceutical ingredients. However the nature of past nucleation inhibition work has been highly empirical in the selection of additives for complexation, and has focused on the development of libraries of viable candidates for specific targets. To the knowledge of the authors, a detailed quantitative study of the induction inhibition effects of a tailor-designed molecular additive on a target molecule has not been established.

This project focuses on characterizing the mechanistic source of nucleation inhibition, differentiating solubility, equilibrium-speciation and kinetic effects. Experiments were carried out carefully accounting for supersaturation and shifts in solubility that took place upon the addition of the complexing agent. The selected model system was based on an experimentally verified complexation pair from a separate study[66] to ensure that strong intermolecular interactions between the relevant species were prevalent in solution. Experiments were carried out with large sample sets in order to fully characterize the probabilistic nature of induction; the data was then reported in terms of representative induction times based on measured nucleation rates, calculated using principles of induction theory[88].

## 4.2 Experimental

### Materials

Benzoic acid and 1,3-Di-o-tolylguanidine, and trifluoroacetic acid were purchased from Sigma-Aldrich (St. Louis, MO) and used as received. Ethanol (200 proof) was purchased from VWR International (Edison, NJ). Microscope 1 mL glass vials with septum polyethylene plug caps were purchased from Fischer Scientific LLC (catalog #033752G). Polytetrafluoroethylene (PTFE) membrane filters with 0.45 µm pores were purchased from VWR International.

### High-Performance Liquid Chromatography

For BA and DOTG detection, the HPLC instrument (Agilent 1260) was equipped with a UV-diode array detector (DAD). The column used was a YMC-Pack ODS-A 150 mm X 4.6 mm i.d. packed with 3 µm particles with 12 nm pore size (YMC America Inc.). The detection wavelength was set at 230 nm for BA and 240 nm for DOTG. The samples were analyzed using an isocratic method with a 30:70 (volume) methanol/water mobile phase containing 0.3 %vol. trifluoroacetic acid, with 10 µl sample injections, a 1 ml/min solvent flow rate, and a column temperature of 30 °C.

### Solubility Curve Determination

Experiments were performed to determine the aqueous solubility of benzoic acid at a reference temperature of 15°C as a function of DOTG concentration. Samples with excess BA were prepared in 10 ml of water with varying amounts of DOTG, heated until complete dissolution was achieved, then cooled to 15°C. After crystals formed in the vials, the samples were stirred for 24 hours at a constant temperature to allow for equilibration. The supernatant from each equilibrated sample was then recovered, filtered using 0.2 µm Acrodisc Syringe Filters, immediately diluted by a factor of 20 in water, and analyzed using HPLC. The resulting concentrations of DOTG and BA were used to develop a solubility curve.

## Design of Nucleation Experiments

Experiments were carried out following an unstirred variant of a procedure for induction time determination described in wok from Joop ter Horst's group[88,92]. A large set of identical nucleation experiments were performed at varying concentrations of BA and DOTG in order to study the effects of the complexation between the molecules on induction time. Samples containing identical solutions of BA/DOTG were prepared in 1 ml shell vials and placed in a Zeiss Axio Observer.Z1m Microscope in groups of 80 at a time. The microscope was equipped with a HAL100 light source and a Z800 Workstation. The samples were fixed in an unstirred stage with a hollow chamber connected to a recirculating chiller for temperature control. Temperature was rigorously maintained at 15°C for all experiments. The microscope was programmed to continuously focus and scan across each of the vials and take a picture at 25X magnification, recording the state of the solution in every sample. Each vial had a picture taken every two minutes. The pictures where visually inspected to determine the timestamp in which the very first crystal nucleus was observed. The measured induction times were used to determine the value of the representative nucleation rates for each set of experimental conditions. The data processing used to determine nucleation rates is discussed in the discussion section within this chapter and explored in further detail in the appendix. This setup was selected due to its high-throughput sampling capacity and its imaging feature, which allows for easy visualization of the onset of nucleation events. The major drawback of this setup lies on its inability to provide agitation for the samples.

## Solution Preparation

To minimize the influence of organic residues and dust particles in the 1 mL shell vials used in the nucleation experiments, all vials were washed twice with 200 proof ethanol and dried in a vacuum oven at 70°C for 1 hour. This was done immediately before each microscope experiment.

The solutions used for the induction time measurements were prepared in DI water with BA and DOTG. The amounts of solute needed to achieve the target supersaturation were measured in an analytical balance and dissolved in a hot stirring plate at 70°C. After full dissolution, the hot solution was filtered through PTFE membranes with 0.45 μm pores and placed into 20 mL glass vials preheated in a 70°C VWR heat block. The 20 ml vials were also cleaned with ethanol and dried the same way as the Shell vials. Aliquots of 500 μL of the hot filtered solution were pipetted into the heated Shell vials and immediately capped to prevent evaporation. After capping, the samples were left 10 minutes on the heating block to equilibrate in temperature, then were rapidly transferred to the microscope setup described earlier and monitored using the microscope image tracking capabilities.

## Selection of Experimental Species

The species selected to carry out the nucleation inhibition experiments are shown in Figure 4-1. The molecules were selected based on the strong existing precedent of molecular interaction and solution complexation between them, as reported by *Weber et al*[66]. Under the selected experimental conditions, the BA becomes anionic, donating its hydrogen from the carboxylic acid moiety. Meanwhile, the DOTG becomes cationic by accepting a proton within its guanidine functionality. The resulting structure leads to a strong attraction between the two molecules, which form a 1:1 charge-neutral complex. The experimental solvent (water) was selected because it has a steep solubility curve for benzoic acid[93], which allowed inducing high supersaturations ($S > 2$) using moderate changes in temperature (55 °C). The ability to generate a strong crystallization driving force using simple cooling was desirable due to the experimental setup employed.
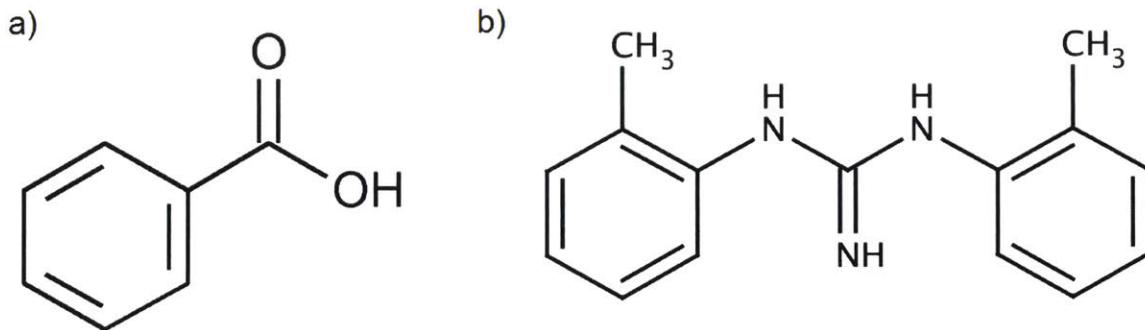
**Figure 4-1:** Structure of the inhibition target, benzoic acid (a), and the selected complexing agent to induce inhibition, 1,3-di-o-tolyl-guanidine (b).

## 4.3 Results and Discussion

It has been extensively reported that complexing agents can have a measurable effect on the solubility of small-molecule organic solutes[12,15]. This precedent, and the necessity of carefully monitoring solubility and supersaturation in induction time experiments made it necessary to determine the correlation between BA solubility in water at 15 °C (experimental conditions), and the concentration of DOTG additive present. Solubility was determined as explained in the experimental section, and the resulting solubility correlation is shown in Figure 4-2 below.
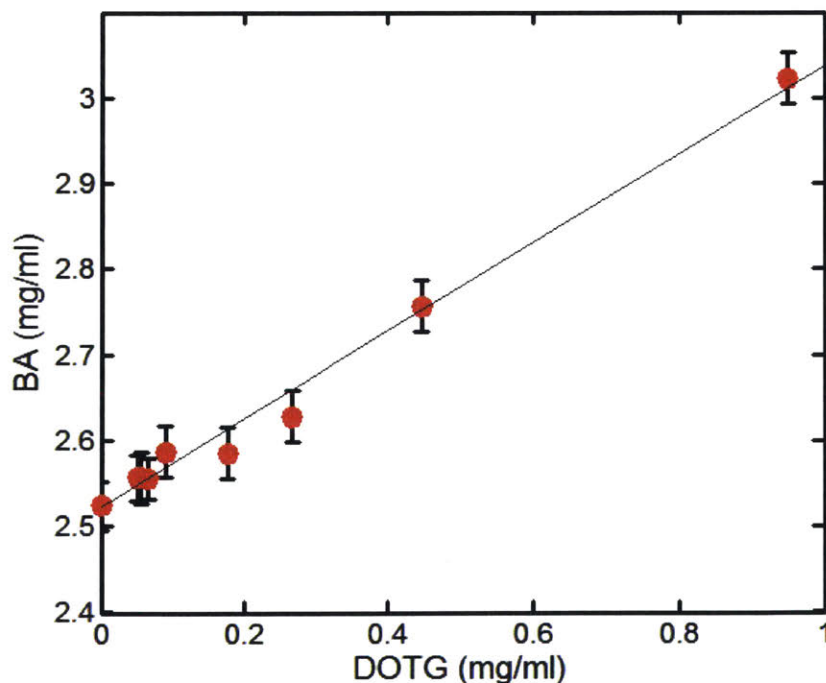


**Figure 4-2:** Effect of DOTG concentration in BA saturation concentration in water at 15 °C. The correlation line was calculated using least squares fitting: $C_{BA} = 0.51 C_{DOTG} + 2.52$. The standard deviation of the slope is $\sigma_m = 0.03$, the standard deviation of the intercept is $\sigma_b = 0.01 \, mg/ml$.

The data in Figure 4-2 clearly demonstrates an overall trend of increasing BA solubility with increasing DOTG concentration. This is in agreement with previous complexation results[12,15], and with past observations of strong interaction between DOTG and BA[65,66]. If the correlation shown

in Figure 4-2 were transformed into a molar basis, the slope becomes exactly unity. This indicates that for each mole of complexing agent added, an extra mole of BA can be dissolved. If we consider this correlation from the perspective of equilibrium chemistry, this implies that the formation of the BA-DOTG complex behaves as if it had an infinitely large equilibrium constant (e.g the complexation process goes to perfect completion). This behavior can be explained by equilibrium chemistry when two facts are taken into account: the interaction between DOTG and BA is very strong (high binding equilibrium constant), and the concentration of BA is high (in excess), meaning that the equilibrium is shifted towards the formation of the complex. The "equilibrium chemistry" section in the appendix explores the mathematics behind the chemistry, explaining this observed result.

An additional result of interest that becomes apparent from the saturation experiment is that the solubility of DOTG was also raised significantly above its natural value. The solubility of DOTG in pure water at the experimental temperature is approximately 0.07 mg/ml (0.29 mMolar). However, in the presence of BA the amount of DOTG that could be successfully dissolved approached 1 mg/ml (4.18 mMolar). This further demonstrates the strong solubilization effect of complexation.

The probability calculations that describe the stochastic process of nucleation follow the theory described in the work of Jiang and ter Horst[88]. A detailed derivation of the equations that allow for the calculation of nucleation rate and induction time is provided in the appendix. The most important expression used in this work is equation 4-1, shown below, which correlates the probability of any crystal to have nucleated ($P^*$) at a given time ($t$), in a vessel with a fixed volume ($V$), with the system's average nucleation rate ($J$).

$$\ln(1 - P^*) = -JV(t - t_g) \qquad\qquad 4\text{-}1$$

The variable $t_g$ is a correction factor related to the growth time required for a visible crystal to be formed before the actual induction time is detected. To apply equation 4-1, it is assumed that $t_g$ is small when compared to experimental sampling times. For this chapter, experiments were designed to ensure that crystals grew to visible sizes rapidly after nucleating. Ensuring this key behavior was an experimental design consideration for the selection of the appropriate base supersaturation in the control group. The interval of data collection of 2 min was much larger that the time a crystal required to grow to become visible, validating the utilization of equation 4-1. This was confirmed by estimating crystal growth rates using image analysis of the crystals in the microscope as they grew. The crystal growth analysis is discussed in detail in the appendix.

The value of $P^*$ can be determined experimentally. Using equation 4-1, the average nucleation rate can be calculated by fitting the probability data into a line, whose slope will be $-JV$. The inverse of the value $JV$ has units of time per number, and it is representative of the induction time of the system with a static set of initial conditions. This quantity will be referred to as $\tau$, and will be used to quantitatively represent and compare the speed of nucleation of a BA sample. The greater the value of $\tau$, the slower the nucleus formation in the experiments.

To measure $P^*$, a large set of identical crystallization experiments have to be made, monitoring the time in which the first crystal (nucleation) is formed. Using this data set, it is possible to establish the probability of a crystal to have nucleated at a given time, linearize the data, and obtain an experimentally determined value for a representative induction time. A visual example of the

images employed to determine nucleation time can be seen below in Figure 4-3.A graphical example of the data collection and analysis process is provided in Figure 4-4.

A reference supersaturation of uncomplexed BA had to be selected to generate a control data set. Trial supersaturations were tested following the procedure described above. An optimal control data set was required to display zero nucleation events in the first set of pictures collected (corresponding to time "0"), and to yield observable crystals as quickly as possible after the starting point in order to generate statistically significant probability vs. time curves. Induction time experiments were carried out as described in the materials and methods section. The first set of data collected corresponded to the control group without any complexing agents. The supersaturation that met the desired performance criteria was found to be S=2.14, having no observable nucleation events in the initial set of pictures (t = 0 minutes) and displaying a large number of visible crystallized vials quickly thereafter. The data obtained from 554 vials is summarized in Figure 4-4. The images shown in the figure illustrate the steps of data processing and conversion required to measure nucleation rate and induction time.
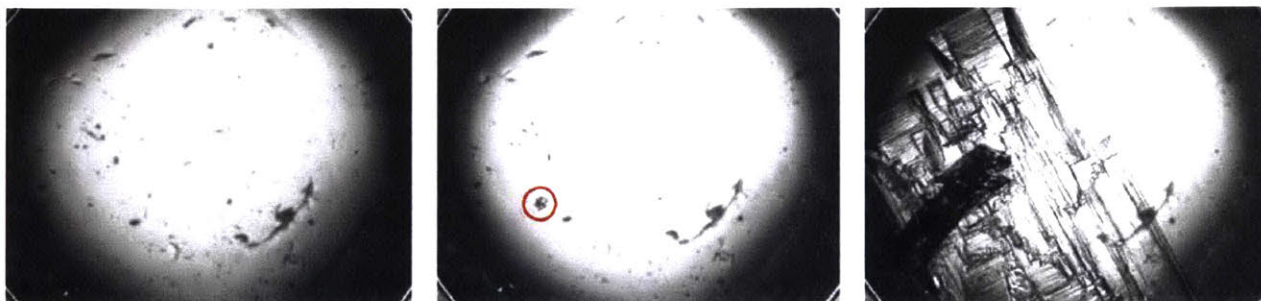


**Figure 4-3:** Sample data point of a crystallization experiment in the microscope setup described and explored in this investigation. The pictures depict the progression of the nucleation and growth of a crystal in one of the 80 vials that are monitored simultaneously in a given experiment. This data correspond to a control experiment with no DOTG additive. Left: initial conditions, no time elapsed; center: first crystal nucleus observed (circled for emphasis); right: final crystal size after experimental completion (180 minutes).

The data from Figure 4-4 evidences an important non-linearity that was observed for long experimental times, as shown in sub-image c). All experiments (with or without complexing agent) carried out in the setup described in the experimental section displayed this behavior. The exact underlying cause for this observed curvature in induction over time is unknown, although non-linearity has been reported in other published works studying nucleation[89,94]. In similar work, non-linearity has been attributed to complex nucleation mechanisms and to two-step kinetics of crystal formation, however, similar molecules have been observed to conform to the linear model for equivalent timescales[92], for measurements carried out under stirred solutions. The true cause for the non-linearity observed in Figure 4-4 is likely imperfections in the experimental setup, such as glass-vial adsorption of solute, or moisture condensation within the vials. These problems could potentially affect the supersaturation within the vials and skew the experimental data over long times, and they are difficult to fully characterize and control. Nevertheless, a linear period was consistently observed at early times for all experimental samples taken in this study. Given that the focus of this work is the comparative effect on induction times in the presence of additives, this ideally-behaving portion of the probability correlation was considered to be an adequate representation of induction times for each given system, and was used to compare representative nucleation rates. All subsequent induction data is reported in terms of the early-time behavior

(equivalent to Figure 4-4d). The reported linear behavior allowed monitoring the nucleation distributions within the range$-0.1 < \ln(1 - P) < -0.2$. The exact time at which linearity ceased for each experiment varied and was dependent on the degree of inhibition of each sample, given that slower crystallizing samples reached smaller values of $\ln(1 - P)$. All linear fits were constrained to pass through the origin to match the theory discussed in equation 4-1.
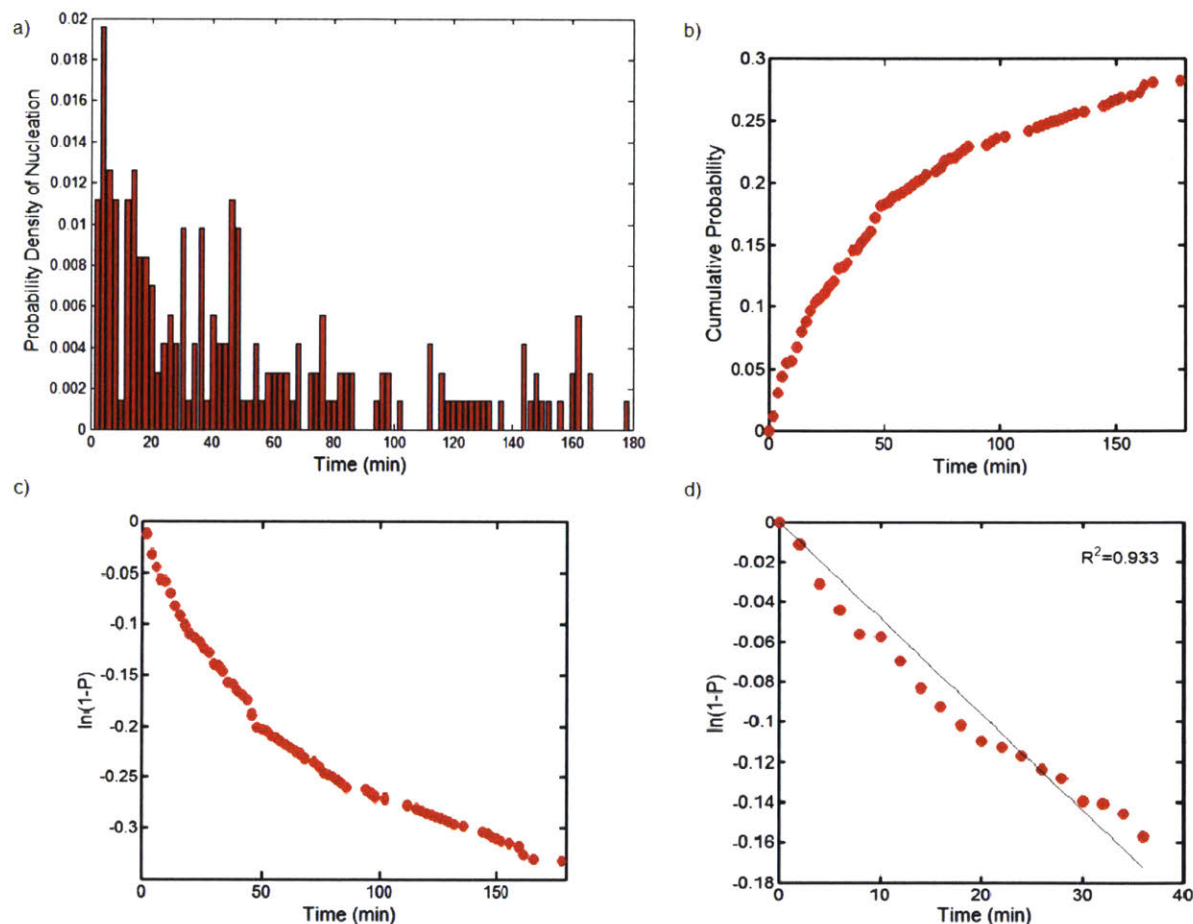


**Figure 4-4:** Induction time microscope results for pure benzoic acid in water at 15 °C with S=2.14. Data shown for 554 experimental vials. a) Probability density distribution of a nucleation event to be observable as a function of time; b) Cumulative probability function as a function of time; c) Logarithm conversion of the cumulative probability to obtain nucleation rates (equation 4-1); d) Magnification of the early times of plot "c", which were observed to be linear. Nucleation rates reported were measured based on the slopes of lines obtained in plots equivalent to "d". The representative induction time measured for the control was $(JV)^{-1} = 208 \pm 20\ min$.

The effect of DOTG in the solubility of BA was demonstrated by the change in BA solubility reported in Figure 4-2. As a consequence of this phenomenon, any experiment carried out to measure changes in the induction time of BA in the presence of DOTG has to account for the effects of this solubility behavior in the system's supersaturation. Supersaturation in this chapter is defined in equation **4-2**.

$$S = c/c^*$$

4-2

63

In equation **4-2**, $S$ is the reported supersaturation, $c$ is the dissolved concentration of BA in water, and $c^*$ is the saturation concentration of BA in water at a given temperature and concentration of complexing agent (DOTG). To compensate for shifts in solubility in the presence of DOTG (causing an increase $c^*$), an adjusted excess amount of BA had to be added to samples that contained DOTG to ensure that each experiment was carried out at a constant supersaturation, thereby warranting that the measured induction times could be compared with respect to a common set of baseline conditions. The excess amount of BA required to keep constant supersaturation was established using the solubility equation shown in Figure 4-2. The effect of adjusting the supersaturation in the presence of complexing agent is reflected in Figure 4-5, and Table 4-1. The control experiment had a saturation of $S = 2.14$, which corresponds to a concentration of 5.4 mg/ml of BA (44.2 mMolar). The crystallized BA samples obtained from the induction time experiments were confirmed to consistently be the same BA crystal polymorph using Powder X-Ray Diffraction (PXRD), even in the presence of DOTG. A brief discussion of the PXRD data can be found in the appendix.

The data collected indicates that the presence of DOTG slows down the induction time of BA with increasing complexing agent concentration. Additionally, it is clear that even when adjusting the BA concentration to preserve constant supersaturation the measured induction time is greater than the control. The results for samples with 0.01 mg/ml DOTG (0.04 mMolar) show the aforementioned trend on average, however the samples remain close to each other if the 95% confidence error is taken into account. For the samples with 0.05 mg/ml DOTG (0.21 mMolar), a significant nucleation inhibition effect is observed, far greater than the calculated error. The "reference" 0.05 mg/ml DOTG (0.21 mMolar) sample was measured to have an induction time 167% larger than the control. The impact of potential growth inhibition effects caused by DOTG was estimated to be minimal. The linear growth rates of samples in the presence and absence of complexing agent were compared using image analysis of the microscope data. This analysis is shown in detail in the appendix. It was estimated that for all the experiments reported in this chaper, the linear growth rate of the BA crystals was approximately 40 μm/min.

It is of note that the "adjusted" samples crystallized more slowly than the controls, despite having a greater concentration of BA dissolved in each sample to correct for the shift in supersaturation. Furthermore, the results indicate that the adjusted samples behaved more similarly to the "reference" experiments than to the control, indicating the change in inhibition effect observed is not dominated by the differences in solubility, but rather an alternative mechanism that prevents crystal formation.

A set of experiments to measure the induction time of BA was carried out at varying concentrations of DOTG in order to determine the importance of complexing agent amount with respect to observed nucleation inhibition effect. The data obtained is summarized in Figure 4-6 and Table 4-2. The results clearly show that as the concentration of DOTG increases, the inhibition effect on BA nucleation becomes more prominent. The increasing induction time trend with DOTG concentration can be seen in Figure 4-7. All the data shown corresponds to samples that were adjusted to ensure a constant supersaturation of S=2.14. Each reported mass fraction of DOTG is calculated with respect to the original mass of BA in the control prior to the adjustment for solubility shifts.

