2012

POLYMORPH FORMATION OF TOLFENAMIC ACID: AN INVESTIGATION OF PRE-NUCLEATION ASSOCIATION

Alessandra Mattei
University of Kentucky, alessandra_mattei@hotmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Alessandra Mattei, Student

Dr. Tonglei Li, Major Professor

Dr. Jim Pauly, Director of Graduate Studies

POLYMORPH FORMATION OF TOLFENAMIC ACID: AN INVESTIGATION OF PRE-NUCLEATION ASSOCIATION

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Pharmacy at the University of Kentucky

By

Alessandra Mattei

Lexington, Kentucky

Director: Dr. Tonglei Li, Professor of Pharmaceutical Sciences

Lexington, Kentucky

2012

ABSTRACT OF DISSERTATION

POLYMORPH FORMATION OF TOLFENAMIC ACID: AN INVESTIGATION OF PRE-NUCLEATION ASSOCIATION

The majority of pharmaceutical products are formulated as solids in the crystalline state. With the potential to exist in different crystalline modifications or polymorphs, each solid form bears its own physical and chemical properties, influencing directly bioavailability and manufacturability of the final dosage form. In view of the importance of crystalline form selection in the drug development process, it is imperative for pharmaceutical scientists to work arduously on various aspects of polymorphism, ranging from fundamental understanding of the phenomenon at the molecular level to practical utilization of a specific crystalline form. One common feature of organic crystals is the existence of distinct molecular conformations in different polymorphic structures, known as conformational polymorphism. Conformational polymorphs are routinely observed in drug development, produced when crystal growth conditions vary. Crystallization from solution involves nucleation and crystal growth, the mechanisms that influence the polymorphic outcome. The embryonic solute aggregate has been recognized to play a critical role in dictating the final crystal structure, and solution conditions are also known to drastically influence the self-association behavior of solute molecules during crystallization, affecting crystal packing of organic molecules. For the crystal growth of conformational polymorphs, changes in molecular conformation not only determine the growth kinetics, but also influence the nature and strength of interactions present in the crystal structures. How conformation and intermolecular interaction affect each other underlines the intricacy and the wonder of crystal growth of the organic. Thus, the overall goal of this research is to provide the fundamental understanding of the extent to which solution conditions influence the molecular conformation in the solid-state of a model drug, tolfenamic acid. By combining experimental studies with advanced computational tools, this dissertation offers novel insights into solution species during pre-nucleation and molecular packing of conformational polymorphs of tolfenamic acid. In-depth understanding of the underlying connection between molecular conformation and crystal packing will help advance the knowledge required for rational control of crystal growth.

KEYWORDS: polymorphs, nucleation, self-association, molecular conformation, intermolecular interaction

Multimedia Elements Used: TIF (.tif)

Alessandra Mattei
Student's Signature

September 2012
Date

POLYMORPH FORMATION OF TOLFENAMIC ACID: AN INVESTIGATION OF PRE-NUCLEATION ASSOCIATION

By

Alessandra Mattei

Tonglei Li, Ph.D.

Co-Director of Dissertation

Paul M. Bummer, Ph.D.

Co-Director of Dissertation

Jim Pauly, Ph.D.

Director of Graduate Studies

September 2012

Date

To my parents and my brother

ACKNOWLEDGEMENTS

I would like to express my gratitude to several people who have helped and supported me through the course of my graduate studies. First, I would like to thank my advisor Dr. Tonglei Li for his guidance, support, and patience. I would like to thank my dissertation committee: my co-advisor Dr. Paul Bummer, Dr. Bradley Anderson, and Dr. Carolyn Brock for their time, support, and contributions to my development and the advancement of this research.

I would also like to thank Dr. Steven Van Lanen and Dr. Zack Hilt for instrument support. I would like to give a special thanks to Dr. Anne-Frances Miller, Mr. John Layton, Xiaonan Mei, and Warintra Pitsawong for their training and guidance in NMR technique. Many thanks to the Institute for High Performance Computing and the Center for Computational and Computer Sciences at the University of Kentucky for providing the computer resources.

I would like to thank all of the members of Dr. Li lab. Thanks especially to Christin Hollis for her support, help, and friendship. On a personal note, I would like to thank my family for their unconditional love and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

API	Active pharmaceutical ingredient
TFA	Tolfenamic acid
CNT	Classical nucleation theory
PXRD	Powder X-ray diffraction
DSC	Differential scanning calorimetry
HSM	Hot stage microscopy
ATR-FTIR	Attenuated total reflectance-Fourier transform infrared
UV	Ultraviolet
NMR	Nuclear magnetic resonance
NOESY	Nuclear overhauser effect spectroscopy
PGSE	Pulsed gradient spin-echo
DFT	Density functional theory
HOMO	Highest occupied molecular orbital
LUMO	Lowest unoccupied molecular orbital
NBO	Natural bond orbital
MD	Molecular dynamics
RDF	Radial distribution function
QM/MM	Quantum mechanics/molecular mechanics

Chapter 1 – Introduction

Of the many states of matter in which a substance can reside, the solid state is the most commonly encountered and thus the most relevant phase in pharmaceutical development. The majority of pharmaceutical materials, either the active pharmaceutical ingredients (APIs) or the excipients, are produced and stored as solids. Additionally, most common drug products are manufactured and formulated as solid-dosage forms, such as tablets and capsules. Drugs for parenteral application are formulated as lyophilized products and dry powder inhalers are becoming popular for delivery of respiratory products.[1,2] Solid oral dosage forms offer greater chemical stability than liquid formulations, low manufacturing costs, high throughput, and patient compliance.

Based on the order of molecular packing, solids are classified into two major classes of crystalline forms and amorphous forms. In the crystalline state, molecules arrange into a highly regular fashion.[1] This molecular order extends in three dimensions over short and long ranges. In the amorphous state, on the other hand, only short-range order is found in neighboring molecules, and this regularity extends over distances of 1-10 nm. Crystalline forms are the focus of this dissertation. Most pharmaceutical products on the market, like most formulations in development, are in the solid crystalline form. Crystalline solids usually are selected and manufactured as APIs for several reasons, including purity, better physical and chemical stability, and ease of handling. This, however, does not mean that pharmaceutical solids, specifically crystalline solids, do not present their own challenges. Different crystalline forms of the same compound can exist and are termed polymorphs. An overview of polymorphism of organic crystals as it pertains to the

pharmaceutical industry, including current challenges and problems faced by pharmaceutical scientists, follows.

1.1 Polymorphism

Many drug substances and related organic compounds can exist as two or more crystalline phases that have different arrangements and/or conformations of the molecules in the crystal lattice,[3] giving rise to the phenomenon of polymorphism. Polymorphism was recognized as early as 1822 from crystal morphologies of inorganic crystals.[4] Certainly by the 1960s, an increasing realization of the importance of polymorphism had emerged within the pharmaceutical community.[5] Because polymorphs have different internal crystal structures, they may differ significantly in their physical and chemical properties.[1,6] Diamond, graphite, and fullerenes are polymorphic forms (denoted allotropes) of carbon all exhibiting very different properties as a result of the distinct packing of the carbon atoms within the crystal structure. Cocoa butter can crystallize in at least five different crystal structures affecting the perception of the fine quality of prepared chocolate.[7] These are but a few examples to emphasize the effect of differences in crystal structure on the properties of a solid.

Attempts to estimate the frequency of polymorphism in pharmaceutical organic compounds revealed that approximately 80 percent of marketed drug molecules exist in multiple polymorphic forms.[8] Another survey of 245 organic compounds reported that half were polymorphic.[9] Regardless of the statistical percentages, the results support the oft-quoted statement by McCrone that the number of polymorphs discovered for each compound is proportional to the time and effort spent in research on that compound.[10]

Thus, it is not surprising that the widely used analgesic ibuprofen had shown no evidence of polymorph behavior, until a recent discovery under a wide range of crystallization conditions led to a second form.[11] Yet drug polymorphism is a widespread phenomenon and continues to receive extensive academic and industrial attention, as highlighted by many books[1,3,8,12] and special journal issues[13-16] that are devoted to this topic.

Given the strong interest in and importance of multiple crystalline forms, it is not surprising that considerable efforts would be made to predict the possible existence of polymorphs. The ability to predict the number of crystalline forms that can be expected in a given case does not yet exist, although not for lack of effort.[17] The last two decades have seen enough of an increase in computer power to make the computational prediction of organic crystal structures a practical possibility, but polymorph prediction is still a long-term goal.[18]

1.1.1 Conformational Polymorphism

One common feature of organic crystals is the existence of distinct molecular conformations in different polymorphic structures, known as conformational polymorphism.[19] The energy difference among polymorphs of an organic molecule can be 2 kcal/mol or even lower.[19,20] This similarity in energy allows for conformationally flexible molecules to adopt different molecular arrangements in the various crystal structures. Conformational polymorphs are routinely produced when crystal growth conditions vary. The existence of distinctive molecular conformations in different polymorphic structures represents a unique opportunity for understanding the relationship between the structure of a material and the properties in question. 5-methyl-2-[(2-nitrophenyl)amino]-3-thiophenecarbonitrile, known as ROY, is a conformationally

flexible molecule, which assumes distinctive molecular conformations in various packing arrangements.[21] ROY has been crystallized in at least 10 polymorphs, seven with solved structures and three whose structures have not been solved. ROY crystals differ in color. Its crystal color is associated with the molecular conformation adopted by the constituent molecules in each crystal structure.

1.1.2 Disappearing and Concomitant Polymorphism

The interest in polymorphism arises from the opportunity to investigate subtle structure-property relationships, since the solid-state properties of a material are closely linked to its three-dimensional crystal structure. Polymorphism is a source of both fascination and frustration. The latter derives from the difficulty in obtaining a specific polymorphic form consistently and reproducibly — so-called disappearing polymorphs.[22] The phenomenon of disappearing polymorphism has been explained in terms of kinetic and thermodynamic considerations.[22] Initially, a kinetically favored crystal form may be produced as long as a more stable crystal is not isolated. If a thermodynamically more stable polymorph crystallizes, then the ability to obtain the metastable form may be compromised. There are well-documented cases[23,24] of crystal forms that were observed over a period of time and not thereafter, as they were displaced by more stable polymorphs. Often, unintentional seeding is a factor in the phenomenon of disappearing polymorphs.[3] Once a new form appears, the presence of seeds of that form may make the crystallization of the previously obtained polymorph very difficult under the same conditions. A seed may not necessarily be composed of the same molecule as the compound to be crystallized. Dust or other foreign particles can act as seeds in promoting crystallization. In any case, the phenomenon suggests a loss of control over the crystallization process. Substantial efforts

thereby have been devoted to the design of consistent and reliable crystallization processes to advance the trial-and-error experimentation. Real-time analysis using on-line sensors has become valuable in increasing crystal product quality as well as reducing the time to production. Various spectroscopic methods, such as attenuated total reflectance-Fourier transform spectroscopy and Raman spectroscopy coupled with fiber optics, have been used for on-line monitoring during the operation of crystallization processes.[25,26]

The conditions of a crystallization process can often lead to the production of multiple polymorphs called concomitant polymorphs[27] due to an overlap of the occurrence domains for different polymorphs. The domain that allows solid forms to exist concomitantly is rarely known, as many growth conditions (i.e., solvent, supersaturation, temperature, and cooling) influence the crystallization outcome. Crystallization in polymorphic systems is governed by a combination of thermodynamic and kinetic factors. As such, the appearance of concomitant polymorphs may arise either because specific thermodynamic conditions prevail or because the kinetic processes have equivalent rates. A proper understanding of the underlying fundamentals of the processes leading to multiple solid forms could help in two main ways. First, chemists and chemical engineers could develop a more robust process for consistent isolation of a single, pure solid form; second, formulation chemists could select the appropriate crystal form based on structural and thermodynamic information.

1.1.3 Thermodynamic Aspects of Polymorphism

Polymorphs of an organic crystal have different energies and thermal stabilities. The thermodynamic stability relationship between different phases of a compound is governed by the Gibbs free energy (G), which is defined as:

$$G = H - TS \quad (1.1)$$

where H is the enthalpy; T is the temperature; and S is the entropy. At a given temperature, the polymorph with the lowest free energy is the stable polymorph. There is only one thermodynamically stable polymorph under a given temperature and pressure, while the others are referred to as metastable forms. One important distinction is whether one form can transform reversibly to another at the so-called thermodynamic transition point or temperature. At the thermodynamic transition temperature two polymorphs have the same stability. Energy-temperature diagrams are commonly used to describe the thermodynamic behavior of polymorphs (Figure 1.1). Two crystalline forms are classified as enantiotropes if a reversible transition temperature is located below the melting temperatures of both polymorphs (Figure 1.1a). Below the transition temperature, Form I is more stable because the free energy of Form I is lower than that of Form II. In such a case, any transformation from the higher free energy form (Form II) to the lower free energy form (Form I) is exothermic below the thermodynamic transition temperature. Above the transition temperature, Form II is the stable solid phase. A polymorphic system is monotropic if the transition temperature is above the melting points of both polymorphs and the phase transformation can take place only in one direction. Throughout the temperature range, Form I is more stable (Figure 1.1b).

A plot of the Gibbs free energy difference (ΔG) against the absolute temperature (T) gives complete and quantitative information on the stability relationship of polymorphs.[28,29] The Gibbs free energy change and its temperature dependence can be experimentally determined from melting data (melting temperature and enthalpy of

fusion). Thermodynamic rules also have been developed by Burger and Ramberger[30,31] to qualitatively determine the enantiotropic or monotropic nature of the relationships between polymorphs. These rules are referred to as the heat of fusion and heat of transition rules and are based on calorimetric measurements. Other guidelines reported in literature include the density and infrared rules, but they are not generally applicable due to several notable exceptions.[31]

In thermodynamic terms, only at the transition temperature can polymorphs have equal free energy and coexist in equilibrium; at any other temperature there will be a tendency to transform to the most stable structure because of the thermodynamic drive toward minimizing the free energy of the system. This implies that mixtures of two polymorphs have limited lifetimes, with transformation kinetics playing a role in those lifetimes.

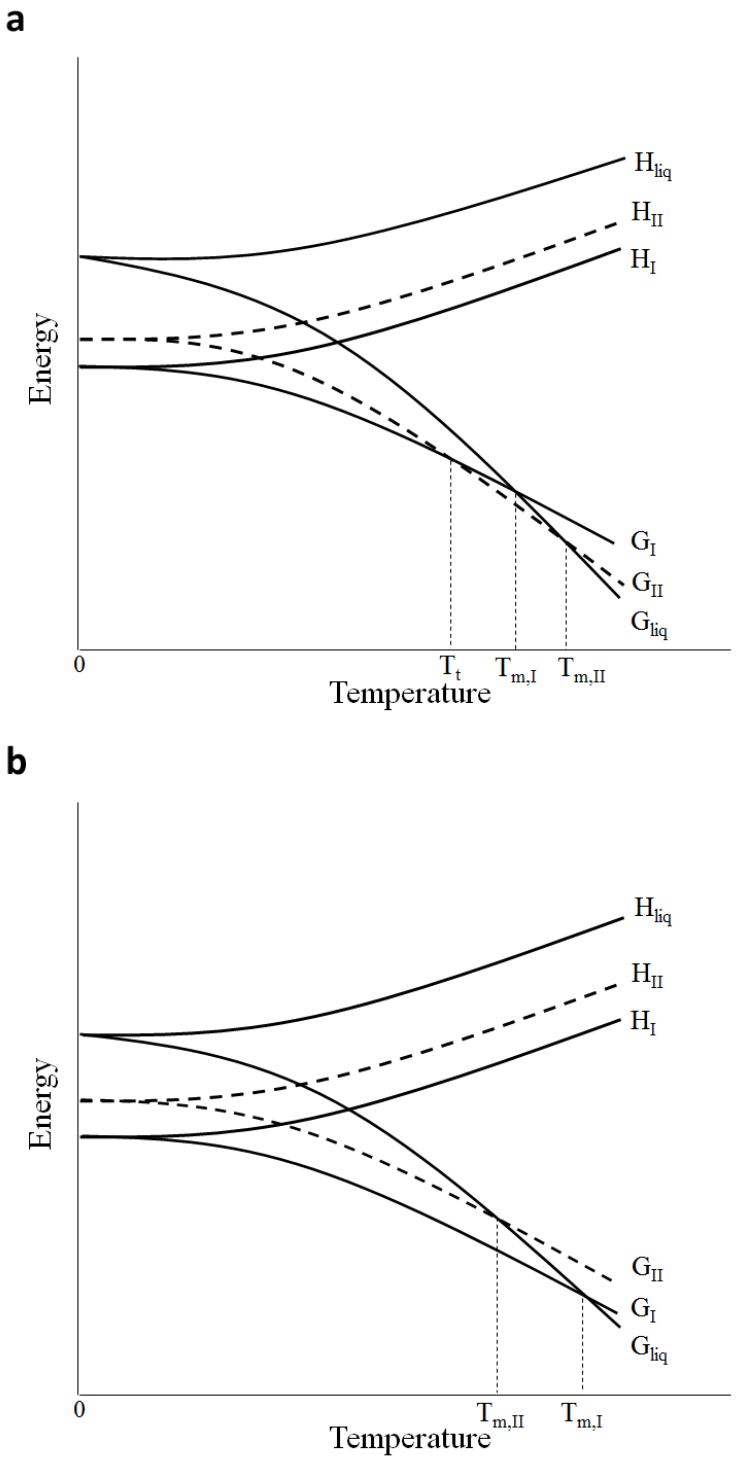

Figure 1.1. Free energy diagrams for enantiotropic (a) and monotropic (b) polymorphic systems. G is the Gibbs free energy, H is the enthalpy, and T is the temperature. Subscripts I, II, and liq refer to polymorph I, polymorph II, and liquid phase, respectively, while subscripts t and m refer to transition and melting temperatures, respectively. (Adapted from Burger and Ramberger).[31]

1.1.4 Polymorph Transformations

The physical stability of crystalline solid forms depends on environmental conditions. During product development, crystalline solids may be exposed to solvents, mechanical stress, or changes in temperature and humidity, each of which can induce phase transformations. As such, it is critical to identify the solid phases and to recognize the transitions between them under relevant conditions. A mechanistic understanding of phase transformations can facilitate rational formulation design and selection of robust processes to ensure consistent product manufacturing and performance.

The rearrangement of molecules into a new structure during phase transformation may or may not involve a solvent or vapor phase. Considering a solid-solid physical transition, a four-step mechanism has been proposed: (a) molecular loosening in the initial phase; (b) formation of an intermediate solid solution; (c) nucleation of the new solid phase; and (d) growth of the new phase.[32] In general, phase transitions via solid-state mechanisms are influenced by factors such as environment (e.g., temperature, pressure, and relative humidity), the physical characteristics of the crystal (e.g., crystal habit, size, and presence of defects), and the presence of impurities.

If crystals grow from and remain in contact with solution, a phase transformation likely takes place via solution by dissolution and recrystallization.[33,34] Solution-mediated phase transformation is a common and effective strategy to transform high-energy, metastable forms to more stable polymorphs, particularly in cases in which concomitant polymorphism[27] occurs as the more stable phase grows at the expense of the less stable form. The phase transformation process involves three main steps: (a) dissolution of the metastable phase; (b) nucleation of the stable form; and (c) crystal growth of the stable

phase.[35] Phase transformation in solution is encompassed by Ostwald's Rule of Stages,[36] which describes a step-wise transformation from the metastable phase to a more stable form. The conversion is driven by the difference in free energy between the metastable and stable solid phases, which is generally reflected as the difference in solubility between the two phases. When a metastable polymorph is suspended in a solvent, dissolution proceeds to its solubility limit. Meanwhile, the dissolved solute creates a driving force for the crystallization of the stable polymorph. As such, the metastable form transforms into the stable polymorph via dissolution and crystal growth processes.

The extent of conversion depends upon the amount dissolved with time, which in turn is related to the solubilities of the forms. Eventually the system reaches its thermodynamically most stable phase, when the stable polymorph is the only one that is present in the final suspension.

1.2 Pharmaceutical Relevance and Implications of Polymorphism

Crystal polymorphism is especially relevant to pharmaceutical development. The arrangement or packing of the molecules in a crystal can and does lead to alteration in the physical, chemical, and mechanical properties of the solid.[1,6] Drug properties vital to the development of a quality drug product are bioavailability and solid-state stability. Solubility and dissolution rate are physical characteristics directly related to bioavailability. Differences in solubility may have implications on the absorption of the active drug from its dosage form[37] by affecting the dissolution rate and possibly the mass transport of the molecules. Higher solubility and faster dissolution rate can lead to measurable increases in bioavailability[38] and presumably therapeutic efficacy. However,

a solid form with higher solubility or faster dissolution rate is metastable and tends to convert to a thermodynamically more stable polymorph over time, as discussed in the previous section. The perceived physical and chemical instability of metastable forms, leading to inconsistent product quality, has limited their use in drug products.

Development of a metastable crystal form can affect the pharmaceutical performance of a drug product, resulting in failures in a marketed product. A well-known example of the importance of controlling the polymorphic form is the antiretroviral drug Ritonavir.[39] A more stable and less soluble crystalline phase appeared in the formulation that failed dissolution testing. Ultimately, the pharmaceutical product was withdrawn from the market because the manufacturing process was no longer able to consistently and reliably produce the desired polymorph. The product was then reformulated using the most stable polymorph.

Despite the potential benefits of higher solubility associated with metastable forms, they are seldom the candidates of choice for formulation development.[2] In general, significant resources are spent in the early development process to find the thermodynamically most stable polymorph.

Selection of a desirable crystal form is an important step in the initial stage of drug development and is required to be in line with the quality-by-design initiatives encouraged by the Food and Drug Administration (FDA) and other regulatory agencies.[40] Polymorph screening is an essential activity. The intent is to uncover all possible crystalline forms and select an optimal solid form suitable for development. However, it is necessary to be aware that polymorph transformations may still occur in common

pharmaceutical processes (i.e., milling, wet granulation, drying, and compression), even when a stable crystal form is selected.[41]

Because of the significance of polymorphism in the drug development process, numerous experimental observations have been made, stimulating interest in various topics ranging from practical utilization of a specific crystal form to fundamental understanding of the phenomenon. There has been great progress in controlling the crystal growth process in order to selectively obtain a desired polymorph or inhibit the growth of an undesired one.[42] To gain control over the polymorph formation, an understanding of crystal growth is crucial. Therefore, the fundamentals of crystallization are discussed extensively in the next section. It is important to provide a general overview of nucleation of crystals from solution and to present challenges and limitations that led to focusing the research, conducted for the purpose of this dissertation, towards the early events of nucleation.

1.3 Nucleation

Crystallization from solution involves nucleation and growth. Nucleation starts from the formation of small embryos of a new crystalline phase in a supersaturated solution. A free energy penalty is generally associated with the appearance of the nuclei, because of the creation of the solid surface. Overcoming the nucleation energy barrier dynamically depends on supersaturation as well as molecular packing of nuclei. A solute will remain in solution until a sufficiently high level of supersaturation is developed to induce spontaneous nucleation. The extent of this supersaturation is referred to as the metastable zone width. In the concentration-temperature relationship (Figure 1.2), the lower line is the equilibrium solubility of the solute while the upper line is the metastable limit, which

is defined as the boundary of stability of a solution. Supersaturated solutions are thereby classified as either labile (if spontaneous nucleation occurs) or metastable (if it does not).[43] In thermodynamic terms, a supersaturated solution is the one in which the solute chemical potential is higher than that of the crystalline solute.

Early events of nucleation play a decisive role in determining the resultant crystal structure.[44,45] A current viewpoint is that solute molecules may form supramolecular, pre-nucleation species, which act as the precursors of crystal growth.[46] The process of nucleation is complex in terms of theoretical description and experimental observation. The size of nuclei and the stochastic nature of the process make measurements of the actual event challenging; yet, nucleation has been recognized and described by various kinetic models, which can provide valuable insight into the nucleation process.

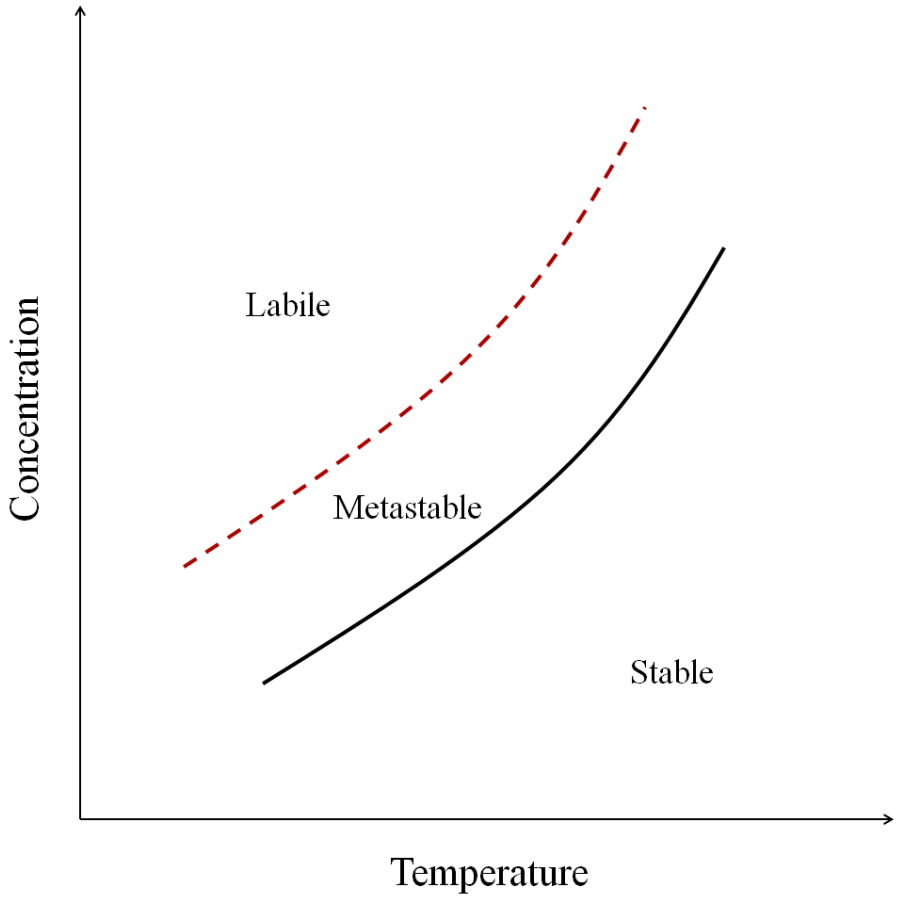

Figure 1.2. Schematic concentration-temperature diagram showing the equilibrium solubility curve (solid line) and the metastable zone limit (dashed line). (Adapted from Mullin)[43]

1.3.1 Classical Nucleation Theory

Classical nucleation theory (CNT) has been widely used in explaining nucleation kinetics.[47-49] In a supersaturated solution, the formation of a new phase requires a free energy change (ΔG), which is given by the sum of the free energy change needed to create a new surface (ΔG_s) and the free energy change for the phase change (ΔG_v), according to the equation:

$$\Delta G = \Delta G_s + \Delta G_v = 4\pi r^2 \gamma + \frac{4}{3}\pi r^3 \Delta G_v \quad (1.2)$$

The change in surface free energy increases with the interfacial tension (γ) between the solid crystal surface and the surrounding solution, as well as the surface of the nucleus. The surface energy change is to be added to the system and is thus positive. On the other hand, the free energy change for the phase transformation is negative because the chemical potential of the crystalline phase is less than that of the solute in solution and is proportional to the volume. Therefore, for a spherical nucleus with radius r, the total free energy change depends on the competition between a decrease in the volume free energy change, which favors growth, and an increase in the surface free energy change, which favors dissolution, as shown in Figure 1.3. In the initial stage of nucleus formation, at a very small radius, the surface free energy term dominates over the volume free energy term and acts to destabilize the nucleus. As a consequence, the nucleus is unstable and tends to dissolve. However, at a sufficiently large radius, the energy reduction from the volume term becomes higher than the energy increase from the increased surface area. Under these conditions, the total free energy decreases continuously and growth becomes energetically favorable, resulting in the crystal formation. The radius where this transition occurs is known as the critical radius (r_c) and the free energy change at the critical radius is called the activation barrier for nucleation (ΔG_{crit}).

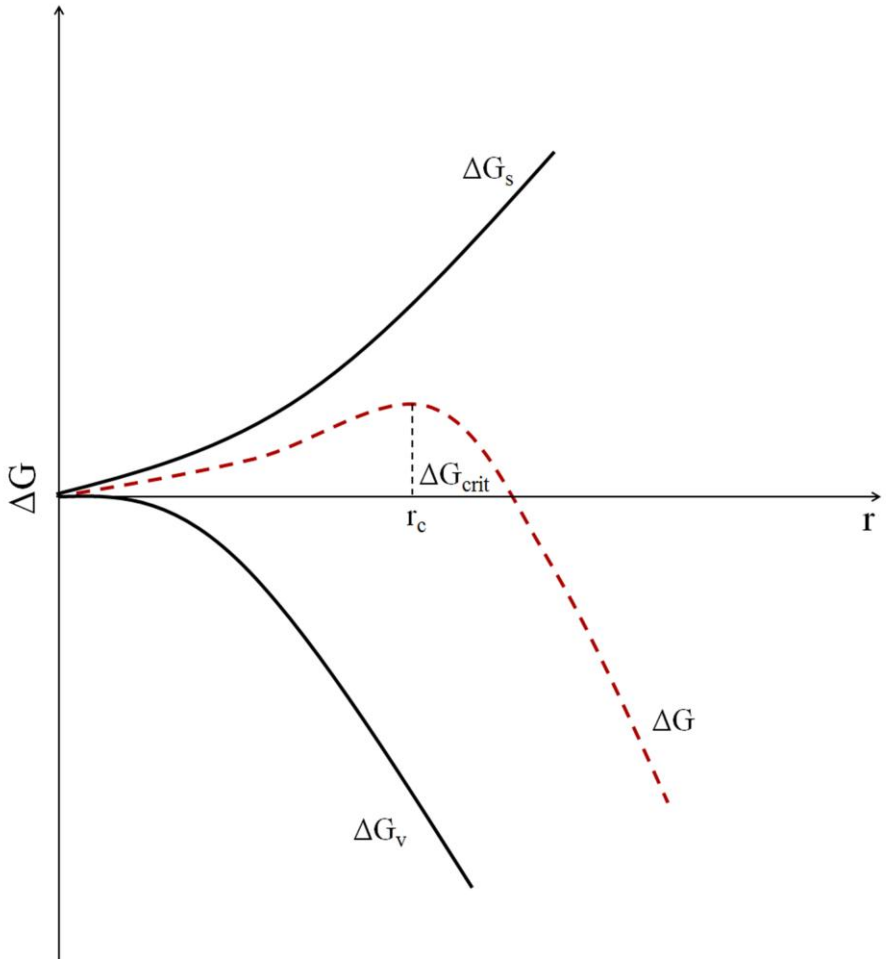

Figure 1.3. Free energy diagram for nucleation. (Adapted from Mullin)[43]

The rate of nucleation is expressed by the Arrhenius equation:

$$J = A\,\mathrm{Exp}\left(-\frac{\Delta G_{crit}}{kT}\right) \tag{1.3}$$

where J is the rate of nucleation, which gives the number of nuclei formed per unit time and per unit volume; *k* is the Boltzmann constant; T is the absolute temperature; and *A* is the pre-exponential factor, which represents the number of molecules per unit volume (N_0) multiplied by the frequency at which the nuclei become supercritical and transform

into crystals (ν_0).[50] In the case of nucleation from condensed phases the attachment frequency (ν_0) is related to the diffusion process, according to the equation:[49]

$$\nu_0 = (48\pi^2 \upsilon)^{1/3} D\, C\, n^{*1/3} \tag{1.4}$$

where υ is the molecular volume; D is the monomer diffusion coefficient; C is the monomer concentration in the bulk solution; and n* is the number of molecules in the critical radius nucleus.

Using the Gibbs-Thomson equation, ΔG_{crit} can be expressed as:

$$\Delta G_{crit} = \frac{16\pi \gamma^3 \upsilon^2}{3(k\, T \ln S)^2} \tag{1.5}$$

where υ is the molecular volume; γ is the solid-liquid interfacial free energy; and S is the supersaturation, defined as the ratio of the actual solution concentration to the equilibrium solubility of the corresponding crystalline form. Note that as supersaturation increases, the free energy barrier to nucleation decreases, and in turn the rate of nucleation increases.

In principle, the described theory has been derived by considering homogeneous nucleation, which occurs in the volume of ideally pure solutions (i.e., solutions containing only solvent and solute molecules). It is nonetheless notoriously difficult to avoid heterogeneous nucleation, which takes place in solutions containing impure molecules and/or foreign microparticles that can provide active centers for nucleation. In addition, the walls of a container can serve as surfaces for heterogeneous nucleation. The thermodynamic approach characterizes crystal growth by macroscopic quantities of

surface and bulk free energies, taking no account of any molecular features of a growing system. It assumes that the formation of embryos takes place by the addition of one monomer at a time. Because nucleation of the new phase is the result of fluctuations that bring together sufficient numbers of molecules to exceed the critical size, only the size criterion is used to decide whether aggregates become critical. Furthermore, CNT can neither identify nor be used to study the different pathways leading from solution to solid crystal.[51] The solution-to-crystal phase transformation involves changes in both density and periodic structure. CNT assumes that the assembly of molecules proceeds in an ordered array as a result of simultaneous fluctuations in density and structure. As such, the molecular structure of a nucleus is the same as that of the final crystal. However, when a crystal's embryo first appears, it may differ from the eventual bulk equilibrium phase, and subsequent changes to the new periodic structure may take place only during later stages of growth.[51]

There are instances in which a crystal's embryo is not the equilibrium phase that resembles the bulk crystal, simply because from a kinetic standpoint the energy barrier leading to a less stable, more disordered state is lower than the one leading to the most stable state.

CNT regards nucleation of a metastable polymorph as bearing a smaller surface energy and subsequently overgrowing the most stable form (i.e., Ostwald's Rule of Stages).[36] Yet there are cases of crystallization in which the most stable forms appear first.[52,53] According to CNT, once an embryo is formed it is committed to a particular polymorph. Yet there are cases in which multiple forms are produced simultaneously.[27] Therefore, the situation as applied to polymorphic systems is by no means as clear as might be inferred

from CNT. Recent studies suggest that pre-nucleation aggregates or clusters exist in solution and may be structurally different from individually solvated solute molecules.[45,54]

1.3.2 Two-Step Nucleation Model

Recognition of the existence of pre-nucleation aggregates, as well as recent crystallization kinetic studies, have set the stage to move beyond CNT.[55,56] Computational simulations[57,58] and experimental studies of proteins[55,59,60] imply the possibility that nucleation comprises at least two stages: aggregation of solute molecules into a disorderly packed ensemble and subsequent reorganization of such a cluster from which crystals nucleate (Figure 1.4).

Formation of a transient phase prior to transforming to the packing structure adopted by the bulk is a commonly observed feature, not only by macromolecular systems,[61] but also by relatively small molecules.[62,63] Indeed, the crystallization of calcium carbonate at room temperature in a pure solution proceeds by the nucleation of an amorphous precursor phase that later transforms to its lower energy crystalline counterpart.

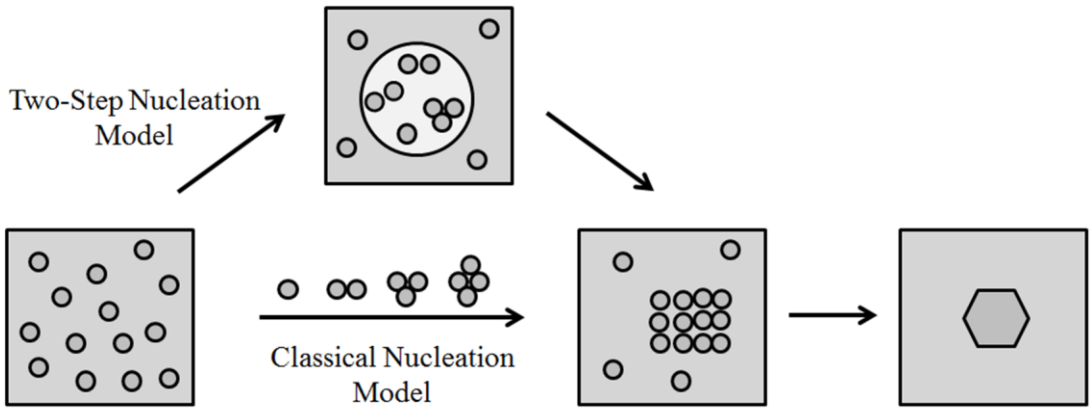

Figure 1.4. Schematic diagram of the two-step nucleation model compared to the classical nucleation model. (Adapted from Erdemir et al.)[60]

The applicability of the two-step mechanism was further deduced by non-photochemical laser-induced nucleation studies of glycine aqueous solutions.[64] As a result of intense laser pulses shone on supersaturated solutions, the nucleation rate increased; in addition, depending upon the laser polarization state, different glycine polymorphs (γ and α) preferentially nucleated. These observations imply alignment of molecules in existing pre-nucleation clusters in the solution, thus aiding the clusters in organizing into a crystal-like entity. If the nuclei form by successive aggregation of molecules in an ordered manner as proposed by CNT, induced alignment of molecules would not cause a significant change in the structure of already ordered aggregates, nor lead to the crystallization of distinct polymorphs.

A kinetic model has been developed to describe the two-step nucleation mechanism.[65] Unlike the kinetic model of CNT, the kinetic model of the two-step mechanism accounts for a correct temperature dependence of the nucleation rate. The incorrect temperature dependence of CNT can be explained by two physical features that are omitted from

CNT.[66] First, the interfacial free energy of a nucleus should depend on curvature. In contrast, the interfacial free energy of a nucleus is considered to be the same as that of an infinite planar surface. Second, the nucleation barrier should vanish at the metastable limit, where the solution loses stability.

The nucleation rate law of the two-step mechanism is based on the assumption that in the supersaturated solution an intermediate disordered cluster is formed with a temperature-dependent and concentration-dependent rate, u_0(C,T). This ensemble can dissociate back into the solution with rate u_1(T) or transform into an ordered crystal nucleus with rate u_2(T). The process can be described by the energy landscape picture (Figure 1.5). The probability of finding the system in state i = 0, 1, or 2 at time t is P_i(t); then a parameter τ that determines the mean time to create one crystalline nucleus in a steady-state process is defined as:

$$\tau = \int_0^\infty \mathrm{t} \left[\frac{d\, \mathrm{P}_2(\mathrm{t})}{d\, \mathrm{t}} \right] d\, \mathrm{t} \tag{1.6}$$

Thus, the parameter τ for the transition from state 0 to state 2 is given by:

$$\tau = \frac{1}{u_0(\mathrm{C},\mathrm{T})} + \frac{u_1(\mathrm{T})}{u_0(\mathrm{C},\mathrm{T})\, u_2(\mathrm{T})} + \frac{1}{u_2(\mathrm{T})} \tag{1.7}$$

The rates u_0, u_1, and u_2 depend on temperature as:

$$u_i(\mathrm{T}) = \mathrm{U}_i \, \mathrm{Exp}\left(-\frac{\mathrm{G}_i}{k\mathrm{T}} \right) \tag{1.8}$$

Therefore, the steady-state nucleation rate (J) can be calculated by:

$$J = \tau^{-1} = \frac{U_0 U_2 \, \mathrm{Exp}\left(\frac{G_0 + G_2}{kT}\right)}{U_0 \, \mathrm{Exp}\left(-\frac{G_0}{kT}\right) + U_1 \, \mathrm{Exp}\left(-\frac{G_1}{kT}\right) + U_2 \, \mathrm{Exp}\left(-\frac{G_2}{kT}\right)} \quad (1.9)$$

where U_0, U_1, and U_2 are pre-exponential factors that account for formation and decay of the transient cluster and formation of the ordered crystalline nucleus, respectively. G_0, G_1, and G_2 represent the energy barriers required for formation of the disordered cluster, decay of the disordered cluster, and formation of an ordered crystalline nucleus, respectively.

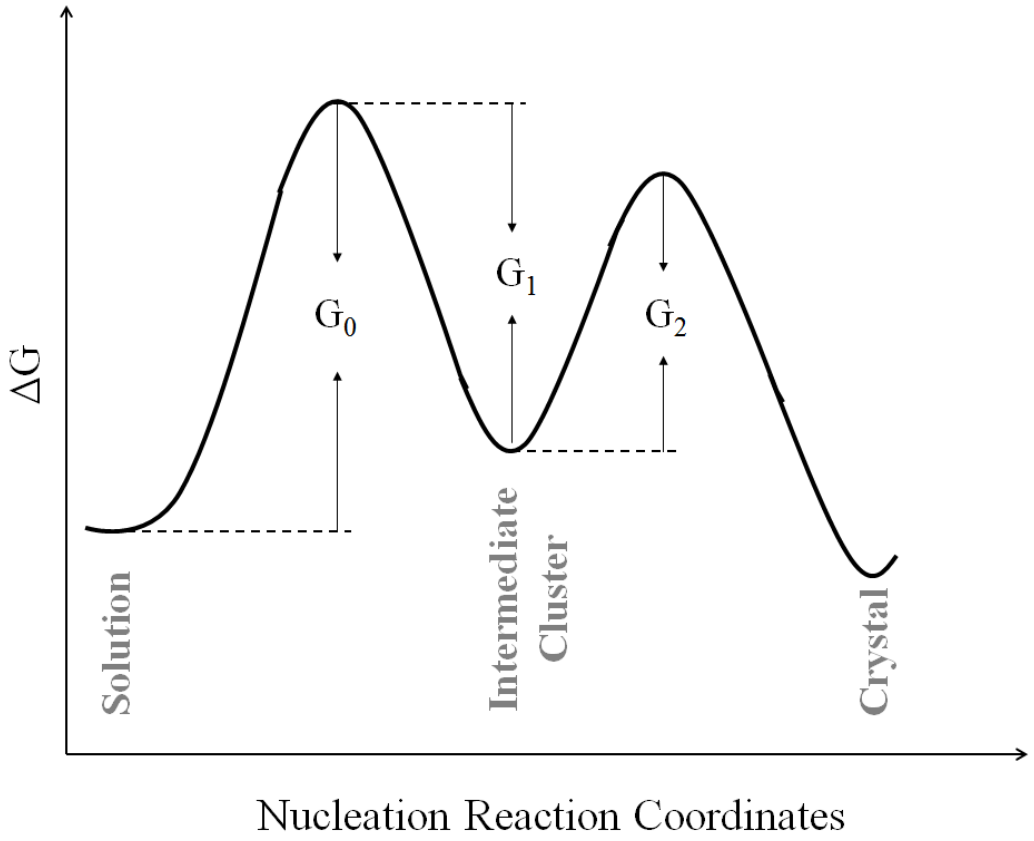

Figure 1.5. Free energy along the pathway for nucleation of crystals from solution. G_0, G_1, and G_2 are the energy barriers for formation of the intermediate cluster, decay of the disordered cluster, and formation of an ordered crystal nucleus. (Adapted from Pan et al.)[65]

A free energy advantage may be gained by forming a transient cluster and subsequent crystallizing, instead of following the classical route directly from individual solute molecules to the formation of structured crystalline nuclei.[67] Correspondingly, the nucleation rate is only determined by the assembling kinetics of transient clusters.[68] Such a theorem contradicts the paradigm of CNT.[56,69]

The two-step nucleation theory helps explain nucleation kinetics. Its applicability relies on the availability of disordered clusters in the homogenous solutions prior to nucleation. While such clusters have been demonstrated for several protein systems and for calcium carbonate solutions, it is likely that not all solutions would support the existence of such a transient phase with poorly defined properties allowing the nucleation of crystals within them. The main criticism of the two-step nucleation model is the physical meaning of the transient, disordered phase. Since this model does not consider the molecular restructure in the intermediate cluster, no insight into formation of distinct crystalline forms is gained.[69] Therefore, the use of this kinetic model will not help address how growth conditions (i.e., solvent, additives, and trace impurities) can affect crystal growth. From a dynamic perspective, association of solution species to form pre-nucleation aggregates is considered an important initial step. Nonetheless, formation of a crystalline phase may occur through successive aggregation of pre-assembled growth units, rather than the monomer addition model, as suggested by CNT, or the structural organization within the disordered phase, as proposed by the two-step nucleation model.[54]

1.4 Mechanistic Understanding of Polymorph Formation

1.4.1 Pre-Nucleation Association

Pre-nucleation aggregates may not necessarily be randomly packed; rather, they may bear a resemblance to the structure adopted by the subsequently crystallized solid phase. The concept of pre-nucleation aggregates was promoted by Etter[44] and Weissbuch.[45] Molecules self-assemble into various states of association in the solution phase; the most stable (i.e., most populated) association of molecules gets carried out in the crystallization process. This has been reported in the growth of 2,6-dihydroxybenzoic acid crystals.[70] It was found that carboxylic acid dimers were present predominately in toluene from which a dimer-based crystal form was produced, while polymeric associated species appeared to dominate in chloroform, leading to a hydrogen-bonded catemer observed in the harvested crystal structure. Similar behaviors were observed for tetrolic acid.[71] Hydrogen-bonded dimers were present in chloroform, giving rise to a dimeric polymorph; dimers were absent in ethanol, from which a catemer-based crystal was obtained.

Molecular simulation approaches, which were applied to study the early stages of molecular association of tetrolic acid, supported the experimental observations.[72] Weak interactions (i.e., low solvation free energy) between the solvent and tetrolic acid molecules prompted two solute molecules to assemble into a hydrogen-bonded cyclic dimer. This explained why crystals obtained from chloroform pack into a dimer-based structure. As solute-solvent interactions became stronger (i.e., high solvation free energy) the formation of dimer motifs from two solute molecules was unfavorable. As such, strong solute-solvent interactions caused tetrolic acid molecules to crystallize into a

catemer-based structure. Similarly, molecular modeling of 5-fluorouracil[73] in nitromethane have shown that initial nucleation of the metastable Form II was directed by the self-association of solute molecules to form hydrogen-bonded dimers. In wet nitromethane and in water, interactions of water molecules with 5-fluorouracil hindered the formation of hydrogen-bonded dimer motifs, leading to the crystallization of the polymorphic Form I.

Solution chemistry of glycine also has been extensively studied.[74-76] Based on observations by atomic force microscopy and grazing incidence X-ray diffraction,[45,77] it is suggested that glycine molecules in neutral aqueous solutions behave primarily as hydrogen-bonded dimers, resulting in a crystal structure in which the dimer motif is retained. Note that the speculation, mainly based on the measurement of the nucleation process at the crystal-liquid interface, was most recently questioned by Huang, et al.[78] Freezing point depression and diffusion coefficient measurements of supersaturated aqueous solutions of glycine have revealed that the solutions are mainly monomeric glycine. The dominant concentration, but not the exclusive presence of glycine monomers, brings into question the common perception of long-lived hydrogen-bonded cyclic dimers as the predominant solution species, serving as units of crystal growth. Yet the finding is consistent with the idea of an equilibrium among the states of association in solution (i.e., monomer and dimer for glycine); nucleation and subsequent crystal growth drive that equilibrium to the product phase.

Together these models suggest a hierarchy of clustering, from individual molecules through dimers to higher-order association species with the critical mass of a nucleus and subsequent crystal formation. Although these efforts provide indirect and/or direct

evidence for the existence of pre-nucleation association, the molecular mechanism of nucleation mainly remains speculative and poorly understood. There is a need to address fundamental questions of how each kind of assembly interacts to give rise to the ultimate production of various crystalline forms. In addition, most of the polymorphic systems studied were organic molecules with different packing motifs in the crystal structure. To date, very little information is available on polymorphic systems exhibiting distinct molecular conformations in different crystal structures. For that reason, molecular details of the specific arrangement of molecules in the associated species and of how molecular conformation affects the self-assembly process are lacking.

1.4.2 Solution Conformation

Determinations of crystal structure of flexible molecules carry the caveat that the crystal lattice conformation is not necessarily identical with that found in solution.[79] Moreover, all energetically accessible conformations of an organic molecule should be present in solution. The population distribution of all conformers is a function of the energy associated with each conformer and the energy gained from the interaction of the molecule with its surroundings.

Saito et al.[80] recently probed molecular association in solutions of *p*-acetanisidide and concluded that molecules associate to form hydrogen bonding motifs similar to those in the crystallizing solid form. The conformation adopted by the molecules within the aggregates was nonetheless different from the molecular conformation found in the crystal. There are many studies showing crystal lattice conformations differing from

solution conformations.[81-83] In such cases, the solution conformation has to adopt the conformation suitable for crystal packing before integrating into the crystal surface.

For systems having low energy barriers between conformers (i.e., lower than 2 kcal/mol), intermolecular interactions can facilitate a change of molecular conformation into one that is suitable for crystal packing. On the one hand, a molecule prefers to take its most energetically favorable conformation; on the other hand, the molecule may have to adjust its spatial arrangement in order to satisfy intermolecular hydrogen bonding. The energy of a conformational change is often comparable with that of intermolecular interaction. Thus, it is very likely that the lattice crystal features a high-energy conformer to accommodate orientation requirements imposed by certain growth synthons; intermolecular interactions supply the energy to stabilize the more highly energetic molecular conformation.[79,84]

It seems plausible to think that when the solution conformation differs from the lattice conformation, the transition between conformers may affect the nucleation kinetics. Specifically, this is important when the energy barrier between conformers is higher than 2 kcal/mol. For such systems, it is believed that multiple conformers in solution decrease the population of the conformer that crystallizes.[85] This in turn can reduce the degree of supersaturation and thus the crystal nucleation rate. The literature contains a few reports relating the difficulty of crystallization to the presence of multiple conformations in solution. For instance, the slow crystallization of conformational polymorphs of alditols (e.g., sorbitol and iditol) has been attributed to the high energy barrier of conversion between conformers.[86] As such, it is likely that only one conformer, referred to as the right conformer, integrates into the crystal surface. Conformers that do not integrate into

the crystal surface may act as impurities. It is believed these impurities preferentially adsorb at specific crystal faces. The adsorbed impurity essentially replaces the host molecules and hinders or slows adsorption of more solute molecules to this face. The change in growth kinetics occurs because the impurity molecule emerges from the crystal surface in a different orientation than the host molecule, disrupting the previous repeating pattern.[87]

1.4.3 Control of Polymorph Formation

The control of the nucleation process is the ultimate means of controlling polymorph formation. Several approaches for controlling nucleation rely on the concept that the crystal structure must reflect the state or states of association in solution. Selective nucleation of polymorphs has been pursued through the use of "tailor-made" additives.

It is believed these tailor-made or structurally similar additives interact with the pre-nucleation aggregates of a specific polymorph and inhibit their growth. Consequently, unaffected phases, although less stable, will grow.[45,88] Davey and coworkers advanced the idea of conformational mimicry for controlling the polymorphism of conformationally flexible molecules.[89] Using rigid additives that mimic the molecular conformation in the stable β polymorph of L-glutamic acid, they were able to selectively prevent its appearance and, thus, crystallize the metastable α polymorph.

Related molecular level strategies for controlling nucleation of organic crystals hinge on rational design of surfaces that can provide sites for a stereochemically induced (i.e., epitaxial) nucleation of a given crystal. Several types of nucleation-promoting surfaces have been studied, including two-dimensional ordered surfaces,[90] molecular single-

crystal surfaces,[91] and polymer surfaces.[92] On these surfaces, polymorph selectivity is thought to be achieved through either a structural lattice matching[93] or intermolecular interactions[94] at the crystal-substrate interface. The first theory relies purely on geometric principles and suggests that matched interactions between the substrate and the matching lattice plane of a growing nucleus lead to reduced activation energy for nucleation, causing selective and oriented growth of a single polymorph. The second theory proposes that chemical interactions direct the nucleation of crystals subsequently grown from the surface. Carter and coworkers reported that self-assembled monolayer substrates, bearing hydroxyl group functionality, selectively stabilize the nuclei of polymorph III of anthranilic acid due to specific hydrogen bonding between contacting planes of the substrate and nucleant.[95] Therefore, the assembly of molecules in solution, as well as in the solid-state, underpins much of the research field in polymorphism. The subject is driven by the understanding of intermolecular interactions and their use to control crystal structures.

1.5 Statement of Problems

Although there has been considerable progress in controlling the nucleation process on a molecular level, understanding of the nucleation mechanism is still challenging both computationally and experimentally. Advances in computational methods, although promising, are not yet sophisticated enough for studying the nucleation of a molecular crystal from solution, which requires a comprehensive consideration of numerous degrees of freedom introduced by solvent, conformation of the molecule, and multiple molecules in the asymmetric unit cell.[96] Also, the length and time scales of the nucleation process

span many orders of magnitude, making it difficult for a given modeling technique to capture all aspects. From the experimental viewpoint, detecting the occurrence of nuclei formation is not an easy task due to limitations in the resolution of currently available instruments. Critical nuclei usually consist of hundreds of molecules and typically are only a few nanometers in size.[43] By the time a crystal nucleus appears, it has already grown to macroscopic size. Thus, the nucleation mechanism can only be inferred from the pre-nucleation solution chemistry or the post-nucleation observation, when crystals are harvested.

The pre-nucleation approach explains polymorph formation from a thermodynamic point of view, while kinetic factors are overlooked. However, polymorph formation is known to be affected by a wide variety of growth conditions, including type of solvent,[97-99] supersaturation,[100] temperature,[101] and use of additives.[53,102] Little is known regarding the mechanism of the self-assembling process of crystallization under various growth conditions. As such, why a unique crystal structure is formed in a specific solvent or why some molecular compounds tend to crystallize in multiple polymorphs concomitantly under the same conditions remains unanswered.

Besides the intricate kinetic effects, nucleation of organic molecules from solution presents additional complexities stemming from flexible molecular conformations and weak intermolecular interactions. The association state of molecules in solution is believed to play a critical role in polymorph formation.[54] However, in conformational polymorphism, how the molecular conformation in solution evolves during nucleation leading to different conformers in final crystal structures remains largely unanswered. The crystallizing molecular species can undergo a degree of conformational change when

transferring from the solution to the solid-state. Any conformational changes raise the question of whether they occur due to packing constraints (e.g., with the incorporation of the molecule into the crystal surface) or whether they are the outcome of rearrangements in the pre-nucleation process.

To date much of the research into the conformations of a molecule has focused on identifying similarities or differences between the crystal lattice and solution conformations,[81,103] without necessarily investigating the properties and magnitudes of the forces that stabilize a particular molecular conformation. One of the most challenging, yet underexplored aspects is the influence of molecular conformation on intermolecular interactions.

Organic crystals are defined as supramolecular entities in which the structural units are repeated regularly and indefinitely in three-dimensions in space. A crystal structure is held by intermolecular interactions whose subtle diversity in strength and directionality is governed by packing motifs and the conformational flexibility of the molecule.[79,104]

These intermolecular interactions define the solid-state structure and determine the polymorphism of an organic system.[105] Still, how intermolecular interactions intervene during the nucleation step, by and large, remains unknown. More importantly, in conformational polymorphism one of the challenges is to improve our understanding of how the balance between conformational energy and intermolecular interaction determines the self-assembly and the precise result of crystallization. Thus, evaluating the extent to which molecular conformation influences intermolecular interactions in conformational polymorphs will provide insights into understanding the molecular mechanism of polymorph formation of organic crystals.

1.6 Objectives

The overarching goal of this dissertation is to gain a molecular-level understanding of the extent to which solution conditions during crystallization influence the conformational polymorphism of drug molecules, using tolfenamic acid as a model solute. Specifically, molecular details of solution conformations of tolfenamic acid and its intermolecular interactions are investigated.

The major hypothesis to be tested is that solution conditions affect the molecular conformation of the model compound necessary for the formation of the most stable associated species during the pre-nucleation self-association process. The following specific aims will be pursued in order to test this hypothesis:

1) Determine if solution conditions in the crystallization medium influence the polymorph formation.

When the molecule in question is conformationally flexible, various crystallization conditions can lead to conformational polymorphs.[3,106] Polymorphism can be controlled by understanding the dependence of crystallization behavior of polymorphs on crystal growth conditions. The extent to which initial solute concentration and temperature govern the formation of polymorphs of tolfenamic acid will be determined.

