STUDIES OF LOCAL STRUCTURE AND NUCLEATION IN BENZOIC ACID DERIVATIVE SYSTEMS

SHARLINDA BINTI SALIM SACHITHANANTHAN

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Chemical and Process Engineering

June 2017

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

ACKNOWLEDGEMENT

First and foremost, I thank my sponsor, Majlis Amanah Rakyat Malaysia (MARA), for funding my PhD studies.

This research has been conducted mainly in The University of Manchester; two years as a registered student and the last two years as a visiting student with the help of Prof Roger Davey. I would like to express my sincere gratitude to Prof Roger for his continuous support of my PhD study and for teaching me the science of crystallisation. His guidance, particularly on the topics of polymorphism, nucleation and solution chemistry, helped me in my experimental works and data interpretation. He also gave me the access to his laboratory and other research facilities within his team, which I highly appreciate. With him, all my problems seem to have the solutions and I can't thank him enough for being there for me.

My sincere thanks also go to Dr Joanna Stevens, who is patience in guiding me on the core level spectroscopy techniques. Working with her had increased my interest and together we explored more on the application of these techniques. Her guidance, motivation and advice helped me throughout my PhD research and in writing this thesis. I thank Prof Sven Schroeder, who initiated this project, Dr Aurora-Cruz Cabeza and Dr Thomas Vetter for their insights regarding the solution chemistry and nucleation kinetics respectively.

I thank Loris Doyle for his support in the lab. He is approachable and always ready to help, which made my works in the lab easier. I appreciate the help of Rebecca Smith and Thomas Brindley as well for taking their time to proofread several chapters and sections in my thesis.

To my number one inspiration, my late mother, Nabeezah. I miss you and hope I make you proud. I thank my late grandmother, Tok Sariah, for being the coolest grandma and her last best wishes as I left to UK. To Sara, thank you for being such a good friend who gave me comfort and made my day with your thoughtful and sweet gestures. God bless all of you.

My supportive family – My father, Salim Sachithananthan, and my siblings, Sharmini, Sharleena, and Surin, thank you for making my yearly trip back at home so wonderful and reviving. Last but not least, my wonderful friends back in Malaysia and here in UK, I could not have made it without all of your support.

ABSTRACT

Nucleation of organic crystals from solution starts from self-assembly of molecules. New and advanced techniques are continuously explored to have an insight into understanding the nucleation mechanism and the local structure in both solution and resulting crystalline forms.

In this thesis, multiple spectroscopic techniques were applied - infrared (IR), Raman, near-edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron (XPS) - that gave an extensive overview of local interactions between molecules. The nucleation rates were calculated based on probability distributions of induction times using multiple micro-reactor setup and revised methodology.

Solid and solution state characterisation of both polymorphs of *p*-nitrobenzoic acid (PNBA), Forms I and II, has been performed and they were found to have similar crystal structures and be enantiotropically related. A comparative examination of substituent effects on local interactions and electronic structures of PNBA, *p*-aminobenzoic acid (PABA), and benzoic acid (BA) has been performed. PNBA has a strong electron withdrawing nitro group, while PABA has a strong electron donating amine group. IR and Raman spectra showed the carbonyl stretching frequency for PNBA is highest, followed by BA and PABA. PNBA has a lower orbital energy for all core 1s levels and lowest unoccupied molecular orbital ($1\pi^*$), than PABA in the solid state. The substituent effect on the carbonyl stretching frequency and oxygen NEXAFS data of PNBA, BA and PABA in acetonitrile and ethanol were in the same order as the solid state. FTIR data for all compounds indicate that solution speciation in acetonitrile is carboxylic acid dimers and solute-solvent complexes, while in ethanol there are just solute-solvent interactions.

The influences of thermodynamic, β, and pre-exponential kinetic factors, A, on the nucleation rate did not correlate with the spectroscopic data. This raised several questions on the accuracy of the obtained β and A from the induction time measurements and the role of hydrogen bonding in the nucleation pathway.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Crystallisation and Polymorphism

More than 90% of active pharmaceutical ingredients (APIs) in drug formulations are present in the crystalline form [1]. The effectiveness of these drugs depends on their bioavailability which, in turn, changes as a function of interdependent physical properties. These include crystal structure, polymorphic form, size and habit. Therefore, controlled crystallisation processes play a crucial part in pharmaceutical manufacturing. Crystallisation is a purification and separation process involving the transition from a disordered system of molecules, ions or atoms in a solution to an ordered three-dimensional molecular array; a crystal [2].

The control of crystal structure polymorphism is a particularly difficult problem, and is already known to occur in more than half of all drugs [1]. Polymorphism is a phenomenon where the crystal structures of two or more solid phases are different, even though the chemical composition is the same. Each crystal structure has specific solid state properties, such as stability, crystal shape, compressibility, density and dissolution rate, which consequently affect the drug performance [3]. Polymorph selection is believed to take place already in the early stages of crystallisation, known as nucleation [4].

Nucleation is the initial process of crystallisation, the beginning transition from molecules or ions in solution to their solid state phase [2]. The driving force required for nucleation to occur is supersaturation. The initial phase transition starts from an aggregation of several molecules, n, into a nucleus of size n. The nucleus will spontaneously continue to grow above a critical nucleus size $n*$, and it will re-dissolve below this size. Capturing and understanding the structure of these extremely small critical nuclei, which can be as small as 1 to 100 atoms, is experimentally a major challenge [5,6]. With the recent advancement advanced in analytical technologies and modelling some progress has been made, and there has been an increase of interest in this field [4,6].

The structure of the nucleus is commonly assumed to be the same as the bulk crystal [4]. Therefore, understanding crystals at the molecular level, specifically the intermolecular interactions that bind molecules together in an ordered arrangement, is crucial. In the absence of strong electrostatic interactions hydrogen bonds are often the strongest intermolecular interaction in organic compounds [7]. Spectroscopy studies in the solution and the solid state have indicated some correlations between the structure of the molecules in solution and in the crystal for some systems [8].

The research in this thesis focuses on gathering more information about the nucleation process for one particular system, *para*-nitrobenzoic acid (PNBA), by combining thermodynamic, kinetic and spectroscopic investigations in solution and in the solid state, and comparing with other previously studied benzoic acid systems. The simplicity yet interesting hydrogen bond interactions for carboxylic acid systems means they provide a particularly worthwhile area of study. This work incorporates both characterisation of the solid-state of PNBA and investigation of how the solvent affects the nucleation process and nature of the hydrogen bond interaction at the carboxylic acid group. In addition, comparison is made to benzoic acid derivatives with different *para*-substituent groups, to study the substituent effect on the hydrogen bond interaction, thermodynamic and kinetics of the crystallisation process. The unique combination of modern and advanced techniques, including multiple reactors and core electron spectroscopies, draws together information across different approaches and length scales from molecules through to crystals.

1.2 Thermodynamic of Solution and Polymorphic Transformation

Crystallisation from liquid solutions is common practice. The process of creating a solution involves the mixing of the solute with the solvent until it is homogenous [2]. Solvents are chosen based on the intermolecular interaction with the solute, which influence solubility. Solutes and solvents are classified into three main categories based on their hydrogen bonds; polar protic, dipolar aprotic and non-polar aprotic. Generally, the solubility of a solute in a solvent is higher if both are in the same class [7]. Solubility data are essential for understanding phase transitions, thermodynamic properties of the solution and polymorphism.

Experimental data of temperature-dependent solubility curves are commonly fitted using the van’t Hoff equation to obtain the thermodynamic properties of the solution – the enthalpy and entropy of dissolution, ΔH_d, and ΔS_d [7,9,10]. Equation 1.1 is the linearised van’t Hoff equation where x is the mole fraction of the solute in the solvent, R is the gas constant, 8.314 J mol^{-1} K^{-1}, and T is the solution temperature. The ΔH_d and ΔS_d are obtained from the slope and the *y*-interception based on the linear regression analysis on the experimental data in a plot of $\ln x$ against $1/T$ [7,9]. Alternatively, by calculating the molar Gibbs free energy of dissolution, ΔG_d ,using Equation 1.2, the ΔS_d at temperature, T , can be calculated based on Equation 1.3 [3].

$$\ln x = -\frac{\Delta H_d}{RT} + \frac{\Delta S_d}{R}$$ **Equation 1.1**

$$\Delta G_d = -RT \ln x$$ **Equation 1.2**

$$\Delta S_d = \frac{\Delta H_d - \Delta G_d}{T}$$ **Equation 1.3**

Polymorphs can be characterised as either monotropic or enantiotropic based on their temperature dependent solubility curves. For a monotropic system, the temperature dependent plots of the free energies of the polymorphs do not cross before all polymorphs have melted, hence any transformation from one polymorph to another below the melting point is irreversible. For an enantiotropic system, the free energy curves cross at least one temperature below the melting points of the polymorphs, hence temperature variation near the crossover points can cause a reversible phase transition. [11]. The stable polymorph is the polymorph which has the lowest free energy. Assuming an ideal solution, the free energy change is proportional to the solubility as shown in Equation 1.2. Therefore, the most stable polymorph should also have the lowest solubility [2,7].

The transition temperature, T_r, is the temperature where all the polymorphs in the system are in equilibrium. The thermodynamic phase transition of a polymorphic system is studied based on the heat of transition, ΔH_T , molar Gibbs free energy of transition, ΔG_T , and entropy of transition, ΔS_T . These thermodynamic properties for the transformation of polymorph A to B are calculated by using Equation 1.4 to 1.7 as below [3].

$$\Delta H_T = \Delta H_s{}^B - \Delta H_s{}^A$$ **Equation 1.4**

$$\Delta G_T = \Delta G_s{}^B - \Delta G_s{}^A$$ **Equation 1.5**

$$\Delta G_T = RT \ln \frac{x_A}{x_B}$$ **Equation 1.6**

$$\Delta S_T = \frac{\Delta H_T - \Delta G_T}{T_r}$$ **Equation 1.7**

Advanced predictive analysis of thermodynamic properties based on the solubility data and thermal analysis is commonly performed. The melting temperature of a metastable polymorph, the correlation between the enthalpy of dissolution to the solubility, temperature and heat capacity and the analysis on the free energy, enthalpy and entropy changes as a function of temperature have proven useful in understanding polymorphic systems [10,12–19]. As an example, the transition temperature of β−PABA to α−PABA at low temperature, 16 °C, is correlated with a high $\Delta S^{\beta \rightarrow \alpha}$ value, which is 3.1 kJ mol^{-1} at 300 K [10]. The *meta*-hydroxybenzoic acid (*m*HBA) polymorphs have been concluded to be monotropically related based on the calculated free energy and enthalpy differences. The free energy difference decreased from 0.87 kJ mol^{-1} at room temperature to 0.50 kJ mol^{-1} at the melting temperature of Form II, while the enthalpy increased from 1.20 kJ mol^{-1} to 2.08 kJ mol^{-1} [17].

1.3 Classical Nucleation Theory

The nucleation process is divided into two groups: primary and secondary. Primary nucleation consists of a liquid phase only and is further categorised into either a homogeneous or heterogeneous process. The homogenous process occurs naturally while foreign particles such as dust, surfaces of objects such as stirring rod, or the wall of the tank will induce the heterogeneous nucleation process. Practically, the homogenous process rarely occurs. On the other hand, secondary nucleation involves the process of seeding by adding crystals (seeds) into the supersaturated solution. This method helps to control the nucleation process as the resulting crystals will inherit the seed's characteristics [7].

Contemporary theories of nucleation include classical nucleation theory (CNT), two-step mechanism theory (TSMT) and pre-nucleation cluster (PNC) theory. CNT was introduced in the 1930s and, to date, is commonly applied in many branches of chemistry [4,7,20]. CNT states that there is a single energy barrier caused by the balance between increasing surface energy and cohesive energy for small particles, which needs to be overcome before spontaneous growth of crystals can take place. Section 1.3.1 and 1.3.2 will elaborate further on the mechanism. TSMT is based on observations in protein crystallisation, where two energy barriers are often observed, one associated with the formation of a dense liquid, and the second with the beginning growth of crystals [21,22]. PNC describes the crystallisation of calcium carbonate through an intermediate PNC or an amorphous phase [23].

1.3.1 Thermodynamic Barrier

Figure 1.1 shows the free energy difference as a function of the nucleus radius which is commonly used when discussing the thermodynamics behind the CNT. The free energy change, ΔG , in creating the nucleus is the sum of the free energy change for the phase transition (liquid to solid) , ΔG_V , and the free energy change for the surface formation (liquid/solid), ΔG_S , is given as in Equation 1.8 [20].

$$\Delta G = \Delta G_V + \Delta G_S \qquad \textbf{Equation 1.8}$$

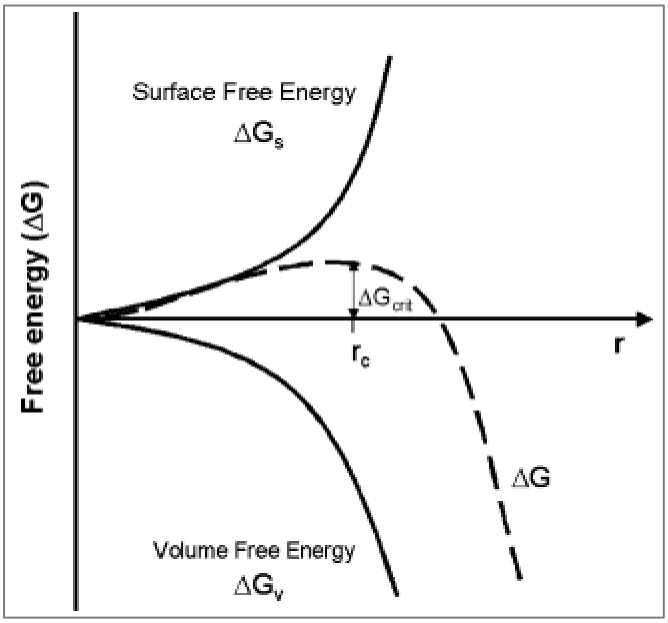


Figure 1.1: Free energy diagram based on classical nucleation theory [20]

Figure 1.2 (a) illustrates an initially homogeneous solution with M number of molecules while, Figure 1.2 (b), illustrates the new state having a cluster with n number of molecules. The difference in free energy is defined as the final free energy of the solution with the cluster minus the initial free energy of the solution [24]. μ_s is the chemical potential of the solution [24].

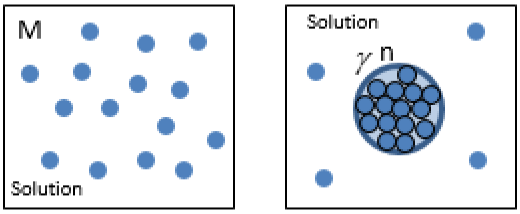


Figure 1.2: (a) An illustration the initial state with M number of solute molecules (blue circle) in a solution state (left) and (b) the final state with a cluster of n size molecules (right) [24].

$$\Delta G = G_{final} - G_{initial}$$ **Equation 1.9**

$$\Delta G = [(M - n)\mu_s + G(n)] - M\mu_s$$ **Equation 1.10**

The formation of the cluster in the final state creates a surface between the cluster and the solution, thus, a surface energy needs to be overcome. $G(n)$ is the free energy of the cluster while $G_{eff}(n)$ is the total surface energy of the cluster are given in Equation 1.11 and Equation 1.12.

$$G(n) = n\mu_c + G_{eff}(n)$$ **Equation 1.11**

$$G_{eff}(n) = \gamma A_c(n)$$ **Equation 1.12**

μ_c is the chemical potential of the crystal, γ (J m^{-2}) is the specific surface energy of the cluster-solution interface and by assuming a spherical shape cluster, the surface area is $A_c(m^2) = (36\pi v_0{}^2)^{1/3} n^{2/3}$ where $v_0 = M / \rho N_A$ is the volume of the molecules in the cluster, M and ρ are the molecular weight (g mol^{-1}) and density (g m^{-3}) of the crystals [5].

Substituting the $G(n)$, $G_{eff}(n)$ and $\mu_s - \mu_c = RT \ln S$ [5] into Equation 1.10, the maximum free energy or work is simplified as in Equation 1.13 where S is the supersaturation ratio. The supersaturation ratio can be calculated as the ratio between the actual concentration and the equilibrium concentration at temperature T [7].

$$\Delta G = -nRT \ln S + \gamma A_c$$ **Equation 1.13**

The maximum free energy, ΔG_{crit}, which is at the critical nucleus radius, r_c, is the maximum energy needed to overcome the nucleation barrier. The ΔG_{crit} is also known as the maximum nucleation work W^* to create a critical nucleus size, n^* [5].

1.3.2 Kinetic Barrier

The nucleation rate, J, which is the number of nuclei formed per unit time and volume, is expressed in the form of the Arrhenius reaction as in Equation 1.14 where k is the Boltzmann constant 1.3805 x 10^{-23} J K^{-1}, T is the temperature in Kelvin (K) while ΔG has been discussed in above section. A is the pre-exponential kinetic term which has been derived in detail by Kashchiev [5,24] but only the final result, as in Equation 1.15 will be discussed here.

$$J = A\exp\left(-\frac{\Delta G}{kT}\right)$$ **Equation 1.14**

$$A = zf^*C_0$$ **Equation 1.15**

$$z = \left(\frac{\Delta G_{crit}}{2\pi kTn^{*2}}\right)^{1/2}$$ **Equation 1.16**

z is the Zeldovich factor calculated as in Equation 1.16, f^* is the frequency attachment of a monomer to the nucleus and C_0 is the concentration of active nucleation sites [24,25]. The C_0 is an unknown parameter for a heterogeneous system as the active nucleation sites is referring to the dust, particles or surfaces in the solution which cannot be quantified [25]. Consequently, the f^* cannot be calculated. The f^* is crucial as it determines the attachment rate of the building units being adsorbed onto the nucleus [4,26]. This further limits the understanding of the molecular arrangement during the nucleation process.

Figure 1.3 illustrates the free energy against the reaction coordinate for the two polymorphs, I and II. G_0 is the free energy of the supersaturated solution while G_I and G_{II} is the free energy to crystallise each polymorph, with polymorph I being more stable than polymorph II. The G^*_I and G^*_{II} are the activated energy related to the complex assembly of the molecules into a specific packing arrangement in creating a new solid phase. This activation energy is embedded in the pre-exponential kinetic factor, A.

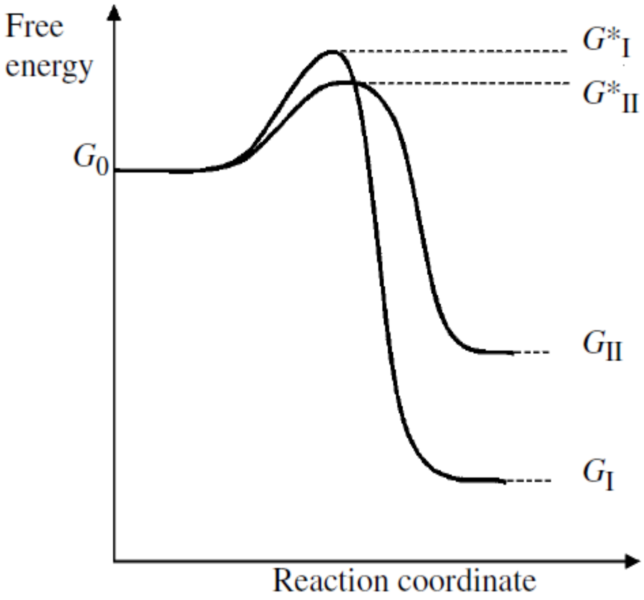


Figure 1.3: Schematic of the activation energy barrier and thermodynamic barrier for polymorphs I and II [11].

The balance between the ΔG and ΔG^* to the nucleation rate, J, is not straightforward. For example, based on Figure 1.3, as the supersaturation for polymorph I is higher, $(G_0 - G_I)$, the critical nucleus size is expected to be larger than for polymorph II and so the activation barrier is bigger. Therefore, kinetically polymorph II is preferred, however, thermodynamically polymorph I is more stable. The overall nucleation rate is a function of temperature (T), supersaturation, S, interfacial energy, γ, the frequency of molecules attachments, f^*, and the concentration of heterogeneous sites, C_0, as shown in Equation 1.17.

$$J = zf^* C_0 \exp\left(-\frac{nkT\ln S + \gamma A}{kT}\right)$$ **Equation 1.17**

A multiplexed micro-reactor setup is a good method for measuring the induction times, t_{ind}, at a constant supersaturation, volume and temperature. The main objective is to obtain the pre-exponential kinetic factor, A, the thermodynamic factor, B, and the nucleation rate, J, in different systems [10,17,26–31]. These kinetic factors are often considered for the interpretation of data in different contexts. For example, the stable α−PABA has been shown to be kinetically favoured over β−PABA to nucleate under a variety of different conditions (temperature, solvent, driving force) [10]. A mixture with a higher proportion of metastable *m*HBA Form II to Form I is kinetically favoured to nucleate in different solvents [17]. Agitation rates have been shown to influence the proportion of these *m*HBA polymorphs [29]. The ratio of the nucleation rate, J , of Form I to Form II decreased from the lower to the higher driving force. The nucleation rate of a filtered and non-filtered solution of isonicotinamide (INA) in ethanol shows a 70% difference in the nucleation rate which is the influence of heterogeneous particles [31].

1.4 Molecular Spectroscopy

Spectroscopy techniques are applied when characterising the solid state form of polymorphs, co-crystals, salts and solvates. Infrared (IR), Raman, ultraviolet-visible (UV-Vis) and nuclear magnetic resonance (NMR) spectroscopy are among the most common techniques used for characterising a compound. Infrared and Raman spectroscopy are based on the vibrational mode of the molecules while UV-Vis provides the transition energy for the excitation of an electron from the highest occupied molecular orbitals (HOMOs) to the lowest unoccupied molecular orbitals (LUMOs) [32]. NMR spectroscopy was less suitable for this project as it requires more complex understanding of the physics of the molecular structure [3]. Nevertheless, this technique is evolving and being applied to a wider scope [33,34].

Spectroscopy studies have been a great tool in probing hydrogen bonds. Understanding hydrogen bonds in depth is important for relating interactions between building units or synthons of a crystal and the solution structure [8,35]. Such building units are taken to be the repetitive unit in a crystal structure [35]. One of the most common examples is a carboxylic acid dimer. The IR spectroscopy studies on carboxylic acids have long established that the local environment influences the frequency of carbonyl (C=O) stretching and hydroxyl (O−H) out-of-plane deformations [36]. For example, a carboxylic acid dimer has an O−H frequency around 950 $\mathrm{cm^{-1}}$and in the monomer around 650 $\mathrm{cm^{-1}}$, while the frequency of C=O is around 1650 $\mathrm{cm^{-1}}$. The stretching frequency is influenced by the functional group, substituents or intramolecular hydrogen bond [36].

Characterising monomer and dimer speciation in the solution state is more complex as there are solute-solute, solute-solvent and solvent-solvent interactions. The carbonyl frequencies of monomers, dimers and complexes for carboxylic acid in various solvents have been studied using IR spectroscopy [37]. The carbonyl frequency for the dimer band was found consistently at the lower and smaller range of frequency, 1692−1700 cm^{-1} while the others speciation at higher frequency band has a wider range at 1698 – 1748 cm^{-1} [36,37]. The wider range is associated with different speciation is solvent dependent.

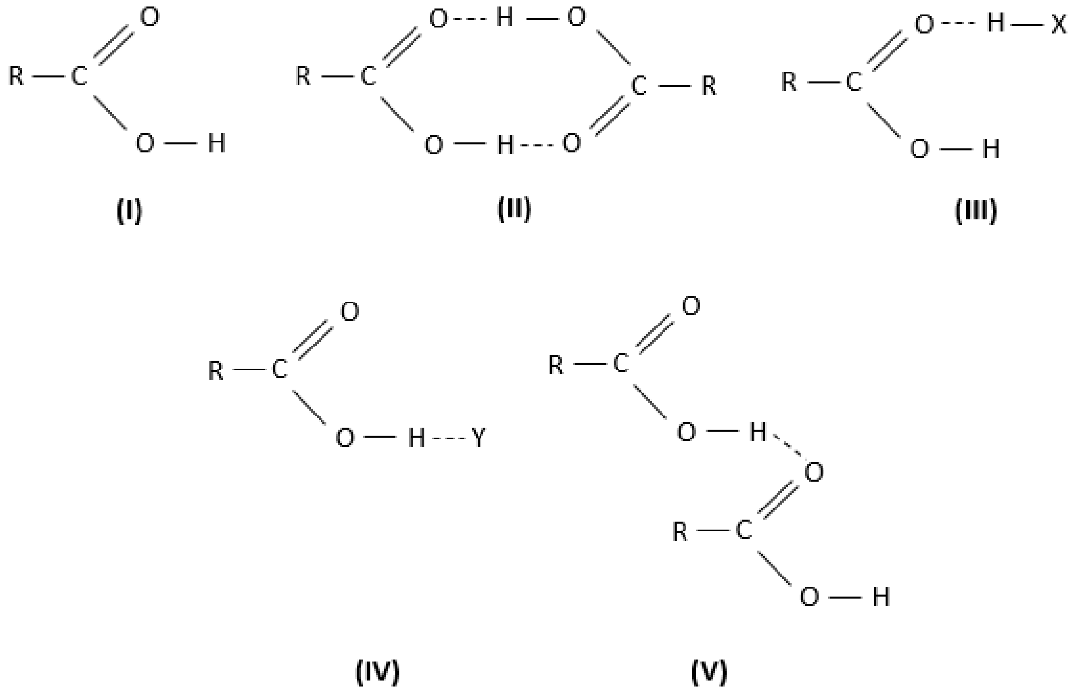


Figure 1.4: The carboxylic acid species may exist as monomer (I), dimer (II), complexes (III) and (IV) or acyclic dimer (V).

The frequency shift of the carbonyl stretching for benzoic acid in different types of solvent suggest five different speciation in the solution as illustrated in Figure 1.4. In a non-polar solvent such as toluene, the existence of monomer **(I)** and dimer **(II)** explained the two carbonyl bands with a frequency difference of more than 40 cm^{-1}. The existence of complex **(III)** and **(IV)** relative to the dimer **(II)** reduces the frequency difference to around 30 cm^{-1} and 20 cm^{-1} for proton donor (H−X) and acceptor (H···Y) solvents such as chloroform and acetonitrile [37]. Concentration dependence studies further support the existence of different speciations [38,39]. The cyclic **(II)** and the chain dimer **(V)** have a small difference in their carbonyl frequencies [40].

The solution speciation has on occasion been linked to polymorphism [8,41,42] and desolvation rate [28,43] but there were equally cases with poor correlation [8,44]. This encourages a wider aspect of studies on the molecular structure. X-ray photoelectron spectroscopy (XPS) and near-edge *x*-ray absorption fine structure (NEXAFS) spectroscopy are core level electron spectroscopy that probes the core electron, 1 s, leading them to be element-specific as well as sensitive to the local chemical/electronic environment [45,46]. An increasing number of studies have been carried out to understand the solid and solution structure by using these methods [34,47–52]. Similarly as in IR spectroscopy, the hydrogen bond interaction is discussed based on the chemical shift in the spectra.

The combination of XPS and NEXAFS data give an overview of the energy level diagram; the 1 s binding energy (BE) is obtained from the XPS, while the LUMOs energy is obtained from the NEXAFS. The sensitivity of these data in probing the hydrogen bond interaction has been shown in the solid state of salts, co-crystals [34,48], polymorphs [50] and in the solution state [49,52]. The characterisation of monomer and dimer interaction has also been discussed [51] however, there is still no strong conclusion. Nevertheless, XPS and NEXAFS are young techniques in the field of crystallisation and solid state characterisation. Therefore, more data from a variety of different systems is needed.

1.5 *p*-Nitrobenzoic Acid (PNBA), Benzoic Acid (BA) and *p*-Aminobenzoic Acid (PABA)

Benzoic acid (BA) and *p*-aminobenzoic acid (PABA) are aromatic carboxylic acids that have been studied widely investigating their thermodynamics, kinetics, molecular structure, and polymorphism [10,15,18,19]. BA has no polymorph and dissolves in non-polar and polar solvents, which revealed an interesting insight into the self-association of BA in the solution state [18,37]. PABA has two polymorphs with different building units; α−PABA has a carboxylic acid dimer and β−PABA has a catemer [15]. This motivates a study of the local environment between these two molecular structures experimentally by using IR, XPS and NEXAFS [50]. The low polymorphic transition temperature at 13°C as well motivates a comprehensive study on the thermodynamic of PABA using both solubility and calorimetric data [10]. The kinetic study of BA and PABA was linked to the desolvation rate of the carboxylic acid monomer and dimer [28].

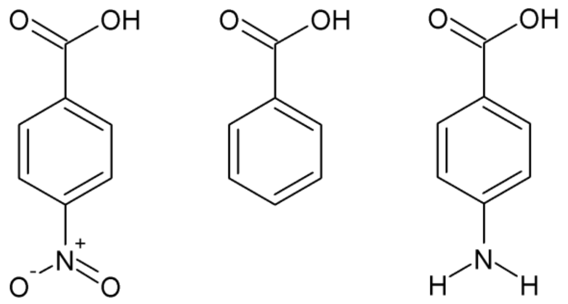


Figure 1.5: The chemical diagram for PNBA (left), BA (middle) and PABA (right)

p-nitrobenzoic acid (PNBA) is an aromatic carboxylic acid with a nitro group. Figure 1.5 illustrates the chemical diagram for PNBA, BA and PABA. PNBA is an interesting compound with an opposite characteristic to PABA. PNBA is a strong electron withdrawing group with a Hammett constant of +0.78 [54] while PABA is a strong electron donating group with a Hammett constant of −0.66 [54]. Unlike BA and PABA, no extensive study has been made on the crystallisation aspect of PNBA and so this will be carried out in this research.

1.6 The Aim of this PhD Research

In the recent years, more techniques have been established in measuring the nucleation rate of an organic compound from the solution [25–28,30,31,53]. This is later derived to obtain the thermodynamic and kinetic factor of the nucleation process (Section 1.3). The thermodynamic data – enthalpy, entropy, and free energy difference of the solution and polymorphic transformation is calculated from the solubility and calorimetric data [10,14]. The molecular structure in the solution and solid state is commonly studied using IR and Raman spectroscopy. The combinations of these studies – kinetics, thermodynamic, and molecular structure have been done on several systems to understand the mechanism of the nucleation process [10,17,28,30,53].

In this PhD research, a similar combination of studies is made - kinetics, thermodynamic and molecular structure but for a new system. The system consists of the benzoic acid derivatives: benzoic acid (BA), *p*-aminobenzoic acid (PABA) and *p*-nitrobenzoic acid (PNBA). The main aim is to understand the nucleation of carboxylic acids by studying the substituents effect of an electron donating and withdrawing group to the hydrogen bond at the carboxylic group and, subsequently, to the thermodynamics and kinetics of the crystallisation process. The amine ($-NH_2$) was the electron donating group used while the nitro ($-NO_2$) was the electron withdrawing group.

This project is divided into five different areas. The first area was to study the polymorphism, solid state characterisation, and thermodynamic properties of PNBA in the solid and solution state (Chapter 3). The second area was to improve the existing methodology of induction time measurements to provide consistency for PNBA in 2-propanol and ethyl acetate (Chapter 4). The third area was to study the limitation of the core level spectroscopy techniques on similar crystal structures (Chapter 5). The fourth area was to combine the vibrational and core level spectroscopy techniques in studying carboxylic acid aggregation, solution chemistry, and substituent effect on the local environment in the solid and solution state (Chapter 6 and 7). Lastly, the nucleation kinetic and spectroscopic data are combined in order to gain insight into the nucleation pathway (Chapter 8).

The novelty of this thesis lies in relating observable PNBA crystallisation kinetics and polymorphism with established and novel molecular structure probes, thermodynamic and kinetic studies. Moreover, the work reported here is only the second example of an application of core level spectroscopies, XPS and NEXAFS, for studying organic crystal polymorphs, and the first to examine the effects of substituents on the electronic structure in the solid and solution state. Third, the influence of pre-exponential kinetic and thermodynamic factor on the nucleation rate of PNBA and other substituted benzoic acids will be examined. Potential energy diagrams highlighting the influence of the nucleation and activation energy barriers will be derived.

1.7 Thesis Chapter Outline

Chapter 1

This chapter introduces the fundamentals of the research topic and relates it to contemporary research in the field. The aims and objectives of the research are stated.

Chapter 2

This chapter briefly describes all the methodologies used for the research in this dissertation.

Chapter 3

This chapter reports the characterization of the polymorph solid state forms of PNBA, focusing on thermodynamic data for each polymorph and for the polymorphic transformation.

Chapter 4

This chapter reports the outcomes of a detailed critical study of the experimental methodologies for measuring induction times for PNBA in 2-propanol and ethyl acetate.

Chapter 5

This chapter describes the use of XPS and NEXAFS in characterising PNBA polymorphs, which, due to the similarity of local structure in the two polymorphs, serves to demonstrate the limitations of these techniques in detecting and characterising polymorphic forms.

Chapter 6

This chapter examines the effect of different para-substituents on the solid state, combining IR, Raman, XPS and NEXAFS spectroscopy data for BA, PABA and PNBA.

Chapter 7

This chapter examines the effect of different para-substituents on the solution state by combining IR and NEXAFS data for BA, PABA and PNBA in ethanol and acetonitrile.

Chapter 8

This chapter combines the thermodynamic, kinetic and spectroscopy studies in the solid and solution state from Chapter 3 to 7.

Chapter 9

This chapter concludes the PhD research and outlines several potential future works.

2. EXPERIMENTAL AND METHODS

This PhD research, as highlighted in Chapter 1, is investigating the thermodynamics, nucleation kinetics and molecular structures of benzoic acid (BA), *p*-aminobenzoic acid (PABA) and *p*-nitrobenzoic acid (PNBA). The thermodynamic studies were based on temperature dependent concentration plots, also known as the solubility curve, which was obtained using the gravimetric analysis method. Nucleation kinetics were studied by measuring induction times in multiplexed micro-crystalliser setup; the Crystal16. The molecular structures in the solid state were studied using vibrational spectroscopy (infrared and Raman) as well as core level electron spectroscopy methods, namely X-ray photoelectron spectroscopy and near-edge X-ray absorption fine structure, XPS and NEXAFS. The molecular structures in the solution state were studied using infrared and NEXAFS spectroscopy. In addition, X-ray diffraction, differential scanning calorimetry and optical microscopy were used in characterising the polymorphs.

2.1 Materials

Benzoic acid (BA), *p*-aminobenzoic acid (PABA) and *p*-nitrobenzoic acid (PNBA) were purchased from Sigma-Aldrich with 99.8 %, 98 % and 99 % purity. PNBA Form I was produced by adding, in excess, PNBA Form II (purchased material) to ethanol or acetonitrile at room temperature. The slurry was stirred for a day, filtered and dried. 2-propanol (anhydrous, 99.5 %), ethanol, acetonitrile (anhydrous, 99.8 %), and ethyl acetate (anhydrous, 99.8 %) were used as solvents.

2.2 Cambridge Structural Database (CSD)

The Cambridge Structural Database (CSD) was established in 1965 and serves as a repository for organic and organometallic crystal structures [55]. This database provides information on X-ray powder diffraction patterns, geometry, structure, correlations, model compounds, packing arrangement and hydrogen-bonding patterns. There are more than 850,000 entries in the CSD as of the year 2017 [55]. There were nine entries in the crystal structure of PNBA dated from 1961 to 2015, with the starting CSD reference code of NBZOAC. Three structures were chosen as references; NBZOAC02 [56], NBZOAC03 [57], and NBZOAC04 [58] based on the single crystal analysis techniques and measured temperature.

Software: Mercury 3.7 and ConQuest 1.18 were used to visualise the molecular structure, calculate the geometry of the structure, obtain the powder pattern, and other information related to the crystal structure determination.

2.3 X-ray Diffraction (XRD)

A crystalline solid is a highly ordered structure of identical molecules arranged in a repeated pattern. The unit cell of the crystal structure contains the information related to the geometry, symmetry and the positions of the atoms [59]. The crystal structure of a crystalline solid form is determined by using X-ray diffraction (XRD). XRD is a non-destructive technique that is based on the unique diffraction pattern of the X-ray beams from the atom [59]. Bragg's Law, as in Equation 2.1, is used to understand the diffraction results, where n is the integer, λ is the wavelength of the incident wave, d_{hkl} is the space between the planes in the atomic lattice, and θ is the angle between the incident and the reflected beam. The final powder XRD pattern will show multiple peaks at different angles and intensities. Each peak represents the index position of the atom in the unit cell. In this thesis, the analysis of the XRD pattern is mainly for polymorph identification.

$$n\lambda = 2d_{hkl}\sin\theta \qquad \textbf{Equation 2.1}$$

Powder XRD patterns were obtained using a Rigaku MiniFlex Diffractometer. The sample was measured from a 2θ of 5–40° at a speed range of 1.5 to 3.0 ° min^{-1} and a step size of 0.03°. The Cu Kα (λ = 1.54 Å) X-ray source has a voltage of 30 kV and a current of 15 mA. If necessary, the sample was ground using a mortar and pestle to ease the task of smoothing out a sample on a sample holder.

Software: OriginPro 2015 was used to replot the powder XRD patterns.

The single crystal XRD pattern of PNBA Form II was determined in the School of Chemistry, The University of Manchester by Dr James Raftery and Dr Chris Muryn. The single crystal XRD measurements were carried out at two temperatures, 100 K and at ambient room temperature, using the Oxford Xcaliber 2 diffractometer, Mo K_α radiation (0.7107 Å) with a graphite monochromator. PNBA Form II was re-crystallised by slow evaporation at 60 °C in acetonitrile to produce crystals with size 0.2 − 0.5 mm.

Software: Mercury 3.7 was used to visualise the molecular and geometry structure of the crystal and extracting the calculated powder XRD pattern. OriginPro 2015 was used to replot this XRD pattern.

2.4 Gravimetric Analysis (Solubility)

The solubility of PNBA Form I and Form II in ethanol, 2-propanol, acetonitrile and ethyl acetate were measured using the gravimetric analysis technique. Purchased PNBA (Form II) was added in excess to 20 mL solvent, and stirred in a jacketed vessel at a controlled temperature using a thermostat. The stirring was stopped after 24 hours, allowing the excess solids to settle at the bottom thus creating a clear solution at the top. 1 – 2 mL of this clear solution was placed in a pre-weighed empty petri dish, m_0, and weighed, m_1. The dried solids in the petri dish were weighed, m_2, after all the solvents had evaporated. The concentration was then calculated by dividing the mass of solids $\left(m_2 - m_0\right)$ by the mass of solvents $\left(m_2 - m_1\right)$.

An additional step was taken in measuring the solubility of the metastable PNBA Form II at 25 °C to 45 °C and PNBA Form I at 55 °C. For example, 2 g of purchased PNBA (Form II) was added to 20 g of solvent. A sample of the slurry was taken at an interval ranging from every 30 min to an hourly basis to measure the powder x-ray diffraction (XRD) pattern. The stirring was stopped and the duration was noted once the transformation (PNBA Form II to Form I) had occurred which was based on the differences in the powder XRD pattern. The same composition of the new slurry was prepared and the stirring was stopped just before the polymorphic transformation occurred. The sampling of the clear solutions was as described earlier. The experiments were performed on both PNBA Form I and Form II from 25 °C to 55 °C at an interval of 5 °C in ethanol and an interval of 10 °C for 2-propanol, ethyl acetate and acetonitrile.

Software: Microsoft Excel 2013 and OriginPro 2015 were used for the statistical analysis and graph plots.

2.5 Slurry Transformation at 25 °C

2 g of purchased PNBA (Form II) was added to 20 g of solvent and stirred at 25 °C. A sample of the slurry was taken at an interval ranging from every 30 min to an hourly basis to measure the powder x-ray diffraction (XRD) pattern. The stirring was stopped and the duration was noted once the transformation (PNBA Form II to Form I) had occurred which was based on the differences in the powder XRD pattern.

2.6 Differential Scanning Calorimetry (DSC)

Differential scanning calorimetry (DSC) was used to detect changes in phase and thermal events. The sample pans were heated or cooled at a fixed rate. A DSC curve shows the melting point and the enthalpy of fusion. Melting and solid phase transition are endothermic, which would show up as a negative peak in the DSC curve, while crystallisation is an exothermic process and would show up as a positive peak [11]. The unit consists of a Mettler TG50, FP85 Thermal Analyser Cell and FP90 Central Processor. Approximately 4 mg of sample was weighed in the aluminium sample pans and sealed with a small hole pierced. The mass of the sample was recorded and the sample was then heated up from 30 °C to 300 °C at a rate of 10 °C min^{-1}. These settings can be adjusted in the temperature profile.

Software: Mettler Toledo FP99A and STARe were used in setting the temperature profile and obtaining the enthalpy and temperature of fusion. OriginPro 2015 was used to replot the experimental data.

2.7 Hot Stage Microscopy

A crystal of PNBA Form I with a size of roughly around 0.1 mm was obtained from a fast evaporation in acetonitrile. The crystal was placed on a hot stage system (LINKAM TMS 93) and the image was captured by an embedded camera (Infinity 1, Lumenera) to the optical microscope with a magnification of 50x. The crystal was then heated from 30 °C to 200 °C at a rate of 1 °C min^{-1} and a video recording was made to monitor the polymorphic transformation.

Software: Infinity analysed and Infinity capture were used to visualise the crystal and to set the video recording while Window Media Player was used to play the recording.

2.8 Induction Time Measurement (Crystal 16)

The induction times for the nucleation process were measured using the Crystal 16. The Crystal 16^{TM} is a multiple reactor system that holds 16 standard glass vials (2 mL) arranged into four separate blocks independently. The process variables include the temperature ranging from −15 °C to 100 °C, stirring speed from 0 to 1250 rpm (using the magnetic bars), heating and cooling rate in between 0 °C to 20 °C min^{-1}, duration and the number of repeats. The turbidity was measured in transmission using a light-emitting diode (LED) and a photo sensor in each vial. Each block was programmed individually. The temperature in each block was calibrated based on the external temperature probe against the set point temperature.

The process profile for the nucleation experiment started by heating up the supersaturated solution at a high temperature, 50 °C, and stirred for an hour at 900 rpm. Then, the turbidity reading was calibrated as 100 % clear solution using the 'tune' setting. Next, the solution was crash cooled at a rate of 5 °C min^{-1} to 25 °C and remained at 25 °C for 8 hours. This was repeated for 5 cycles giving a total of 80 sets of data. The experiment was then repeated at four to six different supersaturations. t_0 is the time the solution reached 25 °C and t_J is the time crystals were first detected in the solution that was based on the time of the first drop in the turbidity reading which is about 99 %. The induction time is the time difference between t_J and t_0. Chapter 4 will discuss this methodology further and the impact on the nucleation data.

Software: Crystal16, Microsoft Excel 2013, MATLAB Version 2016 and OriginPro 2015 were used. The Crystal 16 software generated all the recorded process settings for all the 80 experiments into Microsoft Excel. A MATLAB script written by Dr Thomas Vetter from The University of Manchester was used with permission to generate the final induction times. These data are transferred to Microsoft Excel 2013 for the final nucleation rate calculation. OriginPro 2015 and Microsoft Excel 2013 were used for plotting the graph.

2.9 Infrared Spectroscopy

Infrared spectroscopy can be used to study functional groups in molecular structure through variations of vibrational bands due to bond length and strength (force constant) [32]. Infra-red radiation leads to dipole excitation of characteristic molecular vibrational states [60], resulting in absorption at the corresponding characteristic frequencies The most common vibrational modes are bond stretching and bending as a result of changes in bond lengths and angles. The selection rules governing IR excitation state that a change in electric dipole must take place during the vibration. An example is the stretching vibration of carbon monoxide, CO [60]. IR is commonly scanned in a wide scan-range of 4000 to 600 cm^{-1}. The lower wavenumber means that the less energy is needed to stretch/bend the chemical bonds. For example, a single bond such as C−O has a lower wavenumber than the double bond, C=O. The main interesting factor in applying the IR in this research is the ability to detect the small changes in the wavenumber due to the environment of the bond, such as hydrogen bond interaction[32].

2.9.1 Attenuated Total Reflectance - Fourier-Transform Infrared (ATR-FTIR)

Attenuated total reflectance spectroscopy consists of an internal reflection element (IRE) that has a high refractive index. The sample that has lower refractive index than the IRE, is placed on the surface of this IRE with sufficient pressure to provide good contact. The infrared (IR) light penetrates the sample at a fixed depth of penetration, d_p, where an amount of IR light is absorbed by the sample and the remaining is reflected back as illustrated in Figure 2.1. The Smart Golden Gate, Thermo Nicolet used the diamond crystal as the IRE to measure the IR in the solid state.

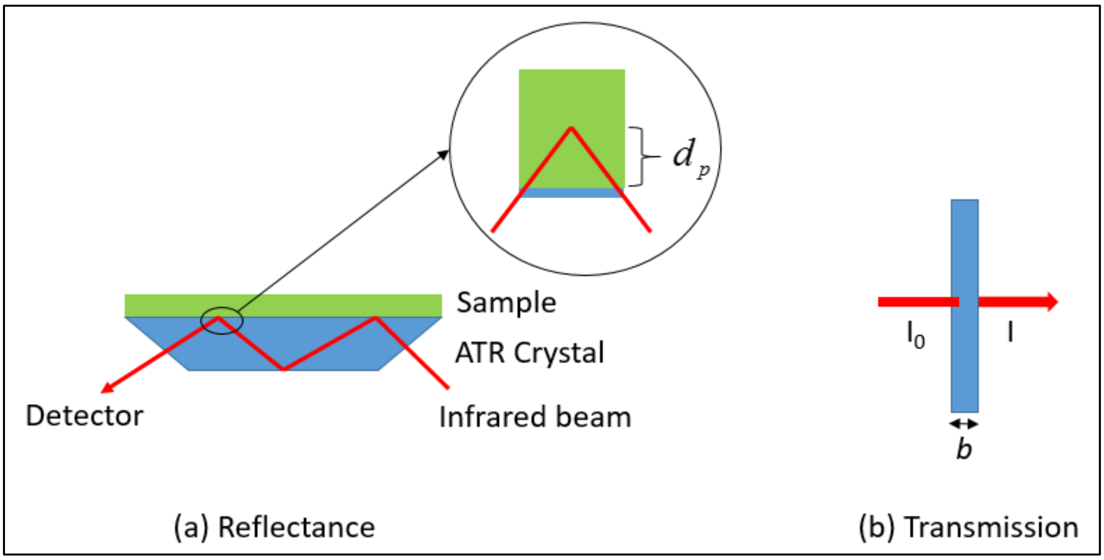


Figure 2.1: The illustration of the differences mechanism for (a) reflectance and (b) transmission IR spectroscopy.

2.9.2 Transmission Fourier-Transform Infrared (FTIR)

Transmission spectroscopy applies the Beer-Lambert Law that the absorbance, A , of the infrared radiation at a specific wavelength is proportional to the molar absorptivity, ε , path length, b , and concentration of the solution, c , as shown in Equation 2.2. Absorbance is a function of the transmitted light, T , which is then the ratio between the intensity of the transmitted light, I , to the intensity of the incident radiation, I_0 , as shown in Equation 2.3 and illustrated in Figure 2.1.

$$A = \varepsilon b c \qquad \textbf{Equation 2.2}$$

$$A = -\ln(T) = -\ln\left(\frac{I}{I_0}\right) \qquad \textbf{Equation 2.3}$$

A demountable cell was used to measure the IR spectra of the diluted solution for PNBA and BA in acetonitrile and ethanol using the Smart Omni-Transmission, Thermo Nicolet. The solution concentration ranged from 0.2 M to 0.0004 M. Figure 2.2 shows the assembly of the demountable cell which consists of the back plate, teflon spacers and gaskets, potassium bromide (KBr) windows, needle plate with two Luer-Lok teflon ports, teflon O-ring and knurled end cap. A thicker spacer has a longer path length [61]. The path length, b , is the distance the light travels through the speciation and as shown in Equation 2.2, the longer path length gives a higher absorption value at the same concentration. 0.1 mm or 0.2 mm thickness of the spacer was used. The solution was filled in the cell by using two syringes; one was filled with the solution and the other was empty. Both syringes were fitted into the Luer-Lok ports, the syringe with the solution was pushed into the cell while the empty syringed was pulled. This minimised excessive pressure on the inlet and secured the cell from damage.

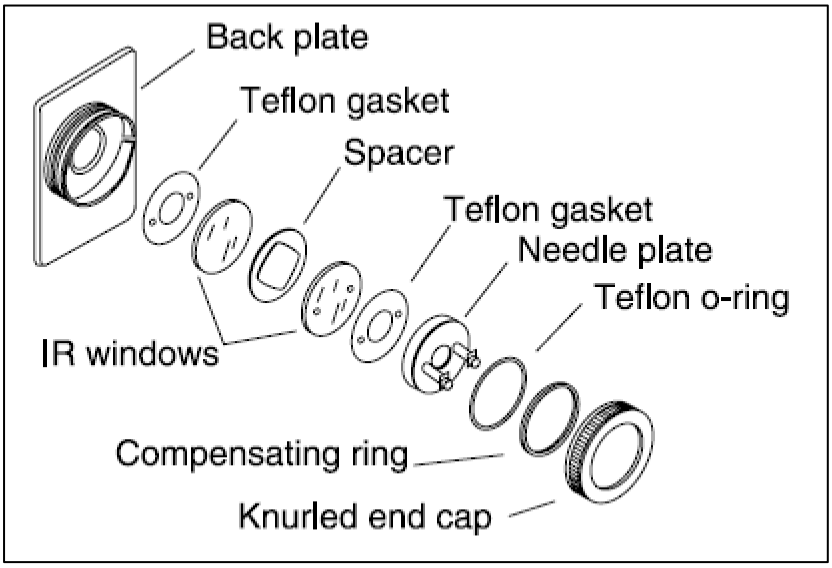


Figure 2.2: The assembly of the demountable cell where the spacer thickness can be varied to achieve different path length [62].

Software for the ATR-FTIR and Transmission FTIR: OMNIC 8 was used for measuring the spectra of the background and sample. The spectra were collected at a resolution of 4 cm^{-1}, 32 scans accumulated in the 4000-650cm^{-1} range. OriginPro 2015 was used to replot the final spectra, determining the peak position and the relative absorbance value.

2.10 Raman Spectroscopy

Raman spectroscopy is similar to the IR spectroscopy, which is based on the vibrational frequencies of the molecule. The mechanism of Raman Spectroscopy, however, is different to IR; Raman is based on the inelastic scattering of a monochromatic light from a laser [60]. The peak characteristic is similar to the IR. Raman scattering is weaker as a portion of the energy is absorbed by the molecules. Symmetric vibrations and covalent bond have a stronger Raman signal than IR [63]. In the IR spectra, the intensity is determined by the change in the dipole moment while in Raman, it depends on the changes in polarizability during the vibration. The polarizability is higher in a high concentrate of loose electrons of molecules, such as molecules with multiple bonds, C=C and O=N=O [63].

A Horiba Jovin-Yvon LabRam 300 confocal Raman microscope with the camera was used with a HeNe laser with an excitation line of 632.82 nm through a 100 to 400 µm range of aperture. The sample was placed on the microscope slide and visualised using microscope objectives of 10, 50x or 100x. Spectra were measured in the range 10-4000 cm^{-1} at 10 seconds per accumulation for 5 accumulations. These settings were adjusted accordingly to the quality of the spectra.

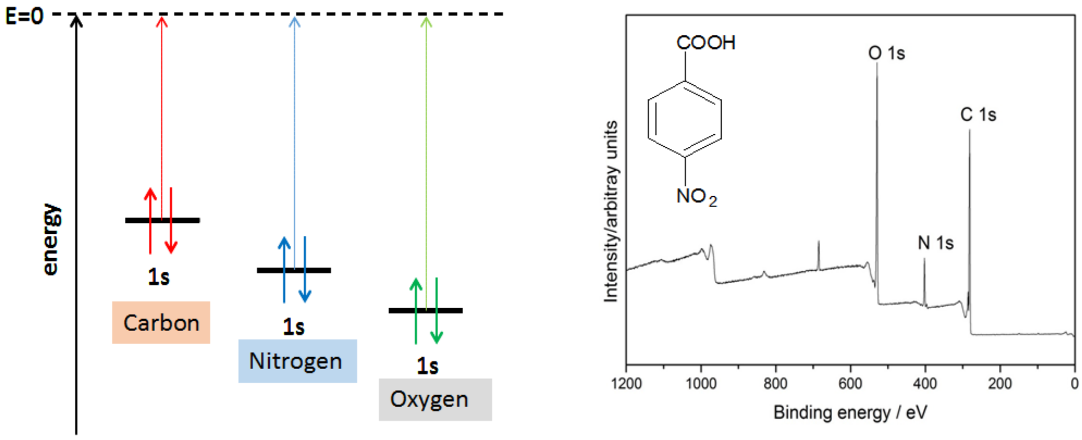
Software: Labspec 6 was used to perform an auto-calibration with silicon sample, visualising the sample, and measuring the spectra. OriginPro 2015 was used to plot the spectra and determine the peak position.

2.11 X-ray Photoelectron Spectroscopy (XPS)

X-ray photoelectron spectroscopy (XPS) is a technique used to identify the elemental composition, electronic structure and chemical environment of a sample, with the signal originating from about 10 nm in depth from the surface [64,65]. XPS provides the binding energy (BE) of the core level electron (such as 1s) by measuring the kinetic energy of the ejected photoelectron (KE) after a known photon energy, $h\nu$ is absorbed, $h\nu$=KE+BE [46]. The BE is also referred as the ionisation potential (IP).