64

**Table 4-1:** Induction time data for BA comparing the effect of DOTG before and after adjusting for changes in solubility, samples labeled "adjusted" have an excess amount of BA with respect to the control to ensure constant supersaturation. Induction time is reported as the inverse of the slope of the lines in Figure 4-5. The reported error is based on the uncertainty of the slope with a 95% confidence. $n$ is the number of crystallization samples that were used to obtain the data for each experiment.

| Sample | $S$ | $C_{BA}$ (mg/ml) | $\tau(min)$ | $n$ |
|---|---|---|---|---|
| Control | 2.14 | 5.400 | $208 \pm 20$ | 554 |
| DOTG 0.01 mg/ml DOTG Reference | 2.13 | 5.400 | $285 \pm 35$ | 290 |
| DOTG 0.01 mg/ml DOTG Adjusted | 2.14 | 5.411 | $256 \pm 18$ | 480 |
| DOTG 0.05 mg/ml DOTG Reference | 2.12 | 5.400 | $769 \pm 61$ | 240 |
| DOTG 0.05 mg/ml DOTG Adjusted | 2.14 | 5.460 | $556 \pm 55$ | 557 |
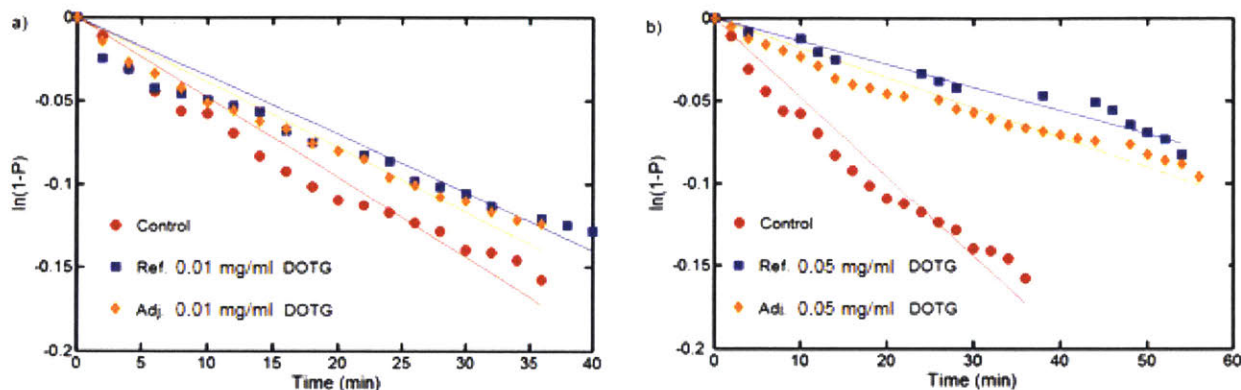


**Figure 4-5:** Induction time probability microscope results for BA in water at 15 °C in the presence of DOTG, comparison of behavior of complexing agent with respect to un-complexed control (red curves), and of effect of supersaturation equivalence adjustment. Blue curves show the induction for samples with no solubility adjustment, while yellow curves show samples with BA amount shifted to ensure constant supersaturation with respect to control. Data shown for samples with 0.01 mg/ml of DOTG (a), and for samples with 0.05 mg/ml DOTG (b). The induction times of these samples are summarized in Table 4-1.
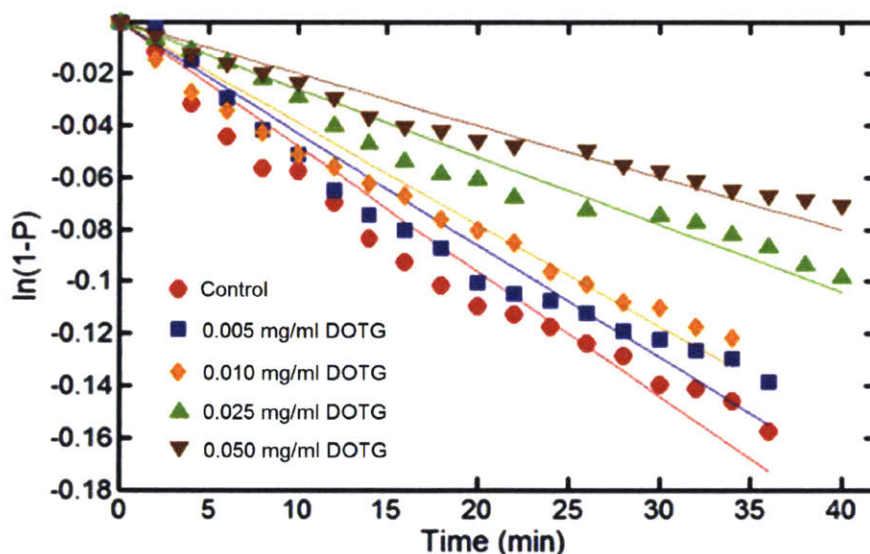


**Figure 4-6:** Induction time probability microscope results for benzoic acid in water at 15 °C, comparison of behavior un-complexed control (red curve), and supersaturation adjusted samples with increasing concentrations of DOTG. The induction time measured for these samples are summarized in Table 4-2.

**Table 4-2:** Induction time data for BA comparing the effect of varying amounts of DOTG at constant supersaturation. Induction time is reported as the inverse of the slope of the linear portion of the plots in Figure 4-6. The reported error is based on the uncertainty of the slope with a 95% confidence. The column for $n$ reflects the number of crystallization samples that were used to obtain the data for each experimental condition.

| Sample | $C_{BA}$ (mg/ml) | $\tau(min)$ | $n$ |
|---|---|---|---|
| Control | 5.400 | $208 \pm 20$ | 554 |
| DOTG 0.005 mg/ml Adjusted | 5.406 | $232 \pm 16$ | 607 |
| DOTG 0.010 mg/ml Adjusted | 5.411 | $256 \pm 18$ | 480 |
| DOTG 0.025 mg/ml Adjusted | 5.430 | $385 \pm 34$ | 459 |
| DOTG 0.050 mg/ml Adjusted | 5.460 | $556 \pm 55$ | 557 |
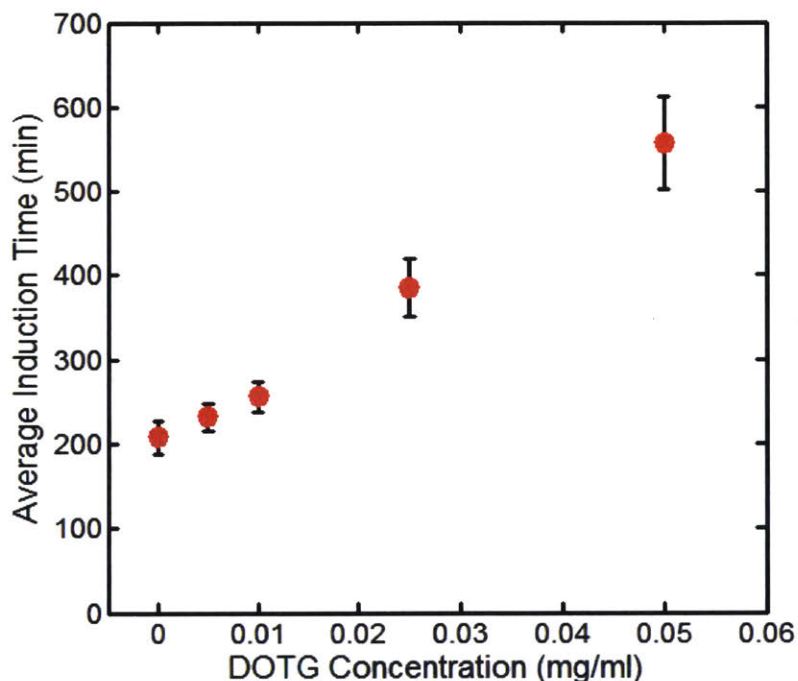


**Figure 4-7:** Calculated induction time microscope results for benzoic acid in water at 15 °C as a function of amount of DOTG added in sample preparation. All samples shown had a supersaturation of 2.14 calculated based on the saturation solubility of BA at each concentration of DOTG. Error bars are presented at 95% confidence.

Based on the data in Table 4-2, when the 95% confidence intervals are taken into account, it is difficult to conclusively establish that there is significant nucleation inhibition at DOTG concentrations below 0.025 mg/ml (0.10 mMolar). However, the measured values do consistently show a monotonically increasing induction time with increasing complex concentration. The measured induction times for the samples with 0.025 mg/ml (0.10 mMolar) DOTG and 0.05 mg/ml (0.21 mMolar) showed a very significant inhibition effect that is well beyond the sample error.

The current operating assumption for the mechanism of complexation is that DOTG and BA make a 1:1 mole complex, as indicated in literature. Additionally, for the system studied in this chapter, it is assumed that in a solution containing BA and DOTG, any complexing agent added to the system is completely utilized in the formation of a solution complex, as implied by the solubility data discussed in earlier results (Figure 4-2). Under these assumptions, upon adding DOTG, the

dissolved concentration of BA is effectively reduced by the equivalent amount of complex formed. For example, based on their respective molecular weights, a BA solution prepared with 0.05 mg/ml DOTG (0.21 mMolar) would experience a removal of 0.026 mg/ml (0.21 mMolar) of BA, since this amount of BA would be "consumed" in the formation of the BA-DOTG complex.

The results shown to this point demonstrate that the inhibition effects observed are not caused by simple supersaturation shifts. To further establish the operating mechanisms of inhibition, an additional set of experiments was carried out with the goal of determining whether the observed inhibition is caused by the speciation removal of BA described above, as would be predicted by equilibrium chemistry, or by an alternative mechanism, which would have to be kinetic. The experiments carried out explored the increase in induction time as the amount of BA within each sample was reduced, without adding complexing agent. The reduction in the BA supersaturation was done in terms of DOTG equivalents: the amount of BA expected to be removed under a given concentration of additive. The results for these experiments are shown in Table 4-3, Figure 4-8, and Figure 4-9. The data is compared to the trend observed for the experiments that were carried out with DOTG, as shown in Figure 4-10. In the upcoming results and discussion, samples are labeled based on the fraction of BA removed with respect to the original control (S=2.14).

The results shown in Table 4-3 and Figure 4-8 follow an expected trend of increasing induction time with decreasing supersaturation as BA is removed. The comparison displayed in Figure 4-10 provides the most striking data, showcasing the significantly stronger nucleation inhibition observed when the DOTG complex is present when contrasted to its equivalent BA complexation removal equivalent. On average, the induction time for the 0.01 mg/ml DOTG (0.04 mMolar) sample was 15% greater than its subtractive equivalent, similarly, the 0.025 mg/ml DOTG (0.10 mMolar) sample was 38% greater, and the 0.05 mg/ml DOTG (0.21 mMolar) sample was 28% greater.

Table 4-3: Induction time data for BA comparing the effect of varying amounts of BA, reported as DOTG equivalents for an assumed 1:1 complexation stoichiometry. Induction time is reported as the inverse of the slope of the linear portion of the plots in Figure 4-8. The reported error is based on the uncertainty of the slope with a 95% confidence. The column for $n$ reflects the number of crystallization samples that were used to obtain the data for each experimental condition. The "-%wt" of BA are with respect to the dissolved BA mass in the control.

| Sample | $C_{BA}$ (mg/ml) | DOTG Equivalent (mg/ml) | $S$ | $\tau(min)$ | $n$ |
|---|---|---|---|---|---|
| Control | 5.400 | 0 | 2.14 | $208 \pm 20$ | 554 |
| - 0.10%wt BA | 5.392 | 0.010 | 2.14 | $222 \pm 13$ | 478 |
| - 0.23%wt BA | 5.385 | 0.025 | 2.13 | $256 \pm 9$ | 479 |
| - 0.47%wt BA | 5.372 | 0.050 | 2.13 | $278 \pm 9$ | 320 |
| - 0.71%wt BA | 5.360 | 0.075 | 2.12 | $294 \pm 12$ | 479 |
| -0.94%wt BA | 5.347 | 0.100 | 2.12 | $435 \pm 19$ | 480 |
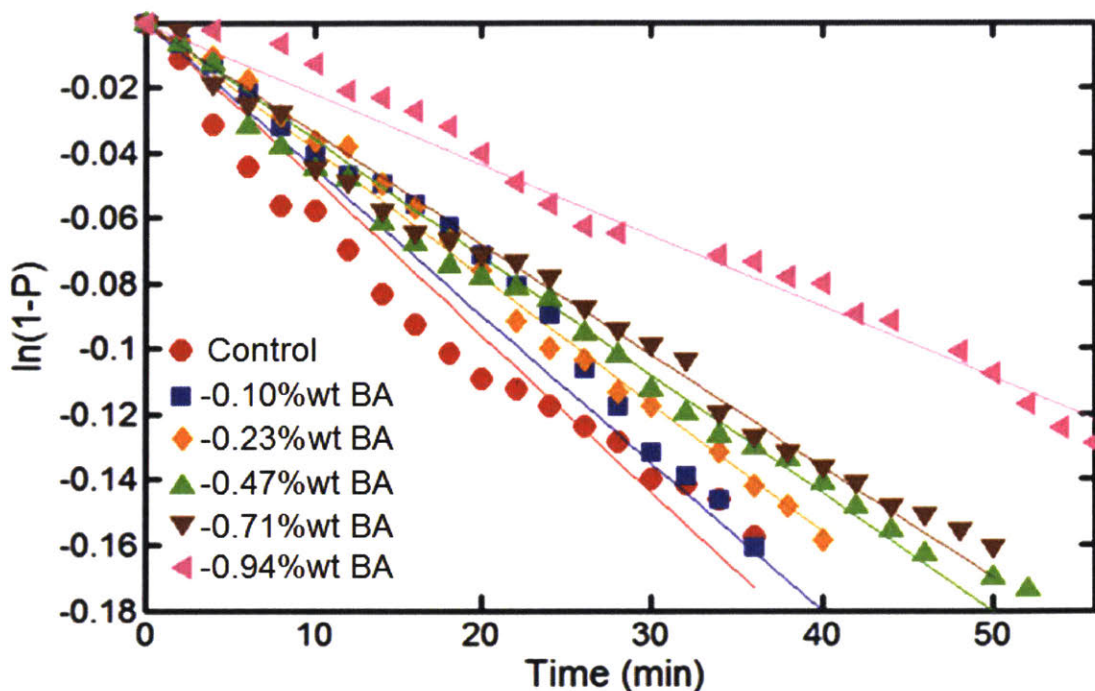
**Figure 4-8:** Induction time probability microscope results for benzoic acid in water at 15 °C, comparison of behavior un-complexed control (red curve), and samples with decreasing amount of BA, reported in terms of mass equivalent of DOTG. The induction time measured for these samples in Table 4-3.
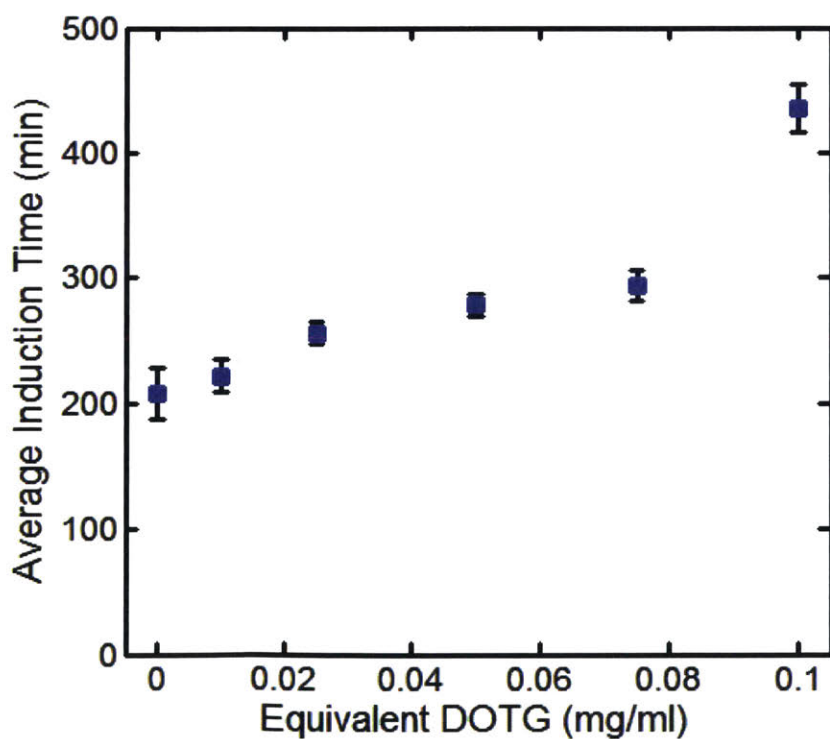


**Figure 4-9:** Calculated induction time microscope results for benzoic acid in water at 15 °C as a function of amount of equivalent DOTG complex mass. These samples were prepared by reducing the amount of BA in preparation. Error bars are presented at 95% confidence.
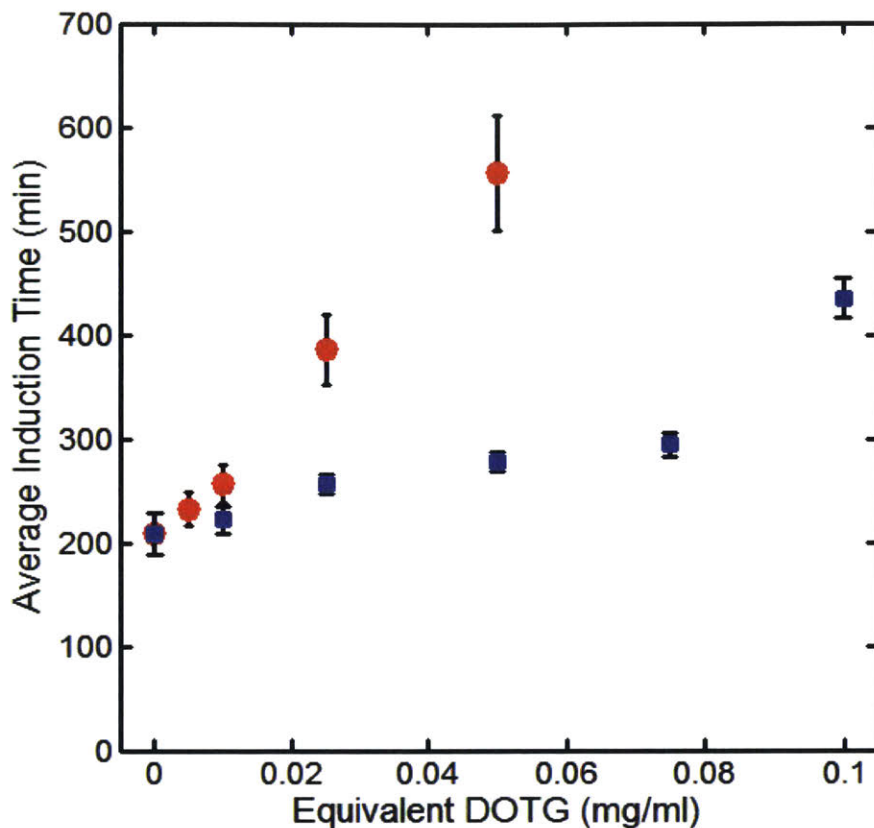
**Figure 4-10:** Calculated induction time microscope results for benzoic acid in water at 15 °C as a function of amount of equivalent DOTG complex mass for samples prepared with DOTG (red curve) and samples prepared by reducing the BA concentration assuming a 1:1 complexation stoichiometry (blue curve). These samples were prepared by reducing the amount of BA in preparation. Error bars are presented at 95% confidence.

The superior extent of the effect of DOTG inhibition when compared to the BA-removed samples is so prominent, that even if a 2:1 complex stoichiometry between BA:DOTG was assumed, it would be insufficient to rationalize the induction times measured for samples prepared with complexing agent based on equilibrium thermodynamics alone. For example, assuming a 2:1 complex, the sample prepared with -0.94%wt BA (-0.43 mMolar) would correspond to only 0.05 mg/ml DOTG (0.21 mMolar) added (instead of 0.1 mg/ml, or 0.42 mMolar, as shown in Table 4-3). However, the experiments show in this case that the BA-removed sample saw an induction time of 435 min, while the sample with DOTG was 556 min. Similarly, the samples with -0.47%wt BA (-0.23 mMolar) would correspond to 0.025 mg/ml DOTG (0.10 mMolar). For these samples, the uncomplexed set had $\tau = 278\ min$, while the complexed set had $\tau = 385\ min$. This behavior remains across the entire range of data explored in this chapter.

The data summarized in Figure 4-5 and Figure 4-10 indicates that the inhibition effect of DOTG is stronger than expected based on solubility and equilibrium chemistry thermodynamics, which implies that the true mechanism behind the measured inhibition has a significant kinetic component as well, which hinders the clustering of ordering of molecules into nuclei. The prevalence of a kinetic inhibition effect over thermodynamics is analogous to behavior observed in other works for polymeric precipitation inhibitors (PPIs)[90,91]. PPIs are known to interact with small molecule APIs without significantly changing their solubility, but greatly extending their

induction times. The exact mechanism for kinetic inhibition of nucleation is currently not well understood, in great part because the exact underlying mechanisms of nucleation itself are not well understood[1,2].

## 4.4 Conclusions

The effect of a solution complexing agent on the nucleation of benzoic acid was explored and quantified employing statistical measurements of induction time. The data clearly showed that the presence of small amounts ($\leq 0.05$ mg/ml, or 0.21 mMolar) of additive was sufficient to induce significant retardation of the onset of nucleation, increasing the representative induction time by up to more than 150% with respect to uncomplexed controls. Adjustments based on solubility and supersaturation were made in order to compare the effect of the complexation additive at a constant supersaturation of benzoic acid. Results indicated that the inhibition effect observed is prevalent at constant supersaturation despite the fact that a greater amount of BA solute was dissolved in each sample with increasing DOTG concentration. The data also indicates that with an increasing complexing agent concentration, the inhibition effect, and hence the induction time, monotonically increase in the range of additive concentrations explored.

Traditional solution complexation models operate under the assumption of a reaction equilibrium between interacting species, rationalizing increases in solute solubility based on the effective removal of that solute during the formation of its intermolecular complex[12,15]. Comparative results in this chapter demonstrate that the impact of DOTG on BA crystallization inhibition is much more pronounced that what is expected from a solute removal mechanism, indicating that there is likely a complex set of molecular interactions in solution that kinetically impede the generation or ordering of crystal nuclei. The exact nature and mechanism of these interactions remain unknown and continue to be a challenging fundamental question, tied to the equally complex and poorly understood mechanisms that govern nucleation itself. Other work has postulated that the presence of tailored complexing agents disrupts the formation of molecular concentration clusters[21], or alters solute diffusivity[90,91], hence disrupting the ordering of molecules into crystal nuclei. Although plausible, these hypotheses remain speculative and difficult to mechanistically validate through experiments.

The findings summarize in this chapter elucidate that tailored intermolecular interactions can be used in order to control the onset of nucleation. These results demonstrate that small-molecule additives have potential as precipitation inhibitors, providing a vast set of molecular tools for further crystallization control in industrial, chemical and pharmaceutical applications.

## 4.5 Appendix

### Induction Time Measurement Theory

The probability calculations follow the theory described in the work of Jiang and ter Horst[88]. The probability of $m$ crystal to form ($P_m$) at a given time $t$ is given by the Poisson distribution shown in equation 4-3.

$$P_m = \frac{N^m}{m!} \exp(-N) \qquad \qquad 4\text{-}3$$

Where $N$ is the average number of crystals that form after the specified time $t$ has elapsed. The value of $N$ can be defined using equation 4-4 shown below.

$$N = JVt_j \qquad \text{4-4}$$

Where $t$ is the elapsed time, $J$ is the average nucleation rate (number per volume per time) and $V$ is the volume of the sample solution.

The probability of <u>no</u> crystals to have formed at a given time $t$ is easily defined using equation 4-3, as shown in equation 4-5:

$$P_0 = \exp(-N) \qquad \text{4-5}$$

Equation 4-5 can be used to define the probability of any crystal to have formed at a given time $t$ ($P^*$), defined in equation 4-6:

$$P_{>0} = P^* = 1 - \exp(-N) \qquad \text{4-6}$$

Substituting equation 4-4 and taking a natural logarithm, equation 4-6 can be easily transformed into a linear relationship with a slope of " $-JV$ ", which is the average nucleation rate per time of the system. This linearized expression is shown in equation 4-7.

$$\ln(1 - P^*) = -JVt_j \qquad \text{4-7}$$

$P^*$ determined experimentally, then, using equation 4-7 the average nucleation rate can be calculated. The inverse of the value $JV$ has units of time per number, and it is representative of the induction time of the system in static set of conditions. The value of $t_j$ is actually a combination of the actual induction time and the time delay required for the first observable crystal to grow. Hence, $t_j$ can be defined as shown below in equation 4-8:

$$\ln(1 - P^*) = -JV(t - t_g) \qquad \text{4-8}$$

For this work, experiments were designed in order to ensure that crystals appeared as soon as possible, after the desired supersaturation was safely established, minimizing the value of $t_g$. All experimental samples displayed a few crystallization events at early enough times to generate lines that fit well with an assumed value of $t_g = 0$. The interval of data collection of 2 min was also much larger that the time a crystal required to grow sufficiently to become visible, minimizing the error from this assumption.

**Equilibrium Chemistry**

The complexation of small molecules in solution has been commonly represented using simple equilibrium chemistry[12,15,65], as shown in equation 4-9. In this example $D$ refers to the limiting reagent, which is the complexing agent.

$$B + D \rightarrow BD \qquad \text{4-9}$$

Were $B$ and $D$ represent the complexing molecules, and $BD$ their complex. The equilibrium association reaction between these molecules can be represented by the equation 4-10.

$$K = \frac{[BD]}{[B][D]} \qquad \text{4-10}$$

Where $K$ is the equilibrium constant and the bracketed terms represent the concentrations of each of the chemical species.

In addition to the equilibrium expression 4-10, the mass balance of each of the chemical species in solution has to be satisfied, as shown in equations 4-11 and 4-12.

$$[B_0] = [B] + [BD] \qquad \text{4-11}$$

$$[D_0] = [D] + [BD] \qquad \text{4-12}$$

Where the values $[B_0]$ and $[D_0]$ correspond to the total amounts of each of the species present within the system. These quantities are known parameters, given that they are defined by how much of each chemical is used in a given experiment.

Case 1. – Strong Binding Constant K.

Re-arranging the mass balance expressions (4-11 and 4-12) to solve for $[BD]$ and substituting into equation 4-10, we obtain:

$$K = \frac{[BD]}{([B_0] - [BD])([D_0] - [BD])} \qquad \text{4-13}$$

$$[B_0][D_0] - [BD]([B_0] + [D_0]) + [BD]^2 = \frac{[BD]}{K}$$

$$0 = [BD]^2 - [BD]\left([B_0] + [D_0] + \frac{1}{K}\right) + [B_0][D_0]$$

$$[BD] = \frac{\left([B_0] + [D_0] + \frac{1}{K}\right)}{2} \pm \frac{\sqrt{\left([B_0] + [D_0] + \frac{1}{K}\right)^2 - 4[B_0][D_0]}}{2} \qquad \text{4-14}$$

For the particular case where $1/K$ is orders of magnitude smaller than the concentrations of the solutes, the following simplification can be made to 4-14:

$$[BD] \approx \frac{([B_0] + [D_0])}{2} \pm \frac{\sqrt{([B_0] + [D_0])^2 - 4[B_0][D_0]}}{2} \qquad \text{4-15}$$

4-15 leads to the following results:

$$[BD] = \frac{([B_0] + [D_0])}{2} \pm \frac{\sqrt{([B_0] - [D_0])^2}}{2}$$

$$[BD] = \frac{([B_0] + [D_0])}{2} \pm \frac{([B_0] - [D_0])}{2}$$

$$[BD]_+ = [B_0]$$

$$[BD]_- = [D_0]$$

Since $[B_0] > [D_0]$, it is impossible to generate $[BD]_+$, meaning that the concentration of the complexation has to be $[BD]_-$ and approach the value of $[D_0]$, which is the limiting reagent.

Case 2. – Excess of solute B

When the solubility of solute B is in excess, we can make a simplifying assumption for equation 4-11, as shown below in 4-16:

$$[B_0] \approx [B] \qquad \text{4-16}$$

Mainly, this expression states that given the excess of B, the total amount of complex $BD$ that can form is relatively small compared to the total concentration $[B_0]$. As such, the overall mass balance can be simplified into 4-16, stating that the concentration of B will remain overall relatively constant. If we re-arrange 4-16 and 4-12 to solve for $[BD]$, substitute into 4-10, and solve for $[D_0]$ we obtain:

$$K = \frac{[BD]}{([B_0])([D_0] - [BD])} \qquad \text{4-17}$$

$$\frac{[D_0] - [BD]}{[BD]} = \frac{1}{K[B_0]}$$

$$\frac{[D_0]}{[BD]} - 1 = \frac{1}{K[B_0]}$$

$$\frac{[D_0]}{[BD]} = \frac{1}{K[B_0]} + 1$$

$$\frac{[D_0]}{\frac{1}{K[B_0]} + 1} = [BD]$$

$$\frac{[D_0]K[B_0]}{K[B_0] + 1} = [BD] \qquad \text{4-18}$$

If we consider that under the current assumptions $K[B_0] \gg 1$, we observe:

$$[D_0] \approx [BD] \qquad \text{4-19}$$

Once again, the calculation indicates that the concentration of complexing agent and of total complex will approach each other, leading to the experimentally observed direct molar correspondence between increase in solubility and complexing agent addition.