2) Determine if solution conditions affect the self-association behavior of solute molecules.

Carboxylic acids can form intermolecular aggregates. In the solution and solid phases, associations mainly involve hydrogen-bonded cyclic dimers.[107-110] In polar solvents, hydrogen bonding with solvent molecules occurs along with dimer formation. This

suggests that solution conditions can disrupt or promote particular hydrogen bonding networks. The effect of changing concentration and temperature on the molecular association of tolfenamic acid will be investigated by ultraviolet (UV) and nuclear magnetic resonance (NMR) techniques.

3) Determine if solution conditions affect the conformation of solute molecules in the solution-state.

Solution conditions can lead to structural changes as imposed by interactions of the molecule with its surroundings. Therefore, it is desirable to probe the effect of solution conditions on the molecular conformation. Molecular conformations existing in solution and possible conformational variation will be monitored by NMR techniques as a function of solute concentration and temperature. Conformations of the model compound in the crystalline state will be analyzed by computational means.

4) Evaluate how the molecular conformation affects hydrogen bonding strength.

Crystal structure investigations[79,84] have shown that hydrogen-bonded crystals may feature high-energy conformers to satisfy orientation requirements and to interact more strongly with neighboring molecules. This suggests that the energy of a molecule's conformation can be correlated with the energy of its intermolecular interactions. Intermolecular interaction strength between tolfenamic acid molecules will be calculated in the gas phase as well as in the condensed phase by using advanced computational tools. The electronic properties of the molecular system will be characterized by utilizing density functional theory concepts in order to understand the fundamental cause that drives one conformer to form stronger interactions with respect to another.

In the following chapters, the results from each specific aim will be presented and discussed. The purpose is to provide an in-depth discussion about experimental observations in the solid and solution states, which drove the need for the development of computational methods to further test and support the hypothesis.

Chapter 2 – Polymorph Formation of Tolfenamic Acid

In this chapter, preparation and characterization of tolfenamic acid polymorphs will be presented. This laboratory published a manuscript that highlighted experimental evidence of the effect of growth conditions, exemplified by supersaturation, on the structural outcome of tolfenamic acid crystals.[111] The published manuscript included a thorough characterization of each polymorph using a combination of analytical techniques. In this chapter, temperature will be added as a kinetic factor to further investigate the influence of growth conditions on the preferential formation of each polymorph via crystallization from solution.

2.1 Introduction

Polymorphism of drug compounds plays a key role in formulation and manufacturing design; screening of polymorph structures is therefore one of the major undertakings in the pharmaceutical industry.[1,8] Because of this, precise control of a specific polymorph is critical for achieving desired solubility and stability requirements. Success in preparing various polymorphs frequently hinges on appropriate control of the kinetics of crystallization. The highly dynamic nature of crystallization may lead to variation and lack of reproducibility in the crystallized product. Also, various growth conditions can cause modifications in the preferential formation of polymorphs during crystallization from solution with the consequences that physical and chemical properties of different polymorphs can affect dosage form performance and manufacturing reproducibility.[3]

Such solution conditions include supersaturation ratio, temperature, solvents, and additives/impurities.[112]

Factors that influence the crystallization pathway have been identified. In some cases, the kinetics and mechanisms of nucleation and crystal growth have been rationalized. The role played by solvents or additives in controlling polymorph formation has been approached according to the theory[98] that the selective adsorption of solvent or additive molecules to certain faces of a polymorph will prevent nucleation or slow growth of that polymorph, such that growth of other polymorphs may be favored. Based on this theory, specific solute-solvent interactions have been shown to direct the polymorph outcome in the crystallization of several compounds, including sulfathiazole,[113] sulfamerazine,[33] and 2,6-dihydroxybenzoic acid.[70]

Empirical rules, such as Ostwald's rule of stages,[36] have often been applied to explain the effect of supersaturation and temperature on polymorph formation of organic molecules. For instance, the effect of supersaturation on the crystallization kinetics of phenylbutazone polymorphs has been explained on the basis of the competition between the rates of nucleation and crystal growth of the polymorphs.[100] Specifically, at low supersaturation, the relative nucleation rates of the polymorphs determined the polymorph selectivity while at high supersaturation the nucleation rates of phenylbutazone polymorphs were almost equal, so that the relative crystal growth rates determined the polymorph outcome.

However, this argument fails to address the effect of conformation on the growth kinetics of different polymorphs. Conformational polymorphs are routinely observed in drug development, and are often produced when solution conditions during crystal growth

vary. Of importance to this research, the conformation assumed by the solute molecule may be affected by growth conditions resulting in distinctive crystalline states. With respect to the kinetics of crystallization, a distribution of conformations in solution could potentially influence the critical nucleation and crystal growth steps. Overall, little is known about the mechanism by which growth conditions affect the polymorph formation of conformationally flexible molecules.

In this research, the model solute selected was tolfenamic acid (TFA, Figure 2.1). TFA belongs to the class of nonsteroidal anti-inflammatory drugs (NSAIDs), derived from N-phenylanthranilic acid. Five structurally characterized polymorphs of TFA have been reported in the literature, displaying conformational polymorphism.[114,115] Crystals of Forms I and II were grown from a solution in ethanol, whereas those of Forms III, IV, and V were obtained using polymers as heteronuclei. Crystal structures of Forms I and II belong to the monoclinic space groups $P2_1/c$ and $P2_1/n$ (a non-standard setting of the space group $P2_1/c$), respectively; both structures contain one molecule in the asymmetric unit (R indices of the crystal structures of Forms I and II determined by single-crystal X-ray diffraction are 5.2 percent and 2.9 percent, respectively).[114] Form III exhibits the same space group as Form I, that is, $P2_1/c$, but with $Z' = 2$, whereas Forms IV and V crystallize in the space group $P\bar{1}$ with $Z' = 3$ and $Z' = 1$, respectively.[115]

In this research, the two most frequently encountered structures, Form I (colorless) and Form II (yellow), were produced and studied. The conformation of the molecules in Form I and Form II differs mainly in the torsion angle τ_1 (-74.9° and -142.6° of Forms I and II, respectively, denoted in Figure 2.1). Similar hydrogen-bonded dimer motifs between neighboring carboxyl groups exist in the two polymorphs (Figure 2.2). The

conformational polymorphism and similar molecular packing made TFA an interesting molecule for studying the influence of growth conditions on the crystallization outcome.

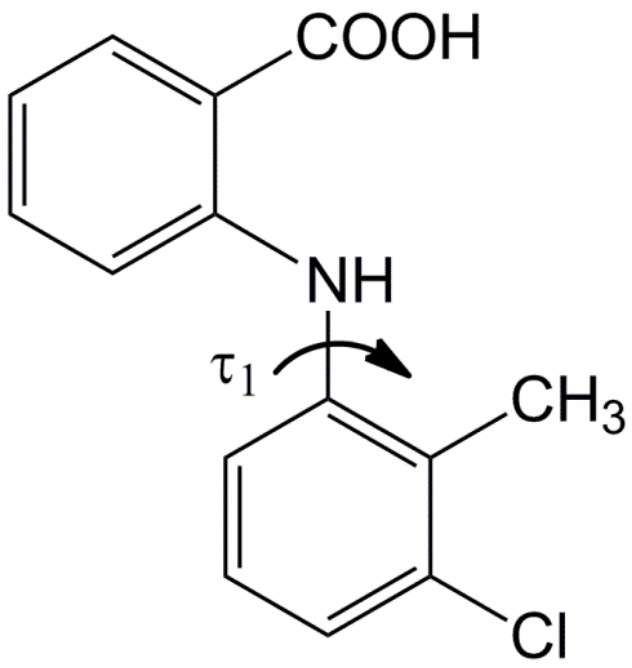

Figure 2.1. Molecular structure of TFA, defining the torsion angle τ_1 differing among the conformers.

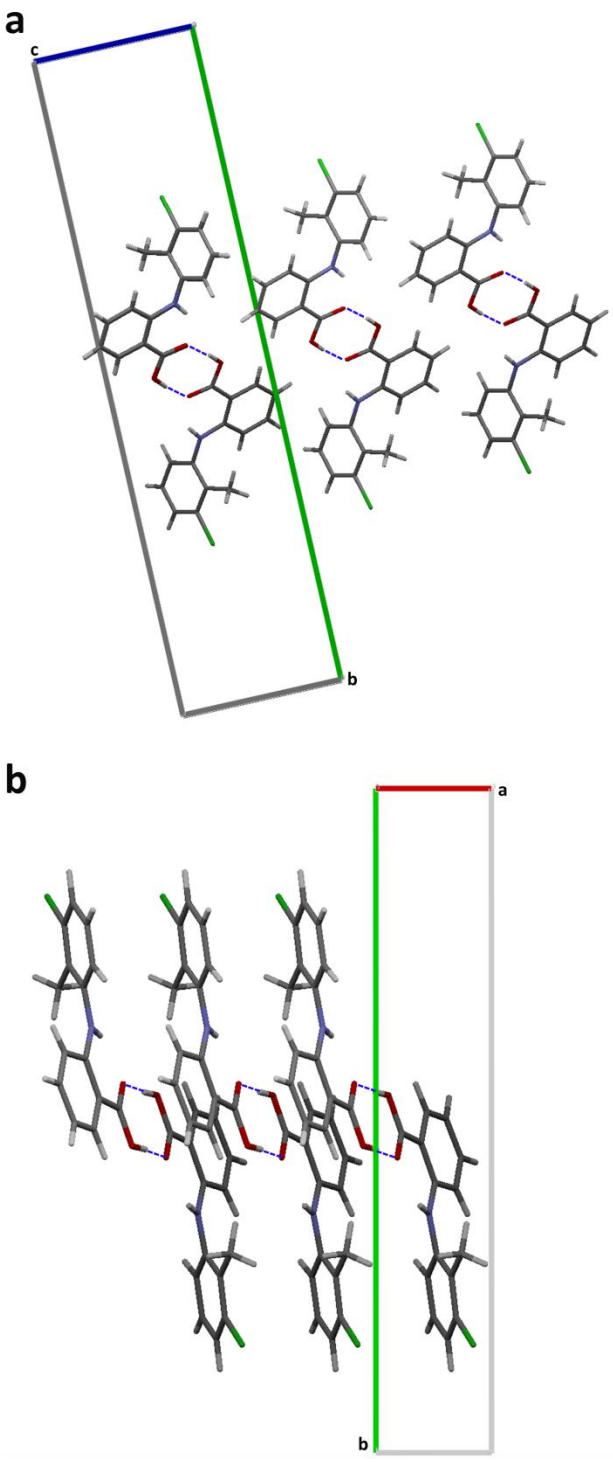

Figure 2.2. Hydrogen bonding dimer motifs of TFA Form I (a) and Form II (b).[114] Hydrogen bonds are denoted as dashed line.

2.2 Materials and Methods

2.2.1 Materials

Tolfenamic acid (TFA, MW = 261.70, > 97 percent purity) was purchased from TCI America (Portland, OR) and was recrystallized from ethanol to obtain Form I. TFA Form II was crystallized by rapid cooling of an ethanol solution to 25° C. The phase purity was verified by powder X-ray diffraction. Ethanol (> 99.9 percent purity) was obtained from Decon Labs, Inc. (King of Prussia, PA) and was used without further treatment. The water content in the organic solvent was evaluated using Coulometric Karl-Fisher titration (Metrohm USA Inc., Riverview, FL) and was determined to be less than 0.1 percent. Polytetrafluoroethylene (PTFE) syringe filters (13 mm and 0.45 μm pore size) were purchased from Millipore, Inc. (Milford, MA) and PTFE syringe filters (0.2 μm pore size) were purchased from Corning Life Sciences (Tewksbury, MA).

2.2.2 Powder and Single-Crystal X-Ray Diffraction

The polymorph purity and identification were assessed by powder X-ray diffraction (PXRD) using a diffractometer (Multiflex, Rigaku Co., The Woodlands, TX) with a Ni-filtered Cu Kα radiation operating at 40 kV and 44 mA. Samples were scanned over a 2θ range of 5 - 40° at the rate of 2 °/min with a step size of 0.02°. Reference powder patterns of TFA crystals were simulated with Mercury 1.4 (The Cambridge Crystallographic Data Centre, Cambridge, UK) based on respective single-crystal structures[114] for comparison with experimental data.

Single-crystal X-ray diffraction was also used to identify TFA solid phases. Data collection was carried out at 90 °K on a Nonius κCCD diffractometer (Madison, Wi) with

Mo Kα radiation (λ = 0.71073 Å). Cell refinement and data reduction were done with the Scalepack and Denzo-SMN programs.[116] Structure solution and refinement were achieved using the Shelxtl package.[117]

2.2.3 Differential Scanning Calorimetry (DSC)

Thermograms of TFA Forms I and II were recorded using a Q20 DSC (TA Instruments, New Castle, DE) equipped with a liquid nitrogen cooling system. The temperature scale and heat flow were calibrated by measuring the onset temperature and the enthalpic response of an indium standard. Samples (3-5 mg) were respectively placed in hermetically sealed aluminum pans and heated from 40 °C to 240 °C at a heating rate of 20 °C/min. In all experiments, the DSC cell was purged with helium at 25 mL/min. Data were analyzed with TA Universal Analysis software (version 4.5A, TA Instruments).

2.2.4 Hot Stage Microscopy (HSM)

Phase transformations of TFA crystals were observed using a polarizing microscope (Olympus BX51 with 10x magnification, Olympus America, Inc., Center Valley, PA) equipped with a hot stage (HSC302, Instec, Inc., Boulder, CO). Single-crystals of pure Forms I and II were sealed in capillary tubes and heated from 25 °C to 240 °C at a heating rate of 20 °C/min. Selected images were taken at various time intervals in order to observe any transformations. The sealed capillary helped to prevent sublimation and simulated the high vapor pressure condition inside a hermetic DSC pan.

2.2.5 Attenuated Total Reflectance-Fourier Transform Infrared (ATR-FTIR) Spectroscopy

Infrared spectra of the crystals were collected with a FTS-7000e FTIR spectrometer (Varian, Inc., Palo Alto, CA) equipped with a deuterated triglycine sulfate (DTGS) detector and a PIKE MIRacle ATR ZnSe accessory. Each spectrum represented 32 co-added scans measured at a spectral resolution of 8 cm^{-1} over the wavenumber range of 4,000-600 cm^{-1}. A background spectrum (without the sample) was collected under the same experimental conditions and was subtracted from each sample spectrum. Spectral data were acquired with Varian Resolutions Pro software.

Mixtures of pure TFA Forms I and II, composed of known ratios of the two polymorphs, were prepared by gently mixing a total weight of 100 mg to obtain a set of standards for the construction of a standard curve for quantitative polymorph analysis. Each calibration sample was analyzed in triplicate.

2.2.6 Solubility Determinations

The solubility of TFA Form I was measured over the temperature range 4 to 55 °C in ethanol. For TFA Form II, only solubility data measured in ethanol at 37 °C were obtained. An amount of the pure solid polymorphs well in excess of their saturation solubilities was allowed to equilibrate with approximately 3 mL of solvent in a 7 mL borosilicate vial. The sealed vial was rotated in an incubator (VWR Scientific, Inc., San Francisco, CA, for measurements at 25 °C, 37 °C, 45 °C, and 55 °C; New Brunswick Scientific, Inc., Enfield, CT, for measurements at 4 °C and 10 °C) set at the desired temperature for at least 72 hours. An aliquot of the saturated solution was then withdrawn

at regular time intervals (1, 2, 4, 6, 8, 12, 24, 36, 48, 60, and 72 hours), filtered through a PTFE filter (0.45 μm pore size), and diluted to an appropriate concentration for spectrophotometric UV analysis at 286 nm (Shimadzu 1800, Shimadzu Scientific Instruments, Inc., Columbia, MD). The maximum concentration of the dissolved metastable form, before declining toward the corresponding plateau value of the stable form, was assumed as its apparent solubility. Equilibration was reached when the concentration of three consecutive measurements differed less than 2 percent. Residual solid phase after equilibration in ethanol solution for 72 hours was identified by PXRD analysis. The solubility measurements of each sample were determined at least in triplicate and errors are reported as standard deviations. Solubility data at 37 °C were used to define the initial supersaturation ratio (i.e., ratio between solute concentration in the solution and equilibrium solubility) with respect to Form I in ethanol in the crystallization experiments at the corresponding temperature.

2.2.7 Grinding and Slurry Experiments

Solvent-drop grinding of pure Forms I and II was carried out using ethanol at 25 °C. In separate experiments, pure Forms I and II were transferred into a mortar and pestle and subjected to grinding while a few droplets of ethanol were added intermittently. Similarly, a 1:1 polymorphic mixture of the two polymorphs was subjected to grinding in the presence of ethanol. The ground samples were analyzed by PXRD.

A slurry of TFA Form II (30 mg) was prepared in a 7 mL borosilicate vial with 1 mL of ethanol. The slurry was then stirred under ambient conditions for 12 hours. The solid phase was examined by PXRD at the end of the equilibration period.

2.2.8 Crystallization Experiments

Crystallization of TFA was performed at 37 °C and 0 °C. Experiments were carried out using a 100 mL three neck flask with an impeller stirring at 100 rpm. Solutions were prepared by dissolving the required amounts of TFA Form I in ethanol at about 60 °C or 25 °C. Following complete dissolution at 60 °C, the solutions at various concentrations (supersaturation) were filtered through a warmed 0.2 μm syringe filter, transferred to the three neck flask, and then cooled rapidly to 37 °C by a temperature-controlled water bath. Solutions dissolved at room temperature were filtered through a 0.2 μm syringe filter, transferred to the three neck flask, and then quench-cooled to 0 °C by use of an ice bath. Liquid-phase samples were removed through a 0.45 μm syringe filter at known time intervals, diluted, and analyzed by UV spectroscopy for the solute concentration measurement. Samples of the solid phase were also taken at predetermined time intervals, dried, and subjected to ATR-FTIR spectroscopy for polymorphic composition analysis. The solid isolated from the suspension was also analyzed with PXRD to identify the form of the solid.

To test the validity of the crystallization method, experiments were conducted by dissolving TFA Form II in ethanol in order to evaluate whether the structure of the starting, dissolved material would affect crystallization kinetics. When supersaturated solutions come into contact with solid surfaces, such as syringe filters, nucleation and/or adsorption may occur. Formation of such induced nuclei may accelerate and/or affect the crystallization behavior of the drug. Thus, to ensure that no (detectable) nucleation occurred during the filtration step and that all particles were dissolved, solutions were maintained at the heating temperature for one hour. The stopper of the round-bottomed

flask was wrapped with Parafilm® to prevent evaporation of the solvent during heating. Subsequently, solutions were cooled without filtration. The rate of stirring of the suspensions was strictly controlled in all the experiments.

Seeding experiments were performed by adding 1 percent (w/w) (based on the total mass of TFA) of dried, ground seed crystals of either Form I or II to ethanol solutions at various supersaturation ratios. After the seeds were suspended in the solutions, the stirrer was turned on and the stirring rate was adjusted to the designated rpm. The solutions were then rapidly cooled to 37 °C or 0 °C, and changes in solution concentration as well as in solid composition were monitored over time, as previously described for the unseeded experiments.

2.3 Results and Discussion

2.3.1 Characterization of Tolfenamic Acid Polymorphs

Both TFA Forms I and II prepared in our laboratory resulted in similar needle-like morphology (Figure 2.3). Shown in Appendix 1 is the growth morphology prediction of TFA Forms I and II. Solubility of TFA Form I was determined in ethanol as a function of temperature (Figure 2.4). TFA solubility increased with temperature. This common trend reflects the expected increase in solubility of organic molecules with temperature as predicted by thermodynamics.[118] No change in the polymorph identity, verified by PXRD, was observed in the solid phases recovered from the solubility study. Solubility profiles of TFA Forms I and II were measured in ethanol at 37 °C and plotted against time (Figure 2.5). The solution concentration when Form II was used as the starting material increased until it reached a maximum around the solubility of the metastable

polymorph, which should be regarded as an "apparent" value because thermodynamic equilibrium could not be reached when measuring the metastable Form II. The concentration declined toward the corresponding plateau value of the stable Form I and remained at that concentration thereafter for 72 hours, showing a typical solvent-mediated phase transformation.[35] The residual solid samples were verified to be Form I regardless of the starting materials. The thermodynamic solubility of Form I in ethanol was then estimated to be 86.1 ± 0.9 mM at 37 °C, while the apparent value of Form II was 95.6 ± 0.8 mM.

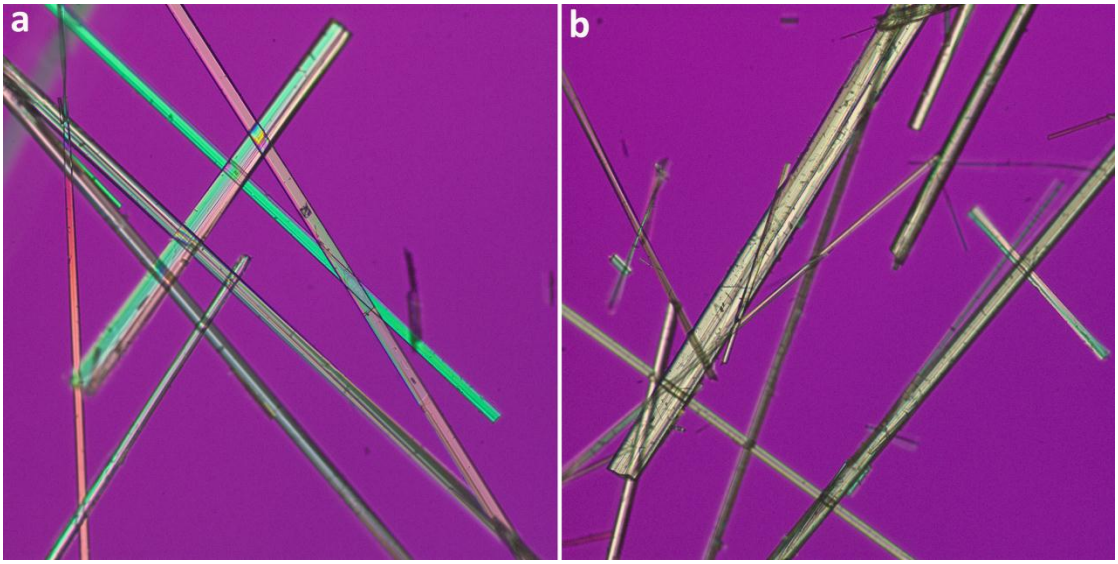

Figure 2.3. Optical micrographs of TFA Form I (a) and Form II (b).

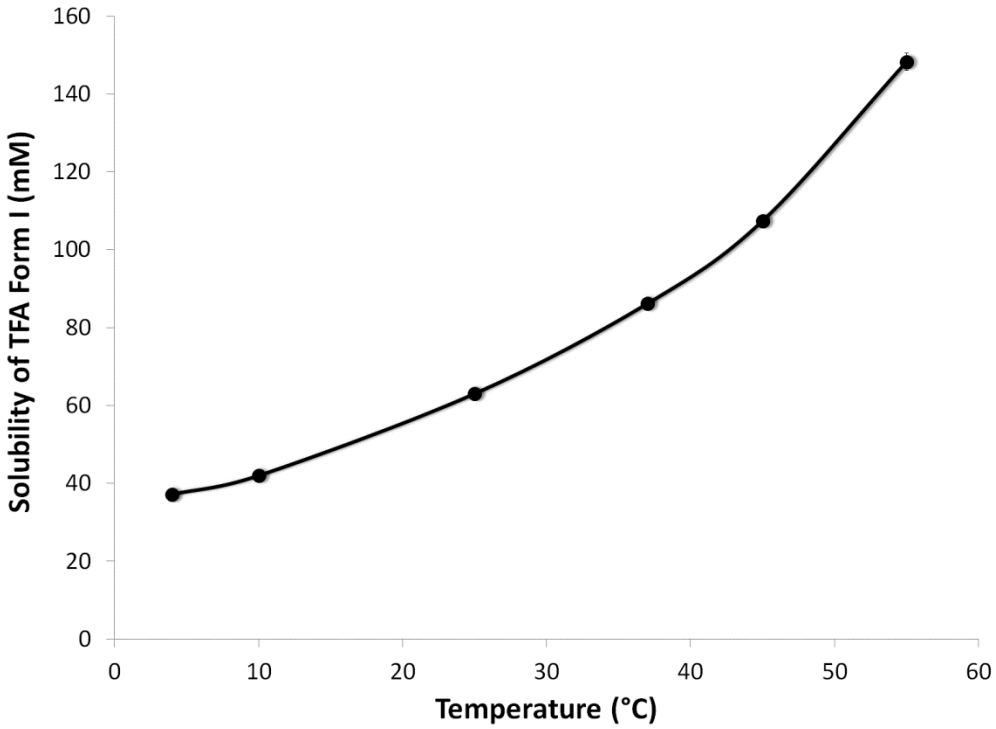

Figure 2.4. Variation of the measured solubility with temperature for TFA Form I.

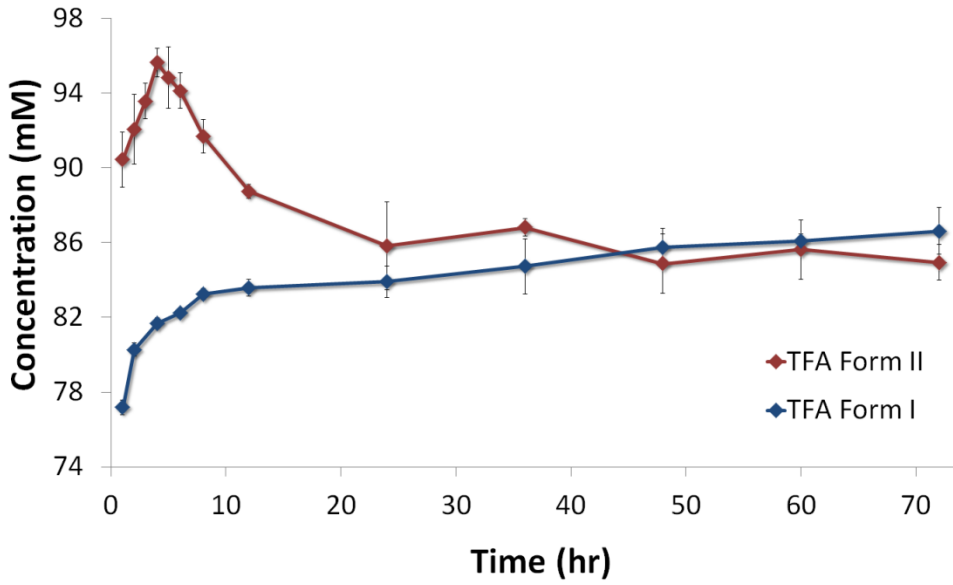

Figure 2.5. Solubility profiles of TFA Forms I and II in ethanol at 37 °C. Error bars represent standard deviation (n=3).

The relative stability of Forms I and II was inferred by the solvent-mediated phase transformation, as well as experiments of solvent-drop grinding, slurry crystallization, and thermal analysis. When Form II was subjected to grinding at room temperature in the presence of ethanol in a mortar and pestle, conversion to Form I occurred after one hour, as monitored by PXRD (Figure 2.6). No phase conversion was observed for the pure Form I. When a mixture of the two polymorphs was subjected to ethanol-assisted grinding, phase transformation occurred readily to give Form I within 20 minutes. In addition, a slurry of pure Form II converted to Form I in eight hours at 25 °C.

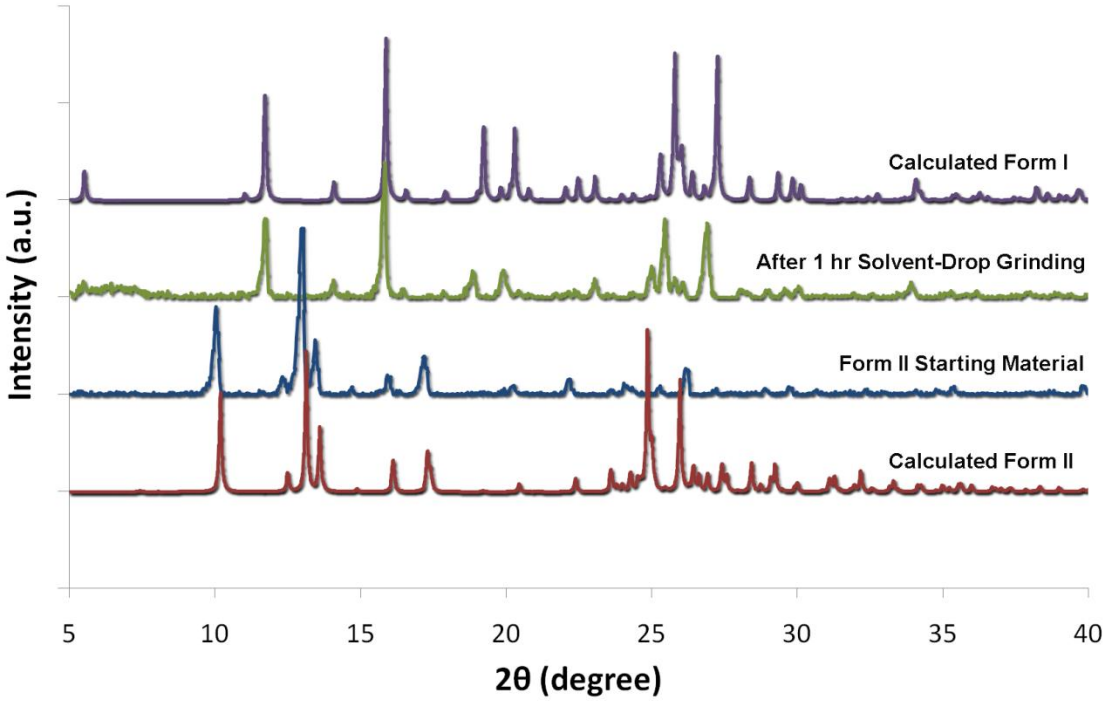

Figure 2.6. Powder X-ray diffraction patterns of TFA Form II subjected to ethanol-assisted grinding. Calculated patterns of Forms I and II are also shown for comparison.

A DSC thermogram (Figure 2.7) of Form I showed melting of the crystal at the onset temperature of 212.5 °C; the enthalpy of melting was 157.1 J/g. Form II showed two endothermic peaks. The first one with an enthalpy of 4.3 J/g and the onset temperature of 141.8 °C appeared to be a solid-solid phase transition to Form I (inset of Figure 2.7) and the second one at the onset temperature of 212.7 °C was the melting of the transformed Form I.

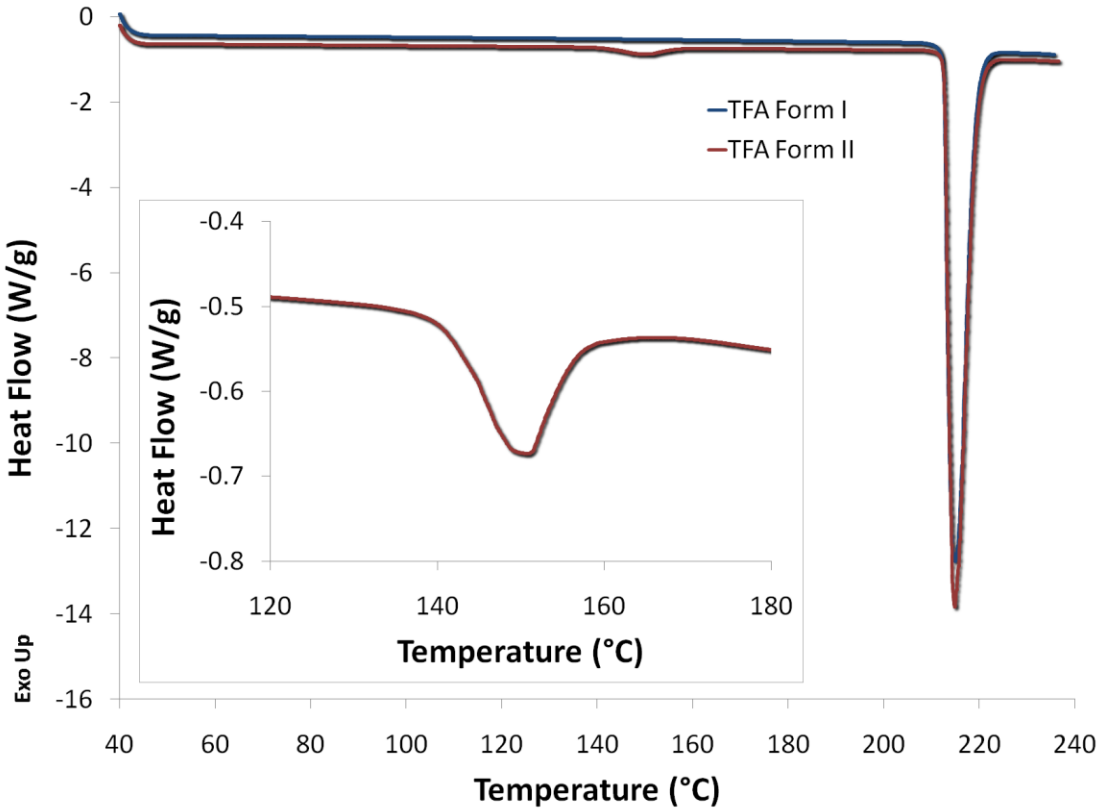

Figure 2.7. DSC thermograms of pure TFA Forms I and II. The inset shows the small endotherm due to phase transition of Form II.

Thermal behavior of Forms I and II was further studied by hot stage microscopy. Their phase behaviors echo the DSC experiments. Form I underwent melting at 213.3 °C with no other phase transitions observed. However, the Form I crystal became opaque and looked damaged. When the Form I crystal was heated, new crystals formed at 204.8 °C due to sublimation (Figure 2.8a). In the encapsulated capillary tube, the water vapor could not escape readily and, thus, built up in the enclosed environment. The resulting high water vapor pressure induced the formation of crystals that were verified to be Form I by single-crystal X-ray diffraction. The crystals were of the monoclinic space group $P2_1/c$ with a = 4.800(1), b = 31.991(1), c = 7.972(3) Å, and β = 105.2(3)°. These values were in agreement with the literature.[114] No visual changes in crystal habit were observed during heating single-crystal of Form II, but the crystal turned colorless and opaque and remained solid throughout, corresponding to a phase change. As such, Form II underwent a solid-solid phase transition at 145.0 °C and converted to Form I, which then started to melt at 211.8 °C (Figure 2.8b). All these experiments indicate that Form I of TFA is the thermodynamically stable polymorph over a temperature range of 25 – 213 °C.

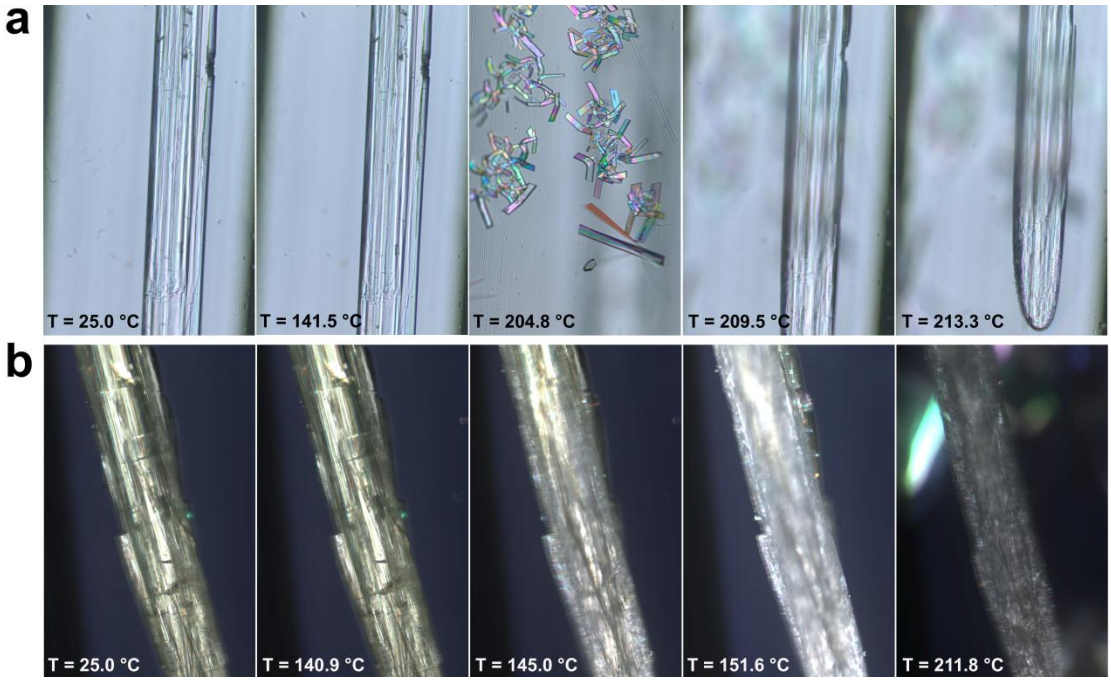

Figure 2.8. Hot stage microscopy snapshots showing melting of Form I at 213.3 °C, formation of new Form I crystals due to sublimation at 204.8 °C (a), and the solid-solid phase transition from Form II to Form I at 145.0 °C (b).

2.3.2 Concomitant Polymorph Formation of Tolfenamic Acid

Crystallization of TFA was conducted in ethanol by cooling from 60 °C to 37 °C. TFA concentration profiles were obtained at supersaturation ratios of S = 1.55, 1.70, and 1.95, respectively (Figure 2.9). The solute concentration, detected by UV absorption, decreased with time, which indicates nucleation and growth of crystals. The process was monitored until the concentration reached progressively stable values. The final solute concentration values were in agreement with the solubility of Form I measured at 37 °C. The figure also marks the apparent solubility of Form II.

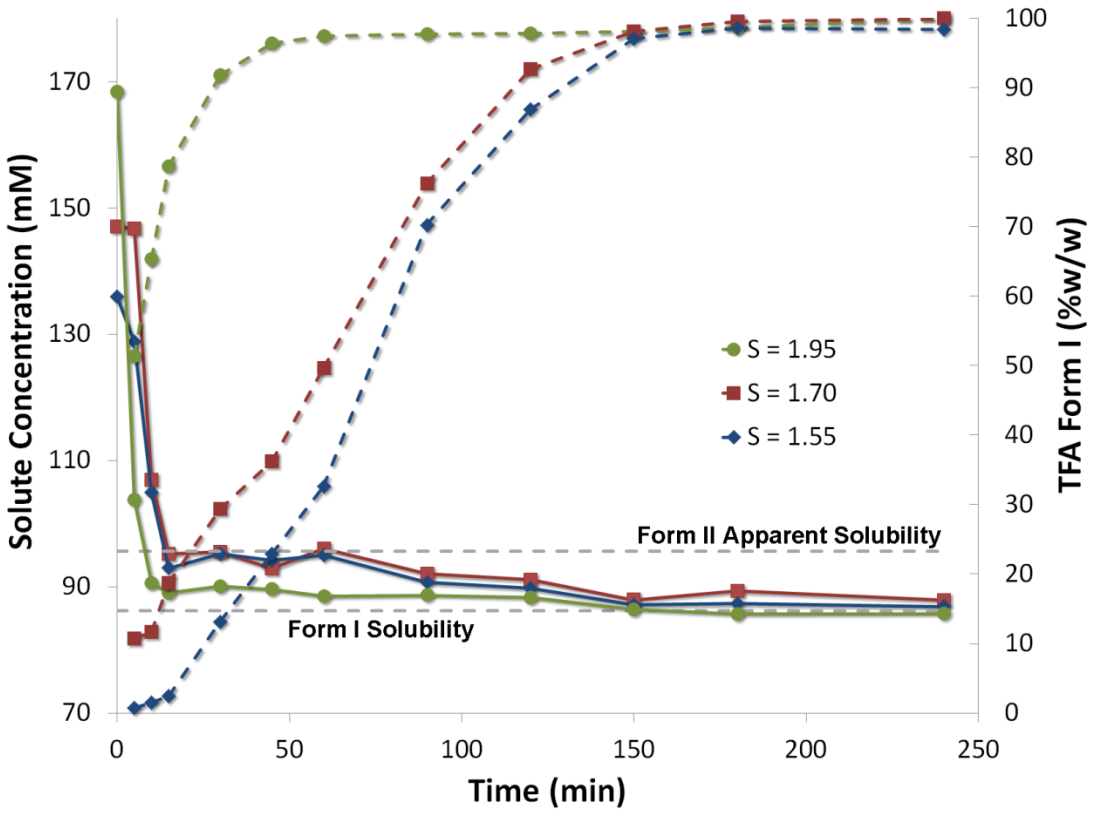

Figure 2.9. Time dependence of concentration measured at 37 °C during rapid cooling of TFA from ethanol at various initial supersaturation ratios (S), 1.55, 1.70, and 1.95. The weight percentages of Form I in the collected crystal samples are plotted correspondingly as dashed lines. Solubility values of the two polymorphs are marked as well.

Crystallized solid samples were collected and analyzed to determine the polymorph composition. PXRD analysis verified the solid-phase identity of these samples (Figure 2.10), obtained at a supersaturation ratio of S = 1.55 and isolated immediately after nucleation. The peak at 10.2° indicates that Form II was present in the sample. In later collected samples, the peaks at 5.5° and 11.7° appeared, suggesting that any crystal of Form II had transformed to Form I. ATR-FTIR spectra of pure Forms I and II, as well as of known mixtures, were obtained (Figure 2.11a) for quantifying the ratio between the two polymorphs. The Form II spectrum exhibited a unique vibrational band at 1,521 cm^{-1}

that clearly distinguished it from the Form I spectrum. The peak at 1,327 cm^{-1} was chosen as a reference due to its insensitivity to the phase composition. The standard curve (Figure 2.11b, R^2 of 0.996) was generated by plotting the ratio of the band intensities against the weight percentage of Form II in the known solid mixture samples. Accordingly, the solid composition of collected solid samples at various time points was determined by FTIR analysis (Figure 2.9, superimposed with the solute concentration profiles). At a lower supersaturation ratio (S = 1.55 or 1.70), the metastable Form II was initially obtained, followed by its transformation to the most stable Form I, likely via a dissolution-recrystallization mechanism (i.e., solvent-mediated phase transformation).[35] When S = 1.55, the solid-phase composition showed no onset presence of Form I and the solute concentration kept at the solubility line of Form II between 30 and 60 minutes. Over the course of the transformation, a decrease in the amount of Form II was associated with a concurrent increase in the quantity of Form I while the concentration shifted to the solubility of Form I. As the supersaturation increased, a mixture of both polymorphs was produced and, particularly when S = 1.95, Form I became dominant and the solute concentration dropped to its solubility line quickly at the beginning of the experiment. The results indicate a greater tendency to crystallize Form I by increasing the initial solute concentration. When S = 1.55 or 1.70, the solvent-mediated phase transformation from Form II to Form I may take dozens of minutes. Instead, when S = 1.95 the appearance of Form I took a few minutes.

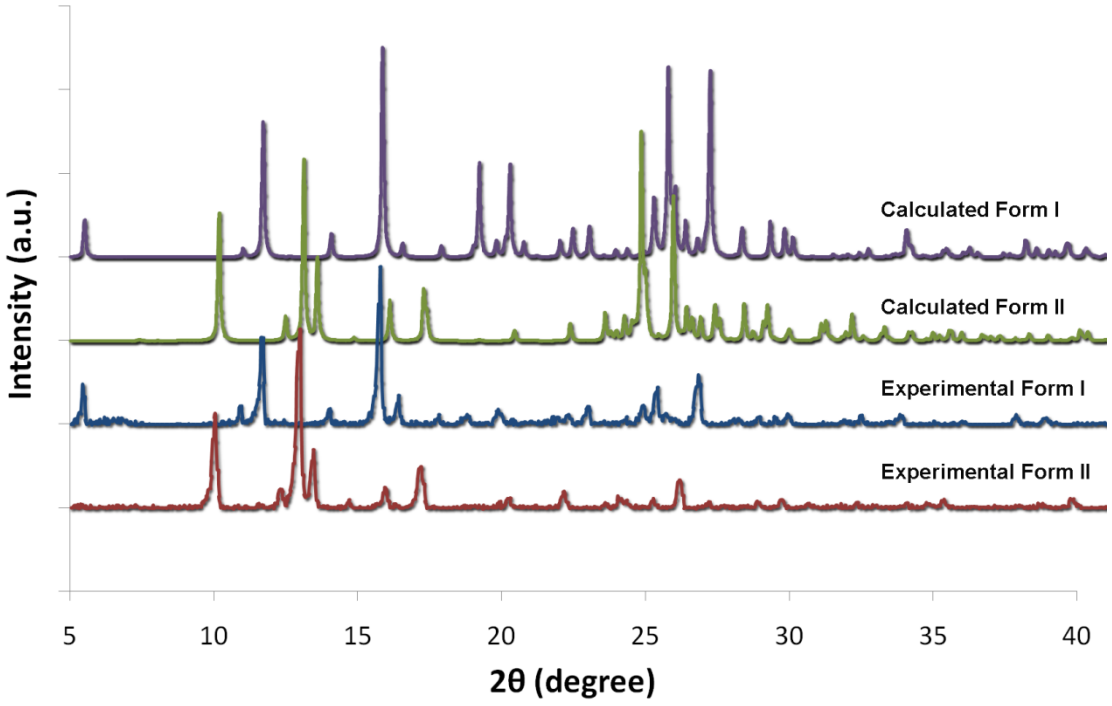

Figure 2.10. Powder X-ray diffraction patterns of TFA Forms I and II obtained at S = 1.55. Calculated patterns of Forms I and II are also shown for comparison.

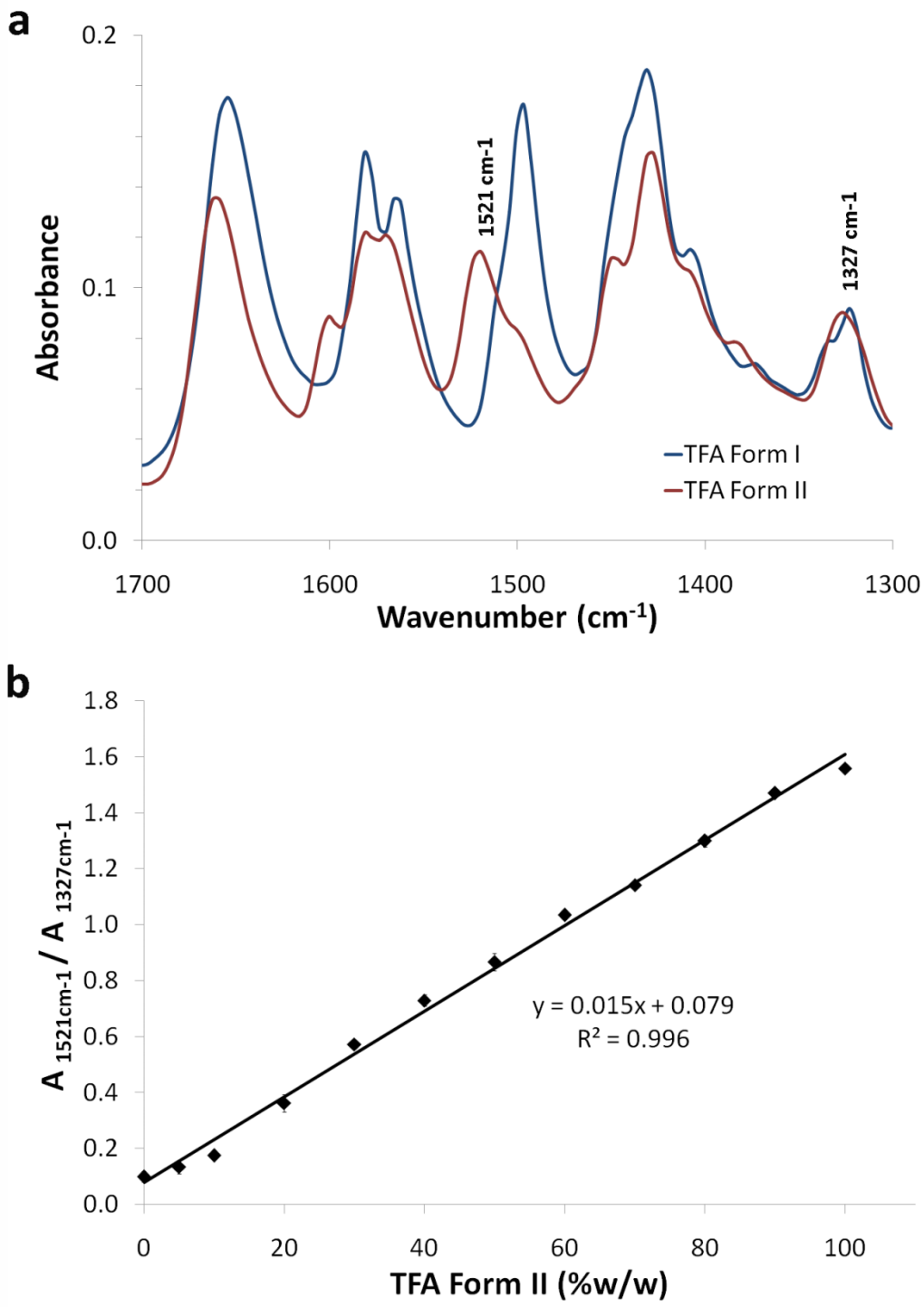

Figure 2.11. Infrared spectra of TFA polymorphs (a) and standard curve of the relative composition of Form II in a mixture of TFA polymorphs (b). Measurements were done in triplicate to determine the standard curve.

Two sets of experiments that were carried out to evaluate the validity of the crystallization procedure included the use of Form II as the starting material and the cooling to 37 °C without filtering the dissolved solutions. When the supersaturation ratio was low, the preferential formation of Form II via crystallization from solution was achieved in both sets of experiments. Using Form II as the initial solid phase, at S = 1.55, the solute concentration initially decreased and then remained constant up to 60 minutes at a concentration corresponding to the apparent solubility of Form II (Figure 2.12a). Similar results were observed when samples were recrystallized from solutions without filtration (Figure 2.12b). When the supersaturation ratio increased, the solute concentration initially dropped rapidly and approached a concentration below the apparent solubility of Form II in both sets of experiments. The solid-phase composition of the sample collected at the beginning of the experiments was about 30 percent (w/w) of Form I (when S = 1.93 for the experiment performed by dissolving Form II and when S = 1.85 for the experiment performed without filtration) (data superimposed in Figures 2.12a and 2.12b). Thus, both Form I and Form II crystallized by increasing the initial solute concentration. The results of these control experiments suggest that neither the initial polymorph of the drug nor the filtration play a noticeable role in the crystallization process. Up to this point, the TFA crystallization behavior can be ascribed to the effect of drug concentration.

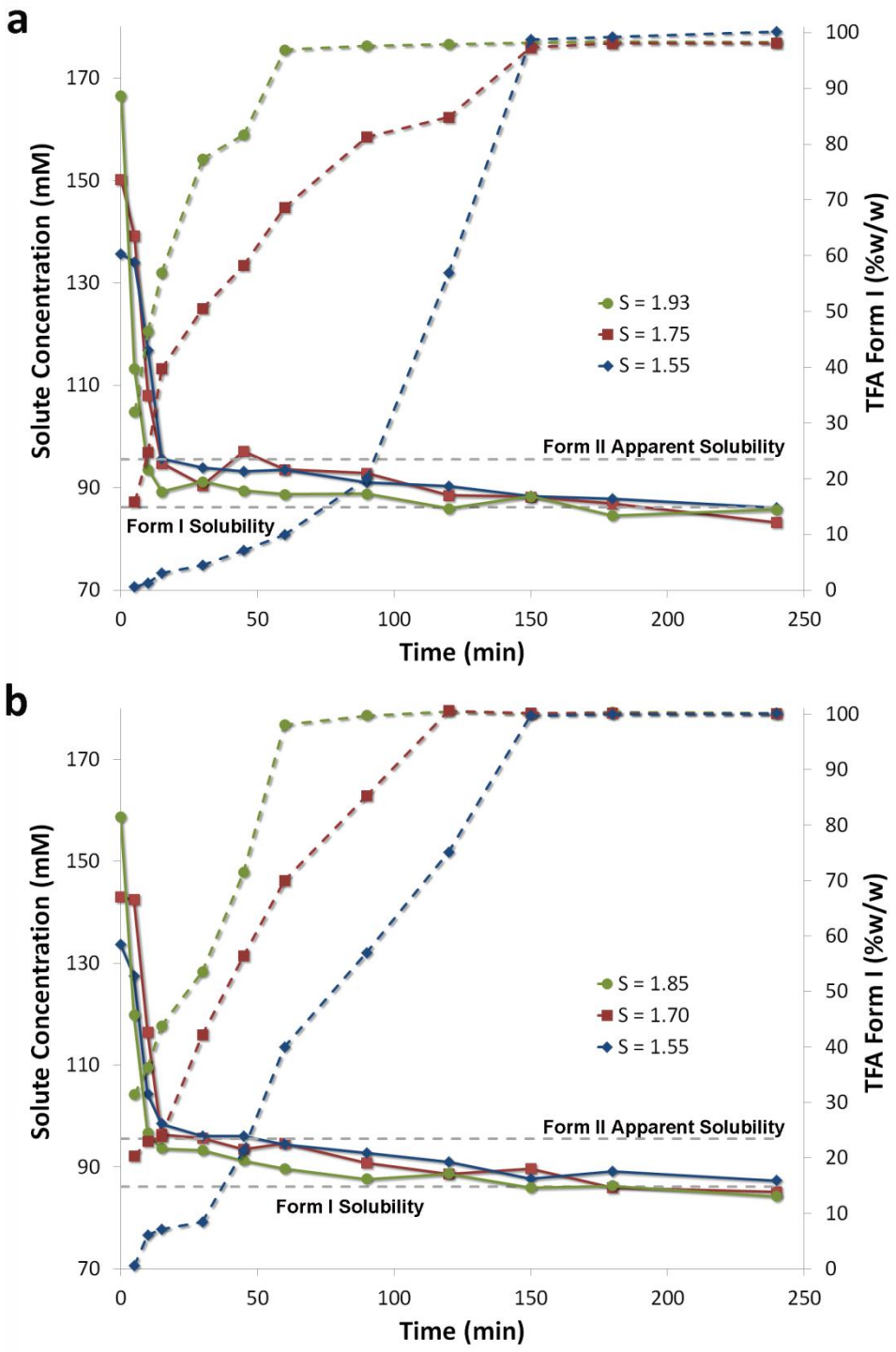

Figure 2.12. Time dependence of concentration measured at 37 °C during crystallization of TFA from ethanol solutions conducted by dissolving Form II at various initial supersaturation ratios (S), 1.55, 1.75, and 1.93 (a) and without filtration at various initial supersaturation ratios (S), 1.55, 1.70, and 1.85 (b). The weight percentages of Form I in the collected crystal samples are plotted correspondingly as dashed lines. Solubility values of the two polymorphs are marked as well.

Although the final outcome of the crystallization experiments was the same (production of Form I), at least two distinct mechanisms of crystal growth are plausible depending on the supersaturation. First, at high supersaturation the appearance of Form I crystals may not be a result of phase transformation from Form II, but rather a consequence of direct nucleation. Form I crystals subsequently may increase the conversion of Form II crystals. Second, the metastable polymorph initially crystallizes at all concentrations, but the conversion to the stable Form I may be faster at high supersaturation than at low supersaturation, meaning that the growth rate of the stable polymorph is concentration-dependent and increases as initial drug concentration increases.