A Kratos Axis Ultra instrument which employed the monochromatic Al K_α source (1486.69 eV) was used to measure the XPS spectra. The X-ray source has a power of 180 W (15 kV and 12 mA). About 1 mg of sample was fixed on the sample holder using double-sided tape. The measurement was carried out at a pressure below 10^{-8} mbar. Survey spectra (wide scans) were measured with an 80 eV pass energy over 1200 to -5 eV using 0.1 eV step and 1000 ms dwell time. High resolution spectra were measured with a 20 eV pass energy within the spectral range of interest, 0.1 eV steps and 1000 ms dwell time.

In addition, the interpretation of the XPS data is detailed here. Figure 2.3 is an example of a wide scan XPS result for PNBA that shows the BE for carbon C 1s, nitrogen N 1s and oxygen O 1s. The BE increases as the core level electrons are more tightly held by the nucleus, and so higher BE occurs for oxygen in the 530 to 540 eV region followed by nitrogen in the 395 to 410 eV region and carbon in the 280 to 295 eV region.



(a) (b)

Figure 2.3: (a) The sketch of energy level diagram for carbon, C 1s, nitrogen, N 1s, and oxygen, O 1s and (b) a wide scan results of PNBA highlighting the position of the C 1s, N 1s and O 1s.

Next, the high-resolution scan at each element provides the chemical environments and state of the atom through chemical shifts. For example, Figure 2.4 shows the three environments of the oxygen atom, C=**O**, N**O**$_2$ and **O**–H. The atom of interest is written in bold and underlined. The BE is influenced by the atom adjacent to the oxygen. The oxygen atom pulls the electron density from the adjacent atom which increases its electronegativity and consequently lowering the BE. Therefore, the BE for O 1s C=**O** is the lowest followed by O 1s N**O**$_2$ and O 1s C–**O**H. Relative to the C–**O**H and C=**O**, the peak area for O 1s N**O**$_2$ is double is because of the two O 1s environment, **O**–N=**O**. The curve labelled as the satellite is a shake-up transition of a valence electron to the unoccupied molecular orbital commonly found at higher BE and at the weak intensity in the aromatic system [66].

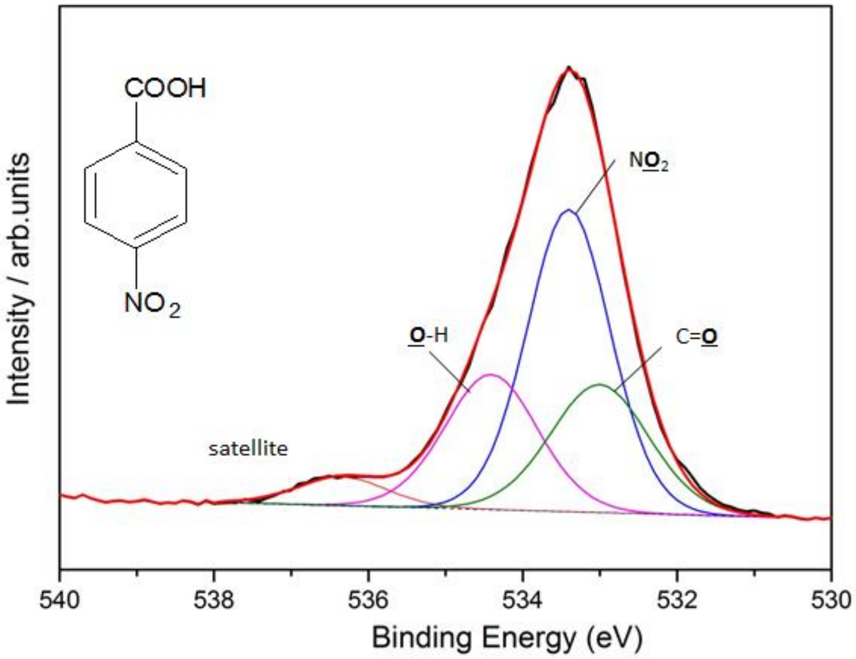
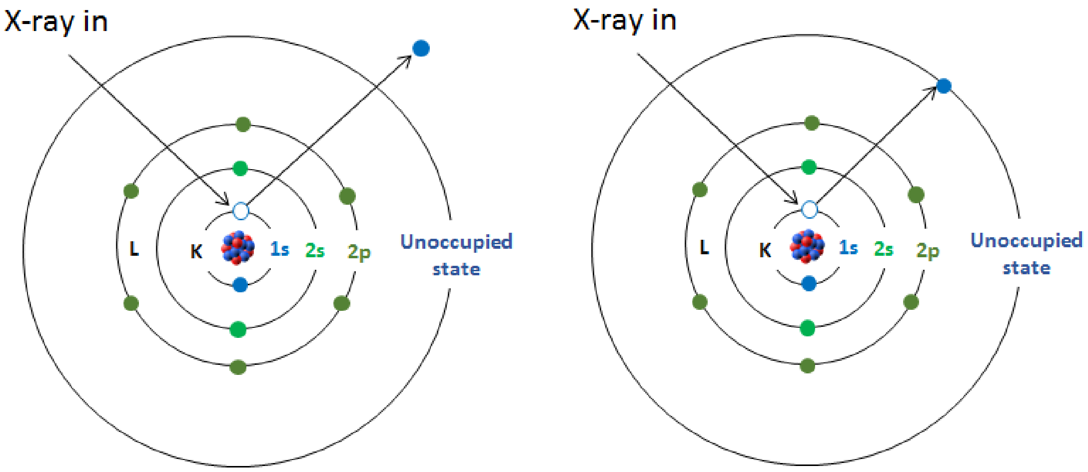


Figure 2.4: A high-resolution scan of for the oxygen and curve-fitting analysis shows the existence of all the chemical bonds, C=O, NO_2 and O-H for PNBA.

Software: CasaXPS was used to analyse the XPS results. A linear background was used and the curve-fitting is made based on the GL (30) line shape, 70 % Gaussian and 30 % Lorentzian. All the curves were calibrated to the C=C peak, ideally at 284.8 eV. OriginPro 2015 was used to replot the whole spectra including the fitted peaks.

2.12 Near-edge X-ray Absorption Fine Structure (NEXAFS) Spectroscopy

Similar to the XPS, near-edge X-ray absorption fine structure spectroscopy (NEXAFS) is used to understand the structure of the molecules. Figure 2.5 illustrates the differences between the XPS and NEXAFS. In XPS, the core level electrons absorb enough energy to be ejected out of the orbitals, 1s → vacuum, while in NEXAFS the core level electrons absorb the energy for the transition to the lowest unoccupied molecular orbital (LUMO), 1s → LUMO. NEXAFS is thus particularly sensitive to anything resulting in variation in electronic structure. This means that NEXAFS detects the presence of the same chemical bonds as in XPS, for example the C=C, C=O, O−H, but additionally the bond lengths between these bonds can also be calculated in favourable cases[67].



(a) XPS mechanism (b) NEXAFS mechanism

Figure 2.5: The illustration of 1 s electron being (a) ejected out of the orbital when the BE is overcome and (b) excited to the unoccupied molecular orbital after absorbing the sufficient energy. (a) and (b) represent the mechanism for XPS and example for NEXAFS.

Figure 2.6 illustrates the difference between XPS and NEXAFS through an orbital energy diagram, with XPS yielding the BE for the 1 s → vacuum and NEXAFS the energy for 1 s → 1 π* transitions. As the XPS and NEXAFS data are measured using different spectrometers, additional experimental relative shifts in the binding and photon energy scales are expected. There is the possibility that the exact relative locations of bands in the XPS and NEXAFS results are ambiguous results as the calibration procedure for the energy scale is different and can also be laborious. Therefore, the XPS and NEXAFS data are either analysed by roughly estimating the shift in energy or using the relative difference rather than the absolute number.

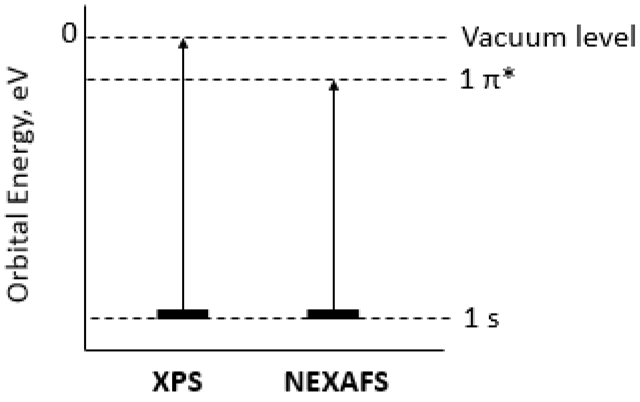


Figure 2.6: The orbital energy sketch of the 1s and 1π* is illustrated by combining the XPS and NEXAFS data.

NEXAFS measurement for the C and N K-edge were performed for PNBA Form I and Form II with partial electron yield (PEY) and an entrance grid of -50V for C and -150V for N on the U7a beamline of the National Synchrotron Light Source (NSLS), New York. O K-edge total electron yield (TEY) was performed on the U49-2 PGM beamline of the BESSY II, Berlin. The solution state for O K-edge partial electron yield (PEY) was measured at the U41-PGM beamline of the BESSY II, Berlin using the liquid microjet technique. Solutions of PNBA and BA in ethanol at 0.10 M and 0.15 M, and PNBA and BA in acetonitrile at 0.02 M and 0.5 M, were prepared and filtered at room temperature. An 18 µm diameter glass nozzle with 0.6 mL min^{-1} flow rate was used.

The interpretation of the NEXAFS spectra is similar to the XPS data. The sharp and narrow peaks at the lower energy, which is the transition of 1 s to the π* molecular orbital are the main interest for this thesis as they can be particularly sensitive to local environment. The broader area of the higher energy is the transition to the σ* orbital. Figure 2.7 is an example of a NEXAFS spectra fitted with peaks. The peaks are assigned to each chemical bond in the same energy order as in XPS data and the ionisation potential is the edge step.

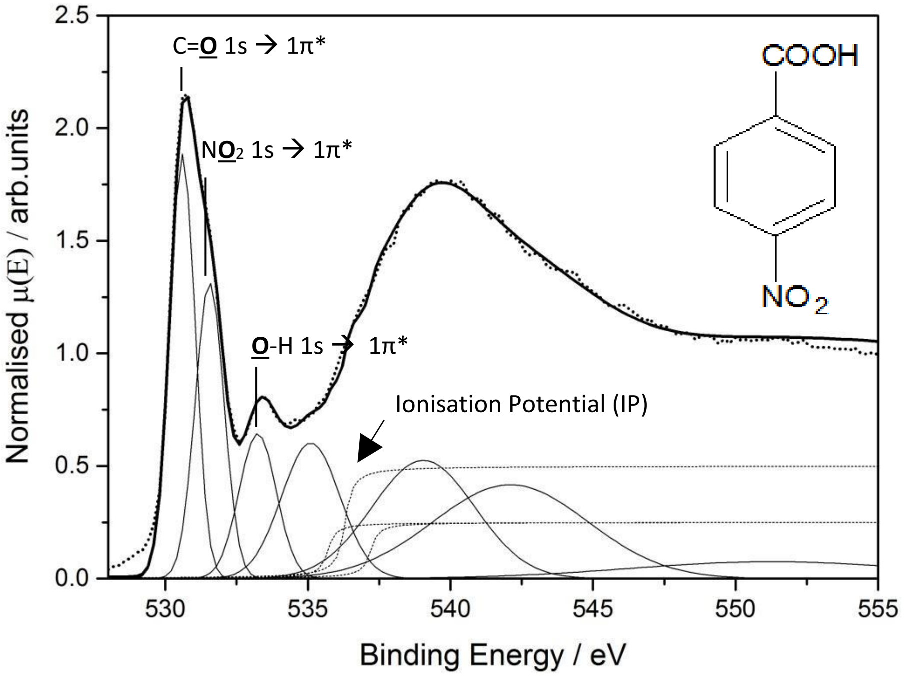


Figure 2.7: A complete curve fitting analysis of the NEXAFS spectra on the O K-edge for PNBA lighting the main peaks of interest.

Software: Athena was used in analysing the NEXAFS spectra. Firstly, the position of the IP or the 1 s BE, was roughly estimated as around 4 eV higher than the XPS. The peaks under the area were then fitted using the Gaussian lines. OriginPro 2015 was used to replot all the spectra and the fitted peaks.

2.13 Gaussian 03 (Computational Chemistry)

Gaussian 03 [68] was used to calculate the predicted molecular orbitals energy based on density functional theory (DFT) for PNBA Form I, PNBA Form II, BA and PABA. These predicted orbital energy calculation guided the interpretation of the XPS and NEXAFS experimental data and visualises the molecular orbitals. The input file for the calculation was a single molecule or a monomer structure from the CSD. The geometry of this monomer was optimised at ground state using the settings of B3LYP / 6-31G (d) before an energy calculation was made. However, in order to differentiate the polymorphs of PNBA Form I and Form II, the energy was calculated without the geometry optimisation.

3. POLYMORPHISM AND ASSOCIATED THERMODYNAMICS OF *PARA*-NITROBENZOIC ACID (PNBA)

Abstract: *p*-Nitrobenzoic acid (PNBA) has two polymorphs, PNBA Form I and commercially available Form II. Solid and slurry/solution state characterisation of both polymorphs has been performed employing powder and single crystal XRD, FTIR, DSC, hot stage microscopy and solubility measurements. The solid state phase transition from Form I to Form II occurs at 122 °C. Solubility and slurry mediated phase transformations between the polymorphs in ethanol, 2-propanol, acetonitrile and ethyl acetate have been performed. The polymorphs are enantiotropically related with a transition temperature at 51 °C, below which Form I is the stable form. PNBA polymorphic transformations are driven by the high entropy factor at 11.6 J mol^{-1} K^{-1} while the enthalpy is at 3.8 kJ mol^{-1}.

3.1 Introduction

Polymorphs of a compound have the same chemical composition, but more than one crystal structure. A change in crystal structure, phase transitions, can occur directly in the solid state as well as when the solid is in contact with its solution[69,70]. Temperature, pressure, relative humidity and mechanical stress can induce solid-state phase transitions[69]. Changes in the polymorphic form will result in different physical properties of a solid, such as solubility, melting temperature, density, dissolution rate and other properties[70]. Therefore, in the pharmaceutical industry, the control of polymorphs is essential in each step, from upstream to the downstream processes [1,69].

X-ray diffraction, thermal analysis, spectroscopy and microscopy methods are commonly used to characterise polymorphs [71]. X-ray diffraction (XRD) in particular reveals the difference of molecule arrangements in different unit cells, while differential scanning calorimetry (DSC) determines the temperature of solid state phase transitions and the melting temperature of the stable form. Vibrational spectroscopy techniques such as infrared (IR) and Raman are sensitive to differences in the molecular level structure as well, while optical microscopy visualises the macroscopic morphology[11,69,71,72]. Solubility measurements distinguish thermodynamically stable polymorphs, which are less soluble, from metastable states, which are more soluble. The thermodynamic parameters for polymorph phase transition in the solution state, i.e., enthalpy, entropy and Gibbs free energy, are calculated from the solubility and calorimetric data [2,7,10,11,14,17] (as discussed in Section 1.2).

In an ideal system, the solubility is predicted by assuming the heat of dissolution is equal to the heat of fusion of the pure solute, ΔH_f, [2,3,7]as given in the van't Hoff Equation (Equation 3.1). This equation is an approximation since it does not account for the temperature effect on ΔH_f which needs the heat capacity, ΔC_p. This means that care must be exercised in comparing measured and calculated solubilities.

$$\ln x_{ideal} = \frac{\Delta H_f}{R}\left[\frac{1}{T_f} - \frac{1}{T}\right]$$ **Equation 3.1**

In interpreting measured solubility data it is common to use the van't Hoff equation as written as Equation 3.2. ΔH_d, the heat of solution for each polymorph is obtained from the slope of the experimental plot of $\ln x$ against one over temperature. Since formally this relationship involves activities coefficient rather than mole fraction, care should be taken when interpreting the values obtained. The molar Gibbs free energy of the dissolution, ΔG_d, and entropy of dissolution, ΔS_d, is then calculated at polymorphic transition temperature, T_r, using Equation 3.3 and Equation 3.4. This is also discussed in Section 1.2.

$$\ln x = -\frac{\Delta H_d}{RT} + \frac{\Delta S_d}{R}$$ **Equation 3.2**

$$\Delta G_d = -RT_r \ln X$$ **Equation 3.3**

$$\Delta S_d = \frac{\Delta H_d - \Delta G_d}{T_r}$$ **Equation 3.4**

The enthalpy and entropy of polymorphic transition, ΔH_T and ΔS_T, are then calculated as the differences between ΔH_d and ΔS_d for each polymorph respectively (Section 1.2). Alternatively, the ΔS_T can be obtained by combining Equation 3.5 and Equation 3.6. ΔS_t is calculated from the slope of the plot $RT\ln(x_{II} / x_I)$ against temperature [10]. Equation 3.7 is then applied to calculate the ΔH_T, as at the transition temperature, T_r, the molar Gibbs free energy of the transition, ΔG_T, is zero.

$$\left(\frac{\partial \Delta G_T}{\partial T}\right)_P = -\Delta S_T$$ **Equation 3.5**

$$\Delta G_T \cong RT \ln \frac{x_{II}}{x_I}$$ **Equation 3.6**

$$\Delta G_T = \Delta H_T - T_r \Delta S_T$$ **Equation 3.7**

p-nitrobenzoic acid (PNBA) is an aromatic molecule with a carboxylic acid and a nitro group, as shown in Figure 3.1. PNBA is used as an intermediate in organic and biological synthesis[73,74]. The nitro group is a strong electron withdrawing group and a poor hydrogen bond acceptor, resulting in preference not to be hydrogen bonded [75,76]. The carboxylic acid has both hydrogen bond donor (O−H) and acceptor (C=O) capabilities.

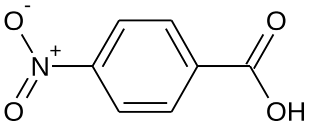


Figure 3.1: The molecular structure of *p*-nitrobenzoic acid (PNBA)

Currently two polymorphs of PNBA are represented in the Cambridge Structural Database (CSD). The earliest crystal structure of PNBA deposited in the CSD was in 1966 [77]. Later structure analysis in 1971 included hydrogen positions [78]. XRD analysis in 1977 was performed on a PNBA single crystal grown from evaporation of ethanol solution. The measurement at room temperature showed that the crystal belongs to an A2/a space group, with unit cell dimensions of *a*=12.918, *b*=5.042 and *c*=21.298 Å[57]. The intramolecular bond lengths of the C=O moiety and the O···O distance across the intermolecular hydrogen bond are 1.234 and 2.660 Å respectively. A similar measurement was performed in 1993 at 120 K [58]. The unit cell dimensions slightly change to *a*=12.857, *b*=5.0272 and *c*=20.997 Å while the bond length for C=O and O···O are slightly shorter at 1.231 Å and 2.648 Å, respectively. The latest deposition was in 2015, using a PNBA crystal that was grown from slow evaporation of acetonitrile solution [75]. Performed at 100 K, and the analysis showed a monoclinic, C2/c space group, with a unit cell dimension of *a*=20.98, *b*=5.03 and *c*=12.85 Å, and bond lengths for C=O and O···O of 1.232 Å and 2.649 Å. The second polymorph of PNBA was first described in 1980 [56]. At 123 K, a monoclinic system with a P_{21}/c space group was found, with a unit cell dimension of *a*=5.403, *b*=5.153 and *c*=24.692 Å. The bond length for

C=O and O···O were at 1.212 Å and 2.628 Å respectively.

In this chapter, the main aim is to understand the polymorphism and the crystallisation process of PNBA. A comprehensive solid state characterisation of both polymorphs in addition to the existing crystallography data is described. Solubility data in different solvents is used to determine the thermodynamic properties and relationship between the polymorphs.

3.2 Materials and Methodologies

The materials used are listed in Chapter 2 - Section 2.1. The methodologies used are powder and single crystal X-ray diffraction (XRD), gravimetric analysis, slurry transformation, differential scanning calorimetry (DSC), hot stage microscopy, infrared and Raman spectroscopy as described in Chapter 2.

3.3 Results and Discussion

3.3.1 Solid State Characterisation of PNBA Form I and II

3.3.1.1 Crystal Structures

The crystal structures of PNBA Form I (CSD Reference ID: NBZOAC03, [57]) and Form II (CSD Reference ID: NBZOAC02, [56]) as obtained from the CSD are shown in Figure 3.3 and Figure 3.3 respectively. PNBA Form I has the space group of A2/*a* while Form II has P2_1/*c*. Both crystal structures have a symmetric carboxylic acid dimer.

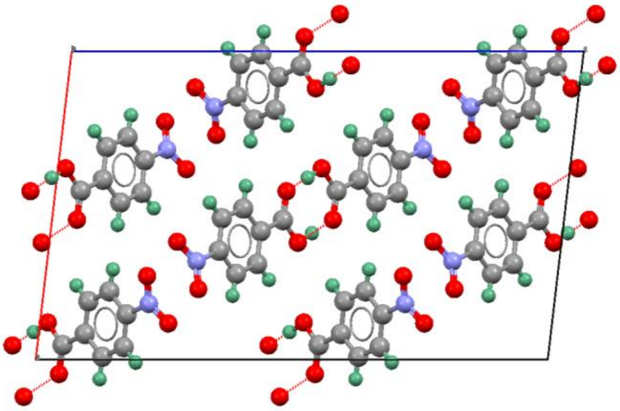


Figure 3.2: The packing of PNBA Form I (NBZOAC03) looking down the b-axis of the unit cell.

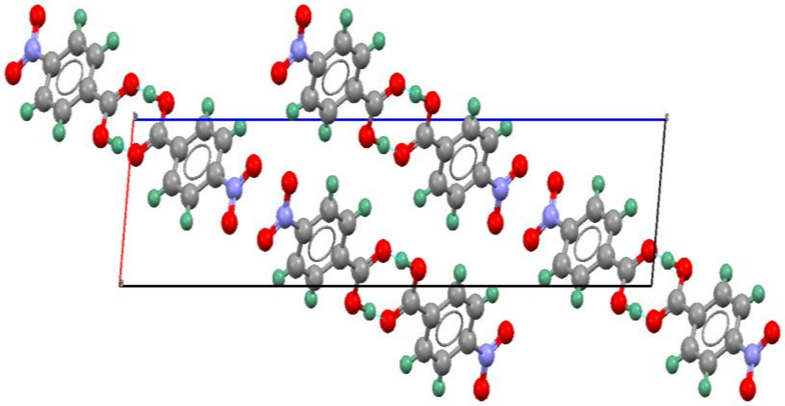


Figure 3.3: The packing of PNBA Form II (NBZOAC02) looking down the b-axis of the unit cell.

Table 3-1 summarises the key bond length and angles for both polymorphs. The bond lengths between the two hydrogen-bonded oxygen atoms, **O**−H···**O,** are 2.66 Å and 2.62 Å for PNBA Form I and II, respectively. The O···O distance is within 2.5-3.0 Å, thus it is classified as a strong hydrogen bond [79]. There is a 0.014 Å difference between the C=O bond lengths, while the bond lengths for C=C, C−N and NO are comparable at 1.4, 1.5 and 1.2 Å for both polymorphs. The torsion angles between the carboxylic acid and nitro group are also similar, differing by only 1° between the two polymorphs.

Table 3-1: A summary of the geometry for the main functional groups (carboxylic acid and nitro) for PNBA Form I and II.

Bond Length	**PNBA Form I, Å**	**PNBA Form II, Å**
C=O	1.23	1.25
O−H···O	2.66	2.62
Angle	**PNBA Form I, °**	**PNBA Form II, °**
COO	123.1	122.3
NO_2	124.2	123.8

The two crystal structures of PNBA Form I and Form II are overlayed as showed in Figure 3.4 (top). This figure shows that both structures are built from an identical hydrogen bonded carboxylic acid dimer. In addition it is clear from the overlay of the packing arrangement that while the middle layer of the crystal packings are similar the top and bottom layers are different. Extracting the middle and bottom layer (Figure 3.4 (bottom)), it is clearly shown again that these structures overlay with each other at the middle layer and the difference between the bottom layer is mainly due to a shift in the arrangement of the adjacent layers.

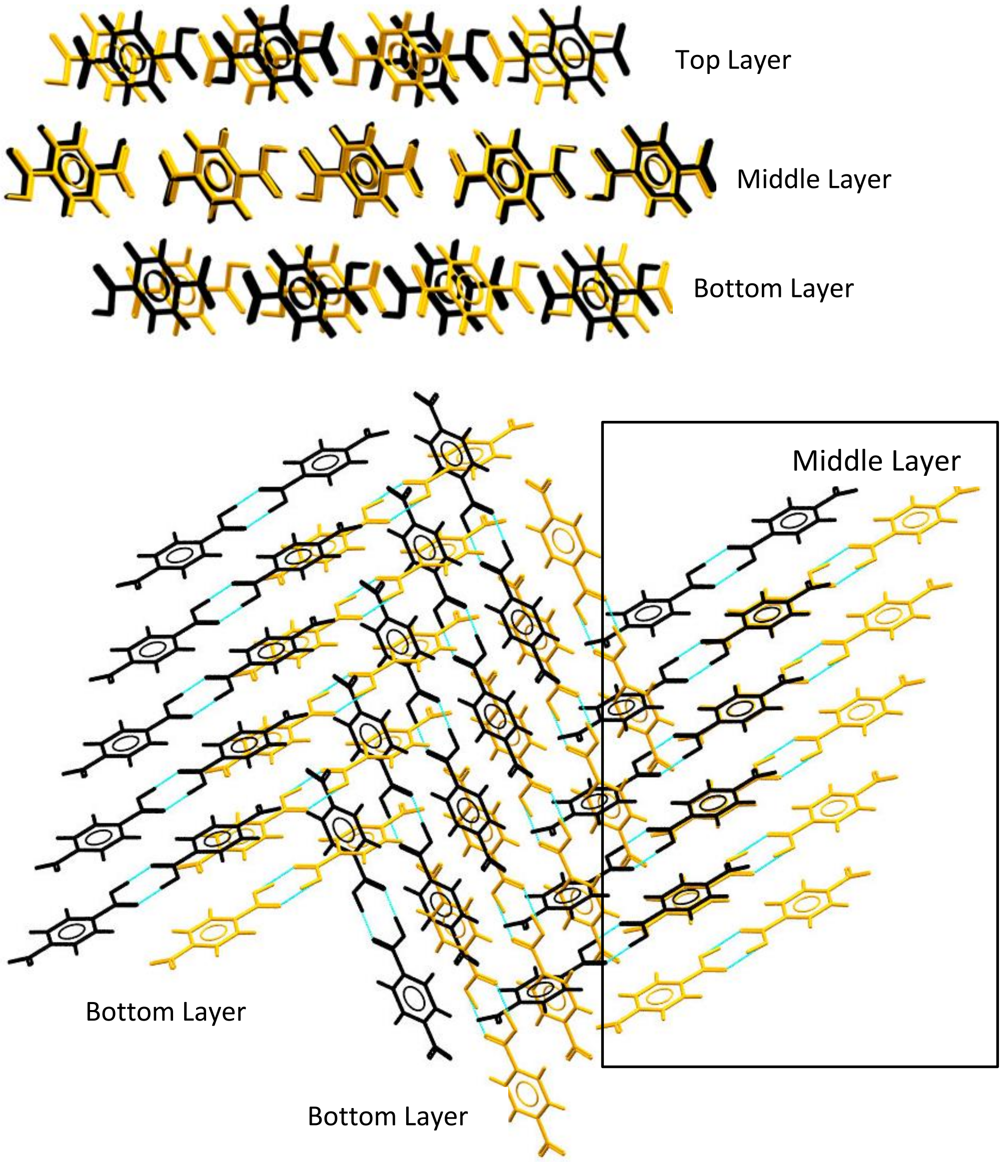


Figure 3.4: (Top) the overlay of PNBA Form I (black, NBZOAC03) and PNBA Form II (orange, NBZOAC02) in three different layers which highlights the difference between the two polymorphs in terms of packing arrangement is at the top and bottom layer. (Bottom) The two structures overlay in the middle layer while at the bottom layer the structures arrangement shift to each other.

3.3.1.2 Powder X-Ray Diffraction (XRD)

The powder XRD pattern of PNBA Form I matches the CSD structure NBZOAC03, as shown in Figure 3.5. The main peaks that characterise PNBA Form I are a sharp peak at 17°, a less intense peak at 25°, a shoulder at 27.5° followed by a sharp peak at 27.8° and another less intense peak at 29.1°.

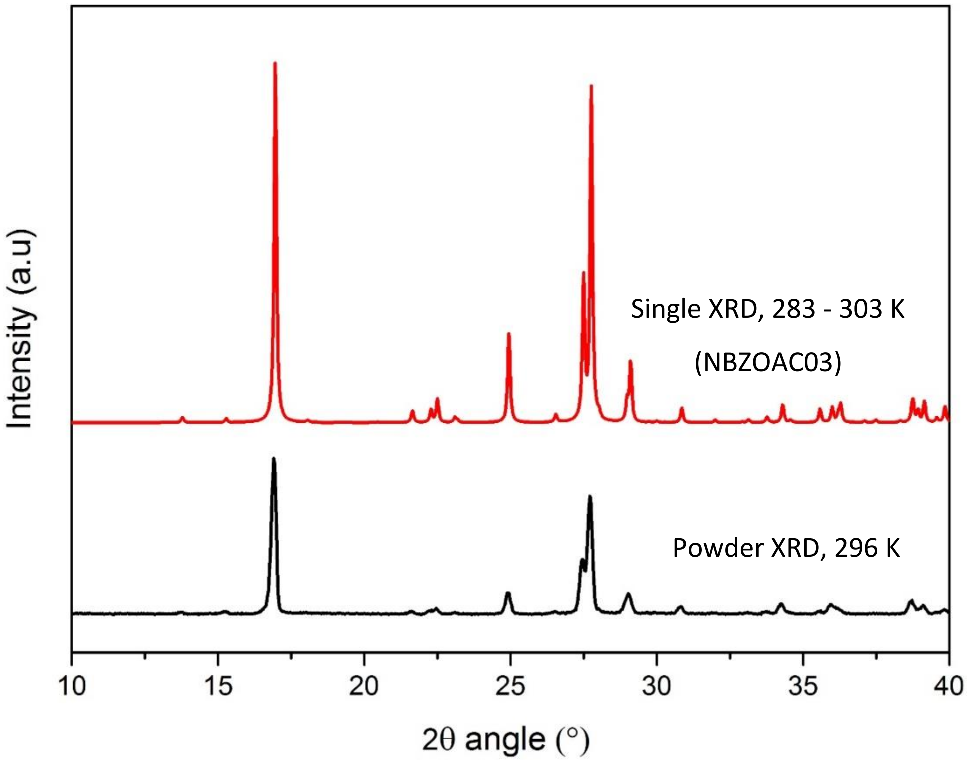


Figure 3.5: Powder XRD pattern of PNBA Form I (black) and NBZOAC03 (red). The characteristic peaks are at 17°, 25°, 27.5°, 27.8° and 29.1°.

The XRD pattern of PNBA Form II shows a temperature dependant effect. Single crystal and powder XRD of PNBA Form II at room temperature (296 K) have a sharp peak at 17.1°, two least intense peaks at 24.9° and 25.2° and three peaks at 26.7°, 27.7°and 28.6°. Single crystal XRD and NBZOA C02 at 100 K and 123 K have similar peaks, but at 17.2°, 25.1°, 25.6°, 27.2°, 27.8° and 29.1°. Figure 3.6 shows that these XRD patterns have the same characteristic peaks but the position shifts by 0.2 to 0.5° at different measurement temperatures.

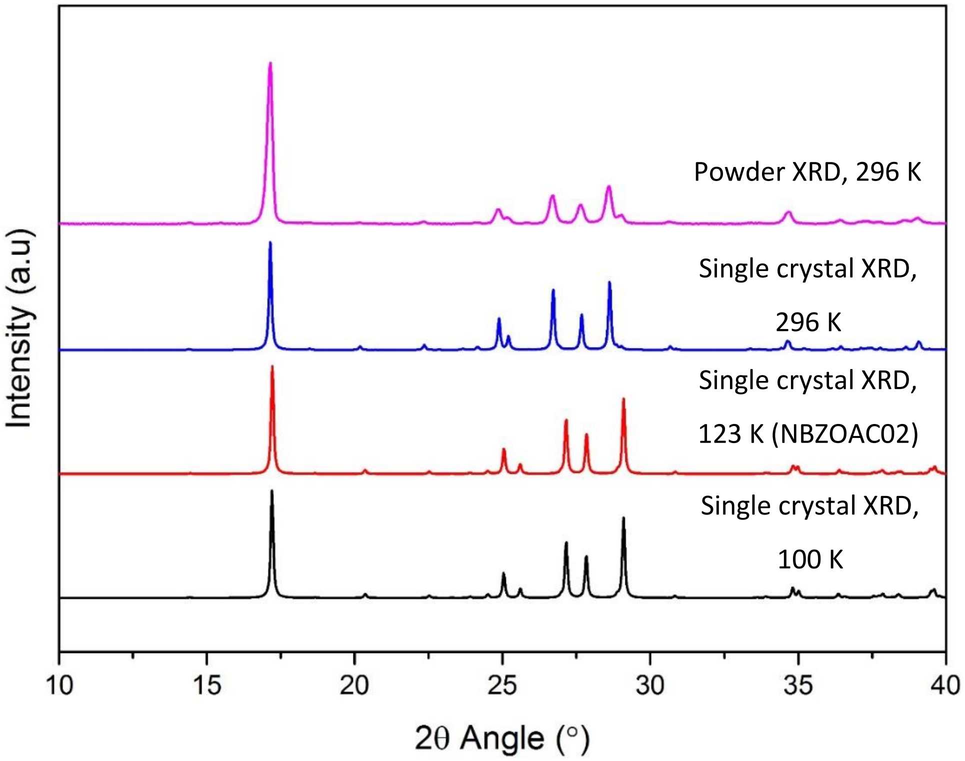


Figure 3.6: Powder XRD pattern (pink) and single crystal XRD at room temperature (blue) and single XRD at 100 K (black) of PNBA Form II compared to NBZOAC02 (red). The peaks position based on powder XRD (pink) are at 17.1°, 24.9°, 25.2°, 26.7°, 27.7°and 28.6°.

3.3.1.3 Fourier Transform Infra-Red (FTIR) and Raman Spectroscopy

Infrared spectroscopy is a common technique used in understanding hydrogen bonded interactions. FTIR spectra of PNBA Forms I and II has similar peak positions but at different intensities, as shown in Figure 3.7. A similar observation is seen in Raman spectra as shown in Figure 3.8. The similarity in the peak positions is anticipated, as the molecular structures of both forms are the same, centrosymmetric carboxylic acid dimer. In the dimer, the O−H out-of-plane deformation peak position is at around 930 cm^{-1} [36]. PNBA Form II has this signal at 924 cm^{-1} while PNBA Form I is at 918 cm^{-1}. Apart from the peak position, PNBA Form II has a broader peak compared to Form I with a shoulder-like broad feature. The higher wavenumber indicates that PNBA Form II dimer has weaker hydrogen bonding. PNBA Form I has higher peak intensity in both FTIR and Raman spectra.

The carbonyl (C=O) stretching has a strong signal in FTIR at 1684 cm^{-1} (Figure 3.7) but a weaker signal in Raman at 1640 cm^{-1} (Figure 3.8). The difference between these two signals, antisymmetric and symmetric stretching, is 44 cm^{-1} and agrees with previous works [80,81] in characterising carboxylic acid dimer. The nitro group vibration is similar between the polymorphs and differs only by 1-2 cm^{-1}. The NO_2 symmetric and anti-symmetric stretching lead to a sharp peak at 1349 cm^{-1} and 1537 cm^{-1} respectively in FTIR as shown in Figure 3.7. In Raman, the symmetric peak is strong at 1353 cm^{-1}, but a weak signal occurs for the anti-symmetric stretch at 1526 cm^{-1} as shown in Figure 3.8.

Another characteristic differentiating the two polymorphs is a lower Raman shift, below 200 cm^{-1}. In the case of benzoic acid, the peaks at 114 cm^{-1}and 122 cm^{-1} are assigned to the O−H···O out-of-plane bending and O−H···O stretching [82]. Here (Figure 3.8), PNBA Form I has a shoulder and a sharp peak at 145 cm^{-1} and 121 cm^{-1} while PNBA Form II has a single peak at 118 cm^{-1}. Even though this may seem a small difference, it has been found to be highly reproducible and this it is a reliable characteristic to differentiate between the two polymorphs.

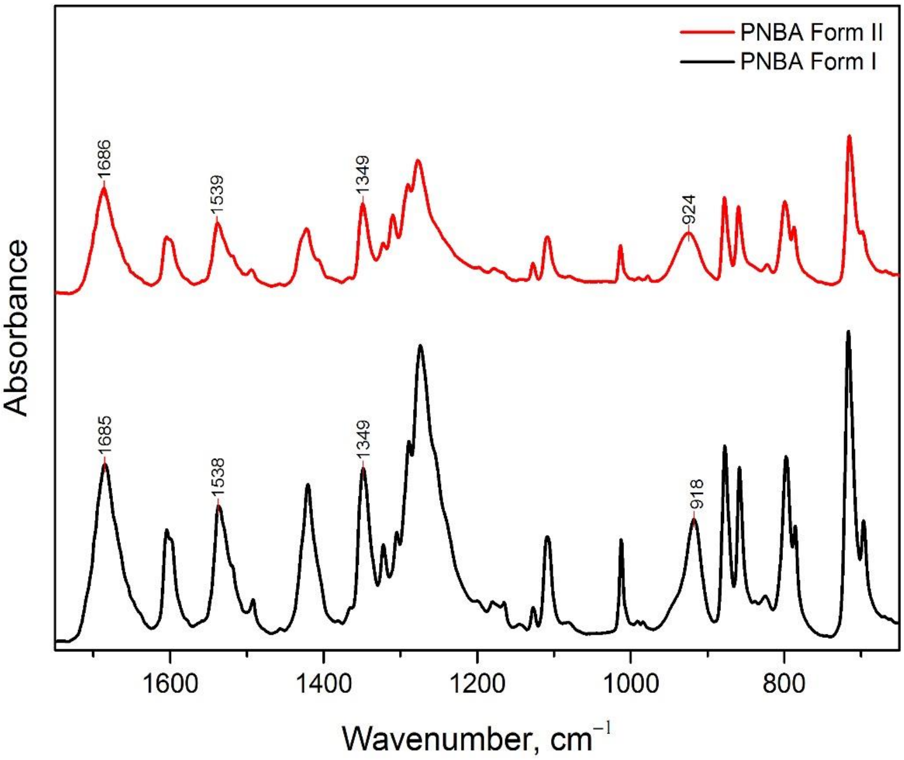


Figure 3.7: FTIR spectra for PNBA Form I (black) and Form II (red) in the 750 to 1750 cm^{-1} region.

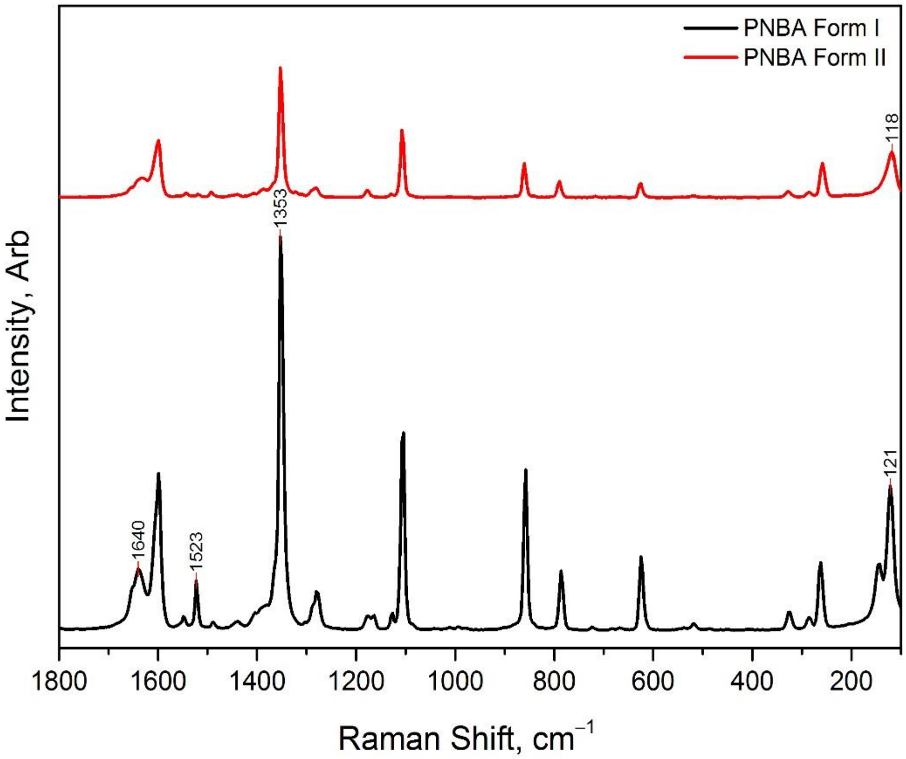


Figure 3.8: Raman spectra for PNBA Form I (black) and Form II (red) in the 100 to 1800 cm^{-1} region.

3.3.1.4 Differential Scanning Calorimeter (DSC)

DSC provides the melting temperature, the heat of fusion and phase transition temperature for the polymorphs. Thermograms for PNBA Form I and II were measured for comparison. The melting temperature was determined from the peak and heat of fusion was based on normalised values. Form I undergoes the phase transition to Form II at 129°C (Figure 3.9 (black)), before reaching the melting point, thus there is no available data on its melting temperature and heat of fusion. The enthalpy and the entropy of transition are 1.96 kJ mol^{-1} and 4.88 J mol^{-1} K^{-1} respectively based on the average of eight data sets.

Form II has a melting temperature of 244.7 °C and a heat of fusion of 29 kJ mol^{-1} (Figure 3.9 (red)). The temperature value compares well with the literature value of 239.2°C, while the previously reported heat of fusion is higher at 36.9 kJ mol^{-1} [83]. The difference might be related to complications arising from the sublimation of PNBA (see next section), which may set up a solid-gas equilibrium before melting, which is likely to be poorly reproducible between calorimetry experiments. Based on the results obtained here the entropy of fusion for PNBA Form II , $\Delta S_f = \frac{\Delta H_f}{T_f}$ is 55.7 J mol^{-1} K^{-1}.

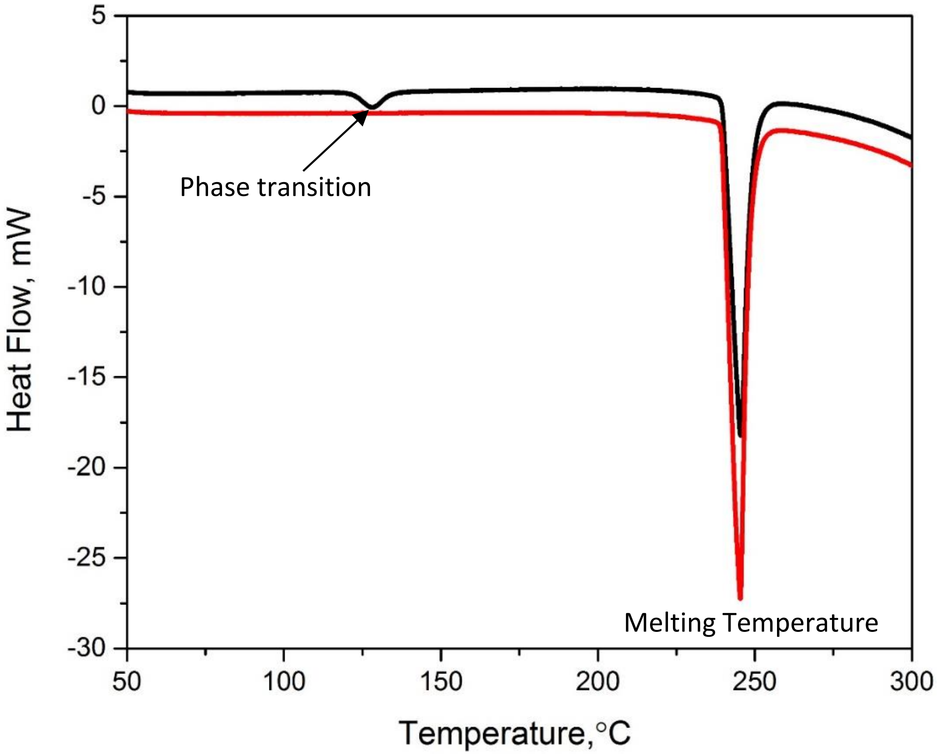
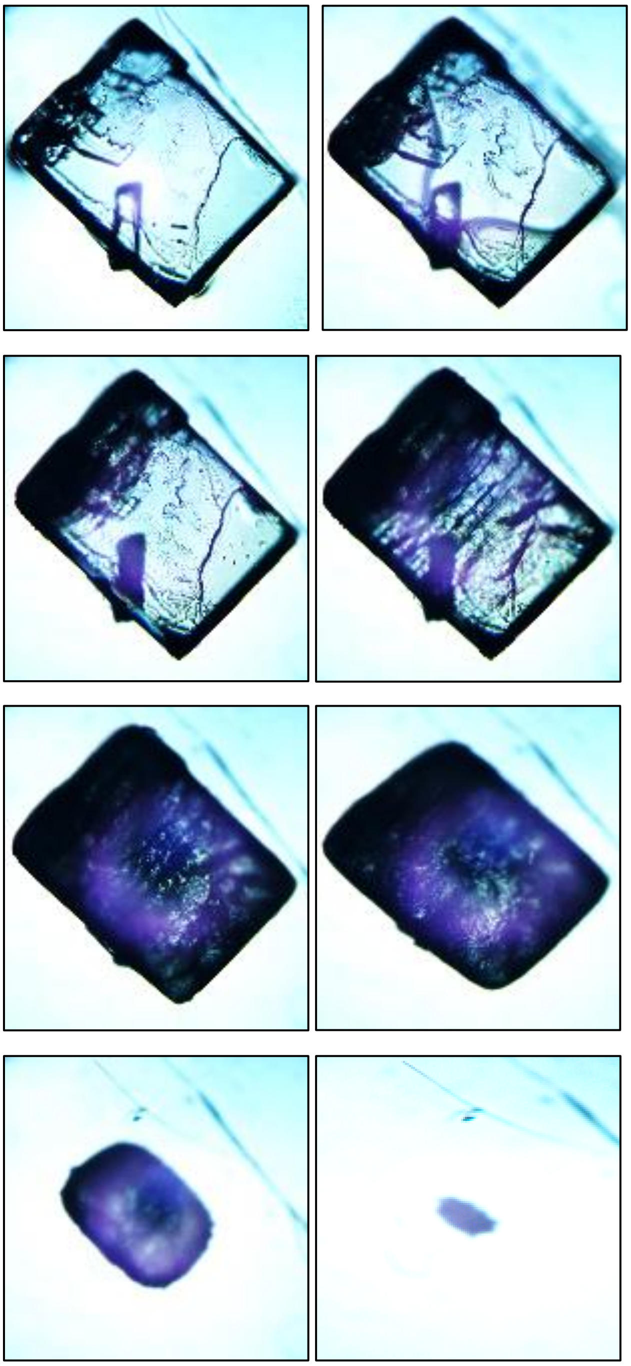


Figure 3.9: DSC pattern for PNBA Form I (black) and PNBA Form II (red)

3.3.1.5 Hot Stage Microscopy

Hot stage microscopy provides a visual view of the PNBA Form I transformation to Form II as the temperature increased. The crystal was heated from room temperature to 190 °C at a rate of 3 °C min^{-1}. At 120 °C to 130 °C, a wavefront moving smoothly through the crystal occurs, resulting in darkening of the areas it has transitioned over, e.g. from top to bottom as shown in Figure 3.10 (c) and (d). This change in appearance is associated with the the transformation of PNBA Form I to Form II. The overall shape of the Form II (Figure 3.10 (e)) crystal remains the same as the initial Form I shape. As the temperature increases to 190 °C, the crystal starts to shrink and evaporates (Figure 3.10 (f), (g) and (h)), suggesting a sublimation process. Repeats made on different crystals as well as accompanying powder XRD measurements at 130 °C confirmed the formation of PNBA Form II.



(a) Left: Starting crystal Form I at 30°C

(b) Right: Crystal Form I at 85°C

(c) Left: Crystal Form I at the top starts to darken at 120°C

(d) Right: Crystal continues to darken to the bottom at 125°C

(e) Left: Crystal Form II at 160°C

(f) Right: Crystal Form II at 180°C starts to shrink.

(g) Left: Crystal Form II continues shrinking at 189°C

(h) Right: Crystal Form II complete sublimating by 190°C

Figure 3.10: Microscopic image of PNBA Form I at 30 °C (a), 85 °C (b), and a phase transformation to PNBA Form II starts to occur at 120 °C (c) and continues at 125 °C (d). At 160 °C (e), the crystal shape remains the same but it darkens and the sublimation process starts to occur evidenced by the shrinking of the crystal as shown at 180 °C (f), 189 °C (g) and 190 °C (h).

3.3.1.6 *In –situ* FT-IR measurement on heated stage

On a separate experiment, the powder form of PNBA Form I was heated from 40 °C to 190 °C and IR spectra were taken at each 5 °C interval. Figure 3.11 shows that as the temperature increases from 40 °C to 95 °C, the O−H out-of-plane peak of PNBA Form I at 917 cm^{-1} starts to shift to the lower wavelength, 913 cm^{-1}, due to the temperature effect [84]. Then, from 100 °C to 190 °C, the peak position shifts to the higher wavelength at 921 cm^{-1} suggesting a phase transition to Form II. In contrast, PNBA Form II shows a continuous temperature effect; the peak at 928 cm^{-1} (40 °C) shifts to lower wavelength as the temperature increases and reaches 918 cm^{-1} at 190 °C. The powder starts to sublime from 160 °C onwards. Figure 3.11 further supports the hot stage microscopy results (Figure 3.10) on the solid state transition of PNBA Form I to Form II which starts at 100 °C.

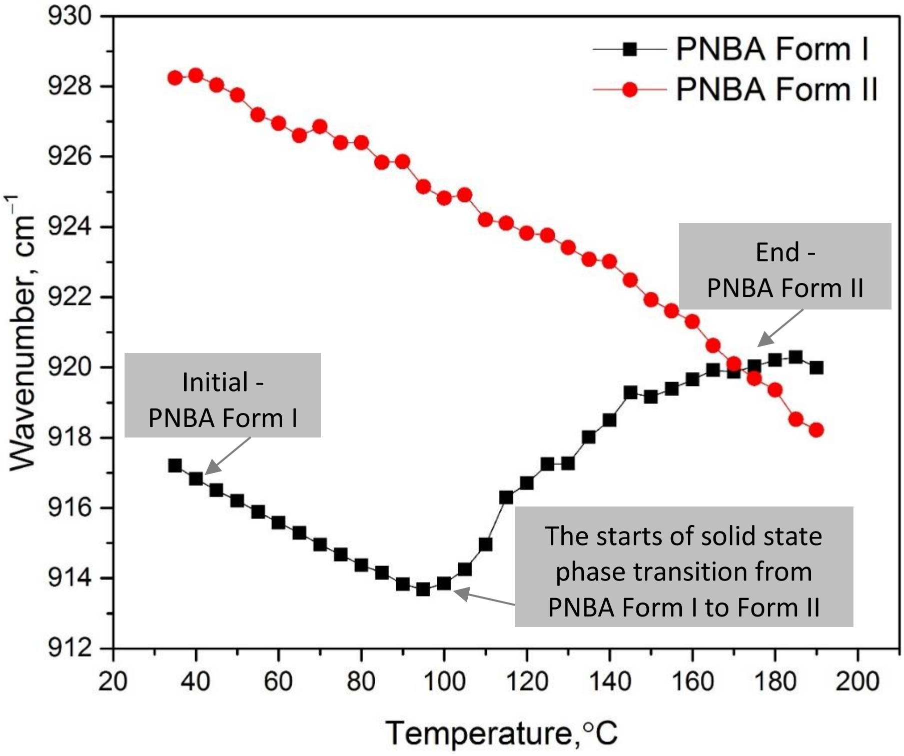


Figure 3.11: The O−H out-of-plane peak position decrease from 917 to 913 cm^{-1} as the temperature increase from 40 to 90 °C and starts to increase to 921 cm^{-1} as the temperature increases from 100 °C to 190 °C.

Figure 3.12 shows the IR spectra between 900 to 960 cm^{-1} at several different temperatures (40, 80, 100, 110, 120, 140 and 180 °C). The initial spectrum at 40 °C (Figure 3.12 (black)) shows the O−H out-of-plane characteristic of PNBA Form I; a sharp peak at 917 cm^{-1} and a small shoulder at 941 cm^{-1}. At 80 °C (Figure 3.12 (yellow)), the peak shifts to 913 cm^{-1} but the small shoulder remains at 940 cm^{-1}. As the temperature increases from 100 °C to 190 °C (Figure 3.12 (blue to red)), the sharp peak position increases towards 921 cm^{-1} while the intensity of the small shoulder increases, resulting in a broad O−H out-of-plane which is the characteristic of PNBA Form II. This shows that a subtle difference between the two polymorphs is at the O−H···O interaction. As the change in O−H···O interaction is small, PNBA Form I transforms to Form II smoothly as shown in Section 3.3.1.5.