**Powder X-Ray Diffraction of Microscope Samples**

Powder X-Ray Diffraction (PPXRD) was performed on all samples using a PANalytical X'Pert PRO diffractometer at 45 kV with an anode current of 40 mA. The instrument has a PW3050/60 standard resolution goniometer and a PW3373/10 Cu LFF DK241245 X-Ray tube. Samples were placed on a spinner stage in reflection mode. Settings on the incident beam path included soller slit 0.04 rad, mask fixed 10 mm, programmable divergence slit and fixed 1° anti-scatter slit. Settings on the diffracted beam path include: soller slit 0.04 rad and programmable anti-scatter slit. The scan was set as a continuous scan: 2θ angle between 4° and 40°, a scan speed of 0.021989 °/s, and a step size of 0.0167113°.

The PXRD was taken for solid samples recovered from microscope experiments with high complexing agent concentrations (0.05 mg/ml DOTG). The obtained crystals were crushed using a mortar and pestle in order to maximize the number of visible crystal faces. The experimental PXRD is overlay and compared with the only known crystal structure PXRD for pure benzoic acid reported within the Cambridge Structural Database (CSDB). The reference structure in the CSDB corresponds to BENZAC01, published by *Bruno et Al*[95].

The data below demonstrates that the observable peaks in the experimental pattern match those of the reference from the CSDB, indicating that the crystallized BA correspond to the known and expected structure, and not to a co-crystal of DOTG.
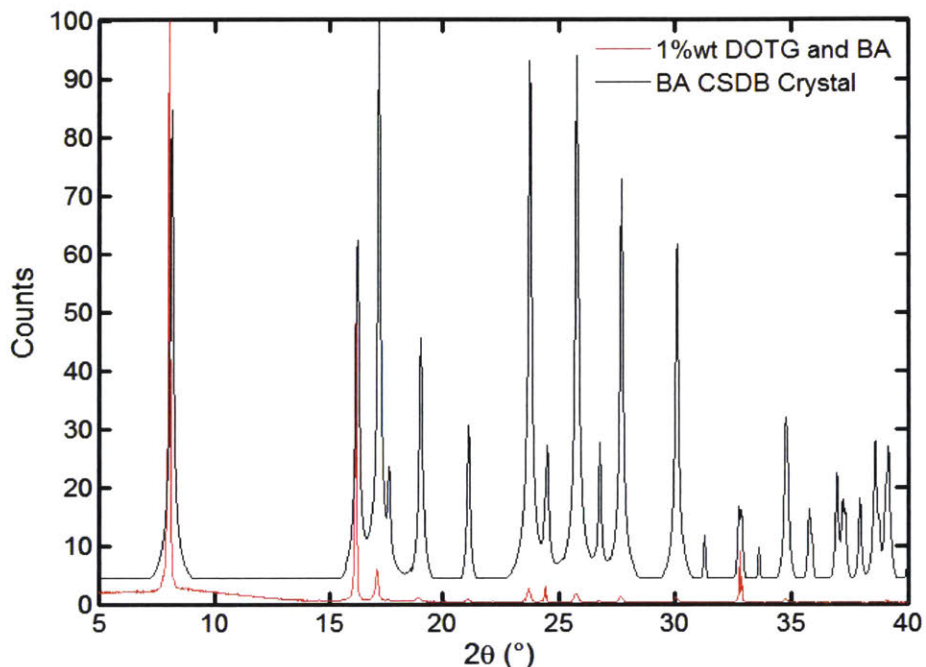


**Figure 4-11:** PXRD overlay of experimental (red) BA sample obtained from induction time experiment with complexing agent present, and reference (black) powder pattern obtained from the Cambridge Structural Database for the only known polymorph of Benzoic Acid.

**Linear Growth-Rate Estimation**

The growth rate of BA crystals in the induction time experiments was estimated using the imaging capabilities of the experimental setup employed with the Zeiss Axio Observer microscope. The microscope's software (AxioVision Version 4.8.2.0) can be used to measure details on images taken during the experiments, as shown in the example figures below.

All crystals monitored consistently displayed crystal growth rates that ranged from 30-40 μm/min. The image analysis shown in the pictures above was carried out for multiple crystal samples obtained at varying concentrations of DOTG. The length of the monitored dimension was plotted as a function of time for each picture analyzed. All samples showed a linear relationship between size and time. The slope of the lines obtained was taken as the measured growth rate. A few examples of such calculations are shown in the figures below.

74
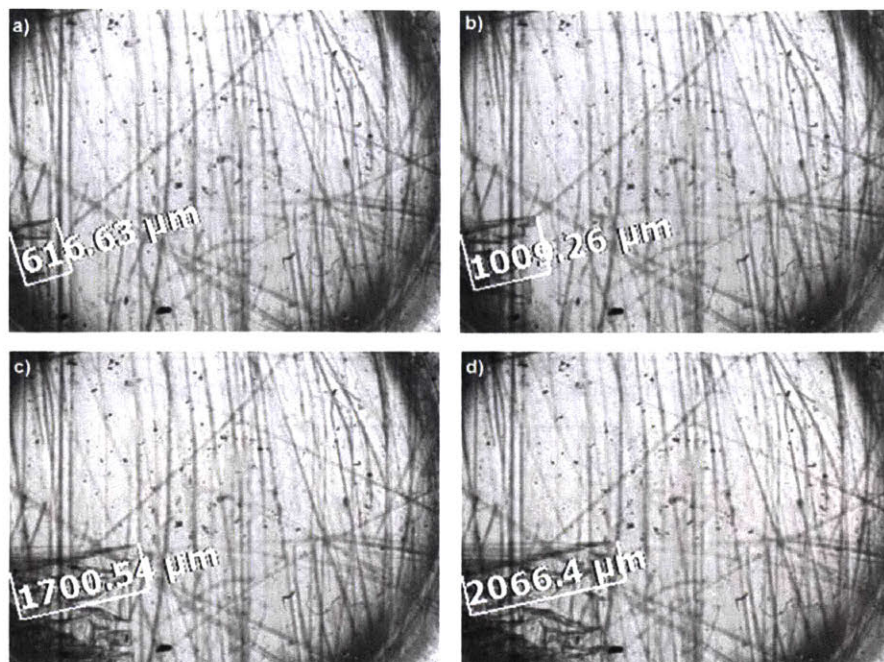
**Figure 4-12:** Demonstration of crystal growth monitoring. The measurements were taken directly from the AxioVision software. The pictures correspond to the same crystal of a control sample of pure BA: a) 18 minutes after induction, b) 28 minutes after induction, c) 48 minutes after induction, d) 58 minutes after induction. The diagonal dimension of the plate crystal was selected because it was the fastest growing, it was easily observed in the pictures and was common to all experiments, usually appearing parallel to the bottom of the experimental vials as shown in this set.
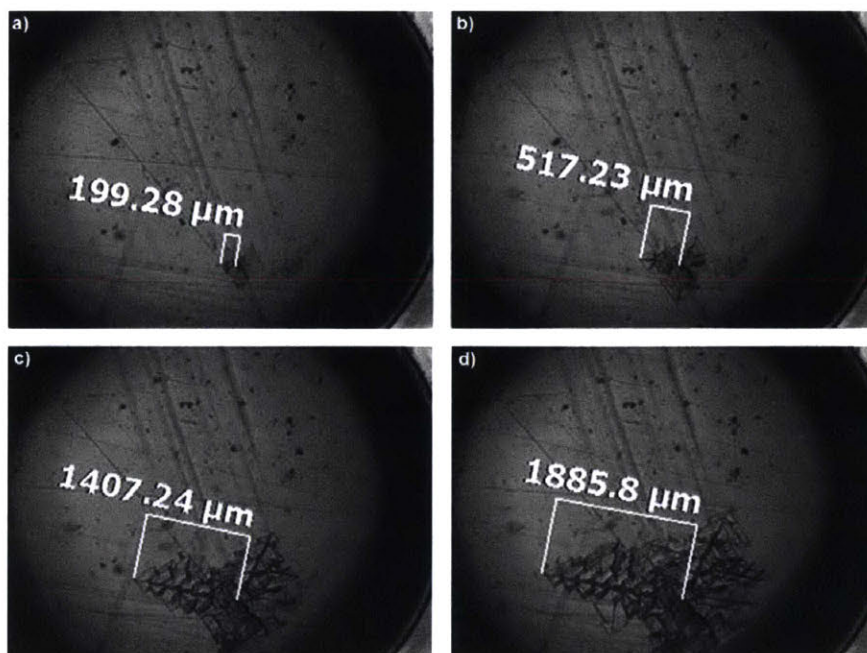


**Figure 4-13:** Second example of crystal growth monitoring. The images were taken directly from the AxioVision software. The pictures above correspond to the same crystal sample of BA with 0.025 mg/ml DOTG and adjusted supersaturation to ensure equivalent conditions to the control: a) 6 minutes after induction, b) 16 minutes after induction, c) 36 minutes after induction, d) 46 minutes after induction.
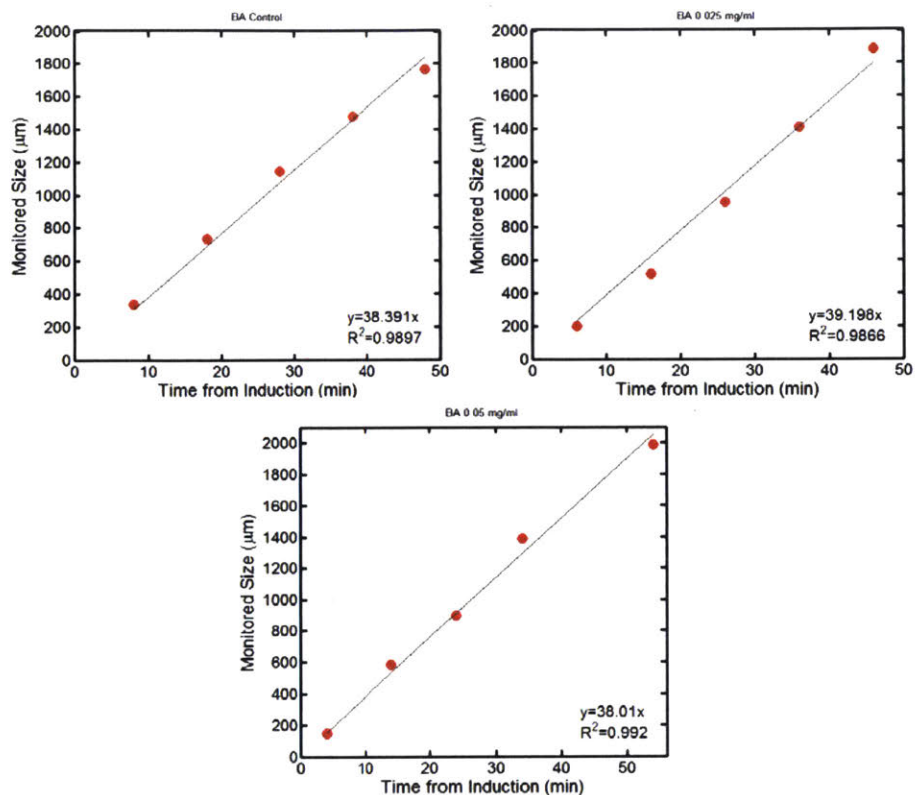
**Figure 4-14:** Example comparison of measured linear growth rates for crystal samples obtained with different quantities of complexing agent in the microscope experiment. For the data shown, samples with DOTG were adjusted to ensure they had the same supersaturation as the control.

The results shown are key in justifying two underlying assumptions for the experimental work and analysis displayed in this chapter:

1. The growth rate of the crystals is sufficiently fast so that the $t_g$ term in the induction equation (Equation 1 in the main manuscript) can be neglected. The interval between pictures taken in the microscope experiments was 2 minutes, which means that after induction, a crystal can be expected to grow 60-80 μm in its largest dimension between samplings. This growth is more than sufficient for detection due to the magnification capabilities of the microscope setup, which allows confidently detection of a crystal as small as 40 μm in all dimensions. Since the time required for a crystal to grow to a visible size is shorter than the time interval between pictures, the time of the first picture with an observable crystal is the most accurate estimation possible of induction time without requiring any correction or adjustment given the experimental setup.

2. The growth rate of crystals was unaffected by the presence of complexing agent at the experimental concentrations explored in this project, indicating that changes in the measured slopes of the induction probability curves were caused by nucleation rather than growth inhibition.

# Chapter 5 : Nucleation inhibition by a dilute, weakly interacting impurity

The contents of this chapter were adapted from work published in a peer-reviewed journal article. The paper has been recently accepted online[96]: Pons Siepermann, C. A.; Myerson, A. S. Inhibition of nucleation using a dilute, weakly hydrogen-bonding molecular additive. *Cryst. Growth Des.* **2018**, *published online*, **DOI:** 10.1021/acs.cgd.8b00367.

## 5.1 Introduction

Recently developed high-throughput experimental methods have led to a breakthrough in the ability to measure nucleation rates of chemical species in a controlled environment. The use of large sample sizes, specialized microscopy, and cycling crystallization setups with Crystal16 units have allowed for the development of large data sets, which allow for the determination of the probability distributions that characterize the naturally stochastic nucleation processes[81,88,89,92,94,97–99]. Resulting from this methodological innovation, many groups have been able to systematically and quantitatively study nucleation rates. This has allowed for the analysis of nucleation effects caused by controlled heterogeneous surfaces and substrates[81,94,97,98], of changing solvents[92], and of polarized light pulses[99]. Similarly, the analysis of induction data has indicated[39,89,94] that nucleation often does not proceed through a one-step classical theory mechanism, but rather through a two-step nucleation process, in which solutions first form droplets where the solute is in high-concentration clusters, which then experience an ordering step, generating the crystal phase[1,3,34,39]. Despite the progress achieved, the understanding of the effect of impurities on nucleation is limited, and although it has been reported[4,51,100], to the knowledge of the authors, it has not been explored systematically through nucleation rate measurements until recently[52].

Classical nucleation theory cannot predict the effect of impurities on nucleation rates, and it has been debated that the presence of small concentrations of an impurity or additive cannot lead to the true inhibition of nucleation[22], despite the existence of evidence to the contrary[51,52,100]. In recent work from our group[52], it was demonstrated that the presence of highly dilute complexing agents could lead to substantial increases in the induction time of a small-molecule organic solute. Furthermore, it was demonstrated experimentally that the nature of the inhibition was driven mostly through an unknown kinetic mechanism, rather than by changes in supersaturation, chemical speciation, or thermodynamics.

In the work herein, the findings on nucleation inhibition of small-molecule organics are expanded further by exploring the effects of a weakly-interacting inhibitor additive on the nucleation rates of 3-nitrophenol (3NP). The chosen partner molecule for inhibition was 3-aminobenzoic acid (3ABA), which forms a 1:1 hydrogen bonded complex with the 3NP solute, as demonstrated in Chapter 2[70]. Besides measuring the change in nucleation rates upon the addition of the 3ABA inhibitor, the strength of interaction between the molecules was measured utilizing isothermal titration calorimetry (ITC) to elucidate the chemistry between the species in solution and rationalize the mechanistic effects of the additive on the solute. A solubility curve was created to determine the effect of the 3ABA additive on the saturation solubility of 3NP at constant temperature. The solubility analysis was necessary to account for any potential changes in supersaturation that could arise due to the formation of a dissolved complex[8,12,15,52] and therefore compensate for such changes in nucleation rate measurements. The purity, crystal growth rate, and polymorphism of the 3NP obtained in the nucleation rate experiments were monitored to establish the potential effects of the complexing agent on these parameters, which could affect in turn the

induction measurements. Finally, the relationship between nucleation kinetics and supersaturation of 3NP at constant temperature was measured in the presence and absence of 3ABA. The experimental data was correlated with the functional form proposed by the classical nucleation theory[36]. The classical model was found to be inadequate to mechanistically rationalize the observed experimental results. As a result, an alternate analysis of the nucleation kinetics was done using a correlation proposed by Vekilov[37,38] based on the two-step nucleation model. The two-step nucleation rate model was found to be able to successfully rationalize the experimental observations. A kinetic mechanism of inhibition was proposed based on the interference of the ordering step during the formation of a nucleus, caused by the presence of the molecular complexes formed between the crystallizing solute and the complexing agent additive.

## 5.2 Experimental Section

### Materials

3-nitrophenol, 3-aminobenzoic acid, trifluoroacetic acid, anhydrous toluene, HPLC water and HPLC methanol were purchased from Sigma-Aldrich (St. Louis, MO). Ethanol (200 proof) was purchased from VWR International (Edison, NJ). All chemicals were used as received. Microscope 1 mL septum cap vials were purchased from Fischer Scientific LLC (catalog #033752G). Filter paper was purchased from VWR International (Edison, NJ).

### High-Performance Liquid Chromatography

For 3NP and 3ABA detection, the HPLC instrument (Agilent 1260) was equipped with a UV-diode array detector (DAD). The column used was a YMC-Pack ODS-A 150 mm X 4.6 mm i.d. packed with 3 μm particles with 12 nm pore size (YMC America Inc.). The detection wavelength was set at 230 nm for both species. The samples were analyzed using an isocratic method with a 55:45 (volume) water:methanol mobile phase containing 0.3 %vol. trifluoroacetic acid, with 10 μl sample injections, a 1 ml/min solvent flow rate, and a column temperature of 30 °C. All sample concentrations were measured against calibrations curves. The calibrations were performed by diluting known amounts (by mass) of the desired solute in a controlled volume of ethanol at room temperature. The curve area of the signal produced by each sample was plotted as a function of the known concentration of solute (mass per volume) of each calibration point. The resulting correlations were linear, following Beer's law, and are shown in the supplemental information.

### Powder X-Ray Diffraction

Powder X-Ray Diffraction (PXRD) was performed on all samples using a PANalytical X'Pert PRO diffractometer at 45 kV with an anode current of 40 mA. The instrument has a PW3050/60 standard resolution goniometer and a PW3373/10 Cu LFF DK241245 X-Ray tube. Samples were placed on a spinner stage in reflection mode. Settings on the incident beam path included soller slit 0.04 rad, mask fixed 10 mm, programmable divergence slit and fixed 1° anti-scatter slit. Settings on the diffracted beam path include: soller slit 0.04 rad and programmable anti-scatter slit. The scan was set as a continuous scan: $2\theta$ angle between 4° and 40°, a scan speed of 0.021989 °/s, and a step size of 0.0167113°.

### Isothermal Titration Calorimetry

ITC experiments were conducted using a TA Instruments Nano ITC calorimeter and followed a methodology similar to the one described by *Weber et al.*[66]. In an ITC experiment, the reference cell was filled with toluene and the sample cell was filled with 3ABA toluene solution (0.486 mM), while 10 μL aliquots of 3NP toluene solution (68.65 mM) were injected from the syringe every

300 seconds with continuous stirring at a speed of 250 rpm. The resultant heat output was subtracted from a blank injection of the same 3NP solution into pure toluene. The data were fitted to an independent binding model, assuming 1:1 stoichiometry using NanoAnalyze software. Given that the binding constants were relatively small, the experiments were carried out in triplicate to ensure reproducibility. The binding constants were measured at 25 °C as a starting reference and at 5 °C to match the nucleation experimental conditions.

## Molecular Species Selection

The work shown in Chapter 4 studied the effect of complexation on nucleation inhibition of benzoic acid (BA) in water, using 1,3-di-$o$-tolylguanidine (DOTG) as an additive[52]. This molecular pair is known to have strong intermolecular interactions between its guanidine and carboxylic acid functional groups[52,66]. These two moieties undergo a proton exchange in solution, which leads to a strong binding energy of $\Delta G > 20 \; kJ/mol$ between them. The strength of this complexation was verified using calorimetry[66]. To further help elucidate the underlying mechanisms that govern inhibition and to explore the universality of the results found in past nucleation inhibition observations[51,52,100], an alternative complexation motif was desirable for rigorous nucleation rate analysis. A new molecular pair was selected and it is shown in Figure 5-1, consisting of 3NP as the solute and 3ABA as the inhibitor additive. The complexation of these molecules was explored and verified in another study from our group[70], both through crystallization experiments and nanofiltration based on molecular size. The selected solvent for all experiments was toluene, which displays a steep solubility curve as a function of temperature for the selected 3NP solute.

**Table 5-1:** Calorimetry results for binding of 3NP with 3ABA in toluene. Results show a weak interaction dominated by strong solvent effects. As expected, the strength of interaction rises significantly as temperature decreases. $K$ is defined as the complexation equilibrium constant as shown in equation 4.

| Temperature (K) | $K$ | $\Delta G$ (kJ/mol) | $\Delta H$ (kJ/mol) | $\Delta S$ (kJ*mol$^{-1}$*K$^{-1}$) |
|---|---|---|---|---|
| 278 | $13.4 \pm 0.7$ | $-6.00 \pm 0.1$ | $-198 \pm 5$ | $-0.69 \pm 0.02$ |
| 298 | $6 \pm 2$ | $-4.3 \pm 0.8$ | $-224 \pm 24$ | $-0.74 \pm 0.08$ |
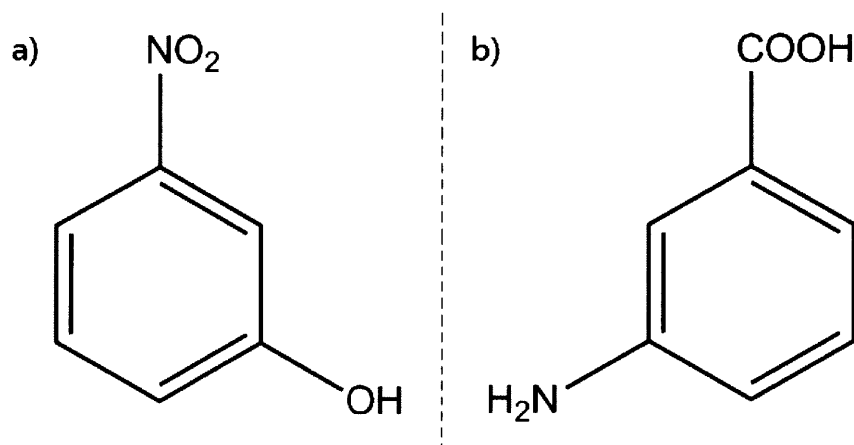


**Figure 5-1:** Structures of the nucleating target solute -3NP- (a), and the complexing agent -3ABA- (b).

The interaction between 3NP and 3ABA stems from the complementary hydrogen bond donating and accepting moieties of each molecule, which are spatially promoted to interact with each other due to their "meta" arrangement around their respective central rings[70]. The interaction between

these groups was verified with ITC and found to be substantially weaker than those between BA and DOTG, displaying an overall binding energy of $\Delta G \approx -6\ kJ/mol$. The calorimetry results from the complexation analysis are summarized in Table 5-1. Data shows the formation of a relatively weak complex (small values of $K$) with a highly exothermic enthalpy of binding, which is countered by strong entropic penalties.

It is likely that the polarized, hydrogen-bonding functional groups (nitro, hydroxyl, carboxylic acid, and amino) of the two complexing molecules are poorly solvated by the toluene, making their mutual interactions in an intermolecular bond highly energetically favorable. However, upon the formation of the complex, not only are the 3NP-3ABA molecules bound by an interaction (assuming a 1:1 complexation stoichiometry), but their functional groups are shielded by their aryl rings. This arrangement could allow for better solvation of the overall complex, potentially leading to the constraint of a greater amount of solvent molecules. The resulting solution would be substantially more ordered than the original uncomplexed parent states. The impact of the entropic penalties is further evidenced by the strong dependence of the overall binding constants to temperature, which are shown in Table 5-1 by the 120% increase in $K$ with just a 20 °C change in experimental conditions. This set of chemical interactions and molecular re-arrangements are difficult to fully confirm without the aid of extensive computational analysis; however, the hypothesized mechanism matches the experimental observations and provides a rationalization for the strong entropic losses shown by the calorimetric data. The contrast in complexation mechanisms and strength of interaction of 3NP-3ABA with that of past systems provides an important comparison that can be used to help further understand of how small impurity concentrations can affect the onset of nucleation. The impact of the change in molecular species is explored further in the discussion.

**Solubility Curve Determination**

Experiments were performed to determine the solubility of 3NP at a reference temperature of 5°C as a function of 3ABA concentration. This solubility curve is necessary to properly account for potential changes in supersaturation induced by the formation of a solution complex during the induction time measurements. Sample vials with excess 3NP were added to 10 ml of toluene with varying amounts of 3ABA, stirred at high temperature until complete dissolution of both solutes was achieved, and then cooled to 5°C. After crystals formed in the vials, the samples were stirred for 24 hours at a constant temperature to allow for equilibration. The supernatant from each equilibrated sample was then recovered, filtered using 0.2 μm Acrodisc Syringe Filters, immediately diluted by a factor of 20 (by volume) in ethanol, and analyzed using HPLC. The resulting concentrations of 3NP and 3ABA were used to develop a solubility curve.