Therefore, to elucidate the governing mechanism responsible for the observed polymorph formation, seed crystals of both polymorphs were respectively introduced into solutions at the onset of cooling (Figure 2.13). Addition of Form I seeds resulted in an abrupt decrease of solution concentration, consistent with crystal growth of the seeds, until the solute concentration eventually reached the solubility of TFA Form I (Figure 2.13a). This was in agreement with the solid-phase composition determined by ATR-FTIR analysis (superimposed in Figure 2.13a), where for all supersaturations the solid-phase was entirely constituted by Form I. Following seeding with Form II crystals (Figure 2.13b), the metastable polymorph first predominated and then slowly converted to Form I. During the polymorph transformation, the solute concentration stayed at the apparent solubility of Form II until the metastable crystals had completely dissolved and the concentration approached the solubility of Form I. Consequently, the relative percentage of the solid Form I increased during the transformation. No significant difference was found between the phase turnover when S = 1.55 or 1.70. When S = 1.95, a small amount

of Form I, about 9 percent relative to Form II, also crystallized upon seeding with Form II, possibly due to nucleation of Form I despite the seeding. Overall, seeding crystals are able to essentially circumvent the primary nucleation step, particularly at low supersaturation. The experiments also rule out the possibility of competitive growth of the two polymorphs that may be affected by the solute concentration. If the two forms possess different concentration-dependent growth rates, seeding or no seeding should not affect the crystallization outcome. This, however, is not what was observed, and the results further confirm that in the unseeded crystallization experiments, the nucleation mechanism is influenced by the solute concentration and that nucleation of the most stable form at high supersaturation apparently disobeys Ostwald's Rule of Stages.[36]

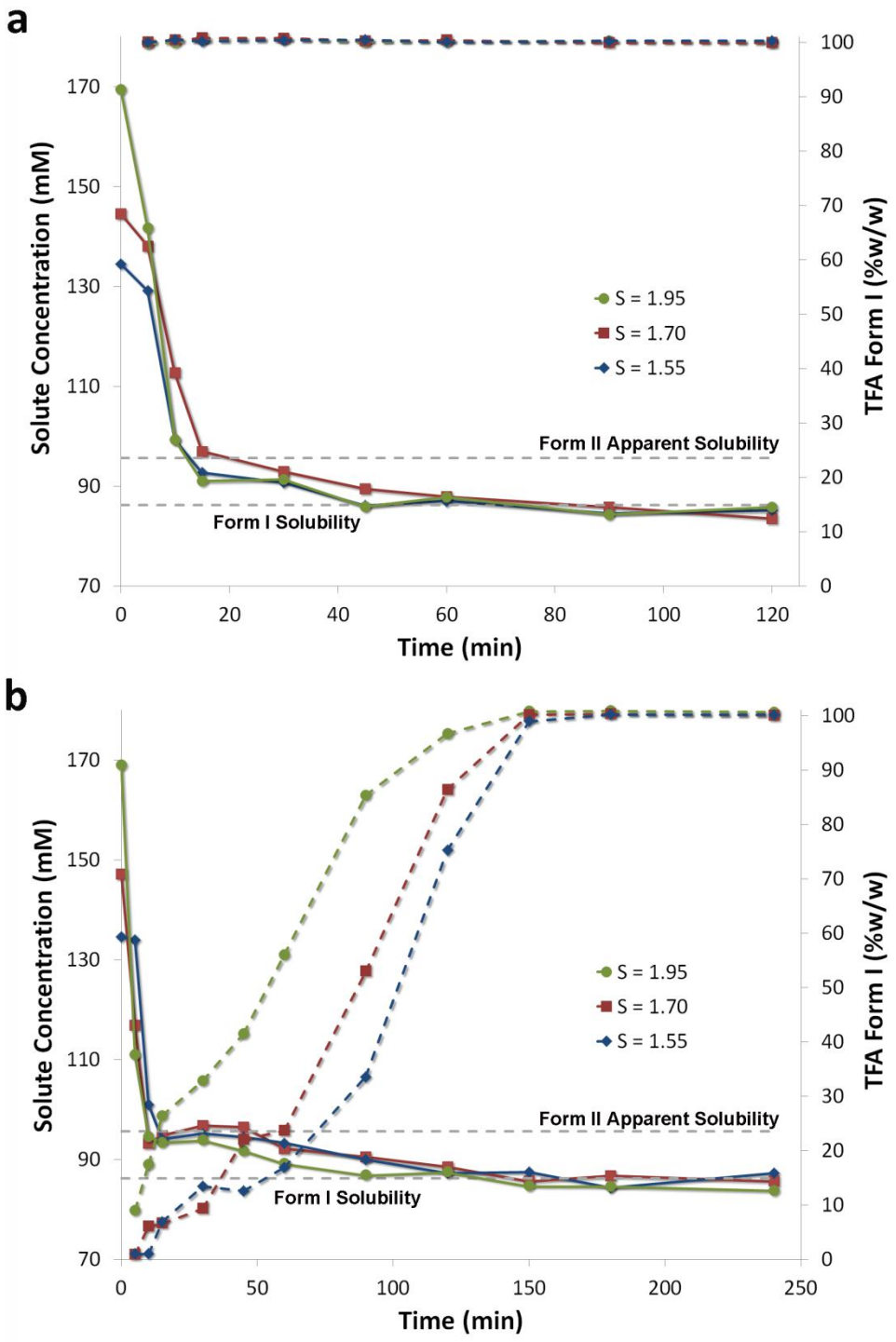

Figure 2.13. Time dependence of concentration measured at 37 °C during rapid cooling of TFA from seeded ethanol solutions at various initial supersaturation ratios (S), 1.55, 1.70, and 1.95 with seeds of Form I (a) and seeds of Form II (b). The weight percentages of Form I in the collected crystal samples are plotted correspondingly as dashed lines. Solubility values of the two polymorphs are marked as well.

Crystallization experiments were also performed by quench-cooling using an ice bath. By rapidly cooling ethanol solutions from 25 °C to about 0 °C, a mixture of the metastable Form II and the stable Form I was obtained irrespective of the initial TFA concentration (Figure 2.14). The solute concentration dropped from the initial value. The decrease in the measured solute concentration was due to an increase in the number of crystals in suspension. However, the decrease in solute concentration was slower for the sample at low initial concentration. The time period in which crystals were first observed was taken as the induction time. In the system at low initial concentration, the induction period was prolonged and no crystal formation occurred for approximately the first 30 minutes of the process. This induction period was reduced as the initial value of solute concentration increased. Information about the polymorphic form of the crystallizing solids was extracted from ATR-FTIR spectra, as previously described. At the initial phase of crystallization, the solid-phase composition revealed the presence of about 13 percent (w/w) and 18 percent (w/w) of Form I at lower and higher solute concentrations, respectively (Figure 2.14, superimposed with the solute concentration profiles). The percentage of Form I slightly increased when the solute concentration increased. The initial dip in the solute concentration profile of both systems was indeed consistent with nucleation of Form I. These nuclei of Form I set the starting point for the transformation process. As Form II crystals grew, the solution concentration decreased below the solubility of Form I and, thus, Form II crystals began to dissolve themselves. The dissolution of Form II created a localized supersaturation that allowed the other crystal form, Form I, to grow at the same time Form II crystals were dissolving. The percentage of Form I in the crystallized product increased as solvent-mediated phase transformation

progressed. Ultimately, the solid composition reached its final state, together with the solute concentration profile, which stabilized at the solubility of Form I.

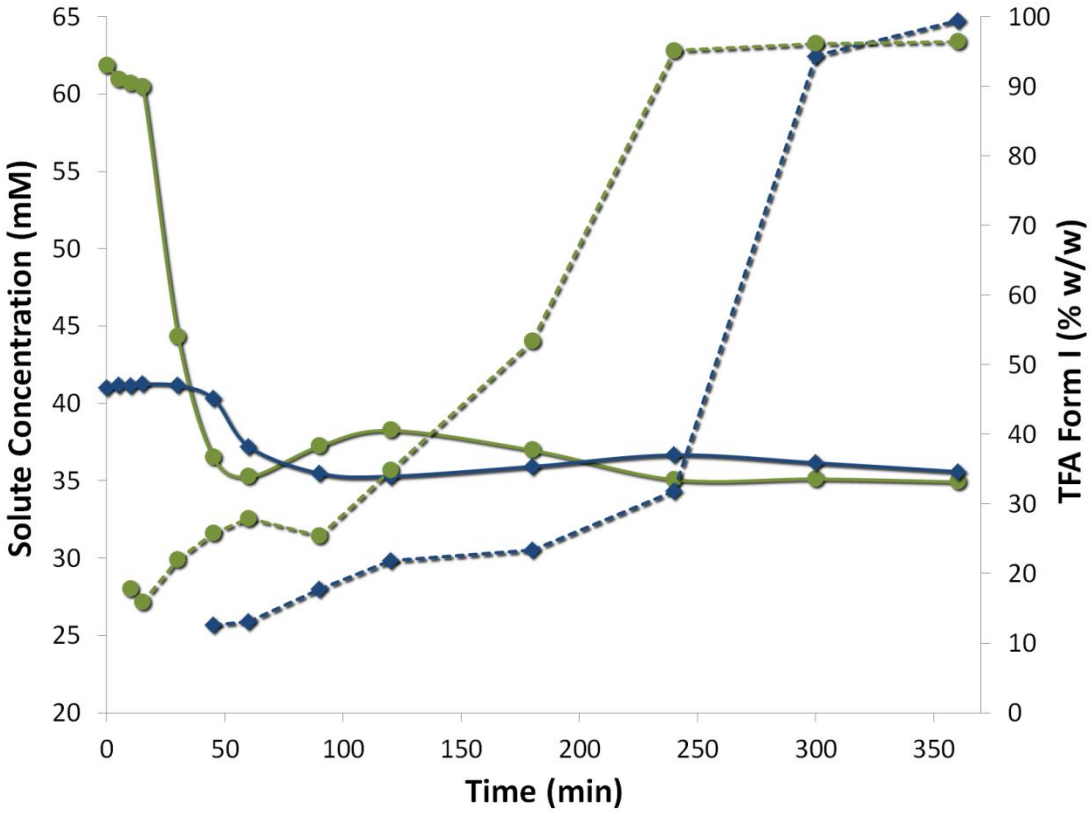

Figure 2.14. Time dependence of concentration measured at 0 °C during rapid cooling of TFA from ethanol at different initial concentrations. The weight percentages of Form I in the collected crystal samples are plotted correspondingly as dashed lines.

When comparing unseeded crystallization experiments performed at 37 °C (Figure 2.9) to those carried out at 0 °C (Figure 2.14), the former systems exhibit much shorter induction periods than the latter systems. The induction period refers to the time that elapses after the creation of supersaturation in solution until a new phase is detected.[119] Due to the bulk nature (i.e., the nuclei cannot be observed directly because of their small dimensions

and, thus, one must rely on the macroscopic clusters) of the off-line analytical measurements described herein, the crystallization measurements at higher temperature are more likely to stem from the already-grown crystals, particularly regarding those when S = 1.95. In this case, direct nucleation and subsequent growth of Form I may occur soon after the onset of cooling, thus giving rise to the dramatic increase in the amount of Form I in the product phase. However, these observations do not invalidate the nucleation mechanism being the determining step of the polymorph formation of TFA. In fact, the crystallization experiments yield an increased amount of the stable Form I when the initial solute concentration increases, whether crystallization occurs at 37 ℃ or 0 °C.

Note that although the isothermal conditions at 0 °C were not tightly maintained by a temperature controlled system during either of the two experiments, the same thermal gradient was applied in the experiments performed at different temperatures (i.e., 37 °C and 0 °C), respectively. For that reason, thermal conditions should not cause changes in the underlying mechanism of the polymorph formation. In order to verify the above statement, seeding experiments were performed using seed crystals of both polymorphs.

After seeding with TFA Form I crystals the solute concentration decreased, until it eventually approached a steady-state concentration (Figure 2.15a), which possibly corresponds to the solubility of Form I at the defined temperature. Accordingly, the solid-phase composition profiles (superimposed in Figure 2.15a) showed that Form I precipitated regardless of the initial supersaturation and persisted over the duration of the crystallization process. This was consistent with crystal growth of the seeds. Thus, Form I seeds were able to suppress the formation of the metastable phase and to produce the Form I crystals exclusively. During the first minutes following the introduction of Form

II seeds (Figure 2.15b), no spontaneous nucleation of the stable Form I was observed in metastable suspensions, which could be attributed to the growth of the metastable crystals already in suspension. Initially, Form II dominated before undergoing solvent-mediated phase transformation to Form I. The time at which the appearance of Form I crystals occurred corresponded to the increase of Form II to Form I turnover that was not significantly different for either of the experiments at different supersaturations. During the polymorph transformation, the solution concentration tended towards a plateau value, indicating that saturation with respect to Form II was reached, until the metastable crystals disappeared to be replaced by the stable crystals and in turn, the solute concentration reached its final level. Form I crystals eventually were isolated.

In the seeded crystallization experiments performed at 0 °C, the observed induction times were substantially shorter at the same initial concentration values when compared to the unseeded experiments. The induction period is a composite of the time required for the formation of nuclei and their growth to a detectable size. In the presence of seed crystals, only growth is expected to occur and this results in shorter induction times being observed. The crystal growth of the already present polymorph, added to the system as seed crystals, is favored irrespective of the initial solute concentration. This implies that the polymorph obtained by unseeded crystallizations is controlled by nucleation.

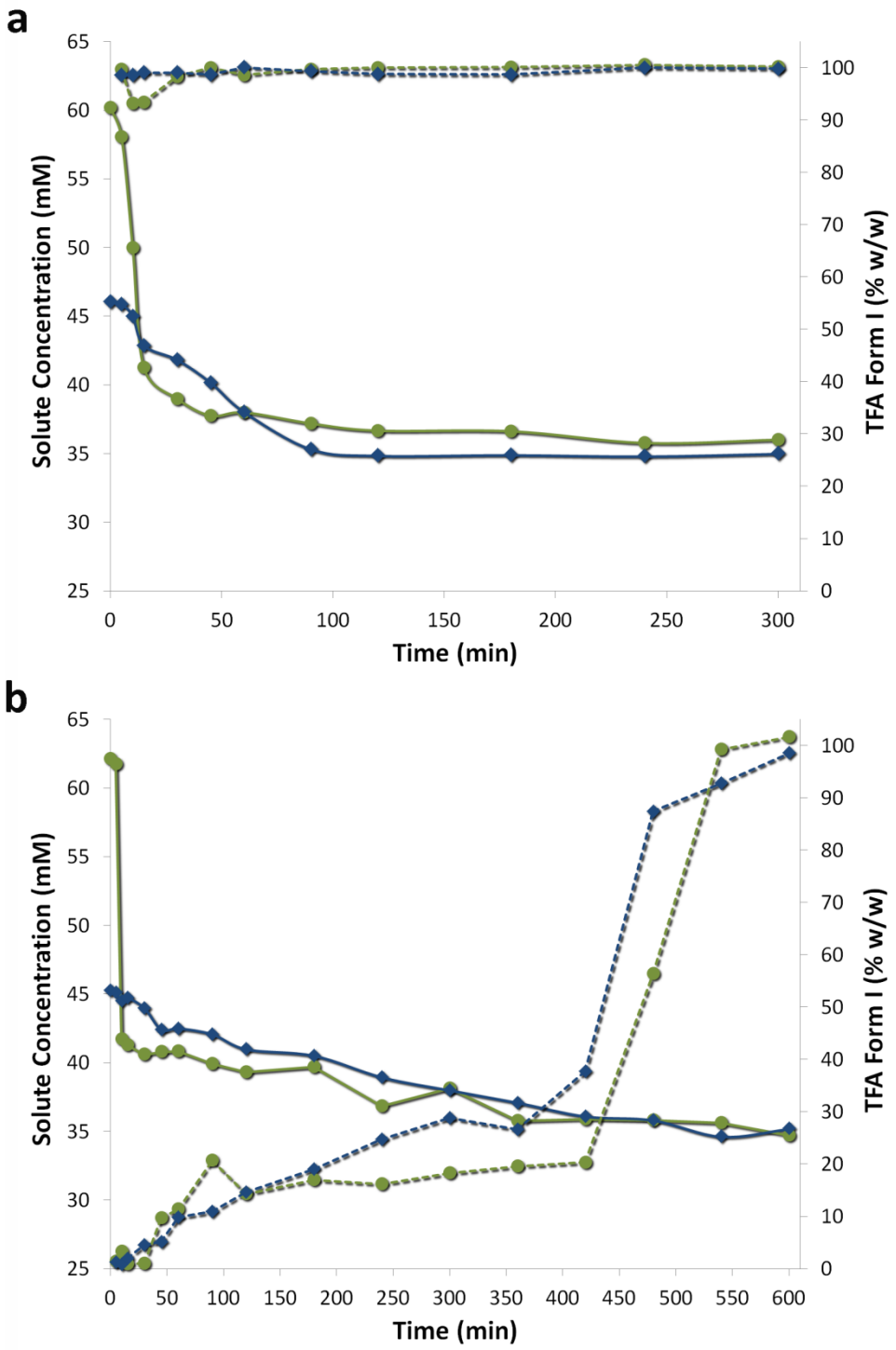

Figure 2.15. Time dependence of concentration measured at 0 °C during rapid cooling of TFA from seeded ethanol solutions at different initial concentrations with seeds of Form I (a) and seeds of Form II (b). The weight percentages of Form I in the collected crystal samples are plotted correspondingly as dashed lines.

2.4 Unanswered Question

In this section, the crystallization results of the model solute, TFA, will be discussed. To assess the effects of solute concentration and temperature on polymorph outcomes, crystallization experiments were conducted at different starting concentrations and temperatures, respectively. An increasing solute concentration resulted in a reduced Form II mass fraction and, thus, biased the polymorph outcome toward the most stable Form I. The biggest question to address is why the most stable polymorph of TFA was obtained when crystallized from highly supersaturated ethanol solution by cooling to either 37 °C or 0 °C temperature. Seeding experiments revealed that the polymorph formation was controlled by nucleation kinetics and not by crystal growth kinetics.

According to CNT, the effect of supersaturation or temperature on the polymorph selectivity is determined by the relative nucleation rates of the polymorphs.[27,43] The TFA polymorph behavior was qualitatively explained by the nucleation theory. The starting point for a discussion of the nucleation kinetics is the free energy diagram for crystallization of a dimorphic system (Figure 2.16).[27] For the considered system, Form I is the most stable polymorph and Form II is the metastable polymorph. There will be a critical radius (r_c) and an activation barrier for nucleation (ΔG_{crit}) corresponding to each crystalline form. The polymorph that bears the smallest surface energy (likely the metastable Form II) crystallizes first. Although the stable Form I may have the greater thermodynamic drive to crystallize, Form II may nucleate first due to its higher nucleation rate. A similar concept should be applied to the TFA system. Accordingly, the driving force for precipitation increases as the solute concentration increases. TFA Form II should crystallize, particularly at high supersaturation, if it bears the lower surface

energy and, thus, the lower activation energy for nucleation. However, direct nucleation of the stable TFA Form I occurred from the crystallization at higher supersaturation.

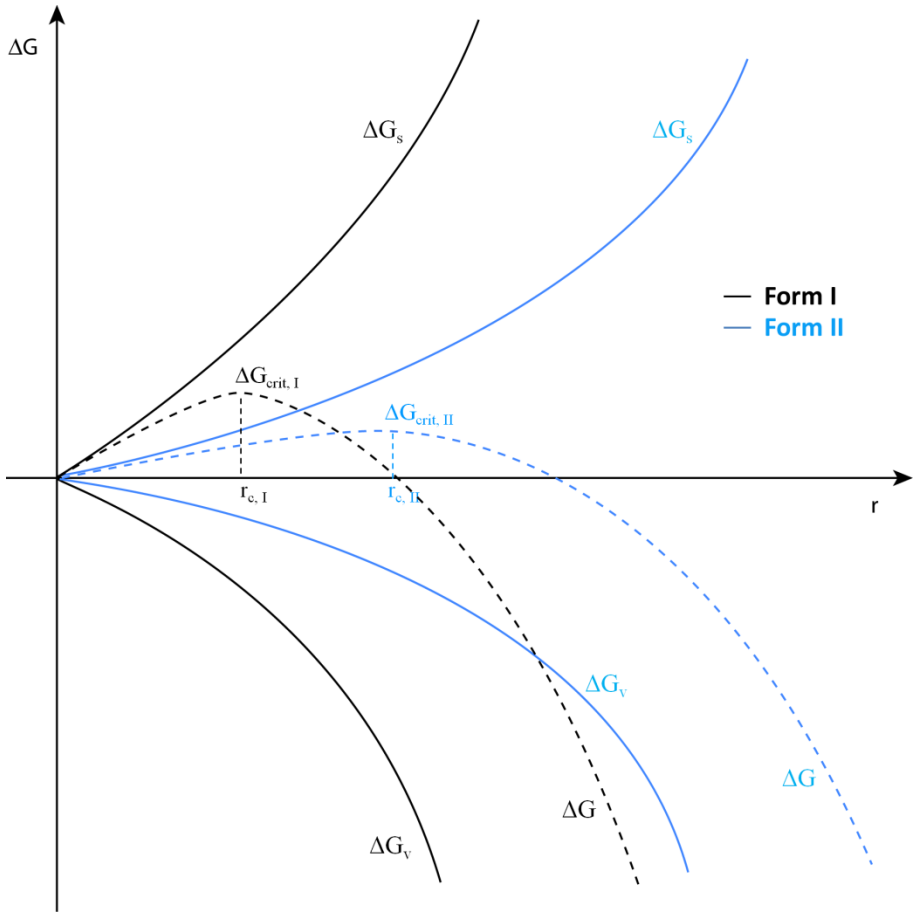

Figure 2.16. Schematic representation of the free energy diagram for nucleation of a dimorphic system, showing the activation barriers for nucleation of Forms I and II. (Adapted from Bernstein et al.)[27]

The induction time often is used as a macroscopic measurement of the nucleation time. The induction time allows for the identification of the underlying nucleation mechanism of a given compound. Also, it allows for a connection to be made between nucleation theory and experimental observations. However, the reproducibility of the measurements

is questionable because it depends on the technique used to detect the formation of a new phase. As such, the measurements could lead to erroneous results during the ex-situ crystallization method employed herein. Also, the induction time is not a fundamental property of the system being studied and may not provide a systematic understanding at the molecular-level of the polymorph formation of TFA. For those main reasons, the induction time was not analyzed in the attempt to determine the relative nucleation rates of TFA polymorphs. Conversely, an understanding of the observed crystallization behavior could be obtained by studying pre-nucleation events in solution. The chemistry of the putative interactions of TFA and its intrinsic properties will be covered in the following chapters.

2.5 Summary

Crystallization of TFA polymorphs in ethanol was found to be affected by growth conditions, such as initial solute concentration and temperature. Both TFA polymorphs crystallize concomitantly under the same conditions. At high temperature and low supersaturation ratio, the metastable Form II was obtained, while at high supersaturation ratio, the most stable Form I started to be produced. At low temperature, an increasing solute concentration resulted in a reduced amount of Form II in the polymorphic mixture.

Results from seeding experiments showed that nucleation is the underlying mechanism responsible for the observed polymorph formation of TFA. However, the effect of solution conditions in the crystallization medium highlighted the limitation of CNT in explaining the nucleation process. Pre-nucleation events may hold the key to achieving a detailed molecular understanding of the polymorph formation of TFA.

Sections of this chapter were previously published.[111] Reprinted with kind permission from Springer Science + Business Media: Pharmaceutical Research, Mattei, A, Li T 2012, Polymorph formation and nucleation mechanism of tolfenamic acid in solution: An investigation of pre-nucleation solute association, 29(2): 460-470. Copyright © 2012 Springer Science + Business Media.

Chapter 3 – Investigation of Solution Chemistry of Tolfenamic Acid

The focus of this chapter is to study the association behavior of tolfenamic acid (TFA) molecules in ethanol solution. The rationale lies in the fact that, prior to the cooling of crystallization, the TFA ethanol solutions were undersaturated but of different concentrations. The extent of self-association as a function of concentration and temperature will be studied. The variation in concentration and temperature may affect how solute molecules behave in solution and consequently result in nucleation of the distinct species. In addition, the effect of concentration and temperature on the conformation of solute molecules in solution will be probed.

3.1 Introduction

Molecules in solution readily form associations through non-covalent interaction, such as hydrogen bonding, dipole-dipole, and electrostatic and hydrophobic interactions. The nature of the associated species in solution is strongly dependent upon the solvent used. The effect of solvent on intermolecular interaction is related to the way the solute molecules interact in solution. For instance, in non-polar solvents, molecules containing a carboxyl group poorly interact with the solvent and solute-solute interactions are favored even at low solute concentrations. On the other hand, molecules containing a carboxyl group in polar solvents highly interact with solvent molecules and solute-solvent interactions are stabilized with respect to solute-solute interactions.[107]

Molecular association in solution also depends on the structure of the carboxylic acid. Acetic acid associated species have been determined to be hydrogen-bonded dimers in

acetonitrile solutions with a self-association equilibrium constant in the range 10^2 kg of solvent/mol.[107] At a low solute concentration, indomethacin does not self-associate to form hydrogen-bonded dimers in acetonitrile and ethanol solutions due to stronger solute-solvent hydrogen bonding, but as concentration increases indomethacin molecules become more closely associated in solution to form dimers via the carboxyl group.[120,121]

Several spectroscopy-based methodologies, including UV, FT-IR, and NMR have been used to investigate the association behavior of small organic molecules in solution.[70,122,123] The studies focus only on the size of the associated species, but not on the structural information at the molecular level. All lack detailed information about the orientation of specific functional groups in the associated species. In this chapter, the formation of associated species will be evaluated by spectroscopic techniques, including UV and NMR. Moreover, the conformation of the interacting molecules in solution will be deduced from qualitative and quantitative Nuclear Overhauser Effect measurements, which can allow the microenvironment of a functional group to be probed. This information will be complemented with size information, derived from pulsed gradient spin-echo NMR experiments.

3.2 Materials and Methods

3.2.1 Materials

TFA was purchased from TCI America (Portland, OR) and was recrystallized from ethanol to obtain pure Form I. Ethanol (> 99.9 percent purity) was obtained from Decon Labs, Inc. (King of Prussia, PA) and was used as received. Potassium chromate (≥ 99 percent) was purchased from Sigma-Aldrich, Inc. (St. Louis, MO). Tetramethylsilane

(TMS), ethanol-d_6 (99 percent D, anhydrous), and D_2O (99.9 percent D) were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA).

3.2.2 UV Spectroscopy

UV spectroscopy measurements were carried out over the concentration range 5.76×10^{-5} – 2.30×10^{-2} M in ethanol at room temperature. Appropriate amounts of the drug were added to volumetric flasks and diluted with absolute ethanol. Each sample was prepared in triplicate. Absorbance spectra of the resultant solutions were recorded in a spectrophotometer (Shimadzu 1800, Shimadzu Scientific Instruments, Inc., Columbia, MD) with quartz cells (New Era Enterprises, Inc., Vineland, NJ) of path length 10, 1, or 0.1 mm at 346 nm, when appropriate. The band at 346 nm corresponds to the n-π^* electronic transition of TFA that is suitable to study the mode of molecular assembly.[124] The path length of a cuvette was calibrated using standard potassium chromate solutions. This was done because of the non-interacting properties of potassium chromate solutions.[125] Values (± standard deviation) thus determined for the path length were 9.99 ± 0.02, 1.01 ± 0.01, and 0.11 ± 0.01 mm. From the absorbance of solutions, apparent molar absorptivities were calculated as a function of solute concentration.

3.2.3 NMR Spectroscopy

All NMR spectra were collected at 600 MHz on an Inova NMR spectrometer (Varian-Agilent, Inc., Palo Alto, CA) equipped with a conventional triple-axis HCN probe. Temperature was controlled by an external unit that pre-conditions the variable temperature (VT) gas (Model TC-84, FTS Systems, Stone Ridge, NY). Data were processed and analyzed using VnmrJ software version 2.2D.

3.2.3.1 ^{1}H-NMR Spectroscopy

TFA Form I was dissolved in ethanol-d_6 at concentrations ranging from 1.5 to 69 mM. The choice of drug concentrations was limited by the signal-to-noise ratio for the lowest concentration and stability of the solution, with respect to crystallization for the highest concentration. ^{1}H-NMR spectra employed a 90° excitation pulse followed by 1.4 s of data acquisition and then a 7.5 s delay for relaxation (T_1 values ranged from 1.6 s up to 2.5 s). Spectra were the result of averaging 16 scans and were acquired at temperatures ranging from 10 °C to 55 °C. When measuring samples at 37 °C and 55 °C, a vortex plug (Wilmad LabGlass, Vineland, NJ) was inserted into a 5 mm NMR tube to prevent solvent evaporation. Chemical shifts were referenced to the methyl protons of ethanol, which resonate at 1.11 ppm relative to the internal TMS at 0 ppm. The solvent signal chemical shift can vary by less than 0.001 ppm/°C,[126] and thus can be reliably used as reference.

3.2.3.2 Pulsed Gradient Spin-Echo (PGSE) NMR

The self-diffusion coefficients of TFA in ethanol-d_6 solutions were measured using the pulsed gradient spin-echo (PGSE) NMR technique.[127] Two magnetic field gradient pulses were applied to examine the effects of translational motion of nuclei on the signal intensity. Self-diffusion coefficients were obtained by a linear least-squares fitting of the attenuation of the NMR echo signal intensity according to the Stejskal-Tanner equation:[127]

$$I = I_0 \exp\left[-(\gamma \delta G)^2 D\left(\Delta - \frac{\delta}{3} \right) \right] \tag{3.1}$$

Here, I and I_0 are the intensities of the resonance in the NMR spectrum measured with and without the magnetic field gradient, respectively; γ is the gyromagnetic ratio of the observed nucleus (it is 4258 Hz/G for ^{1}H); δ and G are the duration and amplitude of the gradient pulses, respectively; Δ is the time period during which diffusion occurs; and D is the self-diffusion coefficient. The duration of the magnetic field pulse gradients (δ) was set to 2 ms. The diffusion time (Δ), equal to *δ+d2+2pw90+d5+d0* in the pulse sequence (Figure A2.1 in Appendix 2), was 500 ms. The gradient amplitude (G) was incremented from 1 to 51 Gauss/cm in 32 linear steps. The pulsed field gradient strength was calibrated with the residual ^{1}H signal in 99.9 percent D_2O, for which the self-diffusion is 1.90 x 10^{-9} m^2/s at 25 °C.[128] Calibration experiments were performed by relating the gradient driver control parameter, usually a digital-to-analog conversion (DAC) value, g, to the resultant gradient strength, G ($G = \kappa g$). The peak intensity data of the residual HDO resonance were analyzed via linear fit to Eq. (3.1) as a function of g^2 to obtain κ using the known diffusion coefficient of D_2O. The gradient calibration constant was found to be (1.14 ± 0.02) x 10^{-2} Gauss/cm/DAC. Solutions of TFA in ethanol-d_6 were injected into a 5 mm NMR tube with a coaxial capillary containing D_2O for calibration. Such an experimental set-up was utilized in order to perform calibration during each set of experiments and to minimize the effect of convection currents caused by small temperature gradients within the sample.[129] Drug solutions were also spiked with TMS (0.4 mM) to be used as an internal standard for diffusion measurements to account for viscosity changes.[130] The number of transients acquired varied from 16 for the supersaturated solutions (65 and 69 mM) to 128 for all of the other concentrations. All diffusion measurements were determined at least in triplicate and obtained with the

sample temperature regulated at 25 °C. Post-acquisition processing included baseline correction before measuring the resonance amplitude of each peak in the data set. To improve the accuracy of the measurements, six peaks corresponding to the TFA molecule were used independently to calculate the diffusion coefficient; thus, the final determined diffusion coefficient represents an average of 18 measurements. A solution containing neat ethanol-d_6 and TMS was also analyzed in order to measure the diffusion coefficients of ethanol and TMS.

3.2.3.3 Nuclear Overhauser Effect Spectroscopy

The Nuclear Overhauser Effect (NOE) corresponds to the fractional change in peak integration of one NMR resonance when a neighboring resonance is saturated. The effect depends strongly on the internuclear distance (r), in that the fractional peak integral change is proportional to r^{-6}. One-dimensional (1D) NOE experiments were performed at TFA concentrations of 1.5, 10.4, and 62.5 mM in ethanol-d_6 at 25 °C. Spectra of 64 scans were acquired using a rectangular pulse to selectively perturb proton frequencies with different mixing times (τ_m) ranging from 0 to 4.6 s. NOE build-up data were fitted to the following exponential function:[131]

$$\mathrm{NOE} = \mathrm{Exp}\left[-(R - \sigma)\,\tau_m\right]\left[1 - \mathrm{Exp}(-2\,\sigma\,\tau_m)\right] \qquad (3.2)$$

where R is the total longitudinal relaxation rate constant of both nuclei; σ is the cross-relaxation rate that represents the rate at which NOE is transferred between two nuclei; and τ_m is the duration of the cross-relaxation between two nuclei. The NOE is a sensitive measure of the distance between observed and perturbed nuclei. As such, the

corresponding internuclear distances were subsequently calculated by comparing cross-relaxation rates σ_{is} with σ_{ref} according to the equation:[131]

$$\frac{\sigma_{is}}{\sigma_{ref}} = \left(\frac{r_{ref}}{r_{is}} \right)^6 \tag{3.3}$$

where σ_{ref} and r_{ref} are the reference cross-relaxation rate and the reference distance, respectively, for a pair of protons separated by a known distance; and σ_{is} and r_{is} are the cross-relaxation rate and the distance, respectively, for a pair of protons of interest whose distance is to be calculated. NOE build-up experiments served as a means of both estimating proton-proton distance and choosing the appropriate mixing time to be used in Nuclear Overhauser Effect Spectroscopy (NOESY) experiments. Because incorrect choice of the mixing time can cause absence of observable signals, it is necessary to ensure the mixing time lies within the linear growth region of the NOE enhancement in order to maximize the NOE signal and the efficacy of the experiment.[132]

Two-dimensional Nuclear Overhauser Effect Spectroscopy (2D-NOESY) was performed on solutions of TFA in ethanol-d_6 at varying concentration and temperature. The pulse sequence used in NOESY experiments consists of three 90° pulses (Figure A2.2 in Appendix 2). The t_1 evolution time was systematically incremented to provide chemical shift information along the F_1 dimension. NOESY spectra for each sample were recorded with a mixing time of 0.8 s; 2074 data points were collected in the F_2 dimension and 200 in the F_1 dimension and 56 scans were time-averaged for each t_1 increment. NOESY spectra were weighted using a Gaussian function in both the F_1 and F_2 dimensions prior to Fourier transformation.

3.2.4 Data Analysis

Non-linear least-squares regression analyses were performed using the software program, Scientist (Micromath Scientific Software, St. Louis, MO). The lack-of-fit test analysis was performed using Scientist and Microsoft Excel. Data are reported as mean ± standard deviation.

3.3 Results and Discussion

3.3.1 Association Behavior of Tolfenamic Acid

3.3.1.1 UV Absorptivity Measurements

To understand the solution chemistry of the system, UV spectroscopy was performed to identify the occurrence and extent of association of TFA solute molecules in ethanol as a function of concentration (below the solubility). By measuring UV absorbance at different concentrations, the molar absorptivity was calculated and plotted as a function of concentration (Figure 3.1). Hypochromic deviation from the Beer-Lambert law is evident, suggesting self-association of solute molecules.[133] The total absorbance may be regarded as a sum of individual species in the solution:

$$\varepsilon C = \varepsilon_m C_m + \varepsilon_n n K_{1,n} (C_m)^n \qquad (3.4)$$

where ε is the apparent molar absorptivity determined from the experimental absorbance; and ε_m and ε_n are the molar absorption coefficients of monomers and *n*-mer aggregates of solute molecules, respectively. In addition, n is the size of *n*-mer aggregates (i.e., 2 for dimer); $K_{1,n}$ is the self-association equilibrium constant of the aggregation; C is the

overall solute concentration; and C_m is the monomer concentration. The equilibrium constant and the molar absorption coefficient of the aggregate species were used as fitting parameters. The best fit of the experimental data based on Eq. (3.4) could then be obtained (solid line in Figure 3.1) with a correlation coefficient of 0.98. The calculated value (± standard deviation) of $K_{1,n}$ was found to be (1.38 ± 0.34) x 10^2 M^{-1} and the corresponding n equals 2. The fraction of molecules present in dimers can be expressed as follows:

$$f_{dimer} = \frac{2K_{1,2}(C_m)^2}{C} \tag{3.5}$$

Consequently, the extent of dimer formation was derived, indicating the trend of dimer association as the solute concentration increases (dashed line in Figure 3.1). Note that the derived model assumes that dimers are the only species in equilibrium with monomer over the concentration range employed. Higher-order association species are plausible. Therefore, to determine the probability of the formation of aggregates of a higher order than dimers, the molar absorptivity versus drug concentration profile was fit to Eq. (3.4) with varying the size of *n*-mer aggregates (i.e., n=2 and 3). A statistical lack-of-fit test analysis was performed for the model fitting and the results of the analysis are shown in Table 3.1. Although both models do not meet the requirements of the lack-of-fit test analysis ($F_{calculated} < F_{critical}$), the rank order of the models based on their calculated F values indicates that the dimerization model (i.e., n=2) provides the best fit for TFA molar absorptivity profile. The evidence obtained from this study is not sufficient to support any one model unequivocally. Nevertheless, formation of higher-order association species seems unlikely or negligible.

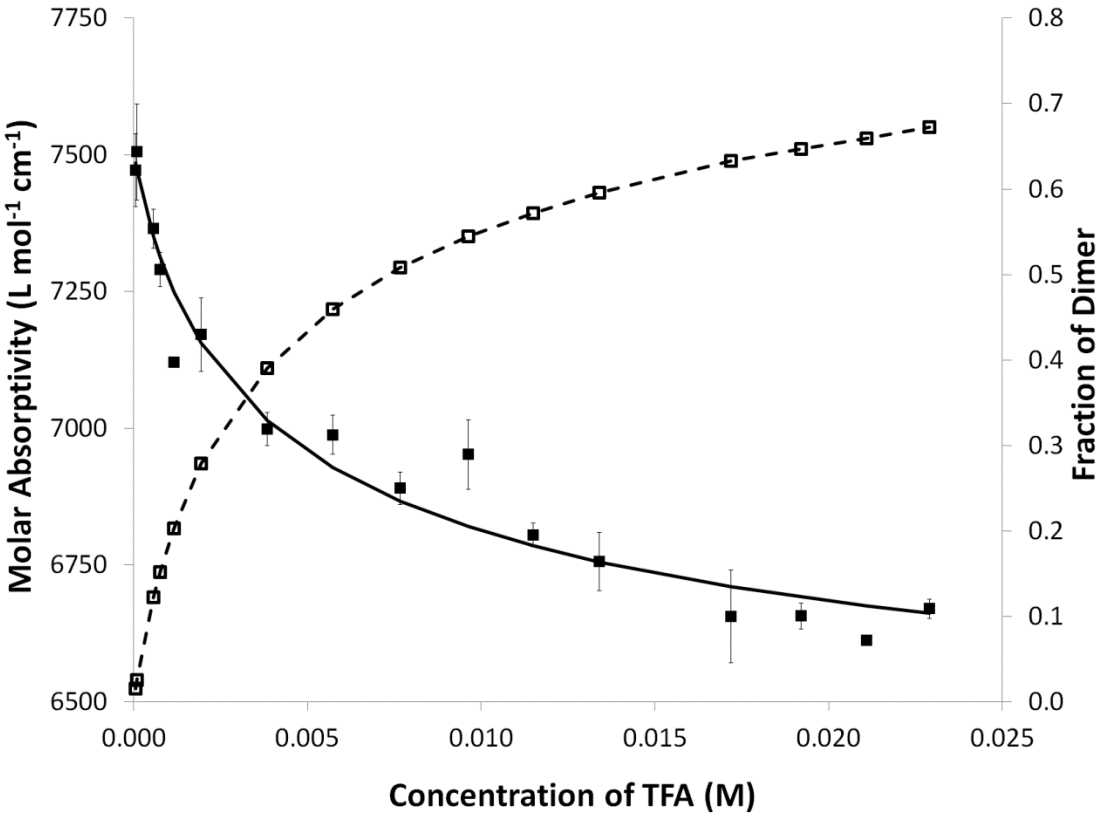

Figure 3.1. Concentration dependence of the molar absorptivity of TFA at λ_{max} = 346 nm. The solid line is a fit of the data according to Eq. (3.4); the dimer fraction is also shown as dashed line versus the total TFA concentration. Error bars represent standard deviation (n=3).

Table 3.1. Lack-of-fit statistical analysis of TFA molar absorptivity profiles with varying the size of *n*-mer aggregates (n).

n	Sum of Squares (Residuals)	Sum of Squares (Lack of Fit)	F Value (Calculated)	F Value (Critical; $\alpha = 0.05$)
2	2.22×10^{-5}	1.47×10^{-5}	4.4	2.0
3	3.90×10^{-5}	3.15×10^{-5}	10.0	

3.3.1.2 NMR Chemical Shift Measurements

The ^{1}H chemical shift is sensitive to subtle changes in the local environment of the molecule and it is often used to study molecular association.[134,135] A molecular association event often manifests itself by perturbations in peak positions in the NMR spectrum. Therefore, ^{1}H-NMR was performed on a series of concentrations of TFA at 25 °C. The aromatic region of the ^{1}H-NMR spectra of TFA at concentrations of 1.5 and 62.5 mM showed a chemical shift variation when the solution was diluted (Figure 3.2). Specifically, the chemical shifts of all protons were shifted to stronger fields with increasing concentration of TFA, suggesting that molecular association was involved. The chemical shift changes of three selected resonances as a function of molar drug concentration were fit to the model that considers a single associated species to be in equilibrium with the monomer, as follows:

$$\delta_{obs} = \frac{\delta_m C_m + \delta_n n K_{1,n} (C_m)^n}{C} \qquad (3.6)$$

where δ_{obs} is the measured chemical shift; δ_m and δ_n are the chemical shifts of the monomer and n-mer aggregate, respectively; C is the total TFA concentration; C_m is the molar concentration of the monomer; n is the size of n-mer aggregates; and $K_{1,n}$ is the equilibrium constant of the self-association. The equilibrium constant and the chemical shift of the aggregate species were used as fitting parameters. Monomer and aggregate species in the solution seem to be in fast exchange with respect to the NMR time scale because either distinct resonances due to both species or line broadening cannot be observed in the proton spectra. Thus, the observed chemical shift can be expressed as the weighted average of the monomer and aggregate chemical shifts. The concentration-

dependent chemical shift profiles (Figure 3.3, solid line) were fit well using n=2 (i.e., the dimerization model). A statistical lack-of-fit test analysis was performed for the model fit and the results indicated that a monomer-dimer model conforms to the requirements of the test, with the F-statistical value less than the F-distribution at the significance level of $\alpha = 0.05$ (Table 3.2). A lack-of-fit test for a monomer-trimer model showed sufficient statistical evidence that there was lack of fit. Thus, the model was considered to be unsuitable for the analysis of the chemical shift data. The self-association equilibrium constants ($K_{1,2}$) obtained independently from the curve fitting of different aromatic protons were consistent and an average value was found to be $(0.31 \pm 0.07) \times 10^2\ M^{-1}$, which is different from that obtained previously in studies of UV measurements. Such a discrepancy in the derived self-association constant may stems from at least two possibilities. First, a different concentration range is employed. A narrow concentration range was used for UV measurements and the estimated association constant may depend upon the concentration range used. Second, the dimerization model proved to have limitations with UV data based on the statistical analysis.

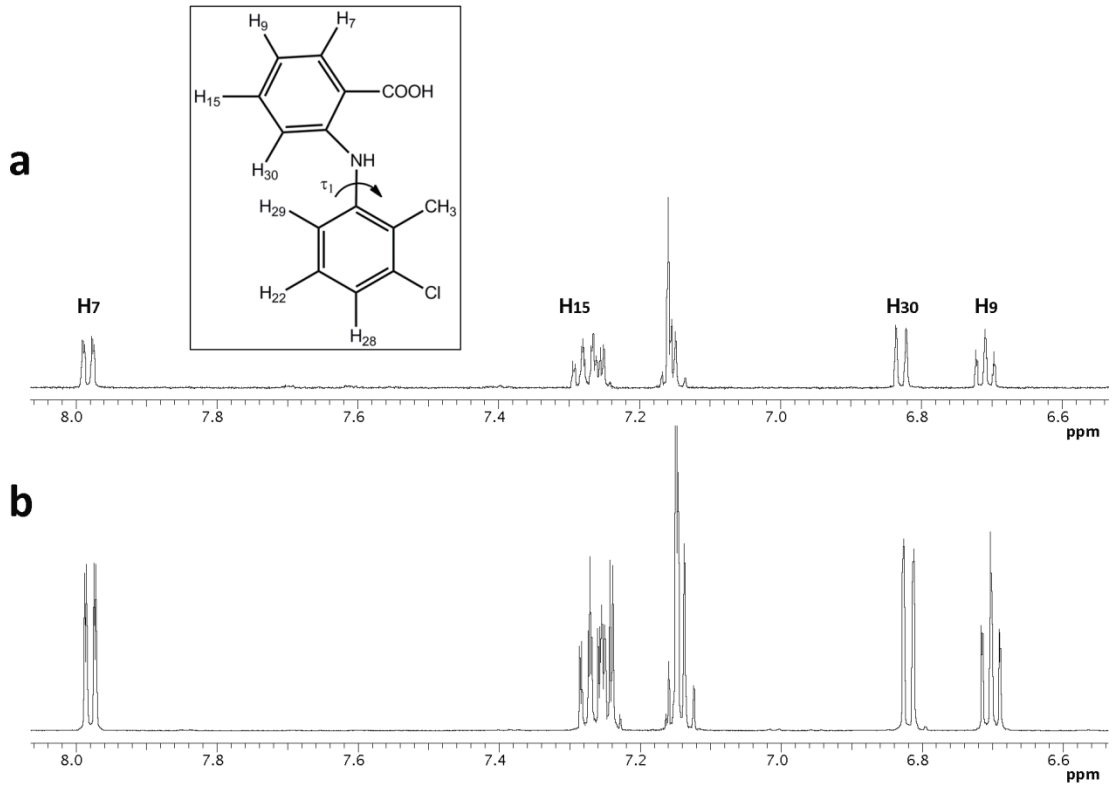

Figure 3.2. ^{1}H-NMR spectra of TFA in ethanol-d_6 at 25 °C recorded at 1.5 mM (a) and 62.5 mM (b); resonances for characteristic peaks are highlighted. Chemical structure and the numbering system of TFA are shown in the inset.

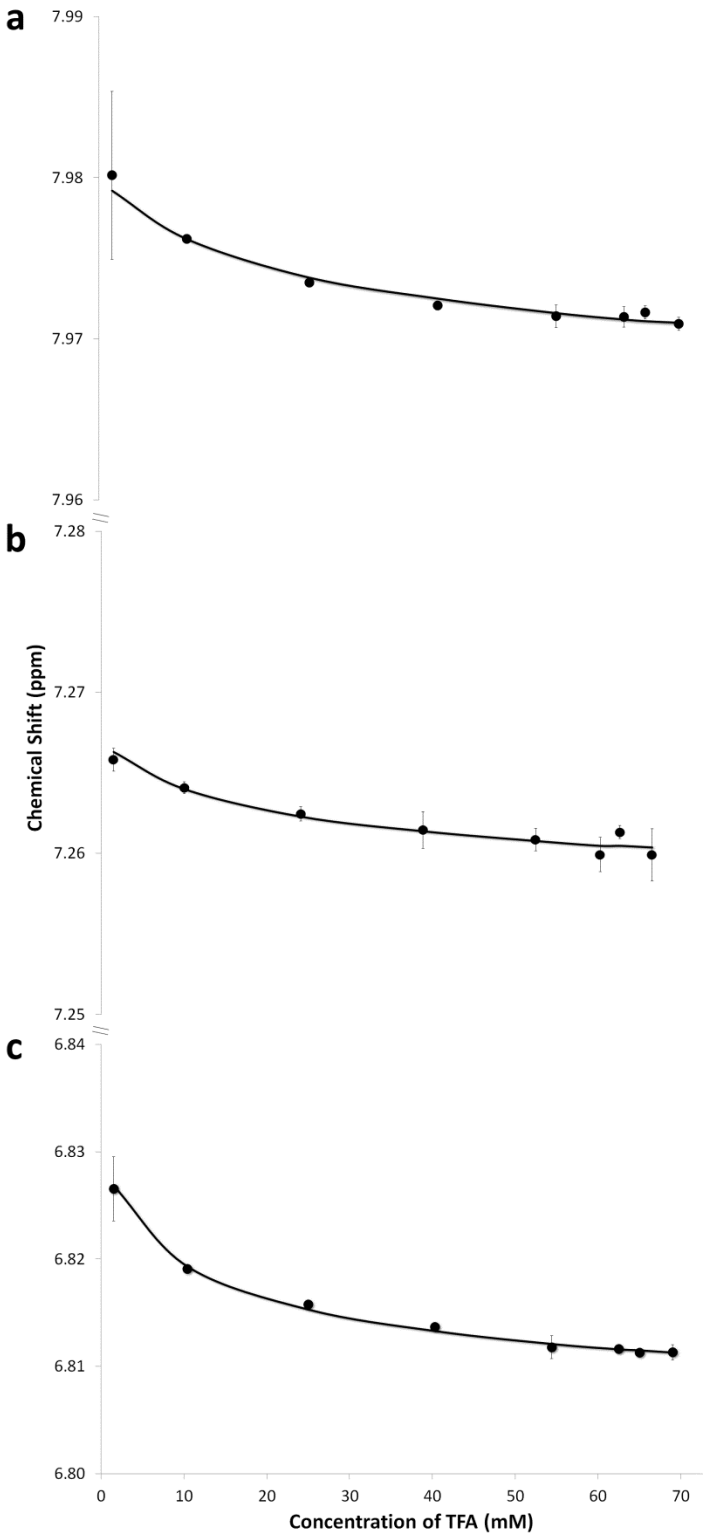

Figure 3.3. Dependences of the chemical shifts for H_7 (a), H_{15} (b), and H_{30} (c) on the concentration of TFA at 25 °C. Error bars show the standard deviation of three independent samples. The solid lines represent the best fit obtained using Eq. (3.6).

Table 3.2. Lack-of-fit statistical analysis of TFA chemical shift profiles at 25 °C with varying the size of *n*-mer aggregates (n). Correlation coefficients for the model fits (R^2) are also reported.

n	Protons	Sum of Squares (Residuals)	Sum of Squares (Lack of Fit)	F Value (Calculated)	F Value (Critical; $\alpha = 0.05$)	R^2
2	H_7	6.16×10^{-5}	4.72×10^{-6}	0.22	2.74	0.99
	H_{15}	1.77×10^{-5}	4.53×10^{-6}	0.92		0.98
	H_{30}	7.96×10^{-5}	1.82×10^{-6}	0.06		0.99
3	H_7	1.26×10^{-4}	6.90×10^{-5}	3.23		0.94
	H_{15}	3.22×10^{-5}	1.91×10^{-6}	3.90		0.89
	H_{30}	1.69×10^{-4}	9.08×10^{-5}	3.11		0.93

Studies were conducted to evaluate how temperature affects the self-association of TFA molecules. Intuitively, if there is an increase in the thermal motion of molecules, there should be a decrease in the probability that the molecules would form dimers. The concentration dependence of the chemical shifts of TFA protons was measured at varying temperatures. Similar to the measurements carried out at 25 °C, when changing temperature, the chemical shifts of the three selected protons decreased as concentration increased (Figure 3.4). The proton H_{30} exhibited a greater change in chemical shift as temperature was varied (Figure 3.4c). This means that H_{30}'s chemical environment is more affected than that of other aromatic protons by the association of TFA molecules. The chemical shifts versus concentration profiles were fit to a dimerization model according to Eq. (3.6) to estimate the self-association equilibrium constant at various temperatures. The estimated association constants of TFA molecules were $(0.47 \pm 0.14) \times 10^2$ M^{-1} at 10 °C, $(0.09 \pm 0.05) \times 10^2$ M^{-1} at 37 °C, $(0.04 \pm 0.04) \times 10^2$ M^{-1} at 55 °C. As such, the fraction of TFA molecules existing as dimer in solution was calculated from Eq. (3.5) and plotted as a function of total drug concentration at different temperatures

(Figure 3.5). At low temperature, the fraction of dimer is relatively high, but significantly reduces with increasing temperature to give a concomitant increase in concentration of the monomer species in solution. The analysis of the temperature dependence of the association constants enabled the thermodynamic parameters of TFA molecular self-association to be estimated. The enthalpy and entropy of TFA self-association were estimated from the van't Hoff plot (Figure 3.6). The magnitude of the enthalpy of formation of TFA dimers in ethanol solution was -45.0 kJ/mol and the entropy was 1.25 x 10^2 J/mol. The negative sign of the enthalpy of dimerization indicates that the formation of TFA dimers in ethanol solution is enthalpically favored. The negative value of the enthalpy of dimerization may be determined by intermolecular interactions, specifically hydrogen bonding, which will be discussed in the following chapters.

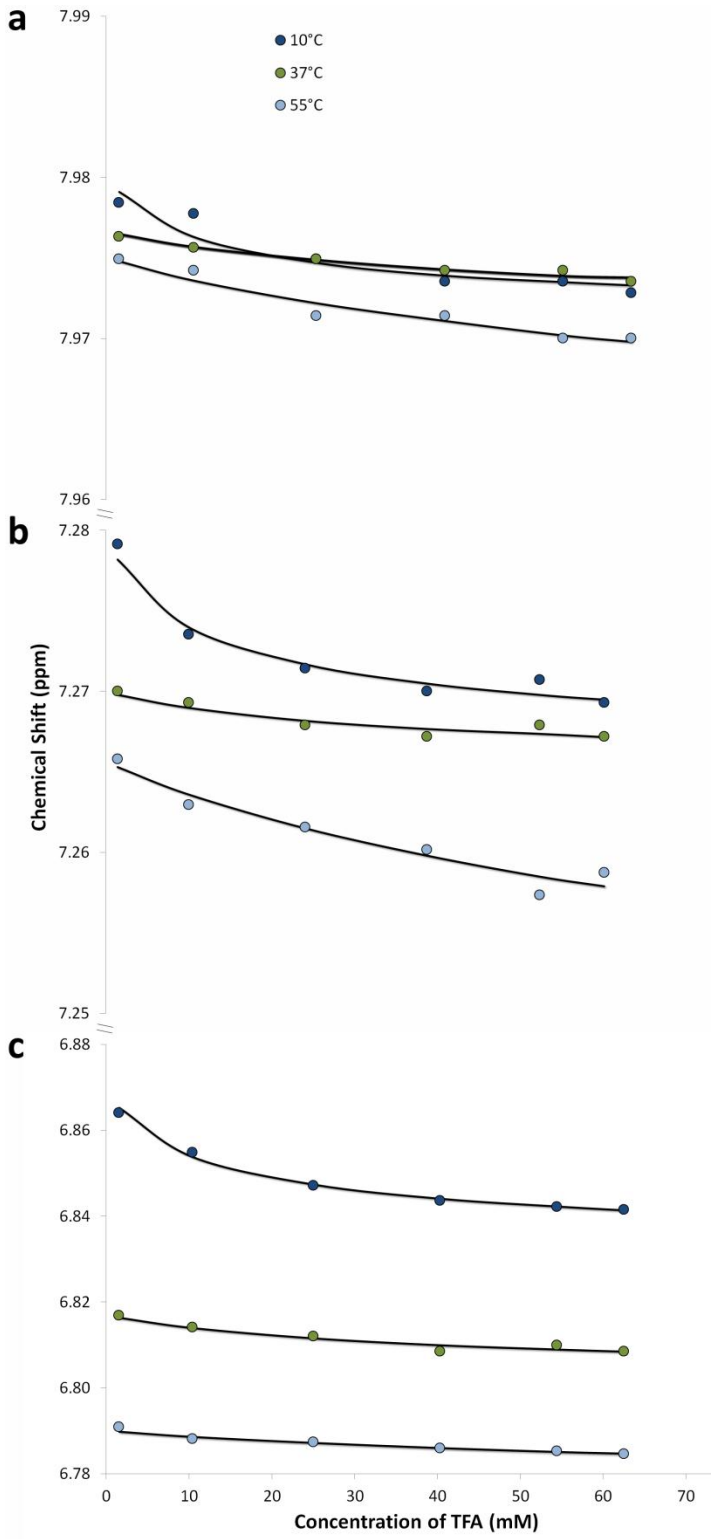

Figure 3.4. Dependences of the chemical shifts for H_7 (a), H_{15} (b), and H_{30} (c) on the concentration of TFA at 10 °C, 37 °C, and 55 °C. The solid lines represent the best fit obtained using Eq. (3.6).