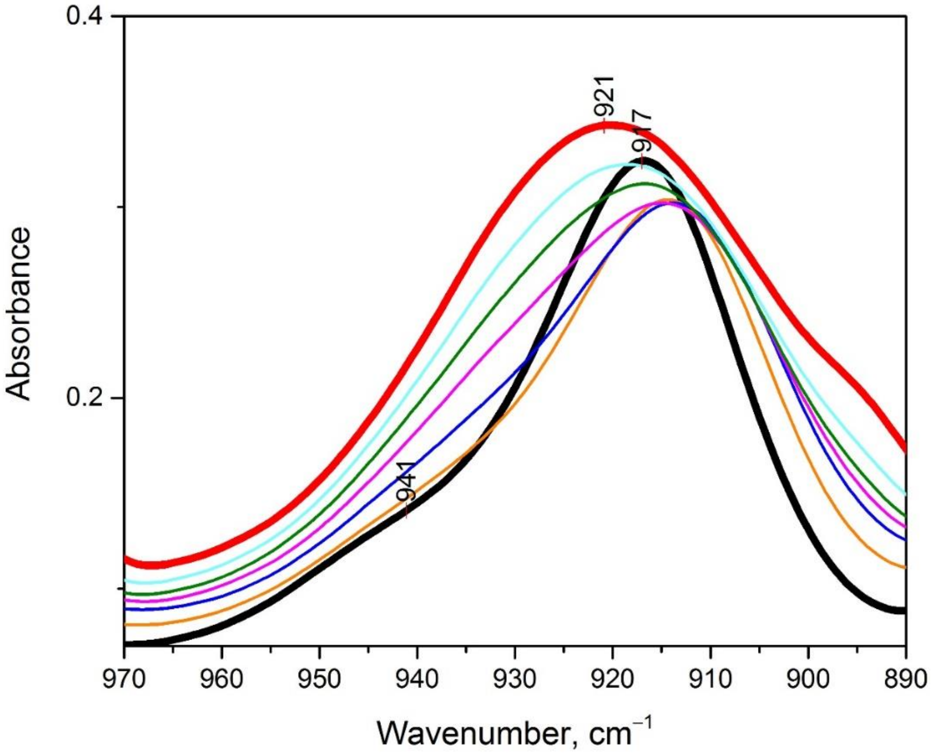


Figure 3.12: The FT-IR spectra in the region of 900 to 960 cm^{-1} at different temperature, 40 °C (black), 80 °C (yellow), 100 °C (blue), 110 °C (pink), 120 °C (green), 140 °C (turquoise) and 180 °C (red).

3.3.2 Thermodynamic of PNBA Form I and II in Solution

3.3.2.1 Solubility of PNBA Form I and II in Ethanol, 2-Propanol, Acetonitrile, Ethyl Acetate

The solubility curves of PNBA Forms I and II in ethanol (concentration versus temperature), at temperatures from 25 °C to 60 °C are shown in Figure 3.13. These solubility curves show that the two PNBA polymorphs are enantiotropically related, with a transition temperature approximately at 50 °C. At a temperature below 50 °C PNBA Form I is the more stable having the lower solubility. In this temperature range Form II will thus transform to Form I. This sits in reversed above 50 °C; Form I is metastable and Form II is stable.

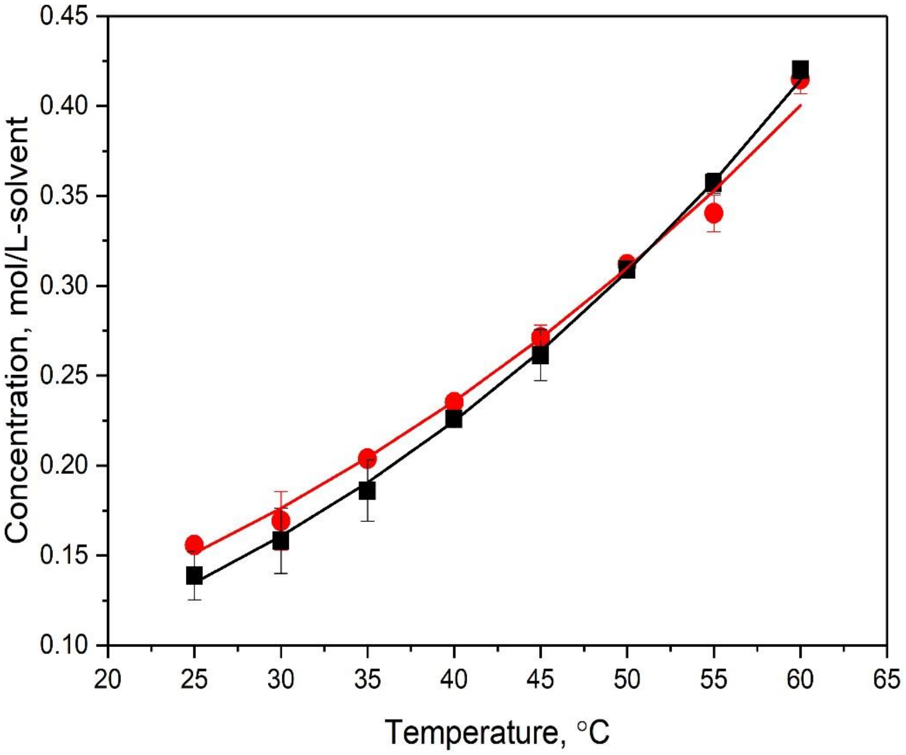


Figure 3.13: Solubility of PNBA Form I (black marker) and Form II (red marker) in ethanol, calculated based on 95 % confidence level. The solid lines are the predicted solubility obtained by fitting the experimental data to the van't Hoff equation.

During the gravimetric solubility measurement, the powder pattern of solids taken from each slurry was measured frequently. This is an important check, to ensure no polymorphic transformation occurred during the equilibration of the solution, which would have affected the result.

Table 3-2 and Table 3-3 summarises the solubility data for PNBA Form I and Form II in ethanol, 2-propanol, ethyl acetate and acetonitrile. The solubility is the highest in ethanol followed by 2-propanol, ethyl acetate and acetonitrile. This solubility data is plotted in Figure 3.14; the solubility trend is similar for ethanol and 2-propanol (increasing significantly with temperature) whereas in ethyl acetate and acetonitrile the changes are not as considerable. Ethanol and 2-propanol have an alcohol group, which has a hydrogen bond donor and acceptor capacity, while ethyl acetate and acetonitrile have only hydrogen bond acceptor capacity. As there is more possibility for solute-solvent interaction in the alcohol, it is not unexpected that the solubility is higher and rises faster with temperature.

Table 3-2: The average solubility data of PNBA Form I and Form II in ethanol and 2-propanol in g PNBA/g solvent and mole PNBA/L solvent with 95% confidence level. Three to five samples were made at each temperature.

Temperature, °C	Solubility, g PNBA /g solvent ***mole PNBA/L-solvent (M)***			
	Ethanol		2-Propanol	
	Form I	Form II	Form I	Form II
25	0.0296 ± 0.0029 ***0.1388 ± 0.0135***	0.0332 ± 0.0006 ***0.1558 ± 0.0028***	0.0208 ± 0.0001 ***0.0974 ± 0.0001***	0.0220 ± 0.0003 ***0.1027 ± 0.0013***
30	0.0339 ± 0.0173 ***0.1582 ± 0.0804***	0.0369 ± 0.0026 ***0.1693 ± 0.0162***		
35	0.0403 ± 0.0037 ***0.1861 ± 0.0171***	0.0440 ± 0.0005 ***0.2038 ± 0.0024***	0.0299 ± 0.0001 ***0.1380 ± 0.0005***	0.0319 ± 0.0002 ***0.1475 ± 0.0009***
40	0.0491 ± 0.0007 ***0.2260 ± 0.0032***	0.0511 ± 0.0003 ***0.2351 ± 0.0016***		
45	0.0572 ± 0.0031 ***0.2614 ± 0.0143***	0.0594 ± 0.0015 ***0.2712 ± 0.0069***	0.0433 ± 0.0004 ***0.1974 ± 0.0020***	0.0455 ± 0.0005 ***0.2075 ± 0.0023***
50	0.0681 ± 0.0007 ***0.3089 ± 0.0033***	0.0687 ± 0.0008 ***0.3119 ± 0.0037***		
55	0.0792 ± 0.0012 ***0.3570 ± 0.0056***	0.0755 ± 0.0023 ***0.3403 ± 0.0103***	0.0656 ± 0.0002 ***0.2953 ± 0.0009***	0.0647 ± 0.0002 ***0.2916 ± 0.0009***
60	0.0939 ± 0.0011 ***0.4201 ± 0.0050***	0.0927 ± 0.0018 ***0.4148 ± 0.0079***		

Table 3-3: The average solubility data of PNBA Form I and Form II in ethyl acetate and acetonitrile in g PNBA/g solvent and mole PNBA/L solvent with 95% confidence level. Three to five samples were made at each temperature.

Temperature, °C	Solubility, g PNBA /g solvent ***mole PNBA/L-solvent (M)***			
	Ethyl Acetate		Acetonitrile	
	Form I	Form II	Form I	Form II
25	0.0162 ± 0.0008 ***0.0870 ± 0.0044***	0.0179 ± 0.0012 ***0.0957 ± 0.0063***	0.0091 ± 0.0005 ***0.0425 ± 0.0025***	0.0104 ± 0.0003 ***0.0485 ± 0.0015***
35	0.0203 ± 0.0016 ***0.1072 ± 0.0083***	0.0218 ± 0.0003 ***0.1150 ± 0.0016***	0.0127 ± 0.0001 ***0.0584 ± 0.0004***	0.0143 ± 0.0014 ***0.0657 ± 0.0066***
45	0.0270 ± 0.0003 ***0.1403 ± 0.0015***	0.0272 ± 0.0008 ***0.1418 ± 0.0040***	0.0177 ± 0.0005 ***0.0803 ± 0.0021***	0.0191 ± 0.0008 ***0.0863 ± 0.0037***
55	0.0358 ± 0.0004 ***0.1837 ± 0.0018***	0.0345 ± 0.0016 ***0.1770 ± 0.0083***	0.0261 ± 0.0003 ***0.1164 ± 0.0012***	0.0251 ± 0.0011 ***0.1120 ± 0.0050***

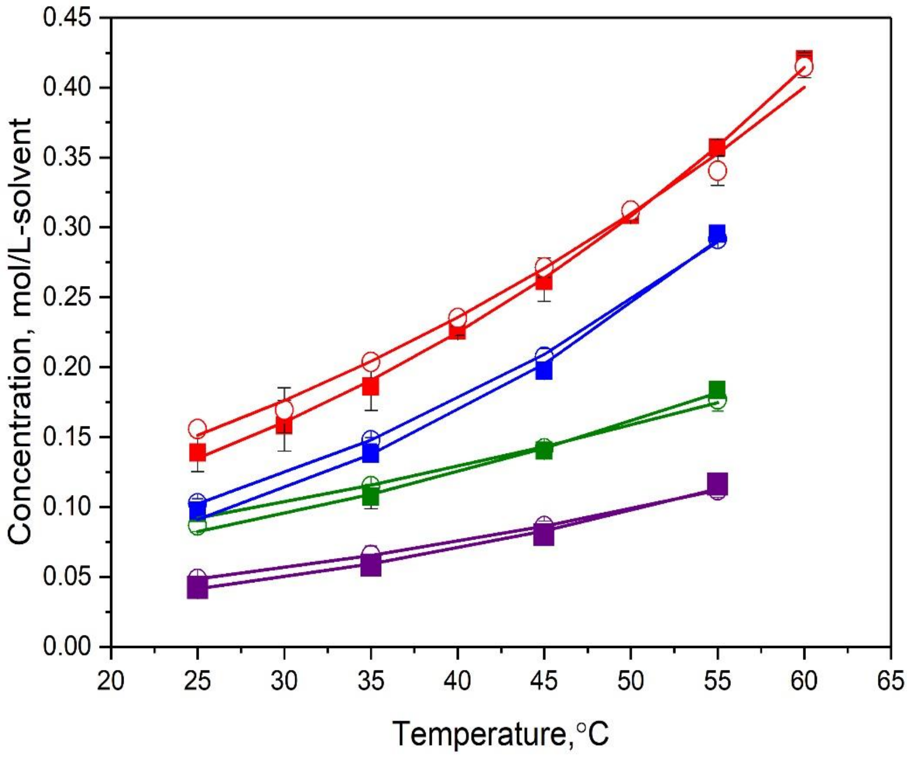
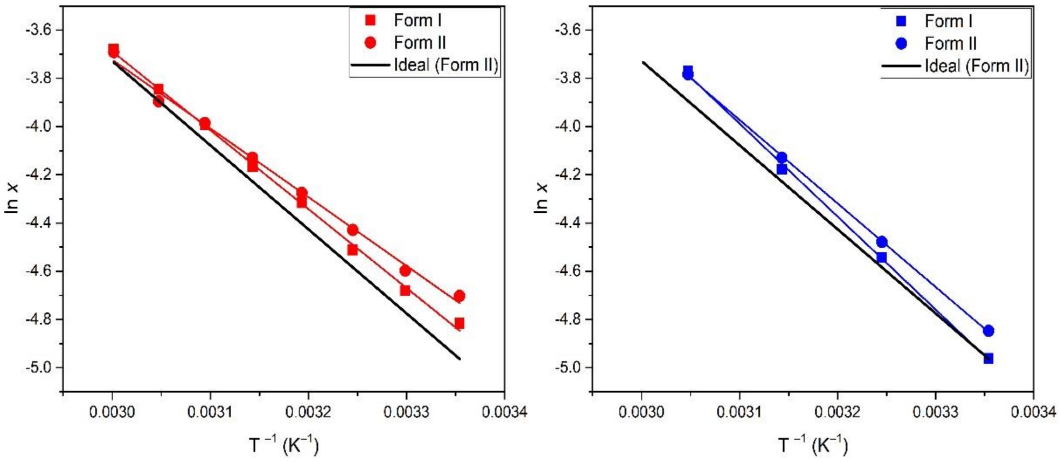


Figure 3.14: Solubility data from the gravimetric analysis (marker) is fitted to the predicted solubility based on van't Hoff equation (solid line) for ethanol (red), 2-propanol (blue), ethyl acetate (green) and acetonitrile (purple).

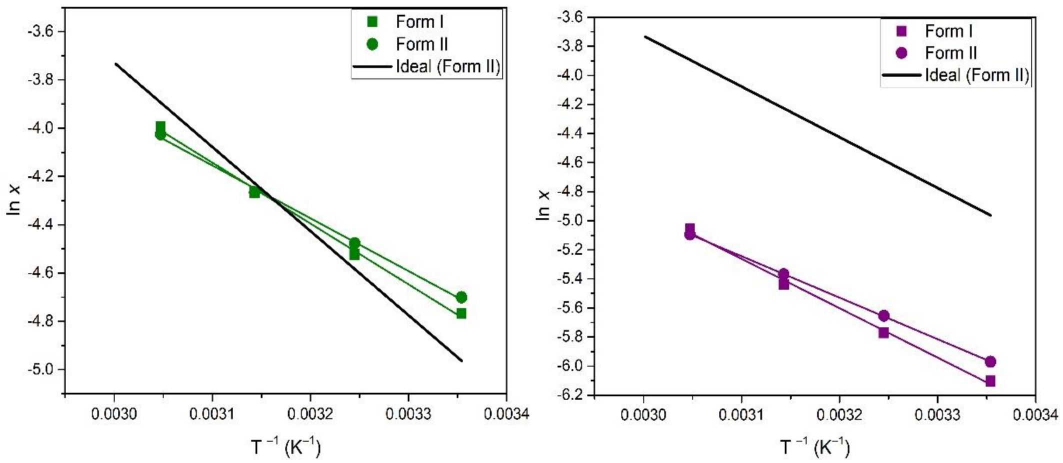
Figure 3.14 shows two sets of solubility data: the experimental data from the gravimetric analysis for PNBA Form I (rectangular marker) and PNBA Form II (circle marker), and the predicted solubility data which are based on the linear van't Hoff plot (presented as the solid lines). The solubility data show that PNBA Form I and Form II are most soluble in ethanol followed by in 2-propanol, ethyl acetate and acetonitrile. Interestingly, this trend aligns with the hydrogen bond donor (acceptor) capability. Ethanol, 2-propanol, ethyl acetate and acetonitrile have hydrogen bond donor (acceptor) properties of 0.83 (0.77), 0.78 (0.95), 0 (0.45) and 0.19 (0.31) respectively [85]. For the alcohols, the hydrogen bond donor property correlates with the solubility in ethanol being higher than in 2-propanol, while the hydrogen bond acceptor property correlates with the solubility of ethyl acetate being higher than in acetonitrile.

Figure 3.15 shows the plot for the ideal solubility (black line) which is based on the enthalpy and temperature of fusion for PNBA Form II (Equation 3.1 and Section 3.3.1.4). In comparison to the ethanol and 2-propanol solubility data, the ideal solubility is lower consistently over the range of temperature (Figure 3.14 (a) and (b)). In ethyl acetate, the ideal solubility is higher than the experimental data at temperatures above 317 K and below this temperature the ideal solubility is lower as shown in Figure 3.15 (c). The ideal solubility is significantly higher than the experimental data in the acetonitrile system, as shown in Figure 3.15 (d).



(a) PNBA in ethanol

(b) PNBA in 2-propanol



(c) PNBA in ethyl acetate

(d) PNBA in acetonitrile

Figure 3.15: Linear regression analysis for plots of PNBA Form I (rectangular marker) and Form II (circle marker) $\ln x$ against temperature, T^{-1}, in (a) ethanol (red), (b) 2-propanol (blue), (c) ethyl acetate (green), and (d) acetonitrile (purple). The black solid line represents the ideal solubility calculated using Equation 3.1 based on DSC data of PNBA Form II.

3.3.2.2 Slurry Transformation

The rate of slurry transformation (PNBA Form II to Form I) is consistently the fastest at 25 °C and slowest at 45 °C for each solvent. The driving force for a polymorphic transformation is the difference in the chemical potential which is correlated to the ratio of mole fraction of PNBA Form II to Form I as in Equation 3.8. Figure 3.17 shows that at 25 °C, the differences in the chemical potential are the biggest (~ 300 J mol^{-1}), thus the polymorphic transformation rate is the fastest. As these differences decreases as the transition temperature was reached, the transformation rate slowed down. For example, at 45 °C and 55 °C, the chemical potential difference is so small that it needs to be left overnight (24 hours) for the transformation to occur. This result further confirms the transition temperature is around 50 °C, as at 45 °C Form II transforms to Form I and at 55°C Form I transforms to Form II.

$$\Delta\mu = RT\ln\left(\frac{x_{II}}{x_I}\right)$$

Equation 3.8

At 25 °C, the rate of slurry transformation (PNBA Form II to Form I) is the fastest in acetonitrile (within 1.5 hours) followed by ethyl acetate (within 6 hours) and ethanol (within 48 hours) then 2-propanol (more than 4 days) at the same experimental conditions

(2 g of PNBA Form II stirred in 20 g of solvent) as summarised in Table 3-4. The powder pattern of solids taken from each slurry was measured frequently. The rate of transformation for PNBA Form II to Form I did not correlate with the solubility, in contrast to *p*-aminobenzoic acid (PABA) system, where the transformation rate was fastest in the higher concentration solution [10].

Table 3-4: Slurry transformation rate for PNBA Form II to Form I at 25 °C in acetonitrile, ethyl acetate, ethanol and 2-propanol.

Solvent	Duration, hours	Concentration at 25 °C, M
Acetonitrile	1.5	0.04
Ethyl acetate	6.0	0.09
Ethanol	48	0.14
2-propanol	>96	0.10

The picture that thus emerges is this: the rate of solvent-mediated transformation is controlled by the dissolution rate of the metastable form and the growth rate of the stable form [2]. As PNBA Form II dissolves in the solution, the solution is supersaturated with respect to the Form I, thus crystals of PNBA Form I nucleate and grow. The growth of PNBA Form I crystals causes more Form II to dissolve to reach the saturation level. This process continues until all the PNBA Form II had dissolved and recrystallization to PNBA Form I is complete. In 2-propanol, the rate of transformation increases significantly when seeded with PNBA Form I, suggesting that the transformation rate in 2-propanol is controlled by the dissolution process of Form II as the growth rate of PNBA Form I is fast. The transformation rate for PNBA Form II to Form I in alcohol is the slowest due to the slow dissolution process of Form II.

3.3.2.3 The Enthalpy and Entropy of Transition for PNBA Form II to Form I

The main aims in this section are to calculate the transition temperature, T_r, and to quantify the contribution of enthalpy and entropy to the polymorphic transformation. The enthalpy and entropy for the transition of PNBA Form II to Form I are obtained using two methods: (1) solubility ratio between the polymorphs, and (2) applying the van't Hoff equation to the solubility data in all the solvents.

Theoretically, at the transition temperature, T_r, the mole fraction solubility ratio of two polymorphs is equal to unity (one) in any solvent [2,10]. A plot of mole fraction solubility ratio for all the four solvents plotted against temperature (Figure 3.16) shows that these ratios increase with temperature. These ratios deviate furthest at 298 K (25 °C) and 308 K (35 °C) but approach unity at 318 K (45 °C) and 328 K (55 °C). The trends are similar in all the solvents. The black marker in Figure 3.16 represent the mean of mole fraction solubility ratio from all the solvents at each temperature. Based on the linear regression for each solvent, at $x_{eq,I} / x_{eq,II} = 1$, the mean transition temperature is 51 ± 6 °C.

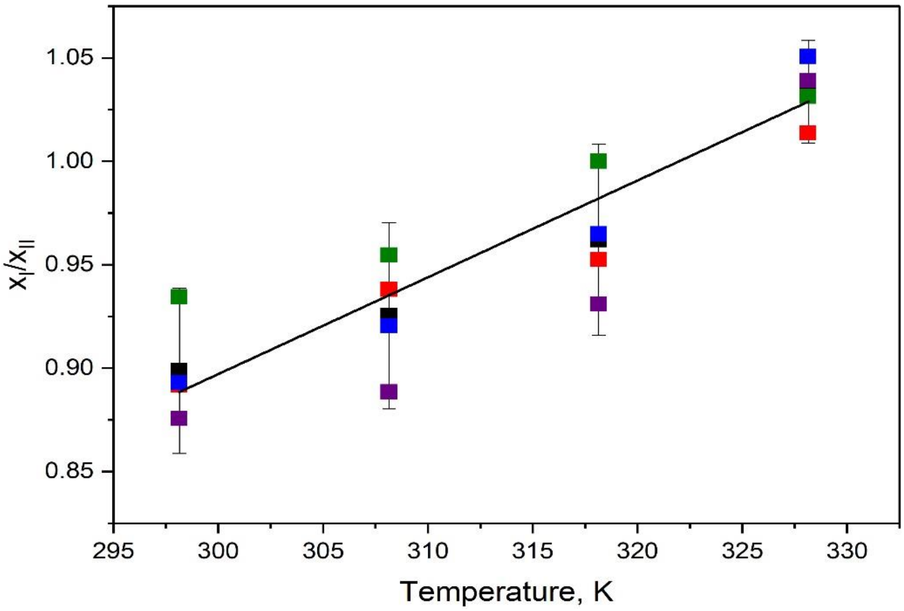


Figure 3.16: Plot of $x_{eq,I} / x_{eq,II}$ against temperatures in ethanol (blue marker), 2-propanol (red), ethyl acetate (green), acetonitrile (purple) and the mean from all the solvents (black marker) and its 95 % confidence level. The linear regression based on the mean value has a Pearson R-value of 98 %.

As discussed in Section 3.1, the slope of the linear regression analysis for the mean of plot $RT\ln(x_{II} / x_I)$ against temperature as shown in Figure 3.17 is the entropy of the transition, which is 11.6 ± 4.2 J mol^{-1} K^{-1}. At the polymorphic transition temperature of 51 ± 6 °C, the enthalpy of the transition is 3.8 ± 3.6 kJ mol^{-1} calculated using Equation 3.7.

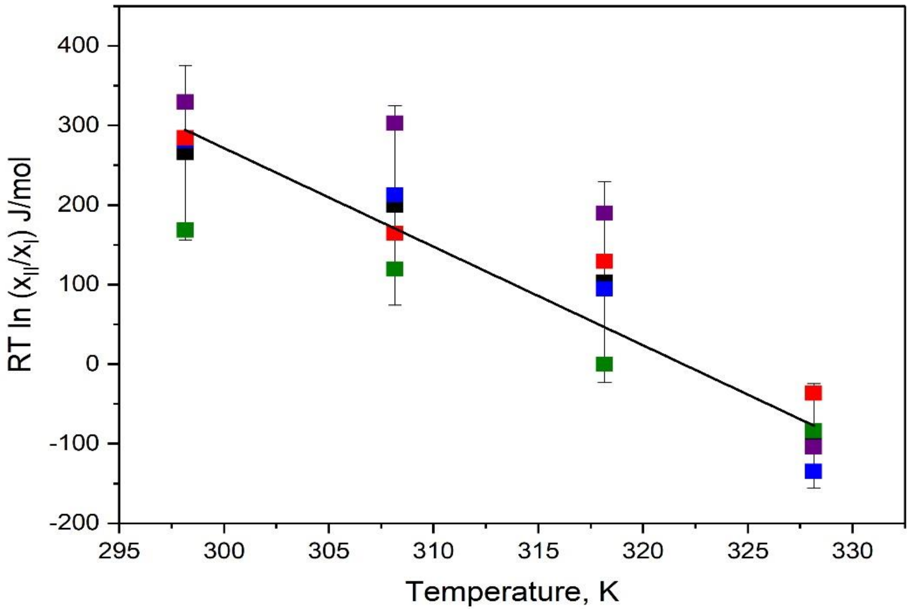


Figure 3.17: Plot of $RT\ln(x_{eq,I}/x_{eq,II})$ at different temperatures in ethanol (blue marker), 2-propanol (red), ethyl acetate (green) and acetonitrile (purple) and a linear regression over all the data. The Pearson R-value for the linear regression for the mean (black marker) is 94 %. The error bars are calculated based on 95 % confidence level of the mean data.

Alternatively, the entropy and enthalpy of the transition are calculated by obtaining first each enthalpy and entropy of dissolution for both polymorphs. The enthalpy of dissolution, ΔH_d , is calculated from the slope of the linear regression analysis for each plot $\ln x$ against T^{-1}as shown in Figure 3.15 for each polymorph in each solvent. The linear regression analysis (solid lines) for each solubility data (markers) has a Pearson R-value above 99 %. ΔG_d and ΔS_d for each of PNBA Form I and Form II are calculated at the polymorphic transition temperature, 324 K, in each solvent – ethanol, 2-propanol, ethyl acetate and acetonitrile using Equation 3.3 and Equation 3.4.

Table 3-5 summarises the ΔH_d , ΔG_d , and ΔS_d for PNBA Form I and Form II in bracket. The ΔH_d and ΔS_d for PNBA Form I are consistently higher than Form II in these solvents. The ΔH_d for PNBA Form I (Form II) is the highest in 2-propanol at 32 (29) kJ mol^{-1}, follow by in acetonitrile at 28 (24) kJ mol^{-1}, in ethanol at 27 (24) kJ mol^{-1} and the lowest in ethyl acetate at 21 (18) kJ mol^{-1}. The ΔS_d for PNBA Form I (Form II) is the highest in 2-propanol at 99 (89) J mol^{-1} K^{-1} followed by acetonitrile with 87 (73) J mol^{-1} K^{-1}, ethanol with 84(73) J mol^{-1} K^{-1} and finally ethyl acetate with 65 (56) J mol^{-1} K^{-1}.

Table 3-5: A summary of the enthalpy, ΔH_d, Gibbs free energy, ΔG_d, and entropy of dissolution, ΔS_d, calculated at the transition temperature, $T_{tr} = 324K$, for PNBA Form I and Form II (in brackets). The ΔG_d is the same for both polymorphs at the transition temperature.

Solvent	ΔH_d, kJ mol^{-1} PNBA Form I (Form II)	ΔG_d, kJ mol^{-1} PNBA Form I (Form II)	ΔS_d, J mol^{-1} K^{-1} PNBA Form I (Form II)
Ethanol	27.2 (23.7)	0.10 (0.10)	84.0 (73.1)
2-propanol	32.0 (28.8)	0.10 (0.10)	98.8 (88.9)
Ethyl acetate	21.0 (18.2)	0.11 (0.11)	64.8 (56.2)
Acetonitrile	28.2 (23.7)	0.13 (0.13)	87.0 (73.1)

The enthalpy of transition, ΔH_T, and the entropy of transition, ΔS_T, are calculated from the differences between ΔH_d and ΔS_d for the polymorphs in each solvent [3] (refer as well Section 1.2) and presented in Table 3-6. Based on Hess's Law, the ΔH_T and ΔS_T are independent of the solvent used [3]. Therefore, the differences observed in Table 3-6 between the solvents are subjected to the experimental error and the lack of consideration of heat capacity and activity coefficient factors as highlighted in Section 3.1. The main experimental error is possibly due to the accuracy of the solubility data for the metastable form. The equilibration duration of PNBA Form II in the solution is shorter and solubility sampling was taken approximately an hour before the transformation to Form I occured. The main advantage of this method is the enthalpy and entropy of dissolution for each polymorph can be calculated. The enthalpy of dissolution for PNBA Form II is comparable to the heat of fusion of PNBA Form II, which is 29 kJ mol^{-1} (Section 3.3.1.4).

Table 3-6: The enthalpy and entropy of transition, ΔH_T and ΔS_T, in ethanol, 2-propanol, ethyl acetate and acetonitrile calculated from Table 3.5. The average ΔH_T and ΔS_T based on van't Hoff equation and solubility ratio.

System	ΔH_T kJ mol^{-1}	ΔS_T J mol^{-1} K^{-1}
Ethanol	3.5	10.8
2-propanol	3.2	9.9
Ethyl acetate	2.8	8.6
Acetonitrile	4.5	13.9
Average (van't Hoff)	**3.5**	**10.8**
Average (Solubility ratio)	**3.8**	**11.6**

Table 3-6 also summarises the average ΔH_T and ΔS_T obtained from using the two methods; the average ΔH_T and ΔS_T calculated from the plot of $\ln x$ against T^{-1} are 3.5 kJ mol^{-1} and 10.8 J mol^{-1} K^{-1}respectively which is similar to the calculated value based on the plot of $RT\ln(x_{II}/x_I)$ against temperature, 3.8 kJ mol^{-1} and 11.6 J mol^{-1} K^{-1}. The average ΔH_T and ΔS_T for PABA are at 2.8 kJ mol^{-1} and 9.8 J mol^{-1} K^{-1} which is calculated based on the solubility ratio method [10] and are similar to PNBA. Both PABA and PNBA polymorphic transformations are driven by the high entropy factor at the transition temperature of 16 °C and 51 °C respectively.

The transition temperature, T_r, is equivalent to $\frac{\Delta H_T}{\Delta S_T}$. Therefore, a higher ΔS_T will result in a lower transition temperature at the same ΔH_T. Table 3-7 summarises the T_r, ΔH_T and ΔS_T for several systems where the polymorph transformation is driven by high entropy. There is no strong correlation between the ΔS_T to the T_r observed. For example, PABA and carbamazepine have similar ΔH_T of 2.8 kJ mol^{-1} and 2.9 kJ mol^{-1} respectively. The difference between their ΔS_T is only by 1.3 J mol^{-1} K^{-1} but the transition temperature difference is 57 °C. Therefore, even though the enthalpy of transition is not the driving force for the polymorph transformation, it still influences the transition temperature. For example, mefenamic acid, sulfathiazole and seratrodast have high entropies of transition but their transition temperatures are also higher than for the other systems at 89, 113 and

83 °C. Therefore, this observation did not support the reasoning that PABA's low transition temperature is driven by the high entropy of transition [10].

Table 3-7: A summary of the transition temperature, T_r, enthalpy of transition, ΔH_T, and entropy of transition, ΔS_T, for several systems.

Solute	T_r, °C	ΔH_T, kJ mol^{-1}	ΔS_T, J mol^{-1} K^{-1}
PNBA	51	3.8	11.6
PABA [10]	16	2.8	9.8
Mefenamic acid [86]	89	4.2	11.6*
Sulfathiazole [87]	113	6.9	17.8*
Seratrodast [88]	83	6.1	17.0*
Acetazolamide [88]	78	2.0	5.7*
Carbamazepine [88]	73	2.9	3.5*

*The entropy of transition is not given in the literature but calculated here based on the given transition temperature and enthalpy of transition.

3.4 Conclusion

PNBA Form I is recrystallized from PNBA Form II (purchased material) at room temperature in these solvents (ethanol, 2-propanol, ethyl acetate and acetonitrile). The slurry transformation rate is the fastest in acetonitrile and the slowest in 2-propanol. The key analytical characteristics that differentiate PNBA Form I and Form II have been shown for XRD, FT-IR, Raman and DSC. The solid state transformation of PNBA Form I to Form II can also be followed using hot stage microscopy while *in-situ* FTIR spectroscopy of PNBA Form I on a heated stage show the changes in the O−H···O interaction during the transformation. PNBA Form II sublimes at a temperature above 180 °C.

PNBA Form I is enantiotropically related to PNBA Form II in ethanol, 2-propanol, ethyl acetate and acetonitrile, with an average transition temperature at 51 °C. Below 51 °C, PNBA Form I is the stable form and Form II is the metastable form. Above 51 °C, PNBA Form I is the metastable form and Form II is the stable form. The solubility of PNBA polymorphs is the highest in ethanol followed by in 2-propanol, ethyl acetate and acetonitrile in the order of hydrogen bond donor and acceptor capability. The average ΔH_T and ΔS_T in all solvents for PNBA system are 3.5 kJ mol^{-1} and 10.9 J mol^{-1} K^{-1}.

4. INFLUENCE OF PROCESS CONDITIONS ON THE NUCLEATION OF *PARA*-NITROBENZOIC ACID (PNBA) FORM II

Abstract: Sensitivity to experimental conditions and procedures, as well as challenges in obtaining consistent experimental results from supersaturation-dependent nucleation rate measurements for PNBA Form II in ethyl acetate and 2-propanol were investigated. The influence of process conditions, such as the stirring speed, volume, preheating temperature, and concentration, on induction times and consequently the measured nucleation rates was studied. A controlled and detailed methodology was established to minimise the influence of preparation and nucleation process conditions, to obtain more reliable nucleation rate data, and consequently better quality thermodynamic and kinetic parameters describing the nucleation process.

4.1 Introduction

Crystal nucleation in solution is the initial process of crystal formation from a supersaturated solution[2]. The formation of crystals starts with an equilibrium of attachments and detachments of molecules forming a cluster[5]. Understanding the self-association of molecules in this early stage of phase separation is important because the self-assembly processes taking place determine molecular structure, e.g., in polymorphs, solvates, salts, and co-crystals [4]. Through this, physical characteristics such as crystal size distributions (CSD) are also affected [89]. As will be shown below, key parameters to understand the preferred molecular level assembly pathways are the kinetic factor, A, and the thermodynamic factor, B, as given by Classical Nucleation Theory (CNT[4]). The kinetic factor contains information on the molecular kinetics while the thermodynamic factor reflects the structure of the critical cluster [4]. Within the CNT framework these factors influence and determine crystal nucleation rates, J, so that experimental studies of nucleation rates can provide information on their dependence on crystallisation conditions, and thus indirectly give information on the nature of the self-assembly processes taking place.

The differences in the chemical potential of a molecule in the solution and solid state, $\Delta\mu$, is the driving force for the nucleation and growth of crystals. At $\Delta\mu > 0$, the solution is supersaturated and the nucleation process can occur. At $\Delta\mu = 0$ and $\Delta\mu < 0$ the solution is saturated and undersaturated, respectively, and nucleation is not possible. Equation 4.1

is the simplified equation in calculating $\Delta\mu$ where R is the Boltzmann constant, T is the absolute temperature and S is the supersaturation ratio [5]. Supersaturation ratio is commonly calculated as in Equation 4.2 where x and $x*$ are the actual and equilibrium concentration in mole fraction. As the supersaturation ratio increases, the probability of nucleation to occur increases as well.

$$\Delta\mu = RT\ln S \qquad \textbf{Equation 4.1}$$

$$S = \frac{x}{x*} \qquad \textbf{Equation 4.2}$$

The nucleation rate, J, can be determined from a statistical distribution of induction times, t_{ind} [27]. The induction time, t_{ind}, is the time difference between the achievement of supersaturation and the detection of crystals. The induction time consists of (i) t_J, the time a single super-nucleus is formed, (ii) t_g, the time for this single super-nucleus to growth and (iii) t_a, the time the crystals are detected once dispersed into fragments as a result of attrition, which is very fast and negligible [90]. As a result, $t_J = t_{ind} - t_g$. The crystals can be detected visually [91] or by detecting changes in the physical properties of the solution using laser beam [92] or turbidity reading [27]. Microfluidic [93] and levitator [94] systems are applied in measuring the induction times from droplets.

Measurement setups based on multiplexed reactors with agitation have become the preferred mode of study of nucleation kinetics over the last decade, due to their similarity to real-world crystallisation systems, including industrial crystallisers [26–28,30,31,95], of which they can be considered a scaled-down version. Typically, a large number of induction times, t_{ind}, is obtained at a constant supersaturation, temperature, and volume, which is then fitted to the cumulative probability distribution [26–28,30,31,95]. The probability, $P(t_{ind})$, to detect crystals at the time t_{ind} are given as in Equation 4.3 and 4.4 where J is the nucleation rate, V is the volume, $M^{+}(t_{ind})$ is the number of experiments in which crystals are detected at the time t_{ind}, and M is the total number of experiments [90]. The nucleation rate is calculated by fitting Equation 4.3 to the experimentally determine $P(t_{ind})$ curves (Equation 4.4).

$$P(t_{ind}) = 1 - \exp(-JV(t_{ind} - t_g))$$ **Equation 4.3**

$$P(t_{ind}) = \frac{M^+(t_{ind})}{M}$$ **Equation 4.4**

The kinetic, A, and thermodynamic, B, factors are then determined from the supersaturation dependent nucleation rate results as in Equation 4.5, which is based on the classical nucleation theory [26–28,31].

$$J = AS\exp\left(-\frac{B}{\ln^2 S}\right)$$ **Equation 4.5**

The methodologies used by different researchers for measuring induction times in solution are similar, but there are significant details. Table 4-1 summarises previously reported process conditions. Typically, a bulk, stock solution was prepared, stirred and heated above the saturation temperature. Brandel and ter Horst chose a small stock volume, 25 mL, and stated that this reduces concentration variation [26] while Mealey et al stirred a 250 – 500 mL stock solution overnight (12 hours) to ensure the solution reached equilibrium [30,53]. Next, the solution was transferred into multiple reactors using either pre-heated bottle top dispensers without filter [26,27,31] or pre-heated syringes with 0.20 µm filters [30,53]. Kulkarni et al found that the nucleation rate of a filtered solution (0.45 µm filters) reduced by 63% compared to the non-filtered in the 1 mL reactor [31].

Once transferred into the multiple reactors, the solutions were stirred at higher speeds, 700 and 900 rpm, in the smaller reactors of 1 – 2 mL [26–28,31] while in the bigger reactors of 20 mL, slower speeds at 200 and 400 rpm are chosen [30,53]. The reasoning for these differences in speed is not clear, however, agitation speed and agitator design do influence the nucleation rate [29,53]. Some researchers then held the solutions at high temperature to ensure all the solids are dissolved, ranging from 40 °C to 60 °C for 30 to 60 minutes [26–28,31]. Mealey et al held the solution at high temperatures overnight to ensure equilibration [30,53]. The high holding temperature and duration are important to minimise the solution history effect [95,96]. Finally, solutions are crash cooled at 5 °C min^{-1} and held for 5 to 8 hours [26–28,31] or immediate manual transport to another water bath takes place [30,53].

Table 4-1: A summary of the process conditions used in several publications in measuring the induction times. The rows that are not highlighted used Crystal 16 setup, which has smaller reactors compared to the setup highlighted in grey.

Reference	**Bulk solution, mL**	**Volume, mL**	**Stirring speed, rpm**	**High holding temperature**	**Number of experiments**
Jiang [27]	50	1	900	-	80
Kulkarni [31]	100	1	700	60 °C	144
Sullivan [28]	120	1/1.5/1.8	900	40 °C	80
Brandel [26]	25	1	700	60 °C	64
Mealey [30]	500	20	400	55 °C	80
Mealey [53]	250	20	200	5 °C + T_{sat}	10-50

The total number of experimental repeats performed differs from 10 to 144. A statistical study by Jiang [27] shows that with 80 experiments, there is an 80 % chance that the measured nucleation rate is within 20 % of the actual nucleation rate. This chance can be increased by increasing the total number of experiments [27]. In addition, Kulkarni excluded experiments that gave induction times smaller than 120 s because the crystals were probably formed during the crash cooling rather than at the constant low holding temperature [31]. Brandel excluded reactors that have mass losses larger than or equal to 0.05 % by measuring the total weight of the vials before and after the experiment [26].

Typical examples of previously reported data are shown in Figures 4.1 and 4.2. In both cases the markers represent the nucleation rate obtained by fitting Equation 4.3 to the experimental data, Equation 4.4. The A and B values are then calculated from the plot of $\ln(J/S)$ against $1/\ln^2 S$ using Equation 4.5. The nucleation rate at each supersaturation can then be re-calculated by re-applying Equation 4.5 using the obtained A and B values. In both figures, the solid lines represent these re-calculated nucleation rate.

Figure 4.1 shows a plot of nucleation rate as a function of supersaturation for PABA in acetonitrile, ethyl acetate and 2-propanol in a range of supersaturation between 1.1 and 1.4. A good fitting is observed between the obtained and re-calculated nucleation rates.

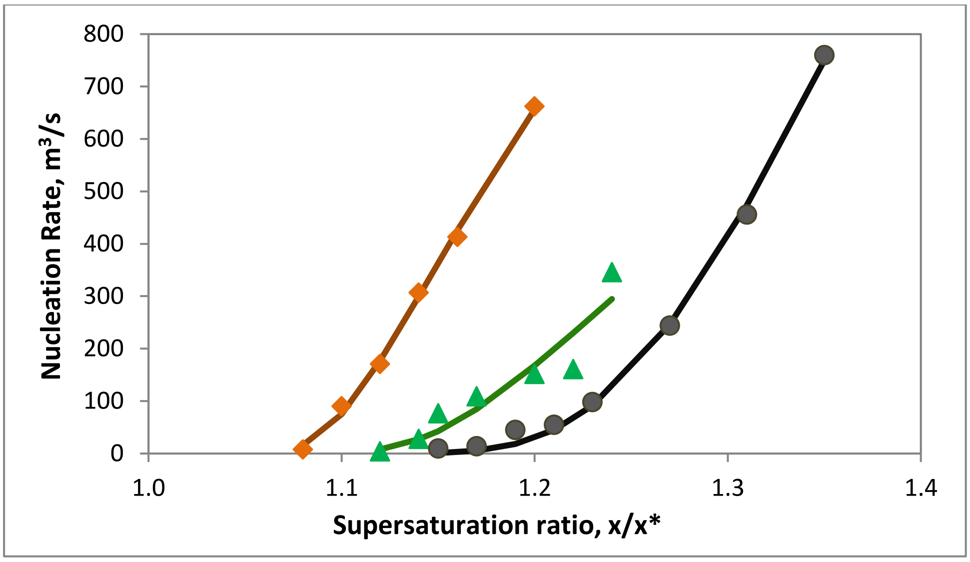


Figure 4.1: Plot of nucleation rate (J) as a function of supersaturation ratio for PABA in acetonitrile (orange), in ethyl acetate (green) and 2-propanol (black). The solid lines are the re-calculated nucleation rate based on Equation 4.5.

Figure 4.2 shows a similar plot for diprophylline (DPL) polymorphs, RI and RII, in dimethylformamide (DMF)[26]. In comparison to PABA, the supersaturation range, $S_{\max} - S_{\min}$, in the DPL system is bigger. The biggest range of supersaturation for the PABA system is 0.2 while in the DPL systems it is 1.5 as shown in Figure 4.2 [26,28]. This range of supersaturation is system dependent and obviously the bigger the range the more points can be measured and hence better estimates of A and B result.

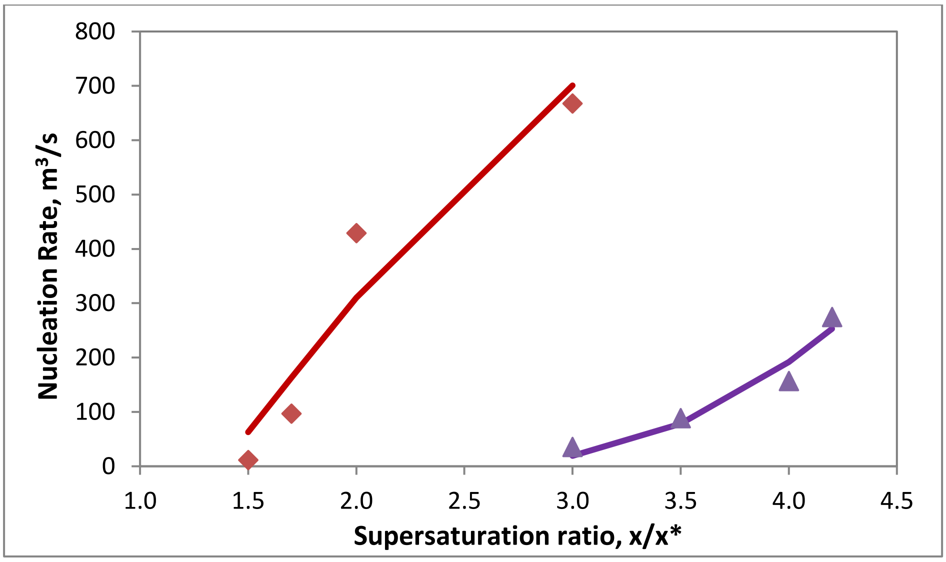


Figure 4.2: Plot of nucleation rate (J) as a function of supersaturation ratio for DPL RII in IPA (red) and DPL RI in DMF (purple). The solid lines are the re-calculated nucleation rate based on Equation 4.5.

Previously it has been found that the differences in the nucleation rate of *p*-aminobenzoic acid (PABA) in 2-propanol, acetonitrile and ethyl acetate and benzoic acid (BA) in toluene are influenced by the kinetic factor, which was then linked to the desolvation rate between dimers and monomers [28]. The nucleation rate of diprophylline (DPL) in two polymorphic systems showed polymorph RII nucleated faster in 2-propanol compared to polymorph RI in DMF driven by the thermodynamic factor [26].

The determination of supersaturation-dependent nucleation rates based on induction time variation is a promising technique to provide the insights into the nucleation process. To complement previous studies of PABA and BA it was therefore decided to examine the related compound, *p*-nitrobenzoic acid (PNBA), in order to study the substituent effects ($-NH_2$, $-H$ and $-NO_2$) on the nucleation rate, kinetic and thermodynamic factors. However, in pursuing this, sensitivity of the induction time variation (nucleation rate) to changes in process conditions for PNBA nucleation from 2-propanol and ethyl acetate was discovered. As discussed earlier, previous literatures reports different process conditions and approaches in measuring the induction times. Therefore, the main aim in this chapter is to define optimum process conditions and a standardised methodology for measuring the induction times for PNBA in 2-propanol and ethyl acetate.

4.2 Methodology

4.2.1 Initial Experimental Set-up

The Crystal 16 apparatus and the basis of induction time measurement are described in Section 2.8. An experimental method was initially set up based on Sullivan's procedure [28] . 120 mL of stock solution was prepared in a conical flask (sealed with parafilm) at each supersaturation by dissolving the appropriate amount of PNBA Form I in the solvent on a heated plate. The heating time and temperature were not controlled - heating proceeded until all solid was dissolved. The stock solution was then dispensed into the Crystal 16 vials using pre-heated plastic pipettes, 1.5 mL for PNBA in ethyl acetate and 1.8 mL for PNBA in 2-propanol. The process procedure for the nucleation experiments starts with heating up all the vials to 40 °C for 30 minutes with stirring at 900 rpm. The solution is then crash cooled at 5 °C min^{-1} to 25 °C and held for 8 hours. This cycle is repeated five times giving a total number of 80 experiments for each supersaturation. The polymorphic form was checked, by powder X-ray diffraction, for several vials at each supersaturation immediately after solids were formed. In the next section, the effect of changing some of the experimental variables was explored.

4.2.2 The Effect of Stirring Speed, 900 rpm versus 1400 rpm

The aim of this experimental section is to study the effect of the stirring speed by comparing results obtained at 900 rpm and 1400 rpm. The nucleation experiment for PNBA in ethyl acetate was prepared as in Section 4.2.1 at supersaturation ratio 1.16, 1.20, 1.24, 1.28, 1.32, and 1.36. These sets of experiments were then repeated at 1400 rpm. In order to confirm the stirring effect, additional two repeats were made for PNBA in ethyl acetate at supersaturation ratio 1.32 at 900 rpm.

4.2.3 Revised Experimental Procedure

The aim of this experimental section is to improve the poor reproducibility trend observed based on the repeats results for PNBA in ethyl acetate at supersaturation ratio 1.32 (Section 4.2.2). Several factors were identified and given as below:-

Table 4-2: The potential issues from the initial experimental set-up and corrective action as in the revised experimental procedure.

Initial Experimental Set-up	Potential Issues	Revised Experimental Set-up
120 mL of stock solution was prepared by measuring the volume of solvent.	The lower accuracy of measuring the volume using measuring cylinder.	80 g of solvent was weighed.
The stock solution was heated on a non-controlled heater plate.	Overheating is possible which caused solvent evaporation.	The stock solution was heated in a jacketed vessel at a controlled temperature, 40 °C.
The stock solution was immediately dispensed into the vials after all the solids were dissolved.	The solution did not reach equilibrium or was not homogenous.	The stock solution was stirred for an hour at 40 °C.
The stock solution was dispensed after the heated plate was switched off.	During dispensing, the stock solution is cooling down which may vary the	The stock solution was dispensed at a constant temperature of 40 °C.

	actual volume of solution dispensed into each vials.	
The stock solution was dispensed using pre-heated plastic pipettes	The variation of the plastic pipette volume with temperature may vary the volume dispensed.	The stock solution was dispensed using plastic pipettes at room temperature.
The process setting in the C16 programme is to hold the solution at 40 °C for 30 minutes.	The solution may not reach equilibrium or be homogenous.	The process setting in the C16 programme is to hold the solution at 40 °C for an hour.

In order to justify the improvement of the revised experimental setup, another two repeats were made for PNBA in ethyl acetate at supersaturation ratio 1.32. 80 g of ethyl acetate was weighed and an appropriate amount of PNBA Form I was dissolved in a sealed jacketed vessel at a controlled temperature of 40 °C for an hour. 1.5 mL of stock solution was dispensed into each vials using plastic pipettes while the solution temperature remained at 40 °C. The process setting in the Crystal 16 is changed from holding at 40 °C for 30 minutes to an hour and remained at 900 rpm. This revised version is then tested for PNBA in 2-propanol at supersaturation ratio 1.59, 1.63, 1.67, and 1.71. For PNBA in 2-propanol, 1.8 mL of solution was chosen to avoid the formation of 'crowning'.

In addition, a simple test was done by comparing the weight of 1.6 mL (1.6 g) of water using a pre-heating plastic pipette (pre-heated in the oven at 60 °C) and at room temperature to justify the variation of volume.

4.2.4 The Effect of Thermal History

The aim of this experimental section is to study the effect of thermal history which is influenced by the high holding temperature. In general, the high holding temperature is chosen to ensure all the solids are dissolved. However, this temperature as well influences the subsequent nucleation rate for certain systems such as *m*-hydroxybenzoic acid [95]. In order to test the sensitivity of PNBA systems to the thermal history effect, the nucleation rate of PNBA was tested at two high holding temperatures, 40 °C and 50 °C.

A stock solution for PNBA in 2-propanol at supersaturation ratio 1.71 was prepared as in Section 4.2.3 with a small change; the stock solution was heated and dispensed into the vials at 50 °C. The process setting in the Crystal 16 was changed from holding at 40 °C to 50 °C for an hour.

4.2.5 Reviewing the Choice of Volume

The aim of this experimental section is to find the optimum volume for the nucleation experiments which takes into consideration the 'crowning' issue and temperature distribution in the solution. The choice of the volume is system-dependent due to the formation of 'crowning'. The maximum volume of 1.8 mL was chosen for PNBA in 2-propanol to avoid 'crowning', which is similar to the behaviour observed for PABA [28]. However, the temperature distribution in the solution is a concern. This is because the upper solution level for 1.8 mL is above the heating/cooling coil and fills up to the neck of the vial as shown in Figure 4.3 (left).

A stock solution for PNBA in 2-propanol at supersaturation ratio 1.71 was prepared as in Section 4.2.4 with another small change; 1.5 mL of solution was dispensed into each vials. 'Crowning' was checked visually and by shaking the vials to observe the falls of any solids into the solution during the duration of the low holding temperature at 25 °C. If the nucleation occurs immediately after shaking the vials, 'crowning' was concluded to form around the neck of the vials. These procedures are then repeated by increasing the volume to 1.6 mL if 'crowning' was concluded to occur at 1.5 mL.

In order to check the temperature distribution, water was used as a medium. Water was dispensed into four vials and the temperature of the water at the upper level of the heating/cooling coil was checked using an external temperature probe and compared to the temperature at the bottom level.

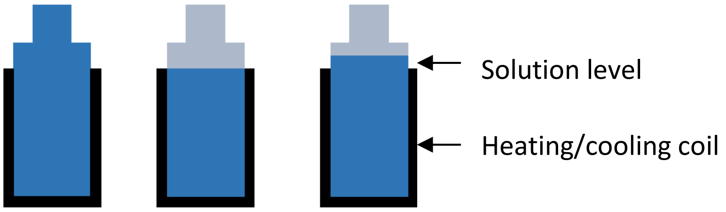


Figure 4.3: A sketch of the solution level and the heating/cooling coil placement at 1.8 mL (left), 1.5 mL (middle) and 1.6 mL (right).