**Design of Nucleation Experiments**

Measurements were carried out following a procedure similar to what is described in the literature for induction time determination[52,81,88,89,92,94,97–99]. A large set of identical nucleation experiments were performed at varying supersaturations of 3NP with and without the 3ABA additive to study the effects of the complexation on induction time kinetics. Samples containing supersaturated solutions of 3NP in toluene were prepared in 1 ml shell vials and placed in a Zeiss Axio Observer.Z1m Microscope in groups of 80 at a time. The microscope was equipped with a HAL100 light source and a Z800 Workstation. The samples were fixed in an unstirred stage with a hollow chamber connected to a recirculating chiller for temperature control. Temperature was rigorously maintained at 5°C for all experiments, and was regularly verified with a thermometer

by measuring the temperature of a reference vial filled with solvent placed in the stage itself. The microscope was programmed to continuously focus and scan across each of the vials and take a picture at 25X magnification, recording the state of the solution in every sample. Each vial had a picture taken every two minutes. The pictures were visually inspected to determine the timestamp in which the first crystal was observed. Based on the stamp record of the picture in which a crystal was first observed, it was possible to determine the time required for each crystallization event to take place. The distribution of induction times from a large data set of samples can be analyzed using equation 1 to determine the time-independent nucleation rate of the experimental system. The derivation and implementation of equation 5-1 has been demonstrated *Jiang et al*[88].

$$\ln(1 - P^*) = -JV(t - t_g) \qquad \text{5-1}$$

Where $P^*$ is the probability of any crystal to have formed in an experiment after an elapsed time of $t$. The value of $P^*$ is obtained directly from the distribution of experimental induction times. $V$ is the volume of the crystallizing solution, $J$ is the nucleation rate, and $t_g$ is a correction designed to account for the time required for a crystal to grow to a detectable size. The representative induction time of an experiment (expected period required for a crystal to form) is equal to $(JV)^{-1}$. The value of $t_g$ is commonly neglected given that it is usually very small compared to the time scale of $t$[81,92,97]. For the measurements in this chapter, the validity of this assumption was confirmed by utilizing the image-capture data from the nucleation experiments. The linear growth rate of the crystals were estimated by measuring the change in dimension of observed crystals as a function of time. The data was processed following the same methodology reported in the literature[52,97]. Analysis shows that, at the supersaturations employed for this investigation, the growth rate of crystals ranged from 20-30 μm/min. Given that data was collected in 2 minute intervals, a nucleated crystal would be expected to grow between 40-60 μm in between measurements. This measured growth is greater than the detection capabilities of a nucleus of the microscope (~30 μm). As a result, the value of $t_g$ is smaller than the data collection interval for these experiments, implying that it can be neglected without incurring errors that are significant with respect to the distributions of the data. A set of sample calculations for growth rates obtained from experimental data is shown in the supplemental information.

After experiments were completed, the vials were inspected to confirm that the solvent level had not noticeably changed. In addition, the mother liquor of samples that had not crystallized after the experimental time had elapsed was recovered as well, diluted in ethanol and analyzed in HPLC to ensure that the supersaturation was consistent with the beginning of the experiments. None of the experiments reported in this chapter displayed either an observable change in solvent level or a detectable change in supersaturation. The product crystals were filtered and collected from the mother liquor after experiments periodically for PXRD and HPLC analysis.

Induction time experiments carried out according to the procedure described in the methodology section have been known to exhibit a tapering-off effect in nucleation rate for long time scales[52,81]. This experimental behavior generates two separate nucleation regimes. The first is for early times, with relatively high crystallization rates that follow the time-independent Poisson distribution of equation 5-1. The second regime displays a significant tapering-off of observed nucleation events, eventually stabilizing towards a second, much slower crystallization rate. This behavior has been consistently seen for multiple experimental conditions[52,81,98] and has been attributed to limitations in the experimental setup, which lead to a change in the nucleation conditions for long experimental time-scales[52]. However, the presence of the well-behaved linear regime at early times

is ubiquitous and reproducible. As a result, the early-time behavior only will be analyzed and discussed within this chapter for all data sets. The corresponding nucleation rates obtained will be those considered as representative of a given set of experimental conditions. The full collection of the experimental data for all conditions tested, including the second crystallization regime, is shown in the supplemental information. Further discussion on the long-term behavior of the nucleation rates is also provided.

**Solution Preparation**

To minimize the influence of organic residues and dust particles in the 1 mL shell vials used in the nucleation experiments, all vials were washed and rinsed with 200 proof ethanol and dried at 80 °C. The solutions used for the induction time measurements were prepared in toluene with 3NP and 3ABA. The amounts of solute needed to achieve the target supersaturation of each experiment were measured in an analytical balance and dissolved in 60 ml of toluene on a hot stirring plate at 80°C. After full dissolution, the hot solution was filtered twice through qualitative filter paper and poured into 20 mL glass scintillation vials preheated in an 80°C VWR heat block. The 20 ml vials were also cleaned with ethanol and dried the same way as the Shell vials. The filtered solution was given 5 minutes in the scintillation vials to equilibrate at temperature on the heating block. Aliquots of 500 μl of the hot filtered solution were pipetted into the heated shell vials and immediately capped to prevent evaporation. After capping, the samples were left for 10 minutes on the heating block to equilibrate to temperature, and were then rapidly transferred to the microscope setup described earlier and monitored for two hours as specified.

**Supersaturation Selection**

Throughout this chapter, supersaturation will be reported in terms of a concentration ratio "$S$", expressed as the dissolved concentration of solute in an experiment, divided by the saturation concentration of the same solute at the crystallization temperature (5 °C), as defined in equation 5-2. While this is an imperfect representation of true supersaturation, $S$ will be employed for its ease of use, and because it is the variable used in the derivation of the nucleation kinetics equation 5-3[36] that has been most widely reported in similar high-throughput nucleation-rate measurement studies[36,88,101]. The quantitative insight that can be extracted from the data analysis performed with this equation is, in the opinion of the authors, mostly referential. Equation 5-3 is an expression based on classical nucleation theory, and it therefore does not properly capture the true mechanisms that take place on non-ideal systems that experience heterogeneous nucleation. Similarly, it ignores the subtleties of the more modern two-step crystallization theory[37,38,89], by assuming that a hard nucleus of a critical size is formed at the onset of nucleation, with identical properties and molecular arrangement to that of the final crystal, and with a well-defined solid-solute interface[36]. Nevertheless, the kinetic parameters ($A$ and $B$) in the classical nucleation rate equation do provide a mathematical basis for comparing the relative nucleation rates that are observed at varying crystallizing conditions. As a result, they will be used in this chapter as a tool to demonstrate the significance of changes in observed nucleation rates as a function of supersaturation upon the addition of a nucleation-inhibiting complexing agent.

$$S = C_{3NP}/C_{3NP}^* \qquad \qquad 5\text{-}2$$

$$J = ASexp(-B/\ln^2 S) \qquad \qquad 5\text{-}3$$

Where $C_{3NP}$ is the concentration of 3NP in solution, and $C_{3NP}^*$ is the saturation concentration of the solute at the crystallizing temperature. The values of $S$ used for the nucleation rate experiments

reported in this chapter were selected based on the limitations and requirements of the microscope setup used for data acquisition described earlier. Given that the vials are unstirred in the cooling stage of the microscope during the experiments, a relatively large value of $S$ is required so that the samples crystallize within a reasonable time frame. A set of trial experiments were carried out using only 3NP in toluene, at varying values of $S$. The results of these trials showed that for $S \approx$ 1.8 a significant portion (~40%) of the 80 vials of a given experiment would crystallize within the first two hours from the onset of supersaturation. At significantly higher supersaturations, a greater fraction of the vials would develop crystals. However, a problem would arise at these higher supersaturations, given that samples would begin crystallizing before the experiments were appropriately set-up and the target measurement temperature (5 °C) was reached. As a result, the selected reference supersaturation was chosen to maximize the amount of useful data that could be obtained with each experiment without compromising the possible onset of false positives and premature crystallizations. The experimental setup was designed to minimize the cooling time to achieve the target supersaturation; however, there are physical and practical limitations that cannot be avoided and force a bound for the upper value of supersaturation that can be worked with.

## Polymorphic Consistency Analysis

The potential effect of the complexing agent on the final product polymorphism was an important consideration for this project. It is a well reported fact that impurities can have significant effects on a crystal-product's solid form[1,7,27]. If the presence of the 3ABA additive induced a change in the 3NP polymorphism during the induction time experiments, any change in nucleation rates could potentially be justified by the modification of the type of crystal that is formed during the experiments. This is true given that different polymorphs can have different nucleation rates[1,2,102]. The consideration of polymorphic consistency is of particular importance when working with 3NP, which has two well-known polymorphs[46,77,103]. 3NP is enantiotropic, having a stable orthorhombic form at room temperature and a metastable monoclinic form[46,77,103]. The two polymorphs exhibit concomitance, which adds a further layer of complexity when tracking the exact purity of the crystals obtained during experiments. A calibration curve was developed using PXRD in order to determine the polymorphic purity of experimental samples obtained from nucleation rate experiments. The methodology employed in the creation of the calibration curves and the purification of the reference pure polymorphs is extensively discussed and explained in the supplemental information.

## 5.3 Results and Discussion

The solubility measurements are summarized in Figure 5-2. The data shows that the addition of 3ABA did not increase the solubility of 3NP significantly within the measurement error. The measured average solubility of 3NP in toluene at 5 °C was 11.05 mg/ml. Higher concentrations of 3ABA and their effect on 3NP were not explored because toluene became quickly saturated with 3ABA, and further additions of the additive led to co-precipitation, rather than co-dilution of 3NP and 3ABA.

The experimental results contrast starkly with the outcome seen in Chapter 4[52], where the addition of the complexing agent led to an equivalent molar increase of the solubility of the base solute. It is a known fact that complexation can have a plethora of effects on the solubility of interacting species[8,12,15]. Traditionally, the formation of strong intermolecular complexes has been associated with increases in solubility for both interacting species as a function of complex concentration[12,15]. In the case of the 3NP-3ABA pair, the ITC data showed that the intermolecular binding constant

between the molecules is extremely weak. It is possible that the low solubility of 3ABA in toluene, which likely occurs due to the rejection of the polar functional groups of 3ABA by the solvent, represents too much of a thermodynamic barrier to be overcome by the interaction with 3NP. This phenomenon would be particularly accentuated when heavy entropic penalties are incurred with complex formation, as was evidenced by the calorimetry data. The resulting chemical environment at high 3ABA and 3NP concentrations likely makes it thermodynamically unfavorable for a greater concentration of 3ABA additive to be solubilized, even with the formation of a complex.
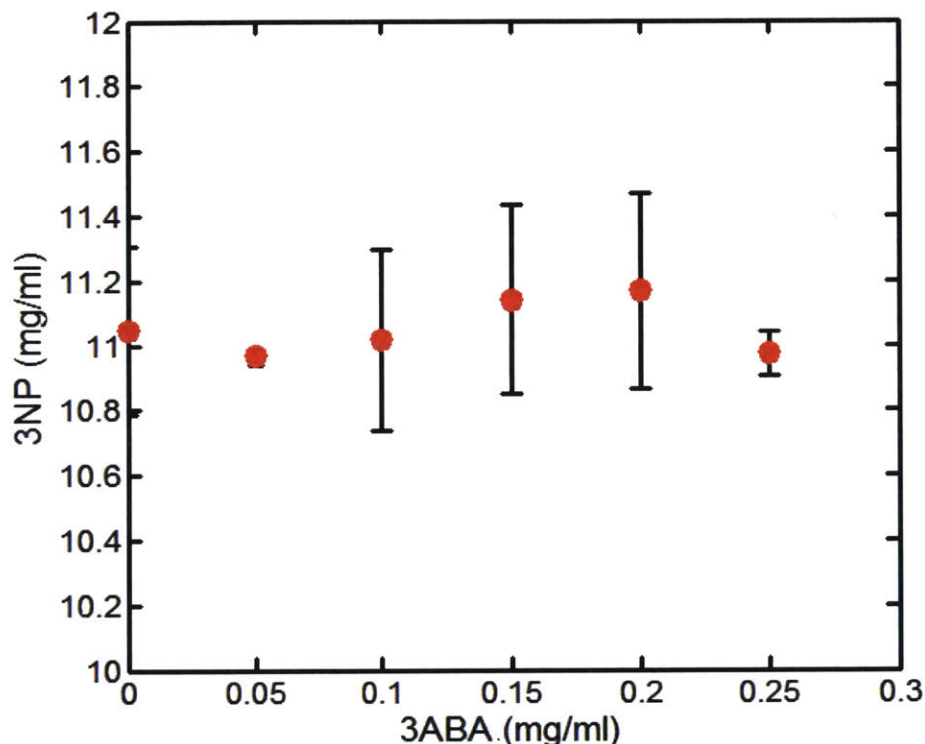


**Figure 5-2:** Solubility curve of 3NP as a function of 3ABA additive concentration. The data shows that within the measured experimental error, the additive has no statistically significant effect on 3NP solubility.

Despite the weak strength of binding between 3NP and 3ABA, a calculation can be made to estimate the equilibrium chemistry between the species. This analysis is necessary to determine if it is thermodynamically feasible for a complex to be formed under the experimental conditions employed during the crystallization experiments. Solution complexation is often approximated as a simple equilibrium process between the relevant species[8,12,15], with an equilibrium constant of $K$, as shown in equation 5-4. The term $C$ refers to the complex between the two parent species.

$$3NP + 3ABA \leftrightarrow C$$

$$K = \frac{[C]}{[3NP][3ABA]} \qquad \text{5-4}$$

In addition to the equilibrium, a mass balance around the total concentration of additive and solute must be satisfied, as shown in equations 5-5 and 5-6. In these equations, concentrations represented

with a zero subscript "$[\ ]_0$" indicate the total amount of chemical species added at the beginning of an experiment, while non-subscript brackets indicate the final concentrations after equilibration.

$$[3ABA]_0 = [3ABA] + [C] \qquad\qquad 5\text{-}5$$

$$[3NP]_0 = [3NP] + [C] \qquad\qquad 5\text{-}6$$

If equations 5-4, 5-5, and 5-6 are combined to solve for the total concentration of complex formed as a function of the value of $K$ and the initial solute concentrations, we obtain expression 5-7. A full algebraic derivation can be found in the supplemental information.

$$[C]^2 - [C]\left([3ABA]_0 + [3NP]_0 + \frac{1}{K}\right) + [3ABA]_0[3NP]_0 = 0 \qquad\qquad 5\text{-}7$$

All the terms in equation 5-7 are known constants or experimental variables, except for $[C]$. Equation 5-7 is a simple quadratic function, and the roots of $C$ can be found. The additive root is aphysical as it yields a value of $C$ greater than $[3ABA]_0$, hence it is discarded. The second root provides a key insight on the extent of complexation as a function of the experimental conditions and binding constants. The simple ratio defined in equation 5-8 represents the percent of additive that is effectively used in the formation of a complex.

$$\xi_C = \frac{[C]_-}{[3ABA]_0} \qquad\qquad 5\text{-}8$$

Where $[C]_-$ is the subtractive root for $C$ in equation 5-7. The concentration $[3ABA]_0$ in the nucleation inhibition experiments reported in this chapter is 0.05 mg/ml, the concentration of 3NP in the control groups is 20 mg/ml, and the value of $K$, as established in Table 5-1, is 12.9. The resulting calculation yields $\xi_C > 0.995$. The reason for this high extent of complexation is the large excess of 3NP with respect to 3ABA, which pushes the equilibrium correlation from equation 5-4 towards the complexed product. The relative inhibitor and solute concentrations used in Chapter 4 are comparable to the ones reported here[52]. The main difference lies in that the binding constant $K$ was several orders of magnitude greater[66]. However, the results shown in equations 5-7 and 5-8 indicate that under the high-excess conditions of crystallizing solute explored in these experiments, the relative strength of interaction has little relevance in the extent of complexation, as long as $K$ is greater than 1. This observation is of particular importance, given that beyond the controlled nucleation inhibition experiments discussed here, industrial crystallization processes can also develop small concentrations of an impurity, which can strongly impact the observed nucleation rates[5]. The presence of small concentrations of an interacting impurity represents a real challenge universally present within applied crystallization.

The results from the equilibrium calculations indicate that, despite the weaker strength of binding evidenced by the calorimetry data and dissolution experiments, the presence of the 3ABA additive at low concentrations leads to the formation of a complex with 3NP. Past results demonstrated that the presence of a complex, even in small concentrations[52], can lead to a strong kinetic inhibition of the nucleation rates of a small-molecule solute. Induction time experiments were carried out with 3NP and 3ABA, as detailed in the methods section. The exact experimental concentrations for the control groups (with and without inhibitor additive), and their respective nucleation rate and induction time results are summarized in Table 5-2. The data of probability of crystallization as a function of time, for the experimental sets shown in Table 5-2 are illustrated in Figure 5-3.

The data was processed and fit to a line as described by equation 5-1. The slope of the linear approximation to each data set is equivalent to $JV$, or $1/\tau$.

**Table 5-2:** Results of induction time experiments. $S$ is the concentration ratio as defined in equation 5-2, $C_{3NP}$ is the concentration of 3NP solute, $J$ is the measured nucleation rate, $\tau$ is the representative induction time, and $n$ is the number of samples used in the induction time experiments.

| Sample | $S$ | $C_{3NP}$ (mg/ml) | $J$ (# sec$^{-1}$ m$^{-3}$) | $\tau$ (min) | $n$ |
|---|---|---|---|---|---|
| Control | 1.810 | 20.00 | $367 \pm 15$ | $91 \pm 4$ | 559 |
| Control with 0.05 mg/ml 3ABA | 1.810 | 20.00 | $110 \pm 1$ | $303 \pm 3$ | 480 |
| 0.05 mg/ml 3ABA Equivalent | 1.805 | 19.95 | $300 \pm 10$ | $111 \pm 4$ | 640 |
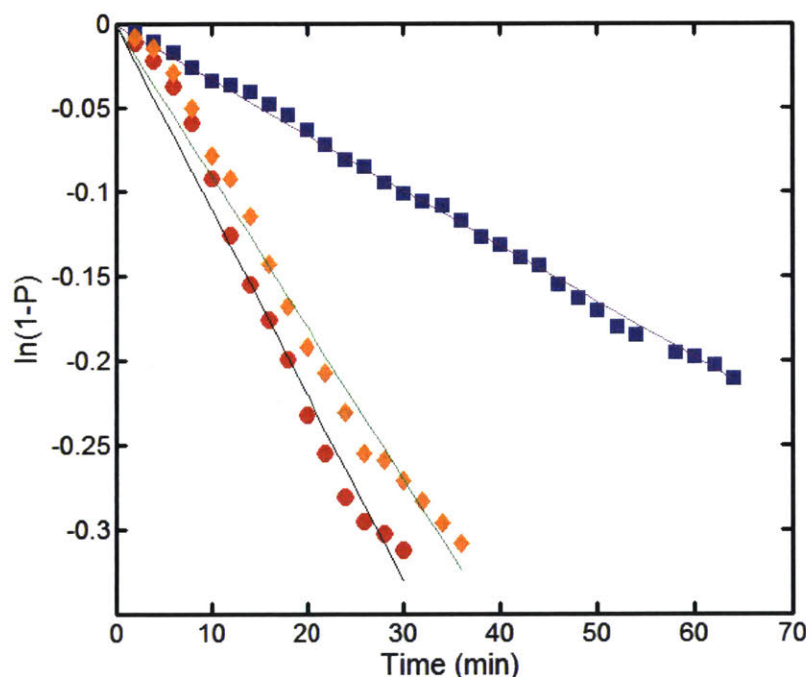


**Figure 5-3:** Comparison of the representative nucleation rate (slope) as a function of crystallization conditions: non-complexed control (red), complexed experiment of equivalent supersaturation to control (blue), and equivalence experiment for the complexed experiment (yellow).

The data in Table 5-2 and Figure 5-3 show that the presence of just a small concentration of 3ABA additive had a large effect on the measured induction time. The results show that with the inhibitor, there is a 233% increase in induction time with respect to the non-inhibited control, even at identical supersaturations. This substantial reduction in nucleation rate was achieved with the addition of only 0.05 mg/ml of 3ABA, which is equivalent to 0.25% of the amount of 3NP dissolved, both in terms of mass and molar concentrations (given the species' similar molecular weights). This observed trend is consistent with past results, obtained with a much more strongly complexing system[52].

The third data set shown in Table 5-2 and Figure 5-3, labeled "0.05 mg/ml 3ABA Equivalent", was obtained to verify the kinetic nature of the inhibition. If the inhibition took place through a thermodynamic mechanism, then the change of the nucleation rate upon the introduction of the 3ABA additive should proceed from a change in supersaturation state. This change would stem

from a reduction of the effective concentration of 3NP, resulting from the equilibrium speciation illustrated by equation 5-4. As a result, the equivalent theoretical "loss" in supersaturation should equate to the amount of complex formed, which would translate into a corresponding reduction in the concentration of free 3NP. Since the amount of 3ABA additive is known, and the total resulting concentration of complex formed is approximately the same as that of the additive (as shown with equations 7 and 8), it is possible to simulate the equivalent reduction in 3NP supersaturation at a given 3ABA concentration. This can be done by performing an experiment without the inhibitor, at a lower supersaturation of 3NP made to match the effective reduction in $S$ stemming from the complexation. The "equivalent" experiment was performed to test this speciation-based mechanism. It is clear from the data that the observed decrease in nucleation rate is not sufficient to account for the effect of the 3ABA inhibitor. As a result, the governing inhibition effect must be the product of a kinetic dampening of the nucleation process. This observation is consistent with past findings from our group[52], which systematically demonstrated that the small-molecule inhibition mechanism explored herein is kinetic and not thermodynamic in nature.

The purity and polymorphism of the crystals obtained for the control groups with and without complexing agent were tested to validate the data collected. HPLC showed no detectable amount of 3ABA on the crystals obtained from experiments containing the inhibitor. The PXRD analysis described in the methods section was employed in order to establish if the polymorphism of 3NP was affected by the presence of 3ABA. Figure 5-4 shows a comparison of the X-ray patterns obtained from the control experiments. Upon visual inspection, it is apparent that the resulting crystals are a mixture of the orthorhombic and monoclinic polymorphs of 3NP. The overall appearance of the patterns with and without complexing agent is identical, hinting that product polymorphism was unaffected by 3ABA. However, the PXRD patterns were integrated and analyzed using the calibration mentioned in the methods and explained in the supplemental information, in order to fully verify that the products were indeed consistent. The summarized average estimations from the calibrated analysis are shown in Table 5-3. Although there is a small difference in the calculated average fraction of orthorhombic polymorph present within the product crystals, the values obtained are well within the measurement error, as shown by the reported uncertainty ranges. Based on the data, here was no statistically significant difference in the product polymorphism due to the presence of 3ABA.

The data in Table 5-3 indicates that the crystals obtained in the nucleation rate experiments were predominantly monoclinic (metastable). This outcome is predictable, given that 3NP preferentially crystallizes in its monoclinic form from toluene. This observation was made during the pure polymorph screening performed as part of the PXRD calibration studies.

**Table 5-3:** Calculated polymorphic composition of the crystallized products from the nucleation rate experiments. Data shown for samples obtained at control conditions ($S$=1.810) with and without 3ABA. Polymorphism is reported in terms of the fraction of the total crystal mass represented by the stable orthorhombic polymorph.

| Condition | % Orthorhombic |
|---|---|
| No Additive | $17 \pm 3$ |
| 0.05 mg/ml 3ABA | $22 \pm 5$ |
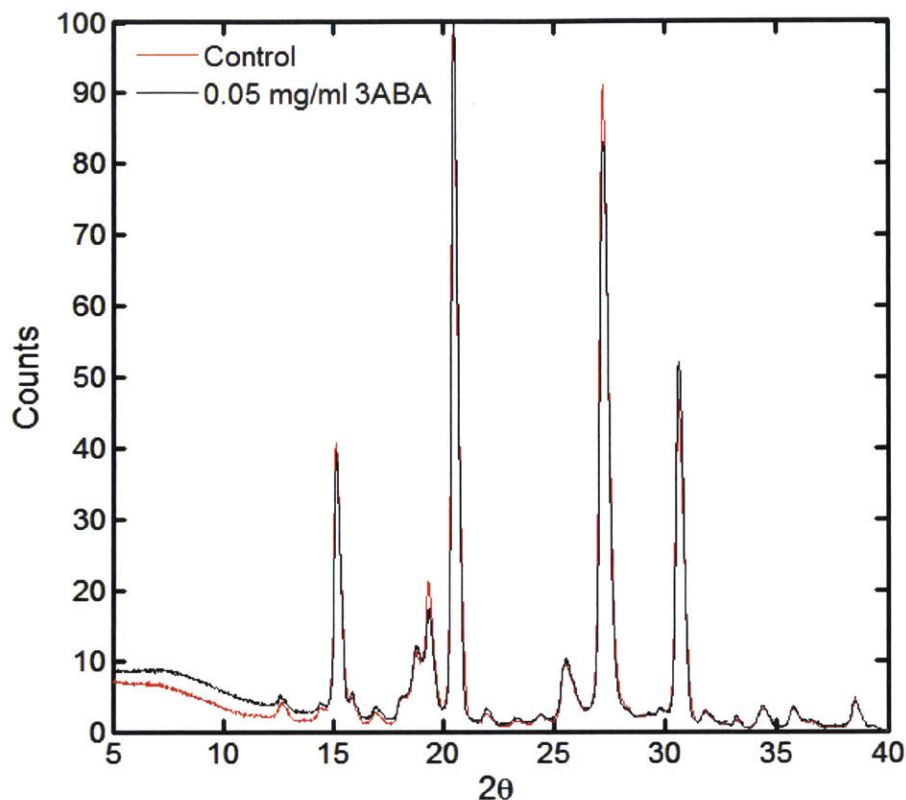
**Figure 5-4:** Overlay of PXRD patterns obtained from microscope nucleation rate experiments. The data shows that the presence of the 3ABA additive has no significant effect on product polymorphism. It is also apparent that the crystals obtained are a mixture of both metastable and stable polymorphs of 3NP.

The growth rate of 3NP was estimated using image analysis from the microscope data. The data showed that there was no significant difference in the observable linear growth rate in the presence of 3ABA. At the control supersaturation ($S$=1.810), the observable growth rate of 3NP was $27 \pm 4$ μm/min, while with complexing agent it was $27 \pm 1$ μm/min. Sample growth-rate calculations and data are provided in the supplemental information. The parity in growth rate between the control groups gives further indication that the shift in the experimental induction times is caused by nucleation inhibition, rather than growth inhibition.

In addition to demonstrating the inhibiting effect of 3ABA, the nucleation kinetics as a function of $S$ were explored for 3NP with and without complexing agent present. This analysis required measuring the nucleation rate of 3NP at varying supersaturations, and observing the change of $J$ as a function of $S$. The model traditionally employed for this correlation is equation 5-3, which despite its theoretical limitations, provides an acceptable reference for analyzing the impact on nucleation incurred by the 3ABA inhibitor. The data sets employed, the experimental conditions, and the general results are summarized in Table 5-4, Table 5-5, and Figure 5-5.