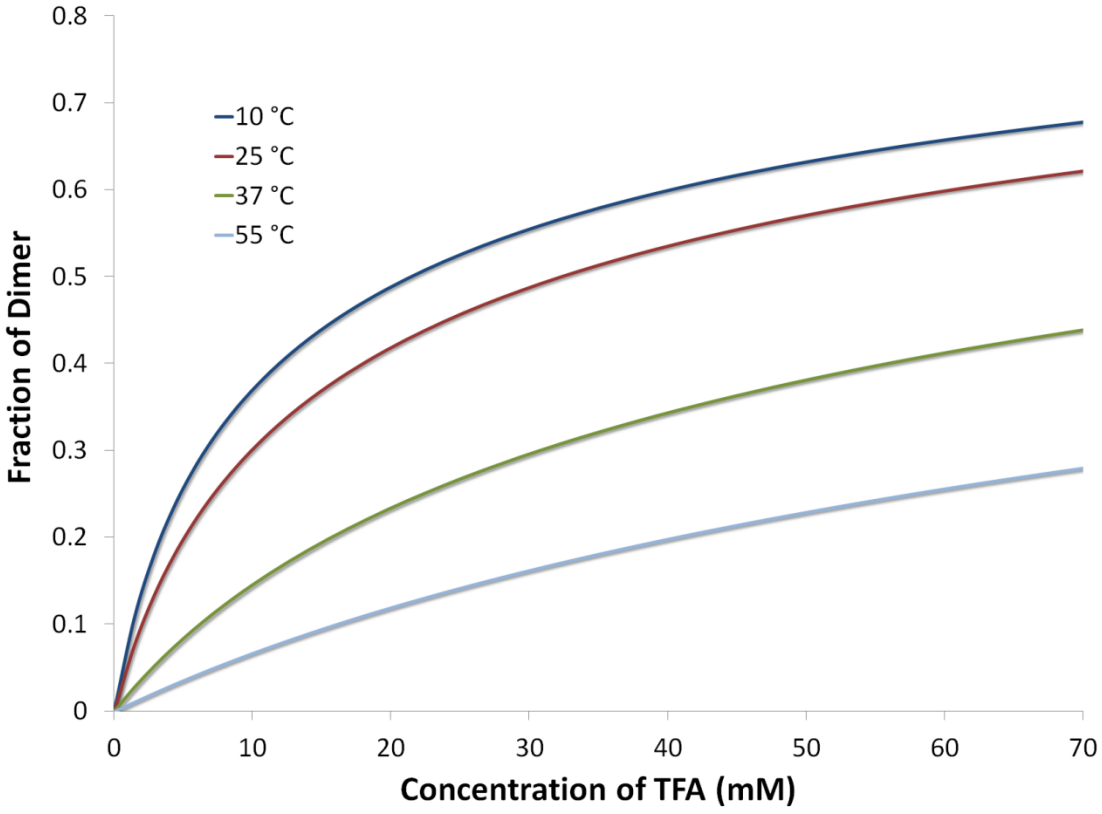

Figure 3.5. Molar fraction of TFA dimer estimated at different temperatures from Eq. (3.5).

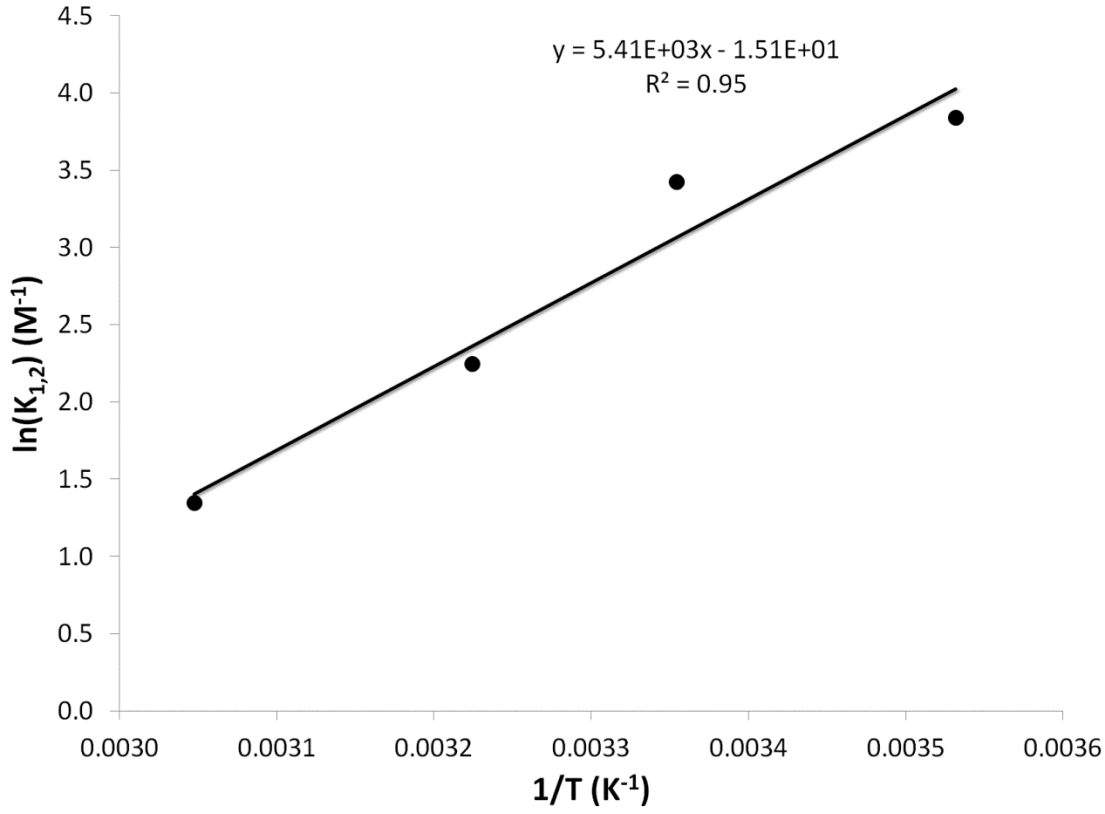

Figure 3.6. Van't Hoff plot of the self-association constant (K1,2) for TFA in ethanol-d_6.

The results suggest that TFA molecules self-associate in ethanol solution to form dimers. The observed trend of the chemical shifts is consistent with two possible structures of dimers. First, the formation of hydrogen-bonded dimers between neighboring carboxyl groups may increase the electron density on the protons of the anthranilic ring and the effect is reflected in a chemical shift decrease of these protons.[121,136] Second, the formation of π-π interactions affords the "ring current" effect on the aromatic protons, which shift to decreased chemical shifts[137] As a result of the major difference in the chemical shift of H_{30} with temperature, hydrogen-bonded dimer is the plausible geometry that can be extracted from the data available. At this point, the results are more consistent with a role for H_{30} in determining the structure of the associated species. To further assess

the structural properties of the associated species, NMR self-diffusion coefficient measurements and NOESY experiments were conducted.

3.3.1.3 Diffusion Coefficient Measurements

The solution behavior of TFA in ethanol-d_6 was investigated by determining the concentration dependence of the self-diffusion coefficient at 25 °C using PGSE NMR methodology. The study relies on the fact that association phenomena, which involve changes in hydrodynamic radii of the molecule, can be monitored through changes in diffusion. The diffusion coefficient is related to the hydrodynamic radius through the Stokes-Einstein equation, assuming the solute to be spherical:

$$D = \frac{kT}{6\pi \eta r} \tag{3.7}$$

where D is the diffusion coefficient; k is the Boltzmann constant; T is the temperature; η is the viscosity of the solution; and r is the hydrodynamic radius.

Hydrodynamic radii calculated from experimental diffusion coefficients have mainly been utilized in structure assignment procedures.[131] It is nonetheless necessary to account for changes in the viscosity of the solution before drawing conclusions about any changes in the hydrodynamic radii. In order to obtain information about size that is independent of the solution viscosity, tetramethylsilane (TMS) was used as the internal standard in the solutions of TFA at different concentrations. This was done because of TMS's non-interacting properties, making it possible to correct the measured diffusion values for any changes in the viscosity.[130]

The measured diffusion coefficients of ethanol and TMS in the reference solution of neat ethanol-d_6 were (1.01 ± 0.01) x 10^{-9} m^2/s and (1.34 ± 0.03) x 10^{-9} m^2/s, respectively. The diffusion coefficient of ethanol was in good agreement with the literature reported value in bulk ethanol if it scaled with solvent viscosity.[138] The diffusion coefficient of TMS was determined as drug concentration increased (Figure 3.7 solid circles). The relatively constant diffusion values of TMS upon increasing drug concentration indicated that the viscosity of the solution was only slightly affected by modification of the solution composition. Indeed, the viscosity values were nearly constant throughout the concentration range used (open circles superimposed in Figure 3.7). The medium viscosity was determined using the Stokes-Einstein equation and considering the hydrodynamic radius of TMS constant for different concentrations. The diffusion coefficient of TFA, obtained after correction for the viscosity variations in the solution (η x D), varied non-linearly with increasing concentration; such a change in diffusion can be mainly attributed to the formation of solute-solute associated species (Figure 3.8).

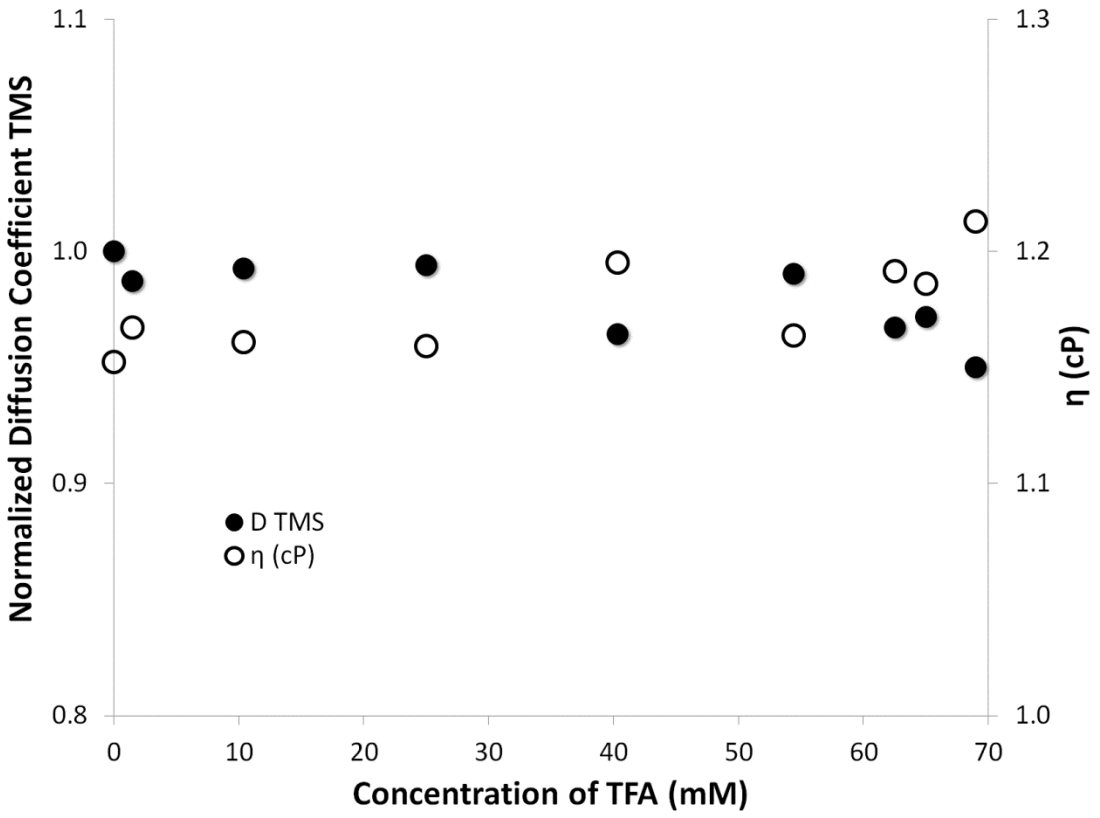

Figure 3.7. Concentration dependence of the diffusion coefficients of TMS in solution of ethanol-d_6 at 25 °C (solid circles). The diffusion coefficients of TMS in TFA-containing solutions were normalized by the diffusion coefficient of TMS in the reference solution of neat ethanol-d_6 and TMS. Viscosity of TFA solutions is shown at 25 °C (open circles).

In all of the studies, the signal attenuation as a function of the pulsed gradient amplitude was fit to the mono-exponential Stejskal-Tanner function (Figure A2.3 in Appendix 2). In literature, a bi-exponential fitting function has been applied when the investigated chemical species are present in two different populations and the molecular exchange between them is slow on the time scale of Δ (i.e., diffusion time).[129,139] The excellent fit of the mono-exponential function to the data can be explained by a rapid chemical exchange between species existing in different association states. For that reason, the

measured diffusion coefficient was considered as a weighted average of the diffusion coefficients of the monomer and dimer.

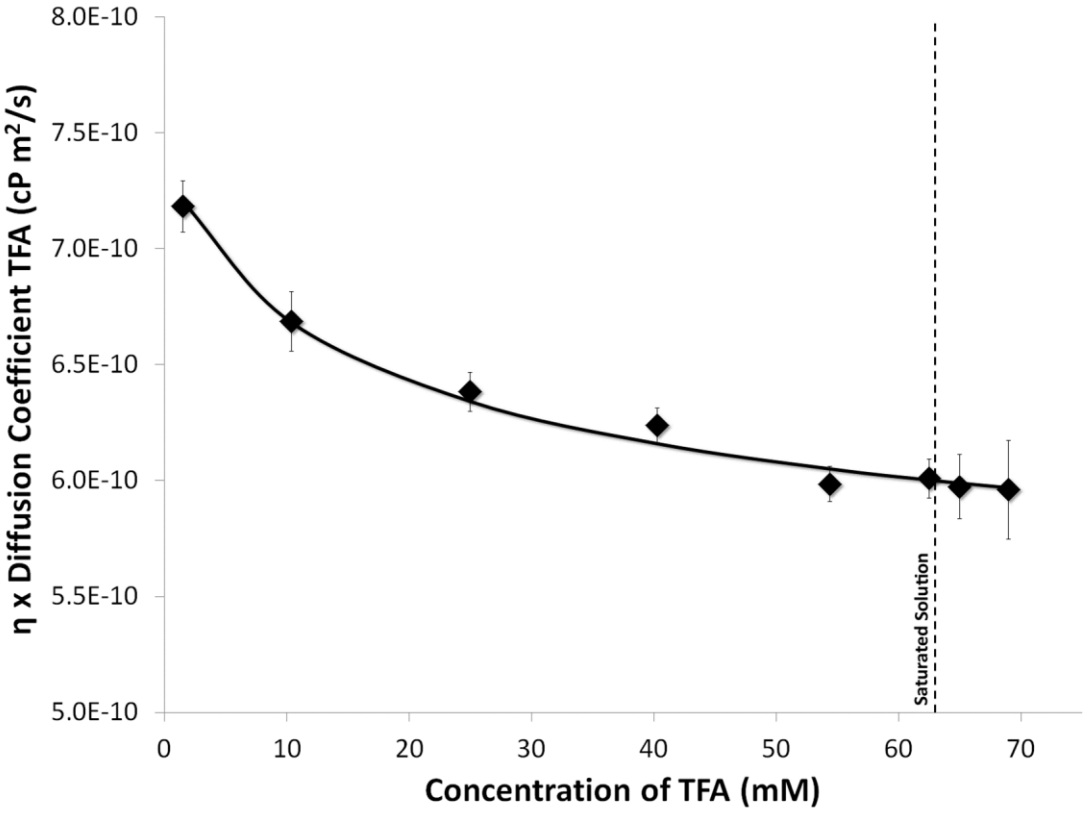

Figure 3.8. Concentration dependence of the diffusion coefficients of TFA corrected by viscosity in ethanol-d_6 solutions at 25 °C. Error bars show the standard deviation (n=18). The solid line represents a fit of the data to Eq. (3.8). The dashed vertical line corresponds to the saturated concentration of TFA in ethanol.

The dependence of diffusion coefficient on concentration was modeled according to the following equation:

$$D'_{obs} = \frac{D'_m C_m + D'_d 2K_{1,2}(C_m)^2}{C} \quad (3.8)$$

where D'_{obs} is the observed diffusion coefficient; and D'_m and D'_d are the diffusion coefficients of TFA monomer and dimer, respectively. Note that the diffusion coefficients, D', are corrected for changes in the solution viscosity. C is the total TFA concentration; C_m is the molar concentration of the monomer; and $K_{1,2}$ is the self-association equilibrium constant, derived from the concentration dependence of the chemical shifts at 25 °C. The diffusion coefficients of the monomer and dimer were used as fitting parameters. The mathematical model gave a good fit (Figure 3.8 solid line) and the corresponding statistical analysis is shown in Table 3.3. The diffusion coefficients of the monomer and dimer were estimated to be (7.38 ± 0.05) x 10^{-10} and (5.10 ± 0.05) x 10^{-10} cP m^2/s, respectively. The ratio of the diffusion coefficient of the dimer to that of the monomer can be used as a signature of the molecular association process. The approach is to assume that monomer has a spherical shape, whereas dimer, resembling two spheres in contact, takes on the shape of a prolate ellipsoid. The ratio for the formation of TFA dimer was found to be 0.69, which is in reasonable agreement with the theoretical value of 0.71 obtained by approximating the monomer-monomer interaction as a hard-sphere molecular contact.[140] The hard-sphere model has been applied to estimate the association state of both proteins[141] and small molecules.[140] The hydrodynamic radii, which are not dependent upon the solution viscosity, were thereby calculated using the Stokes-Einstein equation and plotted as a function of TFA concentration in order to detect the extent to

which the molecular size may change over the concentration range studied (Figure 3.9). The results support that the monomer-dimer model is the predominant association process and that higher-order association species are insignificant based on molecular size measurements.

Table 3.3. Lack-of-fit statistical analysis of TFA diffusion coefficient profiles at 25 °C with n=2 being the size of *n*-mer aggregates. Correlation coefficient for the model fit (R^2) is also reported.

n	Sum of Squares (Residuals)	Sum of Squares (Lack of Fit)	F Value (Calculated)	F Value (Critical; $\alpha = 0.05$)	R^2
2	6.81×10^{-21}	1.31×10^{-21}	0.60	2.74	0.99

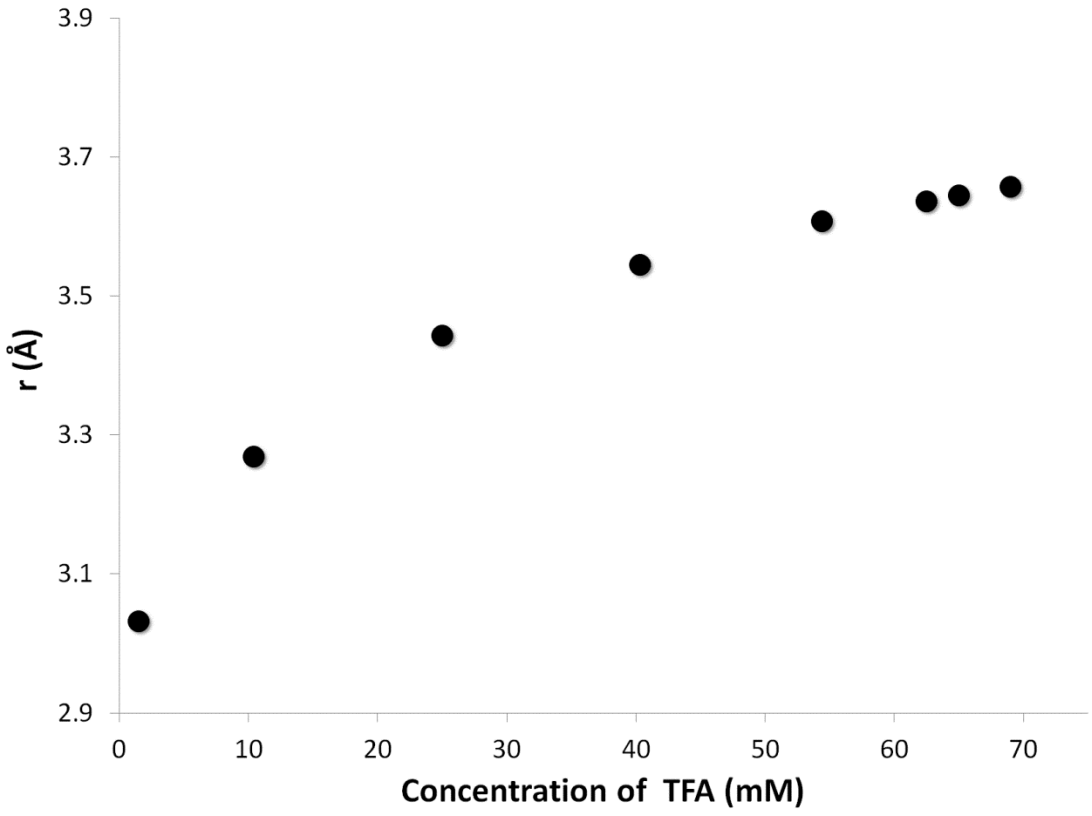

Figure 3.9. Hydrodynamic radii of TFA molecules in ethanol-d_6 solution as a function of drug concentration.

3.3.2 Concentration and Temperature Dependence of Molecular Conformation

TFA exhibits conformational flexibility; the two major conformers of the crystal structures, Forms I and II, differ mainly in the orientation of the chlorinated aromatic ring. NOE methodology is typically employed in structural analysis as a means to characterize the spatial relationship between two nuclei that are located within 5 Å. Since the NOE depends upon internuclear separations, it is also a tool for studying intermolecular interactions. Thus, NOE experiments were performed to seek evidence for

conformational changes in solution and, specifically, to elucidate the spatial arrangement of TFA molecules in the dimer state that is formed in solution.

In 1D-NOE experiments, selective saturation of H_9 produced NOEs on H_{15} and H_7, which belong to the anthranilic ring (Figures 3.10a and 3.10c). Note the clean and well-resolved NOE peaks with a flat baseline. The NOE generated at H_{15} by saturation of H_9 was used as a reference value because these protons belong to a region of the molecule that is considered to be invariant in the two TFA conformations. In the spectrum at 62.5 mM, NOE was observed from H_{30} to the methyl group, which nonetheless showed to be very weak in the spectrum at 1.5 mM (Figures 3.10b and 3.10d). Cross-relaxation rates of H_{15} and H_{30} were then estimated following saturation of H_9 and the methyl group, respectively, by fitting the NOE build-up curves according to Eq. (3.2) (Figure 3.11).

Various factors may perturb NOE intensities, including additional cross-relaxation pathways or spin diffusion, selective polarization transfer from inhomogeneous inversions, and variation in effective rotational correlation time between spins.[142] For organic molecules, tumbling rapidly in non-viscous solvents (i.e., in the extreme narrowing limit) NOEs are roughly independent of the correlation time. TFA described herein falls within the extreme narrowing regime. Also, NOE build-up curves for the selected proton pairs do not exhibit a sigmoidal behavior (i.e., an induction period characteristic of spin-diffusion)[132] in the initial stage of NOE enhancement (Figures 3.11a and 3.11b). Thus, all interactions were treated as direct NOE enhancements. The interproton distance H_9-H_{15} was measured from the reported crystal structures[114] and was used to calculate the distance between protons of interest according to Eq. (3.3). The apparent distance between H_{30} and the methyl protons describes TFA conformational

flexibility and allows the Form I and II conformers to be distinguished as it reflects the orientation around the N-C bond in the chlorinated phenyl moiety. The interproton distance methyl-H_{30} will likely reflect an r^{-6} population-weighted average of the cross-relaxation rates. At the lowest analyzed concentration of 1.5 mM, the time-dependent build-up of NOE effect for the selected H_{30} and methyl protons could not be constructed because the interaction from H_{30} to the methyl protons is too weak or not persistent enough (Figure 3.11b). This indicates that in a solution at low concentration, the selected pair of protons is at a larger distance than the allowable distance for NOE measurement. Molecular conformations in which two protons are at a larger distance than 5 Å are not sufficiently populated or are too short-lived. In the flatter Form II conformer, the distance of closest approach between H_{30} and the methyl protons is 4.5 Å.[114] For the solution concentrations of 10.4 and 62.5 mM, the experimental distance between H_{30} and the methyl protons was found to be about 3.0 Å, which is close to that observed in the crystal structure of the Form I conformer (i.e. 2.9 Å).[114] If flexible molecules exhibiting multiple conformations in solution are interconverting rapidly on the NMR time scale, then conformational exchange will lead to an average of the observed NOEs for each corresponding interproton distance in each conformer. In any case, given that the experimental distance between H_{30} and the methyl protons is comparable to the value corresponding to the Form I conformer in the crystal structure, the precision obtained is significant. As such, a twisted conformation exists in solution.

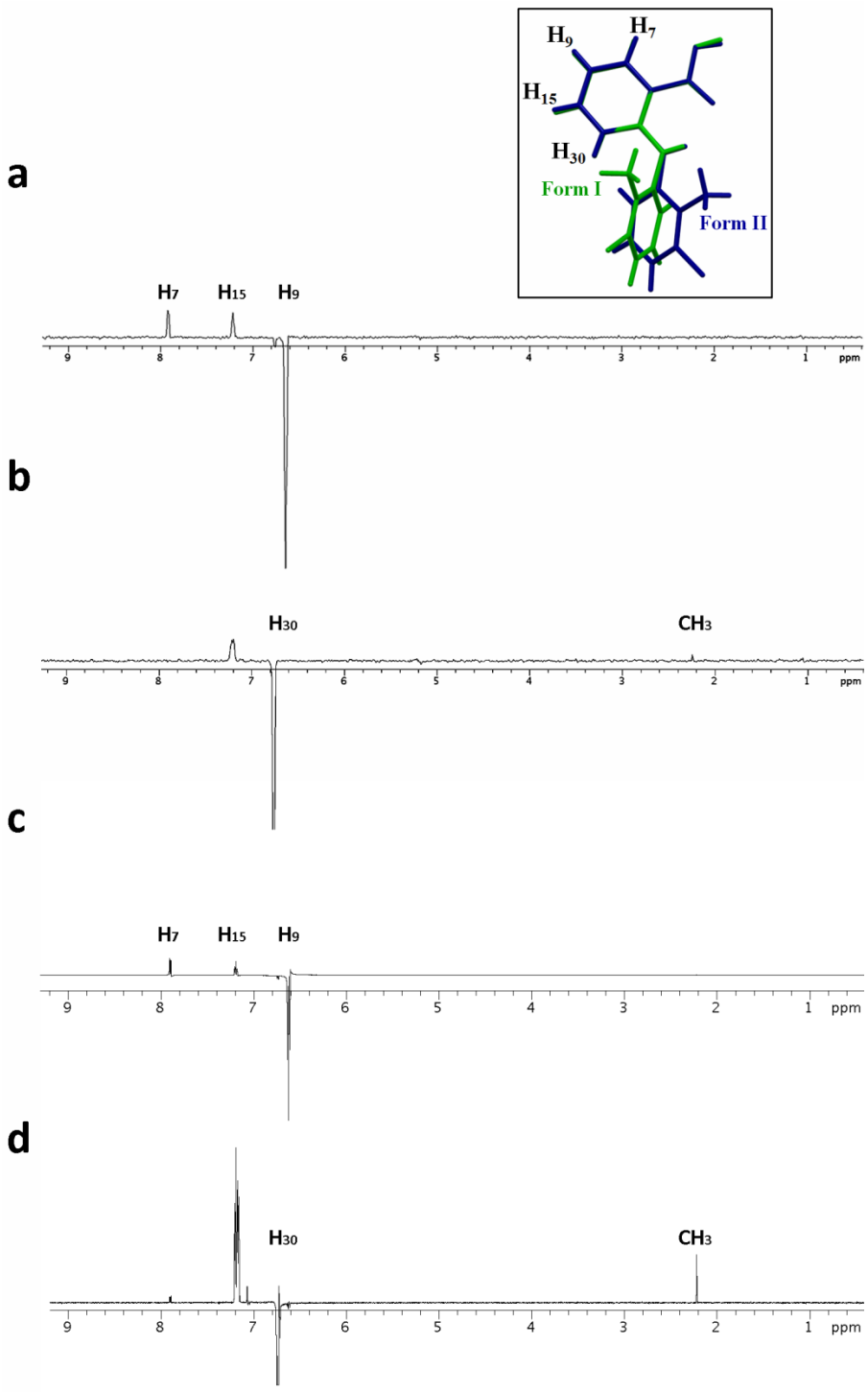

Figure 3.10. One-dimensional NOE spectra of 1.5 mM TFA in ethanol-d_6 (a, b) due to selective saturation of H_9 and H_{30}, respectively, and of 62.5 mM in ethanol-d_6 (c, d) due to selective saturation of H_9 and H_{30}, respectively. The inset shows the conformers of TFA Forms I and II together with the numbering system.

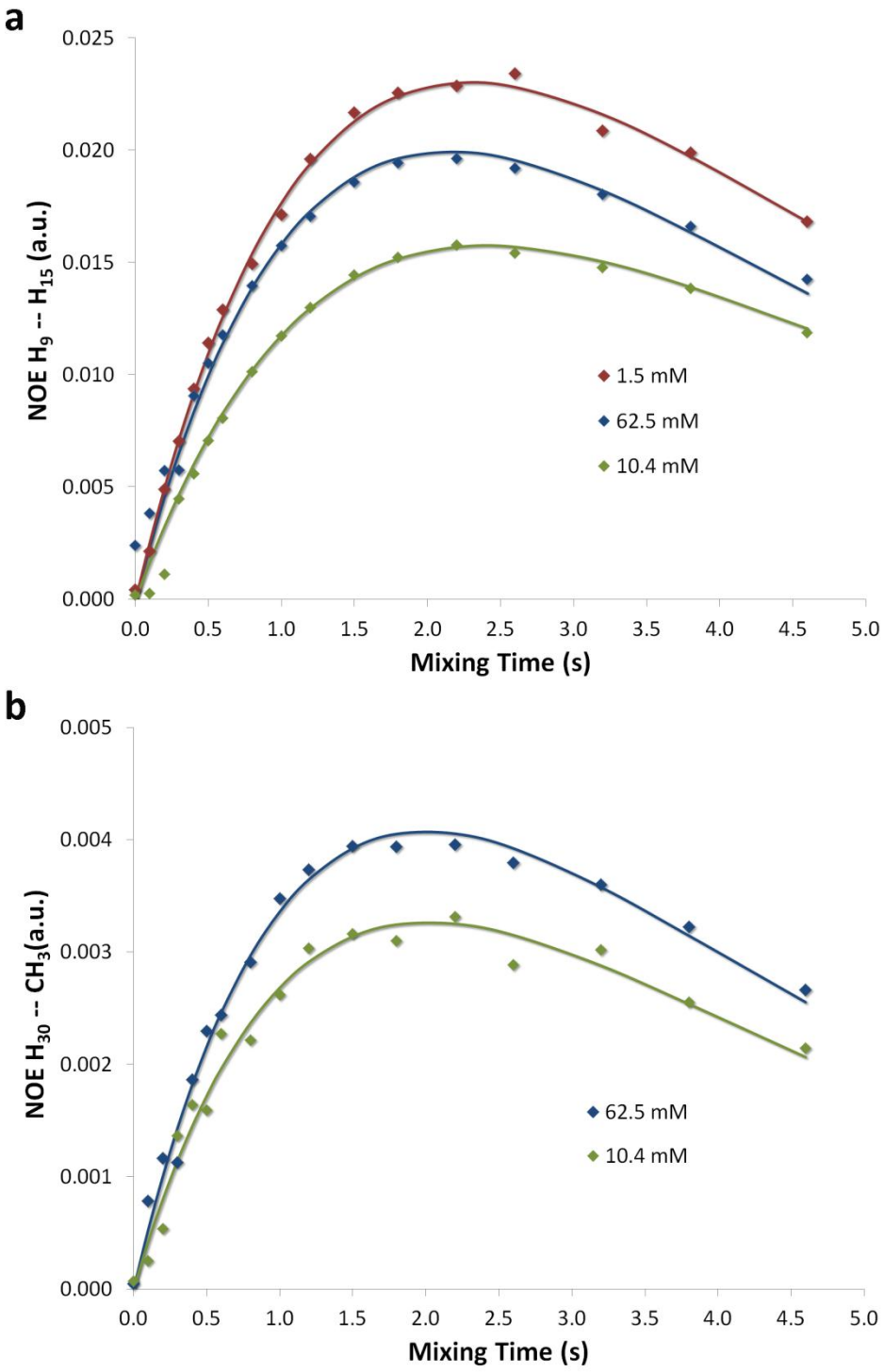

Figure 3.11. NOE build-up curves at different TFA solution concentrations between selected pairs of protons H_9-H_{15} (a) and H_{30}-methyl (b) of TFA as a function of the mixing time. Solid lines represent fits to the data according to Eq. (3.2).

The 2D-NOESY spectra acquired at 25 °C displayed a comprehensive map of close contacts within the molecule (Figure 3.12). All NOE cross-peaks had opposite phases to the diagonal, indicating that these arose from positive NOE enhancements, as expected for a molecule of small size. The 2D-NOESY spectrum of a 62.5 mM TFA solution showed relatively strong correlations between aromatic protons (i.e., H_9-H_{15}, H_9-H_7, H_{30}-H_{15}, and H_{22}-H_{29}) as well as a cross-peak between H_{30} and the methyl protons (Figure 3.12a). The data indicate a favored conformation in which the methyl protons are closer in space to the proton at the anthranilic ring, H_{30}. It is noteworthy that all observed cross-peaks correspond to intramolecular interactions; no additional cross-peaks between aromatic protons of neighboring molecules related to π-π intermolecular interactions are observed (i.e., H_7-H_{30}). The 2D-NOESY spectrum of 1.5 mM solution showed prominent cross-peaks between aromatic protons. Nonetheless, the intramolecular NOE interaction between H_{30} and the methyl protons was significantly weakened (Figure 3.12b). Possible reasons include:

1) the signal-to-noise ratio may be lowered at this low drug concentration;
2) drug molecules, being predominantly in the monomer state, undergo sufficiently rapid rotational motions in solution on the NMR time scale that the distance between H_{30} and the methyl protons cannot remain below 5 Å; and
3) drug molecules preferably adopt a more planar conformation.

To learn whether the observed intramolecular interaction depends on the preferred arrangement of the molecule in solution, 2D-NOESY spectra were acquired at varying concentrations and cross-peak volumes were quantified.

A characteristic peak of TFA Form II (i.e., H_{30}-H_{29}) overlaps with other peaks. As a consequence of the cross-peak H_{30}-methyl being the sole conformational characteristic peak, the outlined NMR studies will only detect the relative ratio with respect to Form I conformer. At different solution concentrations a quantitative interpretation of 2D-NOESY data was accomplished by measuring the cross-peak volumes involving H_{30}-methyl and H_9-H_{15} and the diagonal-peak volumes of H_{30} and H_9. A square region enclosing each pertinent peak was chosen. Spectral noise, and thus the error in peak volume measurements, was estimated by measuring the volume integrals of empty regions of 2D-NOESY spectra and subtracted from the cross- and diagonal-peak volumes. The diagonal-peak volumes were then divided by their values in the solution at the lowest analyzed concentration of TFA. The obtained scaling factor was used to normalize the cross-peak volumes in order to correct for concentration effects. Lastly, the characteristic cross-peak volume corresponding to H_{30}–methyl protons was divided by the reference cross-peak volume for H_9-H_{15} and the ratio was used as indication of the relative ratio of Form I to Form II or of dimer to monomer in solution.

The 2D-NOESY spectra were also acquired at 10 °C (Figure 3.13), 37 °C (Figure 3.14), and 55 °C (Figure 3.15). At 10 °C, the 2D-NOESY spectra of 62.5 and 1.5 mM TFA solutions showed a cross-peak between H_{30} and the methyl protons. However, the specific NOE interaction between H_{30} and the methyl protons was significantly weakened as temperature increased and at 55 °C the interaction was difficult to observe. For each set of temperature cross-peak volumes were quantified as described above.

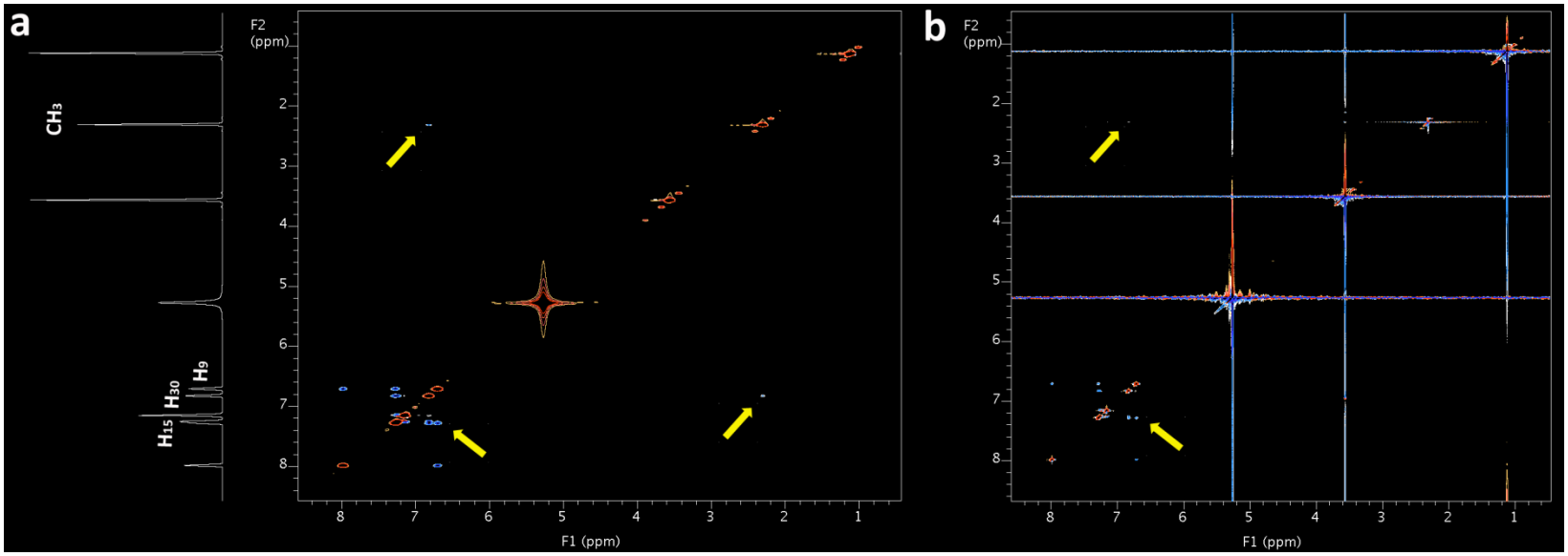

Figure 3.12. 2D-NOESY spectra of 62.5 mM (a) and 1.5 mM (b) TFA in ethanol-d_6 acquired at 25 °C. The interested cross-peak H_{30}-methyl protons and the reference cross-peak H_9-H_{15} are marked by arrows. The corresponding ^{1}H-NMR spectrum is shown at the left side of the 2D-NOESY spectrum.

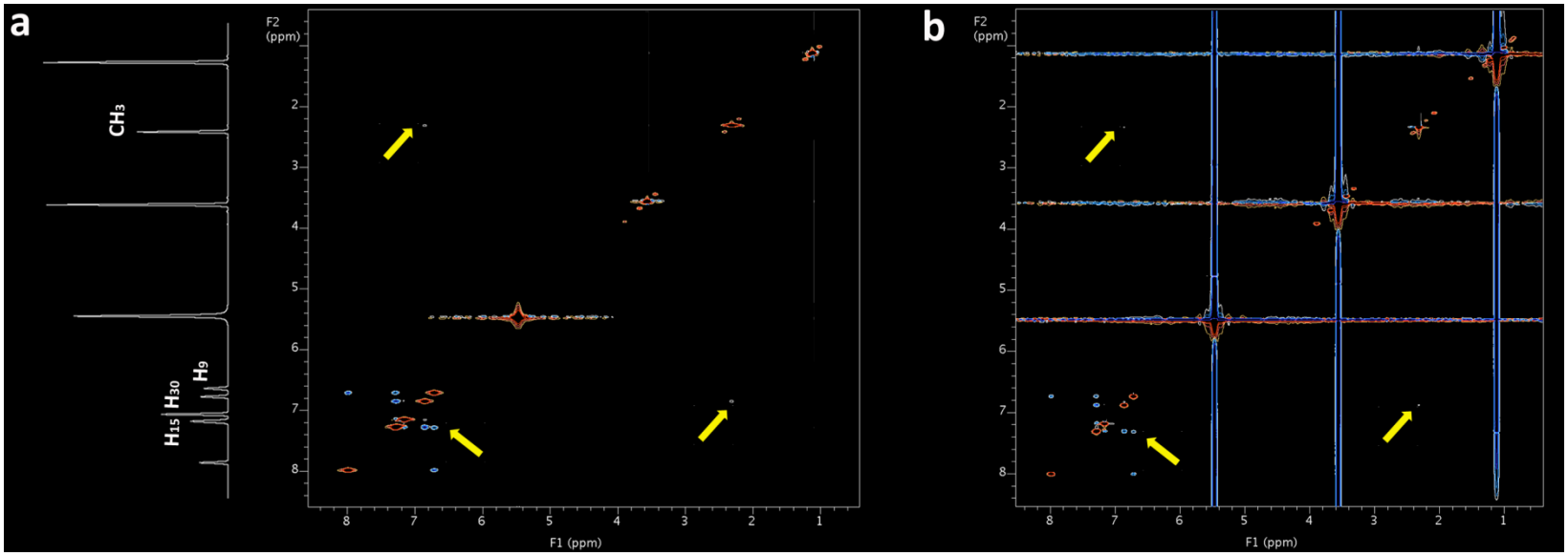

Figure 3.13. 2D-NOESY spectra of 62.5 mM (a) and 1.5 mM (b) TFA in ethanol-d_6 acquired at 10 °C. The interested cross-peak H_{30}-methyl protons and the reference cross-peak H_9-H_{15} are marked by arrows. The corresponding ^{1}H-NMR spectrum is shown at the left side of the 2D-NOESY spectrum.

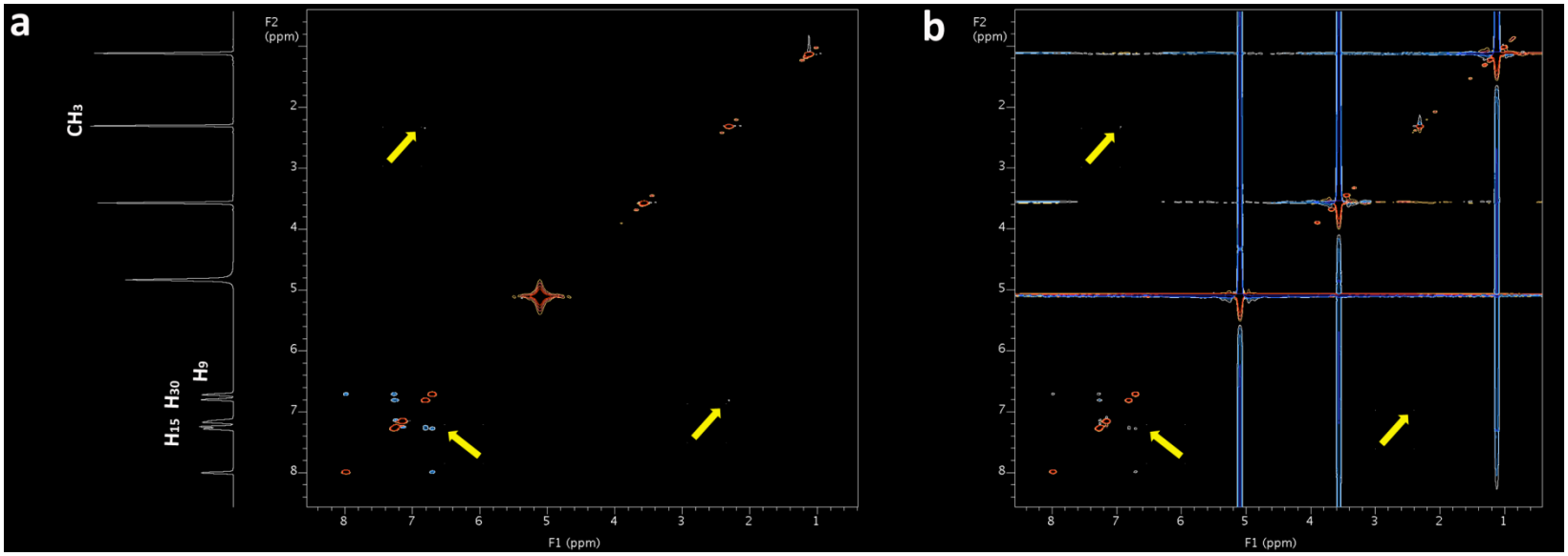

Figure 3.14. 2D-NOESY spectra of 62.5 mM (a) and 1.5 mM (b) TFA in ethanol-d_6 acquired at 37 °C. The interested cross-peak H_{30}-methyl protons and the reference cross-peak H_9-H_{15} are marked by arrows. The corresponding ^{1}H-NMR spectrum is shown at the left side of the 2D-NOESY spectrum.

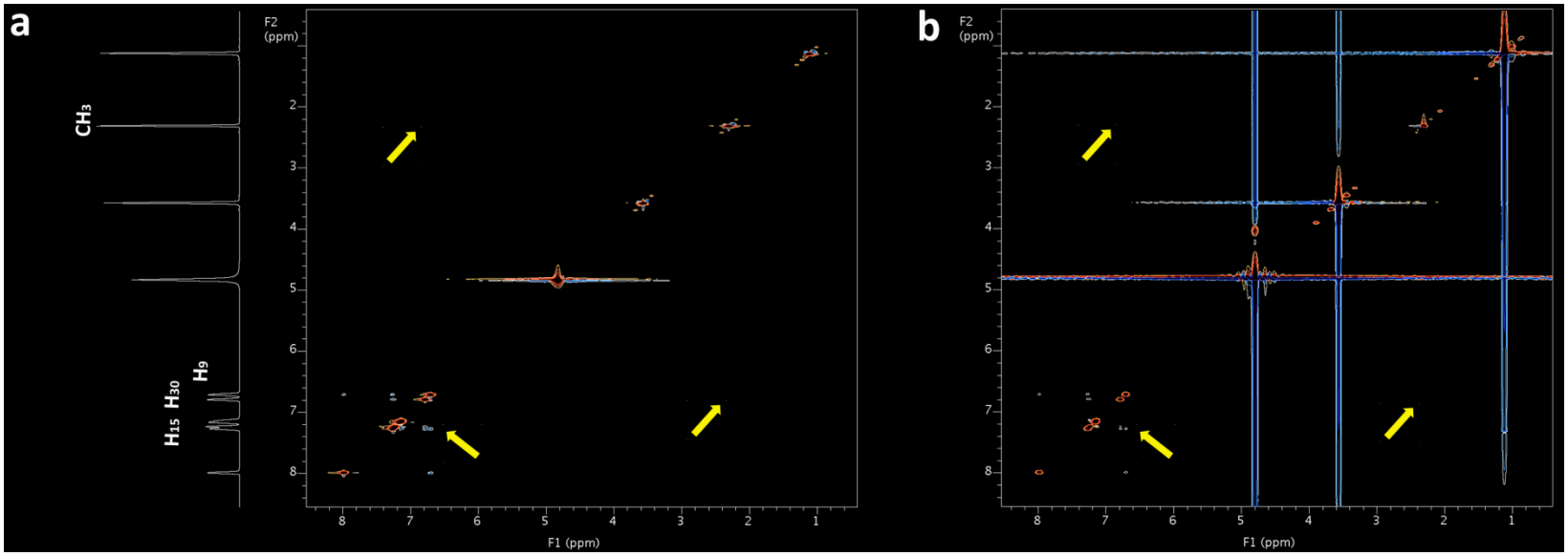

Figure 3.15. 2D-NOESY spectra of 62.5 mM (a) and 1.5 mM (b) TFA in ethanol-d_6 acquired at 55 °C. The interested cross-peak H_{30}-methyl protons and the reference cross-peak H_9-H_{15} are marked by arrows. The corresponding ^{1}H-NMR spectrum is shown at the left side of the 2D-NOESY spectrum.

At 25 °C, when TFA concentration increased, the relative ratio of the Form I conformer or of dimers increased, as indicated by an increase in the cross-peak volume involving H_{30} and the methyl group (Figure 3.16). Similar trends were viewed when the ratio of cross-peak volumes was measured at different temperatures and plotted as a function of TFA concentration. At 10 °C, the relative ratio increased as a function of drug concentration. Also, lowering temperature caused a large variation in favor of the Form I conformer or dimers with respect to the data obtained at 25 °C. At higher temperatures, the cross-peak volume ratio did not change significantly as a function of drug concentration. While there were small differences in the cross-peak volume ratios of the two highest concentrations employed, more prominent trends are observed within each concentration. The results indicate that at low concentration and/or high temperature, molecules prefer to remain monomeric or to adopt a conformation where H_{30} is not in close proximity to the methyl group. When solute concentration increases and/or temperature decreases, molecules form dimers or adopt a conformation where the functional groups (i.e., H_{30} and methyl group) come to close proximity. The twisted Form I conformer is likely to be favored or at least competitive with the Form II conformer.

The conformational behavior is in line with the dimer formation of TFA molecules as a function of concentration and temperature. These experimental observations led us to hypothesize that solution conditions determine self-association of TFA and play a major role in shifting conformational population in favor of the twisted Form I. To assess the significance of solute association versus a possible conformational change, the cross-peak volume ratio versus TFA concentration profiles were fit to a dimerization model according to the equation:

$$V_{obs} = \frac{V_m C_m + V_d 2K_{1,2}(C_m)^2}{C} \quad (3.9)$$

where V_{obs} is the measured cross-peak volume ratio; V_m and V_d are the cross-peak volume ratios of TFA as monomer and dimer, respectively; C is the total TFA concentration; C_m is the molar concentration of the monomer; and $K_{1,2}$ is the self-association equilibrium constant determined from the regression analysis of chemical shift measurements, depending on the temperature. V_m and V_d were used as fitting parameters.

The model fits the experimental data for cross-peak volume profiles of each set of temperature conditions (Figure 3.16 solid lines). Self-association equilibrium constant values, corresponding fitting parameters, and correlation coefficients of the model fits are summarized in Table 3.4. The cross-peak volume ratio of the monomer (V_m) is purely intramolecular whereas the cross-peak volume ratio for the dimer (V_d) can have both intramolecular and intermolecular contributions. The values of V_m range from 0.11 and 0.17 at higher and lower temperatures, respectively. That suggests an H_{30}-methyl distance that is 1.4 to 1.3 times longer than the H_9-H_{15} distance. Specifically, the H_{30}-methyl distance of TFA in the monomer state is 3.50 Å and 3.15 Å at higher and lower temperatures, respectively (Table 3.4). At high temperatures, both small values of the association constants and large values of the cross-peak volume ratio for the dimer (V_d) indicate that TFA molecules exist predominantly in the monomer state. As temperature decreases or concentration increases, TFA molecules form more dimers and the H_{30}-methyl distance becomes shorter. This short H_{30}-methyl distance, and thus, a conformation where the chlorinated aromatic ring is very close to the H_{30} proton at the

anthranilic ring, acquires greater prominence at lower temperature when dimerization also becomes more favorable. Overall, the equilibrium conformer is different at lower temperature than at higher temperature. The ensemble of conformers present at lower temperature either contains more of a conformer with a short H_{30}-methyl distance or has a shorter H_{30}-methyl distance. Note that the molecular conformation within the dimer also seems to be temperature dependent. Because of the low correlation coefficient values, the main limitation of this simple model is that monomer and dimer are assumed to be the predominant species in solution regardless of their respective molecular conformations. Yet the assumption is justified by considering the fast exchange limit between all conformations and interacting molecules, thus resulting in averaged populations. It is difficult to pinpoint the precise molecular nature; however, the salient point is that there exists a correlation between the extent of self-association and solution conformation, wherein a conformer with a shorter H_{30}-methyl distance is present under the solution conditions that favor dimer formation.

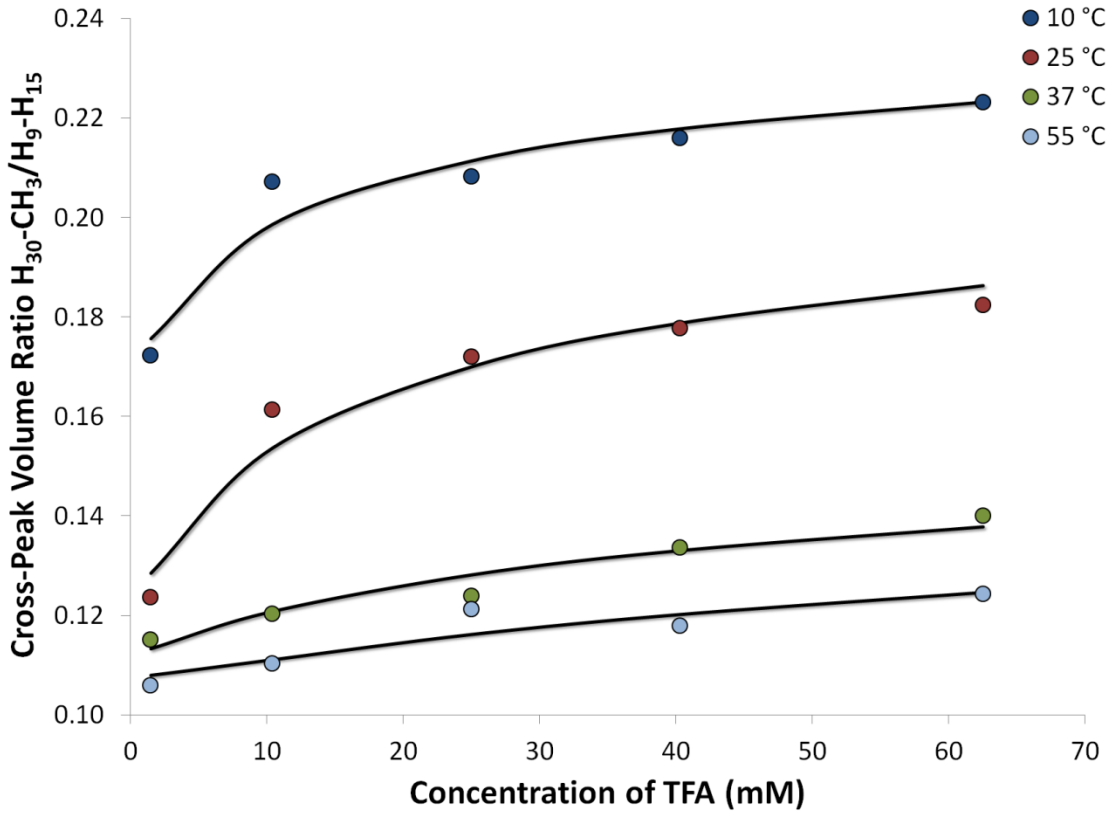

Figure 3.16. Concentration dependence of the ratio between the characteristic cross-peak H_{30}-methyl and the reference cross-peak H_9-H_{15} at different temperatures. The solid lines are fits of the experimental data according to Eq. (3.9) for each set of temperature conditions.

Table 3.4. Self-association equilibrium constants determined from chemical shift measurements at different temperatures, corresponding fitting parameters, calculated distance H_{30}-methyl in the monomer and dimer states, and correlation coefficients obtained by fitting cross-peak volume ratio versus concentration profiles to Eq. (3.9).