4.2.6 The Effect of Sealing the Vials with Parafilm

The aim of this experimental section is to understand the insignificant differences observed for the nucleation rates of PNBA in 2-propanol at supersaturation ratio of 1.75, 1.71, 1.67 and 1.63 using 1.6 mL of solution (method of Section 4.2.5). The main factor to investigate is the consistency of the concentration in each vial throughout the five cycles that takes nearly 48 hours to complete. A stock solution for PNBA in 2-propanol at supersaturation ratio 1.63 and 1.59 was prepared as in Section 4.2.4 and 1.6 mL of solution was dispensed into each vials (Section 4.2.5). The vials were then weighed before and after completing the experiment (48 hours later). The amount of solvent evaporated is then calculated as the difference between these weights. This experiment is then repeated with an additional step; the top of the vials were sealed with parafilm to avoid solvent evaporation.

4.2.7 Actual Concentration Check in Each Vials

The variation of concentration between all 16 vials can be checked by utilising gravimetric analysis. The mass of the empty vial and stirrer (m_{ev}) and the mass of vial, stirrer and solution ($m_{ev+solution}$) were weighed during the sample preparation while the mass of vial, stirrer and dry solids ($m_{ev+solids}$) were taken after completing the experiment by placing the vials in the oven. The concentration is calculated using below equation.

$$c = \frac{m_{ev+solids} - m_{ev}}{m_{ev+solutions} - m_{ev+solids}}$$ **Equation 4.6**

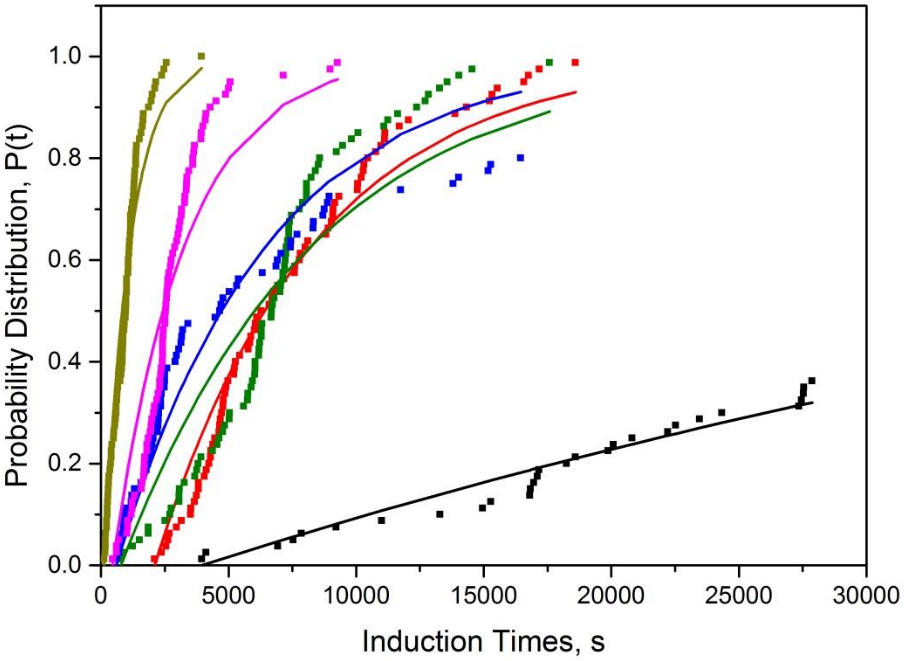
4.3 Results and Discussions

Throughout all these experiment, PNBA Form II nucleates first based on powder x-ray diffraction pattern results. This was checked by taking samples (during the last cycle) immediately when the crystals were detected.

4.3.1 The effect of Stirring Speed, 900 rpm versus 1400 rpm

There are two sets of results related to the stirring speed that will be discussed here; (i) PNBA in ethyl acetate at two different stirring speed, 900 rpm and 1400 rpm, at supersaturation ratio 1.16, 1.20, 1.24, 1.28, 1.32, and 1.36 and (ii) repeats for PNBA in ethyl acetate at 900 rpm for supersaturation ratio 1.32.

Figure 4.4 shows the probability distribution curves of the experimental data (markers) that is fitted with the model (solid lines) for PNBA in ethyl acetate runs at 900 rpm and 1400 rpm. The main observation is the poor fitting between the experimental and model at 900 rpm in comparison to the stirring speed at 1400 rpm.



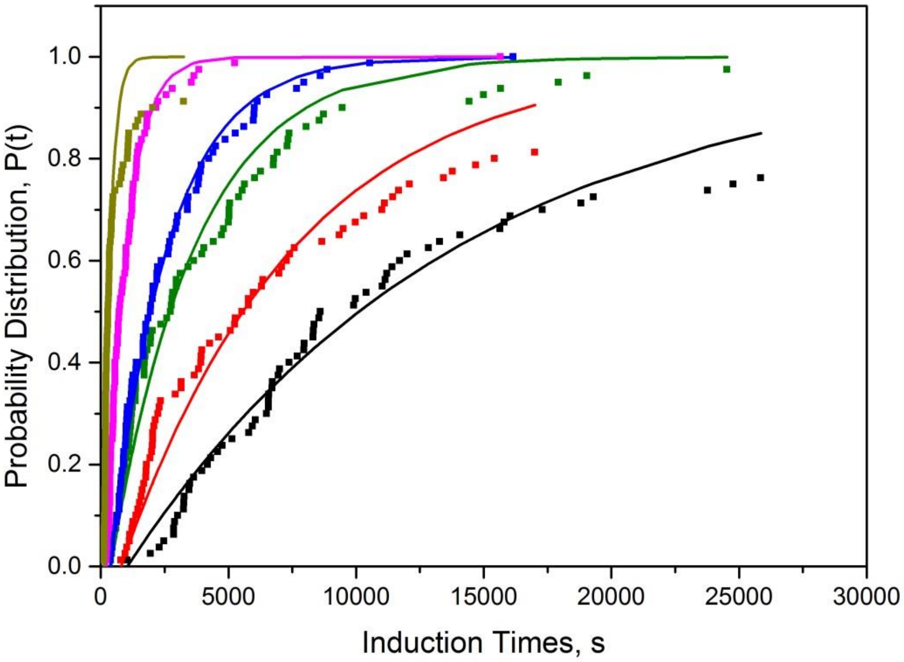


Figure 4.4: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.16 (black), 1.20 (red), 1.24 (green), 1.28 (blue), 1.32 (magenta) and 1.36 (yellow) for PNBA Form I in ethyl acetate at 25°C in 1.5 ml solutions at the stirring speed of 900 rpm (top) and 1400 rpm (bottom). Solid lines are the fits to Equation 4.3.

The summary of the supersaturation dependent nucleation rate and growth time is given in Table 4-3. The nucleation rate at 1400 rpm varies from 81 $m^{-3}s^{-1}$ at *S* 1.16 and consistently increased to 2613 $m^{-3}s^{-1}$ at *S* 1.36. At 900 rpm, the nucleation rate is slower, ranging from 11 $m^{-3}s^{-1}$ to 659 $m^{-3}s^{-1}$ with an inconsistent increment in between the *S* 1.20 (107 $m^{-3}s^{-1}$), 1.24 (88 $m^{-3}s^{-1}$) and 1.28 (112 $m^{-3}s^{-1}$). The growth time, t_g , is fixed as the fastest induction time is slower at 900 rpm than 1400 rpm ranging from 127 s to 3956 s and 122 s to 1508 s respectively.

Table 4-3: The nucleation rate, J , and growth time, t_g , obtained based on the curve fitting between experimental and model.

S, *x/x**	900 rpm		1400 rpm	
	J, $m^{-3}s^{-1}$	t_g, s	J, $m^{-3}s^{-1}$	t_g, s
1.16	11	3956	81	1508
1.20	107	2097	97	798
1.24	88	783	200	389
1.28	112	603	294	373
1.32	235	457	856	227
1.36	659	127	2613	122

The data was then evaluated by checking the randomness of the induction time at each block, vial and cycles as shown in Figure 4.5. Figure 4.5 is an example of the variation check for PNBA in ethyl acetate at supersaturation 1.24 at 900 rpm. Based on the variation, there was no strong justification to exclude any points from this plot. For example, even though Block B Vial 3 showed induction times consistently decreasing from cycle 2 to 5, a systematic cause for such behaviour could not be identified and the data points were retained. The most important feature of this plot is that all blocks showed consistent variation, with no particular vials or cycles giving constant induction times.

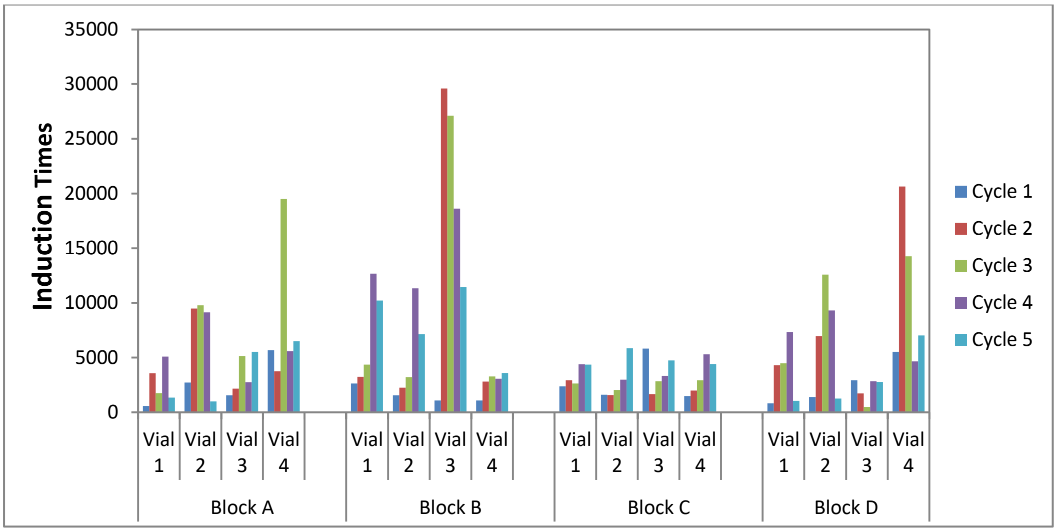


Figure 4.5: A plot of induction times for each block, vial and cycle for PNBA in ethyl acetate at supersaturation ratio 1.28 and stirring speed 900 rpm.

The kinetic and thermodynamic factor, A and B, were obtained from the slope and *y*-intercept from the straight line plot of $\ln(J/S)$ against $1/\ln^2 S$ as shown in Figure 4.6. The value of B differs by only 0.0094. Therefore, the nucleation rate being higher at 1400 rpm is driven by the kinetic factor, 1663 m^{-3} s^{-1} at 1400 rpm compared to 712 m^{-3} s^{-1} at 900 rpm. Even though the stirring effect on the kinetic factor A and nucleation rate were interesting, it was not pursued further as the concern was more with the poor fitting between the experimental and model probability curve at 900 rpm. Moreover, the stirring speed at 1400 rpm was too high causing several stirrers to 'jump' instead of creating homogenous stirred solution.

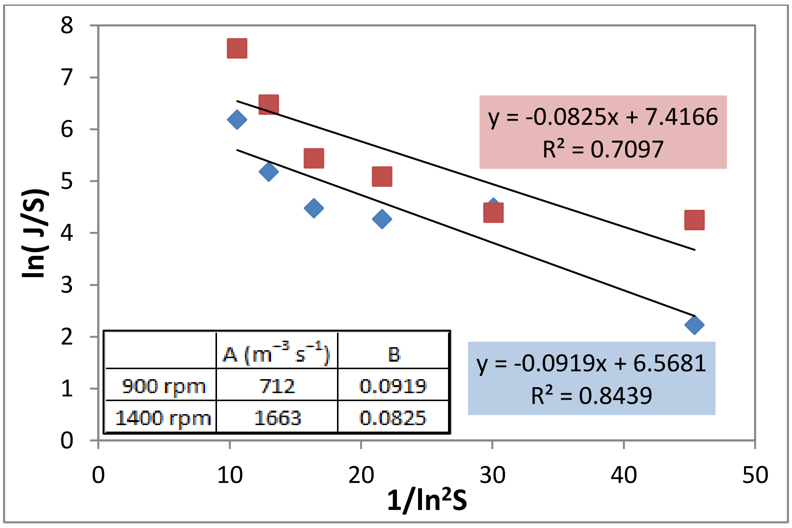


Figure 4.6: Plot of $\ln(J/S)$ as a function of $1/\ln^2 S$ for PNBA in ethyl acetate at 1400 rpm (red) and 900 rpm (blue). The inset table summarised the obtained values for A and B.

Figure 4.7 shows the inconsistency of the probability distribution curves for PNBA in ethyl acetate at supersaturation ratio 1.32 for three sets of data at 900 rpm. The nucleation rate for the first set of data is 235 $m^{-3}s^{-1}$ (bright green curve) followed by 168 $m^{-3}s^{-1}$ (dark yellow curve) and 90 $m^{-3}s^{-1}$ (dark green curve). This had raised a new concern which is discussed in the following section. Secondly, the fittings between the experimental and models were good for the additional repeats (dark yellow and dark green) which conclude that the stirring speed at 900 rpm is a good setting to use for PNBA systems as well.

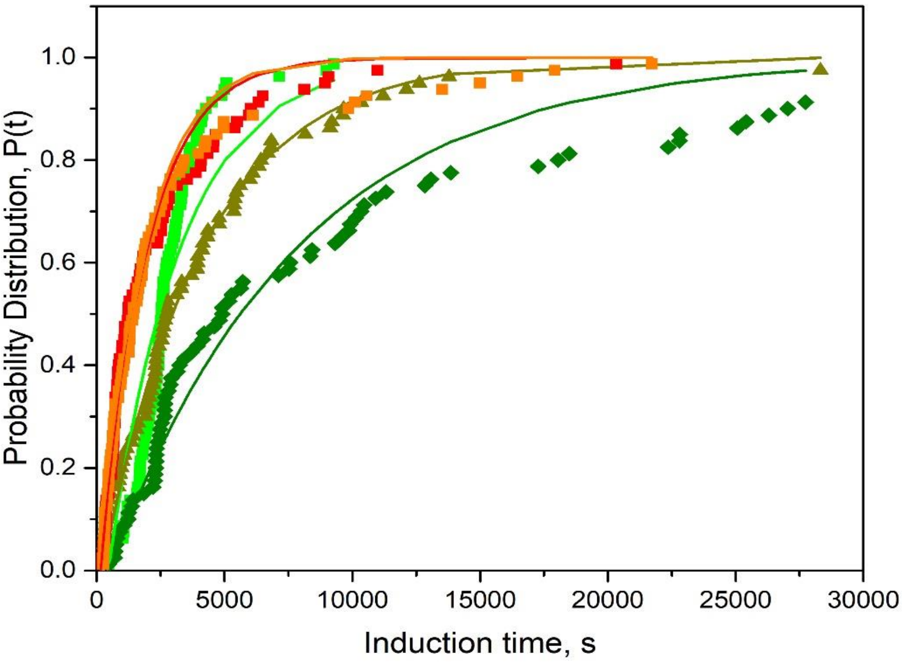


Figure 4.7: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.32 at 900 rpm for PNBA in ethyl acetate using (i) initial experimental set-up - first set of data (green), second repeat (dark yellow) and third repeat (dark green) and (ii) revised experimental set-up – first set of data (red) and second repeat (orange). Solid lines are the fits to Equation 4.3.

4.3.2 Revised Experimental Procedure

Figure 4.7 (red and orange curves) shows two improvements for the probability distribution curves for PNBA in ethyl acetate at 900 rpm after applying the revised method; (i) the good fittings between the experimental data and the model and (ii) two good sets of reproducible data in comparison to the sets of results using 'initial experimental set-up' (bright green, dark yellow and dark green curves). The nucleation rate obtained using this revised experimental set-up is 305 and 330 $m^{-3}s^{-1}$ respectively. The main advantage of the revised methodology is the samples were prepared in a temperature controlled system which reduced the variation of concentration and volume among all the sixteen vials. For example, using water as a medium, 1.6 mL of water dispensed using a non-heated plastic pipettes give an average weight of 1.60 ± 0.002 g while pre-heated plastic pipettes are higher at 1.63 ± 0.02 g.

This method works well for PNBA in 2-propanol at supersaturation ratio 1.59, 1.63, 1.67, and 1.71 as shown in Figure 4.8 where the fittings between the experimental and model are good. The nucleation rate for supersaturation at 1.59, 1.63, 1.67 and 1.71 are 11, 17, 497 and 713 $m^{-3} s^{-1}$ respectively.

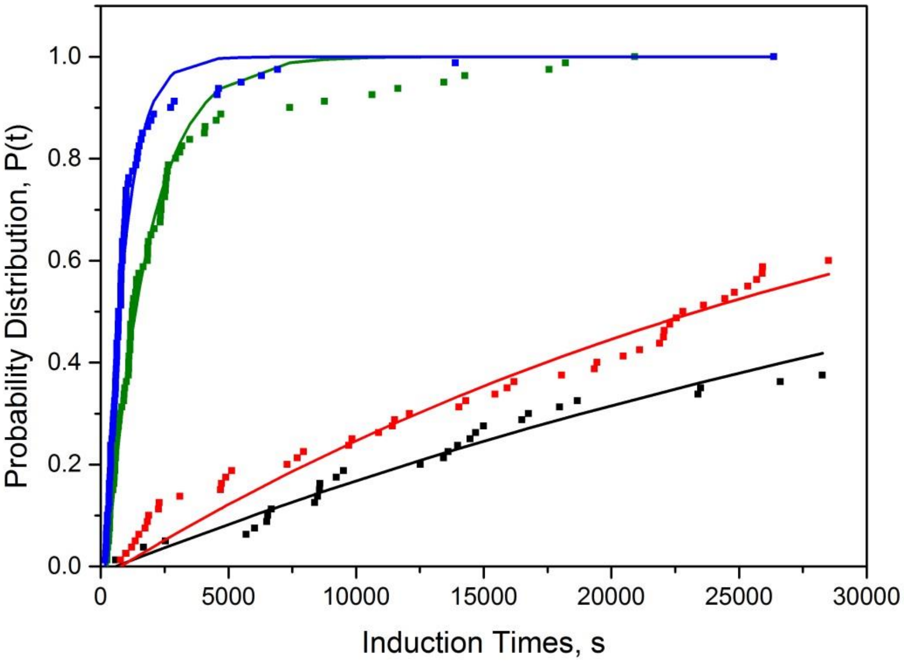
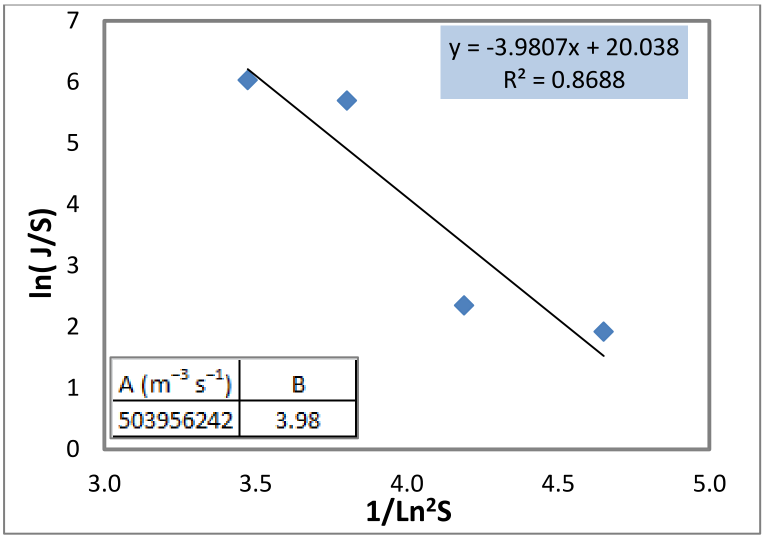


Figure 4.8: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.59 (black), 1.63 (red), 1.67 (green) and 1.71 (blue) for PNBA Form I in 2-propanol at 25dC, 900 rpm and in 1.8 mL solution. Solid lines are the fits to Equation 4.3.

However, the kinetic factor, A, obtained from the y-intercept of the plot as shown in Figure 4.9 is a large and unrealistic value. The run at supersaturation ratio of 1.65 is similar to the run at 1.67, thus the data was not shown here.



A (m^{-3} s^{-1})	B
503956242	3.98

Figure 4.9: Plot of $\ln(J/S)$ as a function of $1/\ln^2 S$ for PNBA in 2-propanol obtained after improvised the preparation of bulk sample. The stirring speed is at 900 rpm. The inset table summarised the obtained values for A and B.

4.3.3 The Effect of Thermal History

Thermal history of the solution is another aspect that influences nucleation kinetics [95,96]. The influence of this factor is minimised or controlled by the preheating temperature and duration, which is before the solution was cooled to achieve supersaturation. An increase in preheating temperature resulted in slower crystallisation kinetics [96]. Figure 4.10 shows that at a higher preheating temperature of 50 °C, the nucleation rate is slower at 285 $m^{-3}s^{-1}$ compared to 713 $m^{-3}s^{-1}$ at 40 °C for PNBA in 2-propanol at supersaturation 1.71. Thus, the pre-heating temperature at 50 °C for an hour was chosen as the optimum condition to minimise the effect of thermal history. This is roughly 10 °C higher than the saturated temperature.

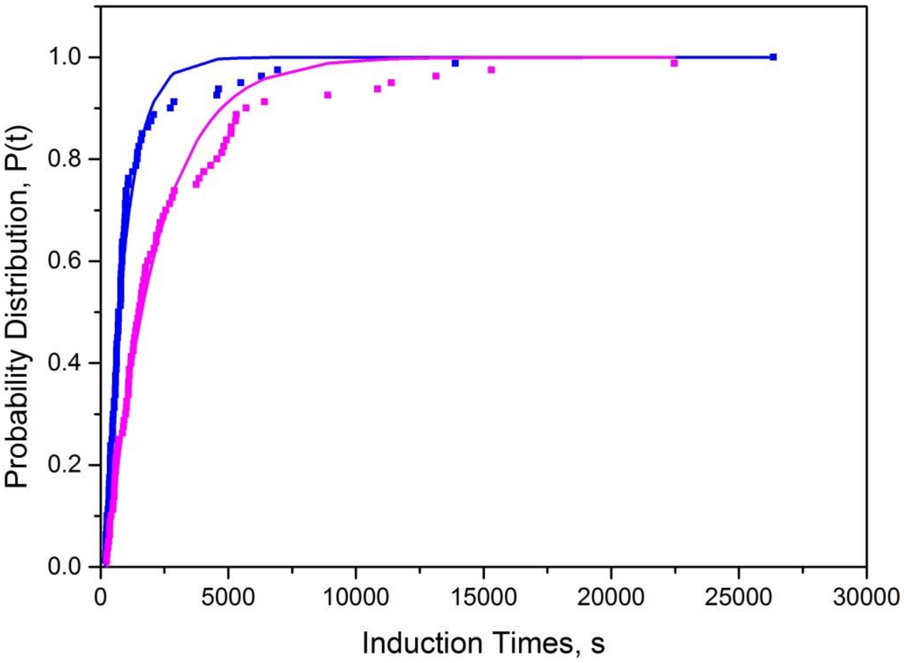


Figure 4.10: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.71 for PNBA Form I in 2-propanol at 25°C, 900 rpm and in 1.8 mL solutions at preheating temperature 40°C (blue) and 50°C (magenta). Solid lines are the fits to Equation 4.3.

4.3.4 Reviewing the Choice of Volume

As 'crowning' still formed at 1.5 mL, the volume of 1.6 mL is used. This is acceptable as the temperature of the solution above and below the heating/cooling coil showed comparable results (using an external temperature probe). However, reducing the volume showed an insignificant difference between the nucleation rates at different supersaturations. At the same supersaturation, four repeats were made (four probability curves with each maximum 80 induction times) in 1.6 mL solutions, the pre-heating temperature at 50°C (an hour) and stirring speed at 900 rpm for PNBA in 2-propanol at supersaturation 1.63, 1.67, 1.71 and 1.75. The results as in Figure 4.11 showed that there is no significant difference in the induction time variation between all the repeats at different supersaturation ratio. The average nucleation rate at supersaturation 1.63, 1.67, 1.71 and 1.75 is 125, 117, 141 and 110 $m^{-3}s^{-1}$.

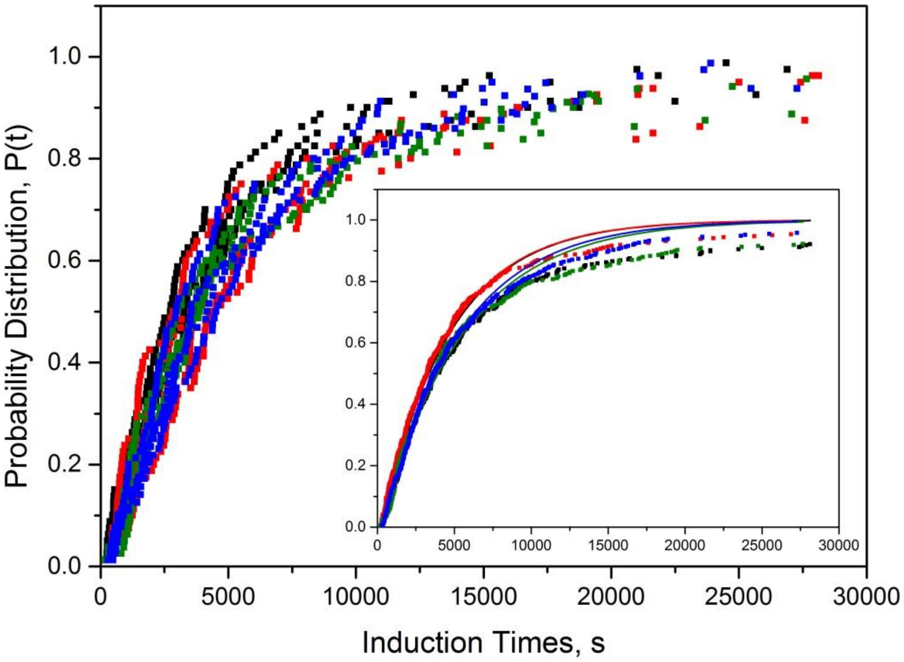


Figure 4.11: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.76 (black), 1.72 (red), 1.68 (green) and 1.64 (blue) for PNBA Form I in 2-propanol at 25°C, 900 rpm, 1.6 mL solutions at preheating temperature 50°C (magenta). Solid lines are the fits to Equation 4.3.. Each supersaturation is repeated four times (320 induction times). The inset figure is the average probability curves at each supersaturation.

4.3.5 The Effect of Sealing the Vials with Parafilm

Concentration is the most important factor in determining the supersaturation dependent nucleation rate. The possibility of solvent evaporation during the experiment (48 hours) was verified by measuring the mass of the solution before and after the experiment. Without sealing the vials, the total number of vials with solvent evaporation exceeding 0.05 % were 12 and 7 at supersaturations 1.63 and 1.59 respectively. This number reduced significantly to 3 and 2 after sealing the vials with a layer of parafilm.

In a sealed (unsealed) conditions, the nucleation rate at supersaturation ratio 1.60, 1.64 and 1.68 is 130 (93), 296 (154) and 364 (115) $m^{-3}s^{-1}$. Theoretically, the evaporation of solvent should result in a higher nucleation rate, but this is not observed for the unsealed vials. This phenomenon is not yet understood. Nevertheless, any mass loss of more than 0.05 % was excluded, referring to the guidelines given by Brandel and ter Horst [26].

4.3.6 The Best Set of Result for PNBA in 2-Propanol

The nucleation rates of PNBA in 2-propanol were then measured using the optimised procedure: volume 1.6 mL, preheating temperature ats 50°C (an hour), stirring speed at 900 rpm and sealing all vials (excluding mass loss more than 0.05 %) for supersaturation ratio 1.43, 1.47, 1.51, 1.55, 1.59, 1.63 and 1.67 are 18, 42, 79, 101, 180, 297 and 364 $m^{-3}s^{-1}$ respectively. A summary of the nucleation rate and growth time at each supersaturation is given in Table 4-4; the nucleation rate increased and growth time reduced with increasing supersaturation ratio. The fitting of the experimental probability curve to the model is good as shown in Figure 4.12. Finally, a satisfactory experimental procedure has been developed.

Table 4-4: A summary of the nucleation rate and growth time for PNBA in 2-propanol with sealing all the vials.

S	J, m^3/s	t_g, s
1.67	364	145
1.63	297	195
1.59	180	290
1.55	101	525
1.51	79	975
1.47	42	1315
1.43	18	1685

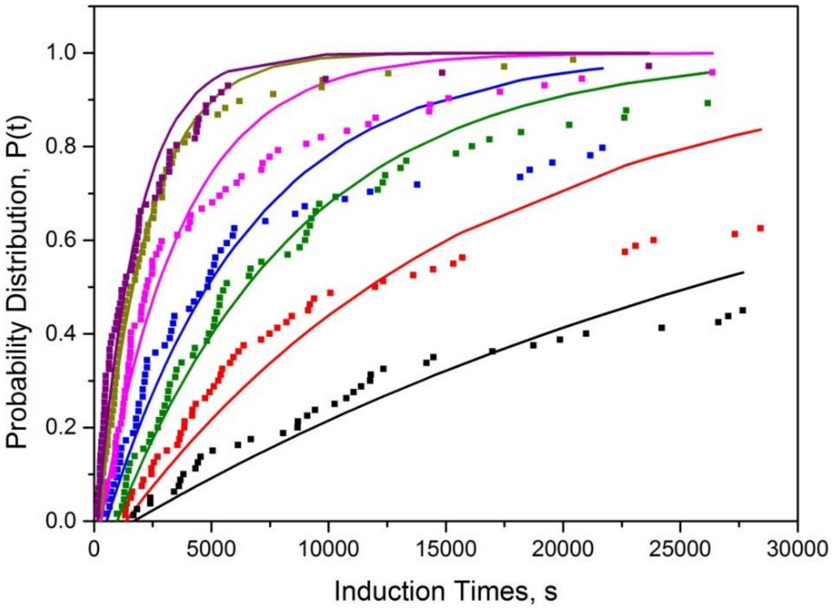


Figure 4.12: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.43 (black), 1.47 (red), 1.51 (green), 1.55 (blue), 1.59 (magenta), 1.61 (yellow) and 1.67 (purple) for PNBA Form I in 2-propanol at 25°C, 900 rpm, 1.6 mL solution at preheating temperature 50°C and sealed vials. Solid lines are the fits to Equation 4.3.

The kinetic and thermodynamic value, A and B, obtained from the plot as shown in Figure 4.13 are 3864 ± 600 m^{-3} s^{-1} and 0.76 ± 0.1 respectively with the Pearson R-value of 98.8 %.

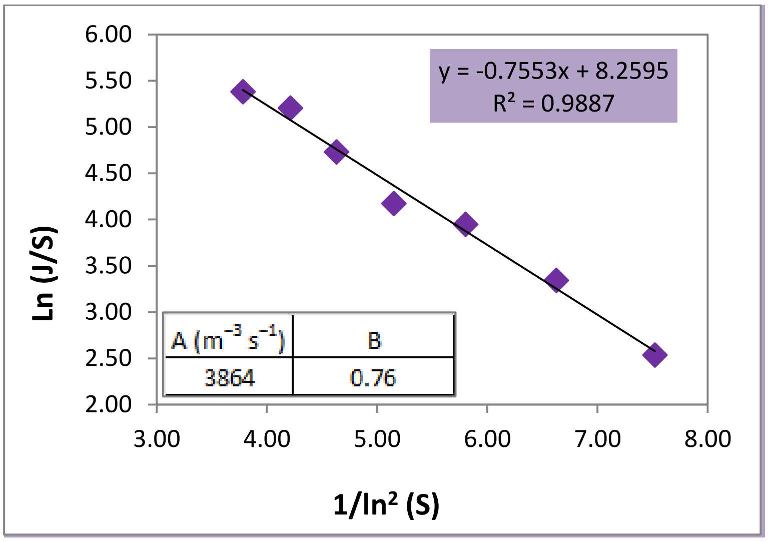


Figure 4.13: Plot of $\ln(J/S)$ as a function of $1/\ln^2 S$ for PNBA in 2-propanol at 900 rpm and 1.6 mL of solution. The inset table summarised the obtained values for A and B.

The plot of the nucleation rate obtained from the induction time variations (marker) and calculated nucleation rate using the obtained A and B (solid line) against supersaturation have good fittings as shown in Figure 4.14.

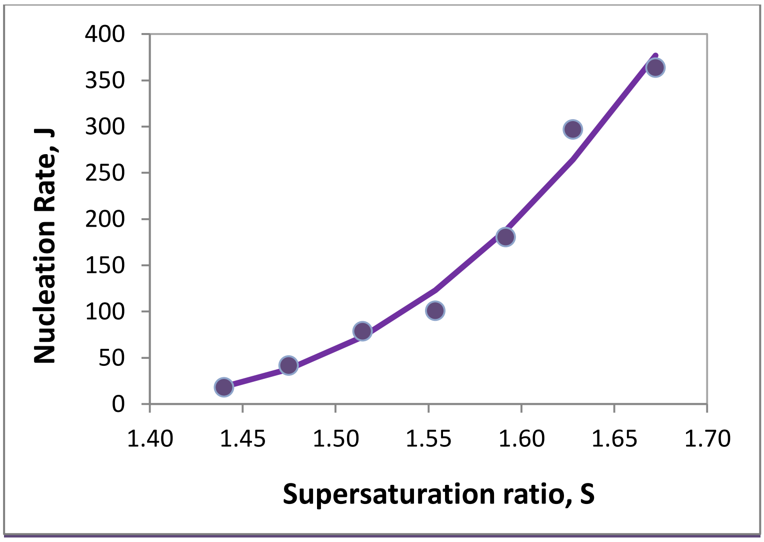


Figure 4.14: Plot of nucleation rate (J) as a function of supersaturation ratio for PNBA in 2-propanol. The solid lines are the calculated nucleation rate based on Equation 4.5.

4.3.7 The Finalised Methodology

Based on all the data recorded above a best methodology can now be recorded. As a preliminary to a full study, a screening test is recommended.

Preliminary Run (Screening Test)

An initial screening test was conducted to determine the volume, minimum and maximum supersaturation, S_{min} and S_{max}, supersaturation interval and preheating temperature for a particular system. The test is performed by preparing each block (4 vials) at different supersaturation, for example at supersaturations of 1.1, 1.2, 1.3 and 1.4. The sample preparation was similar to the one to be described below (under **Full run**) but the bulk solution was reduced to 20 mL. 1.5 mL of solution in each vial was chosen as a start. The other process settings are standardised: five cycles, 900 rpm stirring speed, preheating temperature at 50°C, crash cooling at 5 °C min^{-1} to 25 °C and holding for 8 hours. The screening test is repeated at a different range of supersaturation where the S_{min} should have at least 40/80 samples nucleated, while S_{max} has all samples nucleated immediately above 120 s. Once this range is known, the supersaturation interval is chosen to have at least five sets of supersaturation including the S_{min} and S_{max}. This is to avoid choosing a small interval that gives a small difference in the induction time variation and nucleation rate. The preheating temperature should be at least 10°C higher than the saturated temperature at S_{max} and 10°C below the boiling point of the chosen solvent to minimise the evaporation of the solvent.

Full run

80 g of solvent and the corresponding amount of dissolved solute were placed in a sealed jacketed vessel. The solution was stirred for an hour at a constant controlled temperature of 50 °C using a water bath. The solution (at 50 °C) was dispensed into the vials using pipettes. Pipetting the solution at the same constant temperature will reduce the variation in volume as the solution density is dependent on temperature. Pre-heating the plastic pipettes is not advisable as it affects the actual volume. The vials (reactors) were closed tightly using a rubber top cap and sealed with parafilm. The total weights of the vials were recorded before placing them in the Crystal16 wells. In the crystal 16 setup, the solution was heated up to 50 °C and held for an hour. The solution was then crash cooled at 5 °C/min to 25 °C and held for 8 hours. Five cycles were repeated. 'Crowning' was checked frequently in each cycle visually or by shaking the vials to observe for any solids material

falling down into the solution. The total weights of the vials were recorded at the end of each experiment. The induction times that need to be excluded are induction time less than 120 s [31] and mass loss more than 0.05 % [26]. The polymorphic form was checked for several vials at each supersaturation immediately after solids were formed.

4.3.8 PNBA in Ethyl Acetate

The induction time measurement for PNBA in ethyl acetate are worth a separate mention since they produced confusing results. The initial results were already shown in the Section 4.3.1. The probability curves shown in Figure 4.15 was obtained at 1.6 mL, preheating temperature at 50°C, stirring speed at 900 rpm, the holding time was reduced to 5 hours (due to the fast evaporation rate) and all vials were sealed. The probability curves between the experimental and model fits well, however the nucleation rates did not increase monotonically with the supersaturation ratio. The nucleation rates decrease from 228, 188 and 80 $m^{-3}s^{-1}$ with increasing supersaturation 1.28, 1.32 and 1.36. The 'crowning', solvent evaporation and concentration for each vial (Section 4.2.7) were checked and a significant variation of concentration was found. As an example for targeted supersaturation ratio at 1.24 and 1.32, the actual supersaturation varies from 1.24 to 1.28 and 1.30 to 1.35 respectively.

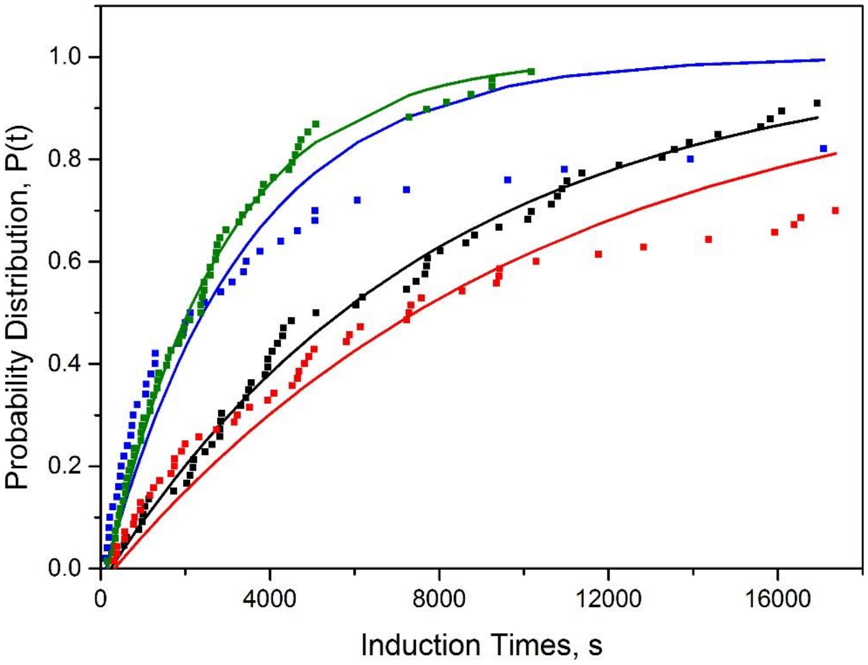


Figure 4.15: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.24 (red), 1.28 (green), 1.32 (blue) and 1.36 (black) for PNBA Form I in ethyl acetate at 25°C, 900 rpm, 1.6 mL solution at preheating temperature 50°C and sealed vials. Solid lines are the fits to Equation 4.3.

This may have resulted from the fast evaporation rate of ethyl acetate, leading to a variation of the concentration during the transfer of 1.6 mL from the bulk solution (at 50 °C) into the vials. In order to reduce this influence, a trial was made by preparing the solution in each vial; for example, 1.60 g of ethyl acetate was mixed with 0.0378 g of PNBA in the vial at room temperature. Eight vials were prepared for each supersaturation 1.32, 1.36, 1.40 and 1.44 with a total of 40 induction times. Again, the 'crowning', solvent evaporation and concentration for each vial were checked and no violation was found. However, the results remain the same as shown in Figure 4.16 that there is no correlation between the supersaturation - 1.32, 1.36, 1.40 and 1.44 and the nucleation rate – 83, 266, 89 and 189 $m^{-3}s^{-1}$. In contrast to the PNBA in 2-propanol positive results as shown in Section 4.3.6, PNBA in ethyl acetate showed negative results with the sealing of the vials.

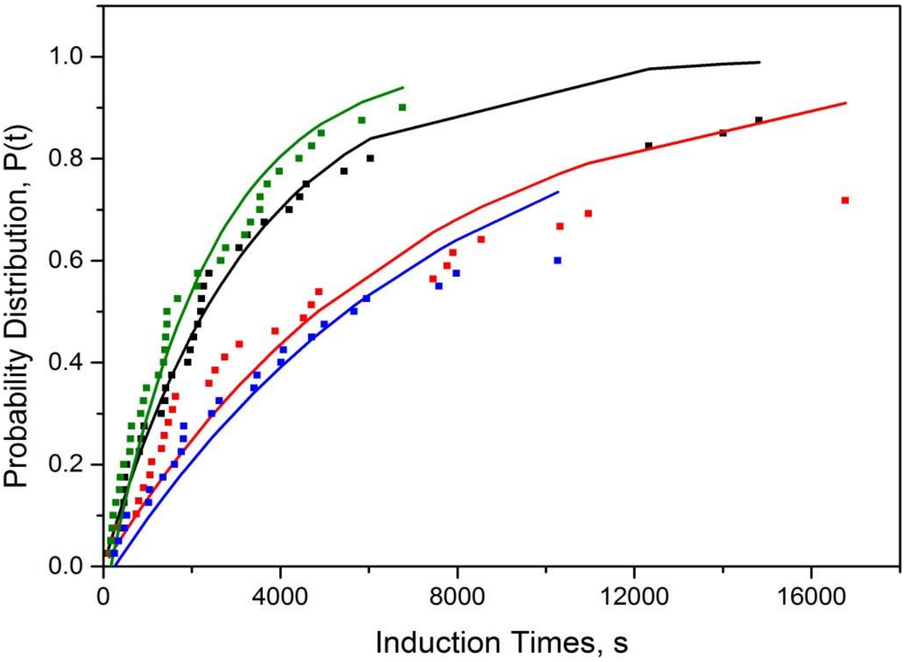


Figure 4.16: Experimentally obtained probability distributions P(t) of the induction times measured at supersaturation ratio 1.32 (blue), 1.40 (red), 1.44 (black) and 1.36 (green) for PNBA Form I in ethyl acetate at 25°C, 900 rpm, 1.6 mL solution at preheating temperature 50°C and sealed vials. Solid lines are the fits to Equation 4.3.

4.4 Conclusion

A preliminary run has been introduced as an essential step in determining the process conditions for nucleation kinetics studies of solute-solvent system in the Crystal 16 apparatus. This screening test determines the minimum and maximum supersaturation range at a starting volume of 1.5 mL and to visually check for any 'crowning'. The fixed parameters are stirring speed at 900 rpm and cooling rate at 5 °C min^{-1}. The minimum supersaturation should have at least 40/80 samples nucleated. The minimum number of probability curves per system is five that is the minimum, maximum and three other supersaturation ratios in between. This reduces the possibility of overlapping probability curves unless such behaviour is inherent in the stochastic nature of the nucleation process for a particular system. The preheating temperature is 10 °C higher than the saturated temperature at maximum supersaturation, provided it is still below the solvent boiling temperature.

The significant variation of induction times, nucleation rates, kinetic and thermodynamic factors with process conditions showed the high sensitivity of the Crystal 16 technique to details in the experimental procedure. However, there is also a possibility that the observed extent of the sensitivity is somewhat specific to the chosen compound, PNBA. In troubleshooting the induction time variation for PNBA, a methodology was established that provides more consistency. This methodology takes into consideration all aspects of sample preparation such as dispensing technique, securing the vials from solvent evaporation and thermal history of the solution. Understanding the nucleation process of PNBA in 2-propanol and ethyl acetate had been a challenge, which highlights the importance of well controlled experimental settings for nucleation experiments.

5. ELECTRONIC STRUCTURE OF *PARA*-NITROBENZOIC ACID (PNBA) CRYSTAL POLYMORPHS: XPS AND NEXAFS STUDIES

Abstract: *para*-nitrobenzoic acid (PNBA) polymorphs, Form I and Form II, are carboxylic acid dimers with only subtle structural differences in the solid state. X-ray photoelectron spectroscopy (XPS), near-edge X-ray absorption fine structure (NEXAFS), and density functional theory (DFT) ground state calculation were combined to examine the electronic structure of both polymorphs experimentally and theoretically. Both polymorphs showed similarity in the binding energy (BE) of the 1s core electrons and the transition energies to the lowest unoccupied molecular orbitals (LUMOs) at the C, N and O K-edges. Small differences arise only from differences in the hydrogen bonding in the carboxylic acid dimer, which becomes evident for Form I in the XPS O 1s (C=**O**) spectrum through a 0.3 eV shift relative to Form II and in ATR-FTIR through the OH out-of- plane wag at a wavenumber −10 cm^{-1} lower.

5.1 Introduction

X-ray photoelectron spectroscopy (XPS) and near-edge X-ray absorption fine structure (NEXAFS) are emerging techniques for probing structure and bonding in the organic solid state, including hydrogen bonding and proton transfer. XPS probes the chemical shifts in the binding energies of atomic core levels. Irradiation with monochromatic X-rays causes ionisation, whereby core electrons (e.g. 1s levels) of a specific element are emitted from the sample by absorption of photons with sufficient energy *hv*. The binding energy (BE) is calculated from the known *hv* and the measured kinetic energy of the emitted photoelectrons [97].

In NEXAFS, the photon energy is continuously scanned across the excitation thresholds (‘absorption edges’) for photoabsorption by atomic core levels, monitoring the associated variations in absorption coefficient for the photons. In the near-edge region of these spectra the core electron (e.g., 1s, then also known as a K−edge spectrum) receive sufficient energy for transitions to unoccupied molecular orbitals [45], but no photoionisation by photoemission takes place. At photon energies beyond the ionisation threshold (commonly referred to as the ionisation potential, IP) photoemission takes place and features structure in the spectra are observed due to constructive and destructive interference of the photoelectron wave through intraatomic and interatomic scattering. The interpretation of XPS and NEXAFS spectra is commonly facilitated by fundamental computational quantum chemical methods such as density functional theory (DFT), such as CASTEP [51]. XPS, NEXAFS and computational methods give complementary information for understanding how local bonding influences crystal structure formation.

XPS and NEXAFS have recently been shown to be sensitive to the effects of intramolecular electronic structure as well as hydrogen bonding. The IP was shown to be sensitive to the strength and the length of hydrogen bond [98,99]. Chemical shifts were observed as well as a result of the substituent and π* delocalisation effects on the hydrogen bond [100]. The features of the spectra of a gas phase (monomer) and solid phase (dimer) molecular species are significantly different, which further highlights again the sensitivity of core level spectroscopy to intermolecular interaction [51].

XPS and NEXAFS are not yet widely used for characterising organic polymorphs, although studies of non-organic polymorphs are established. For example, calcium carbonate polymorphs, such as calcite, aragonite, and vaterite, are characterised by similar XPS C 1s spectra, but the O 1s are significantly different [101]. The surface properties of silica polymorphs, α−quartz and α−cristobalite, in adsorption experiments showed identical peak positions but different peaks ratio of the XPS O 1s and N 1s [102]. The three polymorphs of aluminium silicate - andalusite, sillimanite, and kyanite, have different overall peak position and peaks ratio for the O 1s, which correspond to the differences in local charge density [103]. There were no difference observed in the XPS between the α− and γ−glycine [104]. Predicted calculations on the NEXAFS spectra showed significant difference between the three palladium oxide polymorphs [105].

The sensitivity of NEXAFS and XPS for probing the effects of local environments and interactions in organic polymorphic system was first investigated for *p*-aminobenzoic acid (PABA) [50]. α− and β−PABA have the same chemical composition but different crystal structure. The building unit for α-PABA is carboxylic acid dimer while that for β-PABA is a carboxylic acid interaction with $-NH_2$. Thus, β-PABA has lower electron density on the nitrogen compared to the α-PABA. The difference was captured in XPS and NEXAFS. In XPS, the binding energy N 1s for β-PABA is higher by 0.2 eV and in NEXAFS the excitation of N1s to the 1π* and 3π* is higher by 0.58 eV and 1.60 eV as well as differences in the σ* region so that a unique electronic signature is provided for the polymorphs with NEXAFS, distinguishing between the two forms.

p-nitrobenzoic acid (PNBA) has two polymorphic structures, PNBA Form I and II, with a similar building unit, the carboxylic acid dimer (C=O···H−O). The differences between these polymorphs are so subtle that only small differences in the bending of O−H were detected using vibrational spectroscopy (Raman and ATR-FTIR in Chapter 3). The aim of this study is to further investigate the differences between the two polymorphs by utilising core level spectroscopy techniques, XPS and NEXAFS, assisted by DFT calculations. Given the much subtler structural differences in the PNBA polymorphs compared to PABA, this study was also an exercise in examining the sensitivity of XPS and NEXAFS techniques on structural differences between organic polymorphs.

5.2 Methodology

The XPS and NEXAFS spectra for both forms were measured and the interpretation was assisted by DFT calculations using Gaussian with the B3LYP/631G* functional and basis set. PNBA Form I and PNBA Form II were re-crystallised in solution at 25 °C and 55 °C from the purchased material (Form II).

5.2.1 X-ray Photoelectron Spectroscopy (XPS)

Carbon (C), nitrogen (N) and oxygen (O) 1s spectra were measured using a Kratos Axis Ultra instrument; the monochromatic Al K_α source process settings are 1486.69 eV, 180 W, 15 kV and 12 mA. The measurement was taken in constant analysis energy mode below 10^{-8} mbar with a pass energy of 20 eV, 0.1 eV steps and 1000 ms dwell time per data point. Analysis of the data was performed with CasaXPS software, using a linear background and GL (30) line shape (70% Gaussian, 30% Lorentzian). Samples were referenced to the C=C carbon signal at 286.1 eV. Further details were as described in Section 2.10.

5.2.2 Near Edge X-ray Absorption Fine Structure (NEXAFS) Spectroscopy

C and N K-edge partial electron yield (PEY) were performed with an entrance grid of -50 V for C and -150 V for N at the U7a beamline of the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory, USA. Spectra were normalised by the simultaneously recorded drain current from an in situ gold-coated, 90% transmission grid (I_0) and the energy scale calibrated using the 285.1 eV first π* resonance of an amorphous carbon grid for the C K-edge and 400.6 eV first π* resonance of a titanium nitride grid for the N K-edge. O K-edge total electron yield (TEY) were performed on the U49-2_PGM beamline of the BESSY II, Berlin. Calibration of the beamline energy scale was performed with N_2 gas (TEY) as 1s→2p p transition. Analysis of the data was performed using the Athena software. Further details were as described in Section 2.11.

5.2.3 Density Functional Theory (DFT) Calculation

DFT calculations were performed using Gaussian 90W with B3LYP/631G* for both polymorphic PNBA crystal structures without structural optimisation at the ground state. This is because structural optimisation would not take into consideration polymorphic crystal structures. The calculations were performed on a single molecule (monomer) and two molecules (dimer). The O and H atoms (C=**O**···**H**O) were fixed in order to study the subtle effect of hydrogen bond length on the ionisation potential.

5.3 Results

The XPS, NEXAFS and DFT calculations will be discussed for each chemical species (**C**=C, **C**-N and **C**=O for carbon, C=**O**, **O**-H and N**O**$_2$ for oxygen and **N**O$_2$ for nitrogen) in both PNBA Form I and II.

5.3.1 X-ray Photoelectron Spectroscopy (XPS)

The binding energy scale in the XP spectra was referenced to the C 1s (**C**=C) peak at 286.1 eV instead of the adventitious carbon (contamination) peak at 284.8 eV. This is because the differences in the contamination level for each crystal lead to inconsistency of the chemical shift. The adventitious carbon peak for PNBA Form I is at 285 eV while for Form II it is at 284.8 eV which is still within the expected range 284.7 to 285.2 eV [106].

C 1s XPS for PNBA Form I and II have similar features, with small differences at the contamination peak. The fitted peaks of adventitious carbon (contamination), **C**=C, **C**-N and **C**=O are shown and summarised in Figure 5.1 and Table 5-1. There are no significant differences in the binding energy (ionisation potential) of C 1s (**C**=C), C 1s (**C**-N) and C 1s (**C**=O), which occur at 286.1 eV, 287.3 eV and 290.4 eV respectively. This is followed by the broad shake-up satellite at the higher binding energy of 292.9 eV. The re-crystallised PNBA Form I has higher contamination or possibly having additional contamination relative to the purchased material (PNBA Form II) based on the area under the peak.

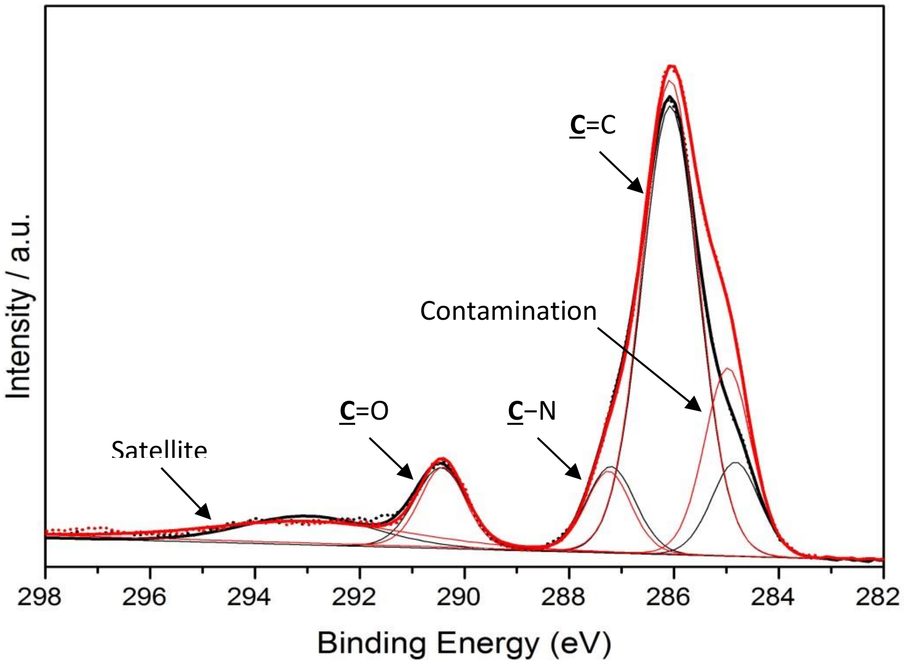


Figure 5.1: Experimental C 1s XPS of PNBA Form I (red), and PNBA Form II (black) and the fitted components peaks, showing the shifts in binding energy (IP) for different chemical species.