**Table 5-4:** Results of induction time experiments as a function of $S$ with no complexing agent present. Variables follow the same conventions as in Table 5-2.

| $S$ | $C_{3NP}$ (mg/ml) | $J$ (# sec$^{-1}$ m$^{-3}$) | $\tau$ (min) | $n$ |
|---|---|---|---|---|
| 1.817 | 20.08 | $497 \pm 36$ | $67 \pm 5$ | 400 |
| 1.810 | 20.00 | $367 \pm 15$ | $91 \pm 4$ | 559 |
| 1.805 | 19.95 | $300 \pm 10$ | $111 \pm 4$ | 640 |
| 1.800 | 19.90 | $260 \pm 10$ | $128 \pm 5$ | 400 |
| 1.791 | 19.79 | $203 \pm 5$ | $164 \pm 4$ | 240 |
| 1.780 | 19.67 | $163 \pm 5$ | $204 \pm 6$ | 240 |

**Table 5-5:** Results of induction time experiments as a function of $S$ with 0.05 mg/ml of 3ABA complexing agent present. Variables follow the same conventions as in Table 5-2.

| $S$ | $C_{3NP}$ (mg/ml) | $J$ (# sec$^{-1}$ m$^{-3}$) | $\tau$ (min) | $n$ |
|---|---|---|---|---|
| 1.855 | 20.50 | $233 \pm 10$ | $143 \pm 6$ | 400 |
| 1.840 | 20.33 | $170 \pm 8$ | $196 \pm 10$ | 320 |
| 1.825 | 20.17 | $133 \pm 4$ | $250 \pm 7$ | 400 |
| 1.810 | 20.00 | $110 \pm 1$ | $303 \pm 3$ | 480 |
| 1.795 | 19.83 | $100 \pm 2$ | $333 \pm 6$ | 320 |

The data shown in Table 5-4, Table 5-5, and Figure 5-5 demonstrate the expected behavior of monotonically decreasing nucleation rates with decreasing supersaturation. The nucleation rate information obtained was processed with equation 5-3, as shown in Figure 5-6. The resulting correlations provide good fits for the nucleation rate data and facilitated the estimation of the representative pre-exponential factor ($A$) and activation energy ($B$) constants for the systems with and without complexing agent. The fitted kinetic nucleation parameters are shown in Table 5-6, with their respective 95% confidence ranges.

**Table 5-6:** Results of kinetic analysis of the nucleation rate data. Parameters obtained from fitting experimental rates to equation 3. The reported range for the pre-exponential factor accounts for the error in ln(A), and represents the calculated 95% confidence interval.

| Condition | Average $A$ (# sec$^{-1}$ m$^{-3}$) | Range $A$ (# sec$^{-1}$ m$^{-3}$) | $B$ |
|---|---|---|---|
| No Additive | $4.79 * 10^8$ | $1.06 * 10^7 - 2.18 * 10^{10}$ | $5.1 \pm 1.3$ |
| 0.05 mg/ml 3ABA | $1.26 * 10^5$ | $7.67 * 10^3 - 2.07 * 10^6$ | $2.7 \pm 1.0$ |
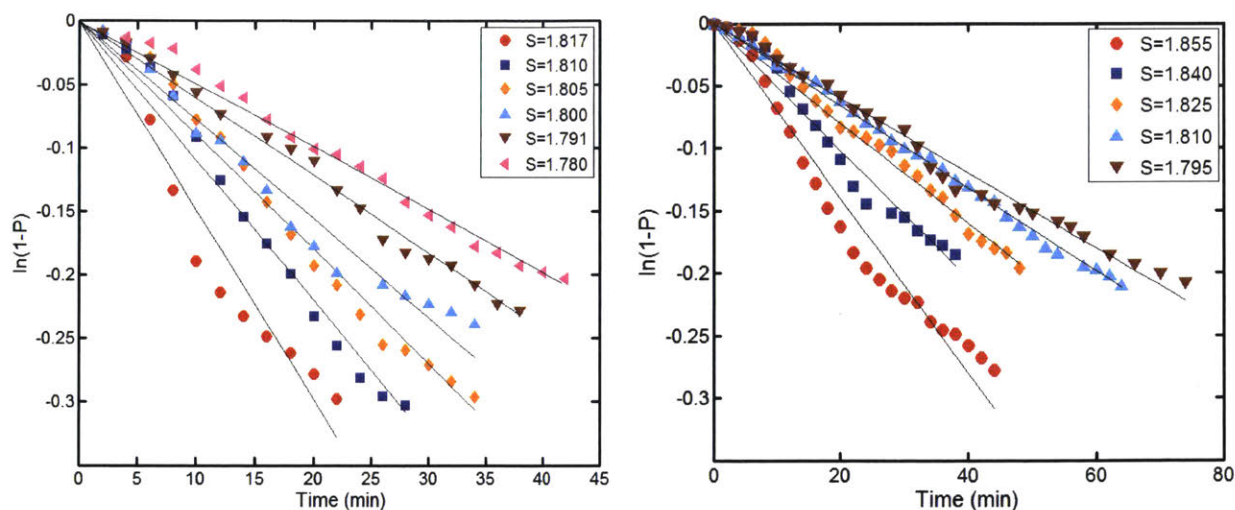
**Figure 5-5:** Results for nucleation rate experiments at varying supersaturations without complexing agent (left) and with 0.05 mg/ml of 3ABA added (right). Data shows a monotonic decrease in the slope (hence in nucleation rate) with decreasing supersaturation, as expected.
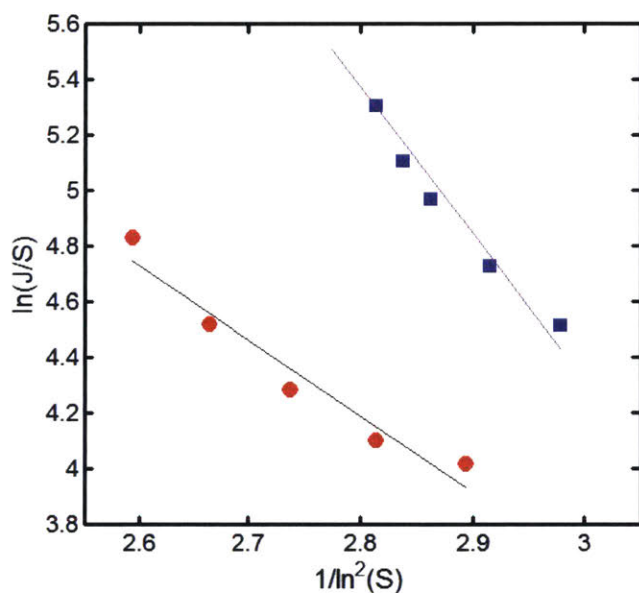


**Figure 5-6:** Nucleation kinetics as a function of concentration ratio ($S$), as predicted by the classical nucleation correlation. Nucleation rates are consistently slower upon addition of a small amount of a complexing agent that interacts through hydrogen bonding with the solute. Data shown for non-complexed samples (blue, $R^2$=0.9621) and samples with 0.05 mg/ml of 3ABA additive (red, $R^2$=0.9471).

The values shown in Table 5-6 indicate a reduction in both $A$ and $B$ parameters for the nucleation experiments that contained 3ABA. The decrease of both $A$ and $B$ implies that there might be competing effects taking place upon the addition of the 3ABA, given that a reduction in $B$ should theoretically increase nucleation rate. However, the calculated reduction in $B$ is only approximately 50% with respect to that of the non-complexed system. This difference in the activation parameter is undermined further when the uncertainty in the calculated variable is considered. Based on the experimental results, the change in $B$ is barely statistically significant. Meanwhile, the change in the pre-exponential parameter $A$ is striking, considering that there is a

90

decrease in more than three orders of magnitude to the average value of $A$ in the presence of 3ABA. Even when the confidence intervals are taken into account, the change in $A$ is well outside the overlap range of the values calculated with and without 3ABA.

Classical nucleation theory defines the pre-exponential parameter $A$ through equation 5-9[36]. In the expression, $v_0$ is the volume of the solute molecules in the crystal phase, $\gamma_e$ is the effective interfacial energy between the nucleus of a forming crystal and the solvent or heterosurfaces around it, $D$ is the diffusivity constant of the solute in the solvent system, $k_b$ is the Boltzmann's constant, and $T$ is the crystallization temperature. It is clear from the classical definition of $A$ that the presence of a weakly-interacting, dilute additive, such as the one explored in this chapter cannot be expected to reasonably reduce the pre-exponential factor by several orders of magnitude, as the experimental data shows. The only parameters that could be potentially affected by the inhibitor are the interfacial energy $\gamma_e$ and the diffusivity $D$. However, changes in $\gamma_e$ should be minimal, given that this parameter is changed by less than an order of magnitude by changing the solvent in which crystallization occurs[92], therefore a less than 1% concentrated additive should have negligible impact. Furthermore, $\gamma_e$ is heavily dominated by the presence of heterogeneous surfaces, such as container walls, which impact the nucleation rate far more than any dilute molecule could[2,36]. Similarly, it is unreasonable to expect that the diffusivity of the crystallizing solute would change by several orders of magnitude due to the presence of the dilute inhibitor, given that it has been shown for nitrophenols that even a change in the solvent system has less than an order of magnitude effect on diffusivity[104]. Based on this simple order of magnitude analysis, it becomes clear that classical nucleation theory cannot rationalize the dramatic change in nucleation rate and kinetics observed by the presence of dilute inhibitors and impurities.

$$A = \left(\frac{4\pi}{3v_0}\right)^{\frac{1}{3}} \left(\frac{\gamma_e}{kT}\right)^{\frac{1}{2}} DC^* \qquad \text{5-9}$$

The correlation proposed from equation 5-3 can be analyzed from the perspective of its merits as a kinetic expression, rather than focusing on the theoretical definition of the $A$ and $B$ parameters, which stems from classical nucleation theory[36]. In general terms, the expression is an Arrhenius-type model which accounts for an activation energy of crystallization and a supersaturation driving force. Alternative models of similar form have been proposed recently[37–39], attempting to capture the more nuanced transitions that take place in a two-step nucleation model. Vekilov et al.[37,38] developed an expression for the nucleation rate $J$, which considers the consecutive transitions and rates required from a homogeneous solution (state 0), to a high-density cluster (state 1), and finally to an ordered state crystal (state 2). The model has been successfully applied to protein crystallization; however, its applicability is believed to extend to other types of chemical systems[38]. The expression for the two-step rate model is shown in equation 5-10.

$$J = \frac{k_2 C_1 T \exp\left(-\frac{\Delta G_2^*}{k_b T}\right)}{\eta_1(C_1, T)\left[1 + \frac{U_1}{U_0}\exp\left(\frac{\Delta G_C^0}{k_B T}\right)\right]} \qquad \text{5-10}$$

The equation defines variables and transition rates in term of each of the three possible states the crystallizing system can take. The numerical subscripts refer to a specific state. $k_2$ is a kinetic constant that accounts for the transition into an ordered crystal state, $C_1$ is the concentration of the crystallizing solute in the high-density cluster phase, $\Delta G_2^*$ is the Gibb's free energy change required

to transition from the cluster phase to the ordered crystal, $\eta_1$ is the viscosity of the solution in the clusters, $U_1$ and $U_0$ are temperature-independent rate constants for the transitions between states 1 and 0 (forwards and backwards), and $\Delta G_C^0$ is the change in Gibb's free energy from the solution state to the high-density cluster intermediate. The expression for equation 5-10 can be simplified to equation 5-11, as shown below[38].

$$J = \frac{\phi k_2 C_1 T}{\eta} \exp\left(-\frac{\Delta G_2^*}{k_b T}\right)$$ 5-11

In equation 5-11, $\phi$ is described as the volume fraction occupied by the clusters[38]. A striking feature of the overall form of equation 5-11 is that it is extremely analogous to equation 5-3. The similarity between the expressions becomes starker if the expression is re-arranged to be a function of supersaturation ($\Delta\mu = \ln S$) at constant temperature. The exponential expression is completely analogous to the classical nucleation form, which is defined as shown in equation 5-12[36]. The values of all the parameters in the exponential term are constants at a fixed temperature, except for the supersaturation[36,38]. As a result equation 5-11 can be re-structured to equation 5-13.

$$\Delta G_2^* = \frac{16 v_0^2 \gamma^3}{3\Delta\mu^2} = \frac{B}{\Delta\mu^2} = \frac{B}{\ln^2 S}$$ 5-12

$$J = \frac{\phi k_2 C_1 T}{\eta} \exp\left(-\frac{B}{\ln^2 S}\right)$$ 5-13

The second approximation is made if we consider the definition of $\phi$[38], shown in equation 5-14:

$$\frac{1}{\phi} = 1 + \frac{U_1}{U_0} \exp\left(\frac{\Delta G_C^0}{k_B T}\right)$$ 5-14

The values of $U_1$ and $U_0$ are of similar order of magnitude[37]. $\Delta G_C^0$ is positive, as long as the high-density clusters are metastable with respect to the homogeneous solution phase. As a result, the expression $U_1/U_0 \exp(\Delta G_C^0/k_b T)$ should be greater than 1. Which allows for the approximation:

$$\frac{1}{\phi} \approx \frac{U_1}{U_0} \exp\left(\frac{\Delta G_C^0}{k_B T}\right)$$ 5-15

Given that $U_0$ is related to the rate of formation of high-density clusters in solution, and it is directly proportional to the solute concentration in the overall solution ($C$), the approximation of equation 5-16 is reached.

$$\frac{1}{\phi} \approx \frac{U_1}{\alpha C} \exp\left(\frac{\Delta G_C^0}{k_B T}\right)$$ 5-16

The parameter $\alpha$ is a constant to reflect the functional form of $U_0$. All the terms in equation are constant at fixed temperature. Given the definition of $S = C/C^*$, we can approximate:

$$\frac{1}{\phi} \approx \frac{U_1}{\alpha S C^*} \exp\left(\frac{\Delta G_C^0}{k_B T}\right) \approx \frac{A''}{S}$$ 5-17

Where $A''$ is lumped term that contains the saturation-independent constants. Combining equations 5-13 and 5-17, we obtain expression 5-18.

$$J \approx \frac{Sk_2C_1T}{\eta A''}\exp\left(-\frac{B}{\ln^2 S}\right)$$
5-18

All of the pre-exponential terms are independent of the concentration of solute in the homogeneous phase ($C$). Which indicates that at constant temperature, the following is true:

$$J = AS\exp\left(-\frac{B}{\ln^2 S}\right)$$
5-19

Where $A$ is a combination of all of the pre-exponential $S$-independent parameters, as defined in equation 5-20.

$$A = \frac{k_2\alpha C_1 T C^*}{U_1\eta_1\, exp\left(\frac{\Delta G_C^0}{k_B T}\right)}$$
5-20

The result of the algebraic analysis performed demonstrates that the two-step nucleation equation rate derived by Vekilov[37,38] can have the exact same functional form as the classical nucleation theory equation, as long as certain system conditions are satisfied and supersaturation is changed at constant temperature. This observation is in excellent agreement with the hypothesis that two-step crystallization can be generally applied to crystallizing systems, rather than just protein crystallizations[3,37,38,81]. Equation 5-19 indicates that the functional dependence of nucleation rate on supersaturation looks identical to the one predicted by classical nucleation theory, as long as the cluster formation is highly metastable and fast. When the clustering step is fast, the rate-limiting process becomes the ordering step into a crystal structure. The resulting process becomes dominated by this ordering step and the overall process appears as if it proceeded through a single classical-like mechanism.

The difference between equations 5-19 and 5-3, is the variables present within the pre-exponential terms, as defined in equation 5-9 and 5-20. Upon inspection of equation 5-20, it is clear that few of the parameters defined can account for the effect of a complexing agent. Given the extremely dilute concentration of the inhibitor, properties that are based on the equilibrium thermodynamics of the bulk solute, such as $U_1$, $\Delta G_C^0$, $\alpha$, and $C_1$, should remain mostly unaffected. This observation is very analogous to what was proposed for the diffusivity and interfacial tension in equation 5-9. Additionally, it has been demonstrated that the viscosity of the liquid phase in the clusters is dependent only on $C_1$ and $T$, which means that the complexing agent has no effect on this property[37,38]. This analysis of the variables implies that the major source for variability in the pre-exponential parameter due to the presence of impurities must come from the ordering constant $k_2$.

As discussed by Vekilov, the two-step equation is much more versatile than the classical one, and can predict non-ideal behaviors in nucleation rates as a function of supersaturation[37,38]. One particular irregular behavior that was observed was that the nucleation rate would change drastically for the crystallization of lysozyme upon the addition of small amounts (<5%) of the additives[37] PEG and glycerol. Upon analysis of the PEG and glycerol data using the two-step model (equation 10), it was observed that most nucleation kinetic parameters remained unchanged (viscosity, volume fraction, $C_1$), and that the thermodynamic energy barrier $\Delta G_2^*$ was either unchanged or only slightly affected. However, the kinetic order parameter $k_2$ drastically changed and explained almost the entirety of the change in measured nucleation rates. Due to the extremely low concentration of additives, the lack of change in the majority of the kinetic and thermodynamic parameters is logical for both the crystallization of lysozyme and for the inhibition experiments

reported herein: less than 1% of weakly-interacting additive should be expected to have minimal effects on the macroscopic properties of a saturated solution. However, the change in $k_2$ is related to the specific transition of solute molecules in a high-density cluster into an ordered crystal state. If we consider the fact that a critical cluster for small-molecules and proteins ranges from the hundreds to the thousands of molecules[105,106], then it is statistically likely for a complex or inhibitor molecule to be present during the ordering of such a cluster, as long as the complex has a concentration in the range of 0.1-1% of the crystallizing solute. We propose that the presence of an interacting additive can disturb and modify the transition and re-arrangement of the solute molecules into the crystal phase, modifying $k_2$, given that even a small disruption of the molecules in a birthing critical nucleus could potentially prevent its proper arrangement. In the case of nucleation promoters, it is likely that the additives help stabilize the solute molecules in their solvent environment, or that their interactions help orient the molecules towards the appropriate geometries of the final crystal form[37]. Conversely, in the case of molecular inhibitors, as the ones discussed herein and in Chapter 4[52], it is likely that the additives and the bimolecular complexes they form act as a jam towards the arrangement of solute molecules. This interference could happen either by physically disrupting the motion of the molecules towards a dense ordered state, where the complex acts as a bulky obstacle that impedes the required arrangement, or by the creation of intermolecular interactions that orient the solute structures unfavorably for the required final form. A schematic illustration of this hypothesized inhibition mechanism is shown in Figure 5-7.
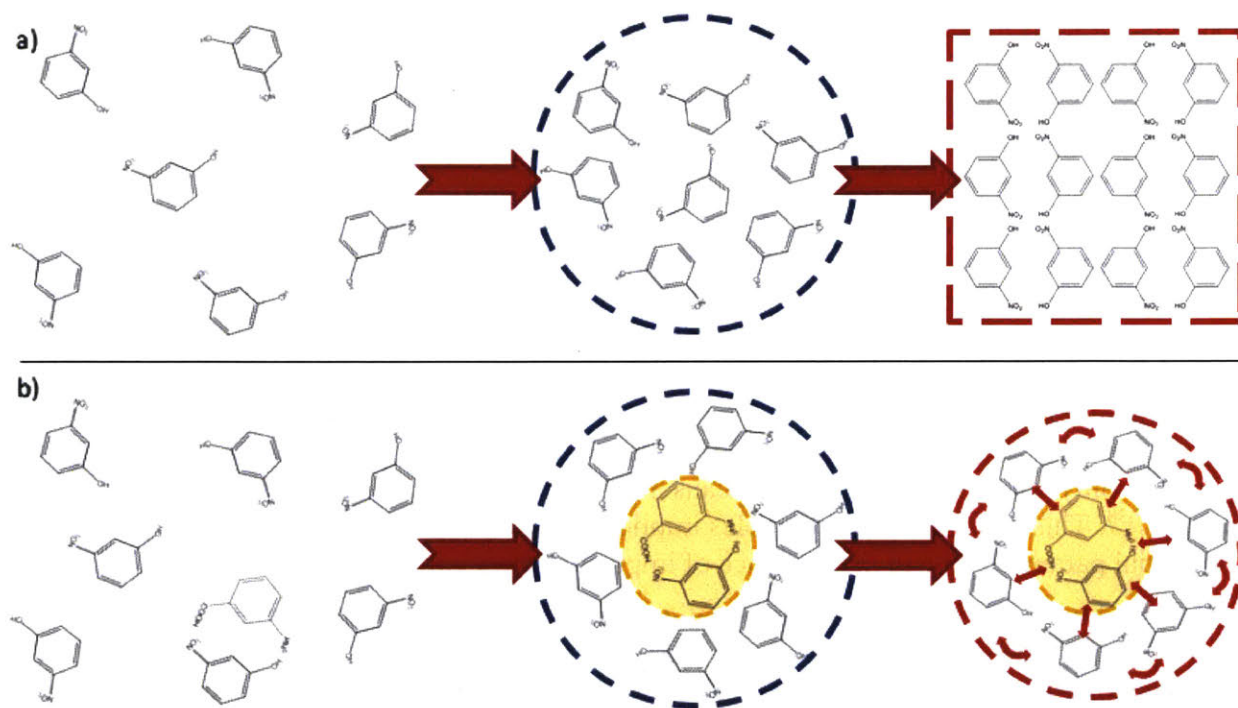


**Figure 5-7:** Proposed mechanism of inhibition caused by weakly-binding, dilute, small-molecule complexing agent, as illustrated by the 3NP-3ABA case study. a) Two-step crystallization model: the supersaturated solution of 3NP leads to the formation of a high-concentration cluster of molecules, which re-arrange themselves to form a crystal. b) Inhibited mechanism: the complex formed in solution is also present within the concentrated clusters, when the molecules attempt to re-arrange themselves, the presence of the bulky and geometrically disruptive complex prevents the ordering into a final crystal.

## 5.4 Conclusion

The experimental results for nucleation inhibition of 3NP using 3ABA verify past findings that the presence of a very dilute hydrogen-bonding complexing agent can lead to substantial reduction of the target solute nucleation rates. For the data shown in this chapter, it was observed that with only 0.25% by molar concentration of additive to solute, the measured induction time for crystallization increased by 233%. The presence of the complexing agent did not affect growth rate, polymorphism, or purity of the final product crystals, which demonstrates that the experimental results were indeed the product of true nucleation inhibition.

Calorimetry data, in conjunction with the nucleation rate measurements, showed that inhibition is not necessarily dependent on the strength of the binding between the inhibitor and the target solute. This is evidenced by the fact that the 3NP-3ABA molecular pair showed a comparable degree of inhibition to the BA-DOTG from similar past work, despite having a weaker intermolecular interaction in solution. The rationale for this equivalence despite weaker speciation and thermodynamics stemmed from a material balance, coupled with the equilibrium constants of binding of the interacting pairs. The calculations showed that due to the low concentrations of the complexing agents, the formation of a complex is extremely favored due to Le Chatelier's principle. As a result, most of the dissolved additive can be expected to form a bound complex with the target solute.

The nucleation rates of 3NP were measured as a function of supersaturation at constant temperature in the presence and absence of the 3ABA complexing agent. The resulting data was represented using the functional correlation proposed by the classical nucleation theory. The results showed that the pre-exponential factor rose dramatically, while the activation energy was minimally affected. The trends observed in the data were also discussed within the context of a two-step nucleation model, and an improved functional form of the nucleation rate equation was proposed, based on the model developed by Vekilov[37,38]. The two-step model was shown to be capable of predicting the functional form for $J$ as a function of $S$ proposed by classical nucleation theory, albeit providing an alternative definition for the relevant kinetic and pre-exponential parameters. Among those parameters was a kinetic ordering constant $k_2$, which had been observed in other studies[34] to change upon the presence of dilute additives, similar to the ones explored in this chapter. A mechanism was proposed to rationalize the observed nucleation inhibition of small molecules, in which the complexed additives form an obstacle in the pre-nucleus high-density clusters. This extends the rate-limiting step to nucleation, which is the re-arrangement of molecules from the liquid state to the ordered crystal form of a nucleus, therefore effectively reducing the ordering parameter $k_2$. This mechanism explains the fact that weakly-bound, dilute impurities can have a massive impact on nucleation rates, and accounts for the fact that due to the low concentrations of inhibitor, most of the kinetic and thermodynamic properties of the crystallizing solution should remain unaffected. It is the opinion of the authors that the results and hypothesis provided herein help elucidate the mechanism through which impurities can modify nucleation events in applied crystallization environments. Furthermore, the data provides evidence that further supports the existence of a two-step nucleation mechanism, even for small-molecule organics, given that the classical nucleation theory cannot predict the effect of kinetic inhibitors such as the ones explored in this chapter.

## 5.5 Appendix

### Equilibrium Speciation Derivation

The derivation is made by combining the equilibrium equation, with the molecular mass balances, shown below.
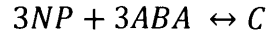
$$3NP + 3ABA \leftrightarrow C$$

$$K = \frac{[C]}{[3NP][3ABA]} \qquad \text{5-4}$$

$$[3ABA]_0 = [3ABA] + [C] \qquad \text{5-5}$$

$$[3NP]_0 = [3NP] + [C] \qquad \text{5-6}$$

The mass balances can be re-arranged to solve for the non-complexed concentrations of the two molecules. Then, the resulting equations can be substituted into equation 5-4. The resulting expression has only 1 unknown: the complex concentration $[C]$.

$$[3ABA] = [3ABA]_0 - [C] \qquad \text{5-21}$$

$$[3NP] = [3NP]_0 - [C] \qquad \text{5-22}$$

$$K = \frac{[C]}{([3ABA]_0 - [C])([3NP]_0 - [C])} \qquad \text{5-23}$$

We re-arrange equation 5-23 to solve for $[C]$:

$$K([3ABA]_0 - [C])([3NP]_0 - [C]) = [C] \qquad \text{5-24}$$

$$K([3ABA]_0[3NP]_0 - [C][3ABA]_0 - [C][3NP]_0 + [C]^2) = [C] \qquad \text{5-25}$$

$$[3ABA]_0[3NP]_0 - [C][3ABA]_0 - [C][3NP]_0 + [C]^2 = \frac{[C]}{K} \qquad \text{5-26}$$

$$[3ABA]_0[3NP]_0 - [C][3ABA]_0 - [C][3NP]_0 + [C]^2 - \frac{[C]}{K} = 0 \qquad \text{5-27}$$

$$[C]^2 - [C]\left([3ABA]_0 + [3NP]_0 + \frac{1}{K}\right) + [3ABA]_0[3NP]_0 = 0 \qquad \text{5-28}$$

Equation **5-28** is a quadratic expression for $[C]$, which can be solved by finding the roots, as shown below.

$$[C] = \frac{[3ABA]_0 + [3NP]_0 + \frac{1}{K} \pm \sqrt{\left([3ABA]_0 + [3NP]_0 + \frac{1}{K}\right)^2 - 4[3ABA]_0[3NP]_0}}{2} \qquad \text{5-29}$$

In equation 5-29, the additive root yields $[C] > [3ABA]_0$, which is aphysical, since it does not satisfy the mass balance of the system. As a result, the final expression for the concentration of complex $[C]$ is:

$$[C] = \frac{[3ABA]_0 + [3NP]_0 + \frac{1}{K} - \sqrt{\left([3ABA]_0 + [3NP]_0 + \frac{1}{K}\right)^2 - 4[3ABA]_0[3NP]_0}}{2} \quad 5\text{-}30$$

The expression can be solved, given that all the equation parameters are known. The value of $K$ was found using isothermal titration calorimetry, while the concentrations of the two interacting species are determined at the beginning of each experiment. As discussed in the main text, given the experimental conditions employed of $K = 12.9$, $[3NP]_0 = 20\ mg/ml$, and $[3ABA]_0 = 0.05\ mg/ml$, equation 5-30 yields $[C] = 0.0498\ mg/ml$, which demonstrates that most of the $[3ABA]_0$ is "consumed" in the formation of a complex. This high extent of complexation takes place due to the large excess of $[3NP]_0$, which drives the equilibrium of the interaction (equation 1) towards completion.