T (°C)	$K_{1,2} \pm SD$ (M^{-1})	$V_m \pm SD$	$V_d \pm SD$	r_m (H_{30}-CH_3) (Å)	r_d (H_{30}-CH_3) (Å)	R^2
10	$(0.47 \pm 0.14) \times 10^2$	0.17 ± 0.01	0.25 ± 0.01	3.15	2.96	0.97
25	$(0.31 \pm 0.07) \times 10^2$	0.12 ± 0.01	0.23 ± 0.01	3.35	3.00	0.98
37	$(0.09 \pm 0.05) \times 10^2$	0.11 ± 0.01	0.18 ± 0.01	3.50	3.13	0.97
55	$(0.04 \pm 0.04) \times 10^2$	0.11 ± 0.01	0.18 ± 0.01	3.50	3.13	0.95

3.4 Conclusion

TFA molecules self-associate in ethanol solutions. Two experimental methodologies, such as UV and NMR, were employed to probe the association behavior of TFA. One limitation from UV absorptivity measurements was the difficulty in drawing conclusions from the data. The main reason was the narrow range of concentration that could be analyzed. This disadvantage made investigating the association of solute molecules by NMR a necessity.

NMR chemical shift measurements of the drug in ethanol solutions of various concentrations indicated that solute molecules exist as monomers at low concentration but form more dimers when the concentration increases. The contribution of higher-order associated species was negligible over the concentration range employed. The size of the aggregates obtained from self-diffusion coefficient measurements was consistent with dimer formation when concentration increased. Prominent changes in NMR chemical shifts of H_{30} proton with temperature and the absence of NOE interproton cross-peaks corresponding to aromatic π-π interactions made hydrogen-bonded dimers the preferred mode of molecular association.

NOE intramolecular interactions were utilized for conformational analysis of TFA in solution. Depending upon concentration and temperature, the methyl group is closer to H_{30} proton at the anthranilic ring. The results suggested the existence of interconversion between TFA conformations in solution. Additionally, the higher cross-peak volume between protons of interest, the greater the extent of dimer formation was. Therefore, the study provided evidence for a short H_{30}-methyl intermolecular distance in TFA dimers in

addition to a temperature-dependent shift of the average H_{30}-methyl intramolecular distance in TFA monomers wherein relatively short distances were observed at low temperature.

Chapter 4 – Effect of Molecular Conformation on the Electronic Properties and Intermolecular Interaction

The goal of this chapter is to decipher the intrinsic relationship between conformational flexibility and intermolecular interaction in TFA polymorphs by utilizing electronic calculations. The focus lies in the recognition of the molecular conformation that has the potential to form stronger intermolecular interactions in TFA crystal structures.

4.1 Introduction

Polymorph screening is increasingly assisted by computational techniques to identify all possible physical forms and to select the most suitable polymorph for drug development.[12,143,144] It is well known that the formation of distinct crystal structures from the same molecule is greatly influenced by thermodynamic and kinetic factors, such as type of solvent[97-99] and use of additives.[53,102] Little is known, however, regarding the fundamental mechanism of the self-assembling process of crystallization under various growth conditions. Still, high-performance computation has and will play an important role in advancing the field of organic crystals to allow examination of molecular events and interactions of crystal packing.

In conformational polymorphs of organic molecules, molecular conformation and intermolecular interaction affect each other from an energetic standpoint. A specific crystal structure reflects a delicate energy balance or compromise between intermolecular forces that are responsible for the structural arrangements and molecular conformations that are adopted by molecules in the crystal.[104] It is known that the energy difference

between possible polymorphs of an organic molecule can be 2 kcal/mol or even lower.[19-21] Thus, a small conformational energy change can result in a totally different crystal packing motif.[145] Conversely, a particular intermolecular interaction arrangement, for instance, imposed by some specific solvent environment, may call for a unique conformation in the resultant crystal.[146] Hydrogen bonding is undoubtedly the most critical force in holding organic molecules in the solid state, not only due to its strength but also to its highly directional nature.[147,148] Understanding of the energy relationship between a molecule's conformation and its intermolecular interactions is imperative to the field of crystal growth and the utmost aim of rationalizing polymorph formation and predicting crystal structures.

Density functional theory (DFT) offers an alternative computational method from the traditional *ab initio* wavefunction techniques for calculating molecular energies and properties. It can provide fairly accurate estimates of molecular energies at much lower computational cost. Based on the idea that the electron density is the fundamental quantity for describing atomic and molecular ground states,[149,150] many electronic concepts have been developed for studying chemical reactivity and molecular interaction. The framework and development of these concepts constitute the so-called conceptual density functional theory.[151-153] By calculating and examining how the electronic structure of a molecular system responds to electronic perturbation (e.g., change of the number of electrons in the system), the intrinsic behavior of the molecule interacting with other systems, physically and chemically, can be uncovered. As the theory bridges the gap between physicochemical properties and underlying structural causes, it allows studies of chemical reactivity and molecular interaction from the viewpoint of electron

density and its derivatives. The conceptual DFT is being actively developed and embraced for studying chemical reactivity.[154-158] A few attempts have been made by our group to characterize intermolecular interactions of organic crystals,[159-161] demonstrating great potential for studying molecular packing.

In Chapter 2, crystallization of TFA Forms I and II from solution was shown to be affected by growth conditions. In Chapter 3, self-association and conformation of TFA molecules in solution were extensively studied based on spectroscopic techniques. It seems plausible that TFA molecules adopt a conformation similar to that found in Form I. In this chapter, electronic calculations and DFT analyses of the conformational polymorphism of TFA will be presented. It was important to use a methodology that could either allow for a more robust evaluation or corroborate experimental results, since a molecular-level understanding of the nucleation mechanism of TFA has yet to be reached. A molecule's electronic properties can provide insights into intermolecular interaction and consequent molecular packing in the solid-state. Therefore, our hypothesis was constructed based on the fundamental relationship between molecular conformation and intermolecular interaction. We believe that during the self-association process TFA solute molecules can take on a conformation suitable for the formation of stronger hydrogen-bonded associated species. The results in this chapter will provide support to our hypothesis by utilizing DFT-based concepts and natural bond orbital (NBO) analysis to characterize electronic structures of the molecule and examine how the conformational flexibility impacts crystal packing.

4.2 Theory and Methods

4.2.1 Theoretical Background of Density Functional Theory (DFT)

Within the framework of conceptual DFT, several electronic concepts have been developed for defining a molecule's interacting capabilities with other molecules.[153,162-164] Two pertinent functions utilized in this chapter are the Fukui function and dual descriptor (the calculation details are given in the next section). The electronic Fukui function, $f(\mathrm{r})$, is defined as the change in electron density, $\rho(\mathrm{r})$, with the change in the total number of electrons, N, at constant external potential, $v(\mathrm{r})$:[152,165]

$$f(\mathrm{r}) = \left(\frac{\partial \rho(\mathrm{r})}{\partial N} \right)_{v(\mathrm{r})} \tag{4.1}$$

The external potential is defined by nuclear charges and positions of a given molecular system. Because of the discontinuity of the number of electrons,[166,167] the Fukui function can be evaluated by finite difference. As such, the nucleophilic Fukui function, $f^{+}(\mathrm{r})$, and the electrophilic Fukui function, $f^{-}(\mathrm{r})$, are introduced as:

$$\begin{aligned} f^{+}(\mathrm{r}) &= \rho^{+}(\mathrm{r}) - \rho^{0}(\mathrm{r}) \approx \rho_{\mathrm{LUMO}}(\mathrm{r}) \\ f^{-}(\mathrm{r}) &= \rho^{0}(\mathrm{r}) - \rho^{-}(\mathrm{r}) \approx \rho_{\mathrm{HOMO}}(\mathrm{r}) \end{aligned} \tag{4.2}$$

In these equations, $\rho^{+}(\mathrm{r})$, $\rho^{-}(\mathrm{r})$, and $\rho^{0}(\mathrm{r})$ represent the electron densities of anionic, cationic, and neutral species of a given molecular system, respectively.[152] The Fukui functions can be approximated as the electron densities of the frontier orbitals (LUMO,

the lowest unoccupied molecular orbital, and HOMO, the highest occupied molecular orbital), because the depletion of electrons generally occurs at the HOMO while the addition of electrons occurs at the LUMO. Examples of these functions for a single TFA molecule clearly illustrate the similarity between the HOMO (Figures 4.1a and 4.1g of Form I and II conformers) and $f^{-}(\mathrm{r})$ (Figures 4.1c and 4.1i) and between the LUMO (Figures 4.1b and 4.1h) and $f^{+}(\mathrm{r})$ (Figures 4.1d and 4.1j). Regions that have large Fukui functions are susceptible to electronic perturbation and bear large polarizability. On the other hand, the electron density (Figures 4.1f and 4.1l) seems to be just indicative of the shape of the molecule.

Another DFT concept is the dual descriptor or second-order Fukui function, $f^{(2)}(\mathrm{r})$,[164] which is defined as the second derivative of the electron density with respect to the number of electrons:

$$f^{(2)}(\mathrm{r}) = \left(\frac{\partial^2 \rho(\mathrm{r})}{\partial N^2} \right)_{v(\mathrm{r})} \tag{4.3}$$

The physical meaning of $f^{(2)}(\mathrm{r})$ is made clear by considering the finite difference approximation:[152,168]

$$f^{(2)}(\mathrm{r}) = f^{+}(\mathrm{r}) - f^{-}(\mathrm{r}) \approx \rho_{\mathrm{LUMO}}(\mathrm{r}) - \rho_{\mathrm{HOMO}}(\mathrm{r}) \tag{4.4}$$

It is shown that $f^{(2)}(\mathrm{r})$ is positive at electrophilic regions that prefer to accept electrons and negative at nucleophilic regions that prefer to donate electrons. Figures 4.1e and 4.1k depict dual descriptor isosurfaces, illustrating both electrophilic and nucleophilic regions

of the conformers in Forms I and II, respectively. Consequently, the dual descriptor can be regarded as the electron distribution between the LUMO and the HOMO.

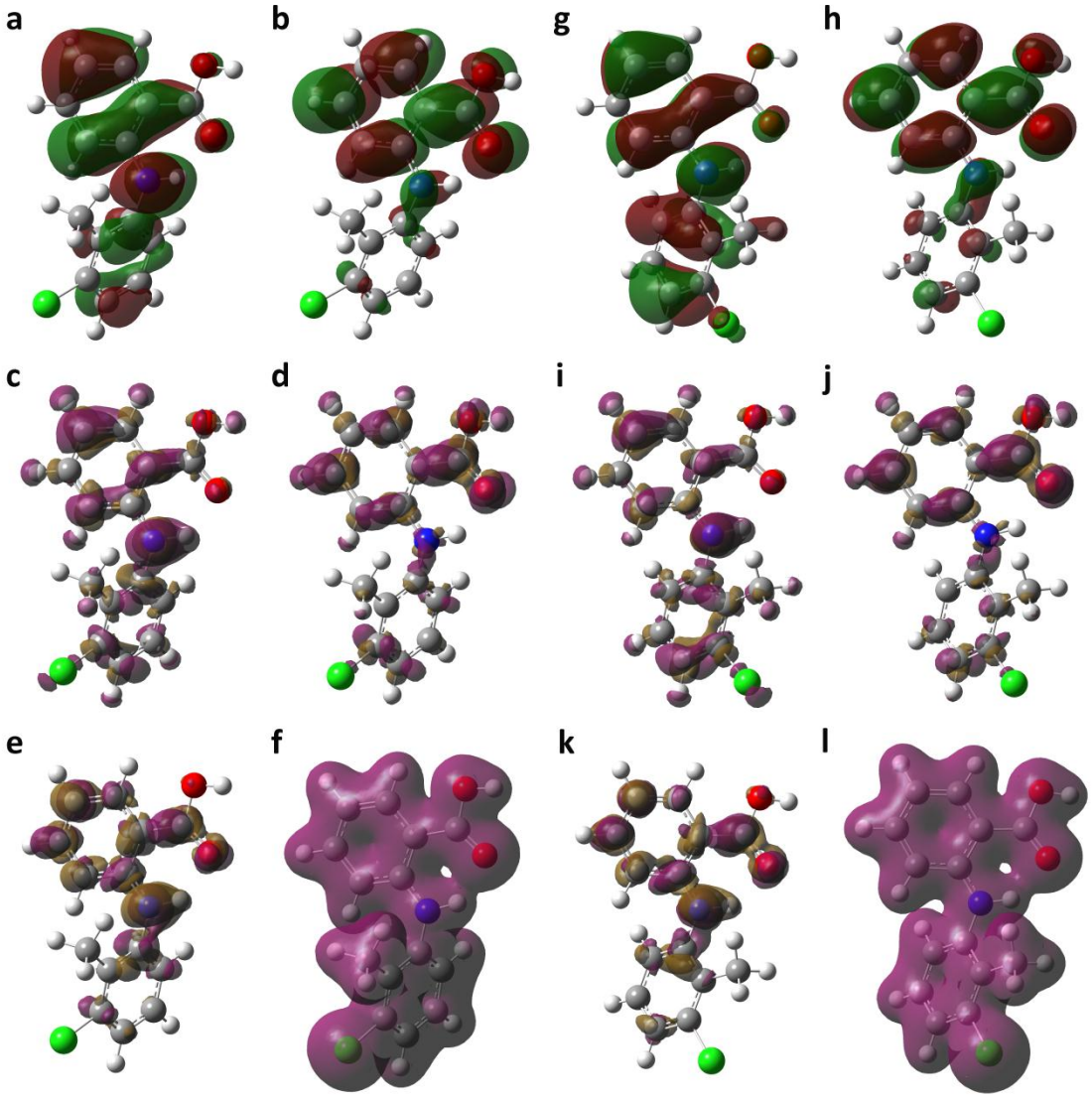

Figure 4.1. Highest occupied molecular orbital (a and g), lowest unoccupied molecular orbital (b and h), electrophilic Fukui function (c and i), nucleophilic Fukui function (d and j), dual descriptor (e and k), and electron density (f and l) isosurfaces of a single TFA molecule taken from Forms I and II.[114] The values of isosurfaces are 0.02 e/bohr^3 for the frontier orbitals and electron density and 0.002 e/bohr^3 for the Fukui functions and dual descriptor, respectively.

Because the Fukui functions and dual descriptor are local functions at every point (r) in space, it is chemically more convenient to examine nucleophilic or electrophilic centers associated with individual atoms. For this purpose, condensed Fukui functions have been proposed,[169,170] analogous to using atomic charges for partitioning the electron density. The condensed Fukui functions can be calculated as:

$$f_\alpha^+ = q_\alpha^0 - q_\alpha^+ = n_\alpha^+ - n_\alpha^0$$

$$f_\alpha^- = q_\alpha^- - q_\alpha^0 = n_\alpha^0 - n_\alpha^- \qquad (4.5)$$

$$f_\alpha^{(2)} = f_\alpha^+ - f_\alpha^-$$

Here, q_α^+, q_α^-, and q_α^0 and n_α^+, n_α^-, and n_α^0 denote the atomic charges and atomic populations on the anionic, cationic, and neutral molecular systems, respectively.[169] Various population analysis schemes, including Mulliken,[171] Hirshfield,[172] and natural bond orbital (NBO),[173,174] can be used to partition or condense electron density into individual atoms. The still popular Mulliken population analysis has a disadvantage in that the results are sensitive to the basis set and that the calculated population can have unphysical negative numbers. The NBO charges prove to be robust in electron population analysis[175] and will be utilized to calculate the DFT properties.

4.2.2 Computational Methodology

4.2.2.1 Conformational Energy

The conformational flexibility of TFA was studied by performing a potential energy scan in the gas phase as a function of the torsion angle, τ_1. Single molecules of TFA extracted from Form I and Form II crystal structures,[114] were fully optimized in order to identify the most stable conformation. The global energy minimum conformer was subsequently used for scanning τ_1 from -180° to +180° with a step size of 2.5°. At each step, all bond lengths, bond angles, and other torsion angles were allowed to be fully optimized. Energy of each conformation at a fixed τ_1 and otherwise fully optimized was then calculated. It was computed in the gas phase as well as in ethanol and tetrachloromethane, which were modeled using the polarizable continuum model (PCM) within the self-consistent reaction field theory.[176,177] The calculations were performed by the B3LYP functional[178] with basis sets of 6-311G(d,p) and 6-311++G(d,p) for the structural optimization and conformational analysis, respectively, using the Gaussian 03 program package.[179]

4.2.2.2 Condensed Fukui Functions

Electronic structure and properties of each τ_1-fixed and otherwise fully optimized conformer were calculated in the gas phase and in solvent media. Condensed Fukui functions were calculated according to Eq. (4.5), based on partial atomic charges by natural bond orbital (NBO) analysis.[174] NBO calculation was also used to obtain donor-acceptor stabilization energies by the second-order perturbation theory.[173,174] The donor-acceptor energy describes quantitatively the electronic bonding-antibonding interaction, providing further intuition about local bonding strengths and intramolecular stability.

Isosurfaces of electronic properties of a single TFA molecule were constructed with GaussView 4.0.[180]

It is worth pointing out that conventional DFT calculation methods cannot fully consider dispersion energy,[181,182] which may account for a considerable portion of intermolecular interactions of organic molecules and crystals.[161,183] Electronic concepts by conceptual DFT, on the other hand, may be better suited for characterizing molecular interactions. In particular, the Fukui function, defined as a second-order derivative of electronic energy, is capable of revealing intermolecular, non-covalent interactions.[161] It is directly related to local polarizability of a molecular system[184] and, more importantly, is associated with interactions due to partially sharing electrons (i.e., hydrogen bonding).

4.2.2.3 Intermolecular Interactions

4.2.2.3.1 Hydrogen Bonding

Hydrogen-bonded dimer motifs exist in the crystal structures of TFA polymorphs. In order to calculate the difference in hydrogen bonding strength between the two Forms, the two dimer configurations were taken directly from the respective crystal structures of TFA polymorphs.[114] Each dimer then underwent constrained optimization with τ_1 and the intermolecular distances between oxygen atoms of the hydrogen-bonded carboxyl groups held constant while all other bond distances, bond angles, and other torsion angles were allowed to be optimized. The optimization was carried out at the level of B3LYP/6-31G(d,p) using the Gaussian 03 package.[179] Root-mean-square (RMS) values of Cartesian coordinates of TFA dimers due to the optimization were insignificant, 0.021 and 0.009 Å for Form I and II dimers, respectively, indicating that the optimization preserved the dimer configurations, particularly the τ_1 angle, from the crystal structures

and corrected minor structural uncertainties, mainly the hydrogen positions that were assigned by single X-ray structure determination.

Intermolecular interaction energies of the optimized cyclic dimers were then calculated in the gas phase at three different levels of B3LYP/6-31+G(d,p), MP2/6-31+G(d,p), and M06-2X/6-31+G(d,p) to ensure that the conclusions were not dependent upon the method. The basis set with diffuse orbitals was selected to improve the description of hydrogen bonding.[185] The basis set superposition error was corrected by the counterpoise method[186] when calculating the intermolecular energies.

The structural optimization of Form I and II dimers was also carried out with only τ_1 held constant while allowing all bond distances, including the intermolecular distances between oxygen atoms of the hydrogen-bonded carboxyl groups, all bond angles, and other torsion angles to adjust. The purpose of this optimization was to verify whether the distances between oxygen atoms could affect intermolecular interaction energy when they were kept fixed during optimization.

In addition, to determine the energy minimum of the hydrogen-bonded dimers the intermolecular distances between oxygen atoms (2.648 Å and 2.644 Å in the crystal structures of Form I and II dimers, respectively) were systematically changed. The geometries for a given system were first optimized with τ_1 and the intermolecular distances between oxygen atoms of the hydrogen-bonded carboxyl groups held constant while all the other parameters were allowed to be optimized. Intermolecular interaction energies of the optimized hydrogen-bonded dimers were then calculated.

4.2.2.3.2 π-π Stacking

Hydrogen-bonded dimers of TFA Forms I and II pack through π-π stacking. In order to calculate the energy of this type of interaction, the geometry of Form I and II dimers interacting via π-π stacking was taken directly from the respective crystal structures of TFA polymorphs.[114] Each dimer underwent optimization of the hydrogen atoms in order to correct minor structural artifacts of the crystal structures. A constrained optimization of the π-π stacking dimers with the main torsion angle fixed could disrupt the parallel arrangement of the molecules in the dimer and, thus, could not reproduce the structure found in the crystal lattice. The optimization was carried out at the level of B3LYP/6-31G(d,p) using the Gaussian 09 package.[187]

Intermolecular interaction energies of the optimized dimers were then calculated in the gas phase at three different levels of B3LYP/6-31+G(d,p), MP2/6-31+G(d,p), and M06-2X/6-31+G(d,p) to evaluate the effects of the computational methods on the calculated energy. Traditional DFT methods, such as the B3LYP functional,[178] perform poorly for π-π stacking interactions. Conversely, a dispersion-corrected DFT method, such as the M06-2X functional,[188] has been shown to give more reliable energy estimations for a variety of systems dominated by dispersion forces, including benzene aggregates. To correct for the basis set superposition error, the counterpoise procedure[186] was applied in the calculation of the intermolecular energies.

4.2.2.3.3 Lattice Energy

Lattice energies of the two TFA polymorphs were calculated with an empirically augmented DFT method.[189,190] A typical DFT method cannot fully consider van der Waals interactions.[181,182,191] To circumvent the limitation, the crystal structures were

firstly optimized with the lattice parameters held constant while allowing fractional coordinates of all atoms to adjust. This approach proved fairly acceptable because of the long-range, weak, and un-directional nature of van der Waals forces that have little impact on fractional coordinates when the lattice parameters are kept fixed during optimization.[189,190] The basis set used was B3LYP/6-21G(d,p); no diffuse orbitals were included due to potential instability introduced to Bloch functions. RMS values of optimized Cartesian coordinates were 0.011 and 0.026 Å for Forms I and II, respectively, demonstrating that the optimization corrected minor structural uncertainties likely induced by hydrogen positions that were assigned by single X-ray structure determination. The single-point energy of optimized crystal structures was then calculated and, according to the optimized atomic coordinates, the van der Waals energy component was included *post priori*. The van der Waals energy was evaluated by a damped analytical model that was based on interatomic distances and empirical parameters.[192,193] The lattice energy was derived as the difference between the total corrected energy of a crystal system and the lowest energy of a fully optimized single molecule in vacuum. A periodical quantum mechanical program, CRYSTAL06,[194] was used for the optimization and energy calculations of TFA crystal structures. Energy convergence for the calculations was set at 10^{-7} Hartree. RMS values of the energy gradient and atomic displacement were set at 0.0001 and 0.0003 atomic units, respectively.

4.3 Results and Discussion

4.3.1 Molecular Conformational Analysis

The molecular conformation in TFA polymorphs differs mainly in the torsion angle τ_1 at the chlorinated aromatic ring portion, not in the anthranilic acid moiety (Figure 4.2). The conformational energy of TFA was plotted against τ_1 (Figure 4.3). The global energy minimum for the single molecule is located at -142°, which corresponds to the conformer in Form II. A shallow minimum at -75° contains the conformer in Form I. As the molecules in TFA polymorphs reside closely to the energy minima (within 2 kJ/mol in gas phase as well as in solvent media), the polymorphism of TFA apparently stems from the conformational variance of the molecule itself. It is worth mentioning that the conformational analysis revealed a very low energy difference between conformers; thus, the three new, recently found polymorphs[115] should not represent a surprise, and the possibility of further polymorphs cannot be ruled out. In fact, the torsion angles, τ_1, of the conformers in these crystals are -138.4° and 126.8° in Form III (Z' = 2), -115.8°, -125.9°, and -134.1° in Form IV (Z' = 3), and -125.1° in Form V; the conformational energy difference among these conformers is again within a few kJ/mol (Figure 4.3). This supports the oft-quoted McCrone argument that the number of polymorphs discovered for a given compound is dependent on the time and effort spent in research on that compound.[10]

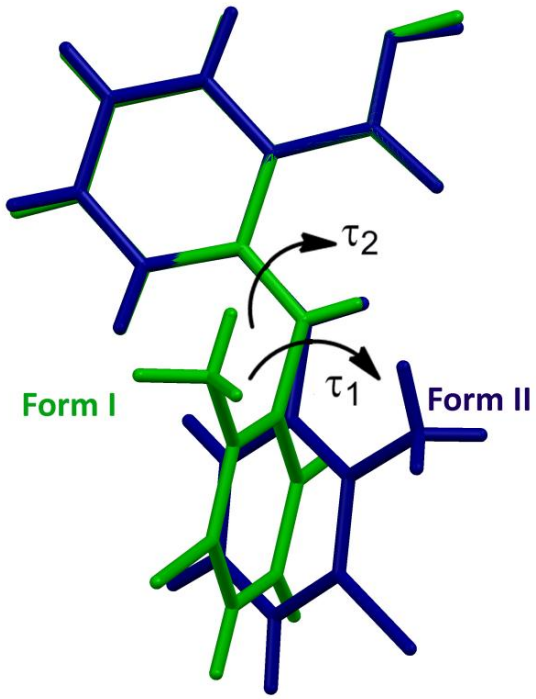

Figure 4.2. Overlay of the two conformers of TFA from Form I, green; Form II, blue. Torsion angles, τ_1 and τ_2, are denoted.

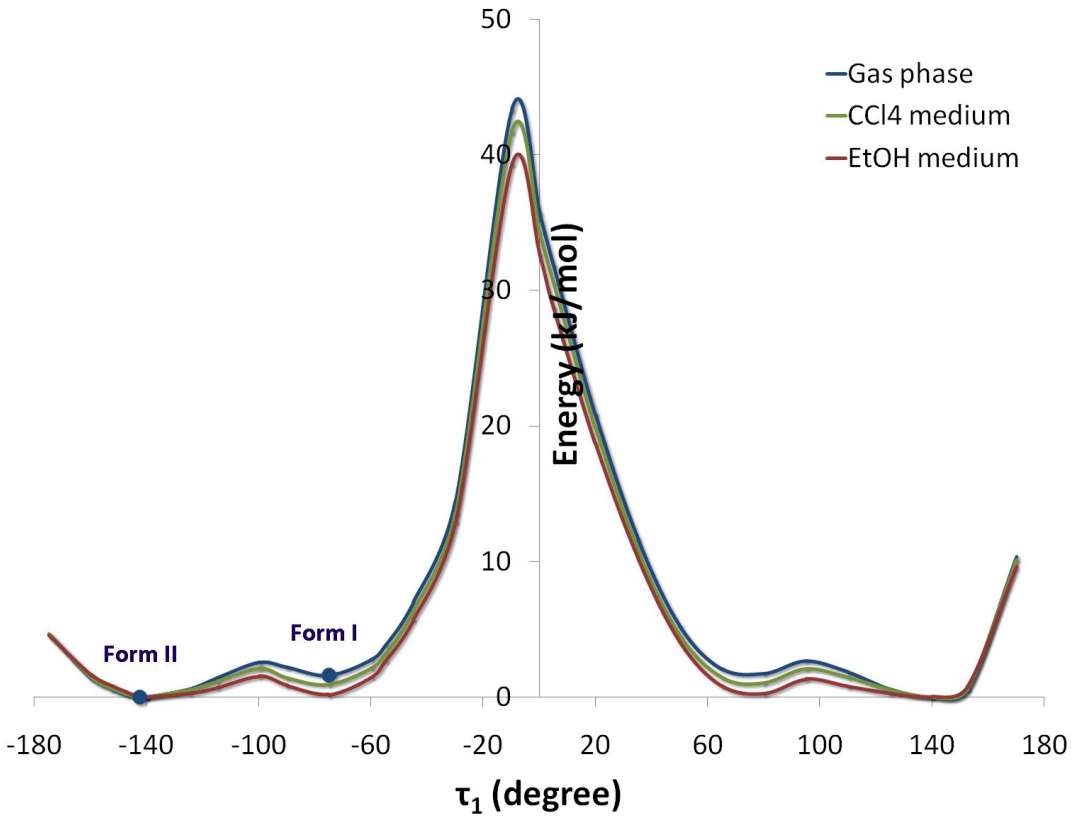

Figure 4.3. Conformational energy of a single TFA molecule as a function of τ_1 in the gas phase, ethanol, and tetrachloromethane. The torsion angles of Forms I and II are marked along the X-axis.

The bonding-antibonding interaction energies were calculated between the lone pair of the amino group and the carboxylated aromatic ring, as well as between the lone pair and the chlorinated aromatic ring. Because of the molecule's symmetry around τ_1, only the energies corresponding to negative values of torsion angle, τ_1, are shown in Figure 4.4. For any conformation, the bonding-antibonding interaction energies between the lone pair and the carboxylated aromatic ring (Figure 4.4a) were significantly stronger than those between the lone pair and the chlorinated aromatic ring (Figure 4.4b). This indicates that the nitrogen atom has a greater tendency to donate electrons to the

carboxylated aromatic ring, thus forming a π-conjugated system in the anthranilic portion. Indeed, the chemical bonds calculated by NBO were $sp^{2.48}$ and $sp^{2.70}$ for the N-C in the anthranilic and N-C in the chlorinated phenyl moieties, respectively. Moreover, the Form II conformer exhibited higher bonding-antibonding interaction energy values between the nitrogen lone pair and chlorinated aromatic ring than Form I (i.e., 117.07 and 79.96 kJ/mol for Forms II and I, respectively). As such, the two aromatic rings bridged by the amino group are almost coplanar in the molecular conformation of Form II (τ_1 = -142.6° and τ_2 = 15.4° in Form II versus τ_1 = -74.9° and τ_2 = 5.4° in Form I). A closer examination of the energy contribution from the π-conjugation as defined by τ_2 (Figure 4.4a) revealed that this stabilizing orbital interaction is determinant in conformations when τ_1 is between -142° and -55°, suggesting that the torsion angle, τ_1, must be in this range in order for the π-conjugation to be retained. The main factor responsible for this behavior is the steric hindrance between the hydrogen at the *meta* position of the carboxylated aromatic ring and the hydrogen or the methyl group at the *ortho* positions of the chlorinated aromatic ring. The steric factor prevents the anthranilic moiety from remaining planar and the extent of non-planarity is illustrated by a systematic variation of τ_2 values as τ_1 changes (Figure 4.5). Significant departure from planarity of the anthranilic moiety occurred with τ_1 lying beyond -142° and -55°, in line with the bonding-antibonding interaction energy (Figure 4.4). Thus, although τ_1 mainly determines the TFA conformational flexibility, strong cooperation is observed between neighboring torsion angles, τ_1 and τ_2. More importantly, the nitrogen lone pair delocalization toward the two aromatic rings accounts for the molecule's conformational distribution.

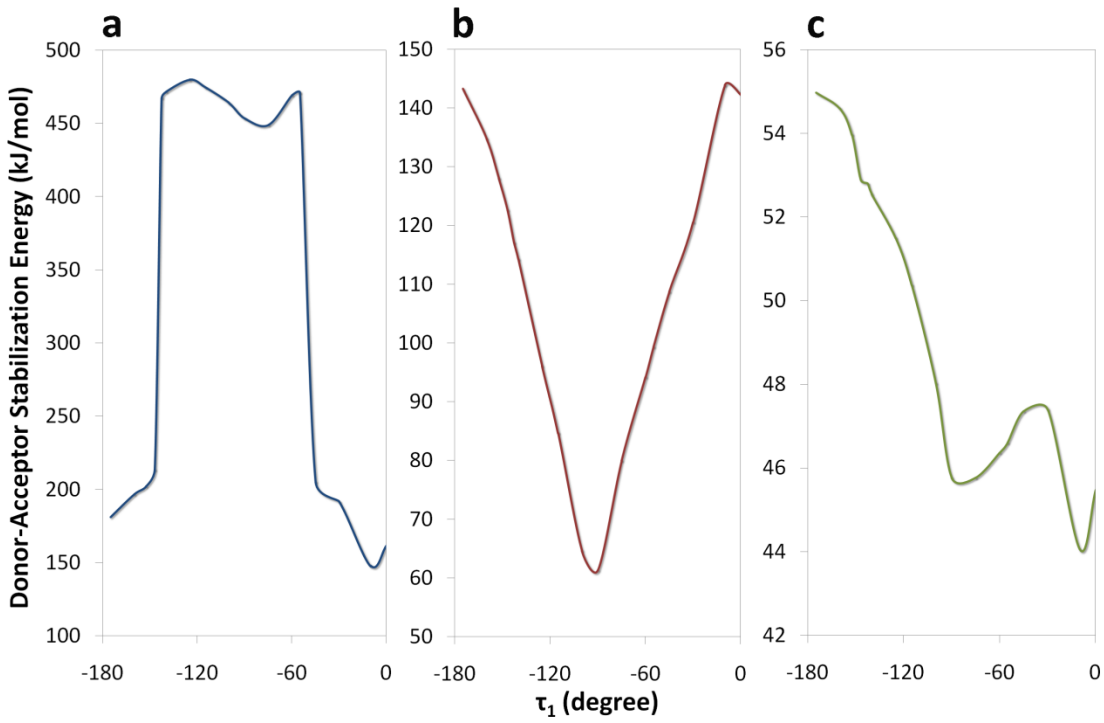

Figure 4.4. Donor-acceptor stabilization or bonding-antibonding interaction energies as a function of τ_1 between the nitrogen lone pair and the carboxylated aromatic ring (a), between the nitrogen lone pair and the chlorinated aromatic ring (b), and between the carbonyl oxygen lone pairs and the amino group (c).

The TFA molecule also forms an intramolecular hydrogen bond between the hydrogen bonding donor amino group and the acceptor carbonyl oxygen. To examine the extent to which this feature influences molecular conformation and structural stability, the bonding-antibonding interaction energies between the carbonyl oxygen lone pairs and the amino group were plotted as a function of τ_1 (Figure 4.4c). The values were 45.77 and 52.80 kJ/mol for Forms I and II, respectively. The larger intramolecular hydrogen bonding contribution in the Form II conformer indicates that its anthranilic acid portion is more constrained by the intramolecular hydrogen bonding, contributing to the coplanarity of the whole molecule. This further implies that the TFA molecule is more inclined to

adopt a planar conformation; such preference arises primarily from the conjugation or delocalization of the lone pair on the amino group to bridge the two aromatic rings and, in turn, forms a stronger intramolecular hydrogen bonding.

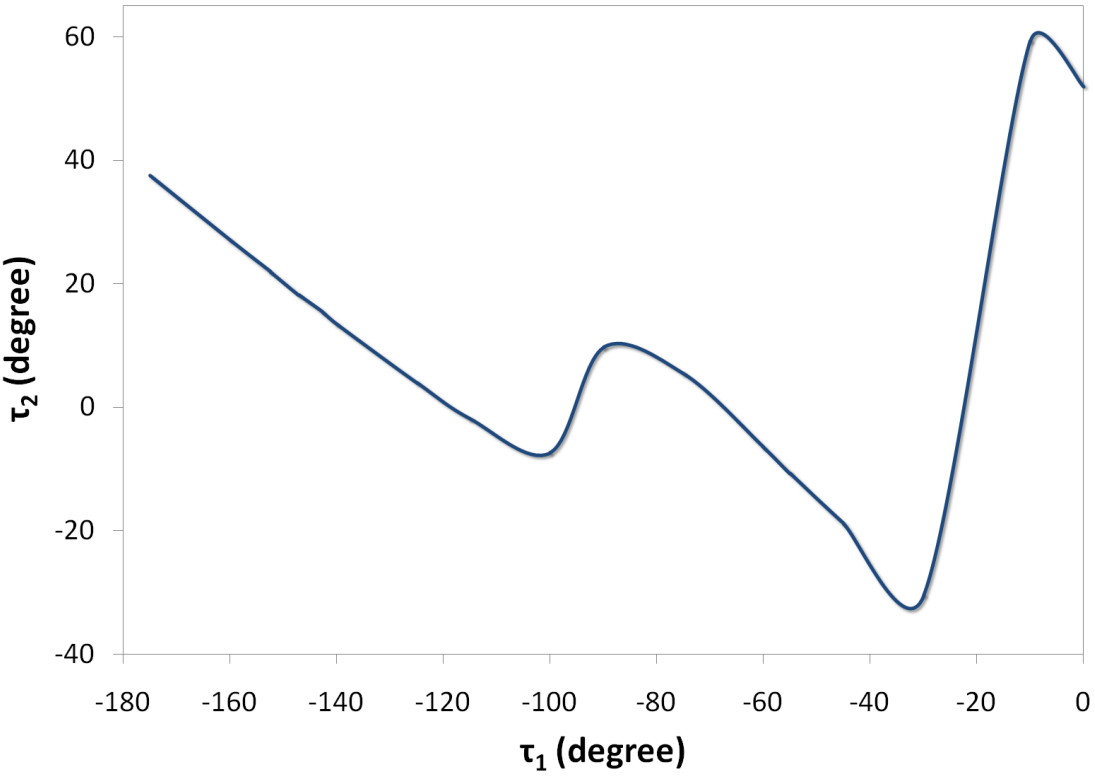

Figure 4.5. τ_2 as a function of τ_1 in respective τ_1–fixed conformers that are fully optimized.

Intermolecular interaction competes with intramolecular hydrogen bonding, resulting in variation in conformation. To study the deterministic effect by the molecule's conformation on the intermolecular interaction strength, condensed Fukui functions were calculated. Dual descriptor magnitudes as a function of the torsion angle, τ_1, were obtained for the carbonyl oxygen and the hydroxyl oxygen atoms (Figure 4.6). In the gas

phase, the dual descriptors of the carbonyl oxygen were found to be 0.060 e in Form I and 0.078 e in Form II, meaning that the atom in Form I is less electrophilic and offers a greater capacity to donate electrons to an electron-deprived hydrogen forming a stronger hydrogen bond (Figure 4.6a). Conversely, being part of a π-conjugated system and being further stabilized by intramolecular hydrogen bonding, the carbonyl group in Form II becomes a relatively poor hydrogen bonding acceptor. The positive value of the dual descriptor of the carbonyl oxygen stems from the local dominance of the LUMO (i.e., lowest unoccupied molecular orbital) over the HOMO (i.e., highest occupied molecular orbital). In fact, the LUMO was mainly delocalized on the anthranilic moiety while the HOMO spanned almost the whole molecule, but less significantly the carboxyl group (Figure 4.1).

In addition, the carbonyl oxygen had the lowest dual descriptor value (i.e., highest capability to donate its valence electrons) when $\tau_1 = \pm 90°$, a position that completely disrupted the π-conjugation between the amino group and the chlorinated aromatic ring. Dual descriptor isosurfaces further clarify the assessment (Figure 4.7). In agreement with the condensed values, a more pronounced nucleophilic region was shown around the carbonyl oxygen of the molecular conformation when $\tau_1 = -90°$ (Figure 4.7b). This region thereby contributes considerably to the intermolecular interactions. In comparison, a larger distribution of the electrophilic nature of dual descriptor enclosed the hydrogen bonding moiety of the conformation when $\tau_1 = -10°$, showing the lowest susceptibility to electron donation (Figure 4.7d). The two conformers in Forms I and II displayed dual descriptors in-between ($\tau_1 = -142.6°$ or $-74.9°$; Figures 4.7a and 4.7c). Therefore, the conformational flexibility around τ_1, the π-conjugation defined by τ_2 (i.e., the planarity of

the anthranilic moiety), and the electronic distribution determine in cohort the most favorable conformers that can strengthen intermolecular contacts.

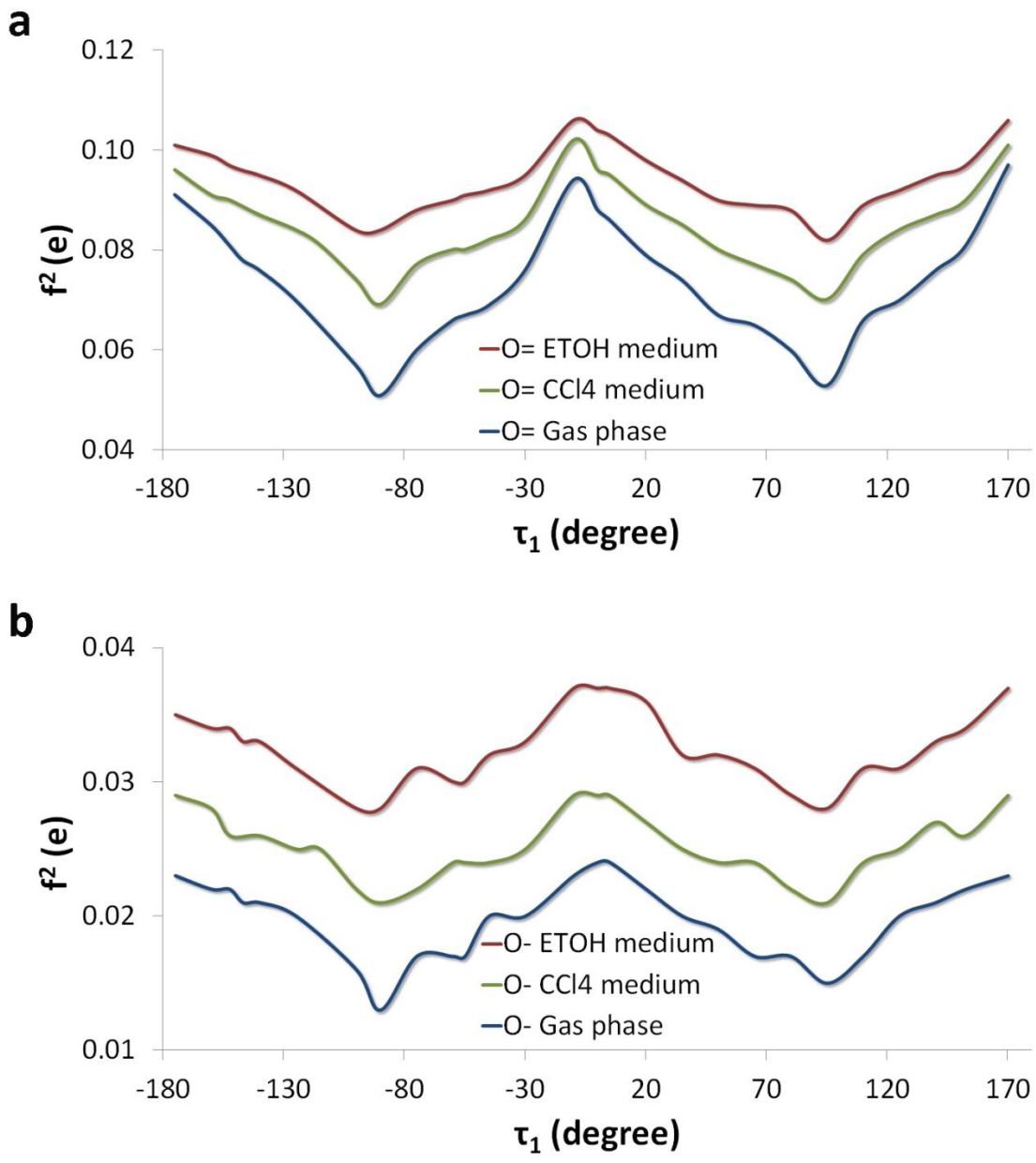

Figure 4.6. Condensed dual descriptors of the carbonyl oxygen (a) and hydroxyl oxygen (b) of a single TFA molecule as a function of τ_1. The values of each conformation were calculated in the gas phase, ethanol, and tetrachloromethane, respectively.

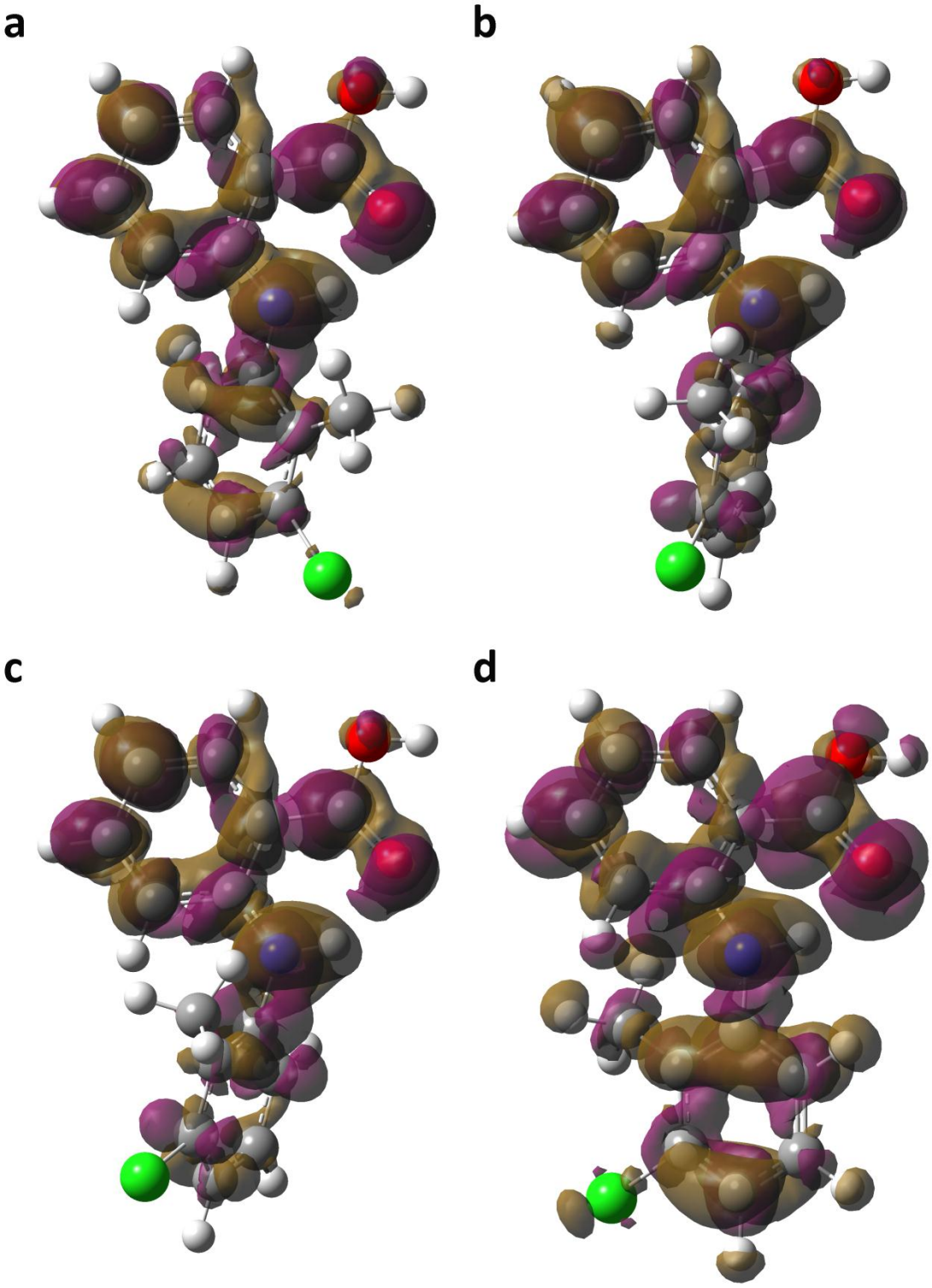

Figure 4.7. Dual descriptor isosurfaces of fully optimized conformers when τ_1 is kept at -142.6° (a), -90° (b), -74.9° (c), and -10° (d). Electrophilic regions are shown in pink and the nucleophilic regions in brown. The values of isosurfaces are 0.001 e/bohr^3.

To further understand the solvent effect on the intermolecular hydrogen bonding, condensed Fukui functions were calculated in two solvent media of different polarities, namely, ethanol and tetrachloromethane. As expected, the dual descriptor values of the carbonyl oxygen increased in the solution phase as compared with the gas phase (Figure 4.6a), indicating that its electron-donating capability becomes weakened upon solvation, specifically in a polar medium. Similarly, the dual descriptors of the hydroxyl oxygen in the gas phase were 0.017 e in Form I and 0.021 e in Form II; they increased by 0.5 percent and 1 percent when computed in tetrachloromethane and in ethanol, respectively (Figure 4.6b). Note that the solvent effect was evaluated implicitly. Nonetheless, the general conclusion regarding the intrinsic connection between the molecule's conformation and intermolecular hydrogen bonding strength holds.

4.3.2 Intermolecular Interaction Energy in the Gas Phase

4.3.2.1 Hydrogen Bonding Interaction

Both crystal structures of TFA polymorphs consist of cyclic, hydrogen-bonded dimers. The intermolecular hydrogen bonding energy of Form I and II dimers, composed of monomers A and B, was computed in the gas phase as follows:

$$E_{\text{hydrogen bonding}} = E_{\text{dimer AB, gas phase}} - (E_{\text{monoomer A, gas phase}} + E_{\text{monomer B, gas phase}}) \quad (4.6)$$

where $E_{\text{hydrogen bonding}}$ is the energy of hydrogen bonding interactions; E_{dimer} is the total energy of the cyclic dimer; $E_{\text{monomer A}}$ and $E_{\text{monomer B}}$ are the individual energies of the monomers. The monomer energies were computed in the basis set of the dimer so that the results were corrected for the mathematical artifact called basis set superposition error.

When performing optimization with the intermolecular distances between oxygen atoms of the hydrogen-bonded carboxyl groups held constant, the hydrogen bonding energy of the Form I dimer calculated by B3LYP was -37.65 kJ/mol per hydrogen bond (or -75.31 kJ/mol for two hydrogen bonds in a dimer), while that of the Form II dimer was -34.60 kJ/mol. Hydrogen bonding energy values calculated by MP2 were -33.72 kJ/mol and -31.08 kJ/mol of Form I and Form II dimers, respectively. M06-2X calculations yielded energy values of -41.15 kJ/mol and -38.19 kJ/mol for Form I and Form II dimers, respectively. Results for the methods tested are summarized in Table 4.1. All of the methods indicate that the dimer of Form I is stronger, by about 3 kJ/mol, than that of Form II. Increased magnitudes of hydrogen bonds calculated by the M06-2X method can be attributed mainly to a substantial improvement of the description of dispersion effects. Nonetheless, the calculations suggest that the conventional DFT method provides fairly reliable estimates of the hydrogen bonding energy. Studies[195-197] employing DFT as a means to study hydrogen-bonded complexes also concluded that the method performs reasonably well in describing hydrogen bonding interactions when compared to *ab initio* wavefunction-based methods and experimental data. For that reason, the DFT methodology will be applied to the characterization of intermolecular interactions in the solution phase that will be described in Chapter 5.

Table 4.1. Hydrogen bonding energy values of Form I and Form II dimers calculated in the gas phase at different levels of theory.

Dimer	$E_{\text{Hydrogen bond, B3LYP}}$ (kJ/mol)	$E_{\text{Hydrogen bond, MP2}}$ (kJ/mol)	$E_{\text{Hydrogen bond, M06-2X}}$ (kJ/mol)
Form I	-37.65	-33.72	-41.15
Form II	-34.60	-31.08	-38.19

When the intermolecular distance was optimized, the hydrogen bonding energy values calculated by B3LYP were -37.73 kJ/mol and -34.94 kJ/mol of Form I and Form II dimers, respectively. The results agree well with those obtained by using the intermolecular distance from the corresponding crystal structures at the same level of theory. This implies that the optimization strategy was able to produce reliable distances between oxygen atoms, close to those in the crystal structures of Forms I and II (i.e., 2.643 Å and 2.650 Å in the optimized Form I and II dimers as compared to 2.648 Å and 2.644 Å in the crystal structures of Form I and II dimers, respectively). Furthermore, the hydrogen bonding energy was examined as a function of the intermolecular distance between oxygen atoms of the cyclic dimers (Figure 4.8). When the intermolecular distance between oxygen atoms decreased, the intermolecular interaction got strengthened and the hydrogen bonding energy decreased. The stability of the cyclic hydrogen-bonded dimers increased with shorter intermolecular distances and the energy of Form I and II dimers reached a minimum located at the same distance as that found in the respective crystal structures (data superimposed in Figure 4.8). This implies that the hydrogen-bonded dimers in the crystal structures correspond to the most stabilizing arrangements. When two molecules are sufficiently close to each other, repulsive

interactions take place accounting for the increase in energy of the system, despite the increasing hydrogen bonding strength.

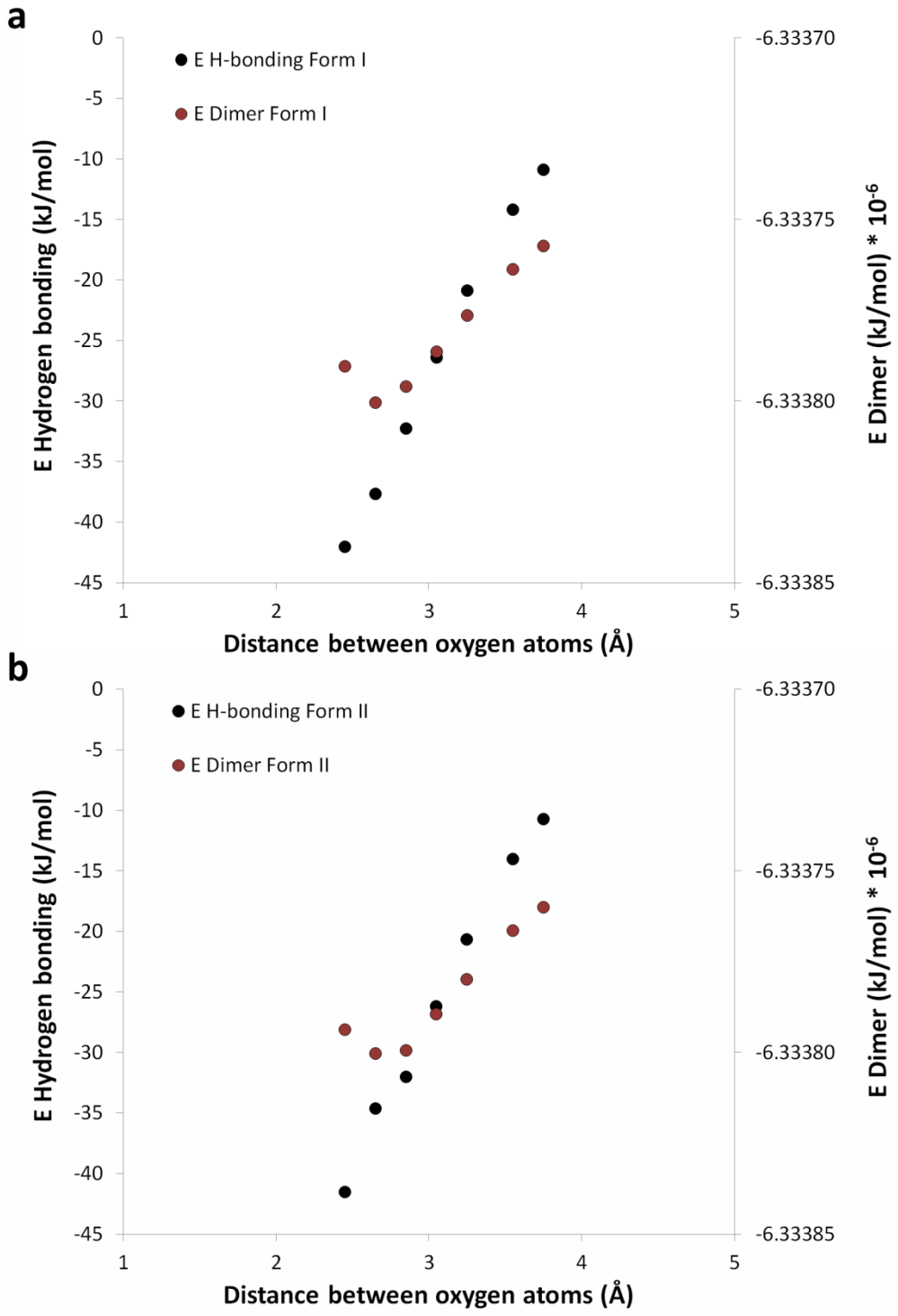

Figure 4.8. Hydrogen bonding energy of Form I (a) and Form II (b) dimers as a function of the intermolecular distance between oxygen atoms of the carboxyl groups. The stability energy of Form I (a) and Form II (b) dimers is also shown versus the intermolecular distance. Energies were calculated at the level of B3LYP/6-31+G(d,p) after constrained optimization.