Table 5-1: Binding energy (IP) of the experimental C 1s XPS of PNBA Form I and II and the difference in binding energy ($\Delta BE_{I-II} = BE_{form\ I} - BE_{form\ II}$)

Peak Assignment	Binding Energy (eV)		
	Form I	Form II(p)	ΔBE_{I-II}
Contamination	284.97	284.83	0.14
C 1s (<u>C</u>=C)	286.07	286.07	0.00
C 1s (<u>C</u>–N)	287.25	287.20	0.04
C 1s (<u>C</u>=O)	290.42	290.45	-0.04
Satellite	292.94	292.97	-0.03

The N 1s XPS spectra are also similar for both polymorphic forms, with the binding energy for the **N**O_2 group at 406.7 eV, as shown in Figure 5.2 and Table 5-2.

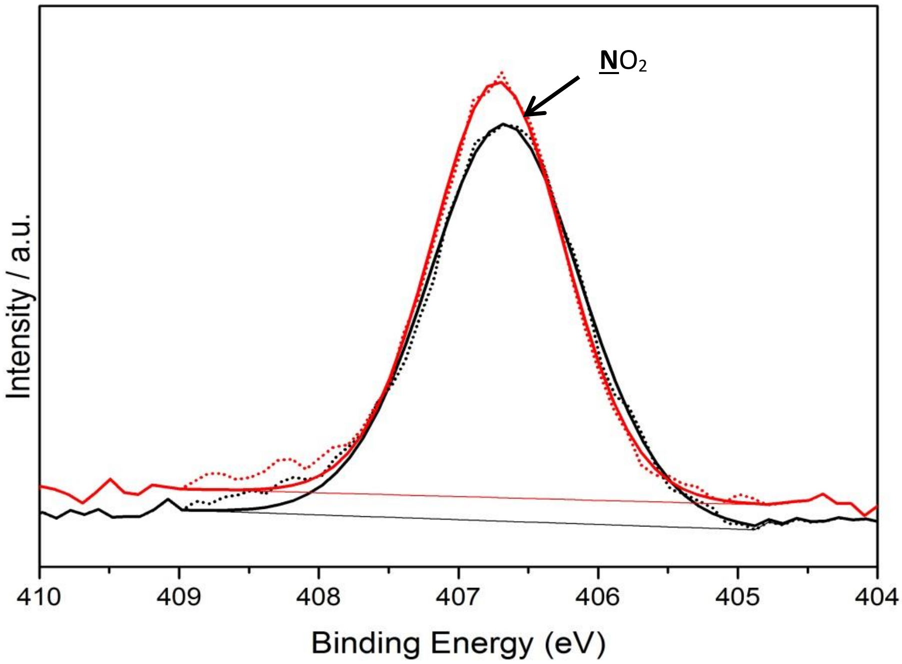


Figure 5.2: Experimental N 1s XPS of PNBA Form I (red), and PNBA Form II (black) and the fitted components peaks, showing the binding energy (IP) for NO_2.

Table 5-2: Binding energy (IP) of the experimental N 1s XPS of PNBA Form I and II and the difference in binding energy ($\Delta BE_{I-II} = BE_{form\ I} - BE_{form\ II}$)

Peak Assignment	Binding Energy (eV)		
	Form I	Form II(p)	ΔBE_{I-II}
N 1s (**N**O_2)	406.72	406.67	0.05

The O 1s spectra in Figure 5.3 show PNBA Form I has a broader shoulder towards lower binding energy in comparison to Form II. The XPS has signals from the three different types of oxygen environment and a shake-up satellite at high energy. Photoemission from the O 1s (**O**−H) oxygen occurs around 534.6 eV, O 1s (N**O**$_2$) at 533.7 eV, and the lowest binding energy peak component arises from the O 1s (C=**O**). The O 1s for **O**−H and N**O**$_2$ signals remain quasi-constant for the two polymorphs, while the ionisation potential for the O 1s (C=**O**) is −0.30 eV lower than PNBA Form II (Table 5-3), which will be discussed further.

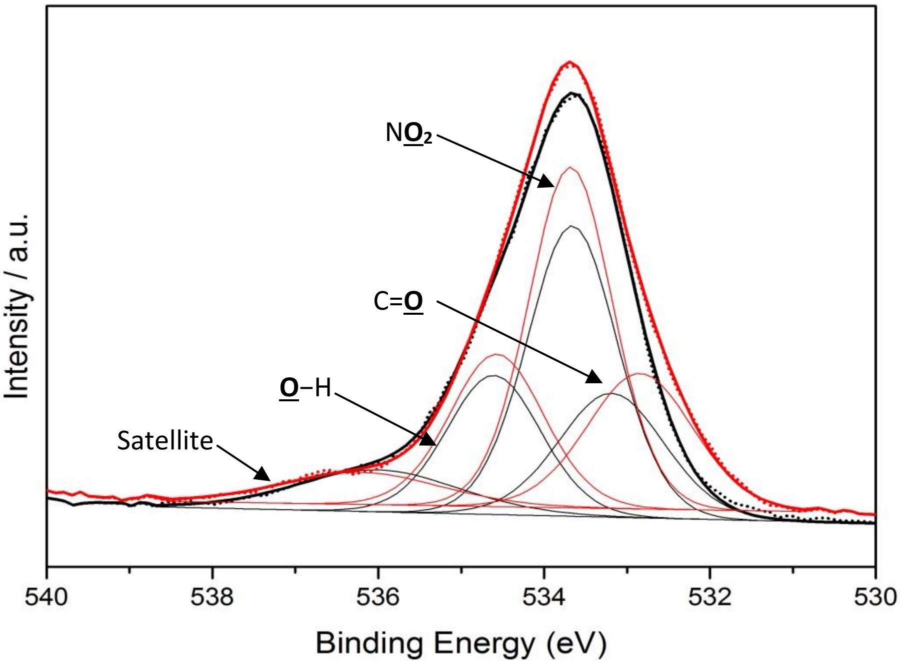


Figure 5.3: Experimental O 1s XPS of PNBA Form I (red), and PNBA Form II (black) and the fitted components peaks, showing the shifts in binding energy (IP) for different chemical species.

Table 5-3: Binding energy (IP) of the experimental O 1s XPS of PNBA Form I and II and the difference in binding energy (($\Delta BE_{I\text{-}II}$ = $BE_{form\ I}$ – $BE_{form\ II}$)), highlighting the shift for the C=O signal.

Peak Assignment	Binding Energy (eV)		
	Form I	Form II(p)	$\Delta BE_{I\text{-}II}$
O 1s (C=**O**)	532.85	533.19	**-0.34**
O 1s (N**O**$_2$)	533.68	533.66	0.02
O 1s (**O**–H)	534.57	534.61	-0.03
Satellite	536.21	536.01	0.20

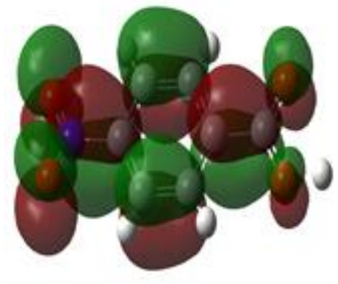
5.3.2 Density Functional Theory (DFT) Calculation

The orbital energies of interest here for XPS and NEXAFS are the core level (1s) and the lowest unoccupied molecular orbitals (LUMOs). Table 5-4 shows the summary of molecular orbital energies for PNBA Form I and II in single molecule (monomer) systems. The lowest energy for the core level 1s (and thus highest binding energy required to eject the electron in XPS) is the oxygen group, followed by nitrogen and then carbon.

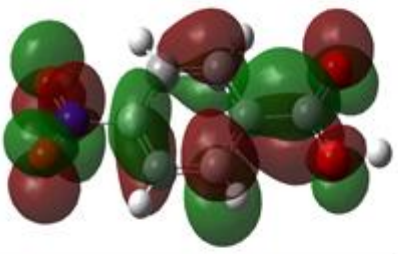
Table 5-4: The calculated molecular orbital energy levels for the chemical species at the lowest occupied (core) levels and the lowest unoccupied molecular orbitals for a single molecule of each of the polymorphic forms of PNBA.

Orbital Number	MO	PNBA Form I		PNBA Form II	
		Assignment	Energy, eV	Assignment	Energy, eV
48	LUMO+4	4π*	1.31	4π*	1.21
47	LUMO+3	1σ*	0.76	1σ*	0.09
46	LUMO+2	3π* (4 ring C)	-1.02	3π* (4 ring C, NO)	-1.02
45	LUMO+1	2π*	-1.06	2π*	-1.13
44	LUMO	1π*	-2.76	1π*	-2.83
12	C 1s	**C**=C	-277.92	**C**=C	-277.79
11	C 1s	**C**=C	-277.98	**C**=C	-277.94
10	C 1s	**C**=C	-278.02	**C**=C	-277.95
9	C 1s	**C**=C	-278.04	**C**=C	-278.05
8	C 1s	**C-**COOH	-278.49	**C-**COOH	-278.49
7	C 1s	**C**-NO_2	-279.27	**C**-NO_2	-279.33
6	C 1s	**C**OOH	-281.32	**C**OOH	-281.38
5	N 1s	$\mathbf{N}O_2$	-396.72	$\mathbf{N}O_2$	-396.73
4	O 1s	C=**O** (COOH)	-520.98	C=**O** (COOH)	-521.00
3	O 1s	$N\mathbf{O}_2$ (OH side)	-522.10	$N\mathbf{O}_2$ (OH side)	-522.10
2	O 1s	$N\mathbf{O}_2$ (C=O side)	-522.14	$N\mathbf{O}_2$ (C=O side)	-522.14
1	O 1s	**O**H (COOH)	-523.23	**O**H (COOH)	-523.29

NEXAFS is determined by the excitation of electrons from the core levels into bound unoccupied orbitals. The visualisation of the lowest unoccupied molecular orbitals (LUMOs) (Figure 5.4) reveals that significant density of states of the LUMO (1π*) and the LUMO+1 (2π*) are present at the nitro group, with the whole orbital delocalised across the nitro, aromatic ring and carboxylic acid moieties. The main electronic structure difference between the polymorphs is associated with the LUMO+2 (3π*). In PNBA Form I, there is no density of states of the LUMO+2 at the nitro group, while in PNBA Form II, the nitro group do participate, albeit weakly, in the orbital (as circled in red in Figure 5.4).



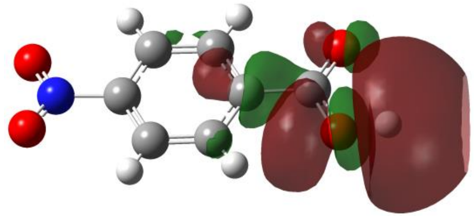
LUMO (1π*)



LUMO+1 (2π*)

LUMO+2 (3π*) - PNBA Form I

LUMO+2 (3π*) - PNBA Form II



LUMO+3 (1σ*)

LUMO+4 (4π*)

Figure 5.4: Visualisation of the five lowest unoccupied MOs (LUMOs) for Form I and II of PNBA based on the DFT calculations.

Transitions from the atomic core 1s level to a bound unoccupied state are only significant when the X-ray absorbing atom participates in the unoccupied MO to which the transition takes place. A scheme of the energy transitions for carbon 1s core level excitation to the four LUMOs is shown in Figure 5.5. Both polymorphic forms have the same calculated ionisation potential and energy levels, except for an additional C 1s transition (<u>C</u>–C → 3π*) for PNBA Form II which is small and negligible.

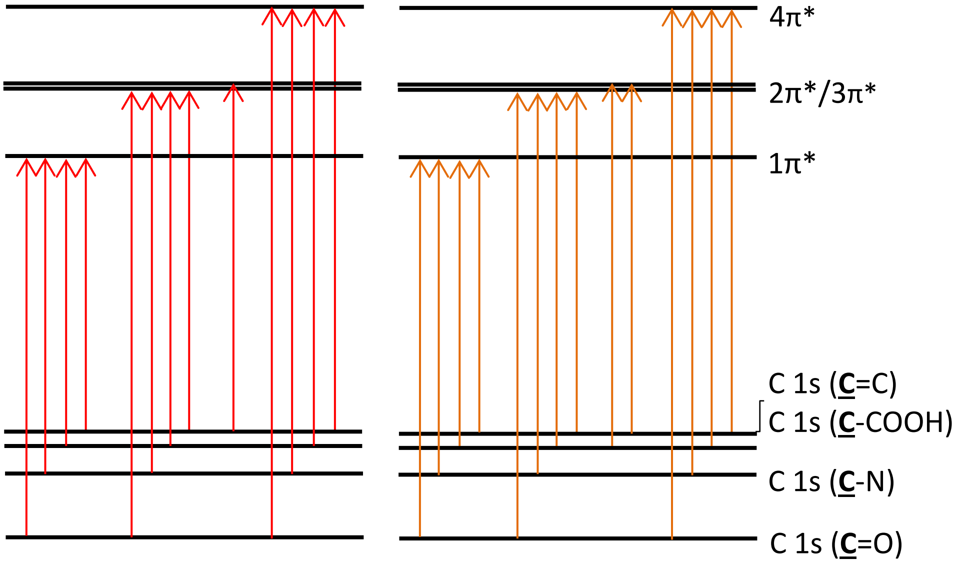


Figure 5.5: Schematic of transitions from carbon core 1s levels to the four LUMOs for PNBA Form I (red) and PNBA Form II (orange) based on DFT calculation on a monomer system.

Excitation of electrons from the carbon core level C 1s (<u>C</u>=O, <u>C</u>-N, <u>C</u>=C and <u>C</u>-COOH) are therefore allowed to 1π*, 2π* and 4π* (Figure 5.4 and Figure 5.5). The difference of the energy levels of <u>C</u>-C, <u>C</u>-N and <u>C</u>=O relative to <u>C</u>=C is +0.4 eV, +1.2 eV and +3.3 eV respectively. Therefore, C 1s (<u>C</u>=O → 1π*) needs the highest energy to excites to LUMO (1π*) at 278.6 eV, followed by C 1s (<u>C</u>-N → 1π*) at 276.5 eV, C 1s (<u>C</u>-C → 1π*) at 275.7 eV and C 1s (<u>C</u>=C → 1π*) at 275.3 eV.

The N 1s schematic energy transitions for nitrogen core level 1s to the four LUMOs show significant differences between the two PNBA polymorphs, as the transition to 3π* is more significant for nitrogen and oxygen than carbon (Figure 5.4). PNBA Form I has a transition from N 1s to LUMO (1π*), LUMO +1 (2π*) and LUMO +3 (4π*), while PNBA Form II has the additional transition to LUMO +2 (3π*) as shown in Figure 5.6. For PNBA Form I, the calculated energy needed for transition from N 1s to 1π*, 2π* and 3π* is 394 eV, 395.7 eV and 398 eV respectively – the successive energy differences relative to the LUMO (1π*) are +1.7 eV and +4.0 eV. For PNBA Form II, the calculated energy needed for transition from N 1s to 1π*, 2π*, 3π* and 4 π* is 393.9 eV, 395.6 eV, 395.7 eV and 397.9 eV respectively – the successive energy differences relative to the LUMO (1π*) are +1.7 eV, +1.8 eV and +4.0 eV.

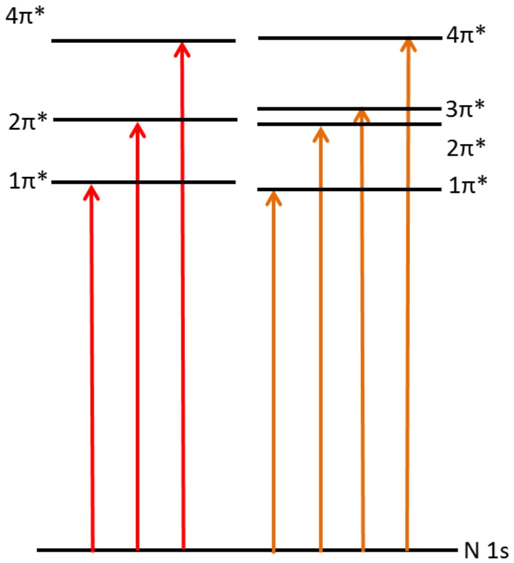


Figure 5.6: Schematic of transitions from N 1s levels to the four LUMOs for PNBA Form I (red) and PNBA Form II (orange) based on DFT calculation on a monomer system.

The O 1s schematic energy transitions from the oxygen 1s core levels to the four LUMOs show both polymorphic forms have the same calculated ionisation potential and energy level except for a tiny and negligible difference (Figure 5.7). PNBA Form II have an additional transition for O 1s (N**<u>O</u>**$_2$ → 3π*) and O 1s (**<u>O</u>**H →3π*). Relative to the O 1s (**<u>O</u>**H →1π*), the energy difference is +1.1 eV for O 1s (N**<u>O</u>**2 →1π*) and +2.3 eV for O 1s (C=**<u>O</u>** → 1π*). The calculated energy transition for the O 1s (OH →1π*), O 1s (N**<u>O</u>**2 →1π*) and O 1s (C=**<u>O</u>** → 1π*) are 518.2 eV, 519.3 eV and 520.5 eV.

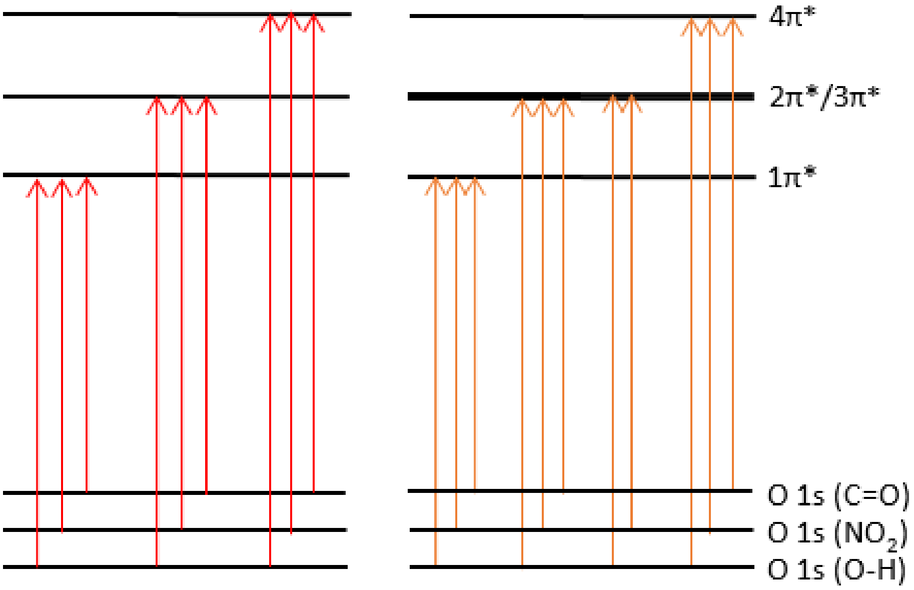


Figure 5.7: Schematic of transitions from N 1s levels to the four LUMOs for PNBA Form I (red) and PNBA Form II (orange) based on DFT calculation on a monomer system.

5.3.3 Near Edge X-ray Absorption Fine Structure (NEXAFS)

The C K-edge spectra (Figure 5.8) show similarity for both polymorphic forms with a sharp double peak at 284.8 eV and 285.7 eV followed by another significant peak at 288.8 eV. These transitions are assigned to C 1s (**<u>C</u>**=C → 1π*), C 1s (C–N → 1π*) and C 1s (C=O →1π*). Based on the DFT calculations, the peaks in-between C–N → 1π* and C=O →1π* are assigned to **<u>C</u>**=C →2π*, **<u>C</u>**=C →3π*, **<u>C</u>**–C →2π* and **<u>C</u>**–N →2π*, and these are represented by fitting the experimental spectra with component peaks (Figure 5.9 and Figure 5.10).

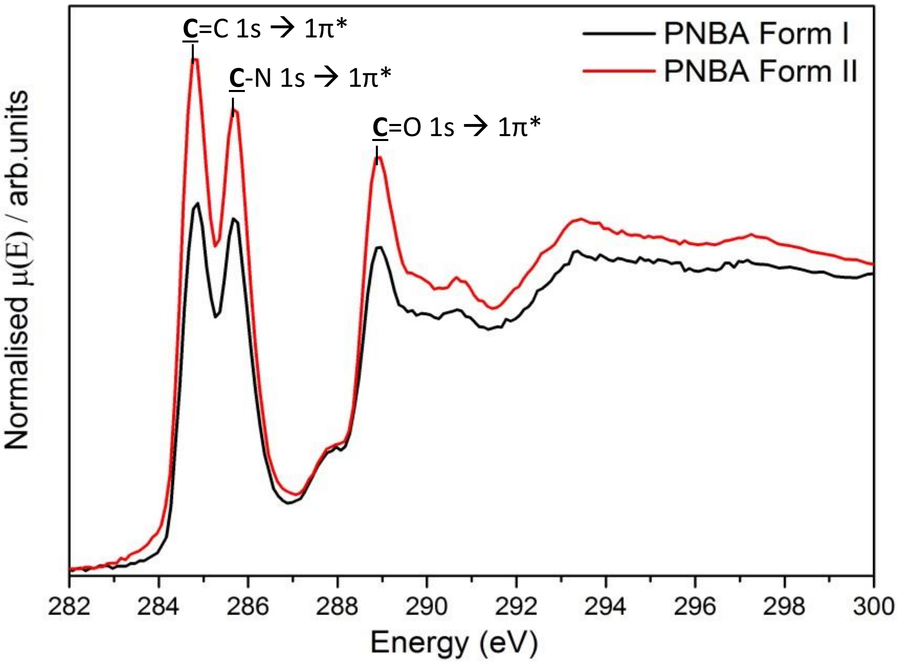


Figure 5.8: Normalised spectra of experimental carbon K-edge NEXAFS of PNBA Form I (black) and Form II (red).

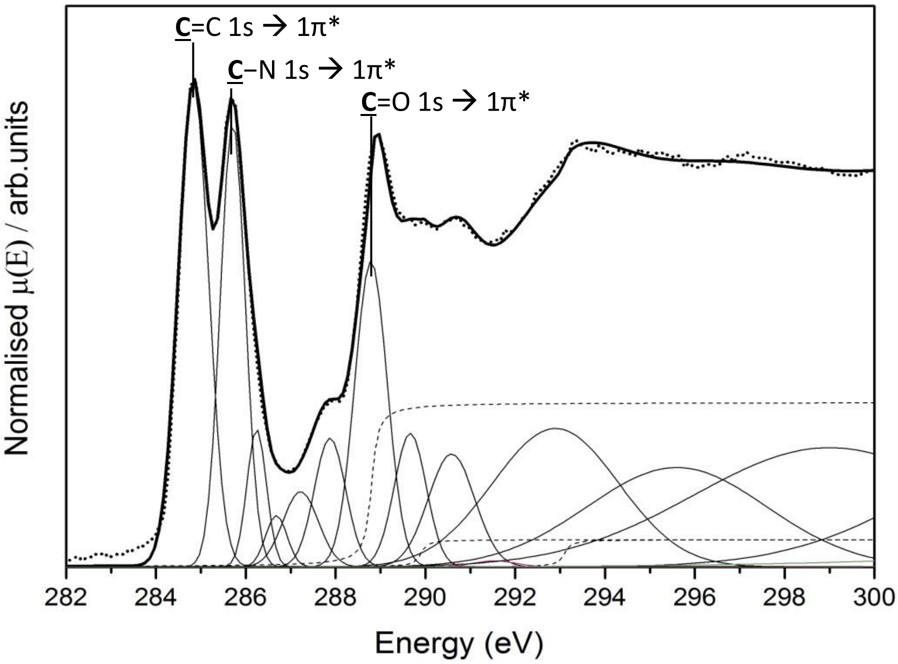


Figure 5.9: Experimental carbon K-edge NEXAFS of PNBA Form I and a model of the spectrum obtained by fitting Gaussians and arctan step functions.

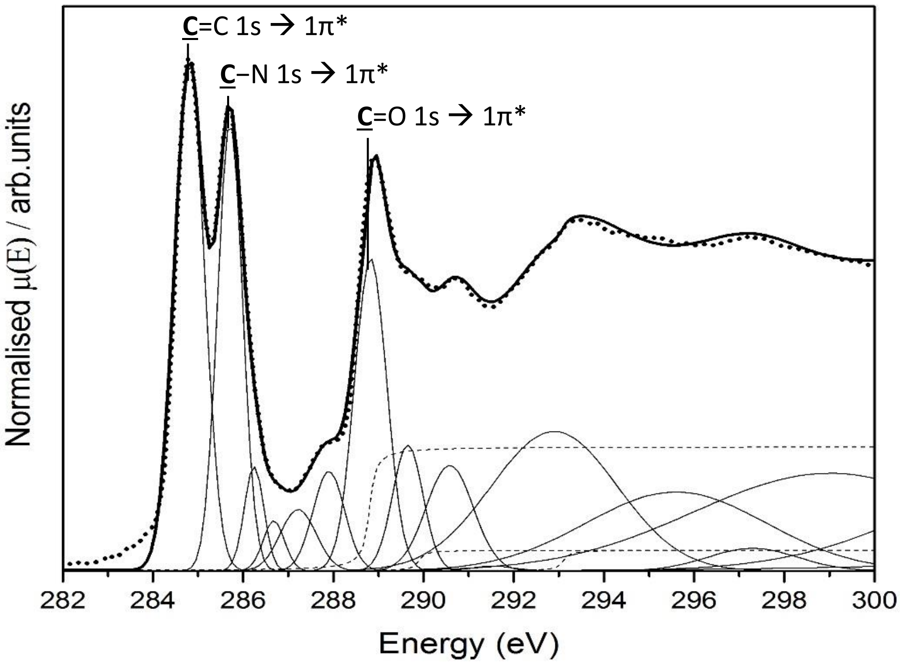


Figure 5.10: Experimental carbon K-edge NEXAFS of PNBA Form II and a model of the spectrum obtained by fitting Gaussians and arctan step functions.

Table 5-5 summarises the transition energies calculated at the ground state (Section 5.3.2) in comparison to the experimental results. The main interest here is the relative energy difference (ΔE*) for each transition relative to C 1s (C=C) → 1π*. ΔE* is based on PNBA Form I and shows a reasonable similarity between the calculated and experimental.

Table 5-5: The comparison of absolute and relative transition energies between DFT calculation and NEXAFS experimental results. The ΔE* is the difference between PNBA Form I transition relative to C 1s (C=C) → 1π*.

Peak Assignments	DFT			NEXAFS		
	Form I, eV	Form II, eV	ΔE*	Form I, eV	Form II, eV	ΔE*
C 1s (C=C) → 1π*	275.27	275.27	0.00	284.81	284.81	0.00
C 1s (C-C) → 1π*	275.73	275.66	0.46	/	/	/
C 1s (C-N) → 1π*	276.50	276.50	1.23	285.71	285.71	0.90
C 1s (C=C) → 2π*	276.97	276.97	1.70	286.24	286.24	1.43
C 1s (C=C) → 3π*	277.02	277.02	1.75	286.68	286.68	1.87
C 1s (C-C) → 2π*	277.43	277.36	2.16	287.22	287.22	2.41
C 1s (C-N) → 2π*	278.20	278.20	2.93	287.88	287.90	3.07
C 1s (C=O) → 1π*	278.56	278.55	3.29	288.79	288.83	3.98

The N K-edge spectra show that both PNBA Form I and II have similar features but differences in intensity (Figure 5.11). The fitted peaks show the transition of N 1s to the LUMO (1π*), LUMO+1 (2π*), LUMO+2 (3π*) and LUMO+3 (4π*) at 404.2 eV, 405.3 eV, 406.8 eV and 408.5 eV respectively for both PNBA Form I (Figure 5.12) and PNBA Form II (Figure 5.13).

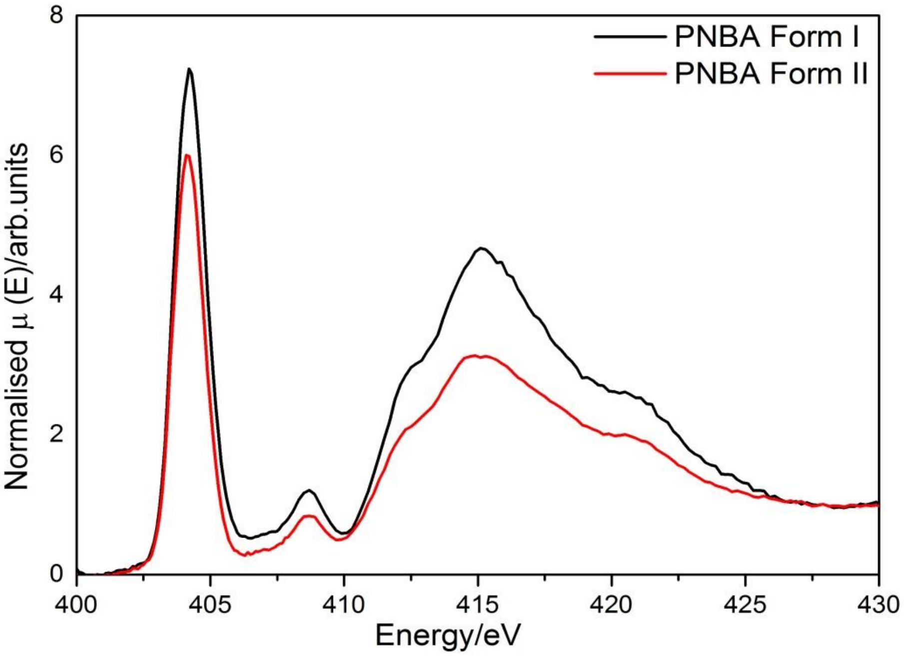


Figure 5.11: Normalised spectra of experimental nitrogen K-edge NEXAFS of PNBA Form I (black) and Form II (red).

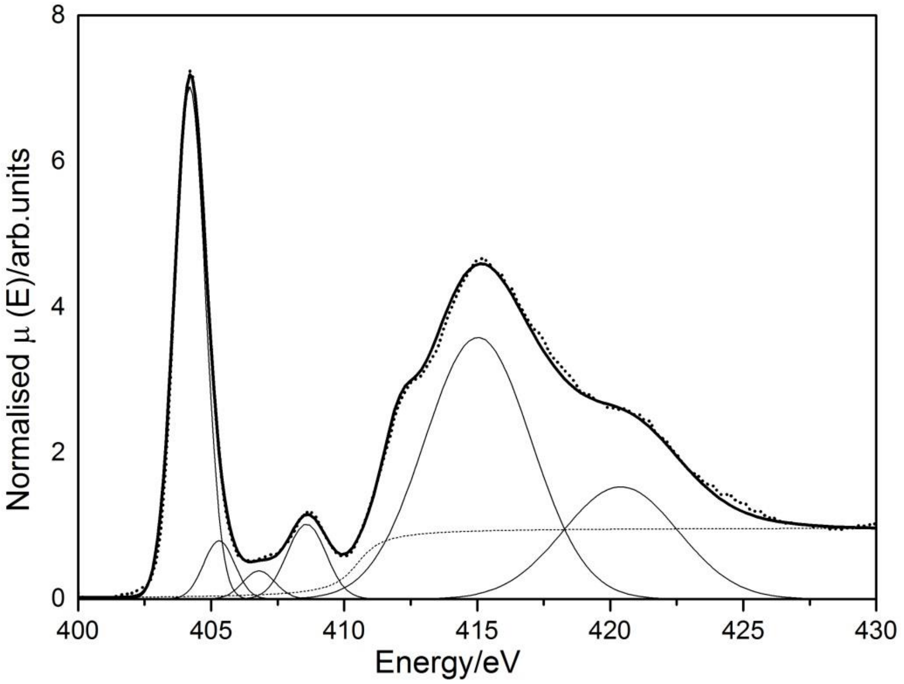


Figure 5.12: Experimental nitrogen K-edge NEXAFS of PNBA Form I and a model of the spectrum obtained by fitting Gaussians and arctan step functions.

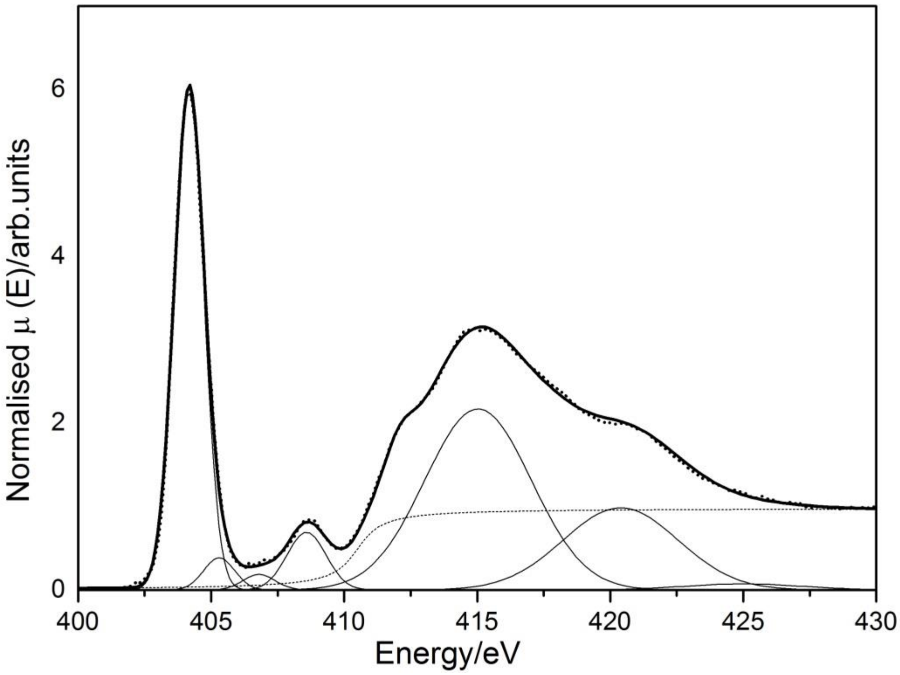


Figure 5.13: Experimental nitrogen K-edge NEXAFS of PNBA Form II and a model of the spectrum obtained by fitting Gaussians and arctan step functions.

Table 5-6 summarises the transition energies calculated at the ground state in comparison to the experimental results. The experimental results for PNBA Form I are not in good agreement with the DFT calculation at the N 1s → 3π*. Nevertheless, based on PNBA Form II the relative energy difference (ΔE_{II}*) shows a reasonable similarity.

Table 5-6: A summary of the transition energy based on DFT calculation and NEXAFS for both PNBA Form I and II.

	DFT			NEXAFS		
N 1s (NO_2) to:	Form I, eV	Form II, eV	ΔE_{II}*	Form I, eV	Form II, eV	ΔE_{II}*
LUMO (1π*)	393.96	393.90	0.00	404.20	404.17	0.00
LUMO +1 (2π*)	395.66	395.61	1.71	405.30	405.30	1.13
LUMO +2 (3π*)	/	395.72	1.82	406.80	406.80	2.63
LUMO +4 (4π*)	398.03	397.94	4.04	408.58	408.54	4.37

The O K-edge spectra for PNBA Form I and II are similar, as showed in Figure 5.14, with two sharp peaks indicating 1s to π* excitations while the broader peak at higher photon energy is the 1s to the σ* excitation. The first peak consist of the two transitions, O 1s C=**O** to the LUMO (1π*) at 530.6 eV and O 1s N**O**$_2$ to the LUMO (1π*) at 531.6 eV, while the smaller intensity peak is for the **O**H to the LUMO (1π*) at 533.2 eV (Figure 5.16 and Figure 5.16).

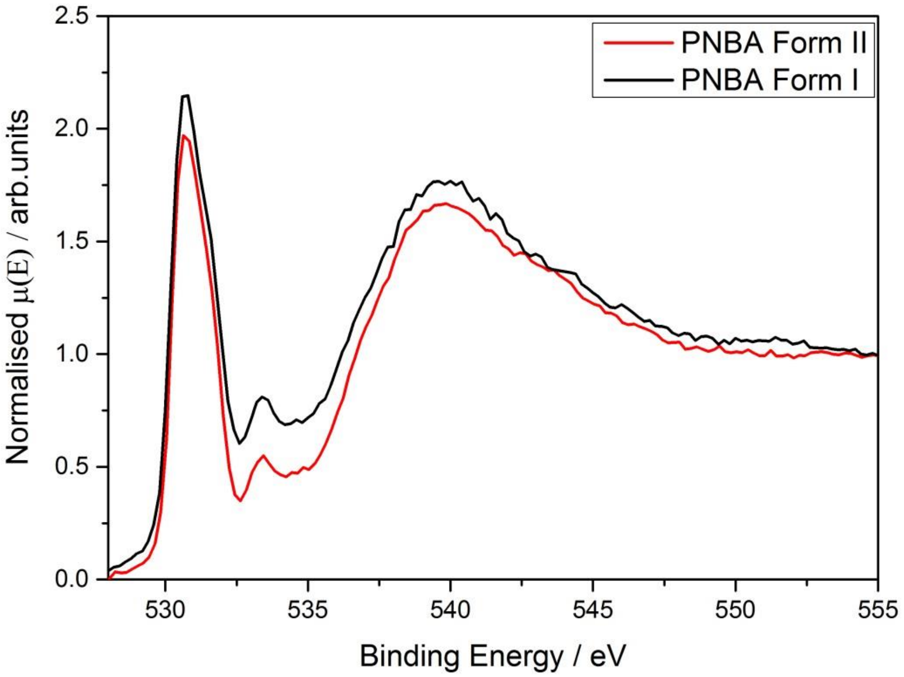


Figure 5.14: Normalised spectra of experimental nitrogen O-edge NEXAFS of PNBA Form I (black) and Form II (red).

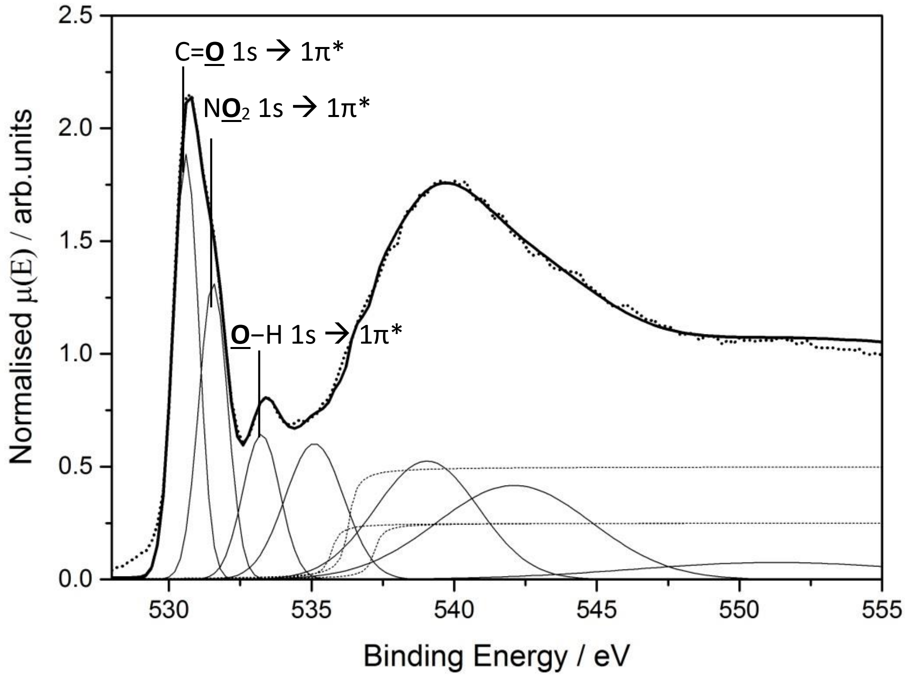


Figure 5.15: Experimental carbon O-edge NEXAFS of PNBA Form I and a model of the spectrum obtained by fitting Gaussians and arctan step functions.

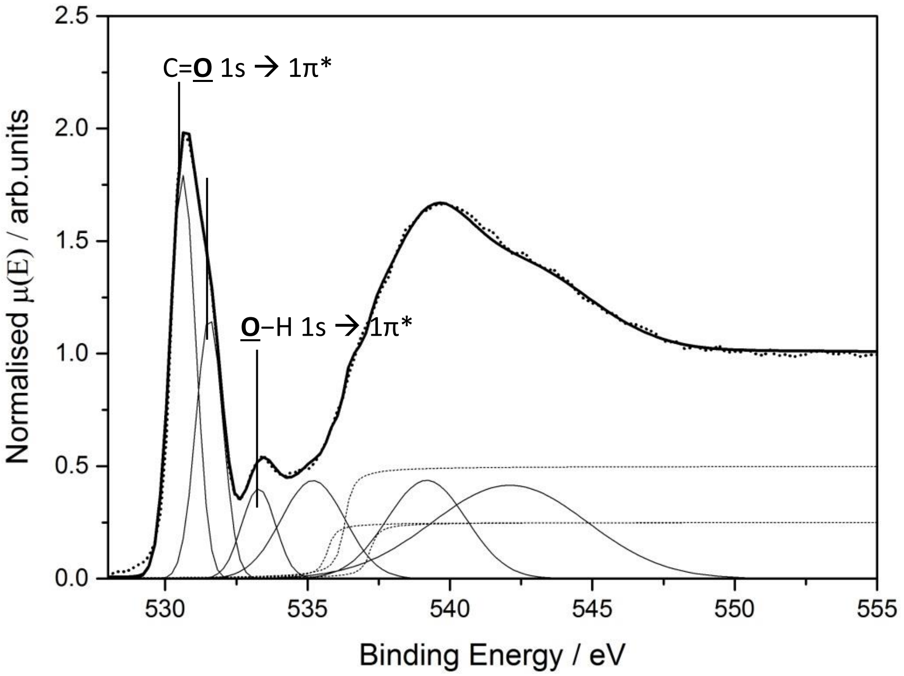


Figure 5.16: Experimental carbon O-edge NEXAFS of PNBA Form II and a model of the spectrum obtained by fitting Gaussians and arctan step functions.

Table 5-7 summarises the transition energies calculated at the ground state in comparison to the experimental results. Again, the relative energy difference ($\Delta E_{II}*$) based on PNBA Form I shows a reasonable similarity.

Table 5-7: A summary of the transition energy of O K-edge based on DFT calculation and NEXAFS for both PNBA Form I and II.

	DFT			NEXAFS		
O 1s	Form I, eV	Form II, eV	ΔE*	Form I, eV	Form II, eV	ΔE*
O 1s (C=**O**) → 1π*	518.21	518.17	0.00	530.63	530.63	0.00
O 1s (N**O**$_2$) → 1π*	519.33	519.27	1.12	531.55	531.55	0.92
O 1s (**O**–H) → 1π*	520.47	520.46	2.26	533.24	533.26	2.61

5.4 Discussion

DFT calculations were made using a relatively basic computational method (B3LYP functional and 6-31G* basis set in Gaussian 09 for a single molecule (gas phase) at ground state) with the main purposes to visualise the molecular orbitals and assist with peaks assignments. Apart from a few differences, the DFT calculation did show reasonable consistency with the experimental results (solid phase) in terms of relative energy differences. There are two major discrepancies between the calculations and experiment, the O1s (C=**O**) binding energy and the transition of N1s (**N**O_2) → 3π*, which will be discussed.

The O 1s (C=**O**) binding energy (XPS) and O 1s (C=**O**) → 1π* transition energy (NEXAFS) showed an interesting combination of results. The O 1s (C=**O**) for PNBA Form I is at −0.3 eV lower than Form II but the transition energy for C 1s (C=**O**) → 1π* is the same at 530.6 eV, indicating that the energy level for 1π* is higher in PNBA Form I than Form II. Similar observation was found in aqueous solution of formic acid and formate [52]. Even though this is a small, subtle difference, it is explainable based on PNBA Form I having a 0.01 Å shorter C=O bond length than Form II. The oxygen at the carbonyl group (C=O) for PNBA Form I is more electronegative than Form II due to the shorter C=O bond length. Thus PNBA Form I has a lower binding energy, 532.9 eV, in comparison to PNBA Form II, 533.2 eV.

The difference of the hydrogen bond strength between the two polymorphs at the carboxylic acid dimer is in agreement with the experimental ATR-FTIR and DFT calculation for the ground state of the dimer structures. The ATR-FTIR of PNBA Form I showed a smaller wavenumber of 918 cm^{-1} compared to Form II at 928 cm^{-1} at the –OH out-of-plane bending (Chapter 3). PNBA Form I has a stronger hydrogen bond as the oxygen at the C=O is more electronegative. For the DFT calculation, the position of the oxygen and hydrogen atom (C=**O**···**H**–O) was fixed to maintain the differences in the bond length. The DFT calculation for the dimers shows −0.14 eV and −0.06 eV differences between the two polymorphs at the O 1s level and O 1s → LUMO (1π*). Figure 5.17 shows the experimental result is closer to the DFT calculation of the dimer than the monomer. This highlights the effects of hydrogen bonding on the electronic structure, and is in line with a calculation showing that with a shorter hydrogen bond length (H···O), the ionisation potential (IP) becomes larger as the hydrogen bonding is stronger [99].

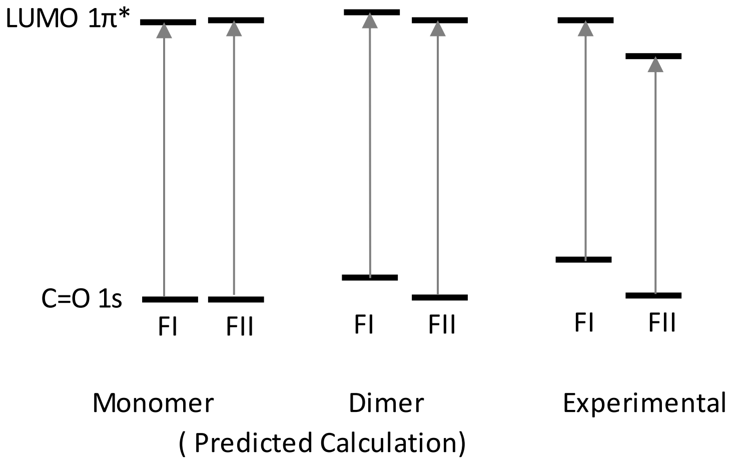


Figure 5.17: The sketch of the energy level at 1s and LUMO (1π*) for DFT calculation in monomer and dimer in comparison to the experimental result.

The predicted DFT calculation in the gas phase for a monomer system showed that PNBA Form II has a small additional peak for the transition of N 1s → 3π*, while experimentally both PNBA Form I and II showed an additional peak, but shifted to significantly higher photon energy than predicted by the DFT calculation. It seems likely that this discrepancy between theory and experiment stems from the use of a monomer model for MO analysis of the spectra. Likewise, for the dimer system, the DFT calculation revealed no transition of N 1s → 3π* for both polymorphic forms as shown in Figure 5.18. Additional work will have to address the origin of this transition.

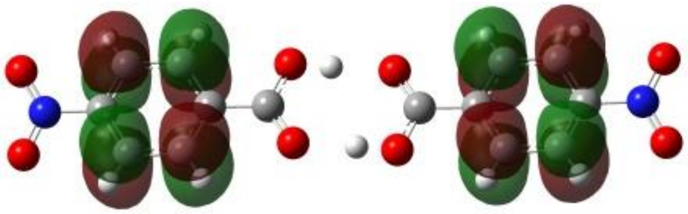


Figure 5.18: Visualisation of the third lowest unoccupied MO (LUMO+3) for a dimer system that is similar for PNBA Form I and II based on the DFT calculations.

A possible explanation for the discrepancy between theory and experiment is that both the monomer and dimer models neglect contributions from the π-π interactions in the crystal structure, which are expected to have the most profound effect on the outermost and diffuse π states, both in terms of shifting its energetic position relative to the ionisation potential and modifying the intensity of the transition. Furthermore, the choice of functional and basis set may affect the results, while the use of ground state calculations neglects the time dependent effects associated with the electronic transition from the initial to the final state. More systematic work is required here, using time-dependent DFT calculations for the solid state.

The DFT calculation may have predicted a small density at the 3π* because of the differences in the bond angle; PNBA Form II has a 2° rotation more at one side of N=O and this is similar to the DFT calculation, predicting only one side of N=O has the electron density as showed in Figure 5.19.

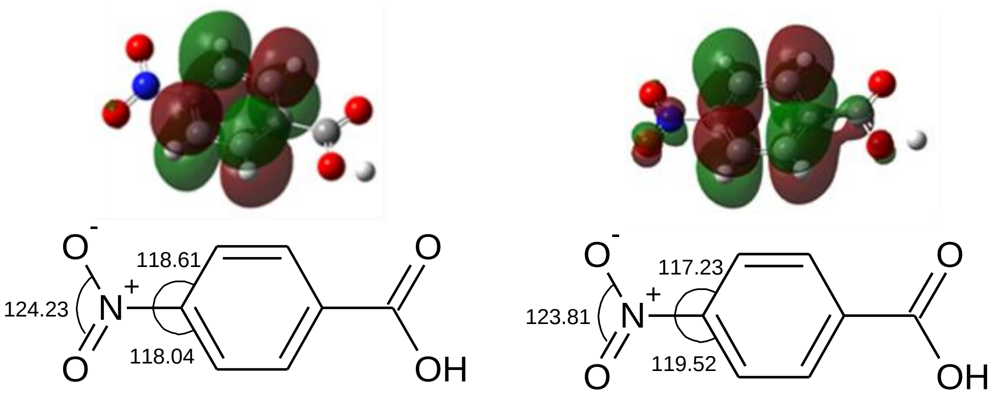


Figure 5.19: A potential correlation between the DFT calculation and the bond angle (°) for PNBA Form I (right) and PNBA Form II (left).

The most significant result here is that the differences between the polymorphic forms of PNBA Form I and II were minor and hardly detectable by XPS and NEXAFS, except for the small differences in the strength of hydrogen bonding at the carbonyl group, C=O. This agrees with the ATR-FTIR and Raman spectroscopy data, which also showed both forms have similar spectra except for the minor differences at the carboxylic acid group in relation to the hydrogen bonding. Therefore, these results suggest that the chemical and conformational state of PNBA in Forms I and II is negligibly different due to the very similar intermolecular interactions and local geometrical structure.

5.5 Conclusion

The influence of intermolecular interactions on the electronic structure based on DFT calculation in a monomer and dimer system correlates with experimental N 1s and O 1s (C=**<u>O</u>**) photoemission results. For a single molecule model, DFT predicts that there is an excitation of N 1s and C 1s (**<u>C</u>**-C) to the 3π*, which visibly diminishes in a dimer system and this was observed in the NEXAFS experimental result. Interestingly, the calculation for the dimer and the XPS/NEXAFS experimental results agree in that ionisation potential of O 1s (C=**<u>O</u>**) is lower while the LUMO (1π*) energy is higher for PNBA Form I.

The electronic structures of PNBA Form I and II extremely similar, which is in agreement with the ATR-FTIR vibrational spectra. A small difference found is PNBA Form I having a weaker hydrogen bond at the carboxylic acid dimer due to a longer hydrogen bond length. This is evident in the XPS O 1s (C=**<u>O</u>**) spectrum for PNBA Form I through a −0.3 eV shift and in ATR-FTIR the stretching of OH out-of-wag-plane (dimerization) is at a wavenumber −10 cm^{-1} lower than for Form II. This chemical shift correlates with PNBA Form I having a shorter C=O bond length at 1.23 Å and a longer hydrogen bond length (C=**<u>O···H</u>**) at 1.6 Å compared to Form II at 1.25 Å and 1.5 Å.

The poor agreement between experiment and theory for what is believed to be the N 1s → 3π* region of the spectra needs further investigation including the influence of π-π interactions in the solid state, the time dependent nature of the transitions and the possible influence of choosing different functionals and basis sets.

6. SPECTROSCOPIC STUDIES OF SUBSTITUENT EFFECTS IN BENZOIC ACID DERIVATIVES IN THE SOLID STATE

Abstract: A comparative examination of the substituent effects on the electronic structures of *p*-nitrobenzoic acid (PNBA), *p*-aminobenzoic acid (PABA), and benzoic acid (BA) has been carried out by X-ray photoelectron (XP), and near-edge X-ray absorption fine structure (NEXAFS) spectroscopy. The variations can be related to infrared and Raman vibrational spectra. PNBA has a strong electron withdrawing nitro group in para position to the carboxylic acid moiety, while PABA has a strong electron donating amine group. The IR and Raman spectra showed the carbonyl stretching frequency for PNBA is the highest, followed by BA and PABA. In line with this, XP and NEXAFS spectra indicate a reduction in electron density at the carbonyl group through conjugation across the aromatic ring, and hence a strengthening of the bond, resulting in a higher stretching frequency. Throughout the molecule, PNBA has a lower orbital energy for all 1s levels as well as a more strongly bound lowest unoccupied molecular orbital (1 π*) compared to BA and PABA.