**Polymorphic Verification Calibration and Methodology**

As discussed in the main body, it was necessary to establish a calibration curve to determine the polymorphic composition of samples crystallized during the nucleation rate inhibition experiments. To generate this calibration curve, it was first necessary to obtain completely pure samples of both 3NP polymorphs. The monoclinic form was obtained by dissolving 16 mg/ml of 3NP in water at high temperature, then allowing the samples to naturally cool to room temperature with constant stirring. The samples were placed overnight in a chilled bath at 5 °C to maximize yield, and stirred for at least 12 hours. The final polymorphic purity was verified using PXRD and compared to a reference x-ray pattern obtained from the Cambridge Structural Database (CSDB). The code for the reference used to verify the monoclinic polymorphism was MNPHOL02 from the CSDB. The orthorhombic form of 3NP is much harder to completely purify, given that crystals obtained from most common bench solvents yield at least a small degree of monoclinic impurity. However, pure orthorhombic samples were obtained by dissolving 40 mg/ml of 3NP in chloroform at 45 °C, then instantly submerging the samples into a 5 °C bath with continuous stirring. Samples were allowed to equilibrate and age overnight for at least 12 hours. Polymorphism was verified as well through PXRD, against the CSDB reference MNPHOL03. A comparison of the purified crystal XRDs with their respective CSDB references are shown in Figure 5-8.

Once purified samples of each of the two 3NP polymorphs were obtained, mixtures of known proportions of each were prepared and analyzed with PXRD. The objective of this process was to determine how unique peaks changed as a function of the polymorphic composition of a mixture of the two forms. The two pure forms were weighed separately, mixed, and finally ground together with a mortar and pestle. The grinding process helped ensure homogeneity in the resulting powder, and eliminate any preferred orientation in the crystal faces that might have formed during the crystallization steps. A representative unique peak was selected for each polymorph and monitored as a function of composition. The two selected unique peaks are shown in. The unique peaks were selected to avoid any overlap with other peaks. For the monoclinic form, the unique peak from 12-14 2θ was chosen, while for the orthorhombic polymorph, the chosen peak was between 22.9-24.1 2θ, as shown in Figure 5-9.
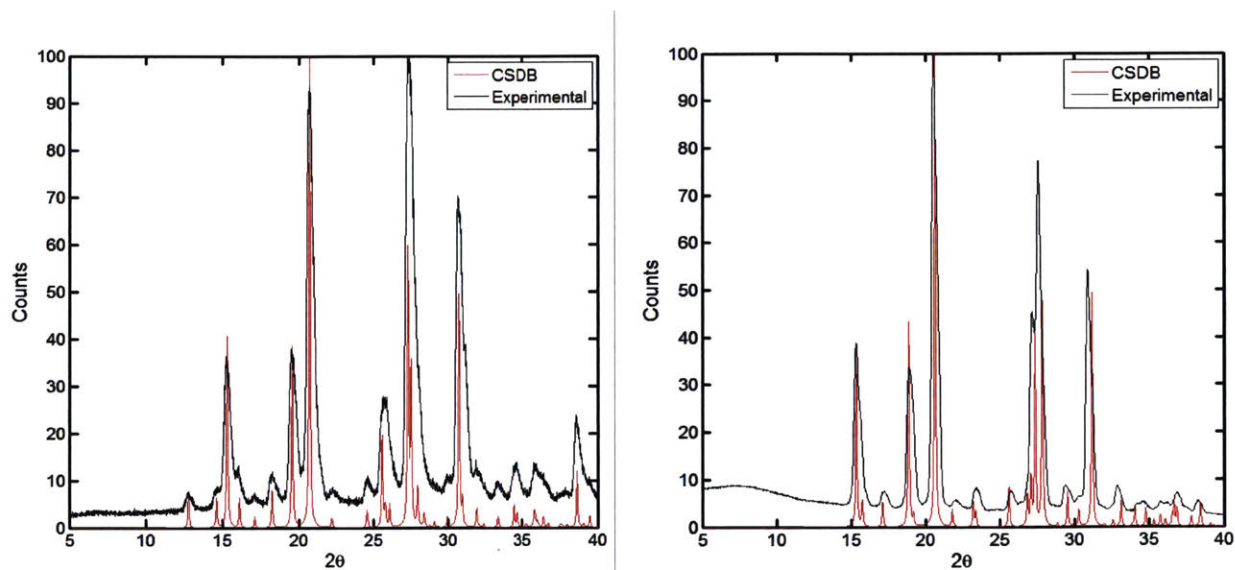
**Figure 5-8:** Comparison of the purified polymorphic forms of 3NP. The metastable monoclinic is shown to the left, while the stable orthorhombic is shown to the right. Both images demonstrate the experimentally obtained samples (black) with the CSDB references (red).
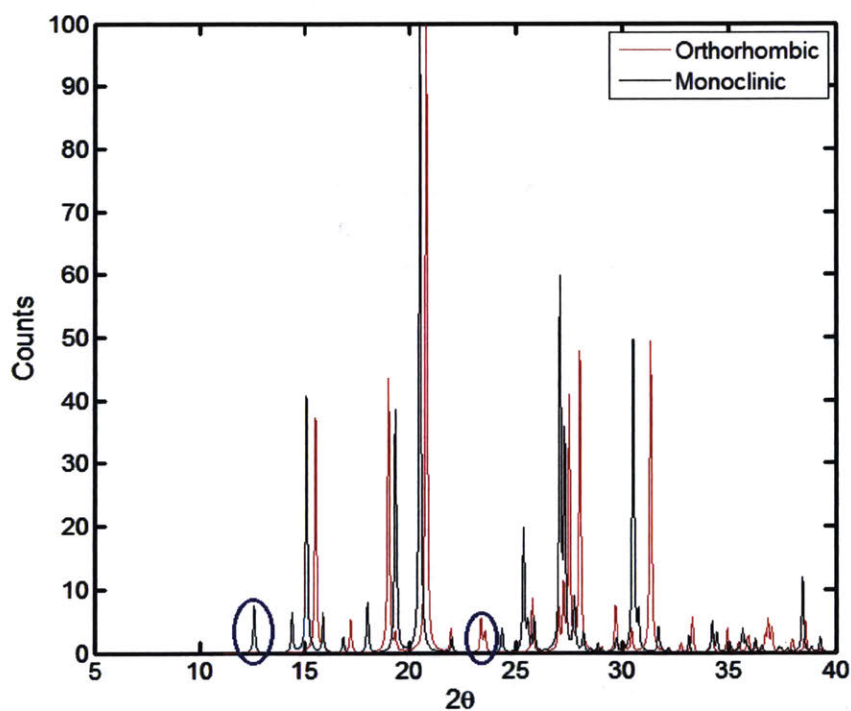


**Figure 5-9:** Overlay of the CSDB polymorphs of 3NP. The comparison of the two PXRD patterns allows determination of non-overlapping unique peaks characteristic to each crystal form. One unique peak was selected for each polymorph and is highlighted on the image with a blue circle.
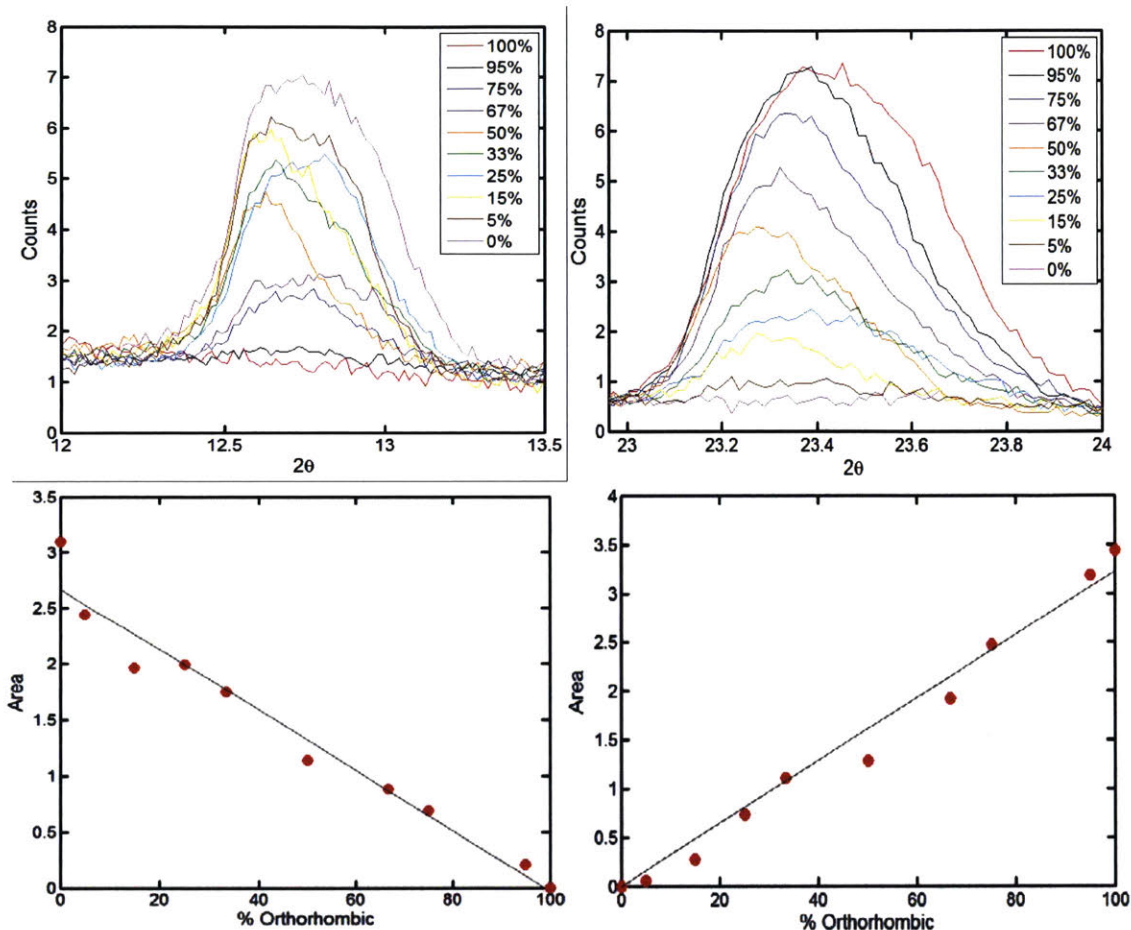
**Figure 5-10:** Calibration of peak area intensity as a function of polymorphic purity. Top: overlay of the XRD patterns used during the calibration. The curves show the effect on the selected unique peaks as the mass percent of orthorhombic polymorph is increased. Bottom: calibration of the integrated area as a function of polymorphic composition for the metastable (left) and stable (right) polymorph unique peaks.

The controlled mixtures of the 3NP polymorphs were used to generate PXRD patterns. Every sample contained 10 mg of the parent powder mixture. The resulting X-ray patterns were normalized to a maximum value of 100 counts. Once normalized, the signal area of the selected unique peaks was integrated and correlated with the polymorphic composition, as a function of the mass percent of orthorhombic (stable) polymorph present within the parent powder. The correlations obtained were strongly linear, as shown in Figure 5-10. As would be expected, the area intensity of the monoclinic unique peak decreases linearly with increasing orthorhombic polymorph content, while simultaneously the area intensity of the orthorhombic unique peak increases linearly.

The resulting correlations can be combined into a ratio of areas, as defined in equation 5-31. The ratio of the areas is a metric that combines the information provided by the integration of both peaks and allows for the utilization of two reference points in the polymorphic determination. The ratio chosen was that of the area of the orthorhombic unique peak ($A_{23\theta}$) divided by the monoclinic unique peak ($A_{12\theta}$). The resulting correlation is shown in Figure 5-11.

$$R = \frac{A_{23\theta}}{A_{12\theta}}$$

5-31
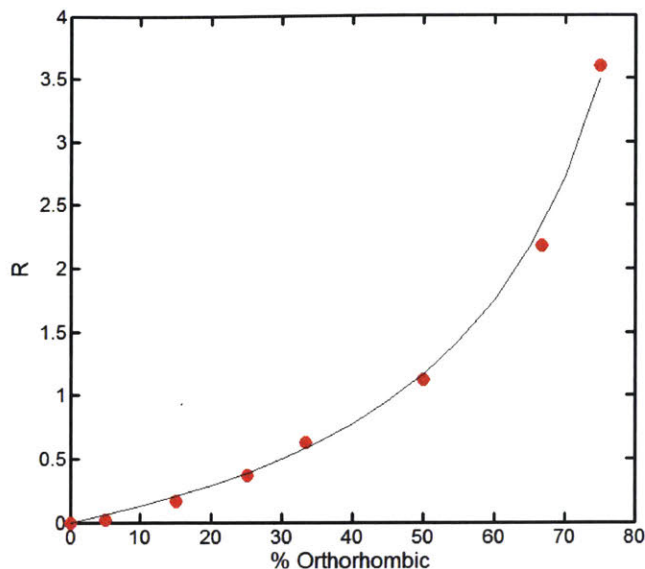


**Figure 5-11:** Calibration of peak area ratio as a function of polymorphic composition. The model fit has $R^2 = 0.99$ and was defined following equation 1.

The correlation obtained was fitted to equation 5-32, where $a$ is a constant of value 1.164, and $O_\%$ is the mass percent of orthorhombic polymorph present within the crystal sample. The equation follows the functional form of the ratio of independent linear functions presented in Figure 5-10. As shown in Figure 5-11, an excellent fit is observed between the experimental data and the calibration model.

$$R = a \frac{O_\%}{(100 - O_\%)}$$

5-32

After the completion of nucleation rate experiments, samples in which crystals had formed were collected. The solids were filtered, dried, and crushed with a mortar and pestle for PXRD analysis. The resulting patterns were analyzed using the calibration curve shown in Figure 5-11. The correlation allowed determining the polymorphic composition of the solid sample and comparing the effect of additives on the final crystal form purity.

**Growth Rate Calculations**

The linear growth rates were estimated using the image-analysis software Axiovision (R 4.8.2 SP2) coupled with the Zeiss Axio Observer.Z1m Microscope. Consecutive pictures of a growing crystals can be scaled and measured to determine the length of the largest dimension of the crystal at each point in time. The dimensions can be plotted as a function of time, and the resulting slope is the best approximation for the growth rate. The resulting plots as well as corresponding data for relevant samples at the control supersaturation ($S = 1.810$) with and without complexing agent are shown. Each set of measurements corresponds to different crystals obtained during the same experimental run. As described on the methods section, each experiment had 80 identical vials, of which 40-50% crystallized. The interval between each image is 2 minutes. Measurements are shown at different stages of growth to illustrate the consistency of the measured growth rates. The onset of nucleation for each crystal was also different, given the stochastic nature of crystallization,

which explains the large discrepancy of time ranges shown. All times shown are in minutes, all dimensions are on micrometers, all slopes in microns per minute.

*3NP Control – No 3ABA:*

| 3NP 1.81 Control – Position 5 | | | 3NP 1.81 Control – Position 8 | | | 3NP 1.81 Control – Position 11 | | |
|---|---|---|---|---|---|---|---|---|
| <u>Images</u> | <u>Time</u> | <u>Size</u> | <u>Images</u> | <u>Time</u> | <u>Size</u> | <u>Images</u> | <u>Time</u> | <u>Size</u> |
| 18 | 34 | 1212.19 | 14 | 26 | 657.89 | 42 | 82 | 545.27 |
| 19 | 36 | 1227.81 | 15 | 28 | 720.95 | 43 | 84 | 593.57 |
| 20 | 38 | 1275.82 | 16 | 30 | 770.08 | 44 | 86 | 654.62 |
| 21 | 40 | 1352.5 | 17 | 32 | 828.98 | 45 | 88 | 696.91 |
| 22 | 42 | 1415.32 | 18 | 34 | 900.95 | 46 | 90 | 742.22 |
| 23 | 44 | 1440.22 | 19 | 36 | 962.02 | 47 | 92 | 835.27 |
| 24 | 46 | 1462.59 | 20 | 38 | 1034.08 | 48 | 94 | 873.74 |
| <u>Slope</u> | 23.49 | | <u>Slope</u> | 31.10 | | <u>Slope</u> | 27.79 | |

<u>Average Growth Rate</u>: 28 ± 4
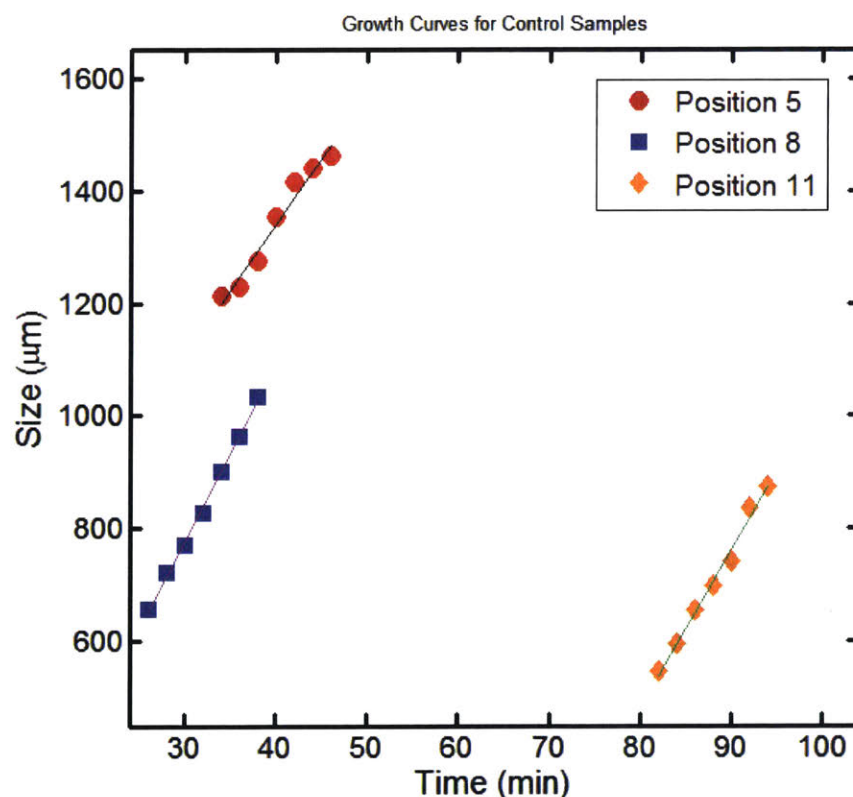


**Figure 5-12:** Growth curves for different samples of the same control crystallization experiment. "Position" refers to the sample number in the microscope stage. Red: position 5, y=23.49x+401.26, $R^2$=0.9667. Blue: position 8, y=31.10x-155.91, $R^2$=0.9973. Yellow: position 11, y=27.79x-1739.8, $R^2$=0.9904. It can be see that the growth rates (slopes) are all very close to each other.

*3NP Control – 0.05 mg/ml 3ABA:*

| 3NP-3ABA 1.81 – Position 1 | | | 3NP-3ABA 1.81 – Position 12 | | | 3NP-3ABA 1.81 – Position 17 | | |
|---|---|---|---|---|---|---|---|---|
| Images | Time | Size | Images | Time | Size | Images | Time | Size |
| 13 | 24 | 277.99 | 33 | 64 | 550.40 | 21 | 40 | 374.26 |
| 14 | 26 | 325.15 | 34 | 66 | 606.32 | 22 | 42 | 425.90 |
| 15 | 28 | 377.12 | 35 | 68 | 650.31 | 23 | 44 | 476.18 |
| 16 | 30 | 417.62 | 36 | 70 | 678.18 | 24 | 46 | 549.43 |
| 17 | 32 | 478.42 | 37 | 72 | 773.21 | 25 | 48 | 575.44 |
| 18 | 34 | 547.34 | 38 | 74 | 822.48 | 26 | 50 | 624.41 |
| 19 | 36 | 596.28 | 39 | 76 | 880.62 | 27 | 52 | 674.81 |
| Slope | 26.43 | | Slope | 27.61 | | Slope | 24.96 | |

Average Growth Rate: 26 ± 1
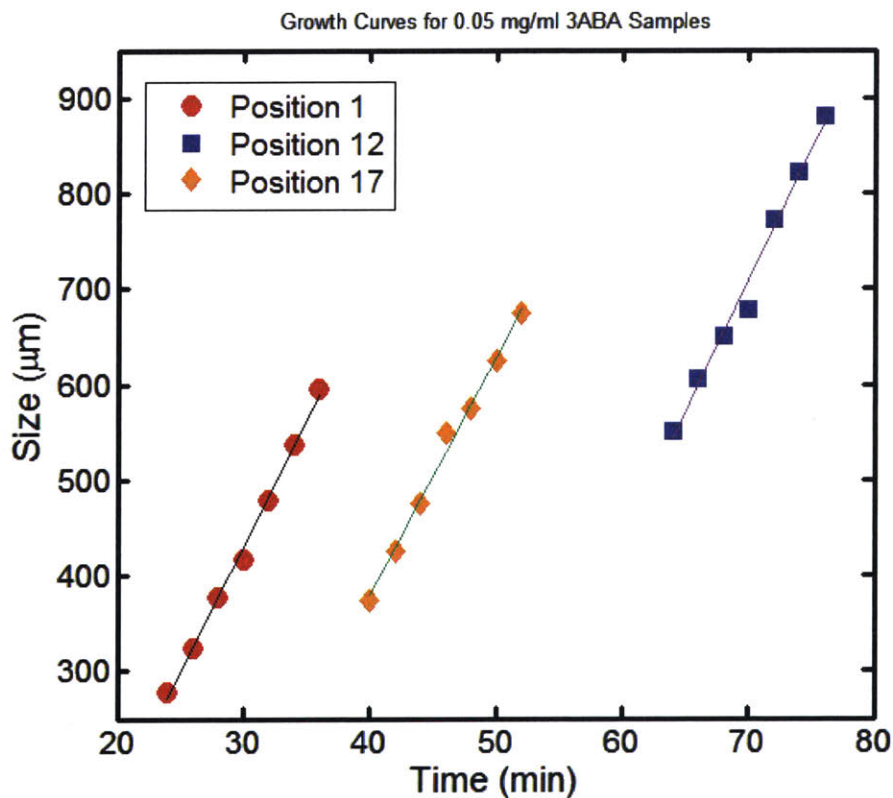


**Figure 5-13:** Growth curves for different samples of the same 0.05 mg/ml 3ABA crystallization experiment. "Position" refers to the sample number in the microscope stage. Red: position 1, y=26.44x+363.16, $R^2$=0.9965. Blue: position 12, y=27.61x-1223.6, $R^2$=0.9862. Yellow: position 17, y=24.96x-619.67, $R^2$=0.9928. It can be see that the growth rates (slopes) are all very close to each other.

## ITC Data

The table below shows the individual data for experiment repetitions used to calculate the binding constants of 3ABA and BA in toluene. The experiments followed the procedure outlined in the main chapter. The conditions of each measurement were identical.

| T (K) | 298 | |
|---|---|---|
| Experiment | K | ΔH (kJ/mol) |
| 1 | 7.00 | -226.95 |
| 2 | 3.55 | -198.24 |
| 3 | 6.73 | -246.88 |
| Average | 5.76 | -224.02 |
| St. Dev. | 1.92 | 24.45 |

| T (K) | 278 | |
|---|---|---|
| Experiment | K | ΔH (kJ/mol) |
| 1 | 14.18 | -200.00 |
| 2 | 12.91 | -201.15 |
| 3 | 13.16 | -192.43 |
| Average | 13.42 | -197.86 |
| St. Dev. | 0.67 | 4.74 |

# Chapter 6 : Polymorphic control of indomethacin using complexation

The contents of this chapter summarize original work that is ongoing and will not be completed by the time of submission of this thesis report. However the contents and discussion shown within are relevant to the thesis topic, and the groundwork that has already been laid out is, in the opinion of the author, worth reporting.

## 6.1 Introduction

Polymorphism is the natural phenomenon through which a molecule can have multiple crystal forms[2,50,107]. Although the chemical structure of the compound involved remains unchanged, its various polymorphs can display significantly different physicochemical properties, including differing solubility, density, reactive stability, and morphology[2,27,49,50,107]. Because of this potential variability in the product properties, polymorph control is a key part of the design and engineering of applied crystallization processes, particularly for specialty applications, such as within the pharmaceutical and food industries. Typically, polymorphism is regulated by adding seeds of the desired form, and by carefully selecting the crystallizing solvent, the initial supersaturation, and the temperature and desupersaturation curves[2]. However, a common challenge for process engineers lies on finding a correct combination of system conditions that can produce the desired crystal form, with a high yield, and without requiring hazardous solvents or high temperatures.

Polymorphism is not only regulated by solute thermodynamics, and crystallization conditions, but also by the presence of impurities[21,108,109]. It is known that impure initial materials, as well as the presence of by-products, and additives can affect the recovered crystal form of a solute. However, peer-reviewed reports discussing this effect are not numerous, and those that do exist lack quantitative analysis of nucleation and polymorphism[45,107,110]. In previous chapters of this thesis, it has been demonstrated that the presence of a weakly-interacting impurity can have massive impact on the nucleation rates of a solute[52,111]. In these past experiments, the observed effect of the inhibitors did not alter the polymorphism of the crystallizing solute, indicating that even for systems with multiple possible crystal forms (chapter 5), all possible nucleation rates were equally inhibited. However, recently completed and unpublished work by *Wijethunga et al.*[112] has demonstrated that functionalized heterosurfaces can lead to changes in the statistical likeliness of a specific polymorph forming at otherwise fixed crystallization conditions. This observation indicates that with the proper chemical targeting, it should be possible to control polymorphism using tailor-made additives, at least for some systems. *Weissbuch et al.* in a review, reported on the use of tailor-made inhibitors to suppress specific polymorphs of small-molecules[109]. Based on this report it stands to reason that complexation can be used as a tool for adding further control during crystallization design, and allowing for selectivity towards a specific form.