4.3.2.2 π-π Stacking Interaction

The crystal structure analysis of TFA polymorphs reveals π-π stacking interactions between aromatic rings of neighboring molecules.[114,115] The distances between aromatic rings are 4.826 Å and 3.836 Å in the crystal structures of Form I and II dimers, respectively. As such, energetic quantification of these interactions was important for comparison with the strength of hydrogen bonding. Interaction energies of Form I and II dimers calculated by B3LYP[178] and MP2[198] methods were positive (i.e., destabilizing). The positive numbers can be due to the limitation of B3LYP and MP2 methods in describing dispersion interactions.[181,182,199,200] Indeed, π-π stacking interactions largely depend on dispersion forces. The M06-2X method may be better suited for computing interaction energies dominated by dispersion forces, although the method is empirical and may overestimate the energy values.[188] The π-π stacking energy of the Form I dimer, calculated by M06-2X, was -16.69 kJ/mol while that of the Form II dimer was -23.80 kJ/mol. The ability of the M06-2X functional to consider dispersion forces allows for an adequate description of π-π stacking interactions between TFA molecules. The results indicate that π-π stacking in the Form II dimer is stronger, by about 7 kJ/mol, than that in the Form I dimer. Still, the energies of π-π stacking are lower in absolute value than those of one hydrogen bond, suggesting that intermolecular hydrogen bonding of TFA dimers is a more relevant contribution compared to π-π stacking. The π-π interaction energy values for the methods tested are summarized in Table 4.2.

Table 4.2. π-π stacking energy values of Form I and Form II dimers calculated in the gas phase at different levels of theory.

Dimer	$E_{\pi\text{-}\pi \text{ Stacking, B3LYP}}$ (kJ/mol)	$E_{\pi\text{-}\pi \text{ Stacking, MP2}}$ (kJ/mol)	$E_{\pi\text{-}\pi \text{ Stacking, M06-2X}}$ (kJ/mol)
Form I	5.03	5.14	-16.69
Form II	5.91	5.17	-23.80

4.3.2.3 Lattice Energy

The lattice energies of Forms I and II were computed to be -137.50 and -135.72 kJ/mol, respectively, supporting the fact that Form I is more stable than Form II. The calculation contradicts the claim that Form II is the stable low temperature polymorph.[114] The calculated lattice energy of Form I is very close to a reported experimental value of the sublimation enthalpy of Form I, 128.4 kJ/mol at 298 K,[201] which yields a derived lattice energy value of -133.4 kJ/mol (see reference 189 for the derivation).

4.3.3 Relationship between Molecular Conformation and Hydrogen Bonding Interaction

The electronic origin of the energy difference between hydrogen-bonded dimers of TFA polymorphs was explored. TFA molecular conformation does have a significant influence on the molecule's electronic structure and, thus, on its intermolecular interacting capabilities. The conformational polymorphism of TFA results from the energy competition between intramolecular π-conjugation and intermolecular hydrogen bonding. By adopting a conformation that is not the most energetically stable, TFA molecules can strengthen intermolecular hydrogen bonding interactions in the crystal. The most stable

conformation in the TFA monomer is almost planar. When solute molecules form hydrogen-bonded dimers, a conformation with reduced delocalization of electrons is favored and a torsional strain in the molecule allows for stronger intermolecular interactions. When considering the conformational variation from the flatter conformer in Form II to the twisted conformer in Form I, the torsional strain raises the energy barrier by about 2 kJ/mol (Figure 4.3 as τ_1 = -74.9° in Form I and -142.6° in Form II). The energy penalty for the rearrangement of the molecular conformation and for the consequent loss of π-conjugation is recovered by stronger hydrogen bonding (*ca.* 3 kJ/mol calculated in gas phase by quantum mechanics). It is thereby reasonable to believe that a unique molecular conformation is a compromise in energy between π-conjugation and intermolecular interaction strength. Understanding molecular interaction and crystal packing necessitates electronic structure calculation and analysis, which can be further facilitated by utilizing DFT and NBO concepts.

4.4 Conclusion

Intermolecular hydrogen bonding interactions are stronger than π-π interactions. The conformational polymorphism of TFA is caused by the competition between delocalization of intramolecular π systems and enforcement of intermolecular hydrogen bonding. From an energetic standpoint, the TFA molecule prefers to remain relatively flat so as to maximize the molecular stability. On the other hand, due to its flexibility and low conformational energy barrier, the TFA molecule can assume a different conformation. The conformational re-arrangement is favored kinetically by a low energy barrier and is driven thermodynamically by the formation of stronger hydrogen bonding upon

dimerization. The impact of the conformational flexibility of TFA on intermolecular interactions is ultimately reflected in its crystal structures. By addressing the molecular and mechanistic aspects underlying the relationship between molecular conformation and intermolecular interaction, DFT concepts can yield considerable insight into the polymorph behaviors of this system.

Sections of this chapter were previously published.[202] Reprinted with kind permission from Elsevier: International Journal of Pharmaceutics, Mattei, A, Li T 2011, Interplay between molecular conformation and intermolecular interactions in conformational polymorphism: A molecular perspective from electronic calculations of tolfenamic acid, 418(2): 179-186. Copyright © 2011 Elsevier.

Chapter 5 – Energy Evaluation in an Explicit Solvent Model

The goal of this chapter is to apply the knowledge obtained in Chapter 4 on the electronic structure of each TFA conformer and intermolecular interaction in the gas phase to determine whether intermolecular interaction strength changes with solvent. The solvent is an important factor in understanding the mechanism of polymorph formation. Therefore, a method to explicitly model solvent molecules will be developed; energy of solute molecules and their intermolecular interaction in solution also will be calculated.

5.1 Introduction

It is established that solvents can significantly affect polymorphism of organic crystals.[97,98] Both the dynamic and thermodynamic properties of solute species are strongly influenced by the microscopic structure and organization of solvent molecules in the surrounding solution. A key to a detailed understanding of the mechanistic role played by the solvent in polymorph formation lies in the ability to model the interaction between solute and solvent molecules. Although the most practical approach for treating solvation effects is often experience combined with trial and error, computational methods can be used to consider solvent effects and help rationalize the experimental procedures.

In recent years, there has been progress in developing techniques for including solvation effects in quantum mechanical calculations. These techniques differ in the level of detail used to describe molecules in solution.[203-205] One possibility to include solvation effects in quantum mechanics calculations and thereby study molecules in solution is based on

continuum solvation models.[206] Rather than describing each solvent molecule explicitly, the solvent is modeled as a continuous medium, characterized solely by a scalar, static dielectric constant. The solute molecule creates and occupies a cavity within this medium. The electrostatic interaction of a solute with the solvent depends upon the charge distribution and polarizability of the solute. The charge distribution of the solute molecule induces polarization of the medium and the resulting induced field in turn polarizes the solute molecule.

The major problem with this continuum model is the inability to properly account for specific solute-solvent interactions, such as hydrogen bonding.[207] The model can only empirically determine solvation free energies by using parameterized dielectric constants and molecular cavity sizes. Also, the properties of the solvent, including its dielectric constant and structure, are not the same in the immediate vicinity of the solute molecule as they are in the bulk.

From a microscopic standpoint, solvation involves the formation of interactions between a solute and a solvent as well as a change in the interactions of the solvent molecules in the vicinity of the solute. Thus, a key to understanding solvation could be the determination of the structure adopted by solvent molecules around the solute. One means of elucidating the solvent shell structure, as well as of minimizing the limitations mentioned above, is to describe the solvent as discrete molecules, with which the interactions are treated explicitly. The ideal description of a solution would be a quantum mechanical treatment of the system, consisting of representative numbers of solute and solvent molecules. However, in practice, the interaction of a solute molecule with hundreds or even thousands of solvent molecules puts a considerable demand upon

computer resources. Molecular mechanical force fields are traditionally used in molecular dynamics simulations of organic and biological systems.[208] They can provide valuable information on molecular structure and thermodynamic properties in solution, such as solvation free energy, which is obtained from the averages of all molecular positions generated during the simulation. In addition, molecular mechanics methods are computationally efficient. Yet changes in the electronic structure of a system require quantum mechanical treatment. Alternatively, the quantum mechanics (QM) and molecular mechanics (MM) could be combined so that the solute molecule, along with the first (or first two) solvation shell is treated by a QM model and the rest of solvent environment is represented by MM force fields.[203] This makes it possible to deal more accurately with hydrogen bonding.

Despite the wide acceptance of the quantum mechanics/molecular mechanics (QM/MM) method, a standard protocol has not been defined yet. There is *a priori* no limit to the level of theory that can be utilized for the treatment of the QM region. Thus, implementations of the QM/MM methods within the framework of *ab initio*[209] or density functional[210] theories have been reported. In this work, density functional theory (DFT) will be used because it has the advantage of including electron correlation effects and because it was utilized to calculate intermolecular interaction energies in the gas phase.

The QM/MM approach was designed with two aims in mind: 1) to get a reliable energetic description of the conformational preference of the TFA molecule in solution and 2) to evaluate intermolecular interactions by taking into account the solvating environment. To accomplish these aims, the first step taken was to perform molecular dynamics simulations, which allowed for the minimization and sampling of the geometries of the

solute molecule necessary for the determination of the energies by QM/MM. When designing the method, it was realized that different specifications, such as charge distribution on molecular conformation, sampling, and size of the discrete QM region, needed to be defined. These critical factors will be evaluated together with the QM/MM capabilities and/or limitations.

5.2 Theory and Methods

5.2.1 Theoretical Background

Before presenting computational details of the calculations, this section will provide a concise review on the combined QM/MM. In classical mechanics, the Hamiltonian function ($\hat{H}$) is used to describe the total energy of the system. In an analogous manner, to calculate the energies and forces on each of the atoms in the QM/MM system, $\hat{H}$ is constructed and the time-independent Schrödinger equation is solved for the energy of the system (E) and the wavefunction for all the nuclei and electrons of the QM region (Ψ), according to the equation:

$$\hat{H}\,\Psi(r,R_{\alpha},R_{M}) = E(R_{\alpha},R_{M})\,\Psi(r,R_{\alpha},R_{M}) \tag{5.1}$$

The wavefunction is a function of the coordinates of the electrons (r) and depends on the positions of both the quantum mechanical nuclei (R_{α}) and the molecular mechanical atoms (R_M). $\hat{H}$ for the QM/MM system is written as follows:[211]

$$\hat{H} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} \tag{5.2}$$

where $\hat{H}_{QM}$ is the Hamiltonian describing the quantum mechanical particles (i.e., electrons and nuclei) of the system and their interactions with each other; $\hat{H}_{MM}$ denotes the MM region; and $\hat{H}_{QM/MM}$ describes the interaction between the QM and MM regions. The expression of $\hat{H}_{QM/MM}$ has contributions from electrostatic and van der Waals interactions. The total energy of the system (E_{tot}) is evaluated by the expectation value of the Hamiltonian and is given as follows:

$$E_{tot} = \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = E_{QM} + E_{MM} + E_{QM/MM} \tag{5.3}$$

where Ψ is the electronic wavefunction of the QM region. The energy of the MM region (E_{MM}) represents the solvent interaction energy and is determined by a standard MM force field. The total interaction energy between the QM and MM regions ($E_{QM/MM}$) consists of contributions from both electrostatic and van der Waals terms. Since electronic structures of the solvent molecules in the MM region are not explicitly represented, the QM/MM van der Waals term is necessary to account for the electronic repulsion and dispersion interactions between both regions.[211] An issue that has not been addressed concerns the treatment of hydrogen bonding interactions of the QM/MM boundary. Nevertheless, what is significant is that the QM/MM approach provides a practical means for quantum mechanical investigations in solution.

5.2.2 Computational Methods

5.2.2.1 Force Fields and Molecular Models

Molecular dynamics (MD) simulations were conducted using the Sander program and the general amber force field (GAFF) in the Amber 9 program suite.[212] TFA Forms I and II in the monomer and dimer states, taken from their respective crystal structures,[114] were built using the xLeap module of the Amber 9 program suite.[212] Single molecules and their respective dimers were optimized with τ_1 and the intermolecular distances between oxygen atoms of the hydrogen-bonded carboxyl groups held constant. The electrostatic potentials (ESPs) around the optimized structures were then utilized to obtain partial atomic charges for each TFA structure by the restricted ESP fitting (RESP) method.[213] Optimization and electrostatic potential calculations were carried out in solvent media of water, ethanol, and toluene with a polarizable continuum solvent model, within the self-consistent reaction field theory.[176,177] Partial atomic charges for molecules in water, ethanol, and toluene were obtained at two different levels of HF/6-31G(d) and B3LYP/cc-pVTZ.[214,215] Two different levels of theory were utilized to evaluate the effect of polarization on the atoms. The polarization effect is important in the condensed phase, where the local electrostatic environment is significantly different than that in the gas phase because of the presence of neighboring atoms. The calculations for obtaining partial atomic charges were performed using the Gaussian 09 program package.[187]

Systems of TFA in water, ethanol, and toluene were constructed by soaking single molecules or dimers, resembling Forms I and II, in a rectangular box. The minimum distances from the solute atoms to the surfaces of the boxes were set to 10 Å, 15 Å, and 20 Å for water, ethanol, and toluene, respectively, with a total of approximately 500

solvent molecules. The TIP3P model[216] was employed for water molecules. The force field parameters for ethanol and toluene were obtained from R.E.DD.B. database.[217]

5.2.2.2 Molecular Dynamics (MD) Simulations

The solvated systems were subjected to energy minimization via 1,000 iterations of steepest-descent and conjugate gradient to eliminate possible bad contacts. The systems were then heated from 0 to 298 °K for 40 ps in the NVT ensemble followed by equilibration runs for 1 ns with a time step of 2 fs under constant pressure (1 atm) and temperature (298 °K). It is generally desirable to perform simulations in the NPT ensemble so that exact temperature control is obtained and experimental conditions may be mimicked.[218] The absence of systematic changes in the densities and the generation of a stable MD trajectory for each of the simulated structures suggested that the equilibration time was sufficiently long. The final desired densities of the simulated structures were 0.99 g/cm^3, 0.79 g/cm^3, and 0.82 g/cm^3 in water, ethanol, and toluene, respectively, in close agreement with the experimental values of solvent densities.[219] Simulations were performed under harmonic restraints on TFA with spring constants of 50 kcal/mol/Å^2 for τ_1 and 40 kcal/mol/Å^2 for distances between oxygen atoms of the hydrogen-bonded carboxyl groups. Moreover, simulations where monomers were allowed to rotate freely were performed. The SHAKE algorithm was used to fix all covalent bonds involving a hydrogen atom.[220] Periodic boundary conditions were employed together with a 10 Å cutoff to treat long-range van der Waals and electrostatic interactions using the particle mesh Ewald method.[221] The temperature was held constant by Langevin dynamics with a collision frequency of 1.0. Constant pressure was maintained using the Berendsen method with a relaxation time of 1.0 ps.[222] The

coordinates of the simulated systems were collected every 50 ps during the equilibration MD stage. MD calculations created molecular trajectories that were used to examine solute conformations using the PTRAJ program implemented in Amber 9.[212]

5.2.2.3 Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations

The final energy-minimized coordinates of each simulated structure were used as an initial geometry to carry out further optimization by using the Own N-layered Integrated molecular Orbital and molecular Mechanics (ONIOM) approach,[223] a QM/MM scheme implemented in the Gaussian09 program.[187] Two layers were defined in the ONIOM calculations. The high layer was calculated quantum-mechanically at the B3LYP/6-31G(d,p) level and the low layer was calculated molecular-mechanically by using the Amber force field as utilized in MD simulations. The QM-treated high layer included the solute and solvent molecules located within 4 Å of the solute, whereas the rest of solvent environment was included in the MM-treated low layer. The minimum model size and, thus, solvent molecules that were included in the QM-treated high layer were inferred by the radial distribution function, calculated after MD simulations.

τ_1 and the intermolecular distances between oxygen atoms of the hydrogen-bonded carboxyl groups were held constant during the QM/MM geometry optimization. For comparison, QM/MM geometry optimizations on TFA monomers were also carried out with no bond length, angle, or torsion angle fixed. Intermolecular interaction energies between solute and solvent molecules, as well as among solute molecules, were then calculated at the B3LYP/6-31+G(d,p):Amber level. The basis set with diffuse orbitals was selected to improve the description of hydrogen bonding. The basis set superposition error was corrected by the counterpoise method[186] when calculating the intermolecular

energies. The energies reported for each structure embedded in water, ethanol, or toluene come from the lowest value obtained over ten snapshots of two independent MD trajectories.

5.3 Results and Discussion

5.3.1 Comparison of the Atomic Charge Methods

The solvent affects the charge distribution in the solute through polarization. As such, an evaluation of the partial atomic charges is crucial. Partial atomic charges for TFA Form I conformer are presented as example. The two sets of charges, obtained at different levels of theory, are highly correlated. Indeed, the correlation coefficient (R^2) is 0.98 in water, ethanol, and toluene (Figure 5.1). This indicates that the two sets of charges for TFA Form I conformer are highly similar in all solvent media.

However, linear regression analysis of the two sets of charges, obtained in different solvent media, yields a slope of 0.75, meaning that the partial atomic charges derived at the level of HF/6-31G(d) are larger in absolute value than those obtained at the higher level of B3LYP/cc-pVDZ. The HF/6-31G(d) method is expected to give large dipole moments and thereby enhance the solute polarity.[214] Such atomic partial charges constitute an appropriate force field model for a solute in a polar solvent, such as water and ethanol, consistent with the role of both solvents in polarizing the solute molecule. Thus, charge set distribution derived at the HF/6-31G(d) level reflects properly the polarity of the environment. In addition, the use of the continuum solvent models in the quantum mechanical calculations made it possible to represent the solvent polarization effects in the point-charge models in a systematic manner. In a nonpolar solvent, such as

toluene, the partial atomic charges obtained at the HF/6-31G(d) level seem to be too polarized. The reason is because the charges derived using such an approach overestimate dipole moments and thereby implicitly include the solvent polarization to some extent. The results are consistent with observations reported in literature, where charges in hydrophobic solvents were as much as 20 – 25 percent lower than those in water.[215,224] As such, in nonpolar solvents the partial atomic charges need to be corrected for the reduced polarization induced by the surrounding solvent. It is suggested that charge distributions derived at a higher level of theory (i.e. the B3LYP functional[178] and the cc-pVDZ basis set) mimic better the condensed phase of nonpolar solvents.[215] Therefore, partial atomic charges for TFA monomers and their corresponding dimers calculated at the level of HF/6-31G(d) in water and ethanol and at the level of B3LYP/cc-pVDZ in toluene were subsequently used in dynamics simulations.

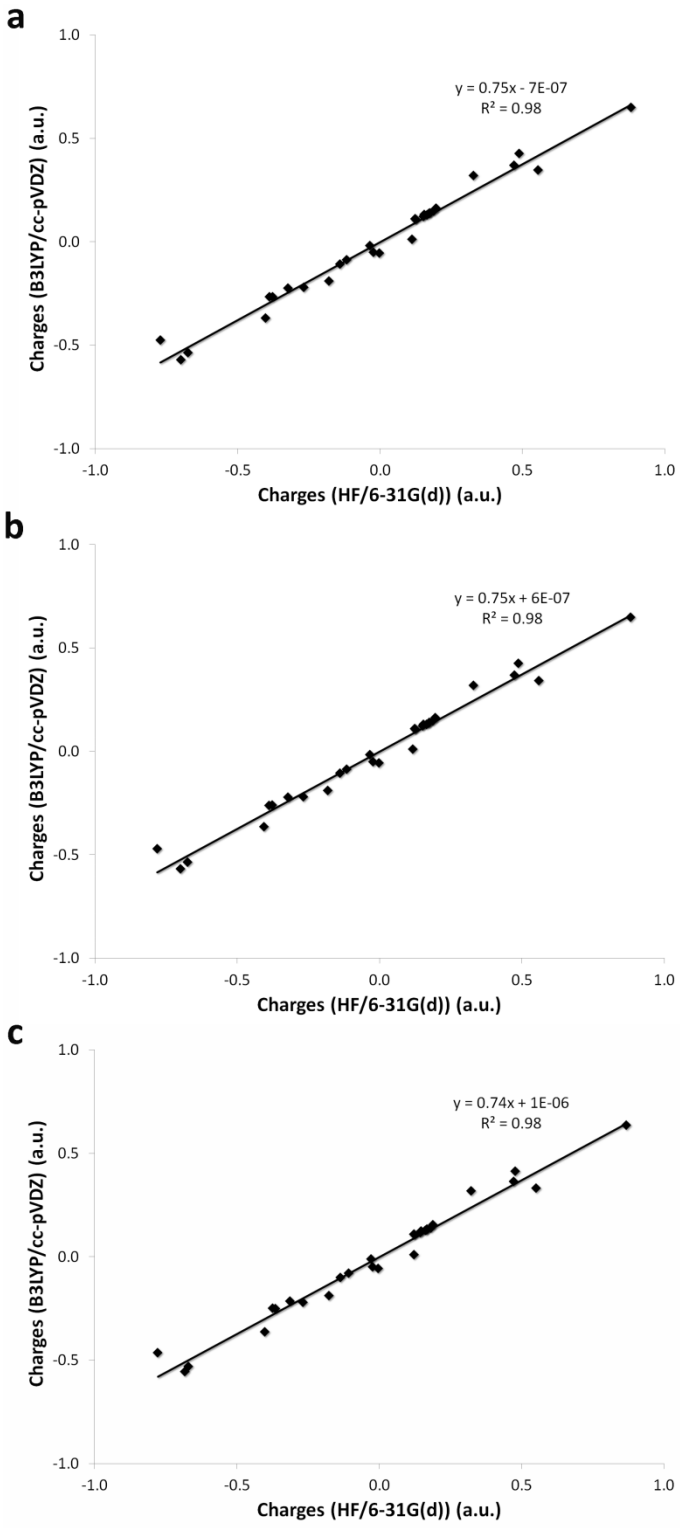

Figure 5.1. Correlation of partial atomic charges for TFA Form I conformer determined by RESP fitting, where the electrostatic potentials were derived at the levels of theory of HF/6-31G(d) and B3LYP/cc-pVDZ in a continuum solvent model, water (a), ethanol (b), and toluene (c).

5.3.2 Molecular Conformations and Dynamics

To understand structural information about the intermolecular interaction between solute and solvent molecules, an examination of the local environment surrounding the solute would be useful to ascertain the relative distribution of solvent molecules. The distribution of solvent molecules was explored by the radial distribution function (RDF), representing the relative probability of finding atoms of a solvent molecule at a certain distance from a specific atom of the solute molecule.[225] RDF plots were generated for the hydrogen atoms in solvent molecules (e.g., water, ethanol, and toluene) with respect to the carbonyl oxygen of TFA (Figure 5.2). A maximum in the probability of locating a hydrogen atom of water or ethanol molecules within 2 Å of the carbonyl oxygen of TFA Form I and II conformers (Figures 5.2a and 5.2b, respectively) indicates that the carbonyl oxygen of TFA interacts with solvent likely via hydrogen bonding. In addition, a second solvation shell by water appears to be located within approximately 4 Å of the carbonyl oxygen of TFA. There are no preferential interactions between TFA Form I or II conformers and toluene. The carboxylic acid group does not form hydrogen bonding with toluene, which has both a poor hydrogen bond donor ability and poor hydrogen bond acceptor ability. RDF plots for hydrogen atoms of solvent molecules with respect to the carbonyl oxygen of TFA Form I and II dimers demonstrate a low probability of finding solvent molecules in the first solvation shell, as indicated by the absence of distinct peaks near the first shell distance of 2 Å (Figures 5.2c and 5.2d). Note that there is no remarkable difference in the solvation structures around the carbonyl oxygen between Form I and Form II conformers as well as between Form I and Form II dimers, when each structure is embedded in water, ethanol, or toluene.

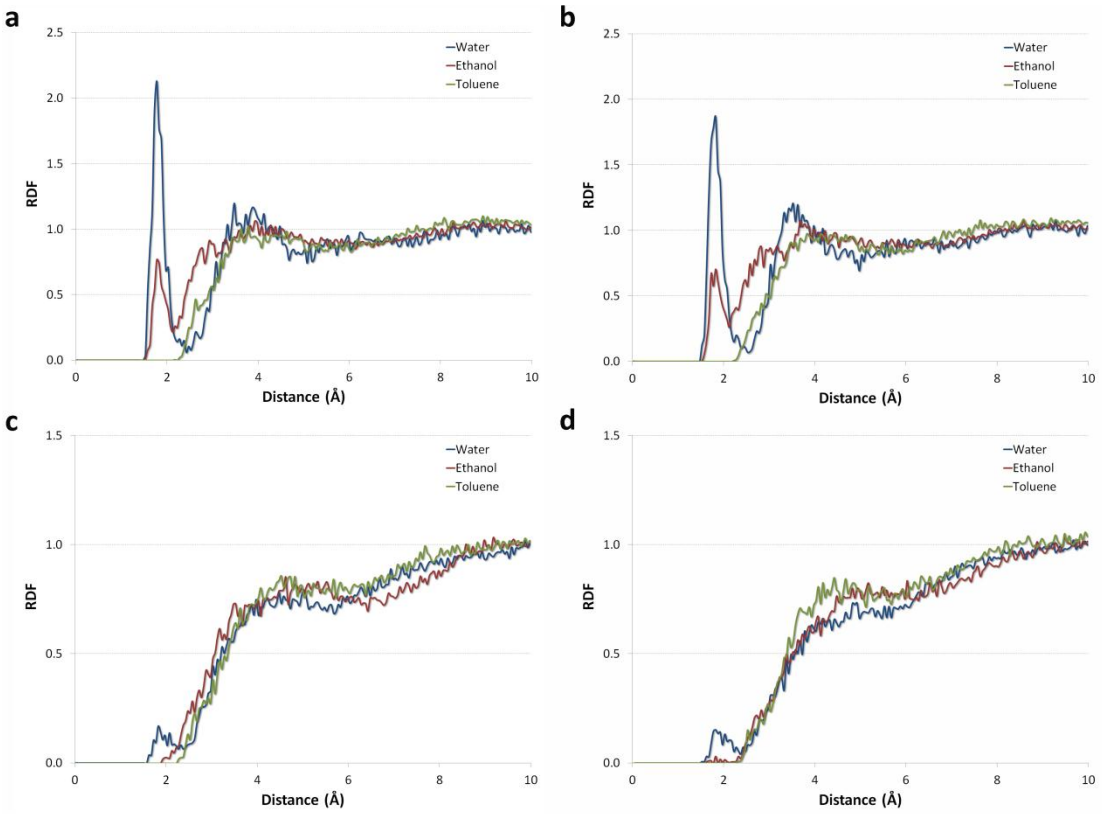

Figure 5.2. Radial distribution function (RDF) plots between solvent molecules and the carbonyl oxygen of TFA Form I conformer (a), Form II conformer (b), Form I dimer (c), and Form II dimer (d) embedded in water, ethanol, or toluene calculated from the acquired MD trajectories.

Thus far the structural properties that affect the drug's ability to interact in the bulk solvent have been discussed. The objective of this work is to be able to determine intermolecular interaction energies using the explicit solvation models. Therefore, the structural models of the solvent around TFA conformers and their corresponding dimers that were obtained from MD simulations were used for the subsequent QM/MM calculations. Key solvent molecules to be included in the layer calculated quantum-mechanically along with the solute molecule were identified based on the analysis of the solute-solvent RDF. It is important to choose a proper size for the discrete model

calculated quantum-mechanically in order to increase the accuracy of the calculations. For that reason, the first two solvation shells surrounding the solute molecule within a distance of 4 Å from the carboxyl group of TFA were included in the QM-treated layer.

Representative snapshots of the structural arrangement of TFA Form I and II conformers (Figures 5.3, 5.4, and 5.5) and Form I and II dimers (Figures 5.6, 5.7, and 5.8) are shown in the analyzed solvent media. During the dynamics run, both Form I and II conformers were also allowed to move freely (i.e., no restraining force was applied). In such a case, TFA conformation, which is described by the torsion angle τ_1, existed mostly in a conformation resembling Form II. The most favorable torsion angle, τ_1, can be found from the peak position in the probability distributions (Figure 5.9). Regardless of the initial conformation, the highest probability occurs for a conformation resembling Form II in water, ethanol, and toluene (i.e., τ_1 near -135°). This behavior is in line with the Form II conformer being the most stable conformation. Representative snapshots of the structural arrangement of the TFA conformer when applying no restraining force are shown in Figures 5.10, 5.11, and 5.12. Further, these examples illustrate the treatment of the solution system by the QM/MM method.

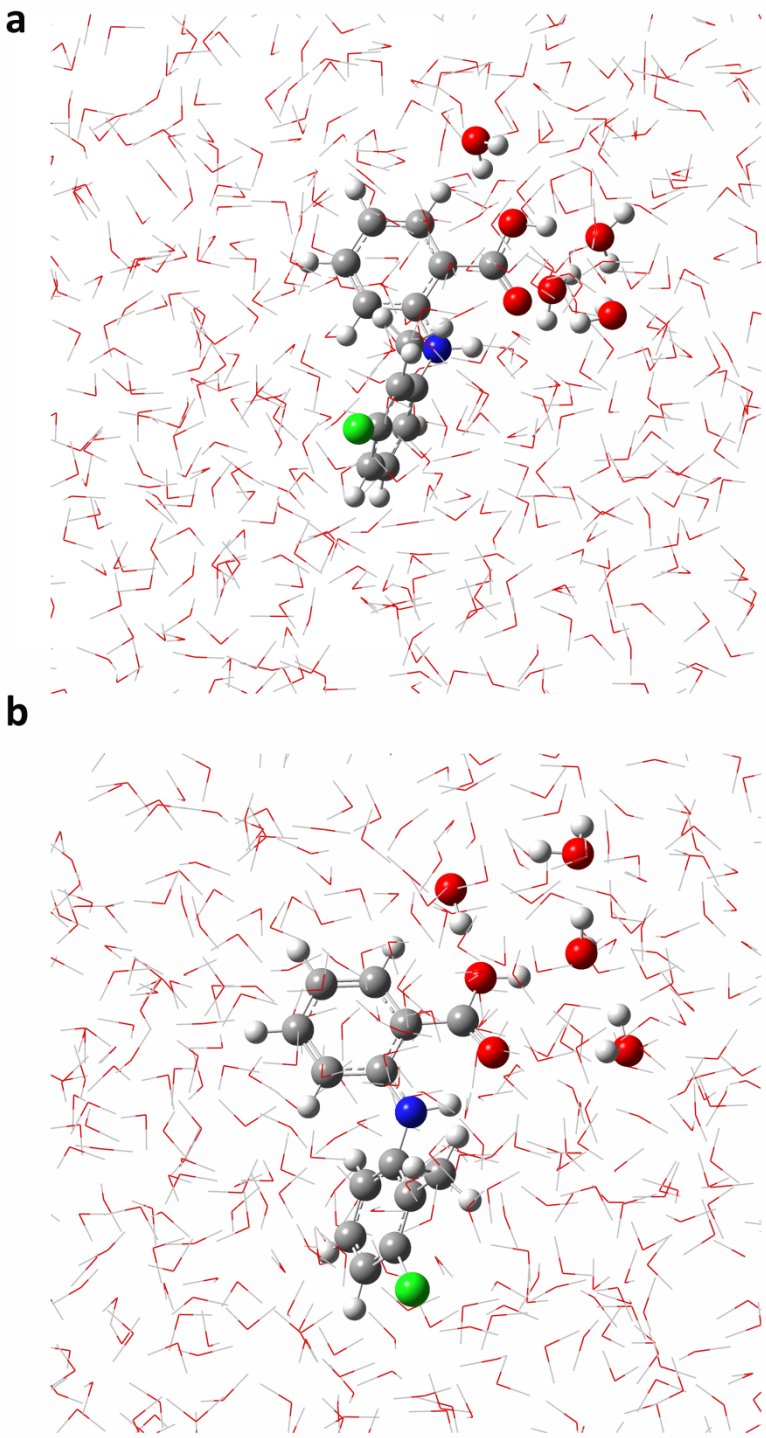

Figure 5.3. Representative snapshots of TFA Form I conformer (a) and TFA Form II conformer (b) in water. The QM-treated layer of the system, including the solute monomer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

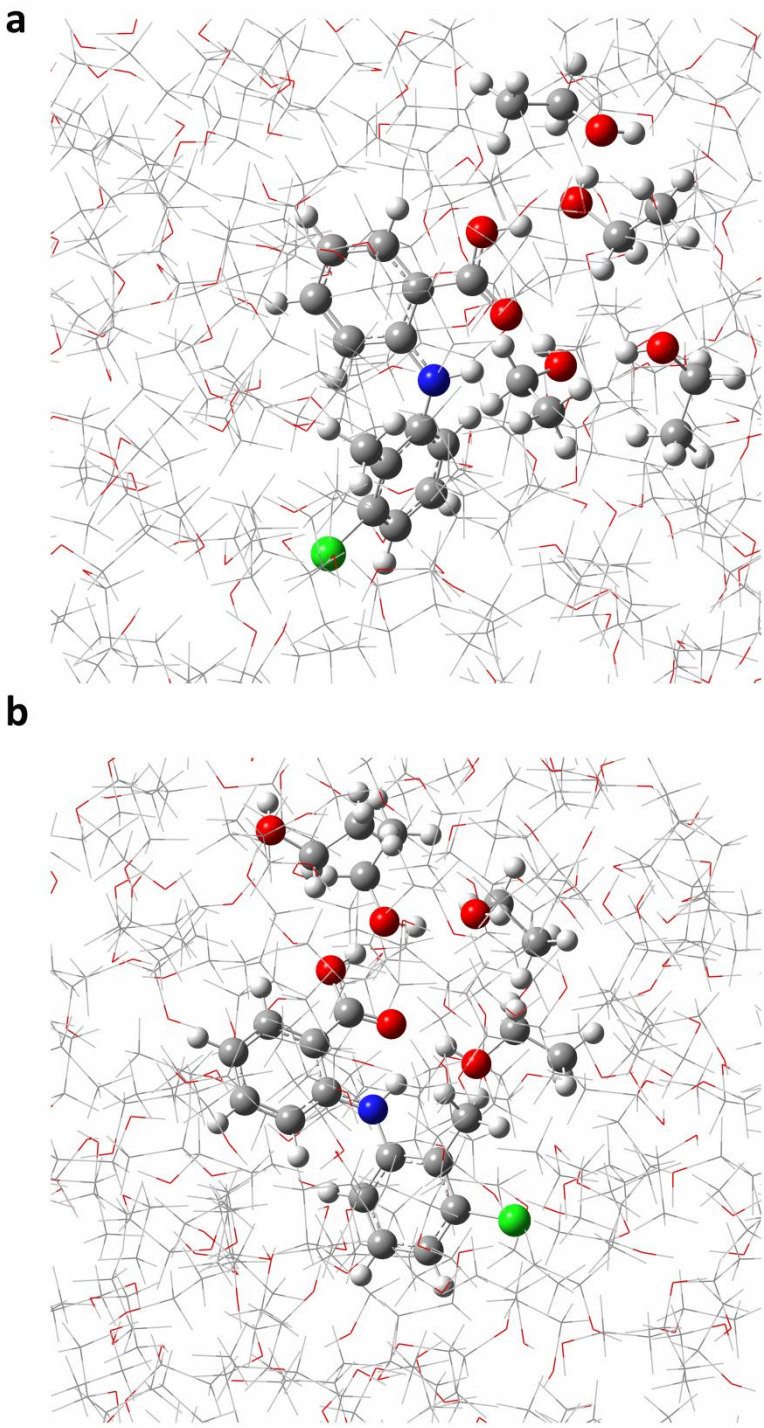

Figure 5.4. Representative snapshots of TFA Form I conformer (a) and TFA Form II conformer (b) in ethanol. The QM-treated layer of the system, including the solute monomer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

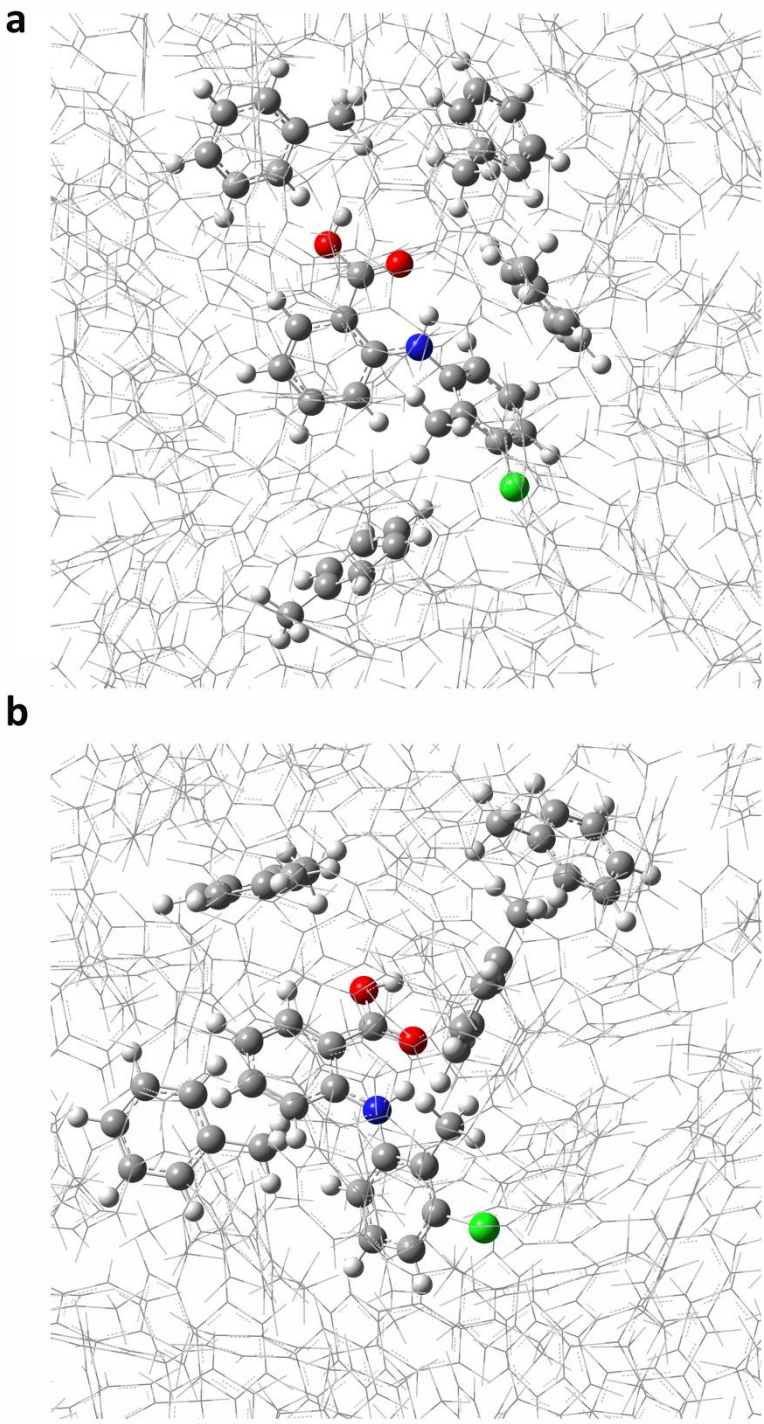

Figure 5.5. Representative snapshots of TFA Form I conformer (a) and TFA Form II conformer (b) in toluene. The QM-treated layer of the system, including the solute monomer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

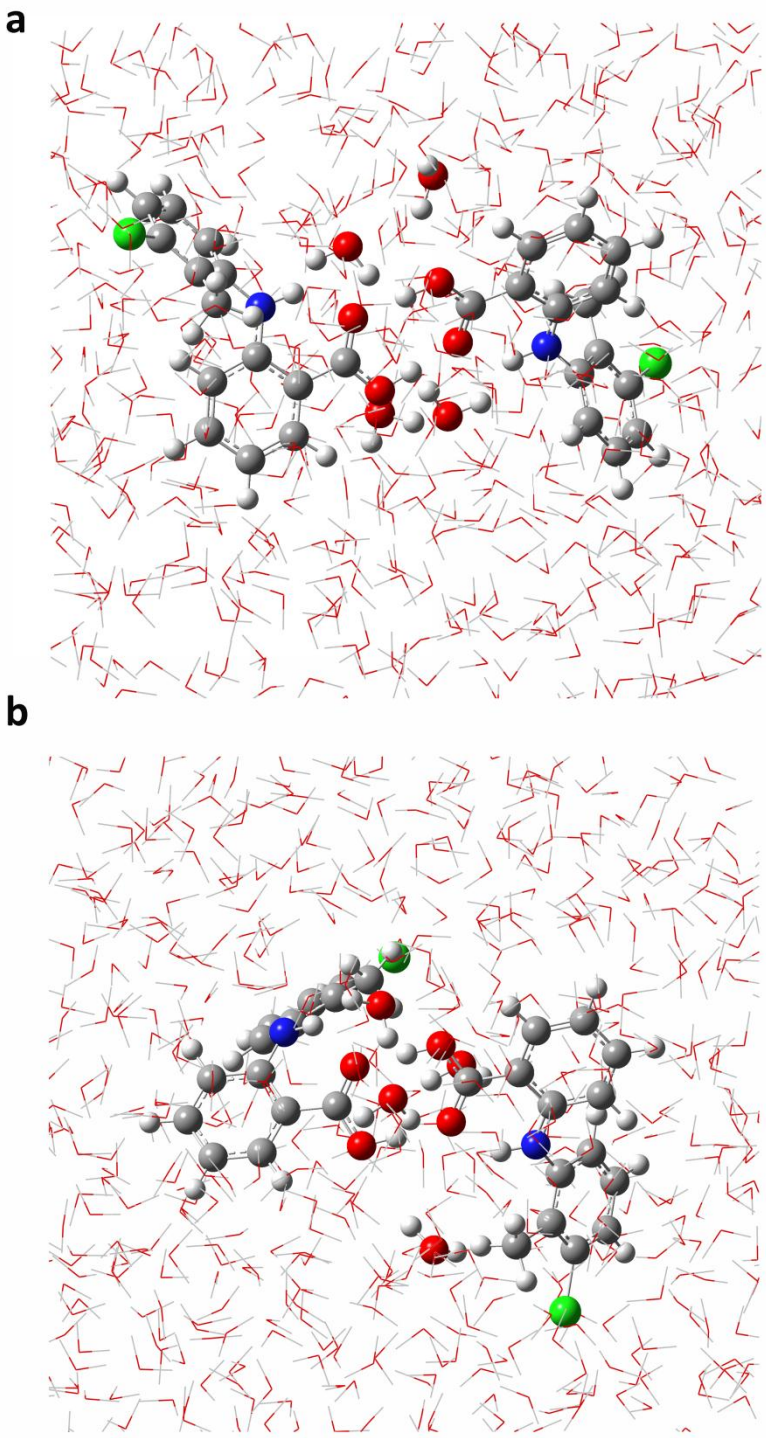

Figure 5.6. Representative snapshots of TFA Form I dimer (a) and TFA Form II dimer (b) in water. The QM-treated layer of the system, including the solute dimer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

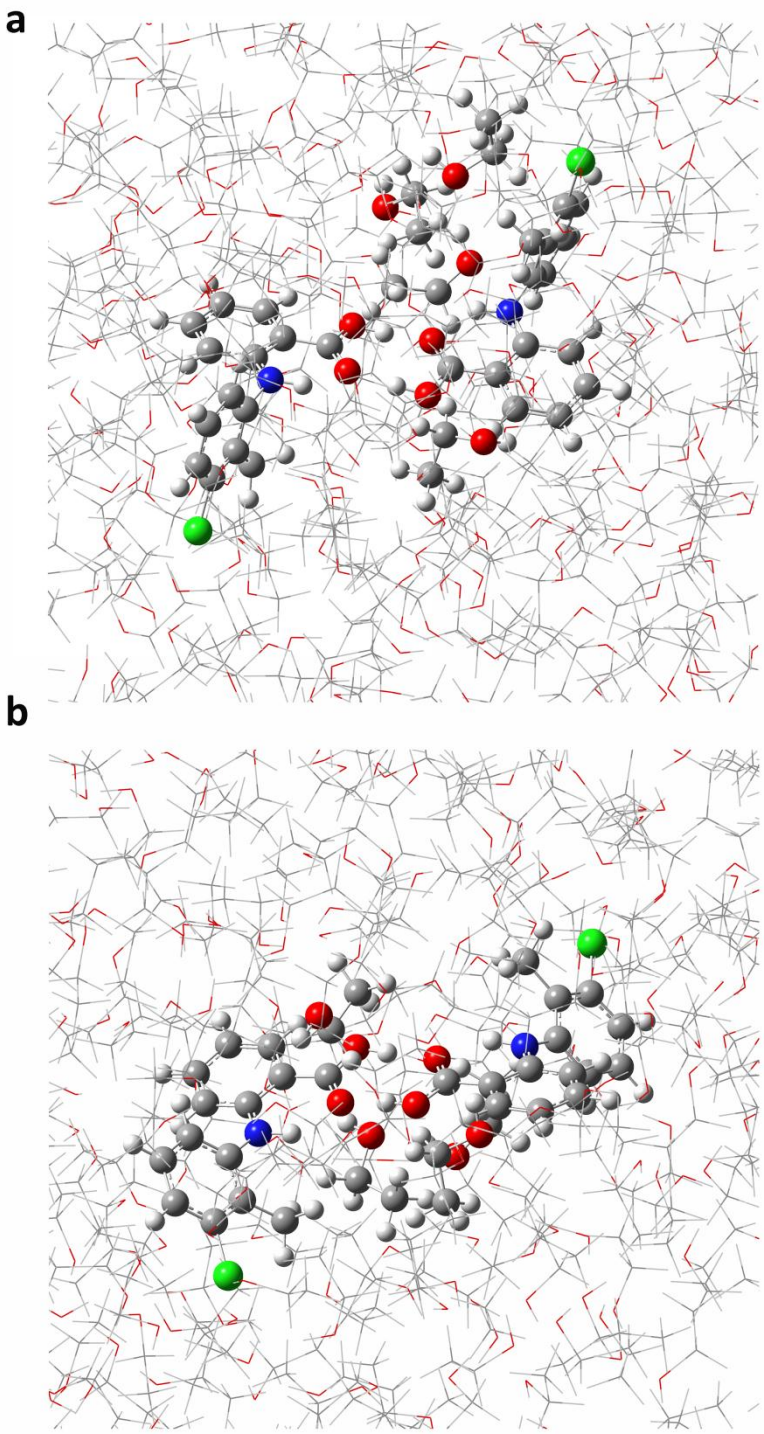

Figure 5.7. Representative snapshots of TFA Form I dimer (a) and TFA Form II dimer (b) in ethanol. The QM-treated layer of the system, including the solute dimer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

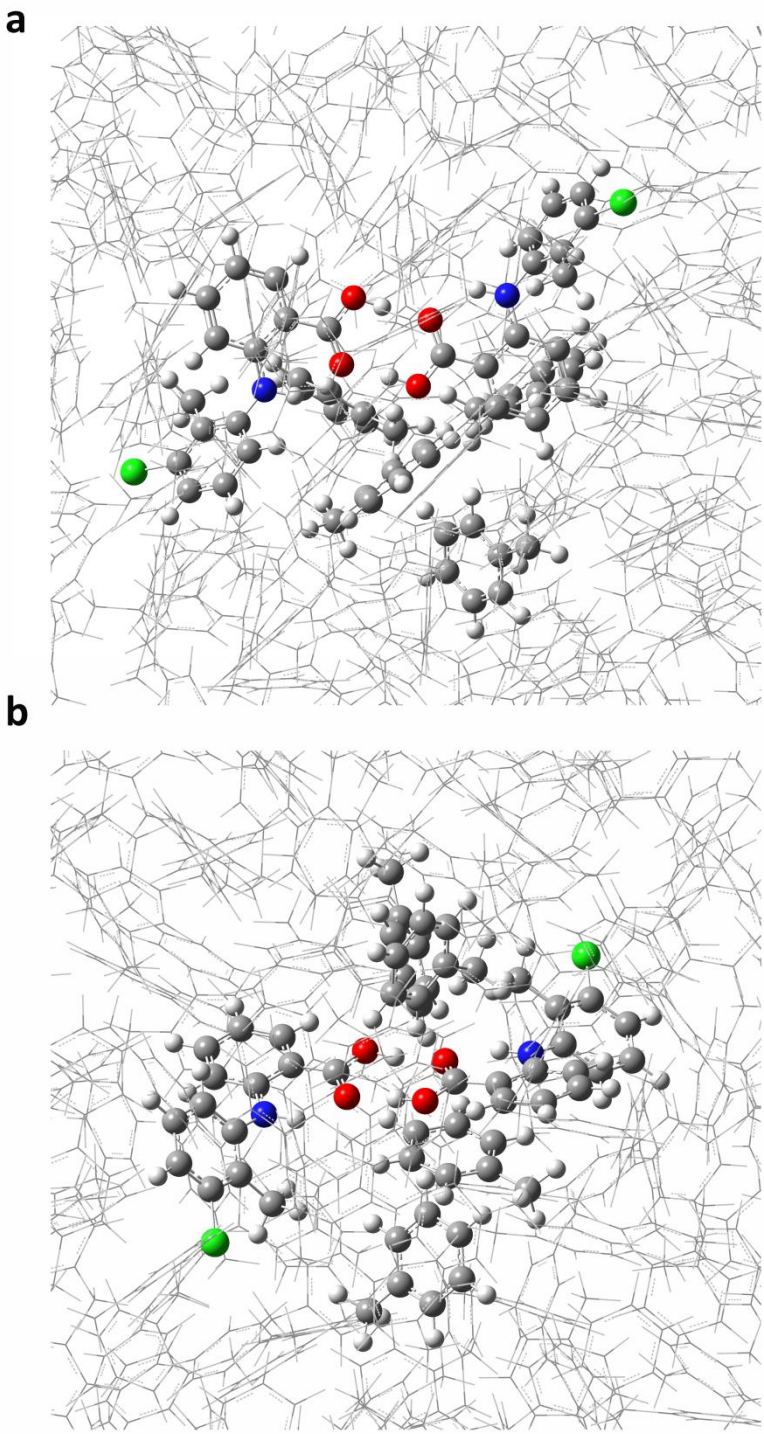

Figure 5.8. Representative snapshots of TFA Form I dimer (a) and TFA Form II dimer (b) in toluene. The QM-treated layer of the system, including the solute dimer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

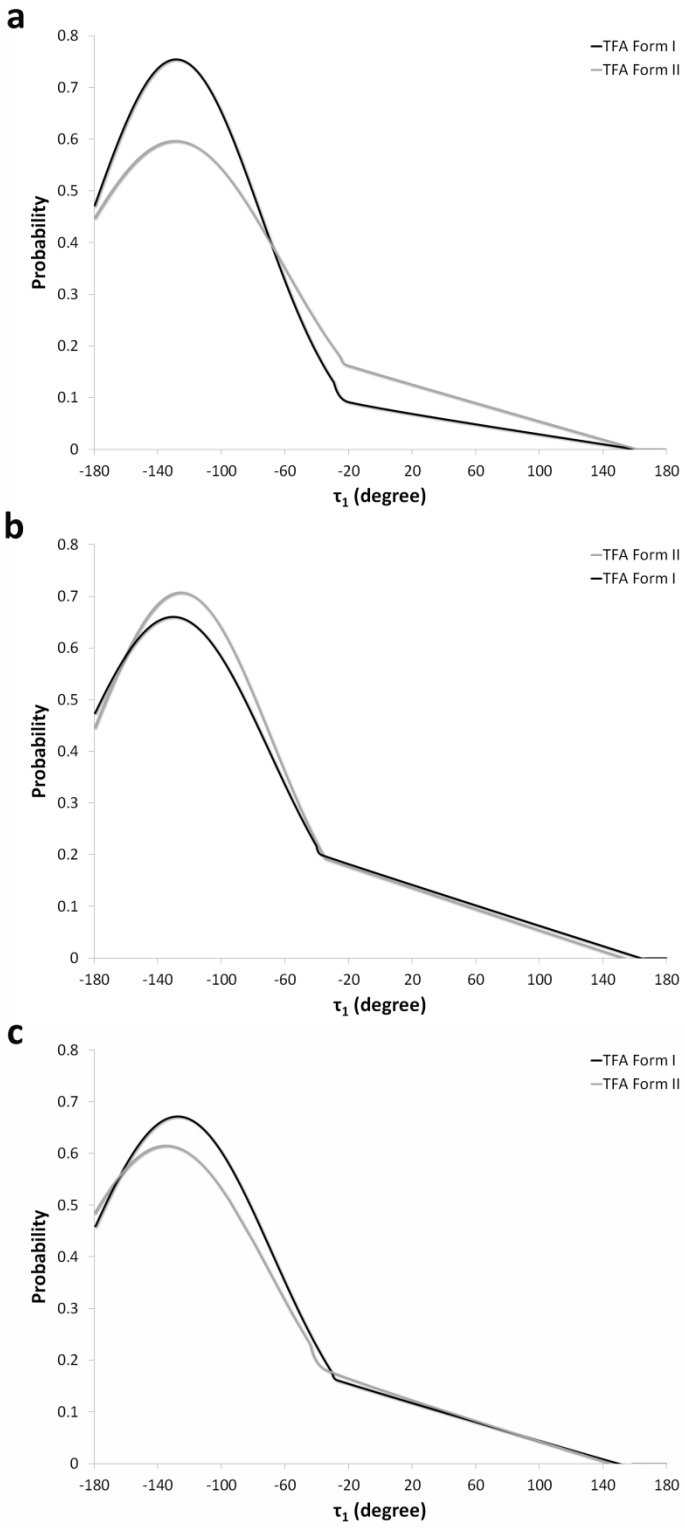

Figure 5.9. Probability distributions of the torsion angle τ_1 for the two TFA fully optimized conformers, Forms I and II, in water (a), ethanol (b), and toluene (c).