6.1 Introduction

Experimental studies of substituents effects on the electronic structures of organic systems have been performed for many years, including nuclear magnetic resonance, vibrational, and core electron spectroscopies [107–114]. Electron withdrawing groups (EWG) and electron donating group (EDG) reduces and increase the electron density in conjugated and aromatic π systems. Understanding the influence of local molecular environment on the electronic structure through studying intermolecular interactions could provide a better insight on many important functions of organic molecules, for example molecular recognition [50,51]. In this context, carboxylic acid (−COOH) groups have been the focus of extensive studies in understanding molecular recognition *via* hydrogen bonding [41,50,51,115–121].

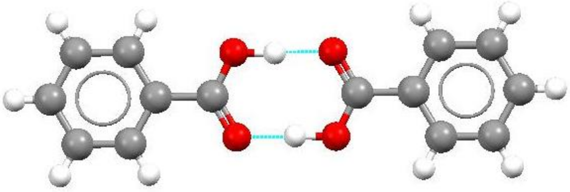
Infrared (IR) radiation is commonly used to study carboxylic acid systems, exploiting the variations in vibrational properties in the a carbonyl (C=O) and hydroxyl (OH) groups [36,81,115]. Vibrational bands are dependent on atomic mass, bond length and strength [32,60]. Raman spectroscopy complements IR in detecting the vibrational properties of the carbonyl band. The IR and Raman bands for the carbonyl asymmetric and symmetric stretch vibrations occur in the region of 1720 – 1680 cm^{-1} and 1680 – 1640 cm^{-1} respectively [122]. IR spectroscopy has become the most prominent technique for detecting carboxylic acid dimers, through an OH-out-of-plane bending frequency in at wavenumbers from 960 to 875 cm^{-1} [36,81,122], and with a difference between the carbonyl stretching frequencies of less than 30 cm^{-1} [81].

X-ray photoelectron spectroscopy (XPS) and near-edge x-ray absorption fine structure (NEXAFS) are emerging core level spectroscopy techniques that provide additional understanding about the carboxylic acid moieties [50,51,113,114]. XPS and NEXAFS probe the electronic structure of the molecules, specifically the core electron (1s) and the unoccupied states below the ionisation potential. In this chapter, the influence of the substituent functional group of three carboxylic acid systems (benzoic acid (BA), *p*-aminobenzoic acid (PABA) and *p*-nitrobenzoic acid (PNBA)) are studied by combining vibrational and core level spectroscopies.

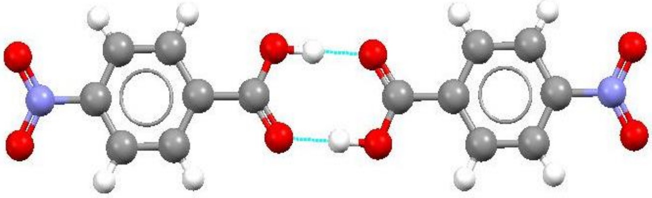
In organic chemistry, the Hammett equation is often used to analyse and predict the influence of *meta* and *para* substituents on the behaviour of benzoic acid systems, relating observable variations in reaction rates and equilibrium constants for reactions with free energies. Relative to BA, PNBA has the nitro ($-NO_2$) group, which is strongly electron withdrawing (Hammett value of +0.78 [123]), while PABA has the amine ($-NH_2$) substituent, which is strongly electron donating (Hammett value of −0.66 [123]). PNBA is the most acidic of the three, followed by BA and PABA with pK_a values of 3.43, 4.20, and 4.65 [124].

The IR vibrational spectra provide carboxylic acid dimer characterisation through the carbonyl stretching and the hydroxyl bending vibration frequencies. XPS and NEXAFS provide information on the electronic structure of each chemical moiety – through the characteristic carbon 1s excitations for $\underline{\mathbf{C}}=C$, $\underline{\mathbf{C}}=O$, $\underline{\mathbf{C}}-N$, the oxygen 1s exciations for $C=\underline{\mathbf{O}}$, $\underline{\mathbf{O}}-H$, $N\underline{\mathbf{O}}_2$ and the nitrogen 1s excitations for $\underline{\mathbf{N}}O_2$ and $\underline{\mathbf{N}}H_2$. In the following it will be examined how the effects of substituents can be related to the information obtained by the three techniques, and how XPS and NEXAFS can be used to obtain further information on specific vibrational features that are widely employed to characterise carboxylic acid dimers.

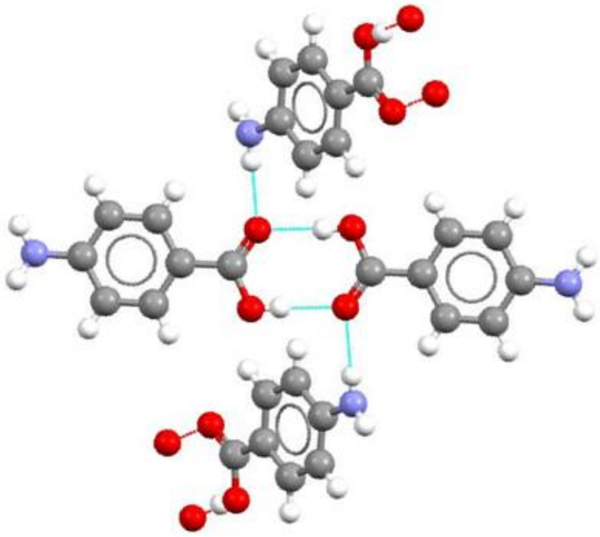
As discussed in chapter 5, PNBA has two polymorphs with very similar crystal structures (PNBA Form I and Form II), while PABA has two polymorphs with quite different local bonding in each crystal structures (α−PABA and β−PABA). PNBA Form I, α−PABA and benzoic acid have the same synthonic building unit in the crystal lattice, namely a carboxylic acid dimer as shown in Figure 6.1. A selected Cambridge Structural Database (CSD) search provides directly comparable crystal structures for BA, PNBA Form I and α−PABA, which were all determined using X-ray diffraction at a temperature around 120 K. The bond lengths for C=O, O−H, and the C=O···H−O or O···O are summarised in Table 6-1. α−PABA has the shortest O−H at 0.8 Å follow by PNBA Form I at 1.0 Å and benzoic acid at 1.1 Å. PNBA Form I has the longest O···O bond length at 2.65 Å, α−PABA at 2.64 Å and BA at 2.61 Å. α−PABA has an additional hydrogen bond length C=O···H−N at 2.1 Å Figure 6.1 (c). The bond length of O···O is preferred as it gave better correlation than the O−H bond length [81].



(a) Benzoic acid (CSD Reference: BENZAC12)



(b) *p*-Nitrobenzoic acid Form I (CSD Reference: NBZOAC04)



(c) α−*p*-Aminobenzoic acid (CSD Reference: AMBNAC07)

Figure 6.1: The crystal structures of benzoic acid, *p*-nitrobenzoic acid and *p*-aminobenzoic acid obtain from the CSD.

Table 6-1: A summary of the C=O, O−H and hydrogen bond lengths for benzoic acid (BA), *p*-aminobenzoic acid (α−PABA) and *p*-nitrobenzoic acid (PNBA Form I).

	BA, Å	α−PABA, Å	PNBA Form I, Å
C=O	1.252	1.239	1.231
O−H	1.080	0.820	0.979
C=<u>O···H−O</u>	2.610	2.641	2.648

6.2 Methodology

A short summary on the methodologies are given here while the full details are given in Chapter 2. PNBA Form I was re-crystallised in acetonitrile at room temperature while the commercially available BA and α−PABA were used as received (Section 2.1). The infrared red spectra were taken using ATR-FTIR with a Thermo-Scientific instrument as described in chapter 2. The spectra was obtained from 32 scans at a resolution of 4 cm^{-1}, in the range of 650 to 4000 cm^{-1} (Section 2.8). The Raman spectra were taken using the Jobin-Yvon Raman microscope described in Chapter 2. The microscope objective used was 100x magnification and the spectra were taken in the range from 10 to 4000 cm^{-1} (Section 2.9). X-ray photoelectron spectra for PNBA Form I and α−PABA were recorded using the Kratos Axis Ultra instrument. High resolution spectra for the carbon, nitrogen and oxygen were taken with 0.1 eV steps (Section 2.10). Peak fittings for α−PABA have been published previously[51]. NEXAFS data for α−PABA and PNBA Form I were performed at the National Synchrotron Light Source (NSLS), New York. Partial electron yield (PEY) for the carbon, nitrogen, and oxygen K-edge were collected at energy resolution of 0.1 to 0.15 eV (Section 2.11). The peak fittings were published previously[51]. Molecular orbital energy calculations for the ground state were performed with Gaussian 09 for a monomer molecule of BA, PNBA and PABA. Structure optimisation was performed before the energy calculations were made, and the B3LYP/6-31G (d) functional and basis set were used (Section 2.12).

6.3 Results

6.3.1 Vibrational Spectroscopy (IR and Raman)

Figure 6.2 shows FTIR results in the region of 800 to 1800 cm^{-1}. The C=O asymmetric stretch and O−H out-of-plane bending frequency are the focus of interest. PNBA has the highest C=O stretching frequency at 1684 cm^{-1}, followed by BA at 1675 cm^{-1}, and PABA at 1659 cm^{-1}. The O−H out of plane bend for PNBA is at 928 cm^{-1}, for BA at 933 cm^{-1} and for PABA at 927 cm^{-1}.

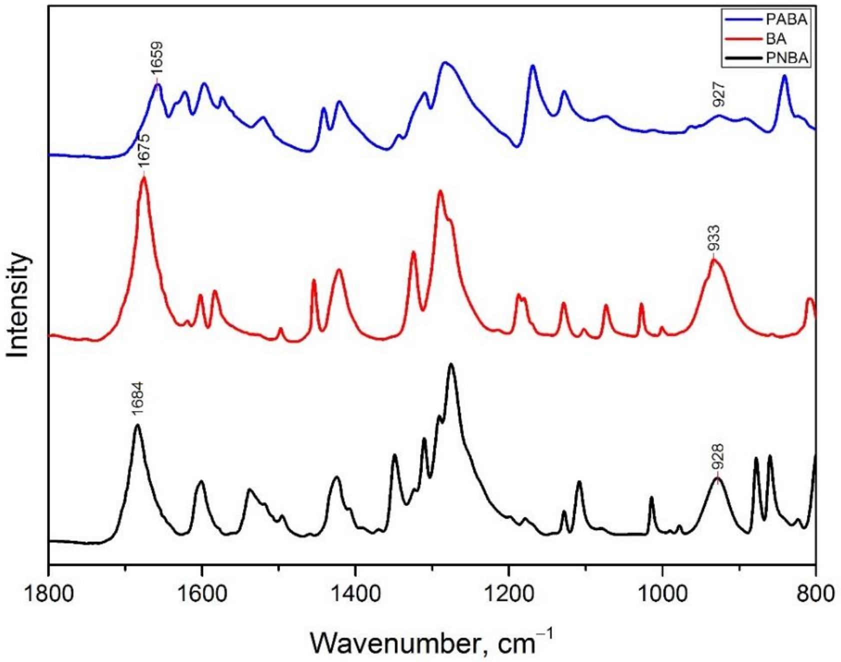


Figure 6.2: Infrared spectra of PNBA (black), BA (red) and PABA (blue) in the 800-1800 cm^{-1} region, with labelled C=O stretching and O-H out-of-plane peak positions.

Figure 6.3 shows the Raman spectra in the region of 1000 to 1800 cm^{-1}. The carbonyl symmetric stretch is weaker; PABA has the weakest signal, a small shoulder at 1622 cm^{-1}, BA at 1634 cm^{-1} and PNBA at 1640 cm^{-1}.

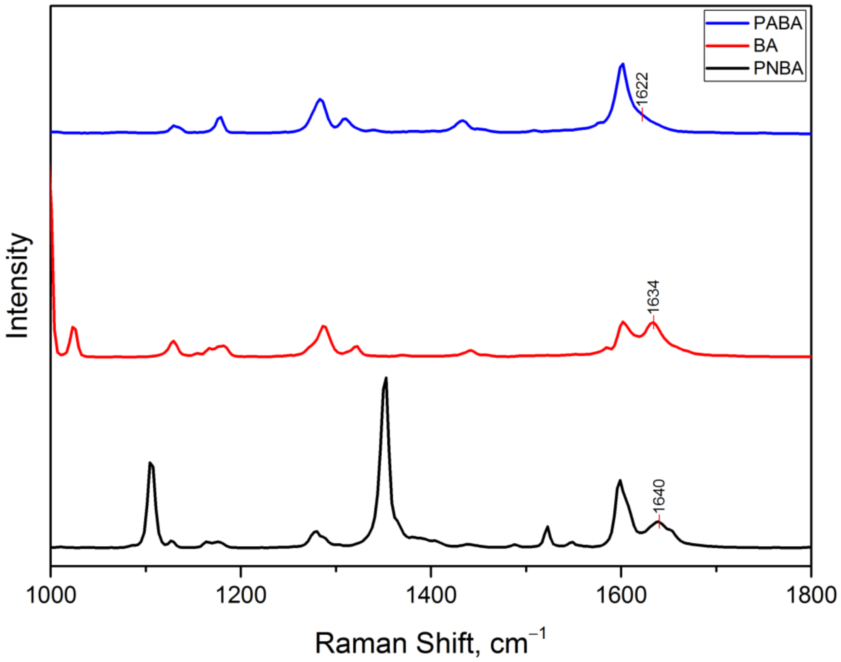


Figure 6.3: Raman spectra of PNBA acid (black), BA (red) and PABA (blue) in the 1000-1800 cm^{-1} region, with labelled C=O stretching peak position. .

6.3.2 Core Level Spectroscopies

The ionisation potential energy and the transition energy from the core electron to the lowest unoccupied molecular orbital (LUMO) are experimentally obtained from the XPS and NEXAFS for PNBA and PABA. There are currently no experimental data for BA, because its high vapour pressure prevents measurements under ultra-high vacuum conditions, so only a predictive DFT calculation was made.

6.3.2.1 Carbon 1s (C 1s) spectra

The C 1s XPS for PNBA Form I indicates a higher ionisation potential than for α−PABA, as shown in Figure 6.4. The fitted peaks of **C**=C, **C**−N and **C**=O are shown and summarised in Figure 6.4 and Table 6-2. Relative to PABA, the chemical shift for PNBA C 1s **C**=O, C 1s **C**−N and C 1s **C**=C are +1.3, +1.1 and +1.6 eV.

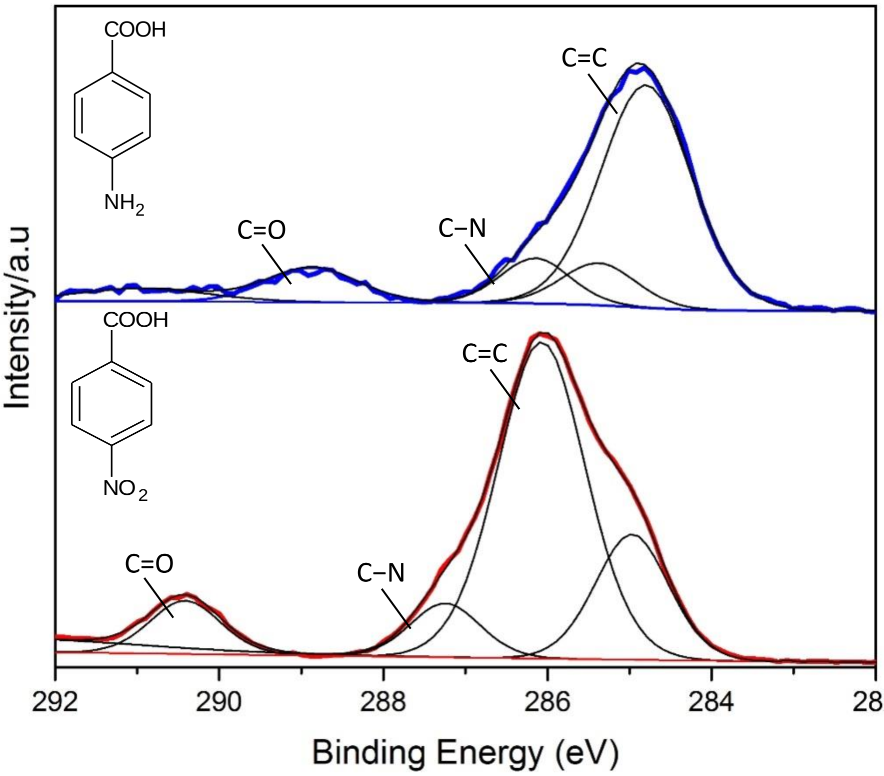


Figure 6.4: Experimental C 1s XPS of PNBA Form I (red), and α-PABA (blue) and the fitted component peaks, showing the shifts in binding energy (IP) for different chemical species.

The C K-edge spectra for PABA and PNBA are significantly different; PNBA has a sharp doublet at 284.8 eV and 285.7 eV, while PABA has a sharp peak at 285.3 eV and a shoulder at 286.63 eV. These peaks are the transition energy for the C 1s **C**=C → 1π* and C 1s **C**−N → 1π*. Both compounds have the same transition energy for the C 1s **C**=O → 1π* at 288.8

eV. Relative to PABA, the chemical shift for PNBA C 1s **C**=C → 1π* and C 1s **C**–N → 1π* are −0.5 eV and −0.9 eV.

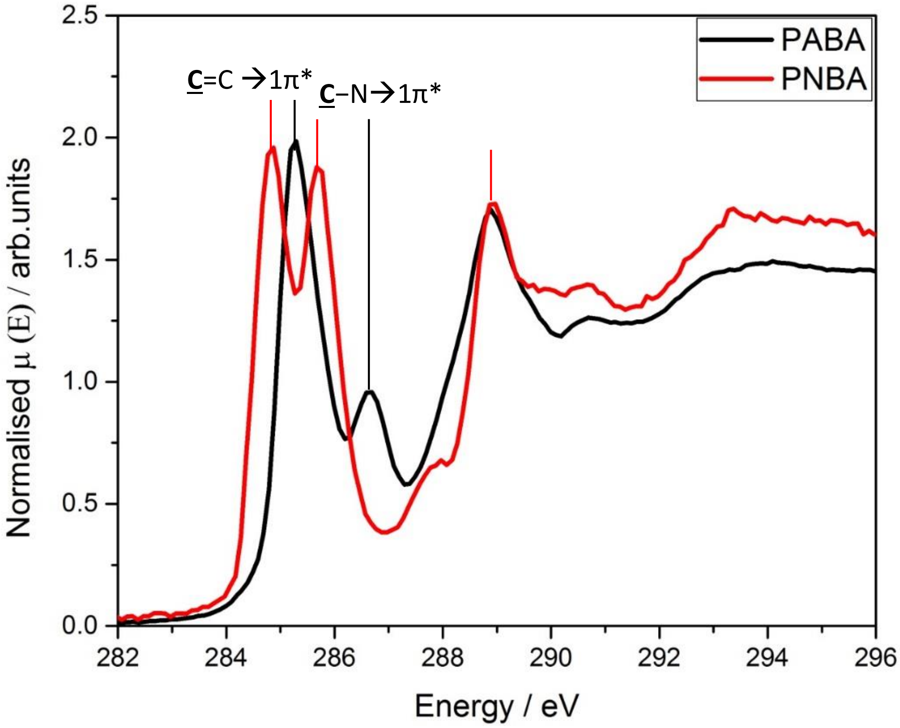


Figure 6.5: Normalised spectra of experimental carbon K-edge NEXAFS of PNBA Form I (red) and α–PABA (black).

Table 6-2 summarises the ionisation potential of C 1s from the XPS data, the transition energy from C 1s to the LUMO (1π*) and the difference in energy between PNBA and PABA. The main observation is the ionisation potential for PNBA is higher than PABA while the transition energy to the 1π* is lower than PABA.

Table 6-2: A summary of the ionisation potential C 1s energy from XPS and the transition energy from C 1s to the LUMO (1π*) from NEXAFS.

	PNBA Form I	α-PABA	$\Delta BE_{PNBA-PABA}$
IP (C=C)	286.07	284.80	+1.27
IP (C–N)	287.25	286.15	+1.10
IP (C=O)	290.42	288.86	+1.56
C 1s (C=C) → 1π*	284.81	285.28	−0.47
C 1s (C–N) → 1π*	285.71	286.63	−0.92
C 1s (C=O) → 1π*	288.79	288.85	−0.06

The predicted ionisation potential (IP) for the C 1s **C**=O, C 1s **C**–N and C 1s **C**=C and the LUMO (1π*) energy level for PNBA, BA and PABA are illustrated in Figure 6.6. The 1s and 1π* energy level is the lowest for PNBA, followed by BA and PABA. Figure 6.7 shows illustrations of the molecular orbitals for the LUMO (1π*) states in the three compounds, indicating that they are similar. Relative to PABA, the calculated chemical shift of the C 1s core level in PNBA are for C 1s **C**=O, C 1s **C**–N and C 1s **C**=C +1.0 eV, +0.7 eV and +1.0 eV, respectively, while the NEXAFS transitions C 1s **C**=O → 1π*, C 1s **C**–N → 1π* and C 1s **C**=C → 1π* are shifted by −1.2 eV, −1.4 eV and −1.1 eV. These results differ somewhat from the experimental values due to the calculation being made for a gas phase (monomer) molecule in the ground state, but they are nonetheless quite similar to the experimental results.

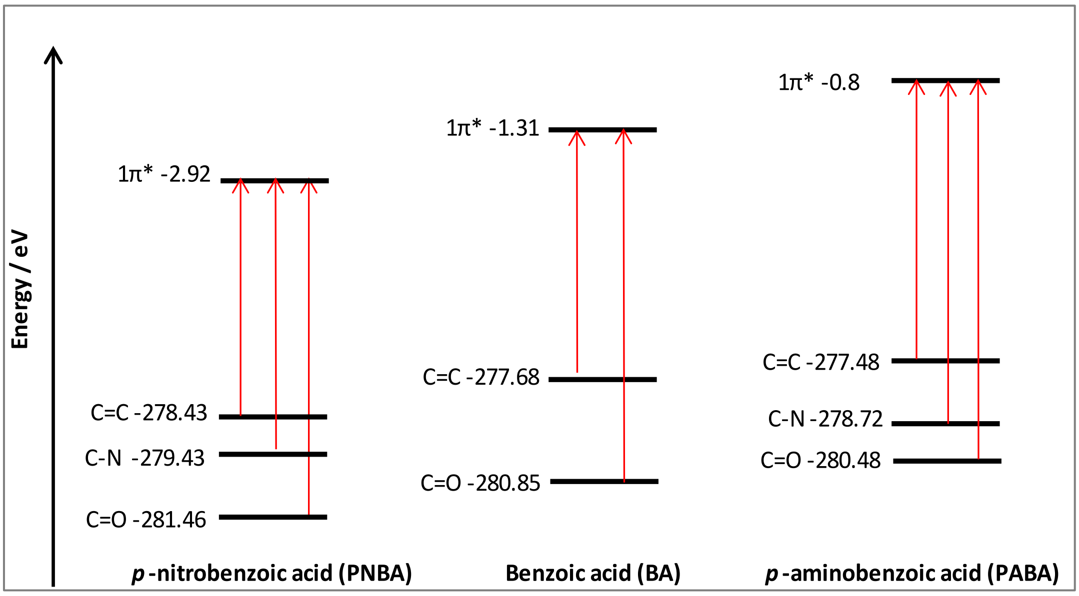


Figure 6.6: Schematic of the calculated transition energy from carbon core 1s level to the LUMO (1π*) for PNBA, BA and PABA.

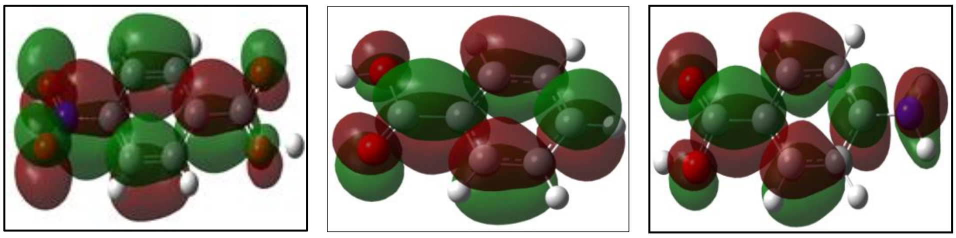


Figure 6.7: The predicted molecular orbital representation for the LUMO (1π*) for PNBA (left), BA (middle) and PABA (left).

6.3.2.2 Nitrogen spectra

The experimental N 1s **N**O_2 XPS spectrum for PNBA Form I indicates a much higher ionisation potential at 406.7 eV than for the α−PABA N 1s **N**H_2 moiety, which appears at 399.6 eV in the spectrum (Figure 6.8).

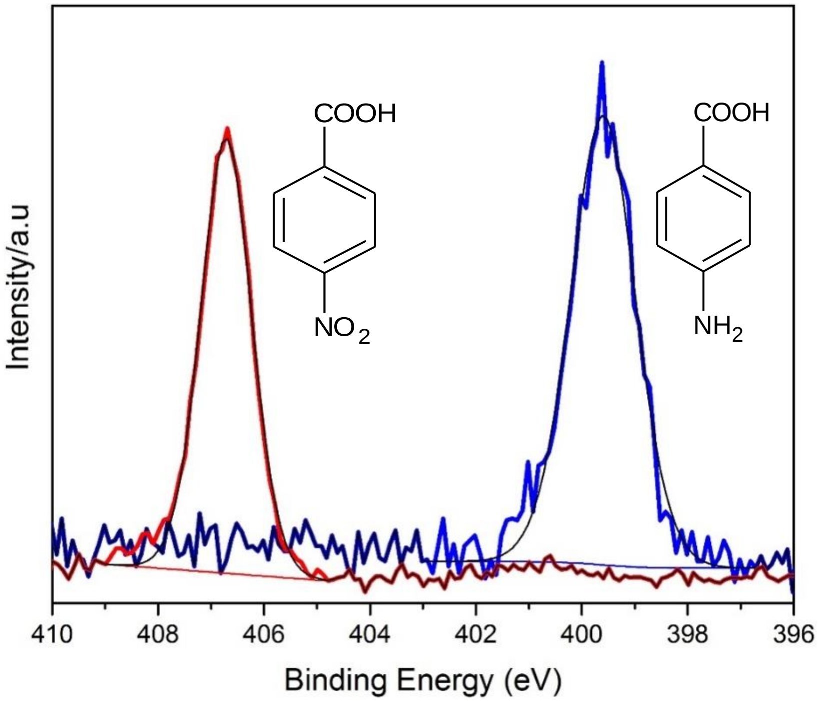


Figure 6.8: Experimental N 1s XPS of PNBA Form I (red), and α-PABA (blue) and the fitted components peaks, showing the shifts in binding energy (IP).

The N K-edge spectra for α−PABA and PNBA Form I are also significantly different, as shown in Figure 6.9; α−PABA has a lower intensity peak for the N 1s **N**$H_2 \rightarrow 1\pi^*$ at 400 eV while PNBA has a sharp and high intensity peak for the N 1s **N**$O_2 \rightarrow 1\pi^*$ at 404 eV.

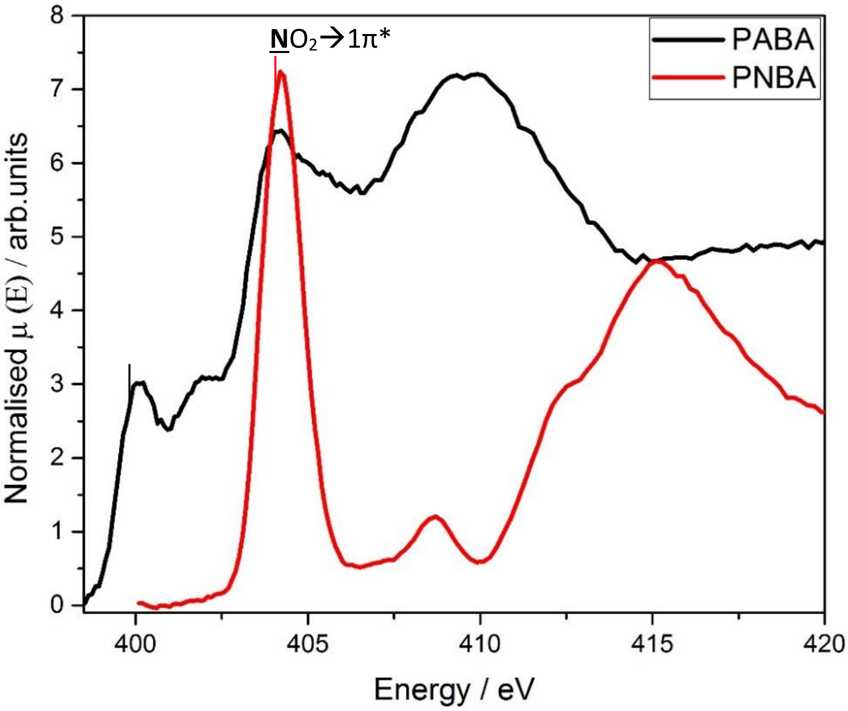


Figure 6.9: Normalised spectra of experimental nitrogen K-edge NEXAFS of PNBA Form I (red) and α–PABA (black).

Table 6-3 summarise the ionisation potential of N 1s from the XPS data, the transition energy N 1s to the LUMO (1π*) and the difference in energy between PNBA and PABA. The main observation is both the ionisation potential and the transition energy to the 1π* is higher for PNBA.

Table 6-3: A summary of the ionisation potential N 1s energy from XPS and the transition energy from N 1s to the LUMO (1π*) from NEXAFS.

	PNBA Form I	α-PABA	$\Delta BE_{PNBA-PABA}$
IP (<u>N</u>H_2)		399.60	+7.12
IP (<u>N</u>O_2)	406.72		
N 1s (<u>N</u>H_2) → 1π*		400.00	+4.20
N 1s (<u>N</u>O_2) → 1π*	404.20		

The DFT-predicted ionisation potential (IP) for the N 1s (<u>N</u>H_2) is lower than N 1s (<u>N</u>O_2) while the LUMO (1π*) energy level is higher for PABA as shown in Figure 6.10. Relative to PABA, the calculated chemical shift for PNBA N 1s <u>N</u>O_2 and N 1s <u>N</u>O_2 → 1π* are +6.0 eV and +3.9 eV which are in good correlation with the experimental data as in Table 6-3.

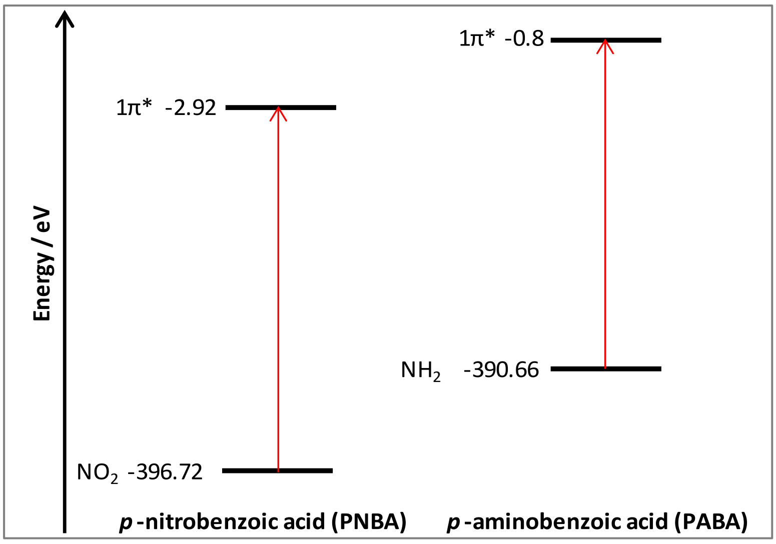


Figure 6.10: Schematic of the calculated transition energy from nitrogen core 1s level to the LUMO (1π*) for PNBA and PABA.

6.3.2.3 Oxygen spectra

The experimental O 1s XPS for PNBA Form I indicates a higher ionisation potential than for α−PABA, as shown in Figure 6.11. The fitted peaks of C=**O**, **O**−H and additional N**O**$_2$ (in PNBA) are shown and summarised in Figure 6.11 and Table 6-4. Relative to PABA, the chemical shift for PNBA O 1s C=**O** and O 1s **O**−H are +0.7 eV and +1.1 eV.

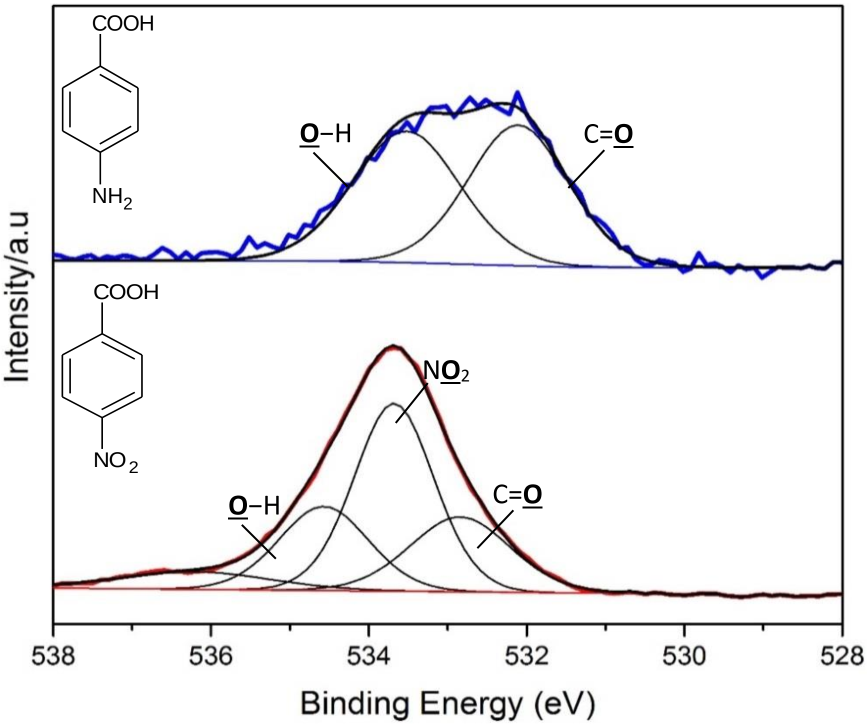


Figure 6.11: Experimental O 1s XPS of PNBA Form I (red), and α-PABA (blue) and the fitted components peaks, showing the shifts in binding energy (IP) for different chemical species.

The O K-edge spectra for PABA and PNBA as in Figure 6.12 shows similar features with PABA at higher transition energy for O 1s C=**O**→ 1π* and O 1s **O**–H → 1π* at 532.1 eV and 533.5 eV. Relative to PABA, the chemical shift for PNBA O 1s C=**O**→ 1π* and O 1s **O**–H → 1π* are −1.5 eV and −1.2 eV. PNBA has an additional transition energy for O 1s N**O**$_2$→ 1π* at 531.6 eV.

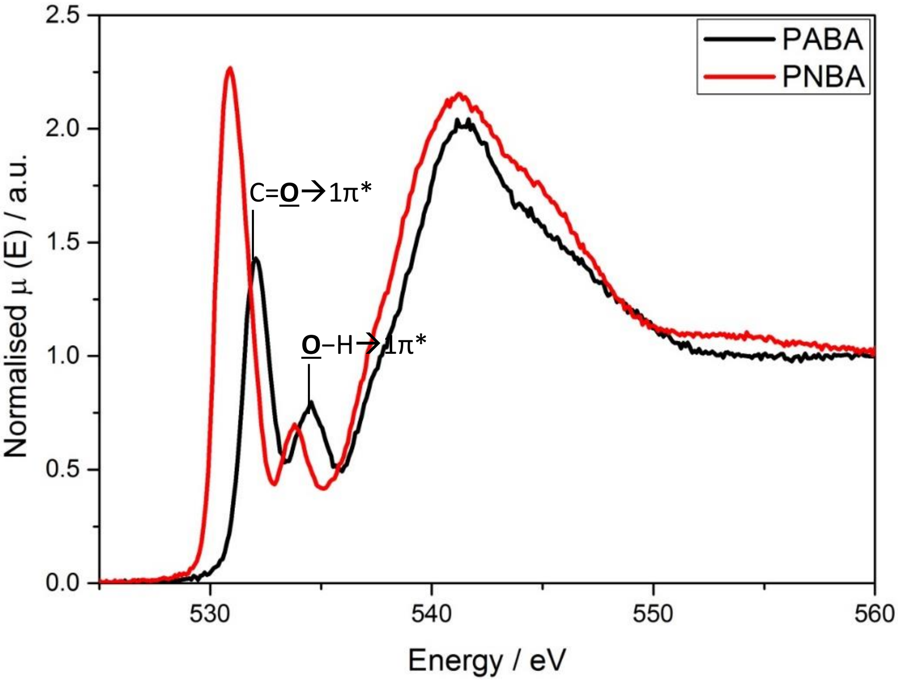


Figure 6.12: Normalised spectra of experimental nitrogen K-edge NEXAFS of PNBA Form I (red) and α–PABA (black).

Table 6-4 summarises the ionisation potentials obtained via O 1s XPS alongside the transition energies from O 1s to the LUMO (1π*), and the associated differences in energy between PNBA and PABA. The main observation is that the ionisation potential for PNBA is higher than for PABA, while the transition energy to the 1π* is lower in PNBA relative to PABA.

Table 6-4: A summary of the ionisation potential O 1s energy from XPS and the transition energy from N 1s to the LUMO (1π*) from NEXAFS.

	PNBA Form I	α-PABA	$\Delta BE_{PNBA-PABA}$
IP (C=**O**)	532.85	532.11	+0.74
IP (N**O**$_2$)	533.68	–	–
IP (**O**-H)	534.57	533.52	+1.05
O 1s (C=**O**) → 1π*	530.63	532.08	−1.45
O 1s (N**O**$_2$) → 1π*	531.55	–	–
O 1s (**O**-H) → 1π*	533.26	534.41	−1.15

The predicted ionisation potentials (IP) for the O 1s C=**O**, O 1s N**O**$_2$ and O 1s **O**–H and the LUMO (1π*) energy level for PNBA, BA and PABA are summarised in Figure 6.13. The 1s and 1π* energy level are lowest for PNBA, followed by BA and then PABA. Relative to PABA, the calculated chemical shift for PNBA O 1s C=**O** and O 1s **O**–H are +1.0 eV and +0.9 eV while the O 1s C=**O**→ 1π* and O 1s **O**–H → 1π* −1.1 eV and −1.2 eV. The relative energy differences are reasonably similar to the experimental results as in Table 6-4.

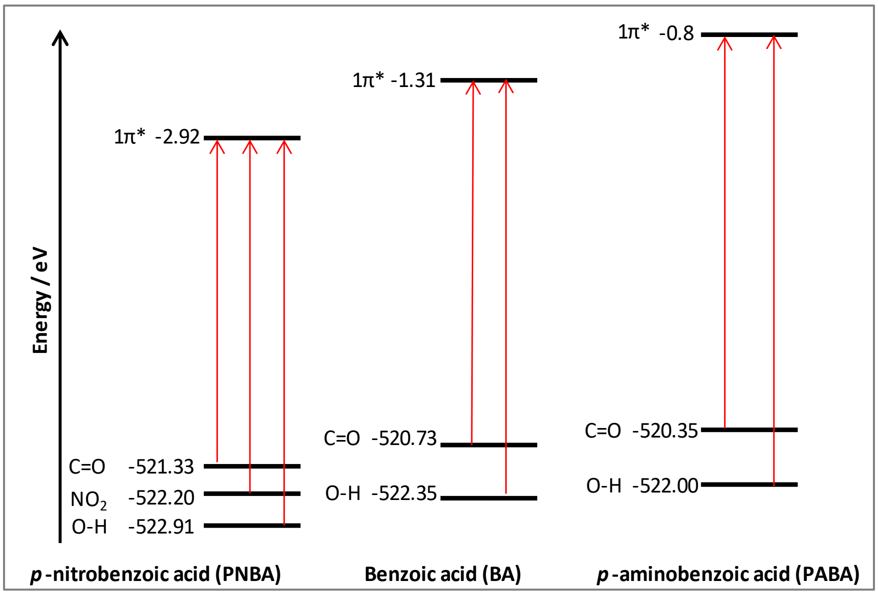


Figure 6.13: Schematic of the calculated transition energy from oxygen core 1s level to the LUMO (1π*) for PNBA, BA and PABA.

6.4 Discussion

IR and Raman showed the carbonyl asymmetric and symmetric stretching frequency is the highest for PNBA (1684, 1640 cm^{-1}), followed by BA (1675, 1634 cm^{-1}), and the lowest for PABA (1659, 1622 cm^{-1}). These shifts are not correlated neither to the C=O nor the O···O bond length (Table 6-1); BA has the longest C=O and the shortest O···O bond length. The O–H out-of-plane bending frequency did not show similar trend. The frequency for PNBA and PABA are the same at 928 cm^{-1} which correlates to the similarity in O···O bond length, 2.64 Å while BA which has the shortest O···O bond length, 2.61 Å, had this bending frequency at 933 cm^{-1}.

The effects of electron withdrawing group (EWG) and electron donating group (EDG) on the carbonyl stretching frequency is significant, but less so on the O–H out-of-plane bending frequency. The −NO$_2$ group deactivates and reduces the electron density at the carboxylic acid (−COOH) while the amine −NH$_2$ group activates and increases the electron density as illustrated in Figure 6.14. The shift in wavenumber is influenced by the C=O bond strength. The C=O resonance structure is stronger with lower electron distribution and weakens with higher electron density, suggesting the stretch of the carbonyl bond for PNBA should appear at higher energy than in PABA. Considering the dynamic dipole moment selection rule for IR vibrational spectroscopy, the weak absorption intensity for α−PABA appears to suggest that there is less change in dipole moment during the vibration, possibly due reduced oxygen atom electron density as a result of hydrogen bonding C=**O···H**−N.

The advantage of using XPS and NEXAFS for characterisation of the molecules is that the effect of substituents is observed at several of the chemical moieties in the conjugated system, i.e., C=O, C=C, C−N, O−H, NO_2, NH_2. The overall C 1s, N 1s and O 1s spectra shows that PNBA has generally higher ionisation potentials throughout than PABA due to the strongly electron-withdrawing nature of the nitro group. Relative to PABA, the chemical shift for C 1s **C**=C, C 1s **C**−N, C 1s **C**=O, O 1s C=**O**, O 1s **O**−H in PNBA are +1.3, +1.1, +1.6 eV, +0.7 eV and +1.1 eV, illustrating very convincingly that the electron density of all atoms across the conjugated system is affected by the presence of the nitro group, including the strong reduction of the electron density at the carboxylic acid moiety, as shown in Figure 6.14 (left). As a result of the reduced electron density, the core electrons experiencing stronger attraction by the nucleus, thus more energy is needed to remove the core electron. For PABA, the electron density throughout the molecule is enhanced through charge transfer from the amine group, as shown in Figure 6.14 (right). As the electron density is bigger, the core electron is held loosely to the nucleus thus a lower energy is needed to remove the core electron.

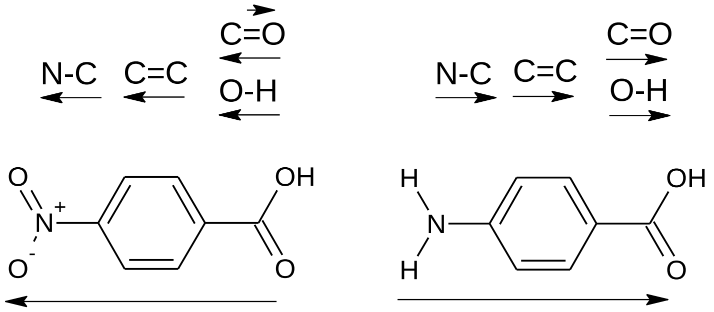


Figure 6.14: An illustration of the electrons being pulled towards the nitro group (left) and electrons being donated towards the carboxylic acid group (right).

The O 1s C=**O** has the smallest relative chemical shift between PNBA and PABA, likely due to the additional influence of the intermolecular hydrogen bond in the crystal structure. PABA and PNBA have similar C=**O···H**−O interaction, but α-PABA has an additional

C=**O···H**−N motif. The multiple interactions at this oxygen moiety thus appear to reduce the relative ionisation potential difference for O 1s C=**O** to +0.7 eV, compare to the differences above +1.0 eV for all other atoms. In addition, the carbonyl itself is a EWG, thus for the C=O in PNBA, the electron density specifically at carbon may be pulled in both directions, resulting a much smaller electron density therefore a higher ionisation potential difference of +1.6 eV relative to PABA.

The NEXAFS results on the excitation of core electron to the lowest unoccupied molecular orbital (1π*) also shows a strong substituent effect for all chemical moieties. Relative to PABA, the chemical shift for PNBA C 1s **C**=O → 1π*, C 1s **C**−N → 1π*, O 1s C=**O**→ 1π* and O 1s **O**−H → 1π* are −0.5 eV, −0.9 eV, −1.5 eV and −1.2 eV. Interestingly, the relative chemical shift is in the opposite direction from the ionisation potential. This is because the LUMO (1π*) energy level is the lowest for PNBA, thus a smaller energy is needed to excite the 1s to the 1π* as illustrated in Figure 6.6, Figure 6.10 and Figure 6.13.

XPS and NEXAFS did not reveal characteristics of carboxylic acid dimer based on the relative difference of ionisation potential C 1s and O 1s and the transition energy for C 1s → 1π* and O 1s → 1π*. A comparison was made between the dimer structure of PNBA Form I and α−PABA to the catemer structure of β−PABA [50], as summarised in Table 6-5 and Table 6-6. A 1 eV difference was previously reported for the chemical shift of C 1s **C**=O relative to the **C**=C and O 1s **O**−H relative to the C=**O** between a calculated monomer and solid phase of α−PABA [51]. A similar observation was made by calculation and experimentally for acetic acid as a monomer and in a cluster phase (dimer), with a 1 eV relative difference between the transition energy, O 1s **O**−H →1π* to O 1s C=**O** →1π* [100]. This highlight the subtle difference between the electronic structures of a dimer and catemer, where both structures are hydrogen bonded, in comparison to a monomer and a dimer, where no such hydrogen bonding takes place.

Table 6-5: Experimental XPS core level shift of β−PABA (catemer) to α−PABA and PNBA Form 1 (dimers) for carbon relative to the C=C and oxygen relative to the C=O.

	Core level 1s binding energy/eV			Relative shift to C=C/C=O	
C 1s	C=C	C−N	C=O	C−N	C=O
β−PABA	284.8	286.0	288.7	+1.2	+3.9
α−PABA	284.8	286.2	288.9	+1.3	+4.1
PNBA Form I	286.1	287.3	290.4	+1.2	+4.4
O 1s	C=O	O−H		O−H	
β−PABA	531.7	533.2		+1.5	
α−PABA	532.1	533.5		+1.4	
PNBA Form I	532.9	534.6		+1.7	

Table 6-6: Experimental NEXAFS transition energy of C 1s and O 1s of β−PABA (catemer) to α−PABA and PNBA Form 1 (dimers) for carbon relative to the C=C and oxygen relative to the C=O.

	1s → 1π* transition energy/eV			Relative shift to C=C/C=O	
C 1s	**C=C**	C−N	COOH	C−N	COOH
β−PABA	285.3	286.6	288.9	+1.4	+3.6
α−PABA	285.3	286.6	288.9	+1.4	+3.6
PNBA Form I	284.8	285.7	288.8	+0.9	+4.0
O 1s	**C=O**	**O−H**		O−H	
β−PABA	531.7	534.3		+2.6	
α−PABA	532.1	534.4		+2.3	
PNBA Form I	530.6	533.3		+2.6	

IR and Raman additionally characterise the carboxylic acid dimer and catemer. The common characteristics of a carboxylic acid dimer based on IR and Raman spectroscopy is the IR O−H out-of-plane peak in between 920 to 950 cm^{-1} and the difference between the carbonyl stretching peak in IR and Raman is above 30 cm^{-1}. The IR shows O−H absorptions at 927, 933 and 928 cm^{-1} while the difference between the carbonyl peak in IR and Raman are 39, 42 and 44 cm^{-1} respectively for PABA, BA and PNBA. β−PABA is a non-carboxylic acid dimer thus did not exhibit the O−H signal in the region of 920 to 950 cm^{-1}and the difference between the carbonyl peak in IR and Raman is 9 cm^{-1}.

6.5 Conclusion

It has been shown that combined IR, Raman, XPS and NEXAFS give a comprehensive view of the relationship between intramolecular electron density variations due to substituent changes, local bonding changes due to intermolecular hydrogen bonding, and vibrational properties, especially at the carboxylic acid moiety. IR frequency and Raman shift for the carbonyl stretching are highest for PABA, followed by BA and PNBA. As the electron density reduces for PNBA, a reduction in electron density at the carbonyl group is evident, through XPS showing that C 1s $\underline{\mathbf{C}}$=O, O 1s C=$\underline{\mathbf{O}}$ and O 1s $\underline{\mathbf{O}}$–H in PNBA have the highest ionisation potential compared to BA and PABA. At the same time, the LUMO energy is differentially lowered, shifting the transitions C 1s $\underline{\mathbf{C}}$=O→ 1π*, O 1s C=$\underline{\mathbf{O}}$→1π* and O 1s $\underline{\mathbf{O}}$–H→ 1π* for PNBA relative to BA and PABA.

These result show how XPS and NEXAFS can be used to experimentally probe the effects of electron withdrawing and donating groups on all chemical moieties in conjugated systems, with C=C, C=O, C–N, O–H examined via the combination of carbon, oxygen and nitrogen core level spectroscopy. Figure 6.15 summarises the 1s and 1π* energy levels, showing how PNBA has the lowest 1s and 1π* energy followed by BA and PABA. It is evident how the electron withdrawing nitro group lowers the LUMO energy relative to BA, while the electron donating group amine group increased it relative to BA.

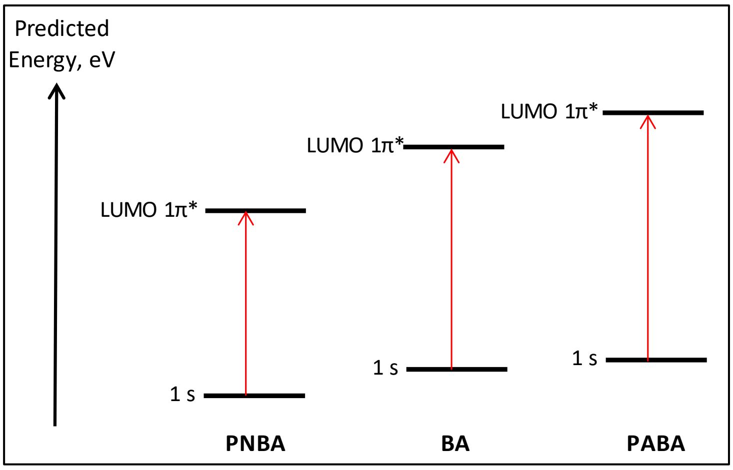


Figure 6.15: Schematic of the transition energy from oxygen core 1s level to the LUMO (1π*) for PNBA, BA and PABA.

The intermolecular interaction at the hydrogen bonded carbonyl group, C=O···H–O, gave a relative energy difference of O 1s C=$\underline{\mathbf{O}}_{\mathrm{PABA-PNBA}}$ difference of about +0.7 eV, which is significantly less than for the other chemical moieties in both systems, which appear with a chemical shift of +1.0 eV. This is likely caused by the additional hydrogen bond in PABA, C=**O···H**−N, which reduces the electron density at the oxygen atom.

IR and Raman have specific features in the spectra that differentiate between a carboxylic acid dimer and catemer. The main identifier is the O−H out of plane peak for dimerise compound is in between 900 to 950 cm^{-1}, accompanied by the difference between the carbonyl peak for IR and Raman being more than 30 cm^{-1}. PNBA, BA and α−PABA exhibits these criteria in comparison to β-PABA which is a catemer form. However, XPS and NEXAFS did not show any specific features that differentiate between a carboxylic acid dimer and catemer. The differences in the electronic structure between a dimer and a catemer appear to be too subtle for detection.

7. SOLUTION CHEMISTRY OF BENZOIC ACID DERIVATIVES BASED ON FTIR AND NEXAFS SPECTROSCOPY

Abstract: Fourier transform infrared (FTIR) and oxygen K−edge near-edge x-ray absorption fine structure (NEXAFS) spectra of *p*-nitrobenzoic acid (PNBA), *p*-aminobenzoic acid (PABA), and benzoic acid (BA) in solution were obtained and compared. The substituent effect on the carbonyl stretching frequency and the transition energy for the C=$\underline{\mathbf{O}}$ 1s → 1π* of PNBA, BA, and PABA in acetonitrile and ethanol were in the same order as in the solid state, indicating dominance of the intramolecular effect of the para substituents over the intermolecular interactions. The FTIR data indicate that solution speciation in acetonitrile to is characterised by a concentration dependent equilibrium between carboxylic acid dimers and solute-solvent complexes, while in ethanol there are just solute-solvent interactions through the stronger hydrogen bond in this system.

7.1 Introduction

Studies of solution chemistry using various spectroscopic techniques, including nuclear magnetic resonance (NMR), Fourier transform infrared (FTIR), Raman, and X-ray absorption technique have previously revealed solute-solute and solute-solvent interactions [4]. Interactions in the solution often correlate with structure features in crystal structures, as in the case of tetrolic acid [115,125], benzoic acid (BA) [115], inosine [126], glycine [127], benzophenone and diphenylamine [128], tolfenamic acid [39,129], and isonicotinamide [130]. Solutions of benzoic acid derivates often show evidence for molecular self-association, most commonly through dimer formation through hydrogen bonding between carboxylic acid groups, which are the same synthonic motif as in the corresponding crystal structures.