As reported by *Weissbuch et al.*[109] and *Davey et al.*[108] chemical systems that exhibit conformational polymorphism can in particular be templated to stabilize one of the specific conformers in solution, leading to the promotion or inhibition of a specific crystal form, based on the selected chemical interaction and the favored conformation. However, the work reported by these researchers was relatively empirical and hypothesized that the inhibition of a particular polymorph took place due to selective attachment to a predominant crystal face of the inhibited polymorph, hindering its growth, and it's nucleation beyond its "critical cluster size". These notions are based on the assumption that nucleation proceeds through classical nucleation theory, which is an outdated model[1], and that growth inhibition is the key to polymorphic nucleation control, which is contradictory when the hypothesized mechanism is described as true "nucleation"

inhibition. Additionally, the data provided does not reflect in depth quantitative analysis of the suppression of nucleation rates of a specific form, but rather, observations on the time required for a solvent-mediated polymorphic transformation to occur, or similar semi-empirical proxies for nucleation. As a result, a thorough analysis of the impact of a conformationally-selective additive on the relative nucleation rate of conformational polymorphs should be performed in order to validate the proposed polymorphic inhibition mechanism.

The project proposed as the subject matter for this chapter, which is currently in its design stage, postulates the detailed study of the nucleation behavior of indomethacin (IMC), and the impact of additives on its polymorphism. Indomethacin is an FDA approved anti-inflammatory drug, with the chemical structure shown in Figure 6-1[112–114]. The molecule is perfect for this study because it displays conformational polymorphism whilst having a relatively simple chemical structure[113–115]. This allows for simplified targeting of its functional groups and facilitates complexing agent selection. The conformational variability of IMC stems from the rotation of the dihedral angles surrounding the amide group, highlighted in Figure 6-1[114]. The resulting molecular motion changes the relative direction of the chlorophenyl group, which can be pointing either in the same direction as the carboxylic acid group (bottom right of Figure 6-1), or as the methoxy group (left of Figure 6-1). These two possible conformations ae illustrated in Figure 6-2.
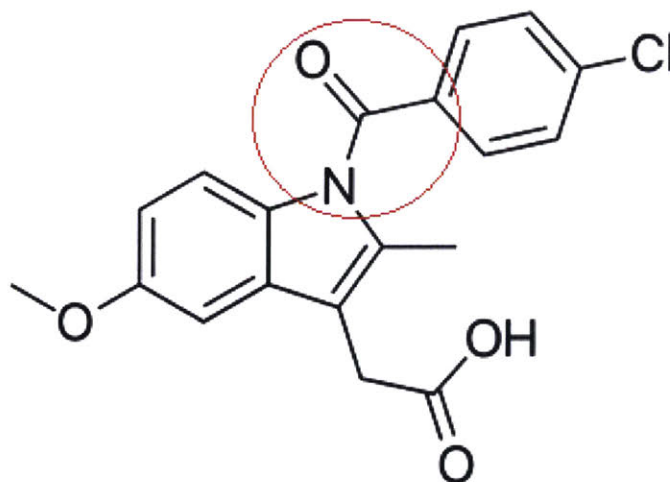


**Figure 6-1:** chemical structure of indomethacin. The highlighted portion of the molecule encompasses the bonds and dihedral angles surrounding the carbonyl group, which experience the rotations that lead to the different possible conformations of the compound.

There are seven known possible polymorphs of IMC; however, only two of them are easily observable and have been successfully characterized[112]. The two predominant forms are the γ and α polymorphs, of which γ is the most stable at all temperatures[113,114]. The repeat unit of the γ form contains only the conformer labeled as "stable" in Figure 6-2, with the carbonyl groups both facing the same direction. Conversely, the unit cell for the α is more complex, containing three interacting molecules of IMC, one of which has the "metastable" conformer (Figure 6-2) of the API. This key difference in conformational composition is the major driver for the design of a complexation motif, given that to suppress the metastable form, a complexing agent should force IMC to remain in its "stable" conformer, hence preventing the molecular arrangement that is necessary to allow the formation of α IMC.
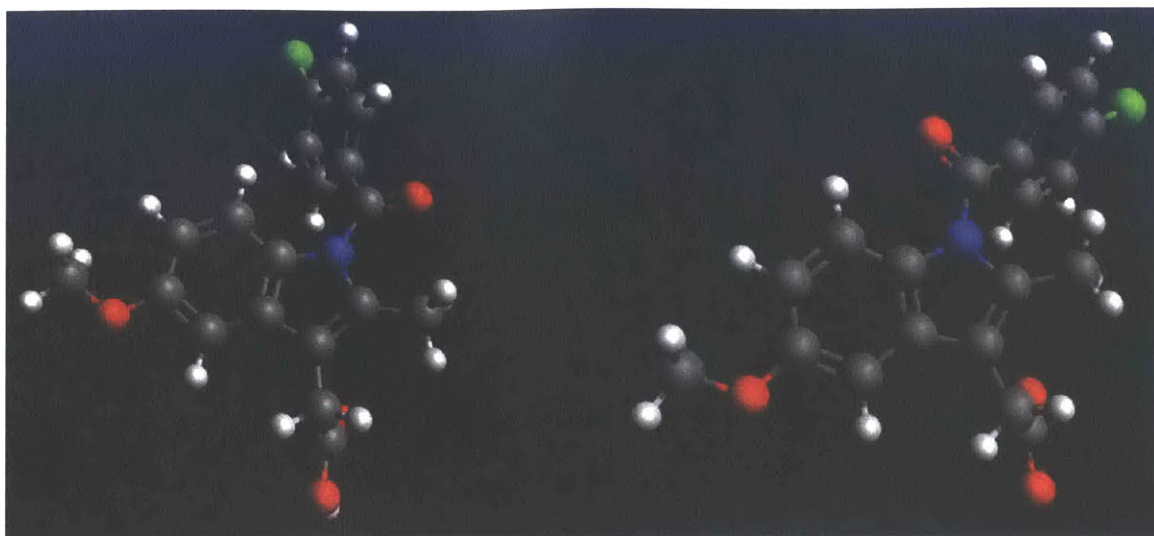
Stable                          Metastable

**Figure 6-2:** structure of the two-major conformations of indomethacin. The difference stems from the rotation of the dihedral angle around the amide group (top of the image). The "metastable" conformer is present only in the $\alpha$ form of IMC, while the "stable" conformer is present in both polymorphs.

The proposed outline of this project is to use high-throughput nucleation experiments with microscopy, as described in chapters 4 and 5, to measure the relative nucleation rates of both polymorphs of IMC and determine how complexing agents affect the preferred polymorphism. In their preliminary work, *Wijethunga et al.*[112] demonstrated that IMC exhibits concomitant crystallization of both polymorphs when it is crystallized in ethanol using cooling crystallization. Their work also demonstrated that the different polymorphs are easy to identify thanks to the microscopy setup, and the characteristically different morphology of each form, as shown in Figure 6-3. Using the image analysis data from the nucleation experiments, it will be possible to measure nucleation rates for both crystal forms under specific crystallizer conditions. Findings from *Wijethunga et al.* [112] showed that the presence of functional heterosurfaces and epitaxial groups affected the relative frequency with which crystals of each form were observed, indicating that the appropriate surface chemistry had the ability to suppress the polymorphism of a specific crystal form. The objective in this project will be to demonstrate that a similar effect can be accomplished using solution complexation to drive the inhibition, rather than foreign surfaces thanks to conformational-locking due to the hydrogen bonding.



**Figure 6-3:** crystallization of different polymorphs of IMC, as shown by the high-throughput microscopy setup. The significantly different appearance of both forms makes them 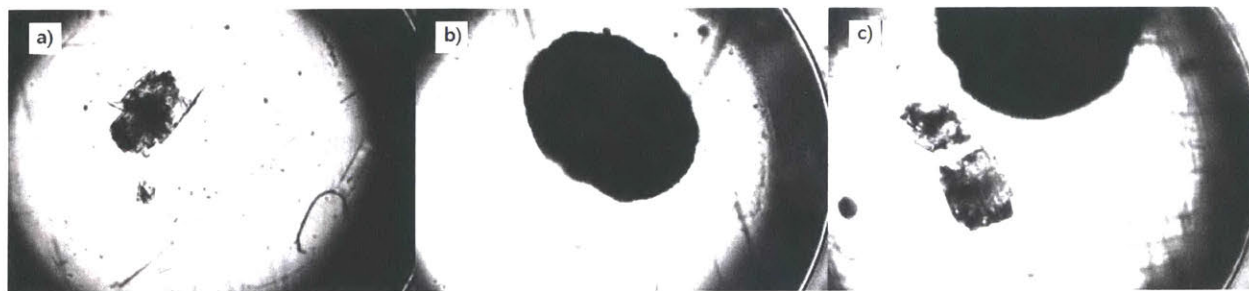easy to differentiate. a) Plate-like crystal of polymorph $\gamma$. b) Clump crystal of polymorph $\alpha$. c) Crystals of both polymorphs appearing simultaneously.

## 6.2 Projected Methods and Results

Experiments will focus on proving the forced changed in conformation caused by the complexing agent, and on the systematic demonstration of the inhibition of the nucleation rate of a specific polymorph of IMC. To achieve this, the steps listed below must be completed:

1.  Find a complexing agent that interacts with IMC and preferentially promotes the nucleation of one of its polymorphs (presumably due to conformational selectivity).

2.  Verify the nature of the complexation interaction and the conformational selectivity in the presence of the additive using calorimetry, spectroscopy, and NMR.

3.  Quantify, using high-throughput induction measurements, the effect of the complexing agent on the nucleation rates of each of the polymorphs of IMC.

All of the steps listed above are crucial to fully test the proposed inhibition hypothesis and elucidate the true inhibition (if observed) mechanism that takes place. Complexing agent selection will be carried out following the procedures outlined in previous chapters of this thesis. Preliminary computational analysis will be carried out to generate a potential complexing agent structure that can lead to the desired selectivity. Second, it will be necessary to characterize the interaction between the complexing agent and IMC, to confirm the formation of a conformationally-selective complex. It is difficult to experimentally demonstrate whether an intermolecular interaction in solution causes the fixation of a conformer, however a combination of calorimetric (ITC), spectroscopic (Raman and IR) and NMR techniques will be employed to confirm the chemical nature of the interactions taking place between the solute and inhibitor. Most likely, the use of two-dimensional NMR to observe the angles and distances of the atoms within the molecule will be necessary to determine the exact conformation of IMC in solution, via the Karplus equation, the Nuclear Overhauser Effect and Residual Dipolar Coupling[116,117]. The resources to perform 2-dimensional NMR analysis will likely require collaboration with other experimental research groups. Establishing whether a complexing agent successfully promotes a conformation is a necessary step for truly explaining potential polymorphic selectivity. Once conformational selectivity and complexation have been demonstrated, the efficacy of tat candidate complexing agent will be confirmed experimentally using nucleation rate measurements with the microscope induction setup. A complexing agent will be deemed "successful" if it produces a significant deviation on the measured relative nucleation rate of the polymorphs with respect to a non-complexed control. All characterization and nucleation rate experiments will be carried out in ethanol for two major reasons: first, ethanol displays concomitant polymorphism of IMC, making it good for studying the equilibrium between the two molecular conformations[112,115]; second, IMC display a sharp solubility curve in ethanol, making it easy to achieve high supersaturations and to force crystals to force for the nucleation rate measurements[112,118,119].

Once characterization is completed, and the steps listed earlier have all been satisfied, it should be possible to unambiguously observe and mechanistically explain selective polymorphic inhibition using conformation-specific solution complexation. As stated at the beginning of this chapter, the work for this project has only recently begun, and will likely be unfinished by the time of submission of this report. However, the experimental outline and relevant mechanistic hypotheses, as well as the literature background that hints at the successful implementation of the proposed approach are topically relevant for this thesis, which is why they have been detailed here.

## Proposed Complexation Motif

Selectivity towards a specific conformer should be most easily achievable for the "stable" configuration, as described in Figure 6-2. In this configuration, both carbonyl groups of IMC must be facing in the same direction. To achieve this, it is important to acknowledge that the two functional groups within IMC most strongly available for intermolecular hydrogen bonding, are the two carbonyl functional groups: the amide (nucleophilic), and the carboxylic acid (nucleophilic and electrophilic)[54,55]. This was confirmed by generating a sigma surface in Cosmotherm17X of IMC, using the TZVP forcefield, as shown in Figure 6-4. As a result, a molecule with hydrogen bond donors and acceptors will likely interact most strongly with these carbonyl moieties, preferably aligning them towards the same direction, as long as the can both simultaneously interact with the complexing agent. To achieve this, the complexation motif illustrated in Figure 6-5 should be implemented. As shown in the image, the desired conformation should be achieved with an interaction partner that has complementary hydrogen bond donor and acceptor groups, spatially located in relative proximity to one another thanks to a common constrained backbone. The backbone will likely be an aromatic ring (aryl), which preferably positions the interacting functional groups in an *ortho* or *meta* arrangement. Once a complexing agent that fits the motif of Figure 6-5 and that can be dissolved in ethanol (the solvent proposed for the induction experiments), it will be possible to perform the analysis outlined to measure and characterize potential polymorphic inhibition.
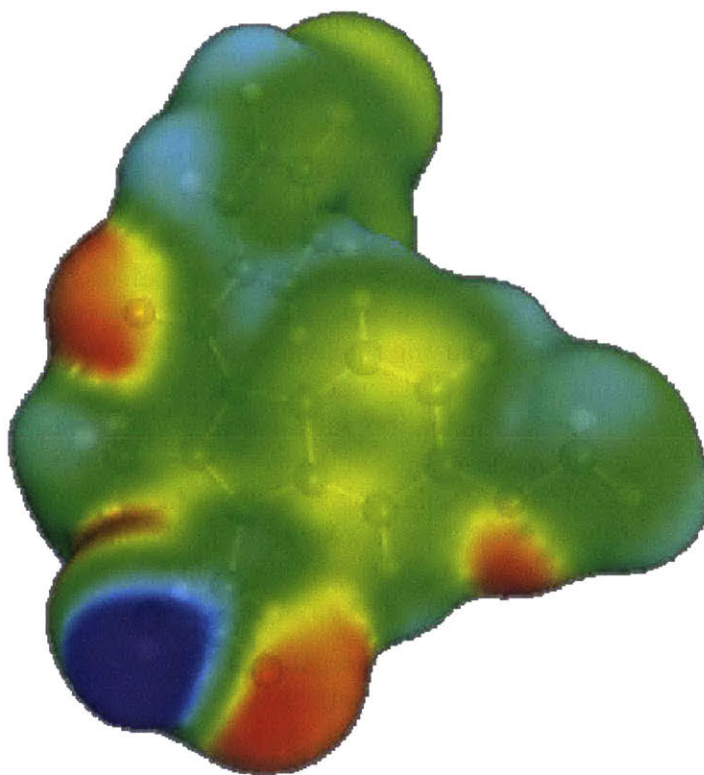


**Figure 6-4:** sigma surface of indomethacin, showing the strongest nucleophilic (red) and electrophilic (blue) atoms of indomethacin. The two carbonyl-containing groups are respectively the most susceptible for intermolecular interactions.
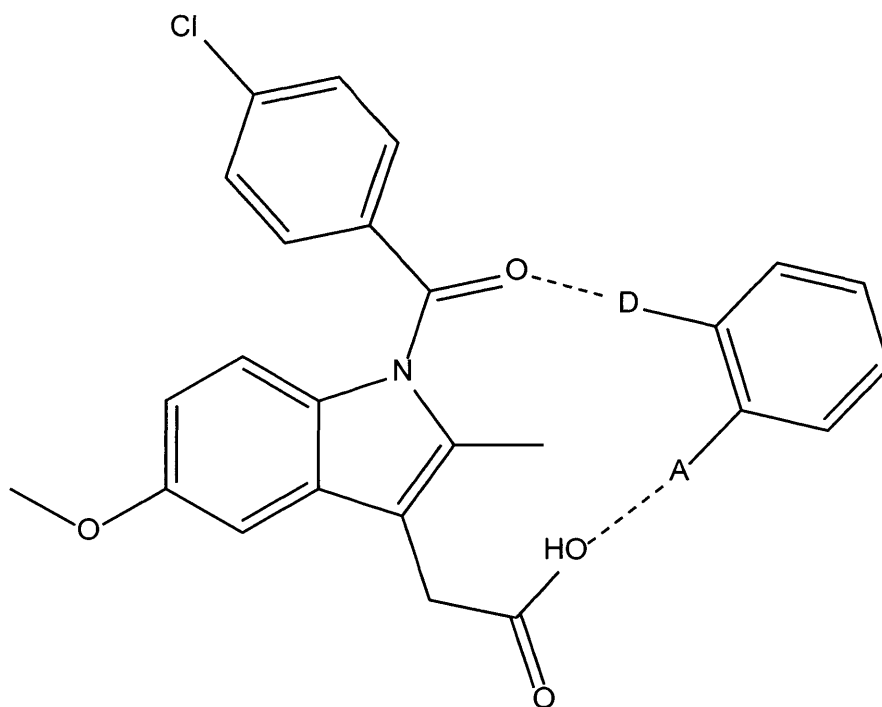
**Figure 6-5:** schematic of the proposed complexation motif needed to sequester the stable conformation of indomethacin, in turn preventing the crystallization of the metastable polymorph. The complexing agent needs a set of hydrogen bond donor and acceptor functional groups that are spatially oriented so that they can force the amide and carboxylic acid groups of IMC to point in the same direction.

# Chapter 7 : Conclusions and Recommendations

The results and analysis provided within this thesis expand the understanding of the effect of impurities, additives, and solution complexation on applied crystallization processes. Specifically, the work herein helped characterize and develop the usage of tailor-made additives for the enhancement of the purification capabilities of crystallization and provided the first systematic demonstration of the effect of impurities and dissolved intermolecular complexes on the nucleation rates of a small-molecule solute. The sections below discuss in detail the most relevant findings and recommendations gained from the whole body of work, split between the two major topics discussed across the various chapters of this report.

## 7.1 Conclusions and outlook for complex-assisted crystallization

The results shown in chapter 2 demonstrated that complex-assisted crystallization can be used to efficiently separate structural isomers, as was shown experimentally by the removal of 3NP from 4NP, using 3ABA. Selectivity of interaction (and as a result, purification) can be achieved as long as there are targetable structural differences between the relative positioning of the hydrogen bond donor and acceptor groups of the purification target and its impurity. The appropriate complexation motif employs a complexing agent that has a constrained backbone, which forces its interacting functional groups to be in relative orientations to one another, so that the intermolecular interactions possible for the additive molecule are spatially favorable towards the impurity and not the crystallizing solute. The findings of this project also demonstrated that complex-assisted crystallization is not uniquely applicable to systems containing molecules with carboxylic acid functional groups, as previously seemed to be the case based on successful past implementations of the technique[65-68].

The project summarized in chapter 3 served as an exploration of the limits of complex-assisted crystallization. The highly ambitious purity requirements, as well as the complexity of the molecules involved in the separation, represented the greatest challenge ever for the technique. Although a small degree of purification was achieved, it was clear that the extreme low impurity concentrations desired for the final product imposed too great of a thermodynamic constraint on the speciation equilibrium required for successful selectivity in the complexation. This limitation was particularly accentuated because the specific functional differences between the chemical structures of the API and its impurity were not conducive for strong intermolecular selectivity towards the latter. These findings, however, are crucial for establishing the operational limits of complex-assisted crystallization and help demonstrate that the technique is most successfully applied for small-molecule systems with a large impurity concentration in the feed stream. This set of conditions indicate that complex-assisted crystallization is likely not best applied on pharmaceutical crystallization processes, but rather in chemical, mining and other applied synthetic processes that have simpler structures and more flexible purity specifications.

There is an extensive amount of valuable work that should be done to further expand complex-assisted crystallization and help it become a viable technique for applied industrial crystallization processes. Some of this work has already begun with the findings of *Vartak et al.* who has demonstrated that complex-assisted crystallization can be successfully implemented continuously[80]. Many challenges arise in the transition from batch to continuous, as considerations around yield maximization and additive accumulation reduction become central in the operational design of the crystallizers. However, the single greatest challenge for complex-assisted crystallization remains the non-universality of the technique, requiring case-by-case design and

selection of suitable complexing agents. The greatest possible contribution that can be developed for the technique would be a method or algorithm (preferably *in silico*) that could efficiently predict the structure of a successful complexing agent based on the chemical structures of a given purification target and its impurity, as well as the crystallization conditions (solvent and temperature). Although *Weber et al.*[66] laid out the preliminary work for a computational strategy that could predict the quality of a complexing agent, their approach was imperfect, not universally applicable, and computationally expensive. As a result, a more streamlined and simplified approach would be invaluable for the future of the technique.

## 7.2 Conclusions and outlook for nucleation inhibition

The data discussed in chapters 4 and 5 provided a fascinating breakthrough in the understanding of the impact of dilute impurities on the mechanism of nucleation of small-molecule solutes. Using high-throughput induction time measurements, it was systematically demonstrated that even a minimal concentration of an impurity that interacts with a crystallizing solute can lead to a massive change in the nucleation rate of said solute. For the cases shown in this thesis, the complexation mechanism was observed to inhibit the onset of nucleation, decreasing nucleation rates by almost a factor of three when only 1 % by mass of the impurity was present within the dissolved solutes. The experimental results demonstrated that the nature of inhibition was kinetic and not thermodynamic, implying that the slight changes in supersaturation and speciation brought about by complexation were not the cause for the nucleation inhibition. An alternative interpretation for the inhibition mechanism was provided, based on careful measurement of the kinetic parameters for the nucleation rates of 3NP in the presence and absence of a small concentration of 3ABA inhibitor. The data was analyzed using classical nucleation theory and the two-step model of nucleation. Order of magnitude analysis and physical interpretation of the model variables led to the conclusion that classical nucleation theory cannot adequately explain the effect of impurities on nucleation rates. The two-step model, however, was found to be capable of rationalizing the observed phenomena by predicting that the kinetic parameter which describes the ordering of high-concentration clusters into a nucleus becomes depressed by the intermolecular complexes. The proposed mechanism states that nucleation inhibition proceeds thanks to the suppression of the ordering step, due to the physical disruption of the molecular movement and required re-arrangement, caused by the bulky intermolecular complexes, as illustrated in Figure 5-7.

Although unfinished, chapter 6 proposed a mechanism for polymorphic control of small molecules using complexation. The underlying mechanism for the technique assumes that crystallization proceeds following the two-step mechanism of nucleation and that the preferred conformation of the solute in the solution dictates the polymorphism of the resulting crystal. Demonstrating this hypothesis requires a unique model system that displays conformational polymorphism, a relatively simple molecular structure, and has the ability to participate in hydrogen bonding with an intermolecular additive. The chosen chemical was indomethacin, which has the added benefit of being an FDA approved API. Successful implementation of this project would require finding a complexing agent that stabilizes one of the two major conformations of the API and leads to a measurable change in the probability of the corresponding polymorph being formed. High throughput experimentation, analogous to the one employed in the induction time measurements, will be used to establish the probability distribution of a given polymorph forming as a function of the presence of specific complexing agents. If complex-induced polymorphic selectivity can be demonstrated, the data would provide further evidence for the two-step mechanism of crystallization. Additionally, the technique could prove useful for applied crystallization systems

111

with well-defined process constrains, given that it would allow for controlling the possible achievable polymorphism in solution without requiring changes in the optimal solvent and temperature conditions required for the crystallization batch.

The findings summarized in chapters 4-6 provide fascinating insights that expand the existing understanding of nucleation as a whole. More detailed work can be done to further characterize and demonstrate the mechanisms that govern how impurities and intermolecular interactions alter the formation of high-concentration clusters and nuclei. Studies of the impact of changes of the concentration of the inhibitor additive on the kinetic parameters of nucleation rate could further prove the hypotheses outlined herein. Additionally, it has been reported that the changes in nucleation rate due to presence of impurity are non-linear, exhibiting a global minimum[5]. This phenomenon could be demonstrated and rationalized using the analysis discussed in chapter 5, based on the competing exponential and pre-exponential behavior of the kinetic parameters in inhibited systems. Finally, the project in chapter 6 should be completed in order to establish whether the proposed strategy for polymorphic control is viable or not. If it is, the use of additives for polymorph selection should be further characterized and demonstrated, so that it can be implemented as a process handle tool for applied crystallization processes.

# Chapter 8 : References

(1) Chen, J.; Sarma, B.; Evans, J. M. B.; Myerson, A. S. Pharmaceutical Crystallization. *Cryst. Growth Des.* **2011**, *11* (4), 887–895.

(2) Myerson, A. *Handbook of Industrial Crystallization*; Butterworth-Heinemann, 2002.

(3) Erdemir, D.; Lee, A. Y.; Myerson, A. S. Nucleation of Crystals from Solution: Classical and Two-Step Models. *Acc. Chem. Res.* **2009**, *42* (5), 621–629.

(4) Févotte, F.; Févotte, G. A Method of Characteristics for Solving Population Balance Equations (PBE) Describing the Adsorption of Impurities during Crystallization Processes. *Chem. Eng. Sci.* **2010**, *65* (10), 3191–3198.

(5) Rauls, M.; Bartosch, K.; Kind, M.; Kuch, S.; Lacmann, R.; Mersmann, A. The Influence of Impurities on Crystallization Kinetics – a Case Study on Ammonium Sulfate. *J. Cryst. Growth* **2000**, *213* (1), 116–128.

(6) Garnero, C.; Aloisio, C.; Longhi, M. R. Ibuprofen-Maltodextrin Interaction: Study of Enantiomeric Recognition and Complex Characterization. **2013**.

(7) Mukuta, T.; Lee, A. Y.; Kawakami, T.; Myerson, A. S. Influence of Impurities on the Solution-Mediated Phase Transformation of an Active Pharmaceutical Ingredient. *Cryst. Growth Des.* **2005**, *5* (4), 1429–1436.

(8) Rodríguez-Hornedo, N.; Nehm, S. J.; Seefeldt, K. F.; Pagán-Torres, Y.; Falkiewicz, C. J. Reaction Crystallization of Pharmaceutical Molecular Complexes. *Mol. Pharm.* **2006**, *3* (3), 362–367.

(9) Higuchi, T.; Bolton, S. The Solubility and Complexing Properties of Oxytetracycline and Tetracycline III: Interactions in Aqueous Solution With Model Compounds, Biochemicals, Metals, Chelates, and Hexametaphosphate. *J. Am. Pharm. Assoc. Sci. Ed* **1959**, *48* (10), 557–564.

(10) Higuchi, T.; Lach, J. L. Investigation of Some Complexes Formed in Solution by Caffeine*: IV. Interactions Between Caffeine and Sulfathiazole, Sulfadiazine, P-Aminobenzoic Acid, Benzocaine, Phenobarbital, and Barbital. *J. Am. Pharm. Assoc. Sci. Ed* **1954**, *43* (6), 349–354.