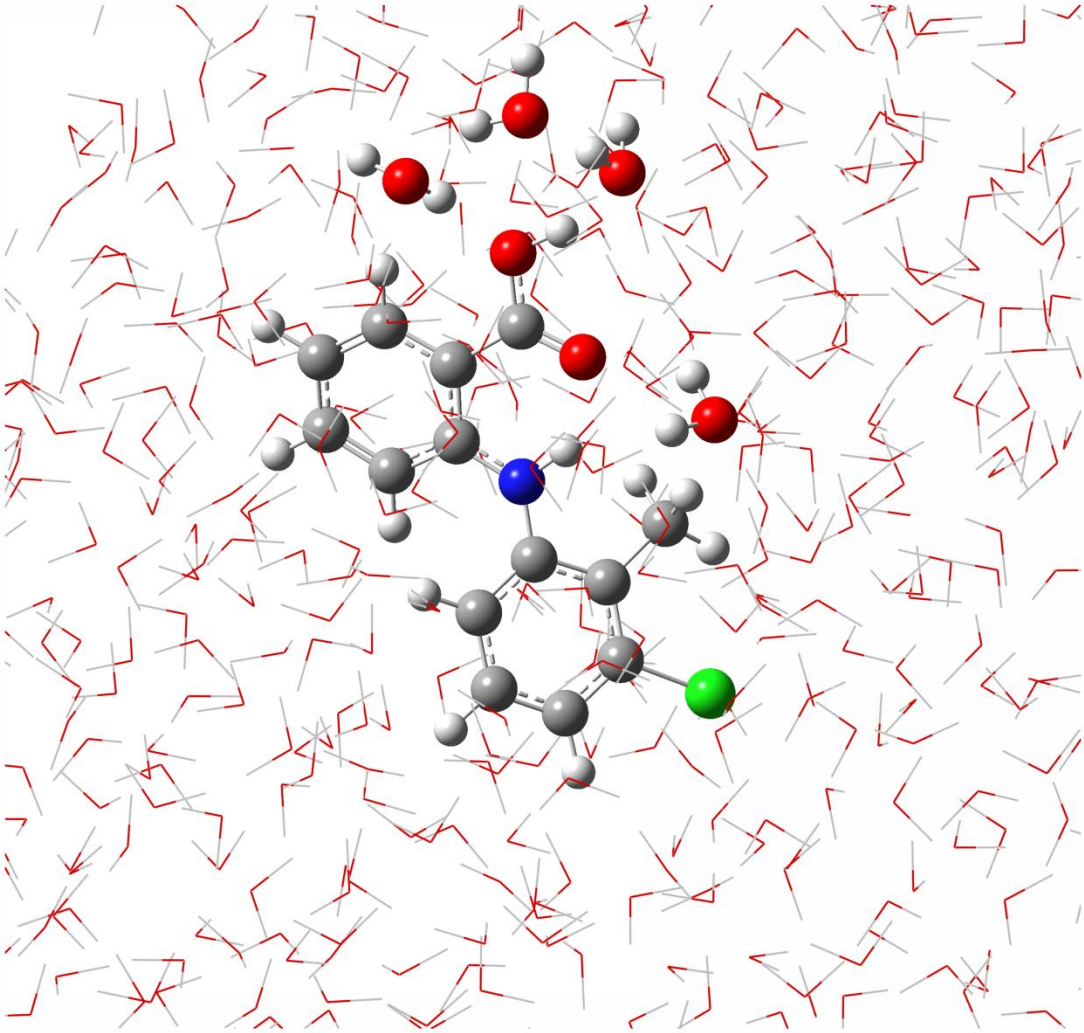

Figure 5.10. Representative snapshot of the TFA fully optimized conformer in water. The QM-treated layer of the system, including the solute monomer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

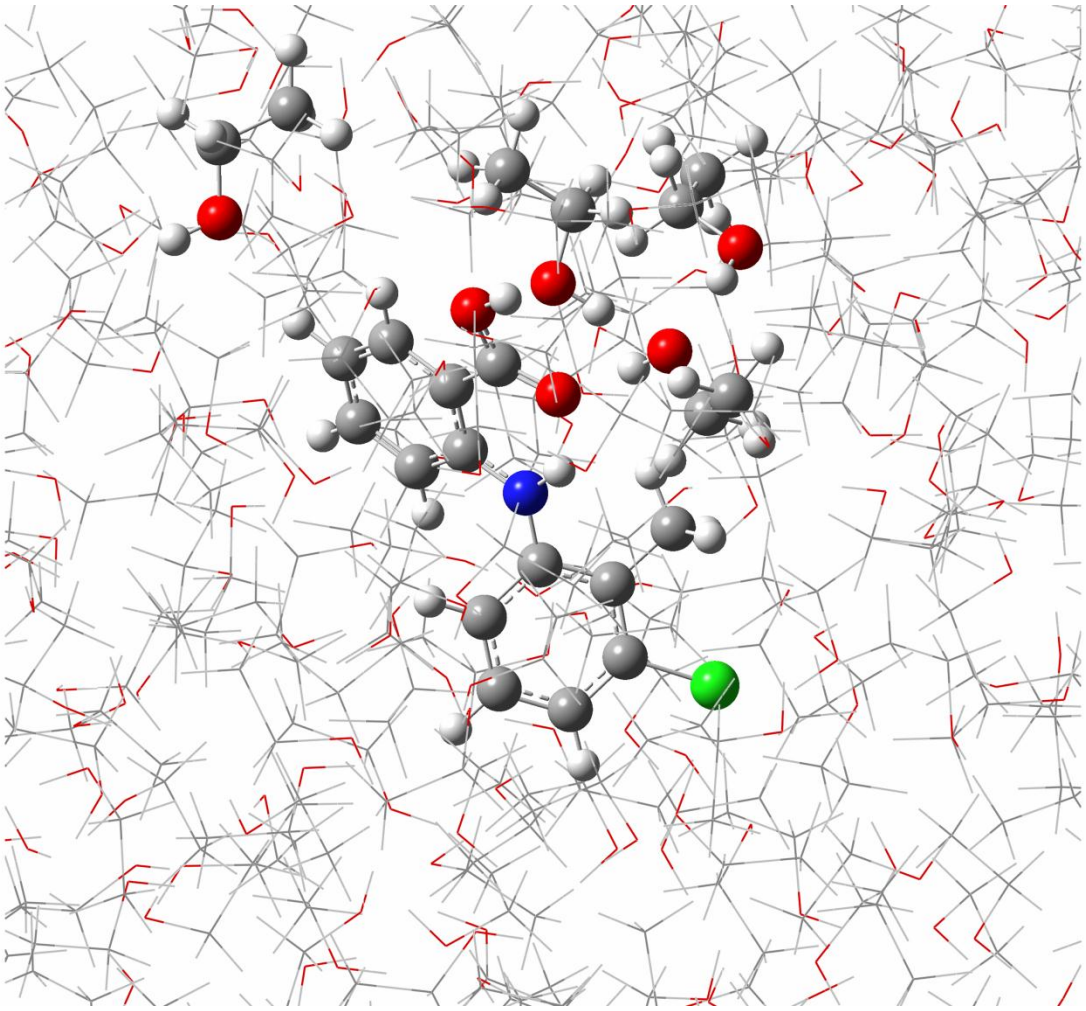

Figure 5.11. Representative snapshot of the TFA fully optimized conformer in ethanol. The QM-treated layer of the system, including the solute monomer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

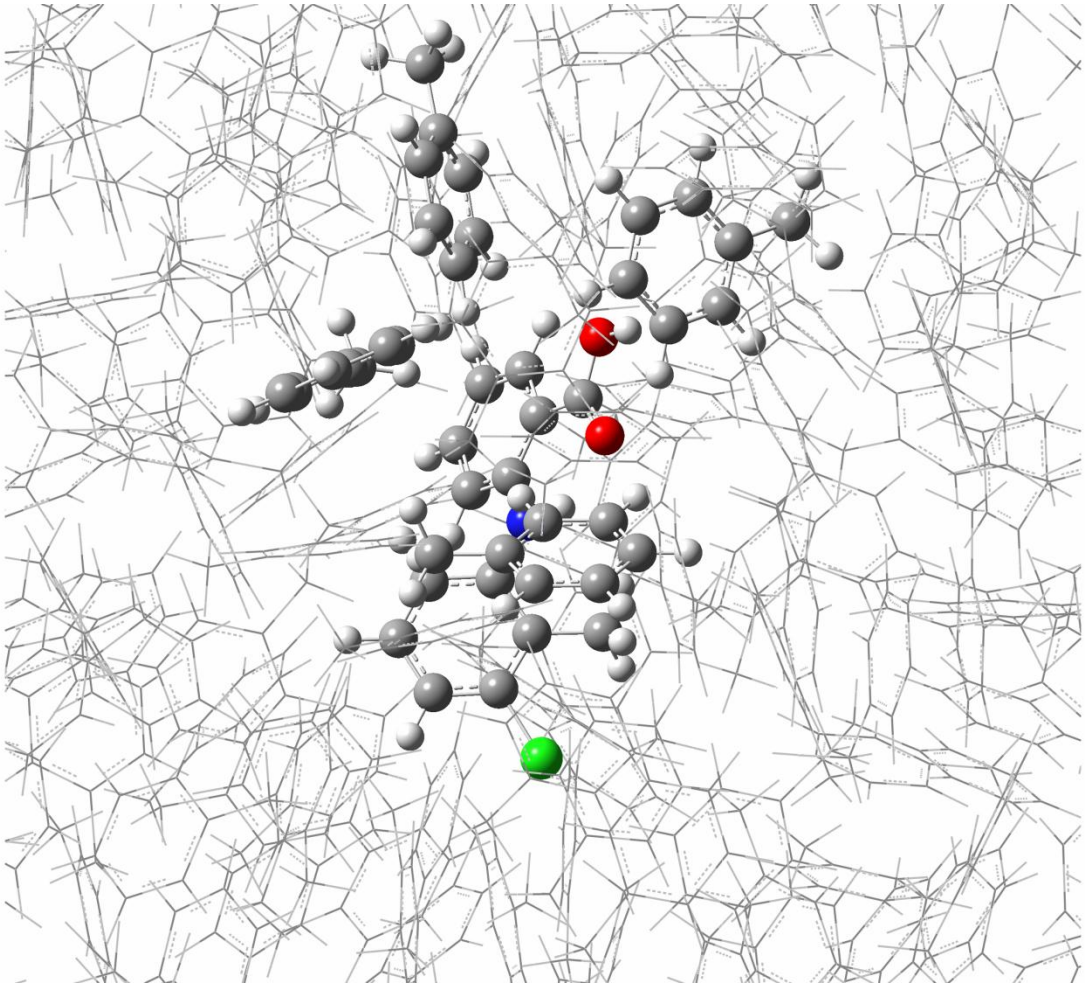

Figure 5.12. Representative snapshot of the TFA fully optimized conformer in toluene. The QM-treated layer of the system, including the solute monomer and solvent molecules within a distance of 4 Å from the carboxyl group, is represented in ball-and-bond model; the MM-treated layer of the system, including the rest of solvent molecules, is represented in stick model.

5.3.3 Sampling Scheme

A central problem for molecular simulations is sampling the conformational space. Such sampling is necessary in flexible molecules to obtain the correct distributions of conformers, which, for example, may vary for a molecule in different solvents. A vast range of methods has been developed and applied to molecular systems in an attempt to enhance coverage of the thermally accessible conformational space. One of the methods is the umbrella sampling simulation, which determines the probability that the system would be in a given conformation.[212] This is accomplished by dividing the pathway with external potentials (which are usually harmonic, hence the name "umbrellas") and sampling each part. Applying an external potential traps the molecule in one region along the reaction coordinates. In addition, the potential works to lower the energy barrier that exists between two conformations. In effect, umbrella sampling allows one to sample a particular section of the reaction pathway in which the molecular conformation is unfavorable.

The efficacy and ability of an approach applied for sampling depends on the goals of the computation. The goal of this work is to calculate the energy of TFA conformers, specifically Forms I and II, in solution by QM/MM. It should be noted that the only purpose of performing MD simulation is to examine the dynamics of the solvent surrounding the solute molecule and the interaction between the solute and the solvent environment. We are only interested in the simulated structures because the total energies calculated by MD are not reliable, since they are the result of parameterized force fields. Therefore, the mode of sampling that was utilized was to collect ten snapshots of each simulated structure from multiple MD trajectories.[226] Specifically, this strategy consisted

of generating trajectories starting from the same initial conformation, but with different initial velocity distributions. Two independent trajectories were generated. A concern guided the systematic search for a reliable sampling scheme: how many snapshots should be extracted in order to achieve a meaningful structural analysis of the results. In order to assess the consistency of the present scheme and the extent of structural stability of the solute conformation generated in the MD trajectories, the time evolution of the torsion angles (i.e., τ_1, τ_2, and τ_3 denoted in Figure 5.13a) and of the energy of the TFA molecule was analyzed. The energy values were calculated by QM/MM. Torsion angle and energy profiles of the Form I conformer in water, ethanol, and toluene are shown as example (Figures 5.13 and 5.14).

Despite the fact that the harmonic restraining force was applied only on the torsion angle τ_1, neither τ_2 nor τ_3 deviates significantly. For the system in water, the overall spread of the torsion angles over time was relatively small, indicating that the TFA conformer did not undergo torsional transitions (Figure 5.13b). As such, the structural features of the Form I conformer were maintained during the simulations, perhaps because the system had settled in a local minimum of the conformational space. Similar results were obtained for the system embedded in ethanol or toluene (Figures 5.13c and 5.13d). Furthermore, the energy of the TFA conformer did not dramatically change among the snapshots acquired from the simulations; all the values were relatively close, falling within 10 kJ/mol (Figure 5.14). The chosen sampling scheme led to ample consistency among snapshots of independent MD trajectories. It would be highly unlikely for the generated trajectories to exhibit structural deviation from the starting structure upon increasing the number of collected snapshots. This work is not intended to be the last word on sampling

efficiency and may represent a simplified approach. However, the question of the effect of sampling on the system stability and on the solute conformation has been addressed. Therefore, because of the reduced variability among the collected snapshots, the lowest energy of the solute molecule in solution was considered. The lowest energy structure was picked because it corresponded to the highest probability to be found within the Boltzmann distribution.

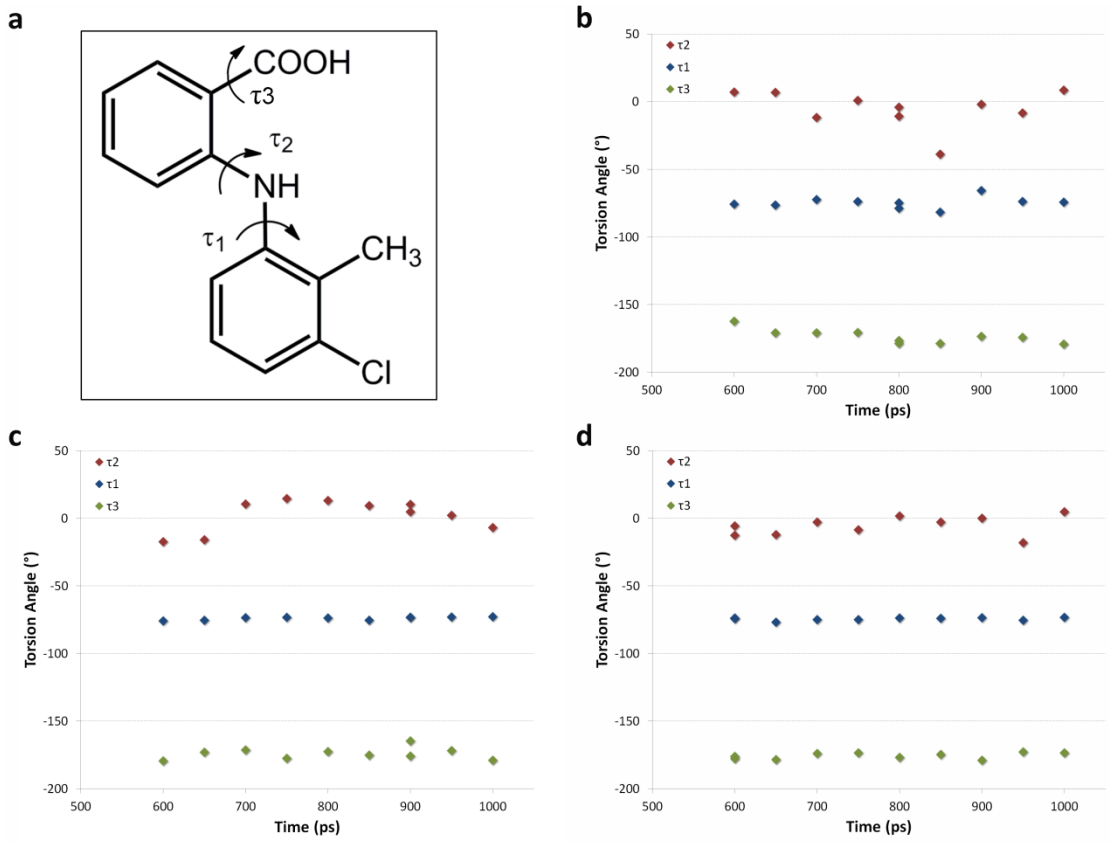

Figure 5.13. Molecular structure of TFA, defining torsion angles τ_1, τ_2, and τ_3 (a). Representative segments of conformational profiles for the three torsion angles of the simulated TFA Form I conformer embedded in water (b), ethanol (c), and toluene (d).

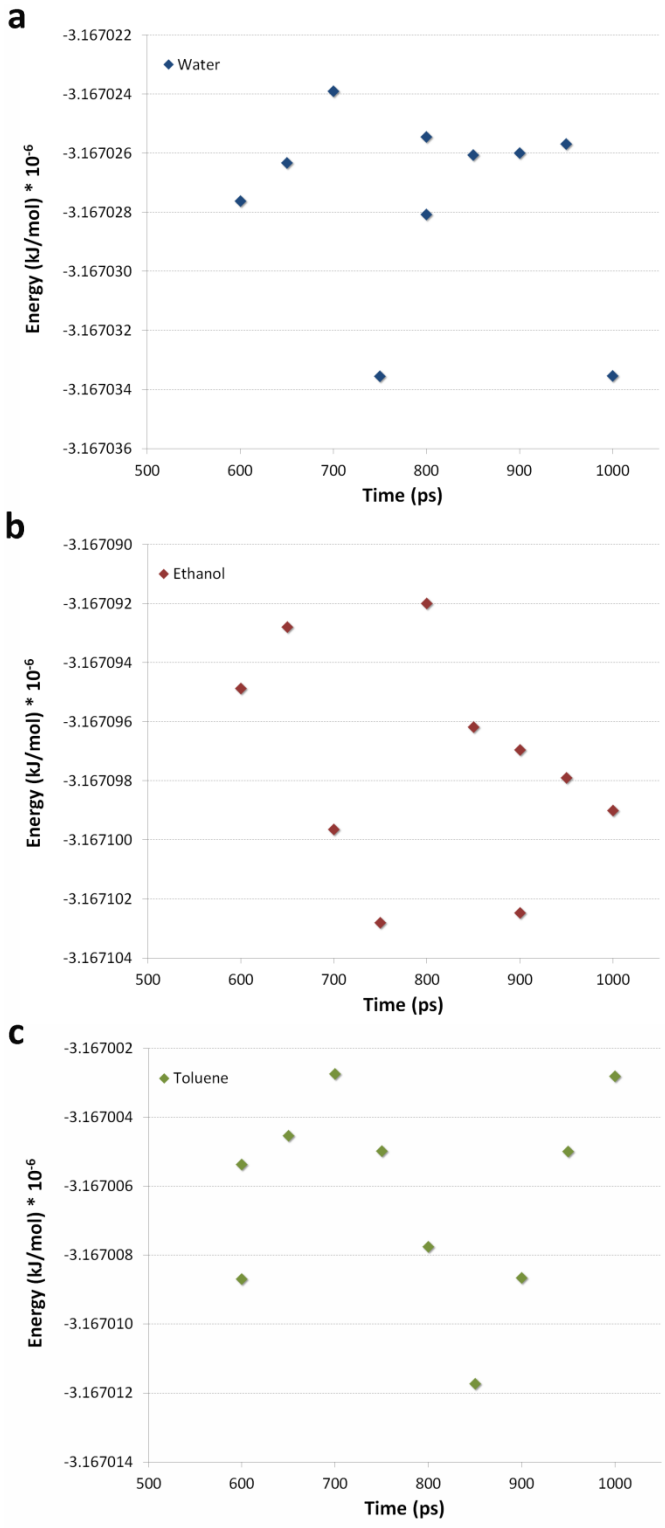

Figure 5.14. Time evolution of energies of TFA Form I conformer embedded in water (a), ethanol (b), and toluene (c). Snapshots were generated from two independent MD trajectories and then the energy values were obtained from QM/MM calculations.

5.3.4 Intermolecular Interaction Energies

The energies of the MM-treated low layer, the QM-treated high layer, and the interaction between the two layers together with the total energy of a system constituted by the Form I conformer embedded in solvent media are shown in Table 5.1. The energies acquired at each level are additive. A careful examination on the calculated total energy of the system suggests that the quality of the result is controlled by the QM-treated high layer that includes the solute and solvent molecules located within 4 Å from the solute. This allows one to focus on the local structural factors and short-range interactions without necessarily considering the long-range interactions for the rest of the solvent environment.

Table 5.1. Calculated energies of each layer defined in ONIOM calculations as well as the total energy for a snapshot of the TFA Form I conformer embedded in solvent media. The energy values are reported in Hartree-Fock (HF).

	E_{Water} (HF)	$E_{Ethanol}$ (HF)	$E_{Toluene}$ (HF)
MM-treated layer	-0.0658	-0.0752	-0.0759
QM-treated layer	-1511.9658	-1826.4633	-2292.5501
Interaction layer	-14.7088	-3.0429	-2.1289
Total system	-1526.7405	-1826.5753	-2294.7550

Once the QM/MM method has been used to determine the energy of the system, the intermolecular interaction energy between solvent and the TFA molecule – more suitably called solvation energy ($\Delta E_{solvation}$) – can be computed according to the following equation:

$$\Delta E_{solvation} = E_{system} - (E_{solute,gas} + E_{solvent}) \quad (5.4)$$

where E_{system} was obtained by the QM/MM method on the solute molecule in a given solvent system; $E_{solvent}$ was obtained by the QM/MM method on the pure solvent; and $E_{solute,\ gas}$ was obtained by the QM calculation on the solute molecule in the gas phase. Solvation energy was calculated for monomers and their respective dimers, resembling Forms I and II, as well as for monomers that were allowed to move freely.

In water, the estimated solvation energy of the Form I conformer was -158.84 kJ/mol, while that of the Form II conformer was -176.40 kJ/mol. In ethanol, the solvation energies of the Form I and II conformers were computed to be -234.49 kJ/mol and -242.74 kJ/mol, respectively. In a nonpolar solvent, such as toluene, the estimated solvation energies of the Form I and II conformers were -136.67 kJ/mol and -142.22 kJ/mol, respectively. When no restraint was applied to the solute molecule, the solvation energies were -176.24 kJ/mol, -247.01 kJ/mol, and -141.62 kJ/mol in water, ethanol, and toluene, respectively. These energy values are close to those computed for the Form II conformer, confirming that during the entire simulation period the TFA molecule resembled the Form II conformational state regardless of the solvent polarity. The results are summarized in Table 5.2.

Table 5.2. Calculated solvation energy values of TFA Form I and II conformers embedded in water, ethanol, and toluene. The energy of a fully optimized molecule (i.e., no restraining force was applied) is also reported in the three solvents.

Conformer	$E_{Solvation, Water}$ (kJ/mol)	$E_{Solvation, Ethanol}$ (kJ/mol)	$E_{Solvation, Toluene}$ (kJ/mol)
Form I	-158.84	-234.49	-136.67
Form II	-176.40	-242.74	-142.22
Fully Optimized	-176.24	-247.01	-141.62

The rank order of solvation energy for TFA conformers is:

$$E_{Solvation, ethanol} < E_{Solvation, water} < E_{Solvation, toluene}$$

By comparing the energy values in ethanol and toluene, the solvation energy is more negative as the solvent polarity increases, as measured by the increase in the dielectric constant. This is mainly because the carboxylic acid in TFA is better solvated in polar solvents than in nonpolar solvents, likely via hydrogen bonding, in tune with RDF plots (Figures 5.2a and 5.2b). However, the solvation energy in water is less negative than that in ethanol, although the polarity of water is higher than that of ethanol. As such, the macroscopic dielectric constants alone are unsuitable measures of the molecular, microscopic interactions.[227] The reason the TFA molecule has a stronger interaction with ethanol and a lower interaction with water might be due to the greater amount of energy needed to make a solvent cavity to accommodate the TFA molecule in water than in ethanol. Other factors, nonetheless, must account for the large negative solvation energy values of both TFA conformers in ethanol. An ethanol molecule has a hydrocarbon group that cannot participate in hydrogen bonds and, thus, intermolecular hydrophobic

interactions may contribute to the structural stability of the solute in ethanol. Also, it is expected that in nonpolar toluene hydrophobic interactions between aromatic rings of the solute and the solvent molecules account for the main contribution of intermolecular interactions to the solvation energy.

The net contributions of each type of interaction (i.e., hydrogen bonding vs. hydrophobic interaction) to the overall stability of the solute molecule in solution could not be inferred. To date, no literature study has been reported on the QM/MM capability to evaluate individual components of the solvation energies in solution. There is also uncertainty regarding the role of individual interactions in favoring one conformation over the other. Overall, the results indicate that in all of the solvents the Form II conformer, which represents the global energy minimum, interacts preferentially with solvent molecules. However, when computing the solvation energy of TFA dimers embedded in water, ethanol, and toluene Form I dimers are more stabilized than Form II dimers. The trend is reflected by lower (or more negative) values of solvation energy for Form I dimers than those for Form II dimers (Table 5.3).

Table 5.3. Calculated solvation energy values of TFA Form I and II dimers embedded in water, ethanol, and toluene.

Dimer	$E_{Solvation, Water}$ (kJ/mol)	$E_{Solvation, Ethanol}$ (kJ/mol)	$E_{Solvation, Toluene}$ (kJ/mol)
Form I	-213.97	-291.19	-264.36
Form II	-188.86	-275.37	-252.35

Solute-solute interactions via hydrogen bonding are expected to compensate for the loss of solvation (i.e., solute-solvent interactions) accompanying dimer formation. The intermolecular hydrogen bonding energy of Form I and II dimers, composed of monomers A and B and embedded in solvent media, was computed as follows:

$$E_{\text{hydrogen bonding}} = E_{\text{dimer AB, solution}} - (E_{\text{monoomer A, solution}} + E_{\text{monomer B, solution}}) \qquad (5.5)$$

The hydrogen bonding strength in the Form I dimer was estimated to be -30.24 kJ/mol, -33.30 kJ/mol, and -31.16 kJ/mol in water, ethanol, and toluene, respectively, while that in the Form II dimer was -28.15 kJ/mol, -30.90 kJ/mol, and -28.14 kJ/mol in water, ethanol, and toluene, respectively. The results, listed in Table 5.4, indicate that hydrogen bonding interaction is more favorable for the Form I dimer as compared to the Form II dimer by approximately 2-3 kJ/mol. The largest difference in hydrogen bonding strength was obtained for the dimer in the toluene phase. Note that the energies reflect the formation of solute-solute interactions per hydrogen bond. Because solvation stabilizes the carboxylic group, the effects of hydrogen bonding interactions between solute molecules are smaller in the solution phase than in the gas phase.

Table 5.4. Calculated hydrogen bonding energy values of TFA Form I and II dimers embedded in water, ethanol, and toluene.

Dimer	$E_{\text{Hydrogen bond, Water}}$ (kJ/mol)	$E_{\text{Hydrogen bond, Ethanol}}$ (kJ/mol)	$E_{\text{Hydrogen bond, Toluene}}$ (kJ/mol)
Form I	-30.24	-33.30	-31.16
Form II	-28.15	-30.90	-28.14

Strong intermolecular hydrogen bonding between solute molecules may compete with intermolecular interactions between solute and solvent molecules, resulting in a conformational change to a structure that is more stable and more prone to form stronger aggregates. The determining factor for the structural rearrangement of a molecule in solution is its energy state. The molecule will adopt a conformation that yields energetically more favorable interactions with the solvent. At the same time, when solution conditions allow for the formation of association species a molecular conformation that maximizes intermolecular hydrogen bonding through dimer formation so as to minimize the energy will be preferred. The energies of Forms I and II in their monomer and dimer states are summarized in Table 5.5. Regardless of the solvent, the Form II conformer is the more stable conformation, confirming the results in the gas phase presented in Chapter 4. The Form II conformer is more stable than the Form I conformer by 12.1 kJ/mol, 3.7 kJ/mol, and 3.9 kJ/mol in water, ethanol, and toluene, respectively. Thus, the Form I conformer, being less stable, adapts to a solvent environment by adopting a more internally strained conformation able to strengthen interactions with another solute molecule. The net effect is that the Form I dimer is more stable in a solution phase. The Form I dimer is more stable than the Form II dimer by 49.1 kJ/mol, 2.4 kJ/mol, and 4.5 kJ/mol in water, ethanol, and toluene, respectively.

The energies of TFA existing as dimer were used to calculate the fraction of the population expected to occur in each configuration at 25 °C in ethanol, in order to corroborate the experimental results. Because the energy difference between the two dimers in the ethanol phase was found to be 2.4 kJ/mol, the Form I dimer would contribute 72 percent of a Boltzmann population. This calculation agrees very well with

the experimental finding of a 62 percent fraction of dimers in the high concentration solution (discussed in Chapter 3), indicating the validity of the energy calculation that accounts for the explicitly modeled solvent molecules.

Table 5.5. Calculated energy values of TFA Forms I and II in their monomer and dimer states embedded in water, ethanol, and toluene. The energy values are reported in Hartree-Fock (HF). The energy difference between Form I and Form II in their monomer and dimer states embedded in water, ethanol, and toluene is also reported in kJ/mol.

Molecular Species	E_{Water} (HF)	$E_{Ethanol}$ (HF)	$E_{Toluene}$ (HF)	ΔE_{Water} (kJ/mol)	$\Delta E_{Ethanol}$ (kJ/mol)	$\Delta E_{Toluene}$ (kJ/mol)
Monomer Form I	-1206.2592	-1206.2854	-1206.2509	12.1	3.7	3.9
Monomer Form II	-1206.2638	-1206.2868	-1206.2524			
Dimer Form I	-2412.4997	-2412.5283	-2412.5214	-49.1	-2.4	-4.5
Dimer Form II	-2412.4810	-2412.5274	-2412.5197			

5.4 Conclusion

Although modeling the solvent as a continuum medium is still used to investigate solvation effects, its shortcomings prompted the need to evaluate the performance of other techniques, in which the solvent is explicitly taken into account. Thus, this chapter presented the computational methodology developed for calculating energies of the solute molecule and its intermolecular interactions in solution. A few observations indicated that the Form II conformer represents the preferred conformation adopted by TFA. First, when allowing either the Form I or Form II conformer to rotate freely, MD simulations revealed a higher probability for the occurrence of the torsion angle τ_1 near the value corresponding to that found in the Form II conformer. The second observation comes from the energy calculation by the QM/MM approach. The Form II conformer forms more favorable interactions with solvent molecules. This accounts for the lower solvation energy of this conformer. Therefore, the Form II conformer is the more stable conformation in a solution phase. The results corroborate those obtained in the gas phase. Also, the preferred conformation of the TFA molecule does not depend on the solvent polarity. However, when the TFA molecule exists in a dimer state the Form I dimer exhibits lower solvation energy and greater stability. The conformational arrangement of the TFA molecule similar to that found in the Form I conformer is essential for the formation of stronger hydrogen bonding between solute molecules.

Chapter 6 – Summary and Discussion

The overall goal of this research is to determine the extent to which solution conditions influence the molecular conformation and how their effects are reflected in the resultant molecular packing arrangements of conformational polymorphs. In this chapter, the goal is to link the experimental results from Chapters 2 and 3 and the computational methods from Chapters 4 and 5. Therefore, all the results carried out through to Chapter 5 will be summarized so that the knowledge learned can be leveraged and the fundamental understanding of the formation of TFA polymorphs can be gained.

6.1 Summary of the Results

TFA has a flexible structure that can adopt various conformations. TFA prefers to take a planar conformation stabilized by π-conjugation, intramolecular hydrogen bonding, and intermolecular interactions between solute and solvent molecules. The TFA molecule can also adopt a more twisted conformation that makes the carboxyl group a better hydrogen bonding acceptor. The electronic properties of the TFA molecule, derived based upon the conceptual density functional theory, provided the underlying structural cause for the carbonyl oxygen in the strained conformer to be more willing to share its electrons, and consequently to be a better hydrogen bonding acceptor. Being local properties, the DFT-based concepts describe the sensitivity of a molecular system to the electronic perturbation that may be caused by interacting with other molecules. A less pronounced distribution of the electrophilic region around the carbonyl oxygen of the twisted

molecular conformation imparts to the chemical moiety the strongest ability for the formation of hydrogen bonding interactions.

Crystal structures of the crystallized Form I and Form II comprise only one conformer per crystal lattice. Form I and II conformers of TFA were considered the two molecular conformations in solution with significant contributions, because they correspond to the local and global minima, respectively, in the conformational energy distribution. Specifically, the calculations performed in ethanol phase revealed that the conformation present in the Form I crystal leads to a local minimum that is 3.7 kJ/mol higher in energy than the global minimum. Accordingly, molecular dynamics simulations provided further evidence that the conformation found in the Form II crystal is in fact more stable. As such, the population distribution of Form I and II conformers, which is a function of the conformational energy associated with each conformer, should favor a higher proportion of these two conformers with respect to other molecular conformations that may exist in solution.

In a polar solvent and in the absence of solute-solute interactions, it might be expected to see TFA molecules in the stable Form II conformation. Particularly when monomers are the dominant solute species, the TFA molecule might assume the most stable conformation that is overall flatter and that resembles the conformation found in Form II. However, the ensemble of TFA conformers present in ethanol solution was characterized by a short distance between the methyl group on the chlorinated aromatic ring and the proton at the *meta* position on the anthranilic ring. Interestingly, a more twisted conformation gained greater prominence at lower temperature when self-association of solute molecules also became more favorable. Either a higher solute concentration or a

lower temperature enhanced solute-solute interactions and, thus, facilitated the formation of molecular association. Intermolecular interaction calculations supported the conclusion that hydrogen bonding interactions provided the driving force for the observed TFA association and that hydrogen-bonded dimers were the favored association mode of interacting molecules. In fact, the interaction energies involved in aromatic association were smaller than those involved in carboxylic acid association.

It is thought that a molecule crystallizes in a conformation close to the global minimum of its potential energy surface. However, molecular conformations that do not correspond to a global energy minimum are observed in the solid state. Studies[81-83] have shown that a molecule can change its conformation to adopt a conformation suitable for association and crystal packing before integrating into the crystal surface. Such a conformational change is related to intermolecular interactions.[79] Biphenyls constitute a class of molecules known to be affected by systematic crystal packing effects.[228] The biphenyl molecule without *ortho* substituents assumes a twisted conformation in the gas phase and in solution, but adopts a more planar conformation in its single known polymorph. Favorable packing arrangements are often associated with increased density, so a change in torsion angle values and, thus, in the conformation of the biphenyl molecule might be correlated with packing efficiency. In other words, a more planar conformation packs more densely than a twisted conformation. The observation is further supported by acene derivatives.[229] Twisted conformations of acene molecules were found to correspond to potential energy minima and be more stable than untwisted conformers, which nevertheless crystallized preferentially and formed more favorable packing interactions in the solid-state. The studies reveal that the distribution of conformers found in crystals

may be different from the distribution expected from the intramolecular potential energy of conformational space. Therefore, crystallizations systematically favor conformations that would be rather improbable in the less-condensed phase.

In concentrated ethanol solutions of TFA, in which solute-solute hydrogen bonding interactions become important, the Form II conformer would form weaker hydrogen-bonded dimers. The hydrogen bonding in Form II is about 2 kJ/mol weaker than that in Form I. However, the stabilization by specific hydrogen bonding interactions can override conformational preferences of a molecule.[146] As such, when solute contacts get strengthened and hydrogen-bonded dimers form, the TFA molecule adopts a conformation that is suitable to form stronger hydrogen bonds.

The conformation of the molecule is susceptible to solution conditions and affects the overall energy state of the system. The planar conformation that reflects the conformation found in the Form II crystal favors intramolecular hydrogen bonding and lowers the interaction energy for the solute with the solvent. The Form II conformer is more stabilized than the Form I conformer by 3.7 kJ/mol, as calculated in ethanol phase. The twisted conformation similar to the conformation found in the Form I crystal allows for strong intermolecular hydrogen bonding. Thus, hydrogen bonding intermolecular interactions can stabilize a conformation that is 3.7 kJ/mol above the lowest energy conformation. Further, the twisted molecular conformation tends to minimize the energy of the dimer state. The Form I dimer is more stabilized than the hydrogen-bonded dimer found in Form II by about 2.5 kJ/mol, as computed in ethanol phase.

The ultimate goal of this research is to investigate whether solution conditions, which influence the molecular conformation occurring in the self-assembled aggregates in

solution, in turn, affect the molecular conformation in the crystalline state, so as to form distinct polymorph structures. Therefore, the results obtained from solution chemistry are compared to the results from crystallization experiments in order to rationalize the polymorph formation of TFA.

Cooling TFA solutions at 40 and 62 mM in ethanol led to the production of Form I and II mixtures, containing approximately 13 percent (w/w) and 18 percent (w/w) of Form I, respectively (Figure 2.14). There was a slight increase of Form I in the composition of polymorphic mixtures as solute concentration increased. Based on NOESY measurements, in solutions at concentrations similar to those in the crystallization experiments and at the lowest analyzed temperature the distribution of conformers favors the twisted Form I conformation, which has a shorter distance between the methyl group on the chlorinated aromatic ring and the proton at the *meta* position on the anthranilic ring (Figure 3.16 and Table 3.4). The internal dynamics corresponding to rapid association/dissociation, as indicated by single peaks with averaged ^{1}H-NMR chemical shifts, and the fast exchange between molecular conformations in solution indicate that the estimated values of proton distances are time-averaged over the possible conformations. However, the results indicate a correlation between the molecular conformation population present in solution and that found in the solid-state. The population of the Form I conformer in solution may reflect the contribution of this conformer to the resultant crystalline polymorph.

The percentage of Form I obtained from samples crystallized at 37 °C was less comparable to the trend of the NOESY measurements obtained at the same temperature. Although the initial concentrations of the crystallization experiments performed at 37 °C

were beyond the concentration range employed in NOESY measurements, it is difficult to explain the dramatic increase up to 50 percent (w/w) of Form I crystals obtained at S = 1.95 (Figure 2.9). One possible reason would arise from faster equilibration rates between conformers at elevated temperatures.[230] If the population of the Form I conformer increases due to a faster equilibration rate, then crystallization of Form I would be expected to be accomplished in higher amounts. In order to estimate the influence of the population of the Form I conformer on crystallization with increasing temperature, the relative proton distances were determined at high temperatures. The trend of the NOESY data at 37 °C indicated a larger intramolecular distance between protons of interest in the population of monomers compared to the distances obtained at low temperatures. Therefore, the large increase of Form I in the solid-state was unexpected. A plausible reason for the large percentage of Form I crystals from solution at high supersaturation may be due to the faster nucleation and crystal growth achieved under these conditions. Thus, any precipitation observed should be considered in the evaluation that the crystal growth had already ensued.

6.2 Proposed Nucleation Mechanism

Given the results of the crystallization experiments, solute species in solution, and intermolecular interaction calculations, the proposed mechanism implies that, when a solution of low concentration gets cooled and becomes supersaturated, the solute monomers start to form nuclei. As the concentration increases, more self-association is warranted and more hydrogen-bonded dimers form in solution. For a thermodynamic reason, solute molecules need to form dimers as strongly as possible and, in the case of

TFA, the molecules have to adjust their monomer conformation in order to form stronger hydrogen bonds (*ca.* 5 kJ/mol increase as calculated in explicit ethanol phase). The energy penalty due to the conformational change (*ca.* 3.7 kJ/mol as calculated in explicit ethanol phase) can be recovered from gaining stronger hydrogen bonds. When the condition of nucleation is met (i.e., an undersaturated solution is cooled), molecular associated species start to form nuclei and initiate crystal growth. Thus, it can be concluded that, when a solution of TFA is rapidly cooled and reaches the nucleation conditions, the strongly interacted hydrogen-bonded dimers act as building blocks for the growth of the colorless Form I crystals. Still, the model does not consider the exclusive presence of dimers in the solution phase, but an equilibrium between the monomer and the dimer states; the nucleation and subsequent crystal growth drive that equilibrium to the product phases, giving rise to concomitant polymorphs. The key concept underpinning the argument is that solute molecules need to take on a conformation suitable for the formation of strong hydrogen-bonded dimers.

The molecular conformation is sensitive to intermolecular interactions during crystal growth. As a concentration increase or a temperature decrease warrants further self-assembly of solute molecules, the effect of solute concentration or temperature on promoting hydrogen-bonded dimer formation is manifested by the conformational change of the solute. The identified specific solute-solute interactions help elucidate the effect of solution conditions in the polymorph formation. The TFA molecule assumes a twisted conformation, which better exposes the carboxyl group in TFA for self-association and, thus, facilitates the formation of the most stable association species and the most stable crystal form. In other words, when the solvated monomers of TFA molecules start to be

convened into dimers, the molecular conformation needs to change in order to form more favorable intermolecular hydrogen bonding interactions in the solid-state. More importantly, this argument implies the utmost role of solute species in governing the polymorph formation from solution.

The association of solute species to form pre-nucleation aggregates has been highlighted as an important initial step in the crystallization process. In addition, the focus of this dissertation has been devoted to the molecular conformation in the state of association of solute molecules in solution and in the final crystalline state. The criterion selected for investigating the relationship between solution conformation and the solid-state outcome was to match the trend of the ensemble of conformers in the solution phase with the fraction of the crystallized polymorph containing that conformer in the crystal lattice. There was a qualitative correspondence between the twisted conformation existing in solution and in the crystal structure. Thus, it might seem reasonable to conclude that the molecular conformation of TFA in solution reflects that in the solid-state. However, differences were observed between NMR measurements and crystallization outcome. Concomitant polymorphs crystallized under the same solution conditions. As a result, a direct correlation between pre-nucleation aggregates and polymorphic solid forms fell short.

In the above analysis, the states of association in the solution phase are considered to be carried out in the crystallization process. The nucleation and crystal growth likely proceed via incorporation of the pre-nucleation aggregates, which serve as growth units, into the crystal structure (Figure 6.1). However, solute molecules do not assume static, fixed configurations and/or arrangements. The fundamental molecular steps that are

associated with nucleation and subsequent growth of organic crystals from solution are dynamic phenomena. Monomers or dimers in solution start to organize themselves into larger associated species. As the pre-nucleation association species grow in size to form critical nuclei, it is necessary to keep in mind that structural changes can still take place during the pathway to a macroscopic crystal. Therefore, rationalizing the polymorph formation based on pre-nucleation association alone may represent an oversimplification of what is really occurring in the crystallization process.

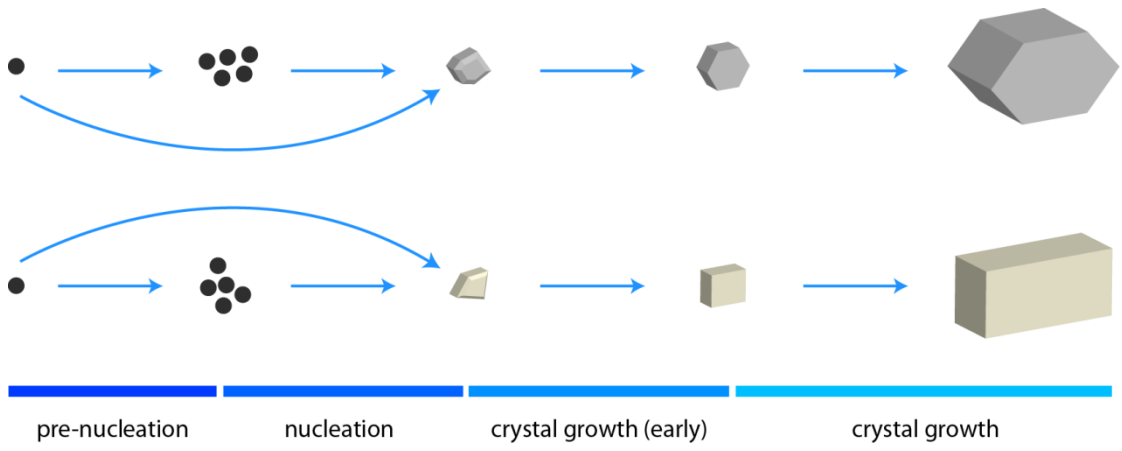

Figure 6.1. Schematic of the nucleation and growth of a polymorphic system, involving the initial association in solution to form precursors (i.e., growth units), followed by formation of nuclei. Nuclei, which attain the critical size, continue to grow until a final crystal is achieved.

From an experimental viewpoint, it would be desirable to be able to follow the evolution of solute species from the rearrangement of pre-nucleation aggregates to the formation of crystalline nuclei that continue to grow. There clearly is a challenge to develop methodologies that will enable the gap between the dynamics of solution entities and the formation of crystals to be bridged. The current methods are limited by their dependence

upon concepts emerging from a molecular prospective. While a direct correlation between pre-nucleation solute species and the final crystallized polymorph was difficult to make in some cases, there is confidence that the study of the conformational preferences of the solute species in solution helped advance the fundamental understanding of the underlying relationship between molecular conformation and crystal packing of TFA conformational polymorphs. Overall, this knowledge could serve as a guide to pharmaceutical scientists for rational control of crystal growth of organic molecules.

Chapter 7 – Conclusion and Future Directions

Studying the nucleation process, which is the earliest stage of crystallization and a primary determinant of the structure and properties of the resultant product, represents a major challenge in polymorphism research. The essential difficulty is that the processes taking place are barely, or not at all, accessible to current experimental methods. The key focus of this dissertation has been a detailed analysis of the solute species in solution at the pre-nucleation stage. The extent to which solution conditions influenced the self-association and the molecular conformation of the model, conformationally flexible solute was rigorously examined at a molecular level. The solution NMR methodology was applied to detect and quantify the presence of associated species and of the equilibrium conformer. The NMR measurements helped gain critical knowledge about the effect of solution conditions on the solute species and on the time-averaged molecular conformation adopted by the solute species. Upon change of solution conditions, TFA molecules self-associated by forming hydrogen-bonded dimers and adopted a twisted conformation. A mathematical model was applied to verify the correlation between the fraction of associated species and the population distribution of conformers.

With this information in hand, the structural cause for the observed conformational distribution in solution was determined by utilizing electronic calculation. Further, intermolecular interactions of the molecular associated species (i.e., dimers), resembling the structural motifs in TFA crystal structures were calculated in the gas phase, as well as in the solution phase. The methodology in a solvent modeled explicitly was developed so there would be good control of the solvent effect on the intermolecular interaction.

Crystallization, which was affected by solution conditions, resulted in the selection of one conformer to be present in the crystal structure. Possible structural changes during the early stage of crystal growth cannot be ruled out. As such, consideration must be given to the studies that could further the polymorph behavior by accounting for structural changes that may occur in the stepwise self-assembly crystal growth process.

The research presented in this dissertation provides a starting point for future studies. A more systematic approach to crystal growth calls for the development of suitable quantitative and noninvasive monitoring techniques. In the pharmaceutical industry, as well as in academia, the on-line monitoring with fiber optic Raman spectroscopy and Fourier transform infrared spectroscopy has been developed. The techniques have been used to achieve effective product quality control of crystal shape and size and to ensure consistent isolation of the desired polymorph. They also have been used to monitor both the solute concentration and the form composition of the crystals throughout the whole crystallization process. As such, more rigorous information on structural changes associated with the crystal growth of polymorphs from solution could be gained with the on-line approach.

With a strong foundation of the contribution of different molecular conformations in the crystallization process, more small organic molecules can be investigated. Systems having a high energy barrier of transition between conformers can be evaluated to further test the effect of solution conditions on the molecular conformation in the solid-state of drug molecules. The studies are aimed at expanding the understanding of the relationship between the conformations of a flexible drug and crystal packing with the long-term goal

of controlling polymorphism based upon the molecular structural properties of the drug compound.

Appendices

Appendix 1. Crystal Growth Morphology

Growth morphology of a crystal is governed by the relative growth rates of its faces. The general rule is that faces that grow fast will become the minor faces. Conversely, faces that grow slow will develop as the major faces.

A1.1 Methods

The morphology of TFA Forms I and II was generated by the Materials Studio 3.5 program (Accelrys, Inc., San Diego, CA) using the single-crystal X-ray data.[114] The crystal habit faces expected to be dominant were identified from the crystal lattice geometry, using the Bravais Friedel Donnay Harker (BFDH) method.[231] The Miller indices of two faces of TFA Form I single-crystals were verified using a powder X-ray diffractometer (Multiflex, Rigaku Co., The Woodlands, TX) with a Ni-filtered Cu Kα radiation (λ = 1.54178 Å). Samples were scanned over a 2θ range of 5 - 30° at the rate of 0.5 °/min with step size of 0.02°.

A1.2 Results

The growth morphology analysis of TFA Form I reveals that the (010) face has the largest morphological importance and it thereby represents the largest part of the total surface area of the crystal. A tiny face for this polymorph of TFA is the (100) face, which could not be measured experimentally (Figure A1.1a). Note that the method may not be totally reliable. PXRD patterns confirm that the predominant face of the Form I crystal is

identified as (010) and a smaller face is (001) (Figure A1.1b). These results demonstrate that the crystal growth is slower along the direction of *b*-axis perpendicular to the (100) face. Conversely, the (001) face grows fast so that it is not a major face and does not appear in the morphology prediction. The hydrogen-bonded dimers form a two-dimensional sheet in the *ac*-plane. The growth morphology results of TFA Form II demonstrate that the largest growth face is (010), as shown in Figure A1.2.

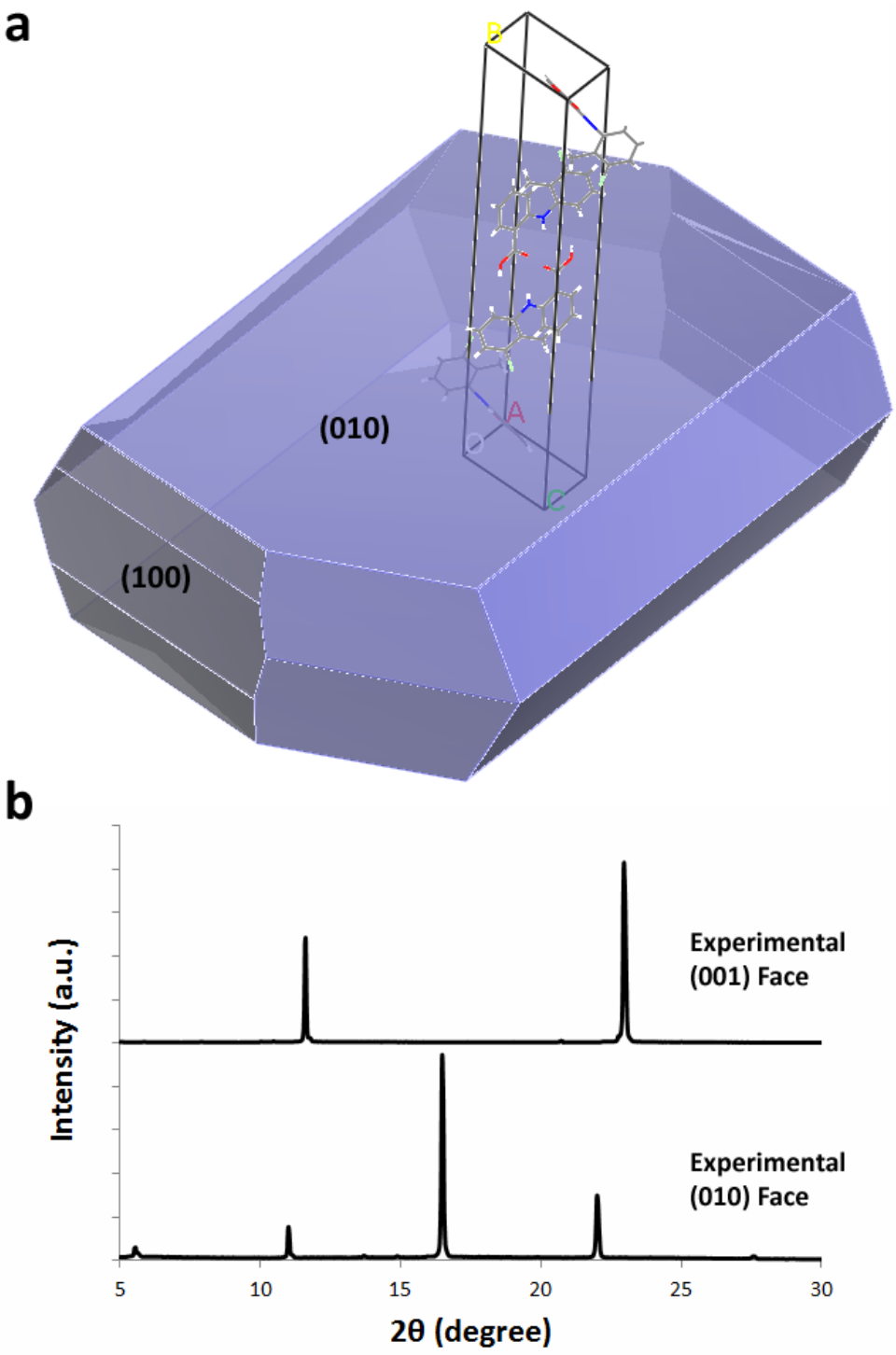

Figure A1.1. Simulated growth morphology and crystal structure of TFA Form I. Crystal face indexing is highlighted (a). Powder X-ray diffraction patterns of the (001) and (010) faces of TFA Form I (b).

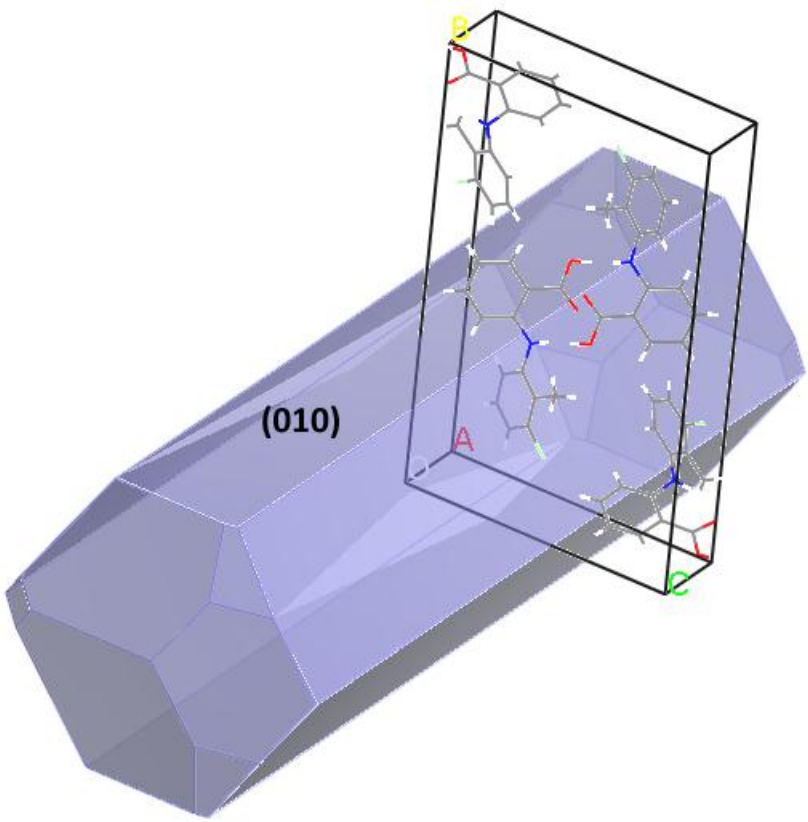

Figure A1.2. Simulated growth morphology and crystal structure of TFA Form II. Crystal face indexing of the predominant (010) face is highlighted.

Appendix 2. NMR Spectroscopy

A2.1 PGSE NMR Method

NMR exploits magnetic field gradients to measure the self-diffusion coefficient. Two magnetic field gradient pulses are employed and they are essential for correlating the translational motion of nuclei to the NMR signal intensity. Figure A2.1 shows a representation of how PGSE sequence measures diffusion. A 90° excitation pulse (pw) brings the magnetization into the XY plane. A field gradient (gt) is applied for a short time (δ=2 ms) so that nuclei experience a phase shift. A pair of 90° pulses can return the magnetization to the XY plane. Thus, the second field gradient can cancel the dephasing produced by the first gradient and refocus all nuclei, provided that no change of position has occurred. If diffusion occurs, the refocusing is not complete. This results in a remaining dephasing, which is proportional to the displacement during the diffusion time. In Figure A2.1 d0 – d5 are acquisition delays. Between the two field gradients the diffusion time is δ+d2+2pw90+d5+d0 during which molecules diffuse.