FTIR has been used for characterising solutes in solution quite extensively, relying on the carbonyl stretch (C=O) and the O−H out-of-plane vibrational features to characterise solution speciation. The O−H out-of-plane is believed to correlate with dimerisation, but it is observed only in non-protic solvents such as chloroform and hexane [81,115,125]. The carbonyl stretching frequencies in carboxylic acid systems are commonly observed as doublets. Each band can be assigned to dimer, monomer, and solvate formation, using the characteristic corresponding carbonyl frequencies [37–39,131] and as well to acyclic dimer [40]. This technique had shown good sensitivity in detecting different speciation in the solution.

Recently, near-edge X-ray absorption fine structure (NEXAFS) spectroscopy has revealed interesting features in the electronic structure of solution species by probing the nitrogen K−edge absorption [47,49]. Observable differences in the N K−edge transitions to unoccupied 1π* and 3π* states distinguished the formation of anionic *p*-aminobenzoic acid (PABA) from its non-ionic crystalline state, while the loss of these transitions was observed for the cationic PABA form, in which protonation of the amino group leads to electronic decoupling of the nitrogen moiety from the aromatic system [47] . Based on these results it was concluded that PABA in methanol was almost exclusively in the non-zwitterionic form, as the N K-edge spectrum exhibited the 1π* and 3π* features of the non-ionic crystalline spectrum and the anionic solute form [47]. As for the cationic form, absence of these features would be expected for the zwitterionic form, in which the amino group would be protonated. Likewise, experimental and computational works on the imidazole system reveal the differences between gas phase, crystal and aqueous solution species [49].

Adding NEXAFS characterisation of PNBA solutions to the still rather limited body of available NEXAFS data for solutions of organic solutes, this chapter describes a comparative study of the influence of hydrogen bonding and substituents on the O K−edge spectrum in the solution state, by comparing to results for the solid state (Chapter 6) and examining solvent effects. The chemical shift between the absorption from the carbonyl (C=O) group in the solid state (Chapter 6) to the solution state is the focus of these investigations. It will be shown that the variation in hydrogen bonding at this centre relates to the changes in the spectral features. Acetonitrile and ethanol were chosen as solvents, as acetonitrile is a strong hydrogen bond acceptor and ethanol a strong hydrogen bond donor. The substituent effects in the solutions of *p*-aminobenzoic acid (PABA), benzoic acid (BA), and *p*-nitrobenzoic acid (PNBA) in acetonitrile and ethanol were also examined through the shift in the carbonyl stretching frequency (*v*C=O) and the oxygen K−edge NEXAFS (C=**O̲** 1s → 1π*).

7.2 Methodology

Only a short summary is given here, while full details are given in Chapter 2 (Section 2.8.2 and 2.11). A demountable cell with the path length of 0.2 mm and KBr windows was used to measure the FTIR spectra of the diluted solution for PNBA and BA in acetonitrile and ethanol using the Smart Omnic-Transmission, Thermo Nicolet. The spectra were obtained from 32 scans at a resolution of 4 cm^{-1}, in the range of 650 to 4000 cm^{-1}. The FTIR spectra were measured for PNBA in acetonitrile (0.002 M to 0.04 M), PNBA in ethanol (0.01 M to 0.16 M), BA in acetonitrile (0.01 M to 0.2 M), and BA in ethanol (0.01 M to 0.2 M). The FTIR spectra for PABA in ethanol and acetonitrile were measured using 0.1 mm path length [132]. O K−edge partial fluorescence yield (PFY) data for the solutions were measured at the U41-PGM beamline of the BESSY II, Berlin using the liquid microjet technique. Solutions of PNBA, BA, and PABA in ethanol at 0.10 M, 0.15 M, and 0.30 M, and of PNBA, BA, and PABA in acetonitrile at 0.02 M, 0.5 M, and 0.25 M were prepared and filtered at room temperature. An 18 µm diameter glass nozzle with 0.6 mL min^{-1} flow rate was used for generating the liquid microjet.

7.3 Results and Discussion

7.3.1 Fourier Transform Infrared (FTIR) Spectroscopy

Two carbonyl peaks were observed for all the three carboxylic acids in acetonitrile as shown in Figure 7.1; BA has the most intense peak at 1725 cm^{-1}and a small shoulder at 1697 cm^{-1}. PABA has peaks at 1711 cm^{-1}and 1679 cm^{-1} [132] while PNBA has the least intense peak at 1734 cm^{-1} and tiny shoulder at 1719 cm^{-1}. The difference between the two peaks is the smallest for PNBA at 15 cm^{-1}, BA at 28 cm^{-1} and PABA at 32 cm^{-1}.

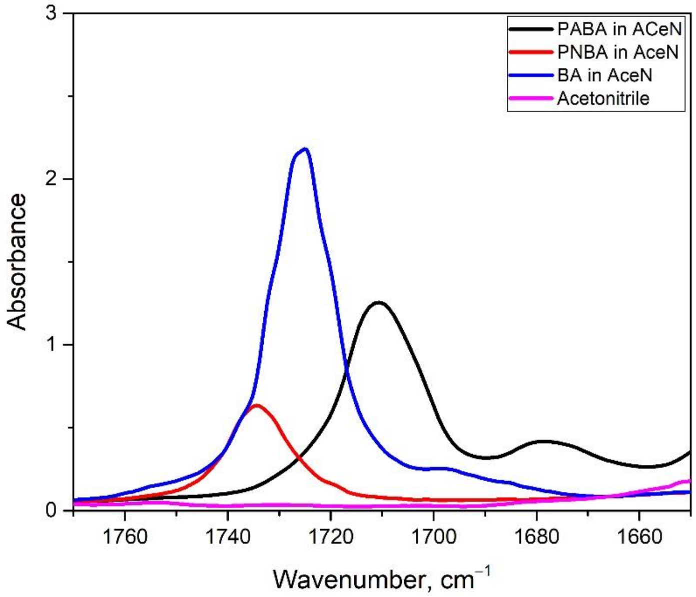


Figure 7.1: The FTIR spectrum of the carbonyl stretching energy for BA (blue), PNBA (red) and PABA [132] (black) in acetonitrile (pink) at 0.15 M, 0.04 M and 0.1 M respectively.

Table 7-1: A summary for both carbonyl stretching energy and the difference between the two peaks in acetonitrile.

	First Peak, cm^{-1}	Second Peak, cm^{-1}
PNBA	1734	1719
BA	1725	1697
PABA	1711	1679

Similar carbonyl peaks were observed for the carboxylic acids in ethanol as shown in Figure 7.2; BA has a less intense peak at 1722 cm^{-1} than at 1700 cm^{-1}, PNBA has a more intense peak at 1725 cm^{-1} than at 1705 cm^{-1} while PABA has a tiny shoulder at 1683 cm^{-1} and a peak at 1715 cm^{-1} [132]. The difference between the two peaks is the smallest for PNBA at 20 cm^{-1}, BA at 22 cm^{-1} and PABA at 32 cm^{-1}.

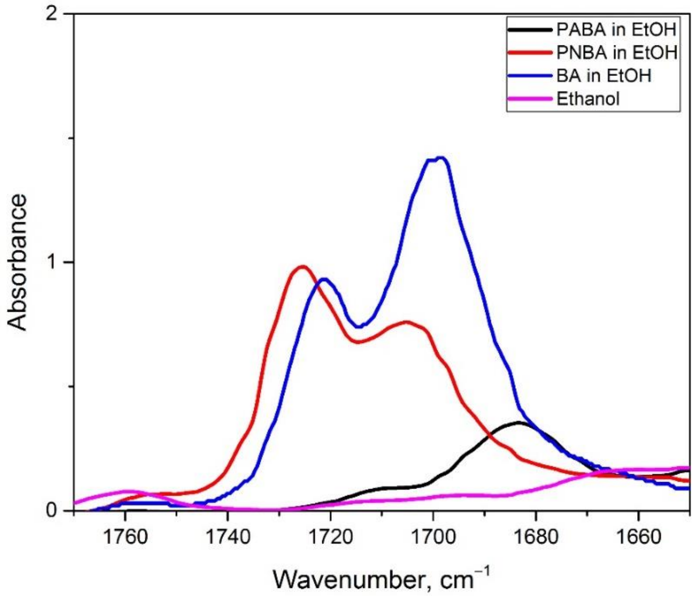


Figure 7.2: The FTIR spectrum of the carbonyl stretching energy for BA (blue), PNBA (red) and PABA [132] (black) in ethanol (pink) at 0.2 M, 0.1 M and 0.1 M respectively.

Table 7-2: A summary for both carbonyl stretching energy and the difference between the two peaks in ethanol.

	First Peak, cm^{-1}	Second Peak, cm^{-1}
PNBA	1725	1705
BA	1722	1700
PABA	1715	1683

The wavenumbers of the carbonyl absorption bands for PABA, BA, and PNBA in acetonitrile and ethanol are in the same order as in the solid state, i.e., PABA has the lowest carbonyl wavenumber followed by BA and PNBA. Therefore, the substituent influence the electron distribution at the carbonyl group in the solution in a similar way as in the solid state; the previous XPS and NEXAFS studies of the solid state indicated that the carbonyl group for PABA has the highest electron density, while PNBA has the lowest, with BA in between. The decrease in the carbonyl stretching frequency is correlated with the strength of hydrogen bonding [133]. Thus, PABA in acetonitrile and ethanol appear to have the strongest hydrogen bond, followed by BA and PNBA.

The solution speciation of these benzoic acid derivatives in acetonitrile and ethanol are different as discussed in Section 1.4. Acetonitrile is a strong hydrogen bond acceptor, while ethanol is a strong hydrogen bond donor and weak hydrogen bond acceptor. Possible interactions at the carboxylic acid are illustrated for acetonitrile (I), ethanol (II), dimer (III) and monomer (IV) in Figure 7.3. The carbonyl stretching frequency would be the highest for the monomer, followed by solute-solvent interaction with solution molecules, the carboxylic acid dimer in solution followed by the dimer in the solid state [37]. BA in acetonitrile was expected to have the speciation of (I) and (III), resulting in a carbonyl stretching frequency at 1726 cm^{-1} and 1697 cm^{-1} [37]. BA and PNBA in carbon tetrachloride (CCl_4) were predicted to have the speciation of (III) and (IV), resulting in a carbonyl stretching frequency at 1696 cm^{-1} and 1742 cm^{-1} for BA [37,131] and 1707 cm^{-1} and 1752 cm^{-1} for PNBA [131].

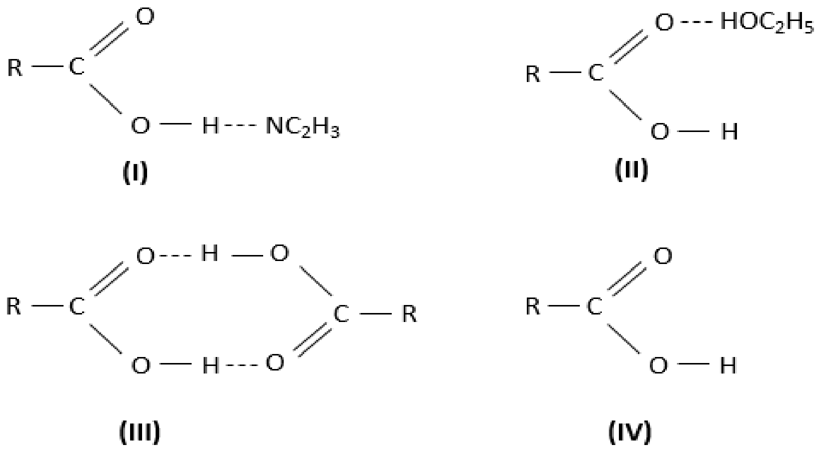


Figure 7.3: The carboxylic acid interaction with the acetonitrile, which is a strong hydrogen bond acceptor (I) and with the ethanol, which is a strong hydrogen bond donor (II), dimer (III) and monomer (IV).

In addition, the relative intensities of the carbonyl doublet band as a function of dilution can be used to characterise the solution speciation [38,39]. Theoretically, the concentration of molecules that are self-associated (dimers) increases with the concentration at the expense of monomer concentration. Therefore, the intensity of the carbonyl band that increases with concentration is the dimer band, while the intensity of the carbonyl band that reduces with concentration is that of the monomer band. These observations are commonly made in non-polar solvents, for example for tolfenamic acid in toluene [39]. In a strongly solvating system, the relative intensities of the carbonyl doublet band as a function of concentration are constant, such as PABA in ethanol [132] and tolfenamic acid [39] in ethanol.

Dilution experiments had previously been performed for PABA in ethanol and acetonitrile [132] and comparison measurements were therefore carried out for PNBA and BA in ethanol and acetonitrile, as shown in Figure 7.4 and Figure 7.5. The carbonyl bands for these systems did not shift in energy in the concentration range from 0.20 M to 0.01 M.

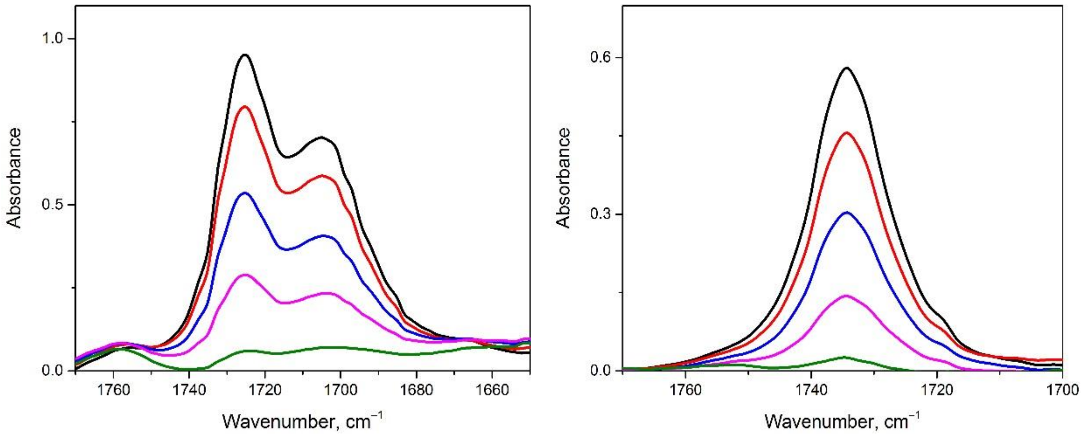


Figure 7.4: Solution IR spectra for dilution experiments for PNBA in ethanol (left) at room temperature – 0.16 M (black), 0.12 M (red), 0.08 M (blue), 0.04 M (pink) and 0.01 M (green) and PNBA in acetonitrile (right) – 0.04 M (black), 0.03 M (red), 0.02 M (blue), 0.01 M (pink) and 0.002 M (green).

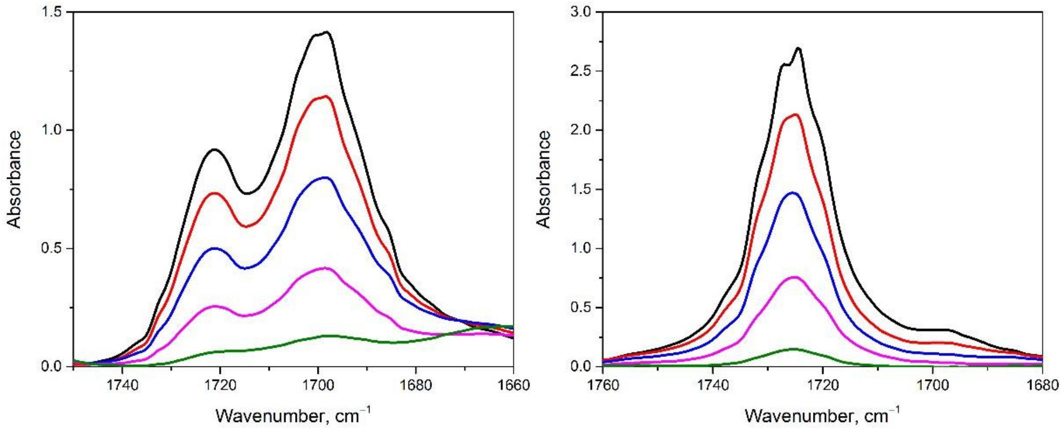


Figure 7.5: Solution IR spectra for dilution experiments for BA in ethanol (left) and in acetonitrile (right) at room temperature – 0.20 M (black), 0.15 M (red), 0.10 M (blue), 0.05 M (pink) and 0.01 M (green).

Table 7-3 shows that the intensity ratio between the carbonyl bands for PNBA, BA and PABA in acetonitrile changes with dilution while in ethanol there are no changes. For all three compounds, both carbonyl bands decrease in intensity in acetonitrile as the concentration is decreased, but the band at lower wavenumber decreases more steeply, approaching zero intensity at low concentrations. The relative decrease difference is most pronounced for BA, followed by PABA and PNBA. The peak that decreases strongly with dilution in acetonitrile also appears to stem from different intermolecular interactions at the carbonyl group, with values of 1718 cm^{-1} for PNBA, 1697 cm^{-1} for BA and 1679 cm^{-1} for PABA.

Table 7-3: A summary of the carbonyl bands intensity ratio for PNBA, BA, and PABA in ethanol and acetonitrile at diluted concentration (Conc.).

PNBA in Acetonitrile			
Conc.	Absorbance at		Ratio
	1734 cm^{-1}	1718 cm^{-1}	
0.04 M	0.58	0.09	6.5
0.03 M	0.45	0.07	6.5
0.02 M	0.30	0.04	6.8
0.01 M	0.14	0.02	8.9

PNBA in Ethanol			
Conc.	Absorbance at		Ratio
	1725 cm^{-1}	1697 cm^{-1}	
0.16 M	0.88	0.63	1.4
0.12 M	0.72	0.52	1.4
0.08 M	0.47	0.34	1.4
0.04 M	0.25	0.19	1.3

Table 7-3 (continue): A summary of the carbonyl bands intensity ratio for PNBA, BA, and PABA in ethanol and acetonitrile at diluted concentration (Conc.).

BA in Acetonitrile			
Conc.	Absorbance at		Ratio
	1726 cm^{-1}	1697 cm^{-1}	
0.20 M	2.65	0.26	10.0
0.15 M	2.09	0.17	12.6
0.10 M	1.44	0.08	19.2
0.05 M	0.74	0.02	49.2

BA in Ethanol			
Conc.	Absorbance at		Ratio
	1721 cm^{-1}	1697 cm^{-1}	
0.20 M	0.92	1.41	0.7
0.15 M	0.73	1.14	0.6
0.10 M	0.50	0.80	0.6
0.05 M	0.25	0.41	0.6

PABA in Acetonitrile			
Conc.	Absorbance at		Ratio
	1710 cm^{-1}	1680 cm^{-1}	
0.10 M	0.33	0.06	5.2
0.08 M	0.28	0.04	6.3
0.01 M	0.20	0.03	8.0
0.001M	0.09	0.01	15.3

PABA in Ethanol			
Conc.	Absorbance at		Ratio
	1715 cm^{-1}	1683 cm^{-1}	
0.23 M	0.06	0.35	0.2
0.17 M	0.05	0.28	0.2
0.11 M	0.04	0.20	0.2
0.06 M	0.01	0.07	0.2

Referring to Table 7-3, PNBA, BA, and PABA in acetonitrile have two carbonyl bands with the difference of less than 40 cm^{-1} and the relative intensities of the carbonyl bands changes with the concentration. The intensity of the carbonyl band that is reduced with the concentration is assigned to the solute-solvent interaction (I), while the one with the increase in intensity is the dimer band. PNBA, BA, and PABA in ethanol have two carbonyl bands with a difference of less than 40 cm^{-1} and there are no changes in these relative carbonyl band intensities with the concentration. Thus, both bands are assigned to solvate formation. The difference between the carbonyl bands for PNBA and BA in CCl_4 are more than 40 cm^{-1}, highlighting the significant difference between a dimer and monomer interaction. Table 7.4 summarise the assignments of the carbonyl bands for PNBA, BA, and PABA in acetonitrile, ethanol and CCl4.

Table 7-4: The solution speciation and the carbonyl stretching energy for PNBA, BA and PABA in acetonitrile, ethanol and CCl_4.

Solute	Solid, νC=O, cm^{-1}	Solvent	Dimer, νC=O, cm^{-1}	Solvates, νC=O, cm^{-1}	Monomer, νC=O, cm^{-1}
PNBA	1684	Acetonitrile	1719	1734	
BA	1675	Acetonitrile	1697	1725	
PABA (α/β)	1659/1687	Acetonitrile	1679	1711	
PNBA	1684	Ethanol		1725, 1705	
BA	1675	Ethanol		1722, 1700	
PABA (α/β)	1659/1687	Ethanol		1683, 1715	
PNBA	1684	CCl_4	1707		1752
BA	1675	CCl_4	1696		1742

7.3.2 Near-Edge X-Ray Absorption Fine Structure (NEXAFS) Spectroscopy

The O K−edge spectra are dominated by the absorption of the ethanolic OH group, but the O 1s (C=**O**) and a weak shoulder signal from the O 1s (**O**−H) from the carboxylic acids could detected in the pre-edge region before the onset of the strong ethanol absorption. Close-ups of the O K−edge spectra of PNBA, BA and PABA in ethanol, in the region of 530 to 533 eV (Figure 7.6), show that PNBA has the lowest energy π* resonance, followed by BA and PABA. PNBA appears to have the lowest intensity due to its low concentration, 0.1 M, in comparison to BA and PABA at 1.5 M and 0.3 M. PNBA has a broader resonance that consists of two peaks, the O 1s (C=**O**) →1π* and O 1s (N**O**$_2$) →1π* at 530.6 eV and 531.2 eV as shown in Figure 7.7. BA has a main peak and a shoulder for the O 1s (C=**O**) →1π* and O 1s (**O**−H) →1π* at 531.3 eV and 533.6 eV while PABA has the O 1s (C=**O**) →1π* peak at 531.6 eV. The **O**−H peak in PNBA and PABA is shielded by the ethanol signal. Relative to BA, the chemical shift of O 1s (C=**O**) →1π* transition for PNBA is −0.7 eV and PABA is +0.3 eV.

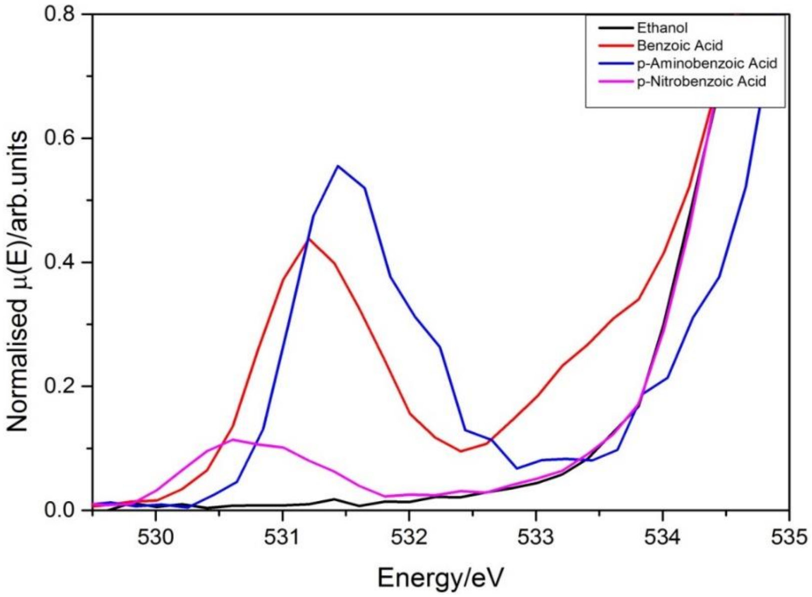


Figure 7.6: A graph of benzoic acid derivatives -BA (red), PNBA (pink) and PABA (blue) in ethanol and the ethanol spectra (black) in the pre-edge region of 530 to 535 eV.

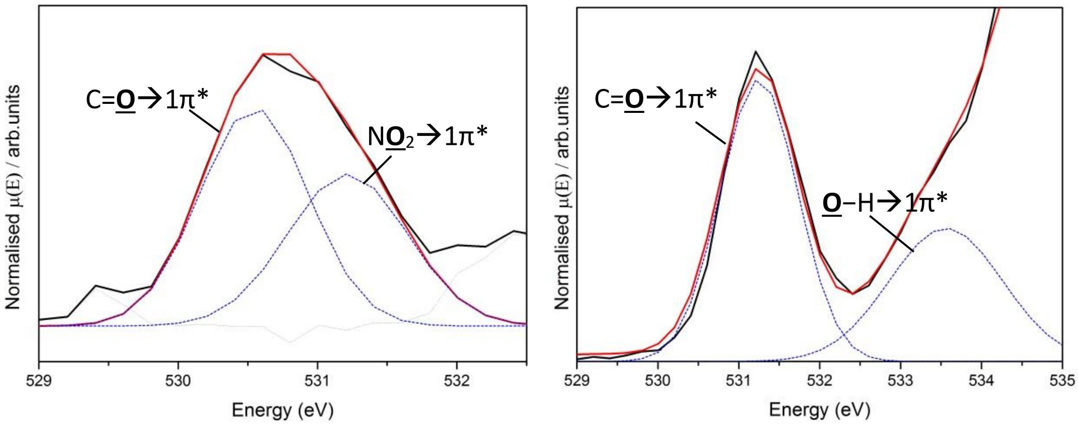


Figure 7.7: A close-up of PNBA (left) and BA (right) in ethanol and the fitted peaks for the O1 C=O, NO2 and O−H transitions.

The O K−edge spectra of PNBA, BA and PABA in acetonitrile, which does not suffer from O K-edge background absorption from the solvent, in the region of 530 to 536 eV (Figure 7.8) shows that PNBA has the lowest energy of the edge onset, followed by BA and PABA. The signal-to-noise quality for PNBA is not as good as for the other two compounds due to its low concentration, 0.03 M. PNBA is characterised by O 1s (C=**O**) →1π* and O 1s (N**O**$_2$) →1π* transitions at 530.5 eV and 531.2 eV. BA and PABA have the O 1s (C=**O**) →1π* at 531.2 eV and 531.9 eV, and the O 1s (**O**−H) →1π* at 533.8 eV and 534.4 eV, respectively, as shown in Figure 7.9. Relative to BA, the chemical shift of O 1s (C=**O**) →1π* transition for PNBA is −0.7 eV and for PABA is +0.7 eV.

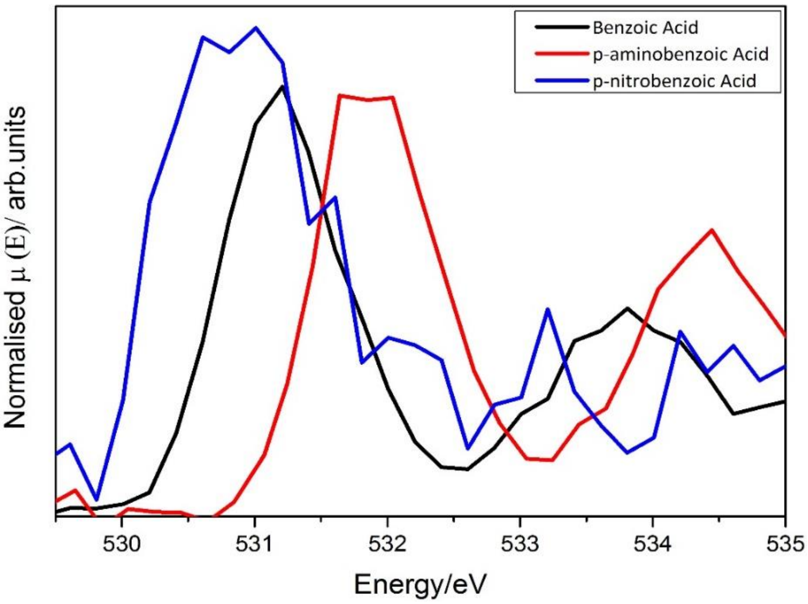


Figure 7.8: A poor yet conclusive signal that PNBA (blue) has the lowest energy in the pre-edge region of 530 to 535 eV follow by BA (black) and PABA (red) in acetonitrile.

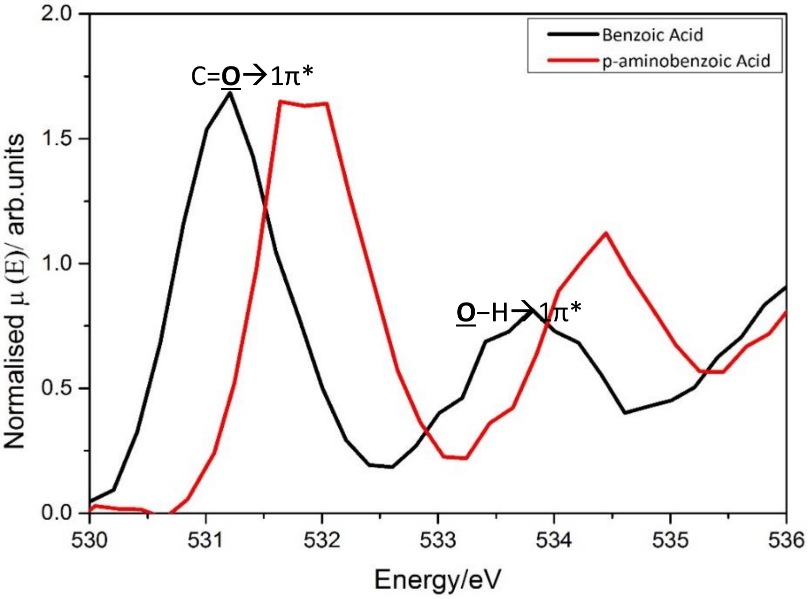


Figure 7.9: A graph of benzoic acid derivatives −BA (black) and PABA (red) in acetonitrile in the pre-edge region of 530 to 535 eV.

The O K-edge spectra in the region of 530 to 536 eV for PABA in the solid forms, α− and β−PABA, and PABA in solution, ethanol and acetonitrile, are shown in Figure 7.10. The O 1s (C=**O**) →1π* energy for α− and β−PABA are 532.1 eV and 531.7 eV while in the solution of acetonitrile and ethanol, they occur at 531.9 eV and 531.6 eV. This chemical shift is driven by the differences in the intermolecular interaction at the carbonyl group (C=O···H). The key finding is the energy of O 1s (C=**O**) →1π* for PABA in acetonitrile, which is in-between that of the two polymorphic solid forms; this will be discussed further in the Discussion section below. The O 1s (**<u>O</u>**−H)→1π* energy for α−, β−PABA and in acetonitrile are at 533.5 eV, 533.2 eV and 534.4 eV, respectively, while in ethanol this signal is not detectable against the strong OH signal from the solvent.

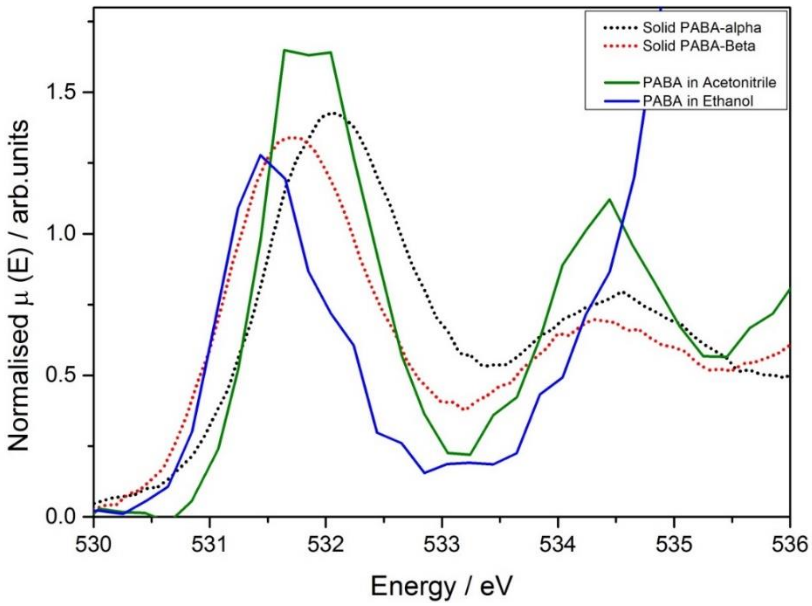


Figure 7.10: Graph of PABA in solid forms, alpha (dotted black) and beta (dotted red) as well PABA in acetonitrile (green) and PABA in ethanol (blue) in the pre-edge region of 530 to 535 eV.

The O K-edge spectra in the region of 530 to 536 eV for BA in acetonitrile and ethanol are shown in Figure 7.11. Solid spectra of BA for comparison are not available, as the solid evaporated under the vacuum conditions of the experiment. The O 1s C=**<u>O</u>** → 1π* peaks for BA in ethanol and acetonitrile are at 531.3 eV and 531.2 eV. The O 1s **<u>O</u>**−H → 1π* peaks for BA in acetonitrile are at 533.8 eV. There is no significant chemical shift (+0.04 eV) between BA and PNBA in ethanol and acetonitrile for O 1s C=**<u>O</u>** → 1π*.

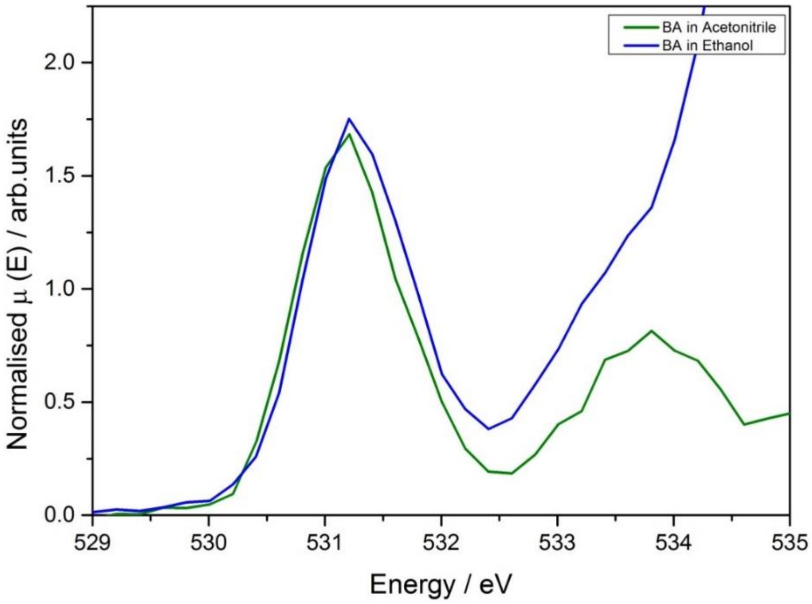


Figure 7.11: Graph of BA in acetonitrile (green) and in ethanol (blue) in the pre-edge region of 530 to 535 eV.

The O K-edge spectra in the region of 529 eV and 533 eV for PNBA in solid form and in acetonitrile and ethanol are shown in Figure 7.12. As PNBA polymorphs have similar spectra, only one of them is shown here (Form II). The O 1s C=**<u>O</u>** →1π* energy for PNBA in the solid, in acetonitrile and in ethanol are at 530.6 eV, 530.5 eV and 530.6 eV, respectively, while the corresponding O 1s N**<u>O</u>**$_2$→1π* are at 531.6 eV, 531.2 eV and 531.2 eV. Relative to the solid form, the chemical shift for O 1s C=**<u>O</u>** →1π* are −0.14 eV and −0.08 eV while O 1s N**<u>O</u>**$_2$→1π* are both at −0.3 eV for PNBA in acetonitrile and ethanol.

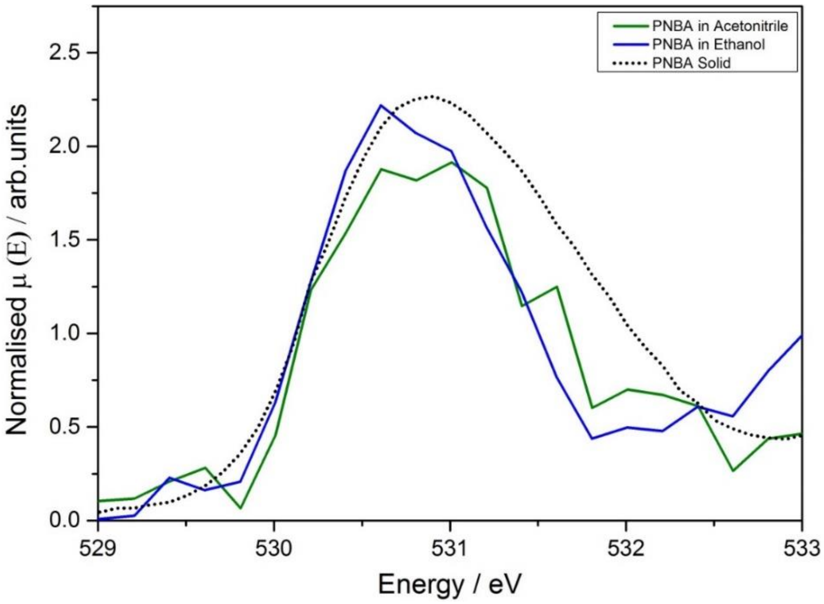


Figure 7.12: Graph of PNBA in solid form (dotted black) as well PNBA in acetonitrile (green) and PNBA in ethanol (blue) in the pre-edge region of 530 to 535 eV.

The substituent effect on the transition of C=**O** 1s → 1π* for PNBA, BA, and PABA in acetonitrile and ethanol is consistent with the observations made for the solid state (Chapter 6); PNBA has the lowest onset of 1s → 1π* transitions, followed by BA and then PABA. Having XPS solution data would be an advantage here, but based on the similarity of the NEXAFS between solution data, gas and solid state it seems likely that PNBA has the lowest ionisation potential and LUMO, followed by BA and PABA in ethanol and acetonitrile.

Table 7-5 summarise the chemical shift for the C=**O** 1s → 1π* from the solid to the solution state for PNBA, BA, and PABA in ethanol and acetonitrile. Two main observations are made in comparing these results to the FTIR. Firstly, NEXAFS detects only a single carbonyl band while FTIR detects two carbonyl bands that are associated with the dimer and solvates interaction. Secondly, NEXAFS shows an insignificant chemical shift from the solid to the solution state and solvent effects while FTIR shows a significant chemical shift.

Table 7-5: The transition energy for O 1s C=O → 1π* and the relative energy difference, ΔE, to the crystal for PNBA, BA and PABA in acetonitrile and ethanol (e.g, $\Delta E_{PNBA} = \Delta E_{PNBA, Acetonitrile} - \Delta E_{PNBA, crystal}$)

Solute	Crystal	Solvents	
	Dimer	Acetonitrile	Ethanol
PNBA	530.6	530.5	530.6
BA		531.2	531.3
α−PABA	532.1	531.9	531.6
β−PABA	531.7		
ΔE_{PNBA}		−0.1	0
$\Delta E_{\alpha-PABA}$		−0.2	−0.5
$\Delta E_{\beta-PABA}$		+0.2	−0.1

XPS solution data and density functional theory calculation would be helpful for giving better insight into the electronic structure and the effect of solvation or self-association to the C=**O** 1s → π* transition energy. This highlight the importance of this experimental data for future computational works. The actual position of C=**O** 1s core level and the LUMO (1π*) may shift between the solid and solution state or between different solvent without shifting the total transition energy C=**O** 1s → π* [52]. The second possibility is the existence of two carbonyl peaks with different intermolecular interaction, as observed in the FTIR work

(Section 7.3.1). Overall a similar picture as for the solid state studies (Chapter 6) emerges, in that the O K−edge NEXAFS appears to be insensitive to detecting the small difference between the two hydrogen bonded systems in comparison to the difference observed between the monomer−dimer system [51,100].

Figure 7.13 shows an illustration that assumes the LUMO (1π*) for PNBA in acetonitrile (III) and in ethanol (IV) are the same as the experimentally obtained α− and β−PABA crystals, (I) and (II). The core level O 1s C=**<u>O</u>** for PABA in acetonitrile and in ethanol are then re-calculated from the experimentally obtained NEXAFS data and the assume LUMO energy. The core level O 1s C=**<u>O</u>** for PABA in acetonitrile (III) is in between α− and β−PABA (I and II) while PABA in ethanol (IV) is above all. PABA in ethanol has the highest electron density (lowest ionisation potential) as the hydrogen bond interaction is the weakest. The existence of a dimer for PABA in acetonitrile may reduce the electron density (higher ionisation potential) as the hydrogen bond is stronger. The differences between the C=**<u>O</u>** 1s → π* for PABA in acetonitrile and ethanol is 0.3 eV. This supports the FTIR works that PABA in acetonitrile has dimer speciation, while in ethanol, there are only solvates.

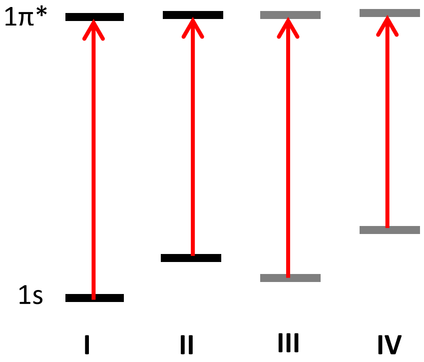


Figure 7.13: A predicted energy level 1s and 1π* for PABA in acetonitrile (III) and in ethanol (IV) based on α− (I) and β−PABA (II) crystals.

Contrasting to the PABA results, PNBA and BA did not exhibit a significant chemical shift between the solvents. A possible explanation is that the electron density at the carbonyl group for these systems is low and thus less sensitive to the small changes in the intermolecular interaction. In contrast, the electron density at the carbonyl in the PABA system is high, thus the ionisation potential energy is sensitive to changes. The solvent effect for PABA differs by 0.3 eV which is still on the lower side, thus for a system with lower electron density at the carbonyl moiety, such as PNBA and BA, it is not surprising that only 0.1 eV changes was observed. Future work should support these hypotheses through validation using computational methods.

7.4 Conclusion

The substituent effect on the carbonyl stretching energy and the transition energy for C=**O** 1s → $1\pi^*$ in solutions parallels that observed for the solid state (Chapter 6). The strongly electron withdrawing nitro group reduces the electron density at the carbonyl centre, while the electron donating amino group increases it. Moving from the solid to the solution state, the C=O resonance structure is strengthened as the electron density at the oxygen acceptor is reduced due to the weaker hydrogen bonding. Therefore in the solution state, the C=O stretching wavenumber is higher than in the solid state, on average by 24 cm^{-1} for the dimers and 43 cm^{-1} for the solvates. PNBA has the highest wavenumber, followed by PABA and BA in both ethanol and acetonitrile solution. PNBA has the highest wavenumber, as the resonance effect is more electron deficient, while PABA has the lowest as the resonance effect is weakened by higher electron density. PNBA forms the weakest hydrogen bond followed by PABA and BA.

The transition energy C=**O** 1s → π^* for PNBA, BA and PABA in the solutions of acetonitrile and ethanol are in the same order as in the solid state (Chapter 6). PNBA has the lowest transition energy followed by BA and PABA in ethanol and acetonitrile. As this is similar to the solid state, the electronic states of these systems should follow a similar energetic pattern. In the solid state, PNBA has the lowest 1s core level binding energy and lowest LUMO energy, followed by BA and PABA. In solution, PNBA exhibits no significant shift to the solid state, while PABA in acetonitrile has shifts of +0.2/− 0.2 eV and in ethanol −0.5/−0.1 eV relative to the α−/β−PABA crystal structures, respectively. Full interpretation of this contrasting behaviour between PNBA and PABA systems should await XPS characterisation of the core level binding energy shifts, to disentangle initial and final state effects on the electron spectra, and possibly accompanying DFT calculations.

Solution speciation in ethanol and acetonitrile was deduced from dilution series via FTIR analysis, and compared to the transition energy of C=**O** 1s → $1\pi^*$ in the solution NEXAFS. Based on relative intensity ratios in the carbonyl peak structure it was concluded that PNBA, BA, and PABA in acetonitrile form both solute-solute (dimer) and solute-solvent associates, while in ethanol there are just interactions with the strongly hydrogen bonding solvent molecules. The NEXAFS data reveal only a single carbonyl band, which does not indicate the existence of two solution species in acetonitrile or ethanol. Assuming the unoccupied molecular orbital energy levels did not change (as previously observed for α− and β−PABA crystal structures), the ionisation potential of PABA in acetonitrile should therefore be lower than in ethanol. These results support the view that dimers exist for PABA in acetonitrile, as the ionisation potential is higher and closer to that in the solid state. In ethanol, the ionisation potential is lower, indicating solvation through hydrogen bonds causes more of an electron density increase at the carbonyl oxygen compared to the solid state.

The NEXAFS data for PNBA and BA did not show any solvent effect. For PNBA this is likely the case because of the dominant influence of the nitro group, which lowers the electron density at the carbonyl group and so the influence of additional solvent interactions is comparatively weak. Future work should examine the experimental observations reported here further through DFT calculations, including the influence of the aromatic ring substituent on the NEXAFS of complexes with solvent molecules, and confirming the interpretation of the changing double-carbonyl vibrational bands as evidence for solute-solute and solute-solvent interactions. Overall, this first attempt at drawing a deeper understanding of solution speciation using 1s → 1π* chemical shift in a solid and solution state NEXAFS raises a number of topics for future investigations.

8. COMPARATIVE ANALYSIS OF NUCLEATION IN BENZOIC ACID SYSTEMS: COMBINING THERMODYNAMICS, KINETICS, AND SPECTROSCOPY RESULTS

Abstract: The nucleation rates of substituted benzoic acids in organic solvents are influenced by the pre-exponential kinetic factor A and the thermodynamic factor B. It has been suggested that the A value can be linked to a microkinetic activation barrier, $\Delta G^{\neq}$, while the B value stems from the well-known nucleation barrier, $\Delta G*$. Higher nucleation rates are associated with high A and low B values. $\Delta G^{\neq}$ was found to be the smallest for PNBA, followed by < PTA < PABA ≅ BA, while the $\Delta G*$ is the lowest in the reverse order PTA ≅ PABA < BA < PNBA. With the exception of BA, the observed order of $\Delta G^{\neq}$ is similar to the carbonyl stretching frequency in the respective solutions, suggesting a link with the vibrational dynamics at carboxylic acid group.

8.1 Introduction

Studies of organic crystal nucleation kinetics based on induction time measurements are often used to obtain mechanistic information. The nucleation rates have been shown to be influenced by many factors, such as the thermal history [95,96], stirrer design [53], agitation rate [29,92], reactor design [92], temperature [134], and solvent effects [28,30,53]. The nucleation rate was calculated by fitting either a lognormal cumulative distribution function [30] or a Poisson distribution function [27]. The thermodynamic factor, B, the pre-exponential kinetic factor, A, and the effective interfacial energy, γ_{eff}, were obtained from the nucleation rate, J, supersaturation based on mole fraction, S, and the induction times, t_{ind}.

The thermodynamic factor, B, pre-exponential kinetic factor, A, and effective interfacial energy, γ_{eff}, were applied in understanding the nucleation process. The metastable polymorph of *m*-hydroxybenzoic acid (mHBA) has lower interfacial energy, a lower thermodynamic factor *B* and a higher pre-exponential kinetic factor *A*. The nucleation rate of this metastable Form II is higher at increasing driving force, thus Form II crystallises more than Form I [29]. The nucleation rate of diprophylline (DPL) Form I in 2-propanol is higher than DPL Form II in dimethylformamide (DMF), which is driven by higher thermodynamic factor thus higher interfacial energy [26]. Solvent effects for the nucleation of salicylic acid in organic solvents exhibited a correlation between the nucleation rate and effective interfacial energy. However, there was no correlation found to the solubility or solvent properties [53]. The solvent effects for the nucleation of risperidone was correlate to the solute-solvent binding energy, which appeared to point to the importance of the desolvation process [30].

The nucleation rate is governed by two energy barriers as illustrated in Figure 8.1; $\Delta G*$is the free energy barrier in creating the surface for the nucleus while $\Delta G^{\neq}$ is the activation energy or the transition state barrier in creating the chemical bonds between the molecules as illustrated below [135]. The nucleation rate based on classical nucleation theory (CNT) is given as in Equation 8.1. $\Delta G*$is embedded in the thermodynamic factor, B, while $\Delta G^{\neq}$ is embedded in the pre-exponential kinetic factor, A. The A and B values are obtained from the *y*-intercept and slope of $\ln(J/S)$ against $1/\ln^2 S$ [90].

$$J = AS\exp\left(-\frac{B}{\ln^2 S}\right)$$ **Equation 8.1**

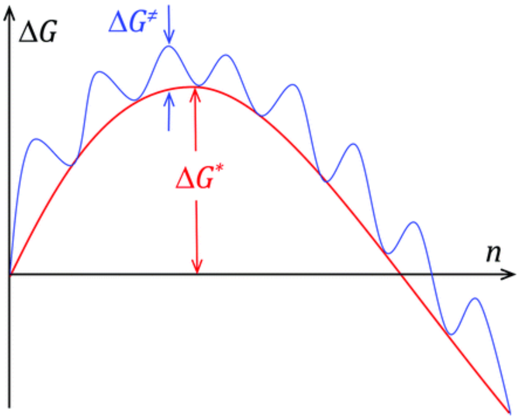


Figure 8.1: A sketch of energy diagram of the free energy in creating the nucleus $\Delta G*$ and the activation energy in creating the chemical bonds $\Delta G^{\neq}$ against the nucleus size, *n* [135].

The main aim of this chapter is to propose the potential energy diagrams in terms of $\Delta G*$ and $\Delta G^{\neq}$, as schematically outlined in Figure 8.1 for several substituted benzoic acids – *p*-nitrobenzoic acid (PNBA), *p*-aminobenzoic acid (PABA), benzoic acid (BA) and *p*-toluic (PTA) in toluene, 2-propanol and acetonitrile – using A and B values determined in the present and in previous work Table 8-1 [136]. Then the possible relevance of the intermolecular interaction at the carbonyl group in understanding the differences in nucleation kinetics will be discussed.

Table 8-1 as well provides the effective interfacial energy, γ_{eff}, and the solubility at 20 °C in mole fraction, x, for all the acids except PNBA at 25 °C. There was no correlation observed between the solubility either to the thermodynamic or kinetic factor [43,136].

Table 8-1: The A, B, and γ_{eff} value were taken from the literature [136] with the 95% confidence interval in the parenthesis.

Solute	**Solvent**	$A \times 10^{-2}$ **(m^{-3} s^{-1})**	$B \times 10$	γ_{eff} **(mJ m^{-2})**	**Solubility,** x
BA	Toluene	13 (8.7-18)	5.2 (4.6-5.8)	4.4 (4.2-4.6)	0.062
BA	Acetonitrile	11 (8.6-15)	1.6 (1.4-1.8)	3.0 (2.8-3.1)	0.033
PTA	Toluene	28 (16-47)	0.32 (0.23-0.40)	1.4 (1.3-1.5)	0.008
PTA	2-Propanol	21 (9-49)	1.1 (0.8-1.3)	2.1 (1.9-2.2)	0.040
PABA	Acetonitrile	13 (9.2-18)	0.28 (0.23-0.32)	1.6 (1.5-1.7)	0.018
PABA	2-Propanol	16 (11-23)	1.2 (1.0-1.3)	2.6 (2.5-2.7)	0.028
PNBA	2-Propanol	37 (16-80)	7.4 (5.9-8.9)	4.6 (4.3-4.9)	0.007

PNBA, BA, and PABA crystals exhibit carboxylic acid dimers as a building unit. The substituents effect the strength of solute-solute and solute-solvent intermolecular interactions crystal at the carbonyl group (C=O), as has been shown using FTIR, XPS and NEXAFS. Table 8-2 and Table 8-3 show the carbonyl stretching frequency of PNBA is the highest, both in the both solid and solution state, indicating a weak hydrogen bond interaction, followed by BA and PABA. The ionisation potential of O 1s (C=O) in PNBA is higher than PABA, proving the electron density at the oxygen group is less in PNBA than in PABA.

The binding energy of O 1s (C=O) → 1π* in the solid and solution is consistent, with PNBA having the lowest binding energy followed by BA and PABA. Based on FTIR analysis, these acids in acetonitrile form both solute-solute (dimer) and solute-solvent associates, while in ethanol there are just interactions with the strongly hydrogen bonding solvent molecules. BA in toluene has both solute-solute (dimer) and monomer molecules. These results had been discussed in detail in Chapter 6 (solid state) and Chapter 7 (solution state). PTA has similar FTIR spectra and characterisation as BA [136]. The FTIR spectra for these acids in ethanol are similar to 2-propanol.

Table 8-2: A summary of the carbonyl stretching frequency (*v*C=O) in the solid state using FTIR, the ionisation potential of O 1s (C=O) using XPS and the binding energy of O 1s (C=O) → 1π* using NEXAFS.

Solute (Solid)	*v*C=O, cm^{-1}	O 1s (C=O), eV	O 1s (C=O) → 1π*, eV
PABA	1659	532.1	532.1
BA	1675		
PNBA	1684	532.9	530.6

Table 8-3: A summary of the carbonyl stretching frequency (*v*C=O) in the solution state using FTIR and the binding energy of O 1s (C=O) → 1π* using NEXAFS in acetonitrile and ethanol.

(Solution)	*v*C=O, cm^{-1}			O 1s (C=O) → 1π*, eV	
Solute	**Acetonitrile**	**Ethanol**	**Toluene**	**Acetonitrile**	**Ethanol**
PABA	1679, 1711	1683, 1715		531.9	531.6
BA	1697, 1725	1700, 1722	1697, 1740	531.2	531.3
PNBA	1719, 1734	1705, 1725		530.5	530.6

8.2 Results and Discussion

8.2.1 Thermodynamic Factor

The nucleation barrier, ΔG^*, is the maximum work to form the critical nucleus size, n^*, and is calculated from the thermodynamic factor B using Equation 8.2 and Equation 8.3. k is the Boltzmann constant, T is the absolute temperature (K) and S is the supersaturation ratio (x/x^*) [90].

$$\frac{B}{\ln^2 S} = \frac{W^*}{kT}$$ **Equation 8.2**

$$n^* = \frac{2W^*}{kT \ln S}$$ **Equation 8.3**

The nucleation rate against supersaturation data for the substituted benzoic acids was obtained from the literature [136] and Chapter 4 (PNBA in 2-propanol). The maximum work, W^*/kT, to create a critical nucleus size, n^*, were calculated using Equation 8.2 and Equation 8.3. Figure 8.2 shows the plot of the nucleation rate for the substituted benzoic acids in acetonitrile, 2-propanol, and toluene as a function of critical nucleus size, n^*. As expected, at higher nucleation rate the critical nucleus size is smaller, while at lower nucleation rate the critical nucleus size is bigger.