(11) Higuchi, T.; Lach, J. L. Study of Possible Complex Formation Between Macromolecules and Certain Pharmaceuticals: III. Interaction of Polyethylene Glycols with Several Organic Acids. *J. Am. Pharm. Assoc. Sci. Ed* **1954**, *43* (8), 465–470.

(12) Higuchi, T.; Pitman, I. H. Caffeine Complexes with Low Water Solubility: Synthesis and Dissolution Rates of 1:1 and 1:2 Caffeine-Gentisic Acid Complexes. *J. Pharm. Sci.* **1973**, *62* (1), 55–58.

(13) Higuchi, T.; Zuck, D. A. Investigation of Some Complexes Formed in Solution by Caffeine†: III. Interactions Between Caffeine and Aspirin, p-Hydroxybenzoic Acid, m-Hydroxybenzoic Acid, Salicylic Acid, Salicylate Ion, and Butyl Paraben. *J. Am. Pharm. Assoc. Sci. Ed* **1953**, *42* (3), 138–145.

(14) Higuchi, T.; Zuck, D. A. Investigation of Some Complexes Formed in Solution by Caffeine. II. Benzoic Acid and Benzoate Ion. *J. Am. Pharm. Assoc.* **1953**, *42* (3), 132–138.

(15) Higuchi, T.; Zuck, D. A. Solubilizing Action of Caffeine on Benzoic Acid. *J. Pharm. Sci.* **1952**, *41* (1), 10–13.

(16) Gans, E. H.; Higuchi, T. The Solubility and Complexing Properties of Oxytetracycline and Tetracycline I. Interaction in Aqueous Solution. *J. Pharm. Sci.* **1957**, *46* (8), 458–466.

(17) Gans, E.; Higuchi, T. The Solubility and Complexing Properties of Oxytetracycline and Tetracycline II**School of Pharmacy, University of Wisconsin, Madison: Interactions in Nonaqueous Solution. *J. Am. Pharm. Assoc. Sci. Ed* **1957**, *46* (10), 587–591.

(18) Kostenbauder, H. B.; Higuchi, T. Formation of Molecular Complexes by Some Water-Soluble Amides II. Effect of Decreasing Water Solubility on Degree of Complex Formation**School of Pharmacy, University of Wisconsin, Madison. *J. Am. Pharm. Assoc. Sci. Ed* **1956**, *45* (12), 810–813.

(19) Kostenbauder, H. B.; Higuchi, T. Formation of Molecular Complexes by Some Water-Soluble Amides I. Interaction of Several Amides with p-Hydroxybenzoic Acid, Salicylic Acid, Chloramphenicol, and Phenol. *J. Am. Pharm. Assoc.* **1956**, *45* (8), 518–522.

(20) Poole, J. W.; Higuchi, T. Complexes Formed in Aqueous Solutions by Sarcosine Anhydride; Interactions with Organic Acids, Phenols, and Aromatic Alcohols. *J. Am. Pharm. Assoc.* **1959**, *48* (10), 592–601.

(21) Addadi, L.; Berkovitch-Yellin, Z.; Weissbuch, I.; van Mil, J.; Shimon, L. J. W.; Lahav, M.; Leiserowitz, L. Growth and Dissolution of Organic Crystals with "Tailor-Made" Inhibitors—Implications in Stereochemistry and Materials Science. *Angew. Chem. Int. Ed. Engl.* **1985**, *24* (6), 466–485.

(22) Davey, R. J.; Schroeder, S. L. M.; ter Horst, J. H. Nucleation of Organic Crystals—A Molecular Perspective. *Angew. Chem. Int. Ed.* **2013**, *52* (8), 2166–2179.

(23) Rimer, J. D.; An, Z.; Zhu, Z.; Lee, M. H.; Goldfarb, D. S.; Wesson, J. A.; Ward, M. D. Crystal Growth Inhibitors for the Prevention of L-Cystine Kidney Stones Through Molecular Design. *Science* **2010**, *330* (6002), 337–341.

(24) McCartney, E. R.; Alexander, A. E. The Effect of Additives upon the Process of Crystallization: I. Crystallization of Calcium Sulfate. *J. Colloid Sci.* **1958**, *13* (4), 383–396.

(25) Crowley, K. J.; Zografi, G. The Effect of Low Concentrations of Molecularly Dispersed Poly(Vinylpyrrolidone) on Indomethacin Crystallization from the Amorphous State. *Pharm. Res.* **2003**, *20* (9), 1417–1422.

(26) Kirkova, E.; Djarova, M.; Donkova, B. Inclusion of Isomorphous Impurities during Crystallization from Solutions. *Prog. Cryst. Growth Charact. Mater.* **1996**, *32* (1), 111–134.

(27) Poornachary, S. K.; Chow, P. S.; Tan, R. B. H. Influence of Solution Speciation of Impurities on Polymorphic Nucleation in Glycine. *Cryst. Growth Des.* **2008**, *8* (1), 179–185.

(28) Granberg, R. A.; Rasmuson, Å. C. Solubility of Paracetamol in Binary and Ternary Mixtures of Water + Acetone + Toluene. *J. Chem. Eng. Data* **2000**, *45* (3), 478–483.

(29) Granberg, R. A.; Rasmuson, Å. C. Solubility of Paracetamol in Pure Solvents. *J. Chem. Eng. Data* **1999**, *44* (6), 1391–1395.

(30) Granberg, R. A.; Bloch, D. G.; Rasmuson, Å. C. Crystallization of Paracetamol in Acetone–water Mixtures. *J. Cryst. Growth* **1999**, *198–199*, 1287–1293.

(31) Valavi, M.; Svärd, M.; Rasmuson, Å. C. Improving Estimates of the Crystallization Driving Force: Investigation into the Dependence on Temperature and Composition of Activity Coefficients in Solution. *Cryst. Growth Des.* **2016**, *16* (12), 6951–6960.

(32) Jackson, K. A. Crystal Growth Kinetics. *Mater. Sci. Eng.* **1984**, *65* (1), 7–13.

(33) Kim, S.; Myerson, A. S. Metastable Solution Thermodynamic Properties and Crystal Growth Kinetics. *Ind. Eng. Chem. Res.* **1996**, *35* (4), 1078–1084.

(34) Gebauer, D.; Cölfen, H. Prenucleation Clusters and Non-Classical Nucleation. *Nano Today* **2011**, *6* (6), 564–584.

(35) Vatamanu, J.; G. Kusalik, P. Observation of Two-Step Nucleation in Methane Hydrates. *Phys. Chem. Chem. Phys.* **2010**, *12* (45), 15065–15072.

(36) Kashchiev, D.; van Rosmalen, G. M. Review: Nucleation in Solutions Revisited. *Cryst. Res. Technol.* **2003**, *38* (7–8), 555–574.

(37) Pan, W.; Kolomeisky, A. B.; Vekilov, P. G. Nucleation of Ordered Solid Phases of Proteins via a Disordered High-Density State: Phenomenological Approach. *J. Chem. Phys.* **2005**, *122* (17), 174905.

(38) G. Vekilov, P. The Two-Step Mechanism of Nucleation of Crystals in Solution. *Nanoscale* **2010**, *2* (11), 2346–2357.

(39) Kashchiev, D.; Vekilov, P. G.; Kolomeisky, A. B. Kinetics of Two-Step Nucleation of Crystals. *J. Chem. Phys.* **2005**, *122* (24), 244706.

(40) GARSIDE, J.; DAVEY, R. J. Invited Review Secondary Contact Nucleation: Kinetics, Growth and Scale-Up. *Chem. Eng. Commun.* **1980**, *4* (4–5), 393–424.

(41) K, B. W.; N, C.; C, F. F. The Growth of Crystals and the Equilibrium Structure of Their Surfaces. *Phil Trans R Soc Lond A* **1951**, *243* (866), 299–358.

(42) Abu Bakar, M. R.; Nagy, Z. K.; Saleemi, A. N.; Rielly, C. D. The Impact of Direct Nucleation Control on Crystal Size Distribution in Pharmaceutical Crystallization Processes. *Cryst. Growth Des.* **2009**, *9* (3), 1378–1384.

(43) Mullin, J. W. *Crystallization*; Elsevier, 2001.

(44) Weissbuch, I.; Addadi, L.; Leiserowitz, L. Molecular Recognition at Crystal Interfaces. *Science* **1991**, *253* (5020), 637–645.

(45) Singh, A.; Lee, I. S.; Myerson, A. S. Concomitant Crystallization of ROY on Patterned Substrates: Using a High Throughput Method to Improve the Chances of Crystallization of Different Polymorphs. *Cryst. Growth Des.* **2009**, *9* (2), 1182–1185.

(46) Bernstein, J.; Davey, R. J.; Henck, J.-O. Concomitant Polymorphs. *Angew. Chem. Int. Ed.* **1999**, *38* (23), 3440–3461.

(47) Cruz-Cabeza, A. J.; Bernstein, J. Conformational Polymorphism. *Chem. Rev.* **2014**, *114* (4), 2170–2191.

(48) Shekunov, B. Y.; York, P. Crystallization Processes in Pharmaceutical Technology and Drug Delivery Design. *J. Cryst. Growth* **2000**, *211* (1), 122–136.

(49) Haleblian John; McCrone Walter. Pharmaceutical Applications of Polymorphism. *J. Pharm. Sci.* **2006**, *58* (8), 911–929.

(50) Caira, M. R. Crystalline Polymorphism of Organic Compounds. In *Design of Organic Solids*; Topics in Current Chemistry; Springer, Berlin, Heidelberg, 1998; pp 163–208.

(51) Ginde, R. M.; Myerson, A. S. Effect of Impurities on Cluster Growth and Nucleation. *J. Cryst. Growth* **1993**, *126* (2), 216–222.

(52) Pons Siepermann, C. A.; Huang, S.; Myerson, A. S. Nucleation Inhibition of Benzoic Acid through Solution Complexation. *Cryst. Growth Des.* **2017**, *17* (5), 2646–2653.

(53) Pina, C. M.; Becker, U.; Risthaus, P.; Bosbach, D.; Putnis, A. Molecular-Scale Mechanisms of Crystal Growth in Barite. *Nature* **1998**, *395* (6701), 483–486.

(54) Hunter, C. A. Quantifying Intermolecular Interactions: Guidelines for the Molecular Recognition Toolbox. *Angew. Chem. Int. Ed.* **2004**, *43* (40), 5310–5324.

(55) Etter, M. C. Hydrogen Bonds as Design Elements in Organic Chemistry. *J. Phys. Chem.* **1991**, *95* (12), 4601–4610.

(56) Ercolani, G.; Schiaffino, L. Allosteric, Chelate, and Interannular Cooperativity: A Mise Au Point. *Angew. Chem. Int. Ed.* **2011**, *50* (8), 1762–1768.

(57) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A. PKa Calculations of Aliphatic Amines, Diamines, and Aminoamides via Density Functional Theory with a Poisson–Boltzmann Continuum Solvent Model. *J. Phys. Chem. A* **2007**, *111* (20), 4422–4430.

(58) Palecz, B.; Piekarski, H.; Romanowski, S. Studies on Homogeneous Interactions between Zwitterions of Severall-α-Amino Acids in Water at a Temperature of 298.15 K. *J. Mol. Liq.* **2000**, *84* (3), 279–288.

(59) R. Desiraju, G. Crystal and Co-Crystal. *CrystEngComm* **2003**, *5* (82), 466–467.

(60) D. Dunitz, J. Crystal and Co-Crystal: A Second Opinion. *CrystEngComm* **2003**, *5* (91), 506–506.

(61) Mohammad, M. A.; Alhalaweh, A.; Velaga, S. P. Hansen Solubility Parameter as a Tool to Predict Cocrystal Formation. *Int. J. Pharm.* **2011**, *407* (1), 63–71.

(62) Thakuria, R.; Delori, A.; Jones, W.; Lipert, M. P.; Roy, L.; Rodríguez-Hornedo, N. Pharmaceutical Cocrystals and Poorly Soluble Drugs. *Int. J. Pharm.* **2013**, *453* (1), 101–125.

(63) K. Duggirala, N.; L. Perry, M.; Almarsson, Ö.; J. Zaworotko, M. Pharmaceutical Cocrystals: Along the Path to Improved Medicines. *Chem. Commun.* **2016**, *52* (4), 640–655.

(64) D. Bond, A. What Is a Co-Crystal? *CrystEngComm* **2007**, *9* (9), 833–834.

(65) Hsi, K. H.; Kenny, M.; Simi, A.; Myerson, A. S. Purification of Structurally Similar Compounds by the Formation of Impurity Co-Former Complexes in Solution. *Cryst. Growth Des.* **2013**, *13* (4), 1577–1582.

(66) Weber, C. C.; Wood, G. P. F.; Kunov-Kruse, A. J.; Nmagu, D. E.; Trout, B. L.; Myerson, A. S. Quantitative Solution Measurement for the Selection of Complexing Agents to Enable Purification by Impurity Complexation. *Cryst. Growth Des.* **2014**, *14* (7), 3649–3657.

(67) Ying Hsi, K. H.; Joseph Concepcion, A.; Kenny, M.; Ahmed Magzoub, A.; S. Myerson, A. Purification of Amoxicillin Trihydrate by Impurity-Coformer Complexation in Solution. *CrystEngComm* **2013**, *15* (34), 6776–6781.

(68) Huai-Ying Hsi, K.; Chadwick, K.; Fried, A.; Kenny, M.; S. Myerson, A. Separation of Impurities from Solution by Selective Co-Crystal Formation. *CrystEngComm* **2012**, *14* (7), 2386–2388.

(69) Gong, Y.; Collman, B. M.; Mehrens, S. M.; Lu, E.; Miller, J. M.; Blackburn, A.; Grant, D. J. W. Stable-Form Screening: Overcoming Trace Impurities That Inhibit Solution-Mediated Phase Transformation to the Stable Polymorph of Sulfamerazine. *J. Pharm. Sci.* **2008**, *97* (6), 2130–2144.

(70) Pons-Siepermann, C.; Huang, S.; S. Myerson, A. Purification of Nitrophenols Using Complex-Assisted Crystallization. *CrystEngComm* **2016**, *18* (39), 7487–7493.

(71) Sidgwick, N. V.; Spurrell, W. J.; Davies, T. E. CXXXII.—The Solubility of the Nitrophenols and Other Isomeric Disubstitution Products of Benzene. *J. Chem. Soc. Trans.* **1915**, *107* (0), 1202–1213.

(72) Kulkarni, G. U.; Kumaradhas, P.; Rao, C. N. R. Charge Density Study of the Polymorphs of P-Nitrophenol. *Chem. Mater.* **1998**, *10* (11), 3498–3505.

(73) Ferguson, S.; Ortner, F.; Quon, J.; Peeva, L.; Livingston, A.; Trout, B. L.; Myerson, A. S. Use of Continuous MSMPR Crystallization with Integrated Nanofiltration Membrane

Recycle for Enhanced Yield and Purity in API Crystallization. *Cryst. Growth Des.* **2014**, *14* (2), 617–627.

(74) Kim, J. F.; Freitas da Silva, A. M.; Valtcheva, I. B.; Livingston, A. G. When the Membrane Is Not Enough: A Simplified Membrane Cascade Using Organic Solvent Nanofiltration (OSN). *Sep. Purif. Technol.* **2013**, *116*, 277–286.

(75) Valtcheva, I. B.; Kumbharkar, S. C.; Kim, J. F.; Bhole, Y.; Livingston, A. G. Beyond Polyimide: Crosslinked Polybenzimidazole Membranes for Organic Solvent Nanofiltration (OSN) in Harsh Environments. *J. Membr. Sci.* **2014**, *457*, 62–72.

(76) Commissioner, O. of the. Search for FDA Guidance Documents https://www.fda.gov/RegulatoryInformation/Guidances/default.htm (accessed Mar 13, 2018).

(77) Wójcik, G.; Mossakowska, I. Polymorphs of P-Nitrophenol as Studied by Variable-Temperature X-Ray Diffraction and Calorimetry: Comparison with m-Nitrophenol. *Acta Crystallogr. B* **2006**, *62* (1), 143–152.

(78) Koehn, F. E.; Carter, G. T. The Evolving Role of Natural Products in Drug Discovery. *Nat. Rev. Drug Discov.* **2005**, *4* (3), 206–220.

(79) Schreiber, S. L. Target-Oriented and Diversity-Oriented Organic Synthesis in Drug Discovery. *Science* **2000**, *287* (5460), 1964–1969.

(80) Vartak, S.; Myerson, A. S. Continuous Crystallization with Impurity Complexation and Nanofiltration Recycle. *Org. Process Res. Dev.* **2017**, *21* (2), 253–261.

(81) Wijethunga, T. K.; Baftizadeh, F.; Stojaković, J.; Myerson, A. S.; Trout, B. L. Experimental and Mechanistic Study of the Heterogeneous Nucleation and Epitaxy of Acetaminophen with Biocompatible Crystalline Substrates. *Cryst. Growth Des.* **2017**, *17* (7), 3783–3795.

(82) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 490–519.

(83) Halgren, T. A. Merck Molecular Force Field. II. MMFF94 van Der Waals and Electrostatic Parameters for Intermolecular Interactions. *J. Comput. Chem.* **1996**, *17* (5–6), 520–552.

(84) Halgren, T. A. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular-Interaction Energies and Geometries. *J. Comput. Chem.* **1999**, *20* (7), 730–748.

(85) Nehm, S. J.; Rodríguez-Spong, B.; Rodríguez-Hornedo, N. Phase Solubility Diagrams of Cocrystals Are Explained by Solubility Product and Solution Complexation. *Cryst. Growth Des.* **2006**, *6* (2), 592–600.

(86) Jayasankar, A.; Reddy, L. S.; Bethune, S. J.; Rodríguez-Hornedo, N. Role of Cocrystal and Solution Chemistry on the Formation and Stability of Cocrystals with Different Stoichiometry. *Cryst. Growth Des.* **2009**, *9* (2), 889–897.

(87) J. Benton, W.; R. Collins, I.; M. Grimsey, I.; M. Parkinson, G.; A. Rodger, S. Nucleation, Growth and Inhibition of Barium Sulfate-Controlled Modification with Organic and Inorganic Additives. *Faraday Discuss.* **1993**, *95* (0), 281–297.

(88) Jiang, S.; ter Horst, J. H. Crystal Nucleation Rates from Probability Distributions of Induction Times. *Cryst. Growth Des.* **2011**, *11* (1), 256–261.

(89) P. Sear, R. Quantitative Studies of Crystal Nucleation at Constant Supersaturation: Experimental Data and Models. *CrystEngComm* **2014**, *16* (29), 6506–6522.

(90) Warren, D. B.; Benameur, H.; Porter, C. J. H.; Pouton, C. W. Using Polymeric Precipitation Inhibitors to Improve the Absorption of Poorly Water-Soluble Drugs: A Mechanistic Basis for Utility. *J. Drug Target.* **2010**, *18* (10), 704–731.

(91) Ozaki, S.; Kushida, I.; Yamashita, T.; Hasebe, T.; Shirai, O.; Kano, K. Inhibition of Crystal Nucleation and Growth by Water-Soluble Polymers and Its Impact on the Supersaturation Profiles of Amorphous Drugs. *J. Pharm. Sci.* **2013**, *102* (7), 2273–2281.

(92) Sullivan, R. A.; Davey, R. J.; Sadiq, G.; Dent, G.; Back, K. R.; ter Horst, J. H.; Toroz, D.; Hammond, R. B. Revealing the Roles of Desolvation and Molecular Self-Assembly in Crystal Nucleation from Solution: Benzoic and p-Aminobenzoic Acids. *Cryst. Growth Des.* **2014**, *14* (5), 2689–2696.

(93) Apelblat, A.; Manzurola, E.; Abo Balal, N. The Solubilities of Benzene Polycarboxylic Acids in Water. *J. Chem. Thermodyn.* **2006**, *38* (5), 565–571.

(94) Knezic, D.; Zaccaro, J.; Myerson, A. S. Nucleation Induction Time in Levitated Droplets. *J. Phys. Chem. B* **2004**, *108* (30), 10672–10677.

(95) Bruno, G.; Randaccio, L. A Refinement of the Benzoic Acid Structure at Room Temperature. *Acta Crystallogr. B* **1980**, *36* (7), 1711–1712.

(96) Pons Siepermann, C. A.; Myerson, A. S. Inhibition of Nucleation Using a Dilute, Weakly Hydrogen-Bonding Molecular Additive. *Cryst. Growth Des.* **2018**.

(97) Tan, L.; Davis, R. M.; Myerson, A. S.; Trout, B. L. Control of Heterogeneous Nucleation via Rationally Designed Biocompatible Polymer Surfaces with Nanoscale Features. *Cryst. Growth Des.* **2015**, *15* (5), 2176–2186.

(98) Diao, Y.; Helgeson, M. E.; Siam, Z. A.; Doyle, P. S.; Myerson, A. S.; Hatton, T. A.; Trout, B. L. Nucleation under Soft Confinement: Role of Polymer–Solute Interactions. *Cryst. Growth Des.* **2012**, *12* (1), 508–517.

(99) Sun, X.; Garetz, B. A.; Myerson, A. S. Supersaturation and Polarization Dependence of Polymorph Control in the Nonphotochemical Laser-Induced Nucleation (NPLIN) of Aqueous Glycine Solutions. *Cryst. Growth Des.* **2006**, *6* (3), 684–689.

(100) Galkin, O.; Vekilov, P. G. Control of Protein Crystal Nucleation around the Metastable Liquid–liquid Phase Boundary. *Proc. Natl. Acad. Sci.* **2000**, *97* (12), 6277–6281.

(101) Kulkarni, S. A.; Kadam, S. S.; Meekes, H.; Stankiewicz, A. I.; ter Horst, J. H. Crystal Nucleation Kinetics from Induction Times and Metastable Zone Widths. *Cryst. Growth Des.* **2013**, *13* (6), 2435–2440.

(102) Ostwald, W. Studien Über Die Bildung Und Umwandlung Fester Körper. *Z. Für Phys. Chem.* **1897**, *22U* (1), 289–330.

(103) Wójcik, G.; Holband, J.; Szymczak, J. J.; Roszak, S.; Leszczynski, J. Interactions in Polymorphic Crystals of M-Nitrophenol as Studied by Variable-Temperature X-Ray Diffraction and Quantum Chemical Calculations. *Cryst. Growth Des.* **2006**, *6* (1), 274–282.

(104) Souvignet, I.; Olesik, S. V. Molecular Diffusion Coefficients in Ethanol/Water/Carbon Dioxide Mixtures. *Anal. Chem.* **1998**, *70* (14), 2783–2788.

(105) Na, H.-S.; Arnold, S.; Myerson, A. S. Cluster Formation in Highly Supersaturated Solution Droplets. *J. Cryst. Growth* **1994**, *139* (1), 104–112.

(106) Moroni, D.; ten Wolde, P. R.; Bolhuis, P. G. Interplay between Structure and Size in a Critical Crystal Nucleus. *Phys. Rev. Lett.* **2005**, *94* (23), 235703.

(107) Lee, A. Y.; Erdemir, D.; Myerson, A. S. Crystal Polymorphism in Chemical Process Development. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2* (1), 259–280.

(108) Davey, R. J.; Blagden, N.; Potts, G. D.; Docherty, R. Polymorphism in Molecular Crystals: Stabilization of a Metastable Form by Conformational Mimicry. *J. Am. Chem. Soc.* **1997**, *119* (7), 1767–1772.

(109) Weissbuch, I.; Lahav, M.; Leiserowitz, L. Toward Stereochemical Control, Monitoring, and Understanding of Crystal Nucleation. *Cryst. Growth Des.* **2003**, *3* (2), 125–150.

(110) Lee, I. S.; Kim, K. T.; Lee, A. Y.; Myerson, A. S. Concomitant Crystallization of Glycine on Patterned Substrates: The Effect of PH on the Polymorphic Outcome. *Cryst. Growth Des.* **2008**, *8* (1), 108–113.

(111) Pons Siepermann, C. A.; Myerson, A. S. Nucleation Inhibition by a Dilute, Weakly Interacting Impurity.

(112) Wijethunga, T. K.; Chen, X.; Myerson, A.; Trout, B. L. Application of Crystalline Substrates for Nucleation Control and Polymorphic Selection of Indomethacin. *Prep.*

(113) Aceves-Hernandez, J. M.; Nicolás-Vázquez, I.; Aceves, F. J.; Hinojosa-Torres, J.; Paz, M.; Castaño, V. M. Indomethacin Polymorphs: Experimental and Conformational Analysis. *J. Pharm. Sci.* **2009**, *98* (7), 2448–2463.

(114) Aubrey-Medendorp, C.; Swadley, M. J.; Li, T. The Polymorphism of Indomethacin: An Analysis by Density Functional Theory Calculations. *Pharm. Res.* **2008**, *25* (4), 953–959.

(115) Andronis, V.; Zografi, G. Crystal Nucleation and Growth of Indomethacin Polymorphs from the Amorphous State. *J. Non-Cryst. Solids* **2000**, *271* (3), 236–248.

(116) Kolmer, A.; Edwards, L. J.; Kuprov, I.; Thiele, C. M. Conformational Analysis of Small Organic Molecules Using NOE and RDC Data: A Discussion of Strychnine and α-Methylene-γ-Butyrolactone. *J. Magn. Reson.* **2015**, *261*, 101–109.

(117) Armand, P.; Kirshenbaum, K.; Goldsmith, R. A.; Farr-Jones, S.; Barron, A. E.; Truong, K. T. V.; Dill, K. A.; Mierke, D. F.; Cohen, F. E.; Zuckermann, R. N.; et al. NMR Determination of the Major Solution Conformation of a Peptoid Pentamer with Chiral Side Chains. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (8), 4309–4314.

(118) Ruidiaz, M. A.; Delgado, D. R.; Martínez, F. PERFORMANCE OF THE JOUYBAN-ACREE AND YALKOWSKY-ROSEMAN MODELS FOR ESTIMATING THE SOLUBILITY OF INDOMETHACIN IN ETHANOL + WATER MIXTURES. *Rev. Acad. Colomb. Cienc. Exactas Físicas Nat.* **2011**, *35* (136), 329–336.

(119) Ruidiaz, M. A.; Delgado, D. R.; Mora, C. P.; Yurquina, A.; Martínez, F. Estimation of the Indomethacin Solubility in Ethanol + Water Mixtures by the Extended Hildebrand Solubility Approach. *Rev. Colomb. Cienc. Quím. - Farm.* **2010**, *39* (1), 79–95.