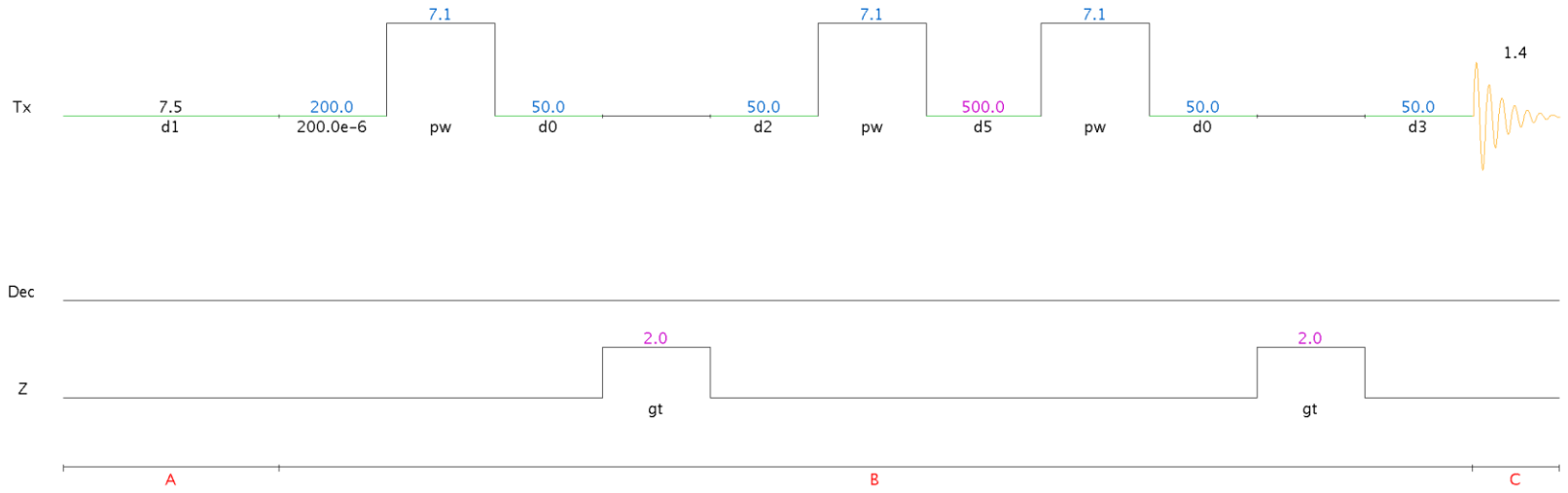

Figure A2.1. Pulse sequence used in PGSE NMR experiments.

A2.2 NOESY Method

The NOESY experiment identifies resonances of protons 4 Å or less from one another. The sequence of NOESY consists of three 90° pulses. The first pulse creates transverse magnetization during the evolution time (d2). The second pulse produces longitudinal magnetization equal to the transverse magnetization component, orthogonal to the pulse direction. The mixing time (τ_m) is kept constant throughout the experiment. The third pulse creates the observable transverse magnetization from the remaining longitudinal magnetization (Figure A2.2). Note a 90° pulse at the beginning of the sequence. This serves to eliminate any net magnetization from the preceding iteration of the sequence and to start each scan with the same amount of magnetization.

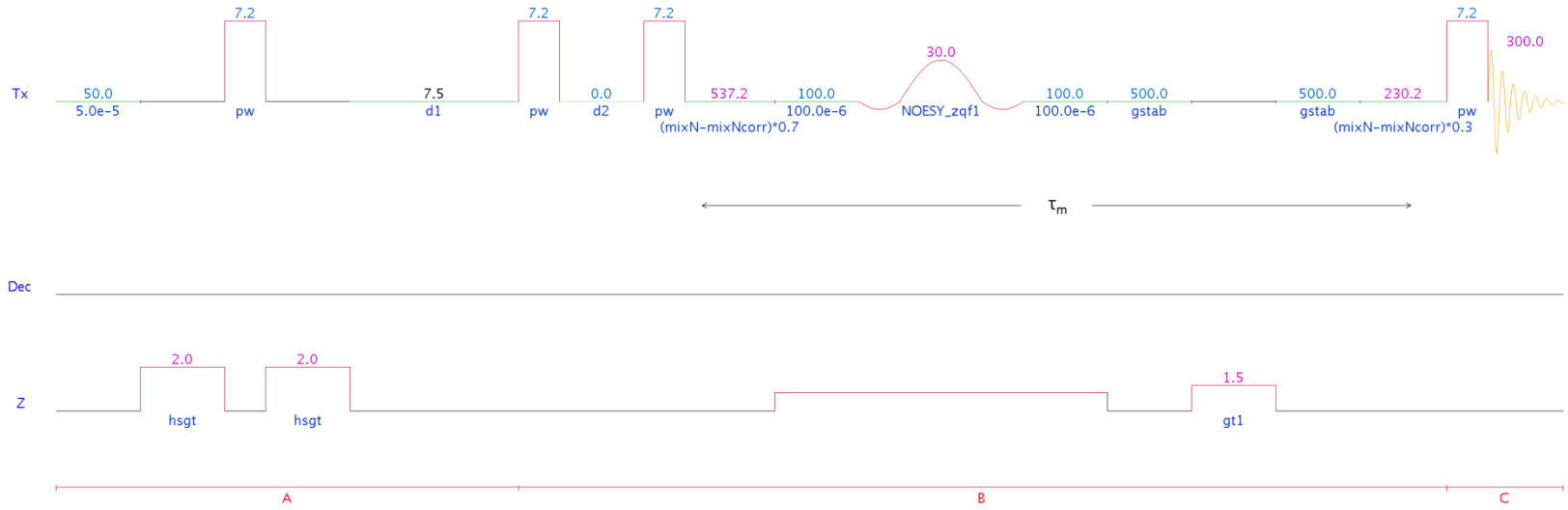

Figure A2.2. Pulse sequence used in 2D-NOESY experiments.

A2.3 Spin-Echo Signal Attenuation Measurements

Information on translational diffusion is encoded by varying the gradient strength for several values. The signal is detected as a free-induction decay, and Fourier transform with respect to time yields a spectrum with the peak intensities (Figure A2.3a). All of the resonances that belong to a pure component will decay exponentially at the same rate with respect to the square of the gradient strength. The signal decay for the nuclei of the drug was ideal (mono-exponential) and gave a good fit to Stejskal-Tanner exponential equation (Figure A2.3b).

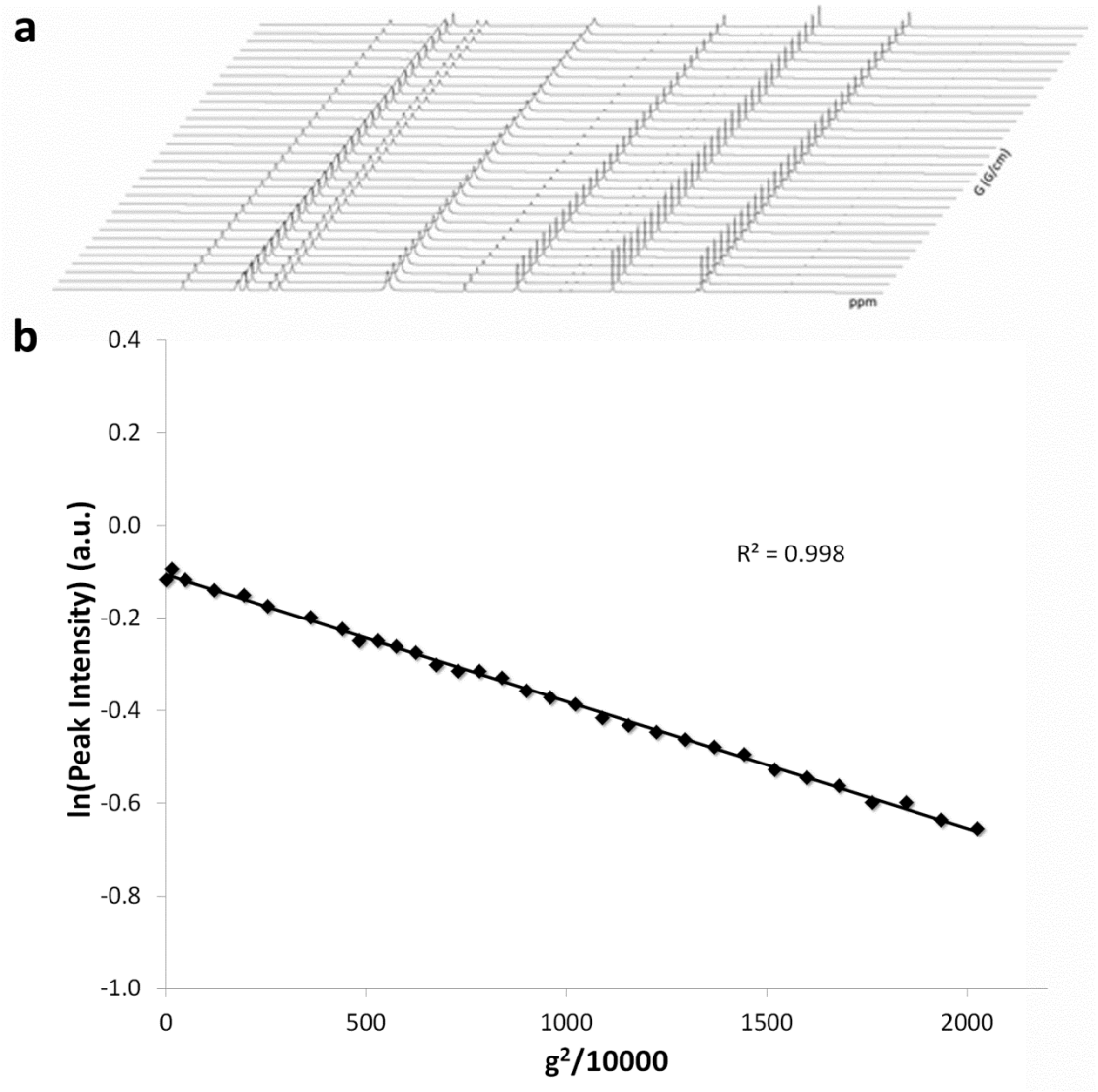

Figure A2.3. Attenuation of the signals with increasing gradient strength when performing PGSE NMR experiments (a) and the representative Stejskal-Tanner exponential plot from the decay of H_{30} peak (b).

Appendix 3. Molecular Dynamics Code

```
#Minimization
&cntrl
nmropt=1,ipnlty=1,
imin = 1,
ntr = 0,
maxcyc = 5000,
ncyc = 2000,
ntpr = 100,
cut = 10.0,
ntb = 1
/
&wt type='REST',istep1=1,istep2=5000,
   value1=1.0, value2=1.0,&end
  &wt type='END', &end
  LISTIN=torsion_mi0.listin
  LISTOUT=torsion_mi0.listout
  DISANG=torsion.f

#MD Heating Step from 0 to 298K
&cntrl
nmropt=1,ipnlty=1,
imin=0,
irest=0,
ntx=1,
ntb=1,
ntr=0,
nrespa=1,
cut=10,
ibelly=0,
ntc=2,ntf=2,
tempi=0,temp0=298,
ntt=3,gamma_ln=1.0,
nstlim=20000,dt=0.002,
ntpr=500,ntwx=500,ntwr=500
/
&wt type='REST',istep1=1,istep2=20000,
   value1=1.0, value2=1.0,&end
  &wt type='END', &end
  LISTIN=torsion_he0.listin
  LISTOUT=torsion_he0.listout
  DISANG=torsion.f
```

#MD Equilibration Step at 298k

```
&cntrl
nmropt=1,ipnlty=1,
imin=0,
irest=1,
ntx=5,
ntb=2,ntp=1,
ntr=0,
nrespa=1,
cut=10,
ibelly=0,
ntc=2,ntf=2,
tempi=298,
temp0=298,ntt=3,gamma_ln=1.0,
nstlim=500000,dt=0.002,
ntpr=500,ntwx=500,ntwr=500
/
&wt type='REST',istep1=1,istep2=500000,
   value1=1.0, value2=1.0,&end
  &wt type='END', &end
  LISTIN=torsion_eq0.listin
  LISTOUT=torsion_eq0.listout
  DISANG=torsion.f
```

References

1. Byrn SR, Pfeiffer RR, Stowell JG. 1999. Solid State Chemistry of Drugs. Second ed., West Lafayette, IN: SSCI, Inc.

2. Qiu Y, Chen Y, Zhang GGZ. 2009. Developing Solid Dosage Oral Forms: Pharmaceutical Theory and Practice. First ed., Burlington, MA: Academic Press.

3. Bernstein J. 2002. Polymorphism in Molecular Crystals. New York: Oxford University Press.

4. Mitscherlich E 1822. Sur la rélation qui existe entre la forme cristalline et les proportions chimique, I. Mémoire sur les arséniates et les phosphates. Annales de Chimie et de Physique 19:350-419.

5. Haleblia.J, McCrone W 1969. Pharmaceutical applications of polymorphism. Journal of Pharmaceutical Sciences 58(8):911-929.

6. Grant DJW. 1999. Theory and origin of polymorphism. In: Brittain HG. Polymorphism in Pharmaceutical Solids. Second ed., New York: Marcel Dekker. p 1-33.

7. Peschar R, Pop MM, De Ridder DJA, van Mechelen JB, Driessen RAJ, Schenk H 2004. Crystal structures of 1,3-distearoyl-2-oleoylglycerol and cocoa butter in the beta(V) phase reveal the driving force behind the occurrence of fat bloom on chocolate. Journal of Physical Chemistry B 108(40):15450-15453.

8. Hilfiker R. 2006. Polymorphism: in the Pharmaceutical Industry. Weinheim, Germany: Wiley-VCH.

9. Stahly GP 2007. Diversity in single- and multiple-component crystals. The search for and prevalence of polymorphs and cocrystals. Crystal Growth & Design 7(6):1007-1026.

10. McCrone WC. 1965. Polymorphism. In: Fox D, Labes MM, Weissberger A, editors. Physics and Chemistry of the Organic Solid State. New York: Interscience Publishers. p 725-767.

11. Dudognon E, Danede F, Descamps M, Correia NT 2008. Evidence for a new crystalline phase of racemic ibuprofen. Pharmaceutical Research 25(12):2853-2858.

12. Brittain HG. 2009. Polymorphism in Pharmaceutical Solids. Second ed., New York, NY: Informa Healthcare USA, Inc.

13. Suryanarayanan R, Byrn SR 2001. Characterization of the solid state. Advanced Drug Delivery Reviews 48(1):1-136.

14. Raw AS, Yu LX 2004. Pharmaceutical solid polymorphism in drug development and regulation. Advanced Drug Delivery Reviews 56(3):235-414.

15. Rogers RD 2004. Polymorphism in crystals. Crystal Growth & Design 4(6):1085-1444.

16. Matzger AJ 2008. Facets of polymorphism in crystals. Crystal Growth & Design 8(1):2-161.

17. Price SL 2004. The computational prediction of pharmaceutical crystal structures and polymorphism. Advanced Drug Delivery Reviews 56(3):301-319.

18. Dunitz JD 2003. Are crystal structures predictable? Chemical Communications (5):545-548.

19. Nangia A 2008. Conformational polymorphism in organic crystals. Accounts of Chemical Research 41(5):595-604.

20. Hollingsworth MD 2002. Crystal engineering: From structure to function. Science 295(5564):2410-2413.

21. Yu L 2010. Polymorphism in molecular solids: An extraordinary system of red, orange, and yellow crystals. Accounts of Chemical Research 43(9):1257-1266.

22. Dunitz JD, Bernstein J 1995. Disappearing polymorphs. Accounts of Chemical Research 28(4):193-200.

23. Henck JO, Bernstein J, Ellern A, Boese R 2001. Disappearing and reappearing polymorphs. The benzocaine: Picric acid system. Journal of the American Chemical Society 123(9):1834-1841.

24. Barsky I, Bernstein J, Stephens PW, Stone KH, Cheung E, Hickey MB, Henck J-O 2008. Disappearing and reappearing polymorphism in p-methylchalcone. Crystal Growth & Design 8(1):63-70.

25. Kee NCS, Tan RBH, Braatz RD 2009. Selective crystallization of the metastable α-form of L-glutamic acid using concentration feedback control. Crystal Growth & Design 9(7):3044-3051.

26. Cote A, Zhou G, Stanik M 2009. A novel crystallization methodology to ensure isolation of the most stable crystal form. Organic Process Research & Development 13(6):1276-1283.

27. Bernstein J, Davey RJ, Henck JO 1999. Concomitant polymorphs. Angewandte Chemie-International Edition 38(23):3441-3461.

28. Yu L 1995. Inferring thermodynamic stability relationship of polymorphs from melting data. Journal of Pharmaceutical Sciences 84(8):966-974.

29. Yu L, Reutzel SM, Stephenson GA 1998. Physical characterization of polymorphic drugs: An integrated characterization strategy. Pharmaceutical Science & Technology Today 1(3):118-127.

30. Burger A, Ramberger R 1979. On the polymorphism of pharmaceuticals and other molecular crystals. I. Theory of thermodynamic rules. Mikrochim Acta [Wien] 72(3-4):259-271.

31. Burger A, Ramberger R 1979. On the polymorphism of pharmaceuticals and other molecular crystals. II. Applicability of thermodynamic rules. Mikrochim Acta [Wien] 72(3-4):273-316.

32. Paul IC, Curtin DY 1973. Thermally induced organic reactions in the solid state. Accounts of Chemical Research 6(7):217-225.

33. Gu CH, Young V, Grant DJW 2001. Polymorph screening: Influence of solvents on the rate of solvent-mediated polymorphic transformation. Journal of Pharmaceutical Sciences 90(11):1878-1890.

34. Zhang GGZ, Gu CH, Zell MT, Burkhardt RT, Munson EJ, Grant DJW 2002. Crystallization and transitions of sulfamerazine polymorphs. Journal of Pharmaceutical Sciences 91(4):1089-1100.

35. Cardew PT, Davey RJ 1985. The kinetics of solvent-mediated phase-transformations. Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences 398(1815):415-428.

36. Ostwald W 1897. Studien uber die Bildung und Unwandlung fester Korper. Z Phys Chem (Liepzig) 22:289-330.

37. Higuchi WI, Lau PK, Higuchi T, Shell JW 1963. Polymorphism and drug availability - Solubility relationships in methylprednisolone system. Journal of Pharmaceutical Sciences 52(2):150-&.

38. Aguiar AJ, Krc J, Kinkel AW, Samyn JC 1967. Effect of polymorphism on the absorption of chloramphenicol from chloramphenicol palmitate. Journal of Pharmaceutical Sciences 56(7):847-883.

39. Chemburkar SR, Bauer J, Deming K, Spiwek H, Patel K, Morris J, Henry R, Spanton S, Dziki W, Porter W, Quick J, Bauer P, Donaubauer J, Narayanan BA, Soldani M, Riley D, McFarland K 2000. Dealing with the impact of ritonavir polymorphs on the late stages of bulk drug process development. Organic Process Research & Development 4(5):413-417.

40. Raw AS, Furness MS, Gill DS, Adams RC, Holcombe FO, Yu LX 2004. Regulatory considerations of pharmaceutical solid polymorphism in abbreviated new drug applications (ANDAs). Advanced Drug Delivery Reviews 56(3):397-414.

41. Zhang GGZ, Law D, Schmitt EA, Qiu Y 2004. Phase transformation considerations during process development and manufacture of solid oral dosage forms. Advanced Drug Delivery Reviews 56(3):371-390.

42. Bernstein J 2011. Polymorphism - A perspective. Crystal Growth & Design 11(3):632-650.

43. Mullin JW. 2001. Crystallization, Fourth. ed., Oxford: Butterworth-Heinemann.

44. Etter MC 1991. Hydrogen-bonds as design elements in organic chemistry. Journal of Physical Chemistry 95(12):4601-4610.

45. Weissbuch I, Lahav M, Leiserowitz L 2003. Toward stereochemical control, monitoring, and understanding of crystal nucleation. Crystal Growth & Design 3(2):125-150.

46. Bernstein J, Etter MC, Leiserowitz L. 1994. Structure Correlations. In: Dunitz JD, Bürgi H-B, Eds. Weinheim: VCH.

47. Becker R, Döring W 1935. Kinetische behandlung der keimbildung in übersättingten dämpfen. Annals of Physics 24:719.

48. Volmer M. 1939. Kinetik der phasenbildung. ed., Steinkopf, Dresden.

49. Kashchiev D, van Rosmalen GM 2003. Review: Nucleation in solutions revisited. Crystal Research and Technology 38(7-8):555-574.

50. Boistelle R, Astier JP 1988. Crystallization mechanism in solution. Journal of Crystal Growth 90(1-3):14-30.

51. Oxtoby DW 2003. Crystal nucleation in simple and complex fluids. Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences 361(1804):419-427.

52. Haisa M, Kashino S, Kawai R, Maeda H 1976. The monoclinic form of *p*-hydroxyacetanilide. Acta Crystallographica Section B-Structural Science 32(4):1283-1285.

53. Lee EH, Byrn SR, Carvajal MT 2006. Additive-induced metastable single crystal of mefenamic acid. Pharmaceutical Research 23(10):2375-2380.

54. Davey RJ, Allen K, Blagden N, Cross WI, Lieberman HF, Quayle MJ, Righini S, Seton L, Tiddy GJT 2002. Crystal engineering - nucleation, the key step. CrystEngComm 4(47):257-264.

55. Vekilov PG 2004. Dense liquid precursor for the nucleation of ordered solid phases from solution. Crystal Growth & Design 4(4):671-685.

56. Vekilov PG 2010. Nucleation. Crystal Growth & Design 10(12):5007-5019.

57. Gavezzotti A 1999. Molecular aggregation of acetic acid in a carbon tetrachloride solution: A molecular dynamics study with a view to crystal nucleation. Chemistry-a European Journal 5(2):567-576.

58. Shore JD, Perchak D, Shnidman Y 2000. Simulations of the nucleation of AgBr from solution. Journal of Chemical Physics 113(15):6276-6284.

59. Galkin O, Pan W, Filobelo L, Hirsch RE, Nagel RL, Vekilov PG 2007. Two-step mechanism of homogeneous nucleation of sickle cell hemoglobin polymers. Biophysical Journal 93(3):902-913.

60. Erdemir D, Lee AY, Myerson AS 2009. Nucleation of crystals from solution: Classical and two-step models. Accounts of Chemical Research 42(5):621-629.

61. Asherie N, Lomakin A, Benedek GB 1996. Phase diagram of colloidal solutions. Physical Review Letters 77(23):4832-4835.

62. Sommerdijk NAJM, de With G 2008. Biomimetic CaCO3 Mineralization using Designer Molecules and Interfaces. Chemical Reviews 108(11):4499-4550.

63. Pouget EM, Bomans PHH, Goos J, Frederik PM, de With G, Sommerdijk NAJM 2009. The initial stages of template-controlled $CaCO_3$ formation revealed by cryo-TEM. Science 323(5920):1455-1458.

64. Garetz BA, Matic J, Myerson AS 2002. Polarization switching of crystal structure in the nonphotochemical light-induced nucleation of supersaturated aqueous glycine solutions. Physical Review Letters 89(17).

65. Pan WC, Kolomeisky AB, Vekilov PG 2005. Nucleation of ordered solid phases of proteins via a disordered high-density state: Phenomenological approach. Journal of Chemical Physics 122(17):174905-174907.

66. Oxtoby DW 1998. Nucleation of first-order phase transitions. Accounts of Chemical Research 31(2):91-97.

67. Lutsko JF, Nicolis G 2006. Theoretical evidence for a dense fluid precursor to crystallization. Physical Review Letters 94:046102.

68. Galkin O, Vekilov PG 1999. Direct determination of the nucleation rates of protein crystals. Journal of Physical Chemistry B 103(49):10965-10971.

69. Chen J, Sarma B, Evans JMB, Myerson AS 2011. Pharmaceutical Crystallization. Crystal Growth & Design 11(4):887-895.

70. Davey RJ, Blagden N, Righini S, Alison H, Quayle MJ, Fuller S 2001. Crystal polymorphism as a probe for molecular self-assembly during nucleation from solutions: The case of 2,6-dihydroxybenzoic acid. Crystal Growth & Design 1(1):59-65.

71. Parveen S, Davey RJ, Dent G, Pritchard RG 2005. Linking solution chemistry to crystal nucleation: the case of tetrolic acid. Chemical Communications (12):1531-1533.

72. Chen J, Trout BL 2008. Computational study of solvent effects on the molecular self-assembly of tetrolic acid in solution and implications for the polymorph formed from crystallization. Journal of Physical Chemistry B 112(26):7794-7802.

73. Hamad S, Moon C, Catlow CRA, Hulme AT, Price SL 2006. Kinetic insights into the role of the solvent in the polymorphism of 5-fluorouracil from molecular dynamics simulations. Journal of Physical Chemistry B 110(7):3323-3329.

74. Myerson AS, Lo PY 1991. Cluster formation and diffusion in supersaturated binary and ternary amino-acid solutions. Journal of Crystal Growth 110(1-2):26-33.

75. Erdemir D, Chattopadhyay S, Guo L, Ilavsky J, Amenitsch H, Segre CU, Myerson AS 2007. Relationship between self-association of glycine molecules in supersaturated solutions and solid state outcome. Physical Review Letters 99(11): 115702-115704.

76. Yu L, Kingman N 2002. Glycine crystallization during spray drying: The pH effect on salt and polymorphic forms. Journal of Pharmaceutical Sciences 91(11):2367-2375.

77. Gidalevitz D, Feidenhansl R, Matlis S, Smilgies DM, Christensen MJ, Leiserowitz L 1997. Monitoring in situ growth and dissolution of molecular crystals: Towards determination of the growth units. Angewandte Chemie-International Edition in English 36(9):955-959.

78. Huang J, Stringfellow TC, Yu L 2008. Glycine exists mainly as monomers, not dimers, in supersaturated aqueous solutions: Implications for understanding its

crystallization and polymorphism. Journal of the American Chemical Society 130(42):13973-13980.

79. Bernstein J, Hagler AT 1978. Conformational polymorphism. The influence of crystal structure on molecular conformation. Journal of the American Chemical Society 100(3):673-681.

80. Saito A, Igarashi K, Azuma M, Ooshima H 2002. Aggregation of p-acetanisidide molecules in the under- and super-saturated solution and its effect on crystallization. Journal of Chemical Engineering of Japan 35(11):1133-1139.

81. Byrn SR, Graber CW, Midland SL 1976. Comparison of the solid and solution conformations of methapyriline, tripelennamine, diphenydramine, histamine, and choline. The infrared-X-ray method for determination of solution conformations. Journal of Organic Chemistry 41(13):2283-2288.

82. Migliaccio GP, Byrn SR 1981. Comparisons of rotamer populations of nialamide, azaperone, and chloroquine in solid-state and in solution. Journal of Pharmaceutical Sciences 70(3):284-287.

83. Byrn SR, McKenzie AT, Hassan MMA, Albadr AA 1986. Conformation of glyburide in the solid-state and in solution. Journal of Pharmaceutical Sciences 75(6):596-600.

84. Hagler AT, Bernstein J 1978. Conformational polymorphism. 2. Crystal energetics by computational substitution. Further evidence for the sensitivity of the method. Journal of the American Chemical Society 100(20):6349-6354.

85. Yu L, Reutzel-Edens SM, Mitchell CA 2000. Crystallization and polymorphism of conformationally flexible molecules: Problems, patterns, and strategies. Organic Process Research & Development 4(5):396-402.

86. Jeffrey GA, Kim HS 1970. Conformations of alditols. Carbohydrate Research 14(2):207-216.

87. Davey RJ 1976. The effect of impurity adsorption on the kinetics of crystal growth from solution. Journal of Crystal Growth 34(1):109-119.

88. Weissbuch I, Addadi L, Lahav M, Leiserowitz L 1991. Molecular recognition at crystal interfaces. Science 253(5020):637-645.

89. Davey RJ, Blagden N, Potts GD, Docherty R 1997. Polymorphism in molecular crystals: Stabilization of a metastable form by conformational mimicry. Journal of the American Chemical Society 119(7):1767-1772.

90. Singh A, Lee IS, Myerson AS 2009. Concomitant crystallization of ROY on patterned substrates: Using a high throughput method to improve the chances of crystallization of different polymorphs. Crystal Growth & Design 9(2):1182-1185.

91. Mitchell CA, Yu L, Ward MD 2001. Selective nucleation and discovery of organic polymorphs through epitaxy with single crystal substrates. Journal of the American Chemical Society 123(44):10830-10839.

92. Price CP, Grzesiak AL, Matzger AJ 2005. Crystalline polymorph selection and discovery with polymer heteronuclei. Journal of the American Chemical Society 127(15):5512-5517.

93. Hooks DE, Fritz T, Ward MD 2001. Epitaxy and molecular organization on solid substrates. Advanced Materials 13(4):227-241.

94. Hiremath R, Basile JA, Varney SW, Swift JA 2005. Controlling molecular crystal polymorphism with self-assembled monolayer templates. Journal of the American Chemical Society 127(51):18321-18327.

95. Carter PW, Ward MD 1994. Directing polymorph selectivity during nucleation of anthranilic acid on molecular substrates. Journal of the American Chemical Society 116(2):769-770.

96. Anwar J, Zahn D 2011. Uncovering molecular processes in crystal nucleation and growth by using molecular simulation. Angewandte Chemie-International Edition 50(9):1996-2013.

97. Threlfall T 2000. Crystallisation of polymorphs: Thermodynamic insight into the role of solvent. Organic Process Research & Development 4(5):384-390.

98. Lahav M, Leiserowitz L 2001. The effect of solvent on crystal growth and morphology. Chemical Engineering Science 56(7):2245-2253.

99. Long S, Parkin S, Siegler MA, Cammers A, Li T 2008. Polymorphism and phase behaviors of 2-(phenylamino)nicotinic acid. Crystal Growth & Design 8(11):4006-4013.

100. Datta S, Grant DJW 2005. Effect of supersaturation on the crystallization of phenylbutazone polymorphs. Crystal Research and Technology 40(3):233-242.

101. Kitamura M, Hironaka S 2006. Effect of temperature on antisolvent crystallization and transformation behaviors of thiazole-derivative polymorphs. Crystal Growth & Design 6(5):1214-1218.

102. Addadi L, Berkovitchyellin Z, Weissbuch I, Vanmil J, Shimon LJW, Lahav M, Leiserowitz L 1985. Growth and dissolution of organic crystals with tailor-made inhibitors - Implications in stereochemistry and materials science. Angewandte Chemie-International Edition in English 24(6):466-485.

103. Eckhardt CJ, Bernstein J 1972. Molecular conformation and electronic structure. The solid state spectrum of a planar anil. Journal of the American Chemical Society 94(9):3247-3249.

104. Etter MC 1990. Encoding and decoding hydrogen-bond patterns of organic compounds. Accounts of Chemical Research 23(4):120-126.

105. Aakeroy CB, Champness NR, Janiak C 2010. Recent advances in crystal engineering. CrystEngComm 12(1):22-43.

106. Bernstein J 1993. Crystal growth, polymorphism and structure-property relationships in organic crystals. Journal of Physics D-Applied Physics 26(8B):66-76.

107. Reeves LW 1959. Studies of hydrogen bonding in carboxylic acids. Transactions of the Faraday Society 55(10):1684-1688.

108. Gavezzotti A, Filippini G 1994. Geometry of the intermolecular X-H...Y (X, Y=N, O) hydrogen-bond and the calibration of empirical hydrogen-bond potentials. Journal of Physical Chemistry 98(18):4831-4837.

109. Bernstein J, Davis RE, Shimoni L, Chang NL 1995. Patterns in hydrogen bonding: Functionality and graph set analysis in crystals. Angewandte Chemie-International Edition in English 34(15):1555-1573.

110. Allen FH, Motherwell WDS, Raithby PR, Shields GP, Taylor R 1999. Systematic analysis of the probabilities of formation of bimolecular hydrogen-bonded ring motifs in organic crystal structures. New Journal of Chemistry 23(1):25-34.

111. Mattei A, Li T 2012. Polymorph formation and nucleation mechanism of tolfenamic acid in solution: An investigation of pre-nucleation solute association. Pharmaceutical Research 29(2):460-470.

112. Kitamura M 2002. Controlling factor of polymorphism in crystallization process. Journal of Crystal Growth 237(3):2205-2214.

113. Khoshkhoo S, Anwar J 1993. Crystallization of polymorphs: The effect of solvent. Journal of Physics D-Applied Physics 26(8B):90-93.

114. Andersen KV, Larsen S, Alhede B, Gelting N, Buchardt O 1989. Characterization of 2 polymorphic forms of tolfenamic acid, N-(2-methyl-3-chlorophenyl)anthranilic acid - their crystal structures and relative stabilities. Journal of the Chemical Society-Perkin Transactions 2(10):1443-1447.

115. Lopez-Mejias V, Kampf JW, Matzger AJ 2009. Polymer-induced heteronucleation of tolfenamic acid: Structural investigation of a pentamorph. Journal of the American Chemical Society 131(13):4554-4555.

116. Otwinowski Z, Minor W. 1997. Denzo-SMN program package. Processing of X-ray diffraction data collected in oscillation model. In: Methods in Enzymology: Macromolecular Crystallography, Part A. New York: Academic Press. p 307.

117. Sheldrick GM. 1990. SHELXTL PC. An Integrated System for Solving, Refining, and Displaying Crystal Structure from Diffraction Data. Fourth ed., Madison: Siemens Analytcal X-ray Instruments.

118. Prausnitz JM, Lichtenthaler RN, Gomes de Azevedo E. 1999. Molecular Thermodynamics of Fluid-Phase Equilibria. Third ed., Upper Saddle River, NJ: Prentice-Hall.

119. Sohnel O, Mullin JW 1988. Interpretation of crystallization induction periods. Journal of Colloid and Interface Science 123(1):43-50.

120. Towler CS, Taylor LS 2007. Spectroscopic characterization of intermolecular interactions in solution and their influence on crystallization outcome. Crystal Growth & Design 7(4):633-638.

121. Lohani S, Nesmelova IV, Suryanarayanan R, Grant DJW 2011. Spectroscopic characterization of molecular aggregates in solutions: Impact on crystallization of indomethacin polymorphs from acetonitrile and ethanol. Crystal Growth & Design 11(6):2368-2378.

122. Davey RJ, Dent G, Mughal RK, Parveen S 2006. Concerning the relationship between structural and growth synthons in crystal nucleation: Solution and crystal chemistry of carboxylic acids as revealed through IR spectroscopy. Crystal Growth & Design 6(8):1788-1796.

123. Spitaleri A, Hunter CA, McCabe JF, Packer MJ, Cockroft SL 2004. A ^{1}H-NMR study of crystal nucleation in solution. CrystEngComm 6(80):489-493.

124. Peral F, Gallego E 1995. Self-association of pyrimidine and some of its methyl derivatives in aqueous solution. Journal of Molecular Structure 372(2-3):101-112.

125. Kelly SM, Jess TJ, Price NC 2005. How to study proteins by circular dichroism. Biochimica Et Biophysica Acta-Proteins and Proteomics 1751(2):119-139.

126. Hoffman RE 2006. Standardization of chemical shifts of TMS and solvent signals in NMR solvents. Magnetic Resonance in Chemistry 44(6):606-616.

127. Tanner JE 1970. Use of stimulated echo in NMR diffusion studies. Journal of Chemical Physics 52(5):2523-2526.

128. Longsworth LG 1960. The mutual diffusion of light and heavy water. Journal of Physical Chemistry 64(12):1914-1917.

129. Antalek B 2002. Using pulsed gradient spin echo NMR for chemical mixture analysis: How to obtain optimum results. Concepts in Magnetic Resonance 14(4):225-258.

130. Cabrita EJ, Berger S 2001. DOSY studies of hydrogen bond association: Tetramethylsilane as a reference compound for diffusion studies. Magnetic Resonance in Chemistry 39(S1):S142-S148.

131. Brand T, Cabrita EJ, Berger S 2005. Intermolecular interaction as investigated by NOE and diffusion studies. Progress in Nuclear Magnetic Resonance Spectroscopy 46(4):159-196.

132. Claridge TDW. 1999. High-resolution NMR Techniques in Organic Chemistry. New York: Elsevier.

133. Rao CNR. 1961. Ultra-Violet and Visible Spectroscopy. London: Butterworths.

134. Veselkov DA, Lantushenko AO, Davies DB, Veselkov AN 2002. The self-association of antibiotic actinocyl-bis(3-dimethylaminopropylamine) in aqueous solution: A H-1 NMR analysis. Russian Journal of Bioorganic Chemistry 28(4):342-347.

135. Macchioni A, Romani A, Zuccaccia C, Guglielmetti G, Querci C 2003. Experimental evidence for the aggregation of (Phen)(2)Pd-2(mu-H)(mu-CO) (+) in solution. Organometallics 22(7):1526-1533.

136. Keinan E, Seth KK, Sahai M, Berman E 1986. Proton NMR studies of self-association in the civet constituent (+)-(S,S)-(cis-β-methyltetrahydropyran-2-yl)-acetic acid. Journal of Organic Chemistry 51(22):4288-4291.

137. Haigh CW, Mallion RB 1979. Ring current theories in nuclear magnetic resonance. Progress in Nuclear Magnetic Resonance Spectroscopy 13:303-344.

138. Easteal AJ 1983. Viscosity and tracer diffusion coefficients, and thermodynamic properties for ethanol + acetonitrile mixtures at 298.15 K. Australian Journal of Chemistry 36(4):665-671.

139. Chen AD, Johnson CS, Lin M, Shapiro MJ 1998. Chemical exchange in diffusion NMR experiments. Journal of the American Chemical Society 120(35):9094-9095.

140. Stokes RH 1965. Tracer diffusion in binary solutions subject to a dimerization equilibrium. Journal of Physical Chemistry 69(11):4012-4017.

141. Price WS, Tsuchiya F, Arata Y 1999. Lysozyme aggregation and solution properties studied using PGSE NMR diffusion measurements. Journal of the American Chemical Society 121(49):11503-11512.

142. Neuhaus D, Williamson MP. 2000. The Nuclear Overhauser Effect in Structural and Conformational Analysis. Second revised ed.

143. Price SL 2009. Computed crystal energy landscapes for understanding and predicting organic crystal structures and polymorphism. Accounts of Chemical Research 42(1):117-126.

144. Aaltonen J, Alleso M, Mirza S, Koradia V, Gordon KC, Rantanen J 2009. Solid form screening - A review. European Journal of Pharmaceutics and Biopharmaceutics 71(1):23-37.

145. Long S, Li T 2009. Controlled formation of the acid-pyridine heterosynthon over the acid-acid homosynthon in 2-anilinonicotinic acids. Crystal Growth & Design 9(12):4993-4997.

146. Reutzel SM, Etter MC 1992. Evaluation of the conformational, hydrogen-bonding and crystal packing preferences of acyclic imides. Journal of Physical Organic Chemistry 5(1):44-54.

147. Steiner T 2002. The hydrogen bond in the solid state. Angewandte Chemie-International Edition 41(1):48-76.

148. Desiraju GR 2007. Crystal engineering: A holistic view. Angewandte Chemie-International Edition 46(44):8342-8356.

149. Parr RG, Yang WT 1995. Density-functional theory of the electronic structure of molecules. Annual Review of Physical Chemistry 46:701-728.

150. Kohn W, Becke AD, Parr RG 1996. Density functional theory of electronic structure. Journal of Physical Chemistry 100(31):12974-12980.

151. Parr RG, Yang WT. 1989. Density Functional Theory of Atoms and Molecules New York: Oxford University Press.

152. Ayers PW, Levy M 2000. Perspective on "Density functional approach to the frontier-electron theory of chemical reactivity". Theoretical Chemistry Accounts 103(3-4):353-360.

153. Geerlings P, De Proft F, Langenaeker W 2003. Conceptual density functional theory. Chemical Reviews 103(5):1793-1873.

154. Huang Y, Zhong A, Rong C, Xiao X, Liu S 2008. Structure, spectroscopy, and reactivity properties of porphyrin pincers: A conceptual density functional theory and time-dependent density functional theory study. Journal of Physical Chemistry A 112(2):305-311.

155. Roy DR, Chattaraj PK 2008. Reactivity, selectivity, and aromaticity of Be-3(2-) and its complexes. Journal of Physical Chemistry A 112(7):1612-1621.

156. Cardenas C, Rabi N, Ayers PW, Morell C, Jaramillo P, Fuentealba P 2009. Chemical reactivity descriptors for ambiphilic reagents: Dual descriptor, local hypersoftness, and electrostatic potential. Journal of Physical Chemistry A 113(30):8660-8667.

157. Feng XT, Yu JG, Lei M, Fang WH, Liu SB 2009. Toward understanding metal-binding specificity of porphyrin: A conceptual density functional theory study. Journal of Physical Chemistry B 113(40):13381-13389.

158. Ugur I, De Vleeschouwer F, Tuzun N, Aviyente V, Geerlings P, Liu SB, Ayers PW, De Proft F 2009. Cyclopolymerization reactions of diallyl monomers: Exploring

electronic and steric effects using DFT reactivity indices. Journal of Physical Chemistry A 113(30):8704-8711.

159. Li T 2007. Understanding the polymorphism of aspirin with electronic calculations. Journal of Pharmaceutical Sciences 96(4):755-760.

160. Aubrey-Medendorp C, Swadley MJ, Li T 2008. The polymorphism of indomethacin: An analysis by density functional theory calculations. Pharmaceutical Research 25(4):953-959.

161. Li T, Ayers PW, Liu SB, Swadley MJ, Aubrey-Medendorp C 2009. Crystallization force - a density functional theory concept for revealing intermolecular interactions and molecular packing in organic crystals. Chemistry-a European Journal 15(2):361-371.

162. Yang WT, Parr RG 1985. Hardness, softness, and the Fukui function in the electronic theory of metals and catalysis. Proceedings of the National Academy of Sciences of the United States of America 82(20):6723-6726.

163. Parr RG, Von Szentpaly L, Liu SB 1999. Electrophilicity index. Journal of the American Chemical Society 121(9):1922-1924.

164. Morell C, Grand A, Toro-Labbe A 2005. New dual descriptor for chemical reactivity. Journal of Physical Chemistry A 109(1):205-212.

165. Parr RG, Yang WT 1984. Density functional approach to the frontier-electron theory of chemical reactivity. Journal of the American Chemical Society 106(14):4049-4050.

166. Ayers PW 2008. The dependence on and continuity of the energy and other molecular properties with respect to the number of electrons. Journal of Mathematical Chemistry 43(1):285-303.

167. Perdew JP, Parr RG, Levy M, Balduz JL 1982. Density functional theory for fractional particle number: Derivative discontinuities of the energy. Physical Review Letters 49(23):1691-1694.

168. Ayers PW, Morell C, De Proft F, Geerlings P 2007. Understanding the Woodward-Hoffmann rules by using changes in electron density. Chemistry-a European Journal 13(29):8240-8247.

169. Ayers PW, Morrison RC, Roy RK 2002. Variational principles for describing chemical reactions: Condensed reactivity indices. Journal of Chemical Physics 116(20):8731-8744.

170. Yang W, Mortier WJ 1986. The use of global and local molecular parameters for the analysis of the gas-phase basicity of amines. Journal of the American Chemical Society 108(19):5708-5711.

171. Mulliken RS 1955. Electronic population analysis on LCAO-MO molecular wave functions. I. Journal of Chemical Physics 23(10):1833-1840.

172. Hirshfield FL 1977. Bonded-atom fragments for describing molecular charge densities. Theoretica Chimica Acta 44(2):129-138.

173. Reed AE, Weinhold F 1983. Natural bond analysis of near-Hartree-Fock water dimer. Journal of Chemical Physics 78(6):4066-4073.

174. Reed AE, Curtiss LA, Weinhold F 1988. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. Chemical Reviews 88(6):899-926.

175. Frenking G, Frohlich N 2000. The nature of the bonding in transition-metal compounds. Chemical Reviews 100(2):717-774.

176. Barone V, Cossi M, Tomasi J 1997. A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. Journal of Chemical Physics 107(8):3210-3221.

177. Amovilli C, Barone V, Cammi R, Cances E, Cossi M, Mennucci B, Pomelli CS, Tomasi J. 1999. Recent advances in the description of solvent effects with the polarizable continuum model. In: Advances in Quantum Chemistry, Vol 32: Quantum Systems in Chemistry and Physics, Pt II, San Diego: Academia Press, Inc., p 227-261.

178. Becke AD 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. Physical Review A 38(6):3098-3100.

179. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, *et al.* 2004. Gaussian 03. Revision C.02. Gaussian, Inc., Wallingford, CT.

180. Dennington R, Keith T, Millam J, Eppinnett K, Hovell WL, Gilliland R. 2003. GaussView, Version 3.0. Shawnee Mission, KS: Semichem, Inc.

181. Kohn W, Meir Y, Makarov DE 1998. Van der Waals energies in density functional theory. Physical Review Letters 80(19):4153-4156.

182. Dobson JF, McLennan K, Rubio A, Wang J, Gould T, Le HM, Dinte BP 2001. Prediction of dispersion forces: Is there a problem? Australian Journal of Chemistry 54(8):513-527.

183. Dunitz JD, Gavezzotti A 2009. How molecules stick together in organic crystals: weak intermolecular interactions. Chemical Society Reviews 38(9):2622-2633.

184. Ayers PW 2007. The physical basis of the hard/soft acid/base principle. Faraday Discussions 135(0):161-190.

185. Mammino L, Kabanda MM 2008. A computational study of the interactions of the phloroglucinol molecule with water. Journal of Molecular Structure-Theochem 852(1-3):36-45.

186. Boys SF, Bernardi F 1970. Calculation of small molecular interactions by differences of separate total energies - some procedures with reduced errors. Molecular Physics 19(4):553-566.

187. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, *et al.* 2009. Gaussian 09. Revision A.1. Gaussian, Inc., Wallingford, CT.

188. Zhao Y, Truhlar DG 2008. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theoretical Chemistry Accounts 120(1-3):215-241.

189. Feng S, Li T 2006. Predicting lattice energy of organic crystals by density functional theory with empirically corrected dispersion energy. Journal of Chemical Theory and Computation 2(1):149-156.

190. Li T, Feng S 2006. Empirically augmented density functional theory for predicting lattice energies of aspirin, acetaminophen polymorphs, and ibuprofen homochiral and racemic crystals. Pharmaceutical Research 23(10):2326-2332.

191. Dion M, Rydberg H, Schröder E, Langreth DC, Lundqvist BI 2004. Van der Waals density functional for general geometries. Physical Review Letters 92(24):246401-246404.

192. Hepburn J, Scoles G, Penco R 1975. A simple but reliable method for the prediction of intermolecular potentials. Chemical Physics Letters 36(4):451-456.

193. Wu Q, Yang WT 2003. A direct optimization method for calculating density functionals and exchange-correlation potentials from electron densities. Journal of Chemical Physics 118(6):2498-2509.

194. Dovesi R, Saunders VR, Roetti C, Orlando R, Zicovich-Wilson CM, Pascale F, Doll K, Harrison NM, Civalleri B, Bush IJ, D'Arco P, Llunnell M. 2006. Crystal06 User's Manual. Universita' di Torino: Torino.

195. Latajka Z, Bouteiller Y 1994. Application of density functional methods for the study of hydrogen-bonded systems: The hydrogen fluoride dimer. Journal of Chemical Physics 101(11):9793-9799.

196. Hao MH 2006. Theoretical calculation of hydrogen-bonding strength for drug molecules. Journal of Chemical Theory and Computation 2(3):863-872.

197. Hunter KC, Millen AL, Wetmore SD 2007. Effects of hydrogen-bonding and stacking interactions with amino acids on the acidity of uracil. Journal of Physical Chemistry B 111(7):1858-1871.

198. Moller C, Plesset MS 1934. Note on an approximation treatment for many-electron systems. Physical Review 46(7):0618-0622.

199. Hobza P, Sponer J, Reschel T 1995. Density functional theory and molecular clusters. Journal of Computational Chemistry 16(11):1315-1325.

200. Tsuzuki S, Luthi HP 2001. Interaction energies of van der Waals and hydrogen bonded systems calculated using density functional theory: Assessing the PW91 model. Journal of Chemical Physics 114(9):3949-3957.

201. Surov AO, Szterner P, Zielenkiewicz W, Perlovich GL 2009. Thermodynamic and structural study of tolfenamic acid polymorphs. Journal of Pharmaceutical and Biomedical Analysis 50(5):831-840.

202. Mattei A, Li T 2011. Interplay between molecular conformation and intermolecular interactions in conformational polymorphism: A molecular perspective from electronic calculations of tolfenamic acid. International Journal of Pharmaceutics 418(2):179-186.

203. Gao J 1996. Hybrid quantum and molecular mechanical simulations: An alternative avenue to solvent effects in organic chemistry. Accounts of Chemical Research 29(6):298-305.

204. Orozco M, Luque FJ 2000. Theoretical methods for the description of the solvent effect in biomolecular systems. Chemical Reviews 100(11):4187-4225.

205. Tomasi J, Mennucci B, Cammi R 2005. Quantum mechanical continuum solvation models. Chemical Reviews 105(8):2999-3093.

206. Cramer CJ, Truhlar DG 1999. Implicit solvation models: Equilibria, structure, spectra, and dynamics. Chemical Reviews 99(8):2161-2200.

207. Tomasi J, Persico M 1994. Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent. Chemical Reviews 94(7):2027-2094.

208. Allen MP, Tildesley DJ 1987. Computer Simulations of Liquids. London: Oxford University Press.

209. Gao J, Freindorf M 1997. Hybrid ab initio QM/MM simulation of N-methylacetamide in aqueous solution. Journal of Physical Chemistry A 101(17):3182-3188.

210. Stanton RV, Hartsough DS, Merz KM 1993. Calculation of solvation free energies using a density functional/molecular dynamics coupled potential. Journal of Physical Chemistry 97(46):11868-11870.

211. Field MJ, Bash PA, Karplus M 1990. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. Journal of Computational Chemistry 11(6):700-733.

212. Case DA, Pearlman DA, Caldwell JW, Cheatham TEI, Wnag J, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Gohlke H, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA. 2006. AMBER 9. University of California, San Francisco.

213. Bayly CI, Cieplak P, Cornell WD, Kollman PA 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. Journal of Physical Chemistry 97(40):10269-10280.

214. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KMJ, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman P 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules Journal of American Chemical Society 117(19):5179-5197.

215. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong GM, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang JM, Kollman P 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. Journal of Computational Chemistry 24(16):1999-2012.

216. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML 1983. Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics 79(2):926-935.

217. Dupradeau F-Y, Cieplak P 2007. The tripos mo2 file format. http://q4md-forcefieldtools.org/Tutorial?Tutorial-1.php#15.

218. Andersen HC 1980. Molecular dynamics simulations at constant pressure and/or temperature. Journal of Chemical Physics 72(4):2384-2393.

219. Weast RC. 1988. Handbook of Chemistry and Physics. Boca Raton, FL: CRC Press, Inc.

220. Ryckaert JP, Ciccotti G, Berendsen HJC 1977. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. Journal of Computational Physics 23(3):327-341.

221. Darden T, York D, Pedersen L 1993. Particle mesh Ewald: An *N*-log(*N*) method for Ewald sums in large systems. Journal of Chemical Physics 98(12):10089-10092.

222. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR 1984. Molecular dynamics with coupling to an external bath. Journal of Chemical Physics 81(8):3684-3690.

223. Dapprich S, Komaromi I, Byun KS, Morokuma K, Frisch MJ 1999. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. Journal of Molecular Structure-Theochem 461-462:1-21.

224. Xiang TX, Anderson BD 2006. Conformational structure, dynamics, and solvation energies of small alanine peptides in water and carbon tetrachloride. Journal of Pharmaceutical Sciences 95(6):1269-1287.

225. Bizzarri AR, Cannistraro S 2002. Molecular dynamics of water at the protein-solvent interface. Journal of Physical Chemistry B 106(26):6617-6633.

226. Auffinger P, Louisemay S, Westhof E 1995. Multiple molecular dynamics simulations of the anticodon loop of $tRNA^{Asp}$ in aqueous solution with counterions. Journal of the American Chemical Society 117(25):6720-6726.

227. Reichardt C. 1988. Solvents and Solvent Effects in Organic Chemistry. Seocnd ed., Weinheim, Germany: VCH.

228. Brock CP, Minton RP 1989. Systematic effects of crystal-packing forces: Biphenyl fragments with H atoms in all four ortho positions. Journal of American Chemical Society 111(13):4586-4593.

229. Pascal RA 2006. Twisted acenes. Chemical Reviews 106(12):4809-4819.

230. Derdour L, Pack SK, Skliar D, Lai CJ, Kiang S 2011. Crystallization from solutions containing multiple conformers: A new modeling approach for solubility and supersaturation. Chemical Engineering Science 66(1):88-102.

231. Donnay JDH, Harker D 1937. A new law of crystal morphology extending the law of bravais. American Mineralogist 22(5):446-467.

Vita

ALESSANDRA MATTEI

Born: June 18, 1977

Birthplace: Grosseto, Italy

EDUCATION

2007 – Present	Ph.D. Candidate in Pharmaceutical Sciences; College of Pharmacy, University of Kentucky, Lexington, KY.
2005 - 2007	M.S. in Chemistry; University of Puerto Rico, Mayagüez, PR.
1996 - 2003	*Laurea* in Pharmaceutical Chemistry and Technologies; University of Pisa, Pisa, Italy.

PROFESSIONAL EXPERIENCES

2008 – 2012	Research Assistant - University of Kentucky, College of Pharmacy.
2007 - 2008	Teaching Assistant – University of Kentucky, College of Pharmacy.
2005 - 2007	Teaching Assistant – University of Puerto Rico, Department of Chemistry.
2004	Intern - Quality Control Area of ABIOGEN PHARMA; Pisa, Italy.

HONORS

- Peter G. Glavinos, Jr., Ph.D. Graduate Student Travel Award Fall 2011

PUBLICATIONS

- **A Mattei**, T Li. (2012) Polymorph Formation and Nucleation Mechanism of Tolfenamic Acid in Solution: An Investigation of Pre-Nucleation Solute Association. *Pharmaceutical Research* 29(2): 460-470.
- T Li, P Zhou, **A Mattei**. (2011) Electronic Origin of Pyridinyl N as a Better Hydrogen-Bonding Acceptor than Carbonyl O. *CrystEngComm* 13: 6356-6360.
- **A Mattei**, T Li. (2011) Interplay between Molecular Conformation and Intermolecular Interactions in Conformational Polymorphism: A Molecular

Perspective from Electronic Calculations of Tolfenamic Acid. *International Journal of Pharmaceutics – Special Issue Entitled "A priori Performance Predictions in the Pharmaceutical Sciences"* 418(2): 179-186.

- S Long, MA Siegler, **A Mattei**, T Li. (2011) Phase Transition from Two Z' = 1 Forms to a Z' = 2 Form of a Concomitant Conformational Polymorphic System. *Crystal Growth & Design* 11(2): 414-421.
- S Long, KL Theiss, **A Mattei**, CD Loftin, T Li. (2010) Solid-State Properties of the Cyclooxygenase-1-Selective Inhibitor, SC-560. *AAPS PharmSciTech* 11(2): 485-488.

PRESENTATIONS

- **A Mattei**, X Mei, A-F Miller, T Li. "Solution NMR Investigation of Molecular Conformation and Self-Association of Tolfenamic Acid and Its Nucleation Mechanism"; Poster AAPS Annual Meeting, Washington, DC, October 2011.
- **A Mattei**, T Li. "Investigating Solution Chemistry of Tolfenamic Acid with Molecular Dynamics and Quantum Mechanics/Molecular Mechanics"; Poster AAPS Annual Meeting, Washington, DC, October 2011.
- **A Mattei**, X Mei, A-F Miller, T Li. "NMR Study of Pre-Nucleation Self-Association"; Oral Presentation PGSRM, Madison, WI, June 2011.
- **A Mattei**, T Li. "Interplay Between Molecular Conformation and Intermolecular Interactions During Crystal Growth: A Case Study of Tolfenamic Acid"; Poster AAPS Annual Meeting, New Orleans, LA, November 2010.
- **A Mattei**, T Li. "Conformational Polymorphism of Tolfenamic Acid: An Analysis by Electronic Calculations"; Poster PGSRM, Columbus, OH, June 2010.
- **A Mattei**, T Li. "Mechanistic Understanding of Polymorphic Selection in Solution"; Poster AAPS Annual Meeting, Los Angeles, CA, November 2009.
- **A Mattei**, T Li. "Polymorph Selection: A Study of Tolfenamic Acid"; Poster PGSRM, West Lafayette, IN, June 2009.
- CP Hollis, **A Mattei**, T Li. "A Surface Energy Model of Eutectic Formation"; Poster AAPS Annual Meeting, Atlanta, GA, November 2008.
- **A Mattei**, S Long, T Li. "Chemical Structure and Crystal Growth: A Study of Diarylamines"; Poster AAPS Annual Meeting, Atlanta, GA, November 2008.

Alessandra Mattei
Student's Signature

September 2012
Date