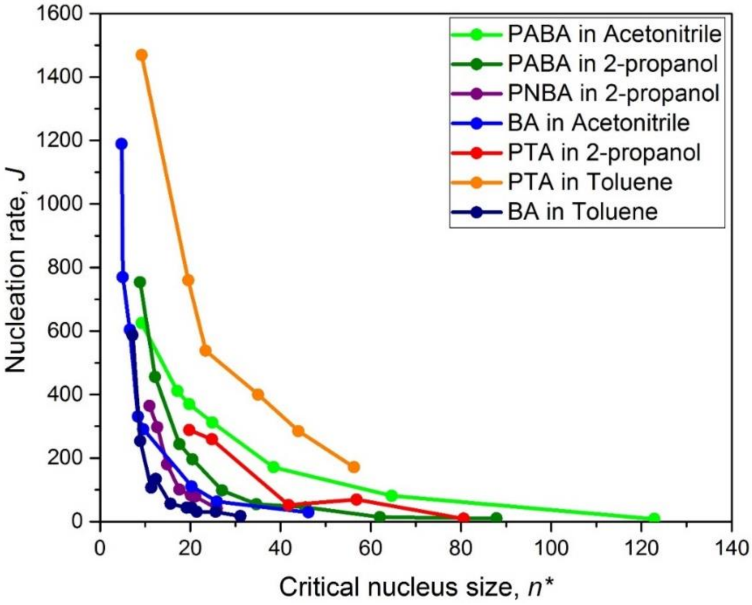


Figure 8.2: The nucleation rate, *J*, as a function of critical nucleus size, *n, for PNBA in 2-propanol (purple), PABA in 2-propanol (dark green), PABA in acetonitrile (light green), PTA in toluene (orange), PTA in 2-propanol (red), BA in toluene (dark blue) and BA in acetonitrile (light blue).**

The maximum work, W^*/kT, to create a critical nucleus size, n^*, of 10, 20, 30, and 40 molecules are shown in Figure 8.3. Figure 8.3 shows that PNBA in 2-propanol has the highest W^*/kT followed by BA in toluene, BA in acetonitrile, PABA in 2-propanol, PTA in 2-propanol, PTA in toluene, and PABA in acetonitrile at each size. A critical nucleus size of 20 molecules was chosen as a realistic reference for the analysis; all the acids show W^*/kT is above 1 kT and lies within the measured range of nucleation rate (Figure 8.2). The main assumption made in this analysis is that the cluster has the same internal structure as the bulk crystal, i.e., in these systems, carboxylic acid dimers.

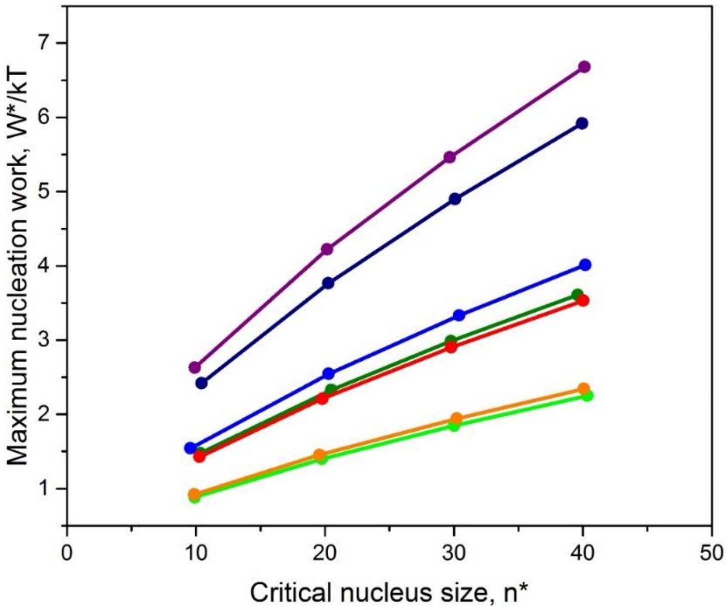


Figure 8.3: The maximum nucleation work, *W*/kT*, as a function of critical nucleus size, *n, for PNBA in 2-propanol (purple), PABA in 2-propanol (dark green), PABA in acetonitrile (light green), PTA in toluene (orange), PTA in 2-propanol (red), BA in toluene (dark blue) and BA in acetonitrile (light blue).**

In a homogeneous system, the maximum work is the sum of the energy of the phase transition and the cluster total surface energy $\gamma A_c(n^*)$. γ (J/m^2) is the specific surface energy of the cluster-solution interface while A_c is the area of the cluster surface. In a heterogeneous system, the effective specific surface energy γ_{ef} (J/m^2), which additionally considers the specific surface energies of the substrate (foreign particles)-solution and cluster-substrate, is used as in Equation 8.4 [5].

$$W^*(n^*) = -n^* kT \ln S + \gamma_{ef} A_c(n^*)$$ **Equation 8.4**

Table 8-4 summarises the calculated S and W^* for $n^* = 20$ based on Equation 8.2, 8.3 and 8.4. The driving force for the nucleation, which is the difference in the chemical potentials is given as, $\Delta\mu = kT \ln S$. These parameters, S, W^*, $-n\Delta\mu$, and $\gamma_{ef} A_c$, are a function of the thermodynamic factor, B, thus they increased in the same order, PABA in acetonitrile $\cong$ PTA in toluene < PTA in 2-propanol $\cong$ PABA in 2-propanol < BA in acetonitrile < BA in toluene < PNBA in 2-propanol as shown in Table 8-4.

Table 8-4: A summary of the calculated S ,$W*$, $-n\Delta\mu$ and $\gamma_{ef} A_c$ at $n* = 20$ using each B value .

System	S , $x/x*$	B	$W*$ x 10^{20} (J)	$-n\Delta\mu$ x 10^{20} (J)	$\gamma_{eff} A_c$ x 10^{20} (J)
PABA in acetonitrile	1.15	0.03	0.6	1.1	1.7
PTA in toluene	1.16	0.03	0.6	1.2	1.8
PTA in 2-propanol	1.25	0.11	0.9	1.8	2.7
PABA in 2-propanol	1.26	0.12	0.9	1.9	2.8
BA in acetonitrile	1.29	0.16	1.0	2.1	3.1
BA in toluene	1.45	0.52	1.5	3.1	4.6
PNBA in 2-propanol	1.52	0.74	1.7	3.5	5.2

Figure 8.4 illustrates the resulting differences in the $\Delta G*$(or W*) barrier for creating the same critical nucleus size. PNBA in 2-propanol requires a relatively high driving force, $S = 1.52$, to overcome a high cluster total surface energy, $\gamma_{eff} A_c$, resulting in a high $\Delta G*$. PABA in acetonitrile and PTA in toluene require the lowest driving force ($S = 1.15$) to overcome the cluster total surface energy.

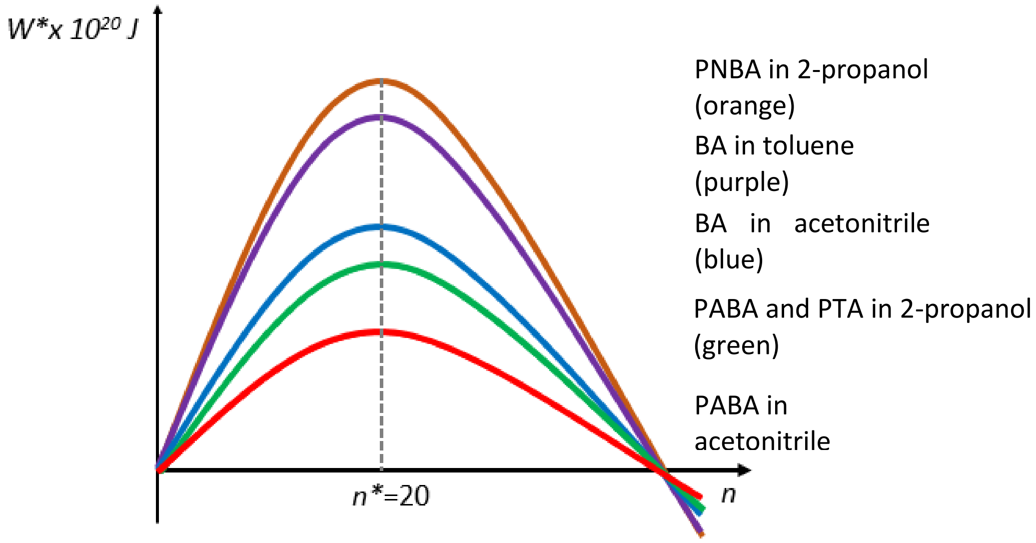


Figure 8.4: An illustration of the differences in nucleation barrier for creating a critical nucleus size of 20 molecules. The lowest energy barrier is for PABA in acetonitrile and PTA in toluene (both in red) followed by PTA and PABA in 2-propanol (both in green), BA in acetonitrile (blue), BA in toluene (purple) and the highest is PNBA in 2-propanol (orange).

Figure 8.5 is a plot of the carbonyl stretching energy (FTIR) for the dimers (solute-solute) and solvates (solute-solvent) against the effective interfacial energy for each solution. There is no correlation between the strength of the intermolecular interactions in the solution to the effective interfacial energy, γ_{eff} , as observed by Mealey et al [30].

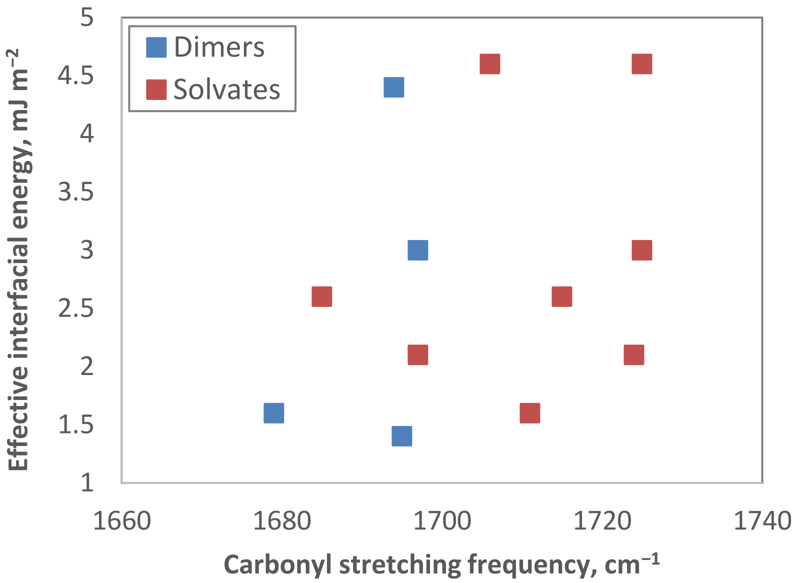


Figure 8.5: A plot of the carbonyl stretching energy for the dimers (blue) and solvates (red) against the effective interfacial energy.

8.2.2 Kinetic factor

The activation energies, $\Delta G^{\#}$, can be obtained from a temperature-dependent nucleation rate analysis [25] which is not covered in this thesis. The $\Delta G^{\#}$ is embedded in the f_0 which is related to the pre-exponential factor A as in Equation 8.5 and Equation 8.6.

$$A = z_0 f_0 C_0$$ **Equation 8.5**

$$f_0 = f \exp\left(-\frac{\Delta G^{\neq}}{kT}\right)$$ **Equation 8.6**

z_0 is supersaturation-independent Zeldovich factor and f is the attachment frequency of building units without an energy barrier [5,25]. The $f_0 C_0$ value is calculated from the A and B value, $f_0 C_0 = A\sqrt{12\pi B}$. f_0 is the supersaturation independent attachment frequency that is not calculated as the concentration of active nucleation sites (heterogeneous system), C_0 is unknown [26]. The order of $f_0 C_0$ value is almost in a similar order as the effective interfacial energy; it seems both are dependent on the unknown parameter of the heterogeneous particles.

Table 8-5: The effective specific surface energy, γ_{eff} , [136] and $f_0 C_0$ calculated from the A and B value.

System	γ_{eff} (mJ/m^2)	$f_0 C_0$ (m^{-3} s^{-1})
PNBA in 2-propanol	4.60	19543
BA in toluene	4.40	5756
pTA in 2-propanol	2.10	5498
PABA in 2-propanol	2.60	3403
PTA in toluene	1.40	3075
BA in acetonitrile	3.00	2702
PABA in acetonitrile	1.60	1336

Nucleation rates to create a critical nucleus size of 20 molecules for a given thermodynamic factor, $B = 0.03$, and pre-exponential kinetic factor, $A = 3700$, were calculated and are summarised in Table 8-6. The systems in the table are ordered according to this nucleation rate $J_{B=0.03}$ from fastest to the slowest. At constant $B = 0.03$, the rates decrease in the order PNBA in 2-propanol > PTA in toluene > PTA in 2-propanol > PABA in 2-propanol > PABA in acetonitrile ≅ BA in toluene ≅ BA in acetonitrile. At constant $A = 3700$, the rates decrease in the order PABA in acetonitrile > PTA in toluene > PTA in 2-propanol > PABA in 2-propanol > BA in acetonitrile > BA in toluene > PNBA in 2-propanol. In general, at $J_{B=0.03}$ the nucleation rate is in the order of PNBA > PTA > PABA > BA while at $J_{A=3700}$ the order is PABA ≅ PTA > BA > PNBA in comparison to the actual J , which is PTA > PABA > BA > PNBA. This shows the influence of the thermodynamic and kinetic factor to the overall nucleation rate.

Table 8-6: A calculated nucleation rate for all systems at $B = 0.03$, $J_{B=0.03}$ and $A = 3700$, $J_{A=3700}$ in comparison to the nucleation rate J calculated at the A and B value.

System	A (m^{-3} s^{-1})	B	$\frac{W^*}{kT}$	J (m^{-3} s^{-1})	$J_{B=0.03}$ (m^{-3} s^{-1})	$J_{A=3700}$ (m^{-3} s^{-1})
PNBA in 2-propanol	3700	0.74	4.2	83	1008	83
PTA in toluene	2800	0.03	1.5	760	760	1004
PTA in 2-propanol	2100	0.11	2.2	288	572	508
PABA in 2-propanol	1600	0.12	2.3	196	436	454
PABA in acetonitrile	1300	0.03	1.4	370	354	1053
BA in toluene	1300	0.52	3.8	44	354	124
BA in acetonitrile	1100	0.16	2.5	111	300	373

In Figure 8.6, an activation energy barrier $\Delta G^{\#}$ sketch on top of the thermodynamic barrier ΔG^* is made based on the A, B and J values. As an example, PNBA in 2-propanol has a high thermodynamic barrier to overcome while PABA in 2-propanol has a lower barrier based on the ΔG^* (Figure 8.6). If the thermodynamic barrier for PNBA were the same as PABA (Figure 8.6, dotted blue line), the nucleation rate would increase to 454 m^{-3} s^{-1} (Equation 8.1) as the kinetic factor A for PNBA is higher at 3700 m^{-3} s^{-1} compared to 1600 m^{-3} s^{-1} for PABA. Hence the activation energy for PNBA ($\Delta G^{\#}{}_{\text{PNBA,2-propanol}}$) would be lower than the corresponding activation energy for PABA.

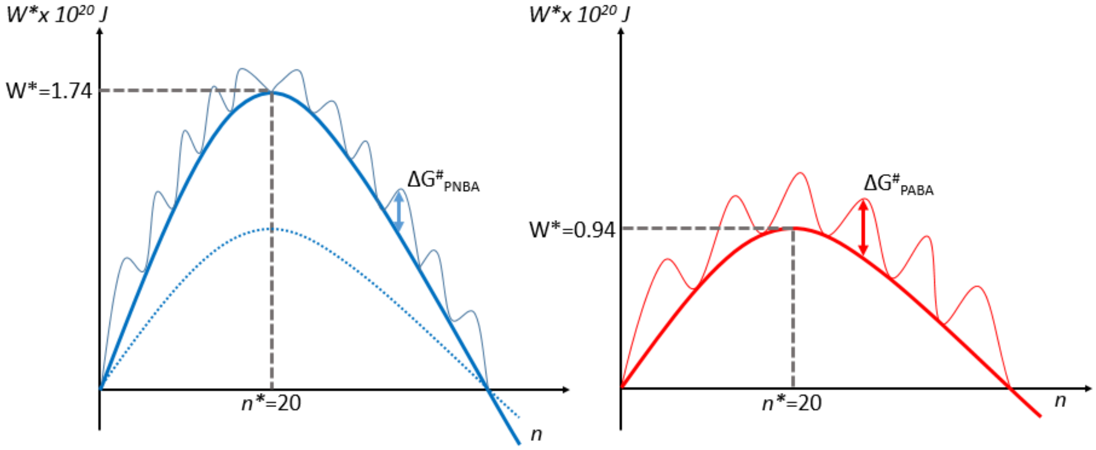
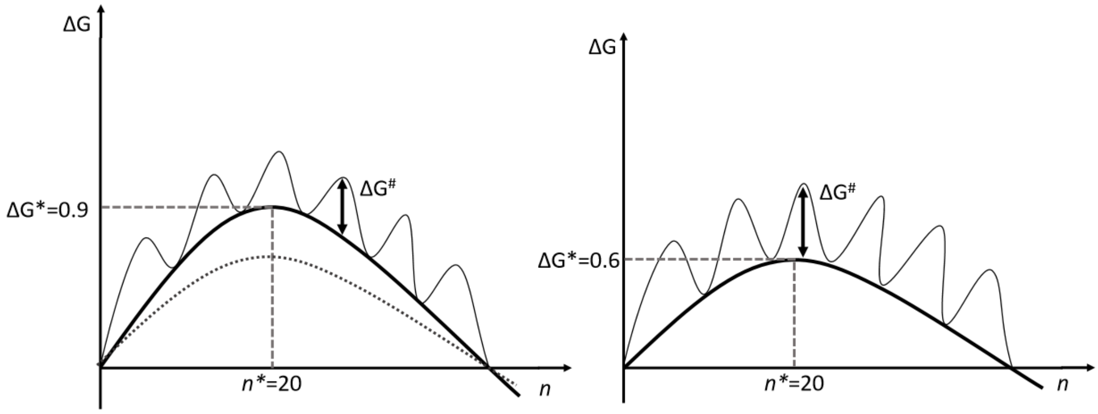


Figure 8.6: A sketch of a possible correlation between the activation energy barrier and the nucleation barrier for PNBA (blue) and PABA (red) in 2-propanol.

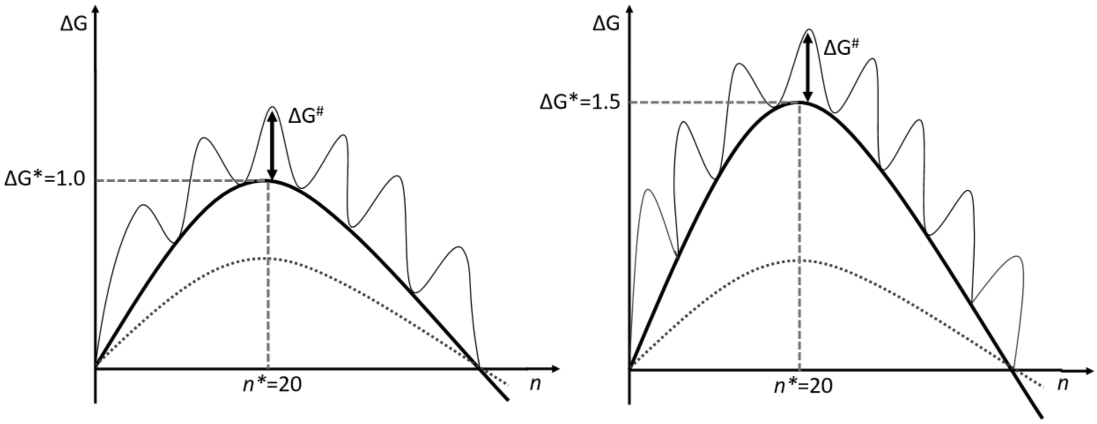
Figure 8.6 illustrates this possible thermodynamic and kinetic barrier, highlighting $\Delta G^{\#}_{\text{PNBA,2-propanol}} < \Delta G^{\#}_{\text{PABA,2-propanol}}$ while $\Delta G*_{\text{PNBA,2-propanol}} > \Delta G*_{\text{PABA,2-propanol}}$. This illustration relates the dependence of the $\Delta G*/B$ and $\Delta G^{\#}/A$ on the nucleation rate. The nucleation rate for PNBA in 2-propanol is the lowest at 83 m^{-3} s^{-1} because it has the highest thermodynamic barrier (highest B), even though the activation energy barrier is the smallest (highest A). Therefore, the nucleation rate is highly dependent on the $\Delta G*/B$ rather than the $\Delta G^{\#}/A$.

Figure 8.7 illustrates the possible energy levels ($\Delta G*$ and $\Delta G^{\#}$) for all the systems. The $\Delta G^{\#}$ is estimated to be smallest for PNBA in 2-propanol, followed by PTA in toluene < PTA in 2-propanol < PABA in 2-propanol $\cong$ PABA in acetonitrile $\cong$ BA in toluene $\cong$ BA in acetonitrile. In general, the $\Delta G^{\#}$ is the smallest in the order PNBA < PTA < PABA $\cong$ BA, which is the opposite of the A value, where the highest is in the order of PNBA > PTA > PABA $\cong$ BA. Nevertheless, this is best validate with an experimental result.



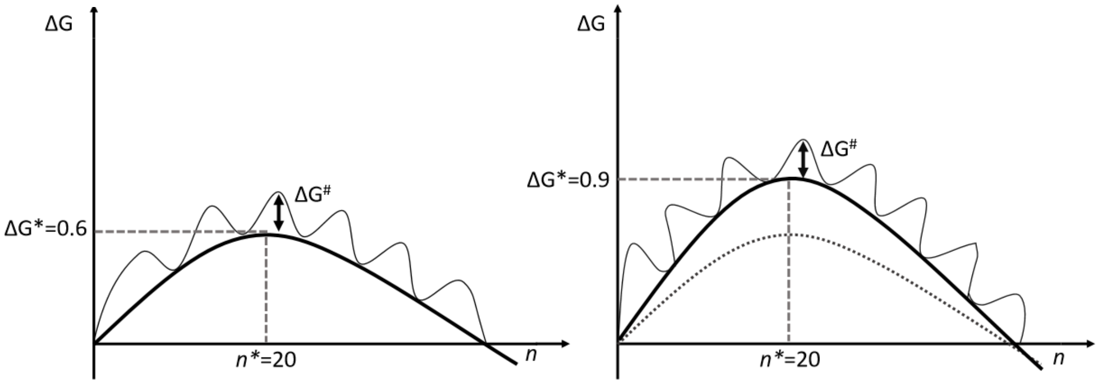
(a) PABA in 2-propanol

(b) PABA in acetonitrile



(c) BA in toluene

(d) BA in acetonitrile



(e) PTA in toluene

(f) PTA in 2-propanol

Figure 8.7: The possible energy diagram correlating the nucleation barrier $\Delta G^{\#}$ and the activation energy barrier , $\Delta G^{\#}$ estimated based on the A and $J_{B=0.03}$ for (a) PABA in 2-propanol, (b) PABA in acetonitrile (c) BA in toluene (d) BA in acetonitrile (e) PTA in toluene and (f) PTA in 2-propanol. Dashed line represent the ΔG^{*} at $B=0.03$.

The influence of solute-solvent interactions on the kinetics is examined based on the pre-exponential factor A and the stretching frequency at the C=O group (FTIR). Excluding BA, there is a strong linear relationship for PNBA, PTA, and PABA at lower carbonyl (C=O) stretching frequency to the $\ln(A)$ as shown in Figure 8.8. This correlates with the differences in $\Delta G^{\#}$; PNBA has a weakly bonded intermolecular interaction (solute-solvent), thus needs a smaller $\Delta G^{\#}$ (bigger $\ln(A)$) in terms of breaking the solute-solvent and then creating a solute-solute interaction. PABA, which is strongly bonded, requires a higher $\Delta G^{\#}$, while PTA is in between PNBA and PABA. This correlation is strong in comparing the solvates in the 2-propanol system and dimers for PTA in toluene and PABA in acetonitrile. The solvates at the higher wavenumber, which have the weakest solute-solvent interaction did not seems to influence the solute-solute formation.

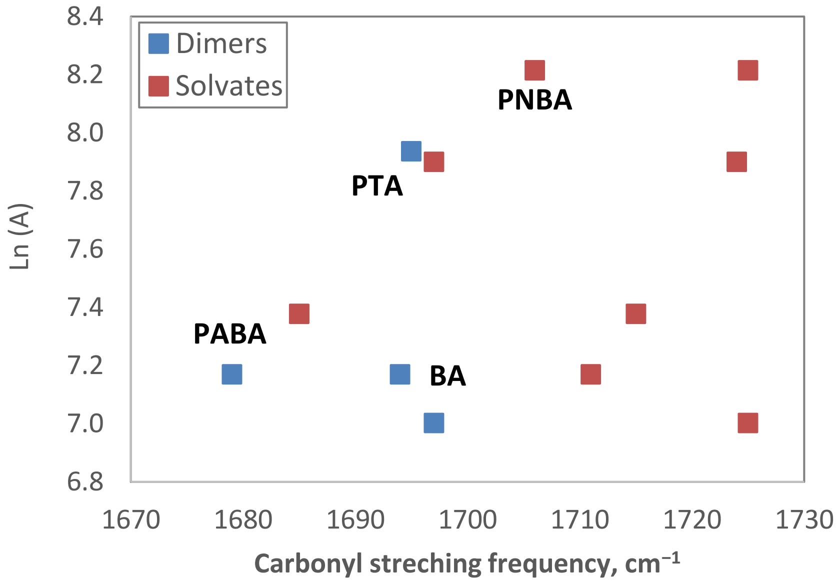


Figure 8.8: A plot of C=O stretching frequency of dimers (blue) and solvates (red) against pre-exponential kinetic factor, ln(A).

In contrast, BA which has a C=O frequency similar to that of PTA has a lower $\ln(A)$ value. BA and PTA are two compounds which chemically (and hence electronically) differ only slightly, through the methyl group vs the hydrogen in *para* position, but exhibit significant differences in terms of solubility, thermodynamic factor, B, pre-exponential factor, A, and the nucleation rate, J. These differences may thus not be primarily related to carbonyl group interactions, as both have the same C=O frequency, indicating the same strength of solute-solvent and solute-solute interaction.

This inconsistency in correlating the intermolecular interaction at the carbonyl group to the $\Delta G^{*}/B$ and the $\Delta G^{\#}/A$ suggest that this interaction is not the main driver for the nucleation process of all carboxylic acids. PNBA and PABA seem to show a consistent result that can be linked to the strength of solute-solvent interaction at the carbonyl group, but the results for BA and PTA do not fit the pattern. Possibly, the π−π interaction plays a more significant role for BA than the interaction at the carbonyl group as shown in the recent studies [136].

8.3 Conclusion

Energy diagrams for ΔG^{*} have been predicted based on the maximum work needed to create a nucleus size of 20 molecules using the thermodynamic factor, B. Then, a prediction of $\Delta G^{\#}$ is made by re-calculating the nucleation rate for each system at constant $B = 0.03$ and $A = 3700$. The ratio between the $\Delta G^{\#}$ to ΔG^{*} were then estimated. The energy diagrams illustrate the interplay between the thermodynamic factor, B ,and the pre-exponential kinetic factor , A, to the nucleation rate. The lower carbonyl stretching frequency for the solvates in PNBA, PTA, and PABA in 2-propanol and the dimers for PABA in acetonitrile and PTA in toluene shows a good correlation to the pre-exponential kinetic factor, A. Therefore, the desolvation and the substituents do seems to affect the nucleation kinetics. However, BA in acetonitrile and toluene did not show similar correlation suggesting that the nucleation process is not only dependent on the hydrogen bond to the carbonyl.

9. CONCLUSION AND FUTURE WORKS

The work on *p*-nitrobenzoic acid (PNBA), benzoic acid (BA), and *p*-aminobenzoic acid (PABA) in this dissertation has further explored the application of core level spectroscopy - near-edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron spectroscopy (XPS). These techniques were applied in studying the electronic structure of PNBA polymorphs and the influence of substituents on the local structure, in the solid and solution state. In addition, a comparison of the results is made to the vibrational spectroscopy techniques – infrared (IR) and Raman. This is one of the novels works in this dissertation: the combination of NEXAFS, XPS, IR and Raman studies has highlighted the unique features of each technique and revealed more information regarding the molecular structures. This was combined with studies of the nucleation process, where the nucleation kinetics data of PNBA and other benzoic acid derivatives [136] were applied for the first time to illustrate the potential energy diagrams that linked the thermodynamic and kinetic energy barriers to the nucleation rate.

The two polymorphs of *p*-nitrobenzoic acid (PNBA) have been investigated in terms of their local interactions in the solid state and nucleation behaviour from solution. Both PNBA Form I and Form II have the carboxylic acid dimer synthon and did not exhibit significant differences as demonstrated by vibrational and core level electron spectroscopies – IR, Raman, NEXAFS and XPS. Infrared spectroscopy, however, did show a subtle difference at the carboxylic acid group, which is correlated with the hydroxyl bond length. The single crystal X-ray diffraction results showed that both polymorphs have similar intermolecular interactions although different crystal packing. Solubility data for both polymorphs was measured in different solvents, showing the polymorphs are enantiotropically related; PNBA Form I is the stable form below 51 °C. PNBA Form II consistently nucleated first in 2-propanol and ethyl acetate, before the transition to Form I at 25 °C. The induction time measurements are highly sensitive to changes in the process conditions, so the experimental methodology was revised resulting in a consistent calculated nucleation rate.

Core level spectroscopies, XPS and NEXAFS, have been applied, discussed, and compared to the established infrared spectroscopy interpretation. The effects of substituents on the local structures were studied for PNBA, PABA and BA. The XP and NEXAFS spectroscopy have the added advantage that the effect of substituents is observed at several of the chemical moieties, for example, C=C, C=O, C−N, O−H while the infrared is only at the C=O in the solid state. NEXAFS and infrared spectra in the solution showed a consistent effect of substituents based on the O 1s (C=**O̲**) → 1π* chemical shift and C=O stretching frequency. However, infrared spectroscopy has the added advantage on characterising carboxylic acid dimer, solvates and monomers in both solid and solution state. Overall, the solid and solution state studies show that PNBA with a strong electron withdrawing group ($-NO_2$) has a weaker hydrogen bonded interaction at the carboxylic acid followed by BA, whereas PABA with a strong electron donating group ($-NH_2$) has the strongest hydrogen bonded interaction.

The potential energy diagrams for the benzoic acid derivatives highlighting the influence of nucleation and activation barrier were derived based on the pre-exponential kinetic factor, A, and thermodynamic factor, B. The nucleation barrier, ΔG^*, is the highest in the order of PNBA < BA < PABA $\cong$ *p*-toluic acid (PTA) while the activation energy barrier, $\Delta G^{\#}$, is potentially the lowest in the order of PNBA < PTA < PABA $\cong$ BA. The solvents (acetonitrile, toluene, 2−propanol) did not influence the outcomes of these energy barriers. The nucleation rate, which is in the order of PTA > PABA > BA > PNBA, is highly dependent on the $\Delta G^*/B$ rather than the $\Delta G^{\#}/A$. With the exception of BA, the activation energy barrier is in the order of substituent effect based on the carbonyl stretching frequency. PNBA, which is weakly bonded, needs a smaller $\Delta G^{\#}$ in creating a solute-solute interaction, while PABA which is strongly bonded needs a bigger $\Delta G^{\#}$.

All these results suggest the essential of studying the nucleation process from the molecular level. The solid and solution state of PNBA, BA, and PABA showed consistent results; PNBA has the weakest hydrogen bonding in the solid and solution state, followed by BA and PABA. The combination of IR, Raman, NEXAFS and XPS spectroscopies gave a more extensive overview of the molecular structure in solid and solution state. With the exception of BA, PNBA and PABA showed a link between the strength of solute-solvent hydrogen bonding to the nucleation rate.

From this thesis, there are several interesting future works to pursue. Advanced density functional theory (DFT) calculations are an essential in more advanced interpretation of XPS and NEXAFS results in the solid and solution state. The DFT calculations would allow further investigation of the influence of monomer, catemer, dimer, and solute-solvent interaction to the local electronic structure. For example, NEXAFS data on its own as in this dissertation raised several questions in comparison to the infrared spectra; possibly the existence of two carbonyl bands underneath the single band area and weaker solvent effects. In addition, the XPS data in the solution could give an insight the effects of the molecular environment in the solution state to the core level (1s)[52]. A comprehensive study combining NEXAFS, XPS, and DFT calculations would provide a detailed overview of the electronic structure in the solution state. Continuing on from the studies of *para*- systems, *meta*- and *ortho*- nitrobenzoic acid and aminobenzoic acid would be interesting systems to study the resonance and inductive effect on the electronic structure for all the chemical moieties. Studying the crystal structures with catemer building units, such as α-oxalic acid, maleic acid, chloroacetic acid and vanillin[81] would also be useful to further investigate whether XPS and NEXAFS techniques have unique features to differentiate a catemer and dimer as the IR and Raman techniques. Finally, these techniques have shown signifant application for characterising organic systems, and could be further applied in studying polymorphs, hydrates, and solvates of organic compounds.

A second strand of future work relates to further understanding of the nucleation process. Temperature dependent nucleation rate measurements could be carried out to obtain the actual value of the activation energy barrier, $\Delta G^{\#}$[25]. These values should then be re-assessed in comparison to the carbonyl stretching frequency to validate the influence of hydrogen bonding interaction at the carboxylic acid in creating solute-solute interaction. The experimentally obtained $\Delta G^{\#}$ value can then be compared to the relative value made in this thesis and used to illustrate the actual energy diagrams. Modelling work to predict the calculated $\Delta G^{\#}$ could also be carried out and compare to the experimental results. Future modelling works can further combine the knowledge of crystal structures, molecular interactions (solute-solvent and solute-solute) and nucleation energy barriers in order to understand the molecular pathways to form crystals.

REFERENCES

1 N. Variankaval, A. S. Cote and M. F. Doherty, *AIChE J.*, **2008**, *54*, 1682–1688.

2 R. J. Davey and J. Garside, *From molecules to crystallizers*, Oxford University Press, **2000**.

3 H. G. Brittain, *Polymorphism in pharmaceutical solids*, Marcel Dekker, **1999**.

4 R. J. Davey, S. L. M. Schroeder and J. H. Ter Horst, *Angew. Chemie - Int. Ed.*, **2013**, *52*, 2167–2179.

5 D. Kashchiev and G. M. van Rosmalen, *Cryst. Res. Technol.*, **2003**, *38*, 555–574.

6 N. Pienack and W. Bensch, *Angew. Chemie - Int. Ed.*, **2011**, *50*, 2014–2034.

7 J. W. Mullin, *Crystallization*, Butterworth-Heinemann, **2001**.

8 R. J. Davey, G. Dent, R. K. Mughal and S. Parveen, *Cryst. Growth Des.*, **2006**, *6*, 1788–1796.

9 D. H. M. Beiny and J. W. Mullin, *J. Chem. Eng. Data*, **1987**, *32*, 9–10.

10 M. Svärd, F. L. Nordström, E.-M. Hoffmann, B. Aziz and Å. C. Rasmuson, *CrystEngComm*, **2013**, *15*, 5020.

11 J. Bernstein, *Polymorphism in molecular crystals*, Oxford University Press, 1st ed., **2008**.

12 F. L. Nordström and Å. C. Rasmuson, *J. Pharm. Sci.*, **2006**, *95*, 748–760.

13 F. L. Nordström and Å. C. Rasmuson, *Eur. J. Pharm. Sci.*, **2009**, *36*, 330–344.

14 F. L. Nordström and Å. C. Rasmuson, *Eur. J. Pharm. Sci.*, **2006**, *28*, 377–384.

15 S. Gracin and Å. C. Rasmuson, *Cryst. Growth Des.*, **2004**, *4*, 1013–1023.

16 A. Maher, D. Croker, Å. C. Rasmuson and B. K. Hodnett, *J. Chem. Eng. Data*, **2010**, *55*, 5314–5318.

17 M. Svard and Å. C. Rasmuson, *Cryst. Growth Des.*, **2013**, *13*, 1140–1152.

18 J. Thati, F. L. Nordström and Å. C. Rasmuson, *J. Chem. Eng. Data*, **2010**, *55*, 5124–5127.

19 H. Yang, J. Thati and A. C. Rasmuson, *J. Chem. Thermodyn.*, **2012**, *48*, 150–159.

20 D. Erdemir, A. Y. Lee and A. S. Myerson, *Acc. Chem. Res.*, **2009**, *42*, 621–629.

21 P. G. Vekilov, *Cryst. Growth Des.*, **2010**, *10*, 5007–5019.

22 P. G. Vekilov, *Nonoscale*, **2010**, *2*, 2346–2357.

23 D. Gebauer and H. Cölfen, *Nano Today*, **2011**, *6*, 564–584.

24 D. Kashchiev, *Nucleation*, Butterworth-Heinemann, **2000**.

25 H. Yang and J. H. Horst, *Crystal nucleation of small organic molecules*, **2016**.

26 C. Brandel and J. H. Ter Horst, *Faraday Discuss.*, **2015**, 199.

27 S. Jiang and J. H. Ter Horst, *Cryst. Growth Des.*, **2011**, *11*, 256–261.

28 R. A. Sullivan, R. J. Davey, G. Sadiq, G. Dent, K. R. Back, J. H. Ter Horst, D. Toroz and R. B. Hammond, *Cryst. Growth Des.*, **2014**, *14*, 2689–2696.

29 J. Liu, M. Svärd and Å. C. Rasmuson, *Cryst. Growth Des.*, **2014**, *14*, 5521–5531.

30 D. Mealey, J. Zeglinski, D. Khamar and A. C. Rasmuson, *Faraday Discuss.*, **2015**, 309.

31 S. A. Kulkarni, S. S. Kadam, H. Meekes, A. I. Stankiewicz and J. H. Ter Horst, *Cryst. Growth Des.*, **2013**, *13*, 2435–2440.

32 S. Duckett and B. Gilbert, *Foundations of spectroscopy*, Oxford University Press, **2000**.

33 R. K. Harris, *J. Pharm. Pharmacol.*, **2007**, *59*, 225–239.

34 J. S. Stevens, L. K. Newton, C. Jaye, C. A. Muryn, D. A. Fischer and S. L. M. Schroeder, *Cryst. Growth Des.*, **2015**, *15*, 1776–1783.

35 G. R. Desiraju, *J. Am. Chem. Soc.*, **2013**, *135*, 9952–9967.

36 L. . Bellamy, *The infrared spectra of complex molecules*, Chapman and Hall, 2nd edn., **1980**.

37 J. Lascombe, M. Haurie and M. Josien, *J.Ch.Phys.Tome*, **1962**, 1233–1245.

38 Y. Fujii, H. Yamada and M. Mizuta, *J. Phys. Chem.*, **1988**, *92*, 6768–6772.

39 W. Du, A. J. Cruz-Cabeza, S. Woutersen, R. J. Davey and Q. Yin, *Chem. Sci.*, **2015**, *6*, 3515–3524.

40 M. Gantenberg, M. Halupka and W. Sander, *Chem. - A Eur. J.*, **2000**, *6*, 1865–9.

41 R. J. Davey, N. Blagden, S. Righini, H. Alison, M. J. Quayle and S. Fuller, *Cryst. Growth Des.*, **2001**, *1*, 59–65.

42 N. Blagden, R. J. Davey, H. F. Lieberman, L. Williams, R. Payne, R. Roberts, R. Roweb and R. Dochertyb, *J. Chem. Soc. Faraday Trans.*, **1998**, *94*, 1035–1044.

43 D. Khamar, J. Zeglinski, D. Mealey and Å. C. Rasmuson, *J. Am. Chem. Soc.*, **2014**, *136*, 11664–11673.

44 C. S. Towler and L. S. Taylor, *Cryst. Growth Des.*, **2007**, *7*, 633–638.

45 G. Hähner, *Chem. Soc. Rev.*, **2006**, *35*, 1244–1255.

46 C. C. Chusuei and D. W. Goodman, *Encycl. Phys. Sci. Technol.*, **2003**, 921–938.

47 J. S. Stevens, A. Gainar, E. Suljoti, J. Xiao, R. Golnak, E. F. Aziz and S. L. M. Schroeder, *Chem. - A Eur. J.*, **2015**, *21*, 7256–7263.

48 J. S. Stevens, S. J. Byard, C. C. Seaton, G. Sadiq, R. J. Davey and S. L. M. Schroeder, *Phys. Chem. Chem.*, **1150**, *16*, 1150–1160.

49 M. J. Thomason, C. R. Seabourne, B. M. Sattelle, G. A. Hembury, J. S. Stevens, A. J.

Scott, E. F. Aziz and S. L. M. Schroeder, *Faraday Discuss.*, **2015**, *179*, 269–289.

50 J. S. Stevens, A. Gainar, C. Jaye, D. A. Fischer and S. L. M. Schroeder, *J. Phys. Conf. Ser.*, **2016**, *712*, 12133.

51 J. S. Stevens, C. R. Seabourne, C. Jaye, D. A. Fischer, A. J. Scott and S. L. M. Schroeder, *J. Phys. Chem. B*, **2014**, *118*, 12121–12129.

52 M. A. Brown, F. Vila, M. Sterrer, S. Thü, B. Winter, M. Ammann, J. J. Rehr and J. A. Van Bokhoven, *J. Phys. Chem. Lett.*, **2012**, *3*, 1754–1759.

53 D. Mealey, D. M. Croker and Å. C. Rasmuson, *CrystEngComm*, **2015**, *17*, 3961–3973.

54 D. H. McDaniel and H. C. Brown, *J. Org. Chem.*, **1958**, *23*, 420–427.

55 The Cambridge Structural Database (CSD) - The Cambridge Crystallographic Data Centre (CCDC), https://www.ccdc.cam.ac.uk/solutions/csd-system/components/csd/, (accessed 6 April 2017).

56 P. Groth, *Acta Chem. Scand.*, **1980**, *A 34*, 229–233.

57 M. Colapietro and A. Domenicano, *Acta Crystallogr. Sect. B*, **1977**, *33*, 2240–2243.

58 M. Tonogaki, T. Kawata, S. Ohba, Y. Iwata and I. Shibuya, *Acta Crystallogr. Sect. B*, **1993**, *49*, 1031–1039.

59 W. Clegg, *Crystal structure determination*, Oxford University Press, **1998**.

60 A. Brisdon, *Inorganic spectroscopic methods*, Oxford University Press, **1998**.

61 P. B. Coleman, *Practical sampling techniques for infrared analysis*, CRC Press, **1993**.

62 Thermo Electron Corporation, in *FT-IR General Accessories*, **2017**.

63 J. Ferraro and K. Nakamoto, *Introductory Raman spectroscopy*, Academic Press Limited, **1994**.

64 D. Briggs and M.P. Seah, *Practical Surface Analysis, Auger and X-ray Photoelectron Spectroscopy*, Wiley, 2nd edn., **1990**.

65 J. S. Stevens and S. L. M. Schroeder, in *Encyclopedia of Physical Organic Chemistry*, **2017**.

66 J. F. Watts, *X-ray photoelectron spectroscopy*, Elsevier Science Ltd, **1994**, vol. 45.

67 J. Stohr, *NEXAFS spectroscopy*, Springer Science & Business Media, **2003**.

68 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev,

A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, **2004**.

69 S. R. Vippagunta, H. G. Brittain and D. J. W. Grant, *Adv. Drug Deliv. Rev.*, **2001**, *48*, 3–26.

70 A. Llinàs and J. M. Goodman, *Drug Discov. Today*, **2008**, *13*, 198–210.

71 G. A. Stephenson, R. A. Forbes and S. M. Reutzel-Edens, *Adv. Drug Deliv. Rev.*, **2001**, *48*, 67–90.

72 D. E. Bugay, *Adv. Drug Deliv. Rev.*, **2001**, *48*, 43–65.

73 S. Shakoor, T. Ahsan, K. Jabeen, M. Raza and R. Hasan, *Res. Microbiol.*, **2010**, *14*, 1644–1646.

74 R. Varala, R. Enugala and S. R. Adapa, *J. Braz. Chem. Soc.*, **2007**, *18*, 291–296.

75 T. S. Thakur and S. S. Singh, *Cryst. Growth Des.*, **2015**, 3280–3292.

76 F. H. Allen, C. A. Baalham, J. P. M. Lommerse, P. R. Raithby and E. Sparrb~, *Acta Cryst*, **1997**, *53*, 1017–1024.

77 T. D. Sakore and L. M. Pant, *Acta Crystallogr.*, **1966**, *21*, 715–719.

78 S. S. Tavale and L. M. Pant, *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.*, **1971**, *27*, 1479–1481.

79 T. Steiner, *Angew. Chem. Int. Ed.*, **2002**, *41*, 49–76.

80 P. A. Vanderhof, R. A. Lalancette and H. W. Thompson, *J . Org. Chem*, **1990**, *55*, 1696–1698.

81 S. Parveen, The University of Manchester, **2007**.

82 L. Colombo, K. Furič, D. Kirin, P. Bleckmann, B. Schrader, W. Meiner, K. Furic, D. KIRIN Institue, R. Boskovic, Z. P. Bleckmann and W. Meier, **1973**, *14*, 105–116.

83 Chemistry WebBook, http://webbook.nist.gov/, (accessed 7 April 2017).

84 R. C. Millikan and K. S. Pitzer, *J. Am. Chem. Soc.*, **1958**, *80*, 3515–3521.

85 Y. Marcus, *Chem. Soc. Rev.*, **1993**, *22*, 409.

86 J. A. Armando and E. Z. Joyce, *J. Pharm. Sci.*, **1969**, *58*, 983–987.

87 M. Kanke and K. Sekiguchi, *Chem.Pharm.Bull*, **1973**, *21*, 878–884.

88 K. U. Rakami, Y. S. Hono, A. H. Igashi, K. U. Memoto and M. G. Odo,

Chem.Pharm.Bull, **2002**, *50*, 263–267.

89 J. Ulrich and C. Strege, *J. Cryst. Growth*, **2002**, 2130–2130.

90 S. Jiang, Delft University of Technology, **2009**.

91 S. Teychené and B. Biscans, *Cryst. Growth Des.*, **2008**, *8*, 1133–1139.

92 N. A. Mitchell, P. J. Frawley and C. T. ’ Ciardhá, *J. Cryst. Growth*, **2011**, *321*, 91–99.

93 L. Goh, K. Chen, V. Bhamidi, G. He, N. C. S. Kee, P. J. A. Kenis, C. F. Zukoski and R. D. Braatz, *Cryst. Growth Des.*, **2010**, *10*, 2515–2521.

94 D. Knezic, J. Zaccaro and A. S. Myerson, *J.Phys.Chem.B*, **2004**, *108*, 10672–10677.

95 F. L. Nordstrom, M. Svard, B. Malmberg and A. C. Rasmuson, *Cryst. Growth Des.*, **2012**, *12*, 4340–4348.

96 M. Kuhs, J. Zeglinski and Å. C. Rasmuson, *Cryst. Growth Des.*, **2014**, *14*, 905–915.

97 P. Van der Heide, *X-ray Photoelectron Spectroscopy: An introduction to principles and practices*, John Wiley & Sons, 1st edn., **2012**.

98 G. Tu, Y. Tu, O. Vahtras and H. Ågren, *Chem. Phys. Lett.*, **2008**, *468*, 294–298.

99 S. Garcia-Gil, A. Arnau and A. Garcia-Lekue, *Surf. Sci.*, **2013**, *613*, 102–107.

100 K. Tabayashi, O. Takahashi, H. Namatame and M. Taniguchi, *Chem. Phys. Lett.*, **2013**, 1–9.

101 M. Ni and B. D. Ratner, *Surf. Interface Anal.*, **2008**, *40*, 1356–1361.

102 C. Tang, J. Zhu, Q. Zhou, J. Wei, R. Zhu and H. He, *J. Phys. Chem.*, **2014**, *118*, 26249–26257.

103 F. S. Ohuchi, S. Ghose, M. H. Engelhard and D. R. Baer, *Am. Mineral.*, **2006**, *91*, 740–746.

104 J. S. Stevens, A. C. De Luca, M. Pelendritis, G. Terenghi, S. Downes and S. L. M. Schroeder, *Surf. Interface Anal.*, **2013**, *45*, 1238–1246.

105 M. Mogi, Y. Inoue, T. Yamamoto and I. Tanaka, *Jpn. J. Appl. Phys.*, **2005**, *44*, 4057–4059.

106 X-ray photoelectron spectroscopy (XPS) reference pages: Using adventitious carbon for charge correcting, http://www.xpsfitting.com/2008/10/using-adventitious-carbon-for-charge.html, (accessed 29 April 2017).

107 H. Neuvonen, K. Neuvonen, A. Koch, E. Kleinpeter and P. Pasanen, *J. Org. Chem.*, **2002**, *67*, 6995–7003.

108 H. Neuvonen and K. Neuvonen, *J. Chem. Soc. Perkin Trans. 2*, **1999**, 1497–1502.

109 R. N. Jones, W. F. Forbes and W. . Mueller, *Can. J. Chem.*, **1957**, *35*, 504–514.

110 F. González-Sánchez, *Spectrochim. Acta*, **1958**, *12*, 17–33.

111 M. S. G. Flett, *Spectrochim. Acta*, **1962**, *18*, 1537–1556.

112 C. C. Turci, S. G. Urquhart and A. P. Hitchcock, *Can. J. Chem.*, **1996**, *74*, 851–869.

113 I. Bâldea, B. Schimmelpfennig, M. Plaschke, J. Rothe, J. Schirmer, A. B. Trofimov and T. Fanghänel, *J. Electron Spectros. Relat. Phenomena*, **2007**, *154*, 109–118.

114 M. Plaschke, J. Rothe, M. Altmaier, M. A. Denecke and T. Fanghänel, *J. Electron Spectros. Relat. Phenomena*, **2005**, *148*, 151–157.

115 R. J. Davey, G. Dent, R. K. Mughal and S. Parveen, *Cryst. Growth Des.*, **2006**, *6*, 1788–1796.

116 V. Sablinskas, M. Pucetaite, J. Ceponkus and L. Kimtys, *J. Mol. Struct.*, **2010**, *976*, 263–269.

117 C. Emmeluth, M. A. Suhm and D. Luckhaus, *J. Chem. Phys.*, **2003**, *118*, 2242–2255.

118 W. Sander and M. Gantenberg, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, **2005**, *62*, 902–909.

119 G. R. Desiraju, *Angew. Chemie Int. Ed. English*, **1995**, *34*, 2311–2327.

120 D. K. Bučar, R. F. Henry, X. Lou, R. W. Duerst, T. B. Borchardt, L. R. MacGillivray and G. G. Z. Zhang, *Mol. Pharm.*, **2007**, *4*, 339–346.

121 A. Lemmerer, D. A. Adsmond, C. Esterhuysen and J. Bernstein, *Cryst. Growth Des.*, **2013**, *13*, 3935–3952.

122 N. B. Colthup, L. H. Daly and S. E. Wiberley, *Introduction to infrared and Raman spectroscopy*, Academic Press, **1975**.

123 L. P. Hammett, *J. Am. Chem. Soc.*, **1937**, *59*, 96–103.

124 D. D. Perrin, E. P. Serjeant, B. Dempsey, A. Albert, A. R. Katritzky, Ed and E. P. M. J. Serjeant, in *CRC Handbook of Biochemistry*, Academic Press CRC Press Chapman and Hall Chapman and Hall Merck & Co, **1965**.

125 S. Parveen, R. J. Davey, G. Dent and R. G. Pritchard, *Chem. Commun.*, **2005**, 1531–1533.

126 R. A. Chiarella, A. L. Gillon, R. C. Burton, R. J. Davey, G. Sadiq, A. Auffret, M. Cioffi and C. A. Hunter, *Faraday Discuss.*, **2007**, *136*, 179–193.

127 D. Erdemir, S. Chattopadhyay, L. Guo, J. Ilavsky, H. Amenitsch, C. U. Segre and A. S. Myerson, *Phys. Rev. Lett.*

128 K. Chadwick, R. J. Davey, G. Dent, R. G. Pritchard, C. A. Hunter and D. Musumeci, *Cryst. Growth Des.*, **2009**, *9*, 1990–1999.

129 A. Mattei and T. Li, *Pharm. Res.*, **2012**, *29*, 460–470.

130 S. A. Kulkarni, E. S. Mcgarrity, H. Meekes and J. H. Ter Horst, *Chem. Commun. Chem.*

Commun, **2012**, *48*, 4983–4985.

131 C. J. . Brooks, G. Eglinton and J. F. Morman, *J. Chem. Soc.*, **1961**, 106–116.

132 R. A. Sullivan, The University of Manchester, **2015**.

133 P. Novak, D. Vikic-Topic, Z. Meic, S. Sekusak and A. Sabljic, *J. Mol. Struct.*, **1995**, *356*, 131–141.

134 C. Lindenberg and M. Mazzotti, *J. Cryst. Growth*, **2008**, *311*, 1178–1184.

135 *Faraday Discuss.*, **2015**, *179*, 155–197.

136 A. J. Cruz-cabeza, R. J. Davey, S. S. Sachitananthan, R. Smith, S. K. Tang, T. Vetter and Y. Xiao, **2017**, *To be published*.